ALLEN TAFLOVE • SUSAN C. HAGNESS

computational electrodynamics

THE FINITE-DIFFERENCE TIME-DOMAIN METHOD

THIRD EDITION

Computational Electrodynamics

The Finite-Difference Time-Domain Method

Third Edition

Allen Taflove
Susan C. Hagness

ARTECH
HOUSE
BOSTON | LONDON
artechhouse.com

Library of Congress Cataloging-in-Publication Data
Taflove, Allen.
Computational electrodynamics: the finite-difference time-domain method.—3rd ed./
Allen Taflove, Susan C. Hagness.
p. cm.—(Artech House antennas and propagation library)
Includes bibliographical references and index.
ISBN 1-58053-832-0 (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 2005
537.6—dc22 2005045273

British Library Cataloguing in Publication Data
A catalogue record of this book is available from the British Library.

Cover design by Igor Valdman

685 Canton Street
Norwood, MA 02062

International Standard Book Number: 1-58053-832-0

10 9 8 7 6 5 4

Contents

Preface to the Third Edition xix

1 Electrodynamics Entering the 21st Century 1

1.1 Introduction 1
1.2 The Heritage of Military Defense Applications 1
1.3 Frequency-Domain Solution Techniques 2
1.4 Rise of Finite-Difference Time-Domain Methods 3
1.5 History of FDTD Techniques for Maxwell's Equations 4
1.6 Characteristics of FDTD and Related Space-Grid Time-Domain Techniques 6
1.6.1 Classes of Algorithms 6
1.6.2 Predictive Dynamic Range 7
1.6.3 Scaling to Very Large Problem Sizes 8
1.7 Examples of Applications 9
1.7.1 Impulsive Around-the-World Extremely Low-Frequency Propagation 10
1.7.2 Cellphone Radiation Interacting with the Human Head 11
1.7.3 Early-Stage Detection of Breast Cancer Using an Ultrawideband Microwave Radar 11
1.7.4 Homing Accuracy of a Radar-Guided Missile 12
1.7.5 Electromagnetic Wave Vulnerabilities of a Military Jet Plane 12
1.7.6 Millimeter-Wave Propagation in a Defect-Mode Electromagnetic Bandgap Structure 13
1.7.7 Photonic Crystal Microcavity Laser 14
1.7.8 Photonic Crystal Cross-Waveguide Switch 15
1.8 Conclusions 16
References 16

2 The One-Dimensional Scalar Wave Equation 21

2.1 Introduction 21
2.2 Propagating-Wave Solutions 21
2.3 Dispersion Relation 22
2.4 Finite Differences 23
2.5 Finite-Difference Approximation of the Scalar Wave Equation 24
2.6 Numerical Dispersion Relation 27
2.6.1 Case 1: Very Fine Sampling in Time and Space 28
2.6.2 Case 2: Magic Time-Step 29
2.6.3 Case 3: Dispersive Wave Propagation 29
2.6.4 Example of Calculation of Numerical Phase Velocity and Attenuation 34
2.6.5 Examples of Calculations of Pulse Propagation 34
2.7 Numerical Stability 39
2.7.1 Complex-Frequency Analysis 40
2.7.2 Examples of Calculations Involving Numerical Instability 43
2.8 Summary 45
Appendix 2A: Order of Accuracy 47
2A.1 Lax-Richtmyer Equivalence Theorem 47
2A.2 Limitations 48
References 48
Selected Bibliography on Stability of Finite-Difference Methods 49
Problems 49

3 Introduction to Maxwell's Equations and the Yee Algorithm

Allen Taflove and Jamesina Simpson **51**

3.1 Introduction 51
3.2 Maxwell's Equations in Three Dimensions 51
3.3 Reduction to Two Dimensions 54
3.3.1 TM_z Mode 55
3.3.2 TE_z Mode 55
3.4 Reduction to One Dimension 56
3.4.1 x-Directed, z-Polarized TEM Mode 56
3.4.2 x-Directed, y-Polarized TEM Mode 57
3.5 Equivalence to the Wave Equation in One Dimension 57
3.6 The Yee Algorithm 58
3.6.1 Basic Ideas 58
3.6.2 Finite Differences and Notation 60
3.6.3 Finite-Difference Expressions for Maxwell's Equations in Three Dimensions 62
3.6.4 Space Region with a Continuous Variation of Material Properties 67
3.6.5 Space Region with a Finite Number of Distinct Media 69
3.6.6 Space Region with Nonpermeable Media 71
3.6.7 Reduction to the Two-Dimensional TM_z and TE_z Modes 73
3.6.8 Interpretation as Faraday's and Ampere's Laws in Integral Form 75
3.6.9 Divergence-Free Nature 78
3.7 Alternative Finite-Difference Grids 80
3.7.1 Cartesian Grids 80
3.7.2 Hexagonal Grids 82
3.8 Emerging Application: Gridding the Planet Earth 85
3.8.1 Background 85
3.8.2 The Latitude-Longitude Space Lattice 86
3.8.3 The Geodesic (Hexagon-Pentagon) Grid 99
3.9 Summary 103
References 104
Problems 105

4 Numerical Dispersion and Stability **107**

4.1 Introduction 107
4.2 Derivation of the Numerical Dispersion Relation for Two-Dimensional Wave Propagation 107
4.3 Extension to Three Dimensions 110
4.4 Comparison with the Ideal Dispersion Case 111
4.5 Anisotropy of the Numerical Phase Velocity 111
4.5.1 Sample Values of Numerical Phase Velocity 111
4.5.2 Intrinsic Grid Velocity Anisotropy 116
4.6 Complex-Valued Numerical Wavenumbers 120
4.6.1 Case 1: Numerical Wave Propagation Along the Principal Lattice Axes 121
4.6.2 Case 2: Numerical Wave Propagation Along a Grid Diagonal 123
4.6.3 Example of Calculation of Numerical Phase Velocity and Attenuation 126
4.6.4 Example of Calculation of Wave Propagation 126
4.7 Numerical Stability 128
4.7.1 Complex-Frequency Analysis 130
4.7.2 Example of a Numerically Unstable Two-Dimensional FDTD Model 135
4.7.3 Linear Growth Mode When the Normalized Courant Factor Equals 1 137
4.8 Generalized Stability Problem 137
4.8.1 Absorbing and Impedance Boundary Conditions 137

4.8.2 Variable and Unstructured Meshing 137
4.8.3 Lossy, Dispersive, Nonlinear, and Gain Materials 138
4.9 Modified Yee-Based Algorithms for Mitigating Numerical Dispersion 138
4.9.1 Strategy 1: Center a Specific Numerical Phase-Velocity Curve About c 138
4.9.2 Strategy 2: Use Fourth-Order-Accurate Explicit Spatial Differences 139
4.9.3 Strategy 3: Use a Hexagonal Grid, If Possible 146
4.9.4 Strategy 4: Use Discrete Fourier Transforms to Calculate the Spatial Derivatives 150
4.10 Alternating-Direction-Implicit Time-Stepping Algorithm for Operation Beyond the Courant Limit 154
4.10.1 Numerical Formulation of the Zheng/Chen/Zhang Algorithm 155
4.10.2 Sources 161
4.10.3 Numerical Stability 161
4.10.4 Numerical Dispersion 163
4.10.5 Additional Accuracy Limitations and Their Implications 164
4.11 Summary 164
References 165
Problems 166
Projects 167

5 Incident Wave Source Conditions
Allen Taflove, Geoff Waldschmidt, Christopher Wagner, John Schneider, and Susan Hagness **169**

5.1 Introduction 169
5.2 Pointwise $\boldsymbol{E}$ and $\boldsymbol{H}$ Hard Sources in One Dimension 169
5.3 Pointwise $\boldsymbol{E}$ and $\boldsymbol{H}$ Hard Sources in Two Dimensions 171
5.3.1 Green Function for the Scalar Wave Equation in Two Dimensions 171
5.3.2 Obtaining Comparative FDTD Data 172
5.3.3 Results for Effective Action Radius of a Hard-Sourced Field Component 173
5.4 $\boldsymbol{J}$ and $\boldsymbol{M}$ Current Sources in Three Dimensions 175
5.4.1 Sources and Charging 176
5.4.2 Sinusoidal Sources 178
5.4.3 Transient (Pulse) Sources 178
5.4.4 Intrinsic Lattice Capacitance 179
5.4.5 Intrinsic Lattice Inductance 183
5.4.6 Impact upon FDTD Simulations of Lumped-Element Capacitors and Inductors 183
5.5 The Plane-Wave Source Condition 185
5.6 The Total-Field / Scattered-Field Technique: Ideas and One-Dimensional Formulation 186
5.6.1 Ideas 186
5.6.2 One-Dimensional Formulation 188
5.7 Two-Dimensional Formulation of the TF/SF Technique 193
5.7.1 Consistency Conditions 193
5.7.2 Calculation of the Incident Field 197
5.7.3 Illustrative Example 201
5.8 Three-Dimensional Formulation of the TF/SF Technique 204
5.8.1 Consistency Conditions 204
5.8.2 Calculation of the Incident Field 210
5.9 Advanced Dispersion Compensation in the TF/SF Technique 213
5.9.1 Matched Numerical Dispersion Technique 214
5.9.2 Analytical Field Propagation 218
5.10 Scattered-Field Formulation 220
5.10.1 Application to PEC Structures 220
5.10.2 Application to Lossy Dielectric Structures 221
5.10.3 Choice of Incident Plane-Wave Formulation 223

5.11 Waveguide Source Conditions 223
5.11.1 Pulsed Electric Field Modal Hard Source 223
5.11.2 Total-Field / Reflected-Field Modal Formulation 225
5.11.3 Resistive Source and Load Conditions 225
5.12 Summary 226
References 227
Problems 227
Projects 228

6 Analytical Absorbing Boundary Conditions 229

6.1 Introduction 229
6.2 Bayliss-Turkel Radiation Operators 230
6.2.1 Spherical Coordinates 231
6.2.2 Cylindrical Coordinates 234
6.3 Engquist-Majda One-Way Wave Equations 236
6.3.1 One-Term and Two-Term Taylor Series Approximations 237
6.3.2 Mur Finite-Difference Scheme 240
6.3.3 Trefethen-Halpern Generalized and Higher-Order ABCs 243
6.3.4 Theoretical Reflection Coefficient Analysis 245
6.3.5 Numerical Experiments 247
6.4 Higdon Radiation Operators 252
6.4.1 Formulation 252
6.4.2 First Two Higdon Operators 253
6.4.3 Discussion 254
6.5 Liao Extrapolation in Space and Time 255
6.5.1 Formulation 255
6.5.2 Discussion 257
6.6 Ramahi Complementary Operators 259
6.6.1 Basic Idea 259
6.6.2 Complementary Operators 260
6.6.3 Effect of Multiple Wave Reflections 260
6.6.4 Basis of the Concurrent Complementary Operator Method 261
6.6.5 Illustrative FDTD Modeling Results Obtained Using the C-COM 267
6.7 Summary 270
References 270
Problems 271

7 Perfectly Matched Layer Absorbing Boundary Conditions
Stephen Gedney **273**

7.1 Introduction 273
7.2 Plane Wave Incident upon a Lossy Half-Space 274
7.3 Plane Wave Incident upon Berenger's PML Medium 276
7.3.1 Two-Dimensional TE_z Case 276
7.3.2 Two-Dimensional TM_z Case 281
7.3.3 Three-Dimensional Case 281
7.4 Stretched-Coordinate Formulation of Berenger's PML 282
7.5 An Anisotropic PML Absorbing Medium 285
7.5.1 Perfectly Matched Uniaxial Medium 285
7.5.2 Relationship to Berenger's Split-Field PML 288
7.5.3 A Generalized Three-Dimensional Formulation 289
7.5.4 Inhomogeneous Media 290

7.6 Theoretical Performance of the PML 291
7.6.1 The Continuous Space 291
7.6.2 The Discrete Space 292
7.7 Complex Frequency-Shifted Tensor 294
7.7.1 Introduction 294
7.7.2 Strategy to Reduce Late-Time (Low-Frequency) Reflections 296
7.8 Efficient Implementation of UPML in FDTD 297
7.8.1 Derivation of the Finite-Difference Expressions 298
7.8.2 Computer Implementation of the UPML 301
7.9 Efficient Implementation of CPML in FDTD 302
7.9.1 Derivation of the Finite-Difference Expressions 302
7.9.2 Computer Implementation of the CPML 307
7.10 Application of CPML in FDTD to General Media 310
7.10.1 Introduction 310
7.10.2 Example: Application of CPML to the Debye Medium 310
7.11 Numerical Experiments with PML 313
7.11.1 Current Source Radiating in an Unbounded Two-Dimensional Region 313
7.11.2 Highly Elongated Domains and Edge Singularities 317
7.11.3 Microstrip Patch Antenna Array 320
7.11.4 Dispersive Media 322
7.12 Summary and Conclusions 324
References 324
Projects 327

8 Near-to-Far-Field Transformation **329**
Allen Taflove, Xu Li, and Susan Hagness
8.1 Introduction 329
8.2 Two-Dimensional Transformation, Phasor Domain 329
8.2.1 Application of Green's Theorem 330
8.2.2 Far-Field Limit 332
8.2.3 Reduction to Standard Form 334
8.3 Obtaining Phasor Quantities Via Discrete Fourier Transformation 335
8.4 Surface Equivalence Theorem 338
8.5 Extension to Three Dimensions, Phasor Domain 340
8.6 Time-Domain Near-to-Far-Field Transformation 343
8.7 Modified NTFF Procedure to More Accurately Calculate Backscattering from Strongly Forward-Scattering Objects 348
8.8 Summary 351
References 351
Project 352

9 Dispersive, Nonlinear, and Gain Materials **353**
Allen Taflove, Susan Hagness, Wojciech Gwarek, Masafumi Fujii, and Shih-Hui Chang
9.1 Introduction 353
9.2 Generic Isotropic Material Dispersions 354
9.2.1 Debye Media 354
9.2.2 Lorentz Media 354
9.2.3 Drude Media 355
9.3 Piecewise-Linear Recursive-Convolution Method, Linear Material Case 355
9.3.1 General Formulation 356
9.3.2 Application to Debye Media 358

9.3.3 Application to Lorentz Media 358
9.3.4 Numerical Results 360
9.4 Auxiliary Differential Equation Method, Linear Material Case 361
9.4.1 Formulation for Multiple Debye Poles 361
9.4.2 Formulation for Multiple Lorentz Pole Pairs 363
9.4.3 Formulation for Multiple Drude Poles 365
9.4.4 Illustrative Numerical Results 367
9.5 Modeling of Linear Magnetized Ferrites 369
9.5.1 Equivalent RLC Model 370
9.5.2 Time-Stepping Algorithm 371
9.5.3 Extension to the Three-Dimensional Case, Including Loss 373
9.5.4 Illustrative Numerical Results 374
9.5.5 Comparison of Computer Resources 375
9.6 Auxiliary Differential Equation Method, Nonlinear Dispersive Material Case 376
9.6.1 Strategy 376
9.6.2 Contribution of the Linear Debye Polarization 377
9.6.3 Contribution of the Linear Lorentz Polarization 377
9.6.4 Contributions of the Third-Order Nonlinear Polarization 378
9.6.5 Electric Field Update 380
9.6.6 Illustrative Numerical Results for Temporal Solitons 381
9.6.7 Illustrative Numerical Results for Spatial Solitons 383
9.7 Auxiliary Differential Equation Method, Macroscopic Modeling of Saturable, Dispersive Optical Gain Materials 387
9.7.1 Theory 387
9.7.2 Validation Studies 390
9.8 Auxiliary Differential Equation Method, Modeling of Lasing Action in a Four-Level Two-Electron Atomic System 394
9.8.1 Quantum Physics Basis 394
9.8.2 Coupling to Maxwell's Equations 398
9.8.3 Time-Stepping Algorithm 398
9.8.4 Illustrative Results 400
9.9 Summary and Conclusions 402
References 404
Problems 405
Projects 406

10 Local Subcell Models of Fine Geometrical Features *Allen Taflove, Malgorzata Celuch-Marcysiak, and Susan Hagness* **407**

10.1 Introduction 407
10.2 Basis of Contour-Path FDTD Modeling 408
10.3 The Simplest Contour-Path Subcell Models 408
10.3.1 Diagonal Split-Cell Model for PEC Surfaces 410
10.3.2 Average Properties Model for Material Surfaces 410
10.4 The Contour-Path Model of the Narrow Slot 411
10.5 The Contour-Path Model of the Thin Wire 415
10.6 Locally Conformal Models of Curved Surfaces 420
10.6.1 Yu-Mittra Technique for PEC Structures 420
10.6.2 Illustrative Results for PEC Structures 421
10.6.3 Yu-Mittra Technique for Material Structures 424
10.7 Maloney-Smith Technique for Thin Material Sheets 427
10.7.1 Basis 427
10.7.2 Illustrative Results 430

10.8 Surface Impedance 432
10.8.1 The Monochromatic SIBC 434
10.8.2 Convolution-Based Models of the Frequency-Dependent SIBC 436
10.8.3 Equivalent-Circuit Model of the Frequency-Dependent SIBC 442
10.8.4 Sources of Error 445
10.8.5 Discussion 446
10.9 Thin Coatings on a PEC Surface 447
10.9.1 Method of Lee et al. 447
10.9.2 Method of Kärkkäinen 450
10.10 Relativistic Motion of PEC Boundaries 450
10.10.1 Basis 451
10.10.2 Illustrative Results 454
10.11 Summary and Discussion 458
References 458
Selected Bibliography 460
Projects 461

11 Nonuniform Grids, Nonorthogonal Grids, Unstructured Grids, and Subgrids **463**
Stephen Gedney, Faiza Lansing, and Nicolas Chavannes

11.1 Introduction 463
11.2 Nonuniform Orthogonal Grids 464
11.3 Locally Conformal Grids, Globally Orthogonal 471
11.4 Global Curvilinear Coordinates 471
11.4.1 Nonorthogonal Curvilinear FDTD Algorithm 471
11.4.2 Stability Criterion 477
11.5 Irregular Nonorthogonal Structured Grids 480
11.6 Irregular Nonorthogonal Unstructured Grids 486
11.6.1 Generalized Yee Algorithm 487
11.6.2 Inhomogeneous Media 491
11.6.3 Practical Implementation of the Generalized Yee Algorithm 493
11.7 A Planar Generalized Yee Algorithm 494
11.7.1 Time-Stepping Expressions 495
11.7.2 Projection Operators 496
11.7.3 Efficient Time-Stepping Implementation 498
11.7.4 Modeling Example: 32-GHz Wilkinson Power Divider 499
11.8 Cartesian Subgrids 501
11.8.1 Geometry 502
11.8.2 Time-Stepping Scheme 503
11.8.3 Spatial Interpolation 504
11.8.4 Numerical Stability Considerations 505
11.8.5 Reflection from the Interface of the Primary Grid and Subgrid 505
11.8.6 Illustrative Results: Helical Antenna on Generic Cellphone at 900 MHz 508
11.8.7 Computational Efficiency 510
11.9 Summary and Conclusions 510
References 511
Problems 514
Projects 515

12 Bodies of Revolution

Thomas Jurgens, Jeffrey Blaschak, and Gregory Saewert **517**

12.1 Introduction 517
12.2 Field Expansion 517
12.3 Difference Equations for Off-Axis Cells 519
12.3.1 Ampere's Law Contour Path Integral to Calculate e_r 519
12.3.2 Ampere's Law Contour Path Integral to Calculate e_ϕ 521
12.3.3 Ampere's Law Contour Path Integral to Calculate e_z 523
12.3.4 Difference Equations 525
12.3.5 Surface-Conforming Contour Path Integrals 528
12.4 Difference Equations for On-Axis Cells 529
12.4.1 Ampere's Law Contour Path Integral to Calculate e_z on the z-Axis 529
12.4.2 Ampere's Law Contour Path Integral to Calculate e_ϕ on the z-Axis 532
12.4.3 Faraday's Law Calculation of h_r on the z-Axis 534
12.5 Numerical Stability 535
12.6 PML Absorbing Boundary Condition 536
12.6.1 BOR-FDTD Background 536
12.6.2 Extension of PML to the General BOR Case 537
12.6.3 Examples 543
12.7 Application to Particle Accelerator Physics 543
12.7.1 Definitions and Concepts 545
12.7.2 Examples 547
12.8 Summary 550
References 550
Problems 551
Projects 552

13 Periodic Structures

James Maloney and Morris Kesler **553**

13.1 Introduction 553
13.2 Review of Scattering from Periodic Structures 555
13.3 Direct Field Methods 559
13.3.1 Normal Incidence Case 559
13.3.2 Multiple Unit Cells for Oblique Incidence 560
13.3.3 Sine-Cosine Method 562
13.3.4 Angled-Update Method 563
13.4 Introduction to the Field-Transformation Technique 567
13.5 Multiple-Grid Approach 571
13.5.1 Formulation 571
13.5.2 Numerical Stability Analysis 573
13.5.3 Numerical Dispersion Analysis 574
13.5.4 Lossy Materials 575
13.5.5 Lossy Screen Example 577
13.6 Split-Field Method, Two Dimensions 578
13.6.1 Formulation 578
13.6.2 Numerical Stability Analysis 580
13.6.3 Numerical Dispersion Analysis 581
13.6.4 Lossy Materials 582
13.6.5 Lossy Screen Example 583

13.7 Split-Field Method, Three Dimensions 583
13.7.1 Formulation 584
13.7.2 Numerical Stability Analysis 589
13.7.3 UPML Absorbing Boundary Condition 590
13.8 Application of the Periodic FDTD Method 594
13.8.1 Electromagnetic Bandgap Structures 595
13.8.2 Frequency-Selective Surfaces 597
13.8.3 Antenna Arrays 597
13.9 Summary and Conclusions 603
Acknowledgments 603
References 603
Projects 605

14 Antennas
James Maloney, Glenn Smith, Eric Thiele, Om Gandhi, Nicolas Chavannes, and Susan Hagness **607**
14.1 Introduction 607
14.2 Formulation of the Antenna Problem 607
14.2.1 Transmitting Antenna 607
14.2.2 Receiving Antenna 609
14.2.3 Symmetry 610
14.2.4 Excitation 611
14.3 Antenna Feed Models 612
14.3.1 Detailed Modeling of the Feed 613
14.3.2 Simple Gap Feed Model for a Monopole Antenna 614
14.3.3 Improved Simple Feed Model 617
14.4 Near-to-Far-Field Transformations 621
14.4.1 Use of Symmetry 621
14.4.2 Time-Domain Near-to-Far-Field Transformation 622
14.4.3 Frequency-Domain Near-to-Far-Field Transformation 624
14.5 Plane-Wave Source 625
14.5.1 Effect of an Incremental Displacement of the Surface Currents 625
14.5.2 Effect of an Incremental Time Shift 627
14.5.3 Relation to Total-Field / Scattered-Field Lattice Zoning 628
14.6 Case Study I: The Standard-Gain Horn 628
14.7 Case Study II: The Vivaldi Slotline Array 634
14.7.1 Background 634
14.7.2 The Planar Element 635
14.7.3 The Vivaldi Pair 637
14.7.4 The Vivaldi Quad 639
14.7.5 The Linear Phased Array 640
14.7.6 Phased-Array Radiation Characteristics Indicated by the FDTD Modeling 641
14.7.7 Active Impedance of the Phased Array 644
14.8 Near-Field Simulations 647
14.8.1 Generic 900-MHz Cellphone Handset in Free Space 647
14.8.2 900-MHz Dipole Antenna Near a Layered Bone-Brain Half-Space 649
14.8.3 840-MHz Dipole Antenna Near a Rectangular Brain Phantom 650
14.8.4 900-MHz Infinitesimal Dipole Antenna Near a Spherical Brain Phantom 650
14.8.5 1.9-GHz Half-Wavelength Dipole Near a Spherical Brain Phantom 652
14.9 Case Study III: The Motorola T250 Tri-Band Phone 653
14.9.1 FDTD Phone Model 654
14.9.2 Measurement Procedures 656
14.9.3 Free-Space Near-Field Investigations and Assessment of Design Capabilities 656

14.9.4 Performance in Loaded Conditions (SAM and MRI-Based Human Head Model) 657
14.9.5 Radiation Performance in Free Space and Adjacent to the SAM Head 659
14.9.6 Computational Requirements 661
14.9.7 Overall Assessment 661
14.10 Selected Additional Applications 661
14.10.1 Use of Electromagnetic Bandgap Materials 662
14.10.2 Ground-Penetrating Radar 663
14.10.3 Antenna-Radome Interaction 667
14.10.4 Biomedical Applications of Antennas 669
14.11 Summary and Conclusions 671
References 671
Projects 676

15 High-Speed Electronic Circuits with Active and Nonlinear Components
Melinda Piket-May, Wojciech Gwarek, Tzong-Lin Wu, Bijan Houshmand, Tatsuo Itoh, and Jamesina Simpson **677**

15.1 Introduction 677
15.2 Basic Circuit Parameters for TEM Striplines and Microstrips 679
15.2.1 Transmission Line Parameters 679
15.2.2 Impedance 680
15.2.3 *S*-Parameters 680
15.2.4 Differential Capacitance 681
15.2.5 Differential Inductance 682
15.3 Lumped Inductance Due to a Discontinuity 682
15.3.1 Flux / Current Definition 684
15.3.2 Fitting $Z(\omega)$ or $S(\omega)$ to an Equivalent Circuit 684
15.3.3 Discussion: Choice of Methods 685
15.4 Inductance of Complex Power-Distribution Systems 685
15.4.1 Method Description 685
15.4.2 Example: Multiplane Meshed Printed-Circuit Board 687
15.4.3 Discussion 688
15.5 Parallel Coplanar Microstrips 688
15.6 Multilayered Interconnect Modeling 690
15.7 *S*-Parameter Extraction for General Waveguides 692
15.8 Digital Signal Processing and Spectrum Estimation 694
15.8.1 Prony's Method 695
15.8.2 Autoregressive Models 697
15.8.3 Padé Approximation 702
15.9 Modeling of Lumped Circuit Elements 706
15.9.1 FDTD Formulation Extended to Circuit Elements 706
15.9.2 The Resistor 708
15.9.3 The Resistive Voltage Source 708
15.9.4 The Capacitor 709
15.9.5 The Inductor 711
15.9.6 The Arbitrary Two-Terminal Linear Lumped Network 711
15.9.7 The Diode 714
15.9.8 The Bipolar Junction Transistor 715
15.10 Direct Linking of FDTD and SPICE 717
15.10.1 Basic Idea 718
15.10.2 Norton Equivalent Circuit "Looking Into" the FDTD Space Lattice 719
15.10.3 Thevenin Equivalent Circuit "Looking Into" the FDTD Space Lattice 721

15.11 Case Study: A 6-GHz MESFET Amplifier Model 723
15.11.1 Large-Signal Nonlinear Model 723
15.11.2 Amplifier Configuration 725
15.11.3 Analysis of the Circuit without the Packaging Structure 726
15.11.4 Analysis of the Circuit with the Packaging Structure 728
15.12 Emerging Topic: Wireless High-Speed Digital Interconnects Using Defect-Mode Electromagnetic Bandgap Waveguides 731
15.12.1 Stopband of the Defect-Free Two-Dimensional EBG Structure 732
15.12.2 Passband of the Two-Dimensional EBG Structure with Waveguiding Defect 732
15.12.3 Laboratory Experiments and Supporting FDTD Modeling 734
15.13 Summary and Conclusions 736
Acknowledgments 737
References 737
Selected Bibliography 740
Projects 741

16 Photonics
Geoffrey Burr, Susan Hagness, and Allen Taflove **743**

16.1 Introduction 743
16.2 Introduction to Index-Contrast Guided-Wave Structures 743
16.3 FDTD Modeling Issues 744
16.3.1 Optical Waveguides 744
16.3.2 Material Dispersion and Nonlinearities 747
16.4 Laterally Coupled Microcavity Ring Resonators 747
16.4.1 Modeling Considerations: Two-Dimensional FDTD Simulations 748
16.4.2 Coupling to Straight Waveguides 750
16.4.3 Coupling to Curved Waveguides 750
16.4.4 Elongated Ring Designs ("Racetracks") 752
16.4.5 Resonances of the Circular Ring 752
16.5 Laterally Coupled Microcavity Disk Resonators 756
16.5.1 Resonances 756
16.5.2 Suppression of Higher-Order Radial Whispering-Gallery Modes 760
16.6 Vertically Coupled Racetrack 761
16.7 Introduction to Distributed Bragg Reflector Devices 765
16.8 Application to Vertical-Cavity Surface-Emitting Lasers 765
16.8.1 Passive Studies 766
16.8.2 Active Studies: Application of the Classical Gain Model 767
16.8.3 Application of a New Semiclassical Gain Model 769
16.9 Quasi-One-Dimensional DBR Structures 770
16.10 Introduction to Photonic Crystals 772
16.11 Calculation of Band Structure 774
16.11.1 The "Order-N" Method 775
16.11.2 Frequency Resolution 778
16.11.3 Filter Diagonalization Method 780
16.11.4 The Triangular Photonic Crystal Lattice 782
16.11.5 Sources of Error and Their Mitigation 784
16.12 Calculation of Mode Patterns 787
16.13 Variational Approach 790
16.14 Modeling of Defect-Mode Photonic Crystal Waveguides 791
16.14.1 Band Diagram of a Photonic Crystal Slab 793
16.14.2 Band Diagram of a Photonic Crystal Waveguide 795

16.14.3 Intrinsic Loss in Photonic Crystal Waveguides 798
16.14.4 Transmission in Photonic Crystal Waveguides 803
16.14.5 Aperiodic Photonic-Crystal Waveguides 806
16.14.6 Photonic Crystal Waveguide Extrinsic Scattering Loss from the Green Function 806
16.15 Modeling of Photonic Crystal Resonators 807
16.16 Modeling Examples of Photonic Crystal Resonators 810
16.16.1 Electrically Driven Microcavity Laser 810
16.16.2 Photonic Crystal Cross-Waveguide Switch 812
16.17 Introduction to Frequency Conversion in Second-Order Nonlinear Optical Materials 813
16.18 PSTD-4 Algorithm 813
16.19 Extension to Second-Order Nonlinear Media 814
16.20 Application to a Nonlinear Waveguide with a QPM Grating 814
16.21 Application to Nonlinear Photonic Crystals 817
16.22 Introduction to Nanoplasmonic Devices 820
16.23 FDTD Modeling Considerations 820
16.24 FDTD Modeling Applications 821
16.25 Introduction to Biophotonics 822
16.26 FDTD Modeling Applications 822
16.26.1 Vertebrate Retinal Rod 822
16.26.2 Precancerous Cervical Cells 824
16.26.3 Sensitivity of Backscattering Signatures to Nanometer-Scale Cellular Changes 827
16.27 PSTD Modeling Application to Tissue Optics 828
16.28 Summary 830
Acknowledgments 830
References 830

17 Advances in PSTD Techniques **847**
Qing Liu and Gang Zhao
17.1 Introduction 847
17.2 Approximation of Derivatives 847
17.2.1 Derivative Matrix for the Second-Order Finite-Difference Method 848
17.2.2 Derivative Matrices for Fourth-Order and N'th-Order Finite-Difference Methods 849
17.2.3 Trigonometric Interpolation and FFT Method 850
17.2.4 Nonperiodic Functions and Chebyshev Method 851
17.3 Single-Domain Fourier PSTD Method 854
17.3.1 Approximation of Spatial Derivatives 855
17.3.2 Numerical Stability and Dispersion 856
17.4 Single-Domain Chebyshev PSTD Method 857
17.4.1 Spatial and Temporal Grids 857
17.4.2 Maxwell's Equations in Curvilinear Coordinates 858
17.4.3 Spatial Derivatives 860
17.4.4 Time-Integration Scheme 861
17.5 Multidomain Chebyshev PSTD Method 861
17.5.1 Subdomain Spatial Derivatives and Time Integration 862
17.5.2 Subdomain Patching by Characteristics 863
17.5.3 Subdomain Patching by Physical Conditions 864
17.5.4 Filter Design for Corner Singularities 864
17.5.5 Multidomain PSTD Results for 2.5-Dimensional Problems 866
17.5.6 Multidomain PSTD Results for Three-Dimensional Problems 868
17.6 Penalty Method for Multidomain PSTD Algorithm 868

17.7 Discontinuous Galerkin Method for PSTD Boundary Patching 873
17.7.1 Weak Form of Maxwell's Equations 873
17.7.2 Space Discretization and Domain Transformation 873
17.7.3 Mass Matrix and Stiffness Matrix 874
17.7.4 Flux on the Boundary 876
17.7.5 Numerical Results for DG-PSTD Method 876
17.8 Summary and Conclusions 879
Appendix 17A: Coefficients for the Five-Stage, Fourth-Order Runge-Kutta Method 879
References 880

18 Advances in Unconditionally Stable Techniques **883**
Hans De Raedt
18.1 Introduction 883
18.2 General Framework 883
18.3 Matrix-Exponential Concepts 884
18.4 Product-Formula Approach 887
18.4.1 The Classic Yee Algorithm as a Particular Realization 887
18.4.2 The ADI Method as a Second Realization 888
18.4.3 Unconditionally Stable Algorithms: Real-Space Approach 889
18.4.4 Unconditionally Stable Algorithms: Fourier-Space Approach 891
18.5 Chebyshev Polynomial Algorithm 892
18.6 Extension to Linear Dispersive Media 895
18.7 Extension to Perfectly Matched Layer Absorbing Boundary Conditions 898
18.8 Summary 899
Appendix 18A: Some Technical Details 900
Appendix 18B: Stability Analysis of Equation (18.17) 902
Appendix 18C: Stability Analysis of Equation (18.19) 904
References 904
Projects 905

19 Advances in Hybrid FDTD-FE Techniques **907**
Thomas Rylander, Fredrik Edelvik, Anders Bondeson, and Douglas Riley
19.1 Introduction 907
19.2 Time-Domain Finite Elements 910
19.2.1 Coupled Curl Equations 910
19.2.2 Wave Equation 913
19.2.3 Equivalences Between Finite Elements and FDTD 917
19.3 Tetrahedral, Hexahedral (Brick), and Pyramidal Zeroth-Order Edge and Facet Elements 918
19.3.1 Tetrahedral Finite Elements 919
19.3.2 Hexahedral (Brick) Finite Elements 921
19.3.3 Pyramidal Finite Elements 922
19.4 Stable Hybrid FDTD-FE Interface 924
19.4.1 Spatial Discretization 924
19.4.2 Time-Stepping on a Hybrid Space Lattice 927
19.4.3 Generalized Newmark Scheme 928
19.4.4 Proof of Stability 929
19.4.5 Alternative Time-Stepping Schemes 930
19.4.6 Extensions of the Hybrid FDTD-FE Concept 931
19.4.7 Reflection at the Interface of FDTD and FE Regions of a Hybrid Space Lattice 931
19.4.8 Scattering from the PEC Sphere 933
19.5 Mesh-Generation Approaches 935

19.6 Subcell Wire and Slot Algorithms for Time-Domain Finite Elements 936
19.6.1 Modeling Thin Wires 936
19.6.2 Modeling Thin Slots 939
19.6.3 Numerical Results for Thin Wires and Slots 941
19.7 Application to Advanced Scattering and Radiation Problems 943
19.7.1 Monostatic RCS of the NASA Almond 943
19.7.2 Bistatic RCS of the Saab *Trainer* Aircraft 945
19.7.3 Input Impedance of the Four-Arm Sinuous Antenna 948
19.8 Summary 949
Acknowledgments 950
References 950

20 Advances in Hardware Acceleration for FDTD
Ryan Schneider, Sean Krakiwsky, Laurence Turner, and Michal Okoniewski **955**
20.1 Introduction 955
20.2 Background Literature 956
20.3 Fundamental Design Considerations 957
20.4 Conceptual Massively Parallel FPGA Implementation 958
20.5 Case Study of Using the FPGA as a Coprocessor 962
20.6 Performance of Custom Hardware Implementations 964
20.7 Fundamentals of Graphics Processor Units 965
20.7.1 Overview 965
20.7.2 Graphics Pipeline 965
20.7.3 Memory Interface 967
20.7.4 Programmable Fragment and Vertex Processors 968
20.8 Implementing FDTD on a Graphics Processor Unit 969
20.8.1 Initialization 969
20.8.2 Electric and Magnetic Field Updates 970
20.8.3 Boundaries 972
20.8.4 Source Excitation 974
20.8.5 Archiving Observation Nodes 975
20.8.6 Multipass Rendering 975
20.8.7 Display 977
20.9 Performance Measurements of the GPU Accelerator 977
20.10 Summary and Conclusions 978
References 978

Acronyms and Common Symbols **981**

About the Authors **985**

Index **997**

Preface to the Third Edition

The first and second editions of this book were published in 1995 and 2000, respectively. We are gratified with their high level of use by both the university and industrial-research communities. The second edition is often the text in senior year and graduate electrical engineering courses in computational electrodynamics, and both editions are frequently cited in refereed journal papers as primary background references for FDTD methods and applications.

This new third edition is extensively revised and expanded. We have had two primary goals in this regard. First, we have worked to update the book's discussions of FDTD theory and applications to keep pace with the continuing, rapid changes in these areas since 2000. 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 further 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.

New Material: Advances in FDTD Theory and Numerical Algorithms

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

- New invited Chapter 17, "Advances in PSTD Techniques," by Qing Liu and Gang Zhao, who have pioneered the theory and application of pseudospectral time-domain computational solutions of Maxwell's equations;

- New invited Chapter 18, "Advances in Unconditionally Stable Techniques," by Hans De Raedt, who, within the framework of the matrix-exponential technique, has pioneered the unification of existing algorithms, as well as the synthesis of completely novel algorithms, for unconditionally stable computational solutions of the time-dependent Maxwell's equations;

- New invited Chapter 19, "Advances in Hybrid FDTD-FE Techniques," by Thomas Rylander, Fredrik Edelvik, Anders Bondeson, and Douglas Riley, who have pioneered the development and application of provably stable hybrids of FDTD and finite-element time-domain techniques;

- New invited Chapter 20, "Advances in Hardware Acceleration for FDTD," by Ryan Schneider, Sean Krakiwsky, Laurence Turner, and Michal Okoniewski, who have led the development of computer hardware / software that promise one order-of-magnitude speedups of FDTD solutions implemented on normal laboratory computers;

- New invited Section 5.9 in Chapter 5, by John Schneider, describing his development of advanced numerical dispersion-compensation techniques for the total-field / scattered-field FDTD wave-source condition;

- New invited Sections 7.7 and 7.9 to 7.11 in Chapter 7 by Stephen Gedney, describing the theory, numerical implementation, and illustrative results of his convolutional PML absorbing boundary condition, the most effective such technique yet for terminating open-region FDTD computational modeling spaces;
- New invited Section 8.7 in Chapter 8 by Xu Li, who has solved the puzzle of why FDTD had previously not properly modeled the backscattering of certain weakly backscattering objects;
- New Sections 9.2.3 and 9.4.3 in Chapter 9 on Drude media, so important for FDTD modeling of metals at optical frequencies;
- New invited Section 9.5 in Chapter 9 by Wojciech Gwarek, who has pioneered the efficient circuit model of linear magnetized ferrites in FDTD simulations;
- New invited Sections 9.6.1 to 9.6.5 in Chapter 9 by Masafumi Fujii, who has led the development of improved FDTD algorithms for nonlinear dispersive media;
- New invited Section 9.8 in Chapter 9 by Shih-Hui Chang on advances in FDTD modeling of quantum-gain materials characterized by a four-level, two-electron atomic system constrained by the Pauli Exclusion Principle;
- New Sections 10.6.1 and 10.6.3 in Chapter 10 on the simple, robust, Yu-Mittra techniques for modeling curved surfaces comprised of either a perfect electric conductor or a dielectric material;
- New invited Section 10.8.3 in Chapter 10 by Malgorzata Celuch-Marcysiak on her simple, robust, ultrawideband equivalent-circuit model of the frequency-dependent skin effect;
- New invited Section 11.8 in Chapter 11 by Nicolas Chavannes on his robust subgridding technique that allows numerically stable local mesh refinement;
- New invited Section 15.7 in Chapter 15 by Wojciech Gwarek, who has pioneered efficient and accurate *S*-parameter extraction from FDTD models of general waveguides;
- New invited Section 15.9.6 in Chapter 15 by Tzong-Lin Wu, who has innovated an efficient FDTD subcell model of the arbitrary two-terminal linear lumped network;
- New invited Sections 16.10 to 16.16 in Chapter 16 by Geoffrey Burr, who has helped to lead the development of FDTD techniques for modeling photonic crystals.

New Material: Advances in FDTD Modeling Applications

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

- New invited Section 3.8 in Chapter 3 by Jamesina Simpson, who describes her FDTD models of around-the-world electromagnetic wave propagation at extremely low frequencies;
- New invited Section 14.9 in Chapter 14 by Nicolas Chavannes, who describes his detailed FDTD modeling case study of the electromagnetic wave performance characteristics of the Motorola T250 tri-band cellphone;
- New Section 14.10.4 in Chapter 14, which describes work by Susan Hagness' group on the application of FDTD modeling of ultrawideband radar techniques for the early-stage detection of breast cancer;
- New invited Section 15.12 in Chapter 15 by Jamesina Simpson, who describes her FDTD models of potential hyperspeed digital interconnects in circuit boards realized by defect-mode waveguides in electromagnetic bandgap structures;
- New Sections 16.17 to 16.21 in Chapter 16, which describes work by Susan Hagness' group in PSTD modeling of frequency conversion in second-order nonlinear optical materials, including photonic crystals;
- New invited Sections 16.22 to 16.24 in Chapter 16 by Geoffrey Burr, who describes recent FDTD modeling applications in nanoplasmonics;
- New Sections 16.25 to 16.27 in Chapter 16, which describe applications of FDTD and PSTD modeling in biophotonics, especially in advancing the detection of early stage cervical and colon cancer.

These new 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.

New Feature: A Web Site Dedicated to This Book

To supplement this third edition, we and the publisher have created a Web site where, with the proper personal identification number (PIN), instructors can download solutions to the homework problems. This PIN also enables downloads of color graphics, videos, and text updates / errata, as generated by the authors. For details, visit www.artechhouse.com or contact the publisher via email at artech@artechhouse.com. While subject to copyright protection, the color graphics, videos, and text updates / errata downloaded in this manner can be freely distributed by the course instructor to his or her students. We believe that this feature greatly enhances the usefulness of this book as an instructional tool.

Structuring University Courses Around This Book

In our respective teaching experiences at Northwestern University and the University of Wisconsin-Madison, we have found that the material covered in this third edition is most appropriate for senior-year undergraduate students who have already taken at least one course in

electromagnetics. Having said that, our personal experience is that even students without the first course in electromagnetics can access most of the material in the first seven chapters, assuming that they have a strong background in vector calculus and computer programming. However, such students would require supplemental assistance to understand the basis of Maxwell's equations.

When used in a semester-length, senior year undergraduate course (i.e., UW-Madison), there is sufficient time to cover the first ten chapters. This includes time for the students to write working FDTD codes in one and two dimensions, with absorbing boundary conditions and total-field / scattered-field grid zoning. When used in a quarter-length, senior year undergraduate course (i.e., Northwestern), there is sufficient time to cover the first seven chapters.

We recommend that the final ten chapters be covered in a second semester or quarter at the graduate level. Some of the advanced new material in Chapters 17, 18, and 19 is appropriate for a special-topics course at the graduate level.

Acknowledgments

In accomplishing this major rewrite and update of the second edition, we gratefully acknowledge all of our contributing chapter authors and coauthors. Their biographical sketches appear in the About the Authors section.

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, during this past year of hyperintensive effort. 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 *fourth edition.*

Allen Taflove, Evanston, Illinois
Susan C. Hagness, Madison, Wisconsin
June 2005

Chapter 1

Electrodynamics Entering the 21st Century

1.1 INTRODUCTION

Maxwell's partial differential equations of electrodynamics, formulated approximately 140 years ago, 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 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 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, 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 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* (EMP). EMP can be sufficiently intense to burn out electrical and electronic equipment on the Earth's surface that are 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 could 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 analyze electromagnetic wave penetration and coupling mechanisms into potential targets of HPM, and to design 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 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.

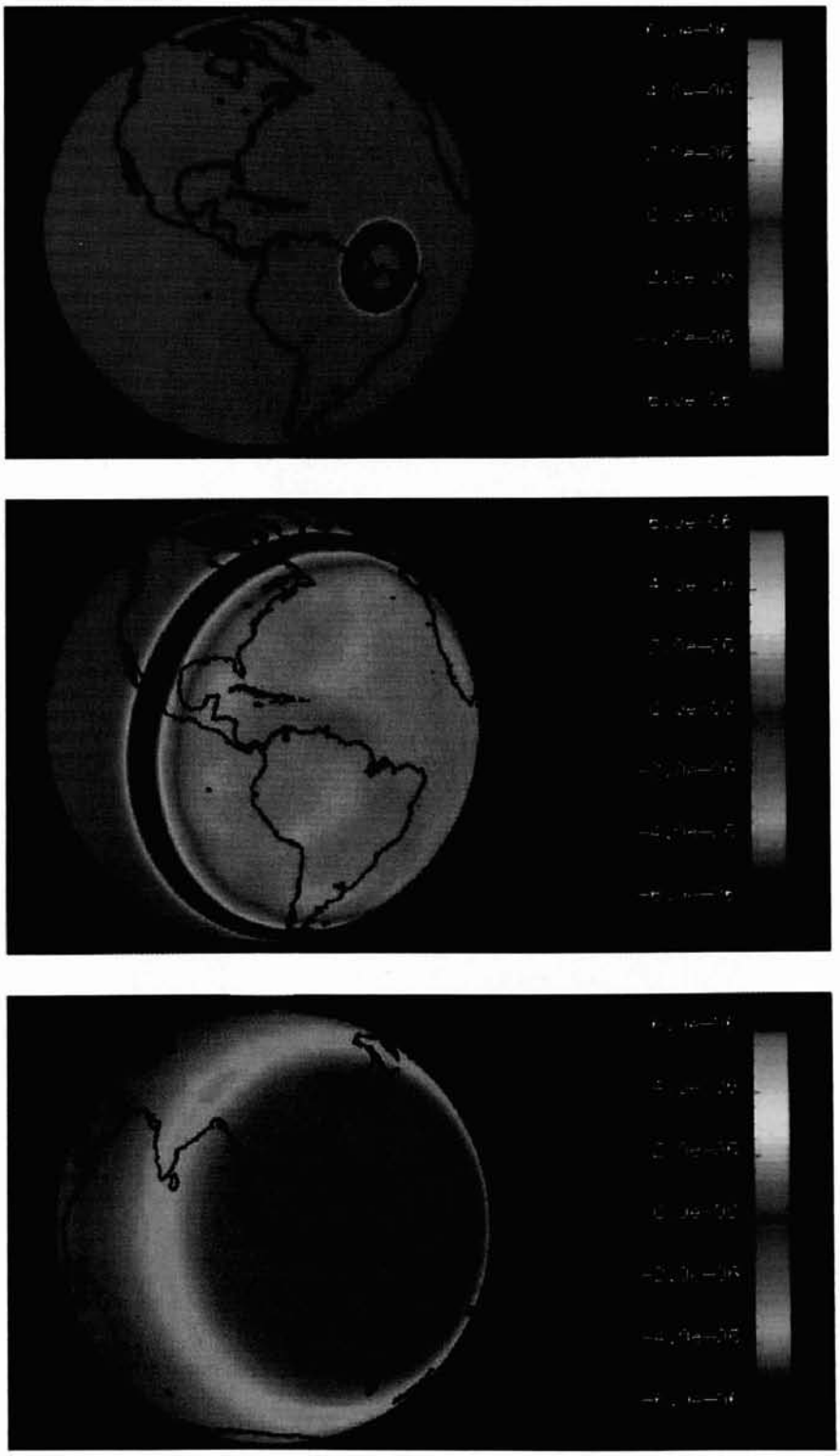


Fig. 1.1 Snapshot visualizations of the FDTD-computed global propagation of an ELF electromagnetic pulse generated by a vertical lightning strike off the coast of South America. All features of the lithosphere and atmosphere located within ±100 km of sea level are modeled in three dimensions with a resolution of ~ 40 × 40 × 5 km. See Chapter 3, Section 3.8 for a complete discussion.

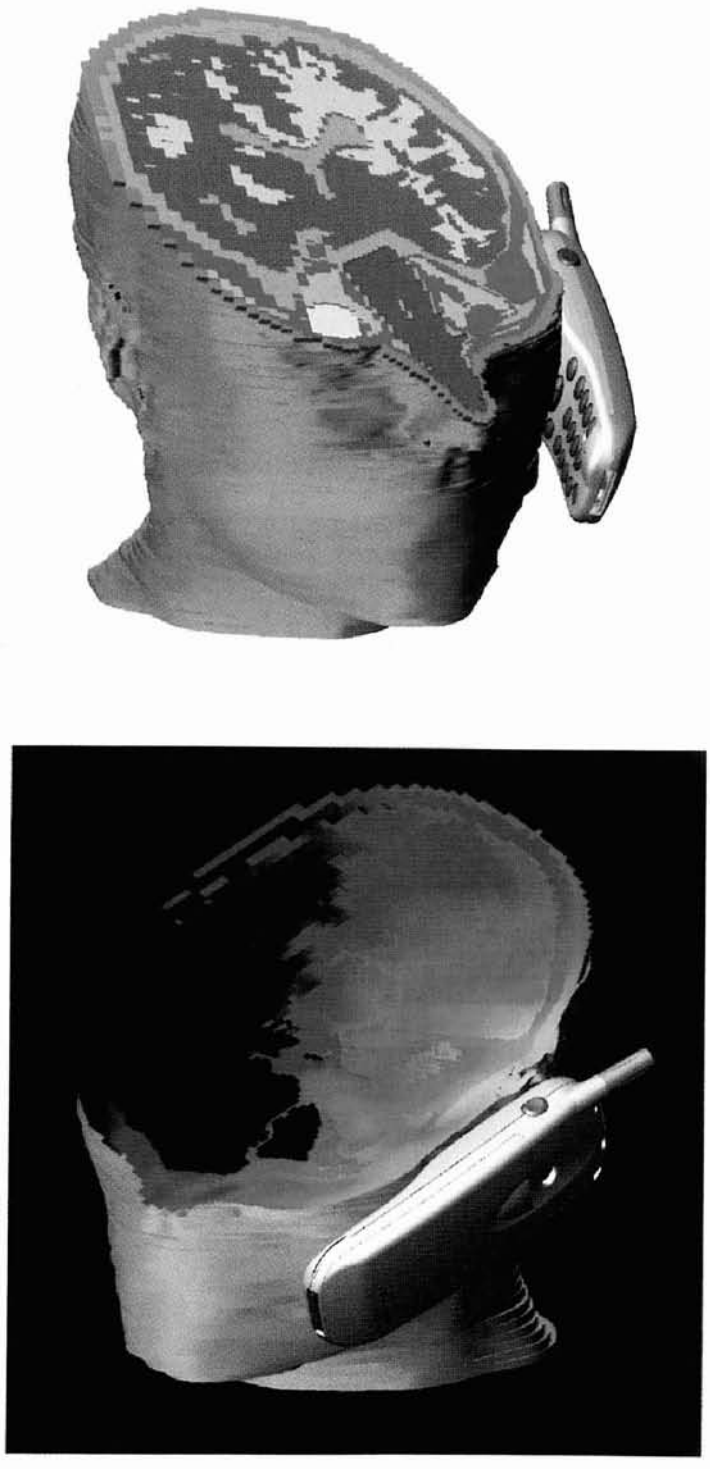


Fig. 1.2 High-resolution FDTD model of the Motorola T250 cellphone depositing radio-frequency power into a 15-tissue numerical model of the human head derived from magnetic resonance imaging (MRI) of a volunteer. The lattice-cell size of the phone model is as fine as 0.1 mm to properly render individual circuit board layers and the helical antenna in the FDTD mesh. The head model is comprised of 121 slices (1-mm thick in the ear region; 3-mm thick elsewhere) having a transverse spatial resolution of 0.2 mm. See Chapter 14, Section 14.9 for a complete discussion.

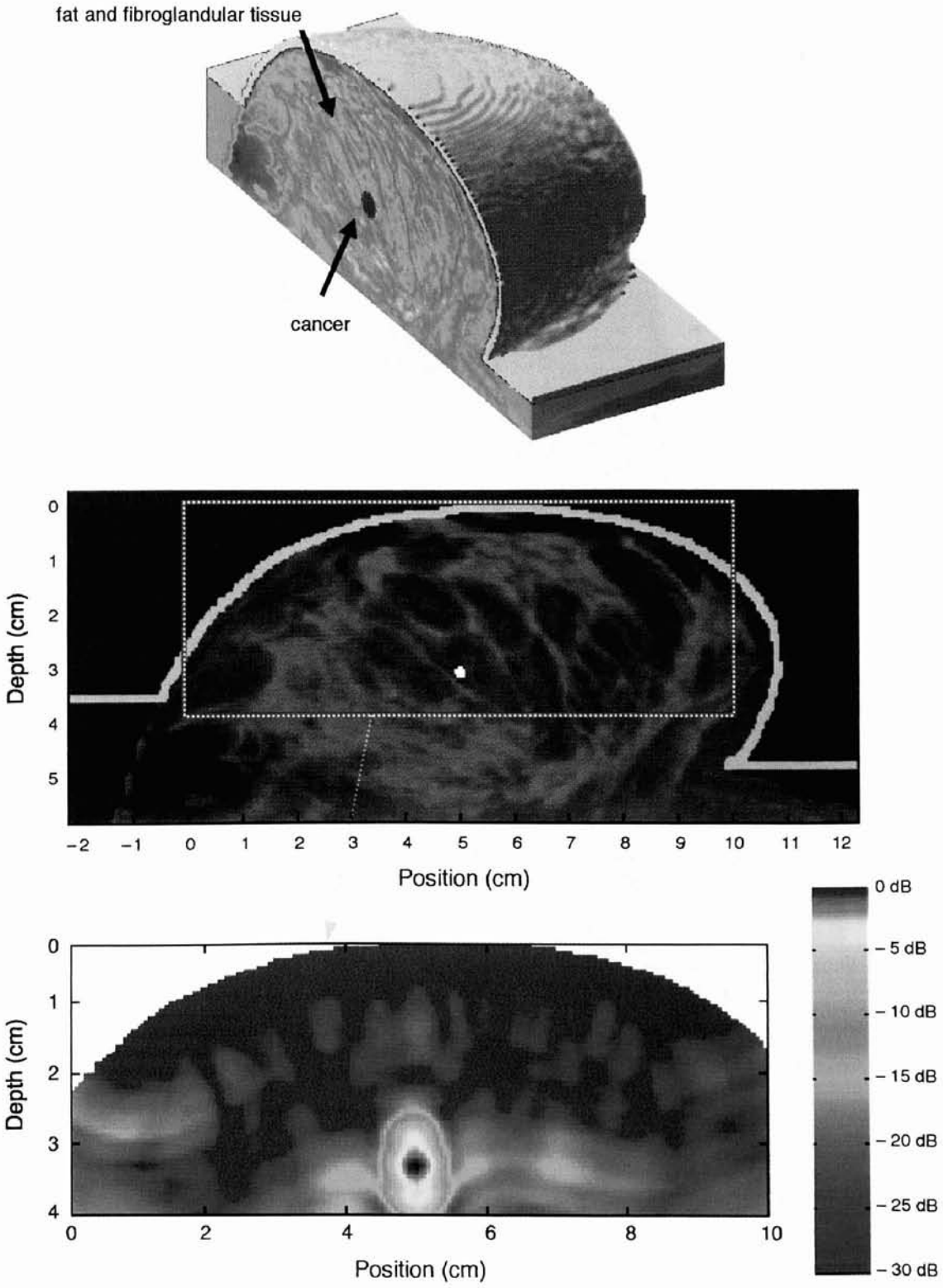
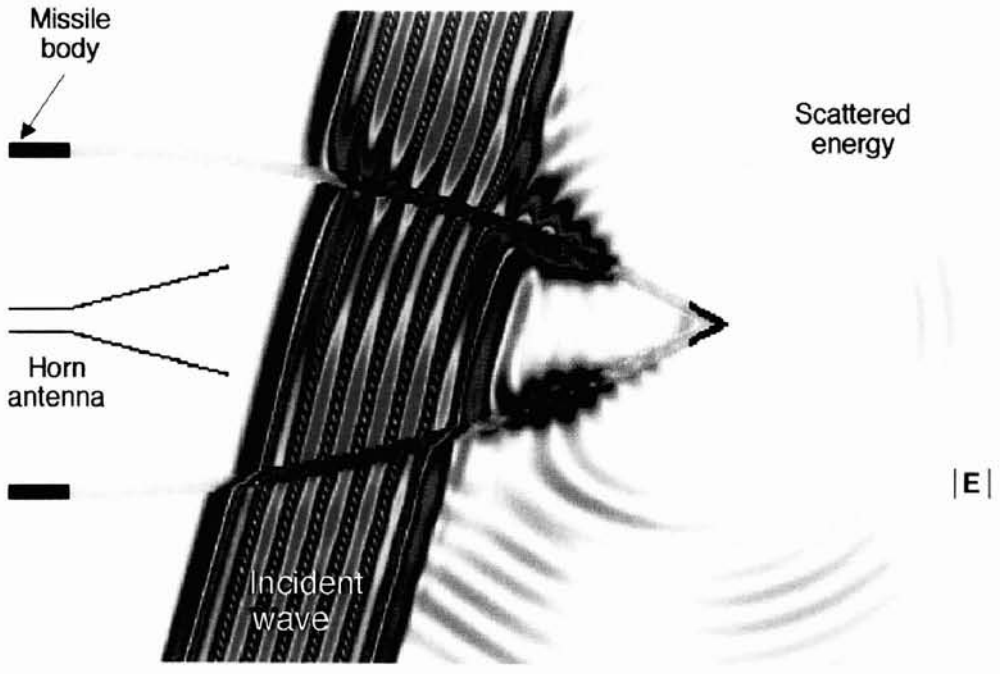
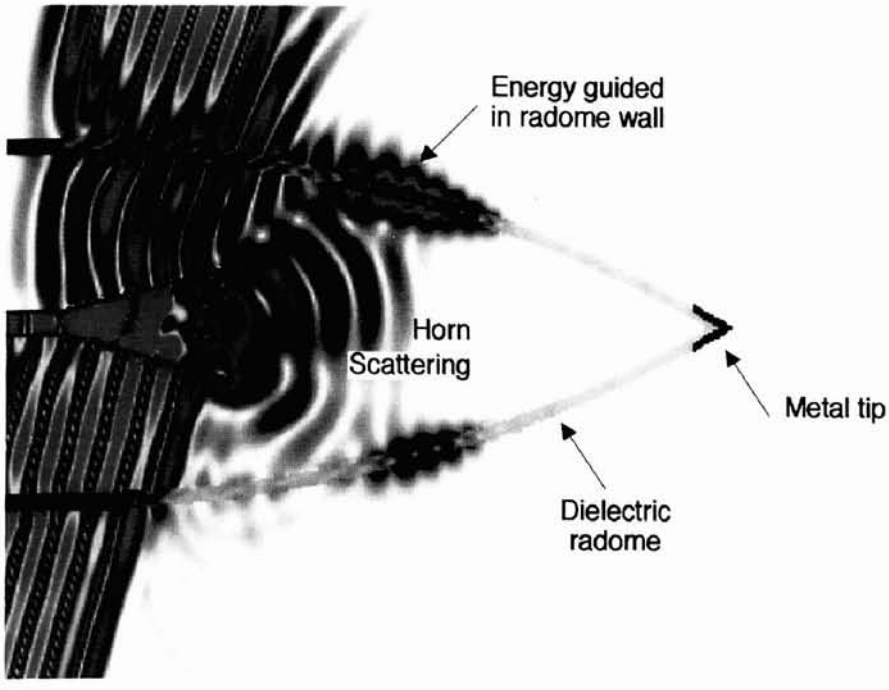


Fig. 1.3 FDTD simulation of ultrawideband microwave detection of a 2-mm-diameter malignant tumor embedded 3 cm within an MRI-derived numerical breast model. The cancer's signature is 15 to 30 dB (30 to 1,000 times) stronger than the clutter due to the surrounding normal tissues. *Adapted from:* Bond et al., *IEEE Trans. Antennas and Propagation*, 2003, pp. 1690–1705, © 2003 IEEE. See Chapter 14, Section 14.10.4 for a complete discussion.

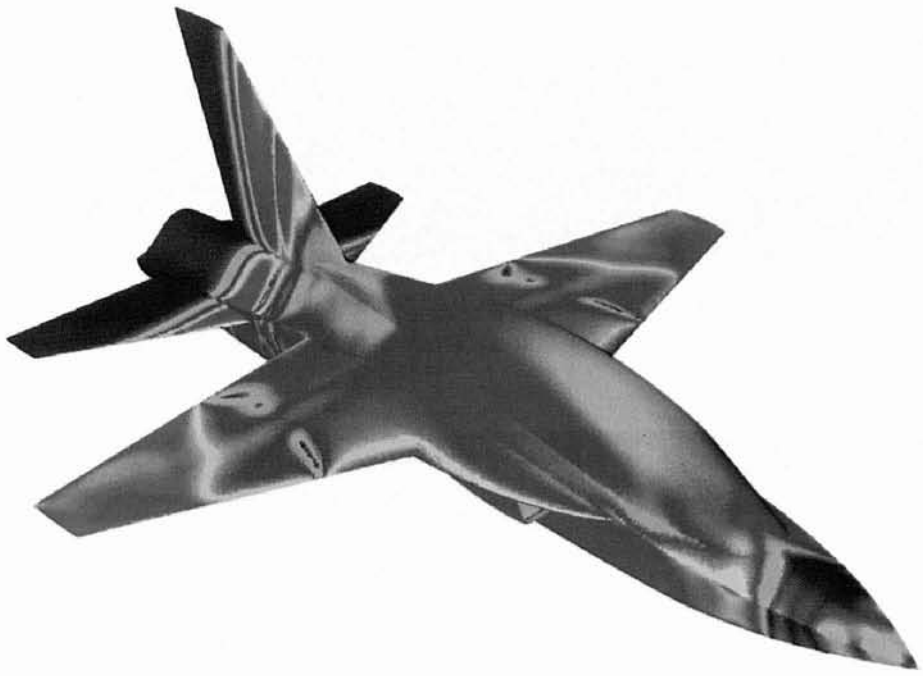


(a) Incident pulse has just entered the radome.



(b) Incident pulse has just encountered the horn antenna.

Fig. 1.4 Snapshot visualizations of the FDTD-computed interaction of a 10-GHz microwave pulse with a missile radome containing a horn antenna. The illuminating plane-wave pulse propagates from right to left at 15° from boresight. See Chapter 14, Section 14.10.3 for a complete discussion.

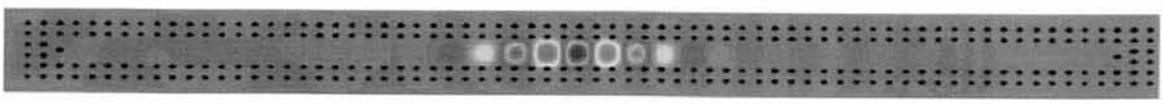


(a) Outer surface of aircraft.

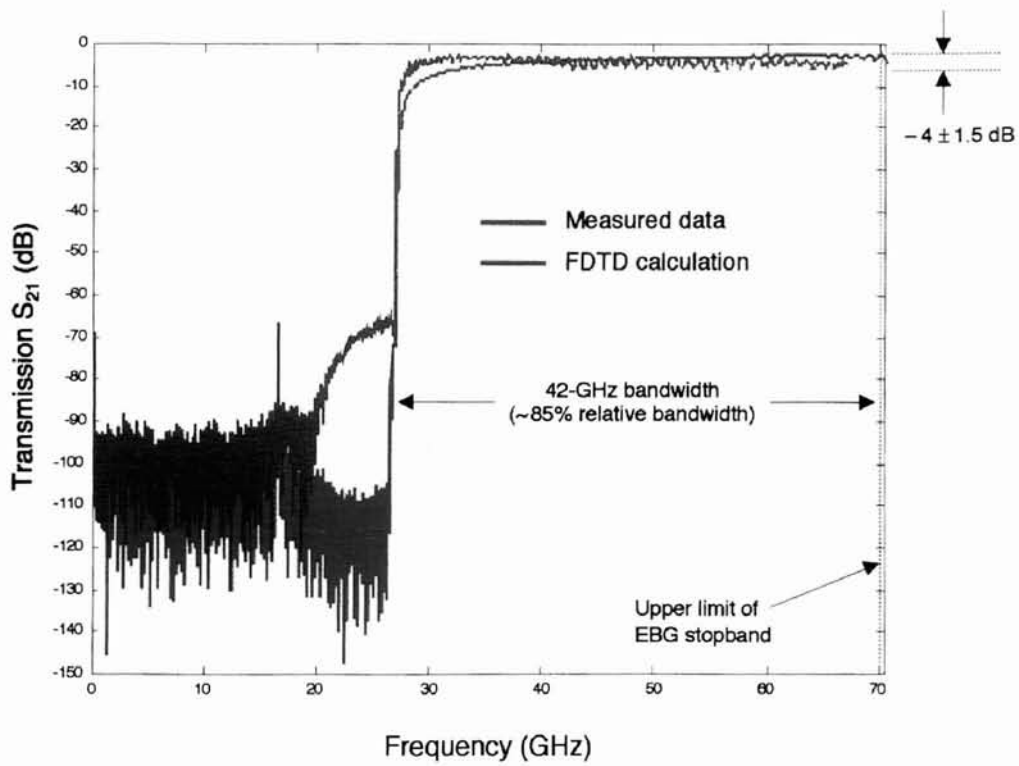
(b) Within a jet engine air inlet.

Fig. 1.5 Snapshot visualizations of the induced surface currents on the Saab *Trainer* calculated by the FDTD-FE hybrid technique for head-on illumination by a horizontally polarized Gaussian plane-wave pulse. The observation time is just as the pulse reaches the aircraft's tail. See Chapter 19, Section 19.7.2 for a complete discussion.

(a) Photograph of the laboratory test article, showing coaxial connectors probe-coupled to a double-row-defect electromagnetic bandgap (EBG) waveguide.

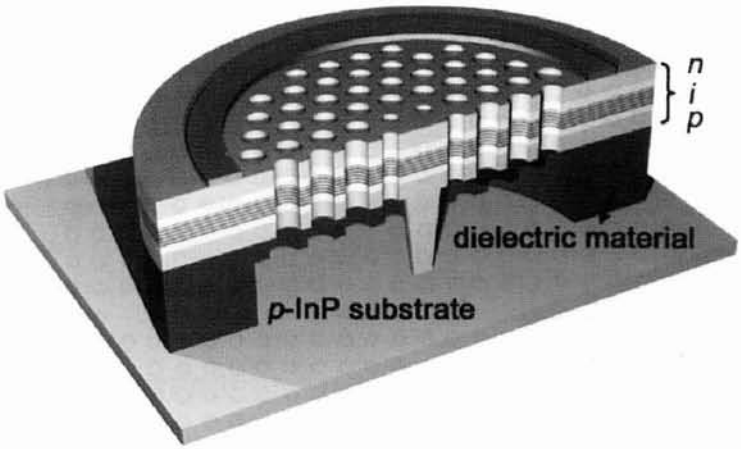


(b) Visualization of the FDTD-calculated propagation of a 20-ps Gaussian-envelope sinusoidal pulse along the EBG waveguide of (a).

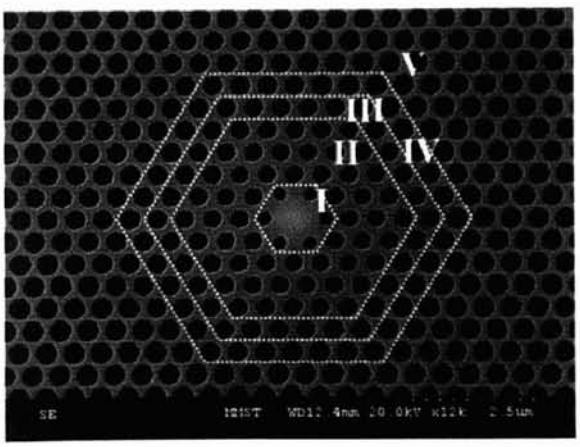


(c) Comparison of FDTD and measured results for the transmission S_{21} of the EBG waveguide of (a).

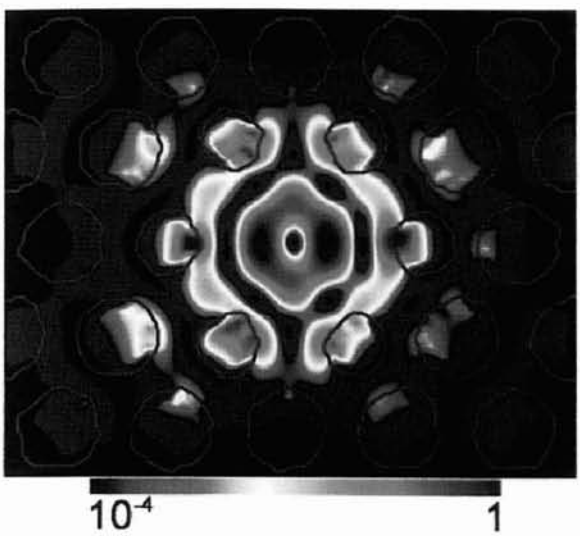
Fig. 1.6 FDTD-designed wireless digital interconnect using electromagnetic bandgap (EBG) technology, achieving a measured 42-GHz bandwidth (28 to 70 GHz). Via pin radius = 0.23 mm; via pin separation = 1.3 mm. *Source:* Simpson et al., *Proc. IEEE Antennas and Propagation Society Intl. Symp.*, © 2005 IEEE. See Chapter 15, Section 15.12 for a complete discussion.



(a)



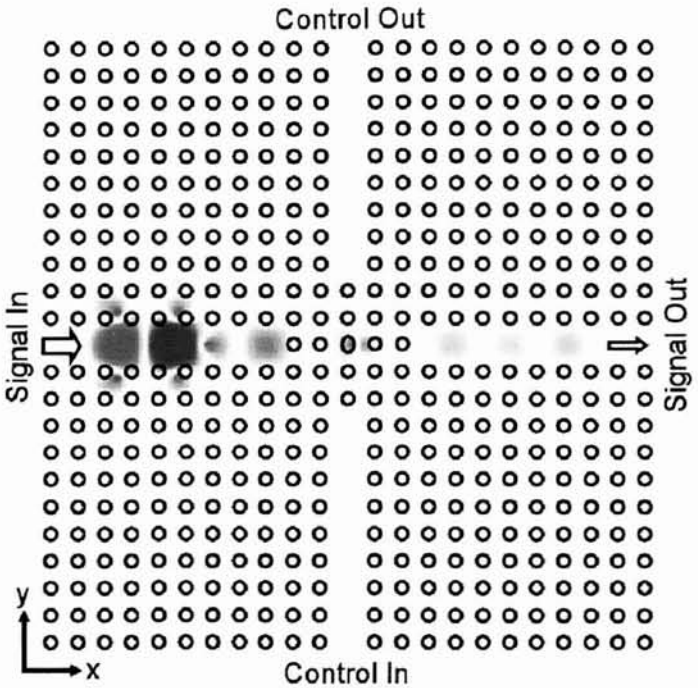
(b)



(c)

Fig. 1.7 Electrically driven, single-mode, low-threshold-current photonic crystal microlaser operating at room temperature: (a) schematic diagram; (b) top view of the fabricated sample; (c) FDTD-calculated E-field intensity profile of monopole mode (log scale). *Source:* Park et al., *Science*, Sept. 3, 2004, pp. 1444–1447. See Chapter 16, Section 16.16.1 for a complete discussion.

(a)



(b)

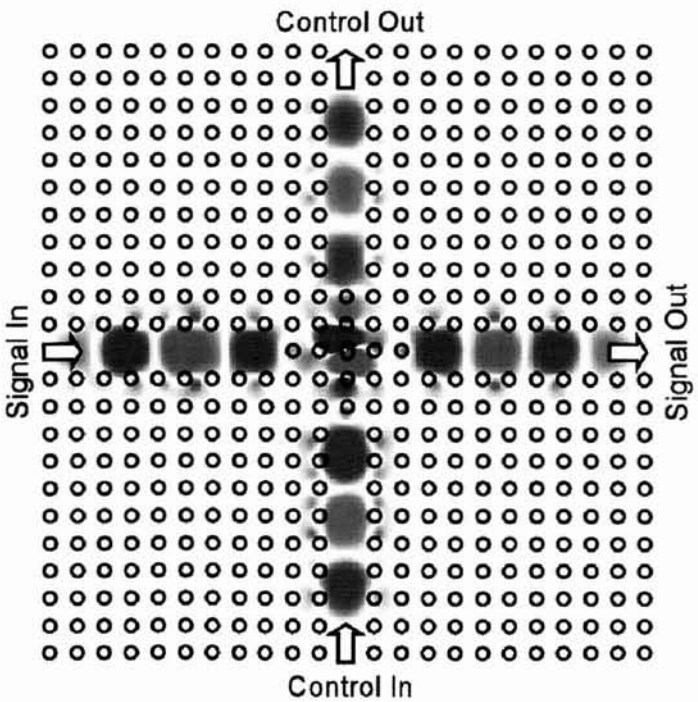


Fig. 1.8 FDTD-calculated electric-field distributions in a photonic crystal cross-waveguide switch: (a) control input is absent, and the signal output is low; (b) control input is present, and the signal output is high. Red and blue represent large positive or negative E-fields, respectively, with the same color scale used for both panels. Black circles indicate the positions of the dielectric rods in the photonic crystal. *Source:* Yanik et al., *Optics Letters*, 2003, pp. 2506–2508. See Chapter 16, Section 16.16.2 for a complete discussion.

While significant progress has been made in solving the ultralarge 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 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, 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 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 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 10 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. Table 1.1 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 Brodwin reported the correct numerical stability criterion for Yee's algorithm; the first sinusoidal steady-state FDTD solutions of two- and three-dimensional electromagnetic wave interactions with material structures [8, 9]; and the first bioelectromagnetics models [9].
1977	Holland [10], and Kunz and Lee [11] applied Yee's algorithm to EMP problems.
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].
1985	Gwarek introduced the lumped equivalent circuit formulation of FDTD [16].
1986	Choi and Hoefer published the first FDTD simulation of waveguide structures [17].
1987, 1988	Kriegsmann et al. and Moore et al. published the first articles on ABC theory in *IEEE Trans. Antennas and Propagation* [18, 19].
1987, 1988, 1992	Contour-path subcell techniques were introduced by Umashankar et al. to permit FDTD modeling of thin wires and wire bundles [20], by Taflove et al. to model penetration through cracks in conducting screens [21], and by Jurgens et al. to conformally model the surface of a smoothly curved scatterer [22].
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. [23], Shankar et al. [24], and Madsen and Ziolkowski [25].

1988 Sullivan et al. published the first three-dimensional FDTD model of sinusoidal steady-state electromagnetic wave absorption by a complete human body [26].

1988 FDTD modeling of microstrips was introduced by Zhang et al. [27].

1990, 1991 FDTD modeling of frequency-dependent dielectric permittivity was introduced by Kashiwa and Fukai [28], Luebbers et al. [29], and Joseph et al. [30].

1990, 1991 FDTD modeling of antennas was introduced by Maloney et al. [31], Katz et al. [32], and Tirkas and Balanis [33].

1990 FDTD modeling of picosecond optoelectronic switches was introduced by Sano and Shibata [34], and El-Ghazaly et al. [35].

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 [36]; studies of beam self-focusing by Ziolkowski and Judkins [37]; the first temporal solitons in two dimensions by Joseph et al. [38]; and the first spatial solitons in two dimensions by Joseph and Taflove [39].

1992 FDTD modeling of lumped electronic circuit elements was introduced by Sui et al. [40].

1993 Toland et al. published the first FDTD models of gain devices (tunnel diodes and Gunn diodes) exciting cavities and antennas [41].

1994 Thomas et al. [42] introduced a Norton's equivalent circuit for the FDTD space lattice, which permits the SPICE circuit analysis tool to implement accurate subgrid 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 [43], which was extended to three dimensions by Katz et al. [44], and to dispersive waveguide terminations by Reuter et al. [45].

1995, 1996 Sacks et al. [46] and Gedney [47] introduced a physically realizable, *uniaxial perfectly matched layer* (UPML) ABC.

1996 Krumpholz and Katehi [48] introduced the *multiresolution 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.

1997 Ramahi [51] introduced the *complementary operators method* (COM) to implement highly effective analytical ABCs.

1997 Dey and Mittra [52] introduced a simple, stable, and accurate contour-path technique to model curved metal surfaces in FDTD space lattices.

1998 Maloney and Kesler [53] introduced several novel means to analyze periodic structures in the FDTD space lattice.

1999 Schneider and Wagner [54] introduced a comprehensive analysis of FDTD grid dispersion based upon complex wavenumbers.

2000, 2001 Zheng, Chen, and Zhang [55, 56] introduced the first three-dimensional *alternating-direction implicit* (ADI) FDTD algorithm with provable unconditional numerical stability.

2000 Roden and Gedney introduced the advanced *convolutional PML* (CPML) ABC [57].

2000 Rylander and Bondeson introduced a provably stable FDTD-FE hybrid technique [58].

2002–2004 Hayakawa et al. [59, 60] and Simpson and Taflove [61, 62] introduced FDTD modeling of the entire Earth-ionosphere waveguide for extremely low-frequency geophysical phenomena.

2002, 2004 Huang [63] and Chang and Taflove [64] introduced a hybrid FDTD-quantum mechanics model of lasing in a four-level, two-electron atomic system governed by the Pauli Exclusion Principle.

2003 DeRaedt introduced the unconditionally stable, "one-step" FDTD technique [65].

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 vector $\boldsymbol{E}$ and magnetic field vector $\boldsymbol{H}$ 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 to 20 samples per λ_0 are needed. 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, *absorbing boundary conditions* (ABCs) are employed at the outer lattice truncation planes. ABCs ideally permit all outgoing numerical wave analogs to exit the computation space 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 $\boldsymbol{E}$ and $\boldsymbol{H}$ 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 [22, 52, 58]. 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 [24]. 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 [25]. 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 preprocessing 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.

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/m^2. Further, let the minimum observable power density of a secondary (scattered) wave be P_S W/m^2, 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. This is due to accumulating numerical dispersion artifacts, including phase-velocity anisotropies and inhomogeneities within the mesh.

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

The 1990s saw the emergence of powerful, entirely new classes of ABCs including Berenger's PML [43]; the *uniaxial PML* (UPML) of Sacks et al. [46] and Gedney [47]; Ramahi's *complementary operator method* (COM) [51]; and the *convolutional PML* (CPML) of Roden and Gedney [57]. These ABCs were shown to have grid 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 *multiresolution time-domain* (MRTD) technique introduced by Krumpholz and Katehi [48], and the *pseudospectral time-domain* (PSTD) technique introduced by Liu [49, 50]. As a result of these advances, the possibility is emerging 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 currently model electromagnetic wave interaction problems requiring the solution of more than 10^9 field-vector unknowns. At this level of complexity, it is possible to develop highly detailed, three-dimensional, volumetric models of both naturally occurring and engineered systems, including, for example:

- The entire Earth-ionosphere system (within ±100 km of sea level) for extremely low-frequency geophysical phenomena such as potential earthquake precursors;
- The entire human body for bioelectromagnetics studies at microwave frequencies;
- Clusters of scores of living cells for biophotonics studies at optical wavelengths;
- Entire aircraft and missiles illuminated by radar up to approximately 3 GHz;
- Entire cellphones, including details of the internal circuit boards;
- Combinations of micron-scale lasers, waveguides, couplers, and resonators forming integrated photonic circuits.

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 is 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

We now present eight examples (with accompanying color visualizations) that illustrate both contemporary and emerging applications of FDTD computational electrodynamics modeling.

- *Impulsive around-the-world extremely low-frequency propagation*
  Can we model the Earth's electromagnetic environment for geophysical studies?

- *Cellphone radiation interacting with the human head*
  Can cellphones be designed to meet government safety standards for microwave exposure while meeting all other design goals?

- *Early-stage detection of breast cancer using an ultrawideband microwave radar*
  Can we design an ultrawideband radar to detect breast cancer earlier than using X-ray mammography?

- *Homing accuracy of a radar-guided missile*
  How do interactions between a missile's horn antenna and radome generate errors in the angular location of a target?
- *Electromagnetic wave vulnerabilities of a military jet plane*
  What is the vulnerability of a military jet to either radar detection or circuit upset due to enemy microwave threats?
- *Millimeter-wave propagation in a defect-mode electromagnetic bandgap structure*
  Within the constraints imposed by existing circuit-board technology, can we design digital interconnects transferring data at rates above 50 Gbits/s?
- *Photonic crystal microcavity laser*
  Can room-temperature operation be achieved with the world's smallest lasers?
- *Photonic crystal cross-waveguide switch*
  Can we design low-power all-optical switches operating above 10 Gb/s?

References are made to later chapters where these applications are discussed in the context of specific FDTD techniques.

1.7.1 Impulsive Around-the-World Extremely Low-Frequency Propagation

Global propagation of *extremely low-frequency* (ELF: 3 Hz to 3 kHz) and *very low-frequency* (VLF: 3 to 30 kHz) electromagnetic waves in the Earth-ionosphere waveguide has a rich history of investigation extending over many years. ELF/VLF propagation phenomena form the basis of submarine communications and remote-sensing investigations of lightning and sprites, global temperature change, subsurface structures, and potential earthquake precursors.

Most theoretical techniques for modeling ELF/VLF propagation within the Earth-ionosphere waveguide are based upon frequency-domain waveguide theory. However, this theory cannot account for the complicated dielectric inhomogeneities represented by the ionosphere, continents, and oceans. Recently, Hayakawa et al. [59, 60] and Simpson and Taflove [61, 62] developed FDTD models for this purpose. In the case of [61], all features of the Earth's lithosphere and ionosphere located within ±100 km of sea level were modeled in three dimensions with a resolution of approximately $40 \times 40 \times 5$ km. Periodic boundary conditions were used in conjunction with a variable-cell grid wrapping around the entire Earth. The model was verified by numerical studies of frequency-dependent propagation attenuation, antipodal propagation, and the Schumann resonance. See Chapter 3, Section 3.8 for a complete discussion.

Fig. 1.1 provides three color snapshot visualizations of the global propagation of the electromagnetic pulse resulting from a simulated vertical lightning stroke at the Equator off the coast of South America. (The complete video of this phenomenon is available for download at http://www.ece.northwestern.edu/ecefaculty/taflove/3Dmovietext@gif.avi.) Slight departures from circular symmetry of the propagating pulse are observed. Analysis reveals that this lack of symmetry is caused by the inhomogeneities of the lithosphere conductivity assumed in the model, rather than by numerical distortions resulting from the gridding technique.

Ongoing work in this area is aimed at calculating signatures of potential *ultralow-frequency* (ULF) precursors of major earthquakes. A second area of interest is the generation and sensing of ULF/ELF signals tailored for the remote detection of large underground ore and oil deposits.

1.7.2 Cellphone Radiation Interacting with the Human Head

Since 1990, there has been 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 safety standard for microwave exposure to the user. Namely, the peak *specific absorption rate* (SAR) for any 1g of tissue must be less than 1.6 W/kg.

FDTD simulations have become a powerful tool for the design of safe, efficient, and 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.

Fig. 1.2 illustrates the SAR component of the case study discussed in Chapter 14, Section 14.9, wherein the Motorola T250 phone is modeled in very fine-grained detail imported from its computer-aided design database. Here, a combination of graded FDTD meshes and local mesh refinement realizes space-cell sizes as small as 0.1 mm. This allows precise rendering of the phone's housing, pushbuttons, multilayer printed circuit board, electrical interconnections, and helical antenna in the FDTD model space. As shown in Fig. 1.2(a), the phone is modeled in the FDTD space lattice as being located immediately adjacent to a high-resolution, 15-tissue numerical model of the human head derived from magnetic resonance imaging of a volunteer. This head model is comprised of 121 slices (1 mm thick in the ear region; 3 mm thick elsewhere), with each slice having a transverse spatial resolution of 0.2 mm. Fig. 1.2(b) provides a color visualization of the FDTD-calculated SAR distribution within the head model.

In related discussions in Chapter 14, Section 14.9, the FDTD model of the T250 phone is validated relative to laboratory measurements of the 1g-averaged and 10g-averaged peak SAR values for a standard physical head phantom. Good-to-excellent agreement between the experimental and FDTD data is observed.

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.

1.7.3 Early-Stage Detection of Breast Cancer Using an Ultrawideband Microwave Radar

The "flip side" of the cellphone human-exposure issue involves potential biomedical applications of electromagnetic waves, wherein the goal is to *deliberately* direct electromagnetic signals into the human body. These applications fall into three broad categories: diagnostic imaging, medical implants, and thermal therapy. An excellent compilation of recent advances in these areas is provided in [66]. The ability of FDTD to model essentially arbitrary configurations of inhomogeneous, lossy, and dispersive dielectric materials makes this method broadly applicable in each of these three areas.

An emerging major application of noninvasive microwave imaging is the detection of breast cancer at an early stage. In particular, the *ultrawideband* (UWB) microwave technology discussed in [67–70] has shown promise for early-stage breast cancer detection. Here, FDTD modeling has been used to simulate the data-acquisition step wherein the breast is illuminated by

UWB pulses generated by an antenna array at the surface of the breast, and the scattered signals received by the antenna array are recorded and processed by space-time imaging algorithms. As illustrated in Fig. 1.3, detailed FDTD modeling of this type indicates that UWB technology has the potential to detect malignant tumors smaller than 5 mm embedded within normal, inhomogeneous breast tissues [69, 70]. These promising results have currently brought UWB breast cancer detection technology to the point of initial preclinical investigations [70].

1.7.4 Homing Accuracy of a Radar-Guided Missile

A long-standing defense-technology problem that can cause unreliable homing of radar-guided missiles is caused by 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.

Fig. 1.4 illustrates how detailed FDTD modeling can provide valuable insights into the complex interactions of the missile body, radome, and internal horn antenna. Here, the E-field within and near a missile radome is visualized at two instants in time for the case of impulsive illumination at 10 GHz. As the incident wave enters the radome, it first generates a radially propagating scattered field due to the action of the radome's metal tip and curved surface. After encountering the horn antenna, a portion of the incident energy is scattered, and the rest propagates down the horn to a matched load. A packet of energy is trapped inside the wall of the dielectric radome. Subsequently, this trapped mode encounters the junction between the radome and the missile body, reflects, and reradiates. Then, a new guided wave forms in the radome wall due to structural scattering from the horn antenna.

It should be clear that this highly spatially detailed, picosecond-by-picosecond exposition of the electrodynamics of this problem cannot be obtained by presently available physical measurements. FDTD computational modeling allows the engineer charged with mitigating targeting errors to literally perform experiments on the computer to test the effectiveness of proposed design alterations.

1.7.5 Electromagnetic Wave Vulnerabilities of a Military Jet Plane

A vital defense-technology problem involves the vulnerability of military aircraft to either radar detection or circuit upset due to enemy microwave threats. With the advent of provably stable hybrid FDTD-FE techniques [58], it is now possible to construct efficient volumetric models of fighter aircraft (detailing the internal stores as well as the external surface shape) for these purposes up to several gigahertz. These models are no longer constrained to approximate complex shapes via Yee staircasing. Instead, the flexibility of modern FE spatial meshing can be applied to the regions of the aircraft model that require this level of detail.

Fig. 1.5 visualizes the induced surface currents on the Saab *Trainer* aircraft and its jet engine air inlet calculated by the FDTD-FE technique of [60] for head-on illumination by a horizontally polarized Gaussian pulse. The *Trainer* is 11m long, 3.4m high, and spans 8m between its wingtips. In results shown in Chapter 19, excellent agreement is obtained for the bistatic RCS of the *Trainer* at 500 MHz for both horizontal and vertical polarizations of the incident wave relative to a fast-multipole-method solution of the phasor-domain electric field integral equation. Only three layers of unstructured FE cells are needed to obtain an accurate geometrical representation of the aircraft, while 4-cm ($\lambda/15$) FDTD "bricks" are used throughout the remainder of the modeling space. Implementing this hybrid FDTD-FE model requires 1.9 GB of memory and 8 hours running time on a single 900-MHz Itanium2 processor.

1.7.6 Millimeter-Wave Propagation in a Defect-Mode Electromagnetic Bandgap Structure

As computer clock rates continue to rise in the microwave frequency range above 3 GHz, problems with signal integrity, cross-coupling, and radiation will eventually render useless the baseband metal-strip circuit-board interconnects used since the 1940s. While replacing metal strips with optical fibers would solve the problem, the required incorporation of optoelectronics would represent a revolution in both chip-making and interconnect technologies.

References [71, 72] propose an alternate solution: bandpass wireless interconnects implemented using "defect-mode" *electromagnetic bandgap* (EBG) waveguides. As studied in [71, 72], EBG structures can be implemented using existing circuit-board fabrication technology as square arrays of copper via pins. One or more rows of pins are removed to create a linear waveguide. Operation at center frequencies above 100 GHz is conceptually feasible as the gain-bandwidth of silicon transistors rises well into this range. Relative to metal strips or optical fibers, such millimeter-wave EBG waveguides would have the following advantages when used for board-level digital interconnects:

- Sufficient high-quality bandwidth (i.e., flat transmission magnitude, linear phase shift, broadband impedance matching, and no multimoding) to support computer processors clocked up to 20 to 30 GHz;
- Fabrication possible using existing circuit-board technology;
- Low signal attenuation due to copper loss;
- Little signal distortion, coupling, and radiation;
- Nearly speed-of-light signal transmission via the usage of low-permittivity dielectric media.

References [71, 72] reported laboratory measurements and supporting three-dimensional FDTD modeling of prototype EBG waveguiding structures having linear double-row defects. These structures were realized using double-sided circuit board with either standard FR4 or low-loss Rogers 5880 as the dielectric material. Copper vias electrically bonded to the upper and lower ground planes served to implement the rows of EBG pins. The waveguides were dimensionally scaled to operate at a center frequency of 10 GHz [71] and 50 GHz [72]. These structures were coupled to input and output coaxial lines using thin, short probes extending transversely across the gap between the upper and lower ground planes and bonded to the opposing ground plane.

Fig. 1.6 provides three visualizations relating to the FDTD modeling and laboratory measurements of the EBG waveguide designed for a center frequency of 50 GHz [72]. This structure spans 8.6 cm between its input and output probes, 0.76 mm between its upper and lower ground planes, and 3.9 mm between the rows of vias bordering the waveguiding defect. The pins have a radius of 0.23 mm and a center-to-center separation of 1.3 mm. Note that the waveguide is bounded on all sides, thereby representing a closed cavity.

Fig. 1.6(c) shows the measured and FDTD results for the insertion loss of this EBG waveguide. A standard Agilent network analyzer having 50Ω source and load impedances was used in the measurements. We see a sharp transition from a stopband of at least −65 dB to the passband above approximately 28 GHz. This passband extends to approximately 70 GHz with a gain flatness of ±1.5 dB, representing a relative bandwidth approaching 90%. Furthermore, the total insertion loss at midband (50 GHz) is only 4 dB. Comparison of this result with that of an identical 12.7-cm-long waveguide shows that approximately 2 dB of this loss is caused by the Rogers 5880 dielectric. Subtracting this dielectric loss from the total loss yields a coupling loss of only approximately 1 dB at each coaxial transition. This indicates the possibility of excellent broadband matching into the EBG waveguiding defect using simple, thin metal probes.

Assuming the availability of suitable low-loss dielectrics to serve as insulating layers within the circuit boards, this technology will be scalable to center frequencies well above 100 GHz. Then, these interconnects will be capable of supporting data rates in the hundreds of gigabits per second, adequate for the elevated computer clock rates expected over the next decade.

1.7.7 Photonic Crystal Microcavity Laser

Photonic crystals are artificial structures that have a periodic variation of the refractive index in one, two, or three dimensions. In essence, these are EBG structures scaled to operate at optical frequencies, hence the descriptor common in the lightwave community, *photonic bandgap* (PBG). Similar to a donor or acceptor state in a doped semiconductor, a small defect introduced into a 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.

Three-dimensional FDTD modeling has recently been used to design an electrically driven, single-mode, low-threshold-current, photonic crystal microlaser that operates at room temperature [73]. Fig. 1.7(a) is a schematic diagram of this microlaser [73]. Here, light is confined to a single defect at the center of a photonic crystal that is fabricated as an array of air holes etched within a 282.5-nm-thick semiconductor slab of refractive index $n = 3.4$. A 1-μm-long InP post placed directly below the defect injects holes, whereas electrons are supplied laterally from a top circumferential electrode. The holes and electrons recombine in six InGaAsP quantum wells within the photonic crystal slab. These quantum wells are designed to have an electroluminescence peak near the communications wavelength of 1,500 nm.

Fig. 1.7(b) is a scanning electron microscope image of the top view of the fabricated photonic crystal slab [73]. The defect-mode cavity is surrounded by five photonic crystal regions I, II, III, IV, and V having the same lattice constant $a \approx 510$ nm, but different air-hole radii: 0.28a, 0.35a, 0.385a, 0.4a, and 0.41a, respectively. Optical electromagnetic fields are confined within the cavity in the vertical direction by total internal reflection at the slab-air interface, and laterally by the action of the photonic crystal bandgap. Therefore, light can escape from this cavity only by either tunneling laterally through the photonic crystal, or exiting vertically by impinging on the slab-air interface at a sufficiently large angle.

Fig. 1.7(c) visualizes the FDTD-calculated E-field intensity profile (log scale) along a planar central cut through the photonic crystal slab [73]. Note that, in creating the FDTD model, actual structural data for the fabricated microlaser was transferred directly from the scanning electron microscope image to the FDTD geometry dataset. This allowed the model to incorporate actual imperfections of the fabrication process. As a result, the FDTD model predicted the actual observed monopole-mode operation of the laser, and reasonably reproduced the field asymmetry resulting from the laser's imperfect fabrication. Cavity quality (Q) factors for this study were in the order of 3,000 for a lasing wavelength of 1,519.7 nm and a modal volume of 0.0587 μm^3. The latter corresponds to 0.684 $(\lambda/n)^3$, which approaches the smallest theoretical value.

The microlaser of [73] discussed here has a flexible geometry that allows fine tuning of its radiation pattern and emission wavelength. The compact size and high spontaneous-emission coupling factor of its defect microcavity also make this laser interesting as a low-noise light source, one day potentially approaching thresholdless operation or a single-photon source. In addition, such microcavity lasers 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.

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

1.7.8 Photonic Crystal Cross-Waveguide Switch

The FDTD method achieves robustness for nonlinear optical problems by retaining the optical carrier and solving for fundamental quantities, the optical $\boldsymbol{E}$ and $\boldsymbol{H}$ 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 $\boldsymbol{E}$ and $\boldsymbol{H}$ boundary conditions at all material interfaces in the time scale of a small fraction of the carrier period, whether or not the media are dispersive and/or nonlinear. As a result, it is almost completely general.

Because of these desirable characteristics, FDTD has been used to study micron-scale all-optical switching in photonic crystals [74]. Photonic crystal microcavities can confine photons to submicron dimensions. Such ultrastrong confinement leads to very high peak power levels which enhance nonlinear processes by many orders of magnitude. This enhancement in nonlinearity can be used to construct ultralow-power all-optical switches.

Fig. 1.8 shows color visualizations of the operation of an all-optical photonic crystal cross-waveguide switch that can operate at 10 Gbits/s with pulse energy levels as small as 100 fJ/pulse. The switching action employs only the intrinsic nonresonant Kerr nonlinearities that are present in semiconductors such as AlGaAs for wavelengths in the order of 1.5 μm (compatible with modern lightwave communication systems). The structure of Fig. 1.8 consists of two photonic crystal waveguides that cross each other at right angles. The microcavity located at the intersection of the waveguides possesses two orthogonal dipole-like cavity modes. Each cavity mode couples to only one of the waveguides, and isolates the signals in the two waveguides. As proven by the FDTD simulations, this accomplishes both spatial and spectral isolation between the signal and the control inputs, even in the nonlinear regime.

Fig. 1.8 illustrates the two states of the device as simulated with FDTD. Initially, the microcavity at the intersection is designed to be out of resonance with the signal input in the absence of the control input. This causes a very low signal transmission, as shown in Fig. 1.8(a). When a control pulse is launched, as shown in Fig. 1.8(b), it causes a strong electric field buildup in the microcavity which shifts its resonant frequency due to the material nonlinearity. The overall result is a bistable transition of the cavity resonance which switches the signal transmission from low to high. Since the switching process occurs via bistable transitions, the device is digital, that is, only low and high transmission states are allowed. As discussed in [74], the same device can function either as a switch or as a memory element, depending upon the design parameters.

The implications of the work shown in Fig. 1.8 may be profound for the realization of "optonics," a proposed successor technology to electronics in the 21st century, which would integrate optical-fiber interconnects and optical microchips into systems of currently unimaginable information-processing capability.

1.8 CONCLUSIONS

Whereas large-scale solutions of Maxwell's equations were motivated in the past primarily by the requirements of military defense, the entire field of computational electrodynamics has shifted toward important applications in communications, computing, and biomedicine. These applications touch all of us in our daily lives. Now, the economic and physical well-being of nations, as well as their military security, is being markedly improved by advances in computational electrodynamics.

A key remaining goal is the computational unification of electromagnetic waves, charge transport, and quantum phenomena to allow modeling the broadest range of problems:

- These problems literally cover the frequency spectrum from the ultralow to the ultraviolet (equivalently, time scales from the better part of a minute to subfemtoseconds).
- These problems cover characteristic distance scales from that of the entire Earth to that of a single atom.

Arguably, FDTD and related space-grid time-domain techniques comprise the most flexible and robust means to solve problems spanning these enormous ranges, especially where geometrical complexity, nonlinearity, and multiphysics dominate. In this spirit, let us proceed with the remaining 19 chapters of this book, and engage our readers in exercising their intellect and energy as they explore the rapidly expanding universe of FDTD computational electrodynamics.

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 Comp. Sci. Engrg.* 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 Propagat.*, 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 Tech.*, 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 Tech.*, Vol. 23, 1975, pp. 888–896.

[10] Holland, R., "Threde: a free-field EMP coupling and scattering code," *IEEE Trans. Nuclear Sci.*, 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. Electromagn. Compat.*, 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. Electromagn. Compat.*, 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. Electromagn. Compat.*, 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. Electromagn. Compat.*, Vol. 24, 1982, pp. 397–405.

[15] Taflove, A., and K. R. Umashankar, "Radar cross section of general three-dimensional scatterers," *IEEE Trans. Electromagn. Compat.*, Vol. 25, 1983, pp. 433–440.

[16] Gwarek, W., "Analysis of an arbitrarily shaped planar circuit — A time-domain approach," *IEEE Trans. Microwave Theory Tech.*, Vol. 33, 1985, pp. 1067–1072.

[17] Choi, D. H., and W. J. Hoefer, "The finite-difference time-domain method and its application to eigenvalue problems," *IEEE Trans. Microwave Theory Tech.*, Vol. 34, 1986, pp. 1464–1470.

[18] 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 Propagat.*, Vol. 35, 1987, pp. 153–161.

[19] Moore, T. G., J. G. Blaschak, A. Taflove, and G. A. Kriegsmann, "Theory and application of radiation boundary operators," *IEEE Trans. Antennas Propagat.*, Vol. 36, 1988, pp. 1797–1812.

[20] 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 Propagat.*, Vol. 35, 1987, pp. 1248–1257.

[21] 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 Propagat.*, Vol. 36, 1988, pp. 247–257.

[22] Jurgens, T. G., A. Taflove, K. R. Umashankar, and T. G. Moore, "Finite-difference time-domain modeling of curved surfaces," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 357–366.

[23] 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 Propagat.*, Vol. 35, 1987, pp. 1160–1173.

[24] 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.

[25] 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.

[26] 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. Biomed. Engrg.*, Vol. 35, 1988, pp. 179–186.

[27] 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 Tech.*, Vol. 36, 1988, pp. 263–267.

[28] Kashiwa, T., and I. Fukai, "A treatment by FDTD method of dispersive characteristics associated with electronic polarization," *Microwave Optics Tech. Lett.*, Vol. 3, 1990, pp. 203–205.

[29] Luebbers, R., F. Hunsberger, K. Kunz, R. Standler, and M. Schneider, "A frequency-dependent finite-difference time-domain formulation for dispersive materials," *IEEE Trans. Electromagn. Compat.*, Vol. 32, 1990, pp. 222–229.

[30] 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 Lett.*, Vol. 16, 1991, pp. 1412–1414.

[31] 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 Propagat.*, Vol. 38, 1990, pp. 1059–1065.

[32] 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 Propagat.*, Vol. 39, 1991, pp. 1203–1212.

[33] Tirkas, P. A., and C. A. Balanis, "Finite-difference time-domain technique for radiation by horn antennas," *Proc. 1991 IEEE Antennas Propagat. Soc. Intl. Symp.*, Vol. 3, 1991, pp. 1750–1753.

[34] Sano, E., and T. Shibata, "Fullwave analysis of picosecond photoconductive switches," *IEEE J. Quantum Electron.*, Vol. 26, 1990, pp. 372–377.

[35] El-Ghazaly, S. M., R. P. Joshi, and R. O. Grondin, "Electromagnetic and transport considerations in subpicosecond photoconductive switch modeling," *IEEE Trans. Microwave Theory Tech.*, Vol. 38, 1990, pp. 629–637.

[36] 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 Lett.*, Vol. 17, 1992, pp. 180–182.

[37] 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 Soc. America B*, Vol. 10, 1993, pp. 186–198.

[38] 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 Lett.*, Vol. 18, 1993, pp. 491–493.

[39] Joseph, R. M., and A. Taflove, "Spatial soliton deflection mechanism indicated by FDTD Maxwell's equations modeling," *IEEE Photonics Tech. Lett.*, Vol. 2, 1994, pp. 1251–1254.

[40] 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 Tech.*, Vol. 40, 1992, pp. 724–730.

[41] Toland, B., B. Houshmand, and T. Itoh, "Modeling of nonlinear active regions with the FDTD method," *IEEE Microwave Guided Wave Lett.*, Vol. 3, 1993, pp. 333–335.

[42] 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 Guided Wave Lett.*, Vol. 4, 1994, pp. 141–143.

[43] Berenger, J. P., "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comp. Phys.*, Vol. 114, 1994, pp. 185–200.

[44] 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 Guided Wave Lett.*, Vol. 4, 1994, pp. 268–270.

[45] 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 Guided Wave Lett.*, Vol. 4, 1994, pp. 344–346.

[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 Propagat.*, 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 Propagat.*, 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 Tech.*, 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, Tech. Rept. 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 Propagat. Soc. Intl. Symp.*, Vol. 1, pp. 122–125.

[51] Ramahi, O. M., "The complementary operators method in FDTD simulations," *IEEE Antennas Propagat. Mag.*, Vol. 39, Dec. 1997, pp. 33–45.

[52] Dey, S., and R. Mittra, "A locally conformal finite-difference time-domain algorithm for modeling three-dimensional perfectly conducting objects," *IEEE Microwave Guided Wave Lett.*, Vol. 7, 1997, pp. 273–275.

[53] Maloney, J. G., and M. P. Kesler, "Analysis of Periodic Structures," Chap. 6 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, A. Taflove, (ed.), Norwood, MA: Artech House, 1998.

[54] Schneider, J. B., and C. L. Wagner, "FDTD dispersion revisited: Faster-than-light propagation," *IEEE Microwave Guided Wave Lett.*, Vol. 9, 1999, pp. 54–56.

[55] Zheng, F., Z. Chen, and J. Zhang, "Toward the development of a three-dimensional unconditionally stable finite-difference time-domain method," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 1550–1558.

[56] Zheng, F., and Z. Chen, "Numerical dispersion analysis of the unconditionally stable 3-D ADI-FDTD method," *IEEE Trans. Microwave Theory Tech.*, Vol. 49, 2001, pp. 1006–1009.

[57] Roden, J. A., and S. D. Gedney, "Convolutional PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media," *Microwave Optical Tech. Lett.*, Vol. 27, 2000, pp. 334–339.

[58] Rylander, T., and A. Bondeson, "Stable FDTD-FEM hybrid method for Maxwell's equations," *Comput. Phys. Comm.*, Vol. 125, 2000, pp. 75–82.

[59] Hayakawa, M., and T. Otsuyama, "FDTD analysis of ELF wave propagation in inhomogeneous subionospheric waveguide models," *ACES J.*, Vol. 17, 2002, pp. 239–244.

[60] Otsuyama, T., D. Sakuma, and M. Hayakawa, "FDTD analysis of ELF wave propagation and Schumann resonances for a subionospheric waveguide model," *Radio Science*, Vol. 38(6), 2003, 1103.

[61] Simpson, J. J., and A. Taflove, "Three-dimensional FDTD modeling of impulsive ELF propagation about the Earth-sphere," *IEEE Trans. Antennas Propagat.*, Vol. 52, 2004, pp. 443–451.

[62] Simpson, J. J., and A. Taflove, "Efficient modeling of impulsive ELF antipodal propagation about the Earth-sphere using an optimized two-dimensional geodesic FDTD grid," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 3, 2004, pp. 215–218.

[63] Huang, Y., *Simulation of Semiconductor Materials Using FDTD Method*, M.S. thesis, Northwestern Univ., Evanston, IL, 2002.

[64] Chang, S.-H., and A. Taflove, "Finite-difference time-domain model of lasing action in a four-level two-electron atomic system," *Optics Express*, Vol. 12, 2004, pp. 3827–3833.

[65] De Raedt, H., K. Michielsen, J. S. Kole, and M. T. Figge, "Solving the Maxwell equations by the Chebyshev method: A one-step finite difference time-domain algorithm," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 3155–3160.

[66] Lazzi, G., O. P. Gandhi, and S. Ueno, (eds.), *IEEE Trans. Microwave Theory Tech.: Special Issue on Medical Applications and Biological Effects of RF/Microwaves*, Vol. 52, 2004, pp. 1853–2083.

[67] Hagness, S. C., A. Taflove, and J. E. Bridges, "Two-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: fixed-focus and antenna-array sensors," *IEEE Trans. Biomed. Engrg.*, Vol. 45, 1998, pp. 1470–1479.

[68] Hagness, S. C., A. Taflove, and J. E. Bridges, "Three-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: design of an antenna-array element," *IEEE Trans. Antennas Propagat.*, Vol. 47, 1999, pp. 783–791.

[69] Bond, E. J., X. Li, S. C. Hagness, and B. D. Van Veen, "Microwave imaging via space-time beamforming for early detection of breast cancer," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 1690–1705.

[70] Li, X., S. K. Davis, S. C. Hagness, D. W. van der Weide, and B. D. Van Veen, "Microwave imaging via space-time beamforming: Experimental investigation of tumor detection in multilayer breast phantoms," *IEEE Trans. Microwave Theory Tech.: Special Issue on Medical Applications and Biological Effects of RF/Microwaves*, Vol. 52, 2004, pp. 1856–1865.

[71] Simpson, J. J., A. Taflove, J. A. Mix, and H. Heck, "Computational and experimental study of a microwave electromagnetic bandgap structure with waveguiding defect for potential use as a bandpass wireless interconnect," *IEEE Microwave Wireless Components Lett.*, Vol. 14, 2004, pp. 343–345.

[72] Simpson, J. J., A. Taflove, J. A. Mix, and H. Heck, "Advances in hyperspeed digital interconnects using electromagnetic bandgap technology: Measured low-loss 43-GHz passband centered at 50 GHz," *Proc. IEEE Antennas Propagat. Soc. Intl. Symp.*, Washington, D.C., 2005.

[73] Park, H.-G., S.-H. Kim, S.-H. Kwon, Y.-G. Ju, J.-K. Yang, J.-H. Baek, S.-B. Kim, and Y.-H. Lee, "Electrically driven single-cell photonic crystal laser," *Science*, Vol. 305, 2004, pp. 1444–1447.

[74] Yanik, M. F., S. Fan, M. Soljacic, and J. D. Joannopoulos, "All-optical transistor action with bistable switching in a photonic crystal cross-waveguide geometry," *Optics Lett.*, Vol. 28, 2003, pp. 2506–2508.

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

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \tag{2.1}$$

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

$$u(x, t) = F(x + ct) + G(x - ct) \tag{2.2}$$

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

$$\frac{\partial u}{\partial t} = \underbrace{\frac{dF(x+ct)}{d(x+ct)}}_{F'} \cdot \underbrace{\frac{\partial (x+ct)}{\partial t}}_{c} + \underbrace{\frac{dG(x-ct)}{d(x-ct)}}_{G'} \cdot \underbrace{\frac{\partial (x-ct)}{\partial t}}_{-c} \tag{2.3a}$$

$$= cF'(x+ct) - cG'(x-ct)$$

$$\frac{\partial^2 u}{\partial t^2} = c^2 F''(x+ct) - c(-c)G''(x-ct) \tag{2.3b}$$

$$= c^2 F''(x+ct) + c^2 G''(x-ct)$$

$$\frac{\partial u}{\partial x} = \underbrace{\frac{dF(x+ct)}{d(x+ct)}}_{F'} \cdot \underbrace{\frac{\partial (x+ct)}{\partial x}}_{1} + \underbrace{\frac{dG(x-ct)}{d(x-ct)}}_{G'} \cdot \underbrace{\frac{\partial (x-ct)}{\partial x}}_{1} \tag{2.4a}$$

$$= F'(x+ct) + G'(x-ct)$$

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

and substitute into the scalar wave equation (2.1)

$$c^2 F''(x+ct) + c^2 G''(x-ct) = c^2\left[F''(x+ct) + G''(x-ct)\right] \tag{2.5}$$

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) written in phasor form:

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

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

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

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

$$-\omega^2 = c^2 \cdot \left(-k^2\right)$$

$$\omega^2 = c^2 k^2 \quad \rightarrow \quad k = \pm \omega / c \tag{2.7b}$$

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 $v_p = \omega / k$:

$$v_p = \frac{\omega}{k} = \pm c \tag{2.8}$$

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, that is, $\omega = \omega(k)$, 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

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

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:

$$u(x_i + \Delta x)\big|_{t_n} = u\big|_{x_i, t_n} + \Delta x \cdot \left.\frac{\partial u}{\partial x}\right|_{x_i, t_n} + \frac{(\Delta x)^2}{2} \cdot \left.\frac{\partial^2 u}{\partial x^2}\right|_{x_i, t_n} + \frac{(\Delta x)^3}{6} \cdot \left.\frac{\partial^3 u}{\partial x^3}\right|_{x_i, t_n}$$

$$+ \frac{(\Delta x)^4}{24} \cdot \left.\frac{\partial^4 u}{\partial x^4}\right|_{\xi_1, t_n} \tag{2.10a}$$

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:

$$u(x_i - \Delta x)\big|_{t_n} = u\big|_{x_i, t_n} - \Delta x \cdot \left.\frac{\partial u}{\partial x}\right|_{x_i, t_n} + \frac{(\Delta x)^2}{2} \cdot \left.\frac{\partial^2 u}{\partial x^2}\right|_{x_i, t_n} - \frac{(\Delta x)^3}{6} \cdot \left.\frac{\partial^3 u}{\partial x^3}\right|_{x_i, t_n}$$

$$+ \frac{(\Delta x)^4}{24} \cdot \left.\frac{\partial^4 u}{\partial x^4}\right|_{\xi_2, t_n} \tag{2.10b}$$

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

$$u(x_i + \Delta x)\big|_{t_n} + u(x_i - \Delta x)\big|_{t_n} = 2u\big|_{x_i, t_n} + (\Delta x)^2 \cdot \left.\frac{\partial^2 u}{\partial x^2}\right|_{x_i, t_n} + \frac{(\Delta x)^4}{12} \cdot \left.\frac{\partial^4 u}{\partial x^4}\right|_{\xi_3, t_n} \tag{2.11}$$

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

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

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:

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

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:

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

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

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

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

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

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

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

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:

$$\begin{array}{ccccccc} u_i^{n+1} & = & u_{i+1}^n & + & u_{i-1}^n & - & u_i^{n-1} \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \begin{bmatrix} F(x_i + ct_{n+1}) \\ + G(x_i - ct_{n+1}) \end{bmatrix} & = & \begin{bmatrix} F(x_{i+1} + ct_n) \\ + G(x_{i+1} - ct_n) \end{bmatrix} & + & \begin{bmatrix} F(x_{i-1} + ct_n) \\ + G(x_{i-1} - ct_n) \end{bmatrix} & - & \begin{bmatrix} F(x_i + ct_{n-1}) \\ + G(x_i - ct_{n-1}) \end{bmatrix} \end{array} \tag{2.18a}$$

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

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

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

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

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

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

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

$$e^{j[\omega(n+1)\Delta t - \tilde{k}\, i\Delta x]} = \left(\frac{c\Delta t}{\Delta x}\right)^2 \left\{ e^{j[\omega n\Delta t - \tilde{k}(i+1)\Delta x]} - 2e^{j[\omega n\Delta t - \tilde{k}\, i\Delta x]} + e^{j[\omega n\Delta t - \tilde{k}(i-1)\Delta x]} \right\}$$

$$+ \; 2e^{j[\omega n\Delta t - \tilde{k}\, i\Delta x]} - e^{j[\omega(n-1)\Delta t - \tilde{k}\, i\Delta x]} \tag{2.21a}$$

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

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

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

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

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:

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

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

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), which is 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

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:

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

$$= \frac{1}{\Delta x}\cos^{-1}\left[1 - \frac{1}{2}\left(\frac{\omega}{c}\right)^2(\Delta x)^2\right] = \frac{1}{\Delta x}\cos^{-1}\left[1 - \frac{1}{2}(k\Delta x)^2\right] \tag{2.24a}$$

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:

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

We see that the numerical wavenumber $\tilde{k}$ in (2.24b) reduces to 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

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

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

Performing the indicated cancellation, we obtain

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

Just as in Case 1, $\tilde{k}$ reduces to 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

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

$$= \frac{1}{\Delta x}\cos^{-1}\left\{1 + 4\left[\cos\left(\frac{k\Delta x}{2}\right) - 1\right]\right\} \tag{2.26a}$$

where again $k = \omega/c$. Further noting that $k = 2\pi/\lambda_0$, and letting $\Delta x/\lambda_0 = 0.1$, we proceed:

$$\tilde{k} = \frac{1}{\Delta x}\cos^{-1}\left\{1 + 4\left[\cos\left(\frac{2\pi}{\lambda_0}\cdot\frac{\Delta x}{2}\right) - 1\right]\right\}$$

$$= \frac{1}{\Delta x}\cos^{-1}(0.8042) = \frac{0.63642}{\Delta x} \tag{2.26b}$$

Defining a numerical phase velocity $\tilde{v}_p$ by analogy to the phase velocity for the case of the analytical wave equation [see (2.8)]:

$$\tilde{v}_p = \omega / \tilde{k} \tag{2.27a}$$

we have

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

The numerical phase velocity for this example is seen to be 1.27% less than 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 $[(100 - 98.73)/10] \cdot 360°$, 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. 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 $[(200 - 199.378)/20] \cdot 360°$, or 11.196°, again reduced by a factor of about 4:1.

General Solution

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

$$S = c\Delta t / \Delta x \tag{2.28a}$$

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

$$N_\lambda = \lambda_0 / \Delta x \tag{2.28b}$$

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

$$\tilde{k} = \frac{1}{\Delta x} \cos^{-1}\left\{1 + \left(\frac{1}{S}\right)^2 \left[\cos\left(\frac{2\pi S}{N_\lambda}\right) - 1\right]\right\}$$

$$= \frac{1}{\Delta x} \cos^{-1}(\zeta) = \frac{1}{\Delta x}\left[\frac{\pi}{2} - \sin^{-1}(\zeta)\right] \tag{2.29a}$$

where

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

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

$$N_\lambda|_{\text{transition}} = 2\pi S/\cos^{-1}\left(1 - 2S^2\right) \tag{2.30}$$

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 $\tilde{v}_p$ vary 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)

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

The numerical phase velocity is given by

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

Since

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

we can rewrite (2.32a) as

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

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

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

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:

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

Substituting (2.34) into (2.29) yields

$$\tilde{k} = \frac{1}{\Delta x}\left[\frac{\pi}{2} + j \ln\left(j\zeta + \sqrt{1 - \zeta^2}\right)\right]$$

$$= \frac{1}{\Delta x}\left[\frac{\pi}{2} + j \ln\left(j\zeta + j\sqrt{\zeta^2 - 1}\right)\right] \tag{2.35a}$$

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

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

$$= \frac{1}{\Delta x}\left\{\frac{\pi}{2} + j \ln\left[\left(-\zeta - \sqrt{\zeta^2 - 1}\right) e^{-j\pi/2}\right]\right\} \tag{2.35b}$$

Upon taking the natural logarithm, we further obtain

$$\tilde{k} = \frac{1}{\Delta x}\left\{\frac{\pi}{2} + j\left[\ln\left(-\zeta - \sqrt{\zeta^2 - 1}\right) - j\frac{\pi}{2}\right]\right\}$$

$$= \frac{1}{\Delta x}\left\{\pi + j \ln\left(-\zeta - \sqrt{\zeta^2 - 1}\right)\right\} \tag{2.35c}$$

This yields

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

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

$$\tilde{v}_p = \frac{\omega}{\tilde{k}_{\text{real}}} = \frac{\omega}{(\pi / \Delta x)} = \frac{2\pi f \Delta x}{\pi}$$

$$= \frac{2 f \lambda_0}{N_\lambda} = \frac{2}{N_\lambda} c \tag{2.37a}$$

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

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

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

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

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

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

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

$$N_{\lambda,\min} = 2S \leq 2 \tag{2.38c}$$

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

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

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]:

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

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 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 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 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 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 $N_\lambda \to 1$. In this limit, as well, the attenuation 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 $N_\lambda >> 10$, 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.

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." In fact, this ringing represents time-retarded propagation (i.e., $\tilde{v}_p < c$) of the sparsely sampled, high spatial-frequency spectral content generated by the step discontinuities.

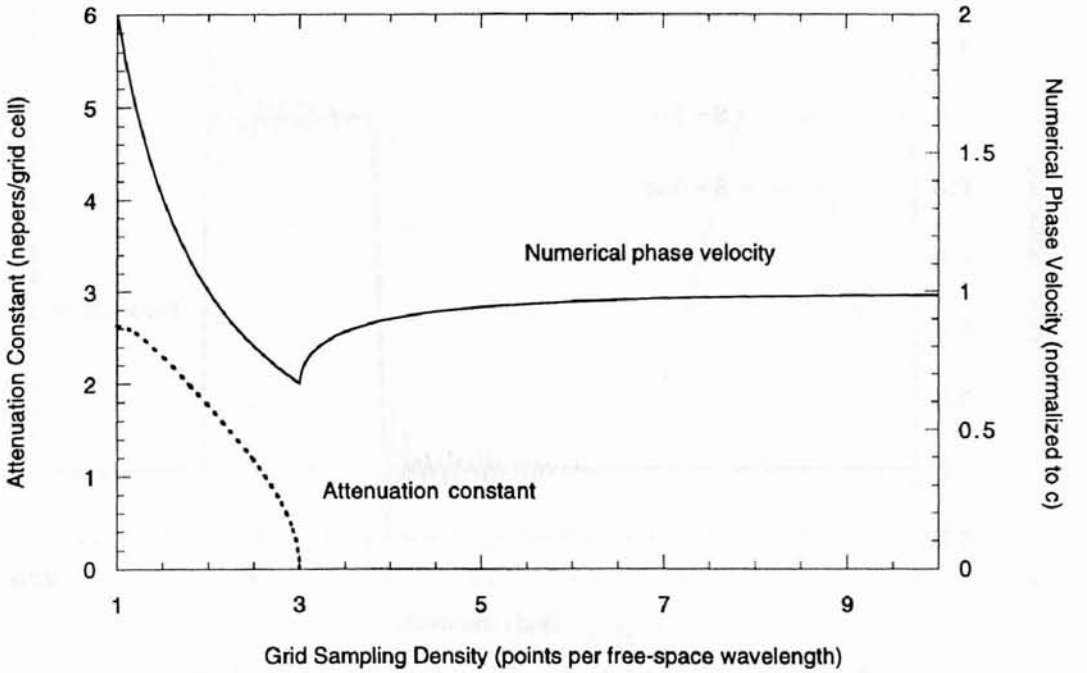


Fig. 2.1 Variation of the normalized numerical phase velocity $\tilde{v}_p/c$ and attenuation per grid cell $\alpha\Delta x$ as a function of the grid sampling density $(1 \le N_\lambda \le 10)$ for a Courant stability factor $S = 0.5$.

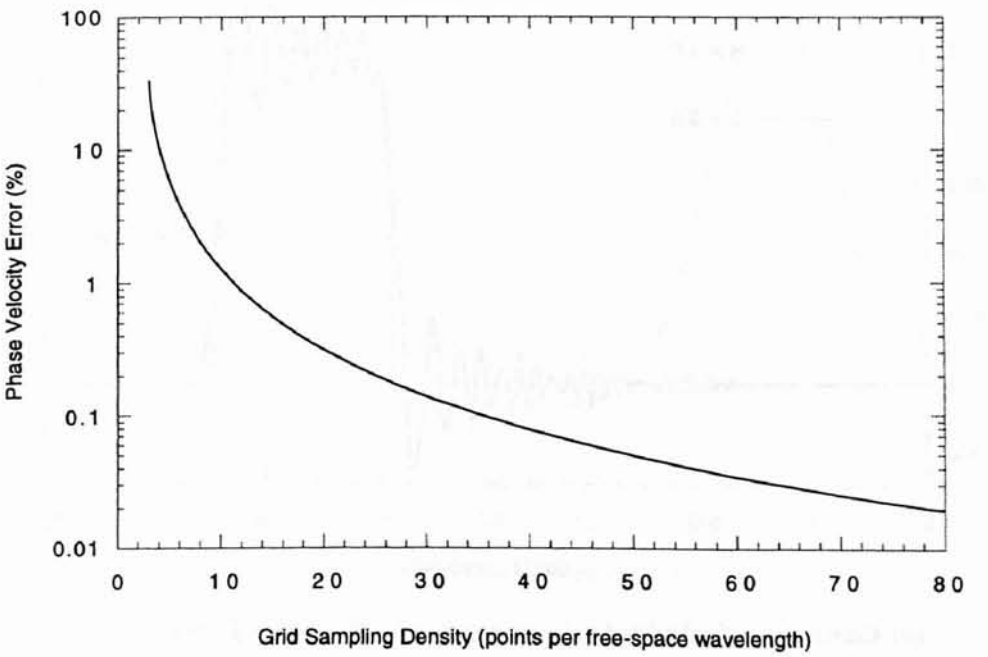
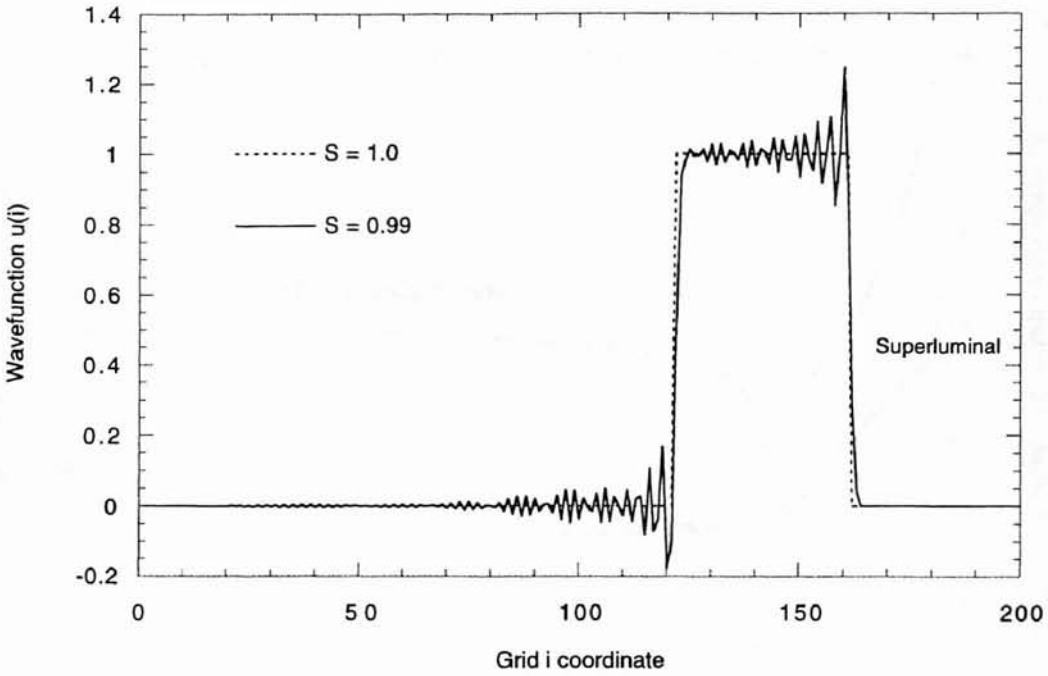
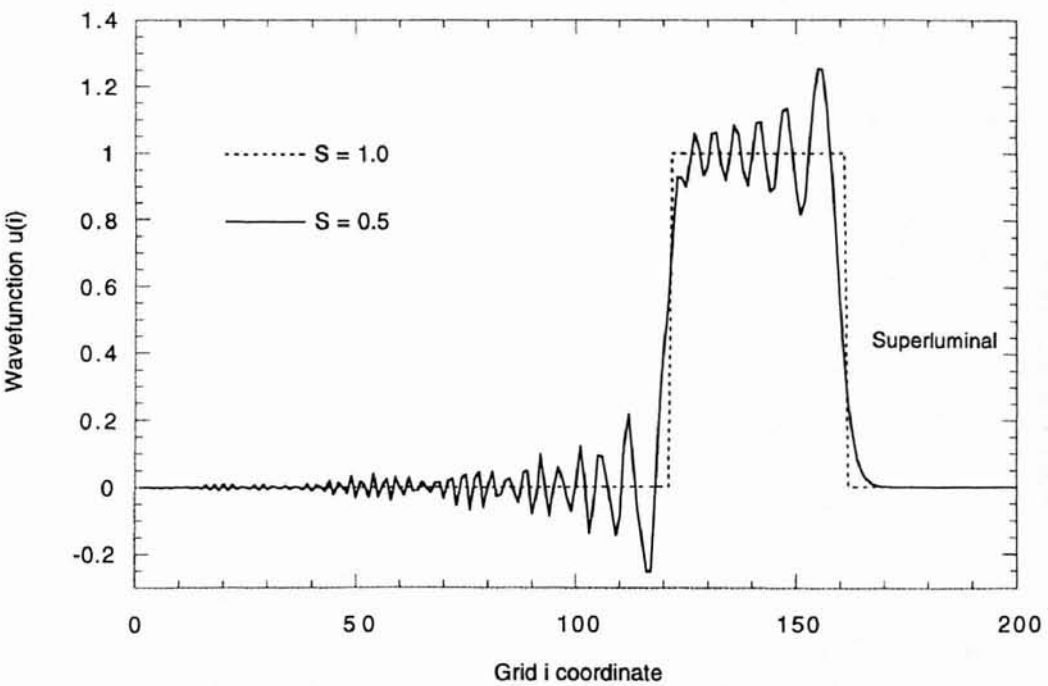


Fig. 2.2 Percent numerical phase-velocity error relative to the free-space speed of light as a function of the grid sampling density $(3 \le N_\lambda \le 80)$ for a Courant stability factor $S = 0.5$.

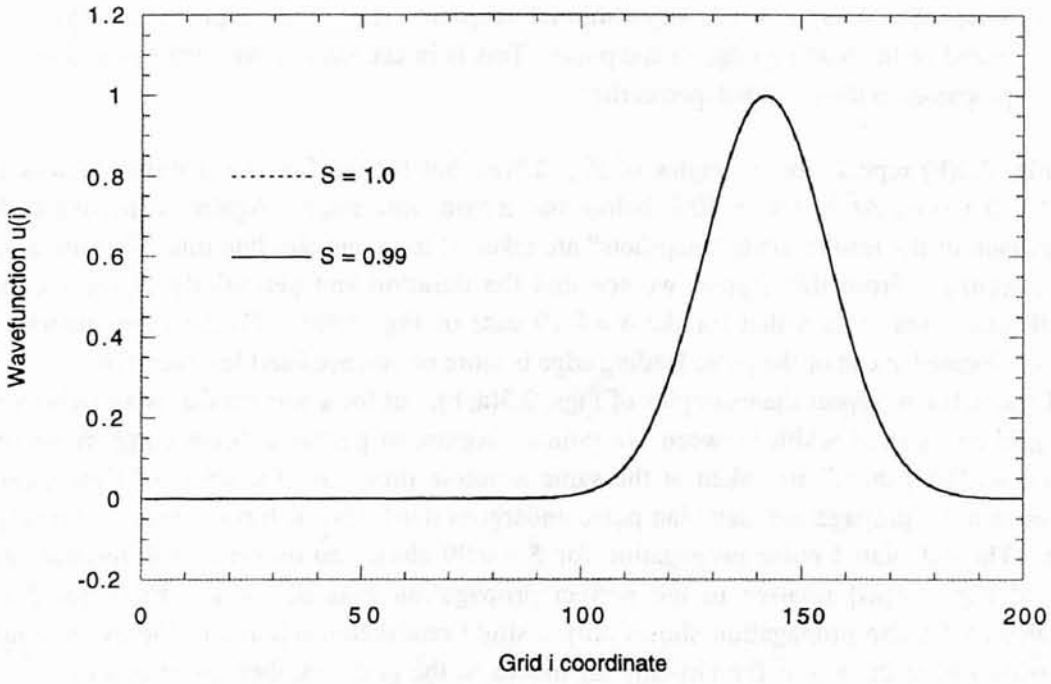


(a) Comparison of calculated pulse propagation for $S = 1$ and $S = 0.99$.

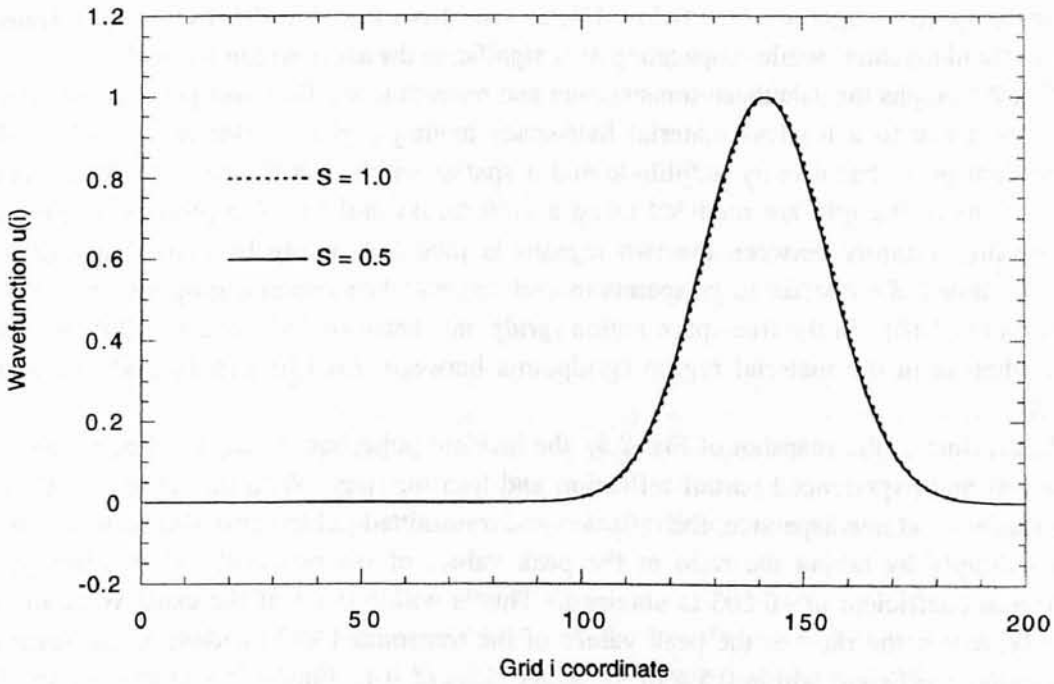


(b) Comparison of calculated pulse propagation for $S = 1$ and $S = 0.5$.

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

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, b) repeat the examples of Figs. 2.3(a, 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.

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 gridpoints in each region when implementing the time-stepping algorithm of (2.16). In the free-space region (gridpoints between $i = 1$ and $i = 139$) we assume $S = 1$, whereas in the material region (gridpoints 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 gridpoint.

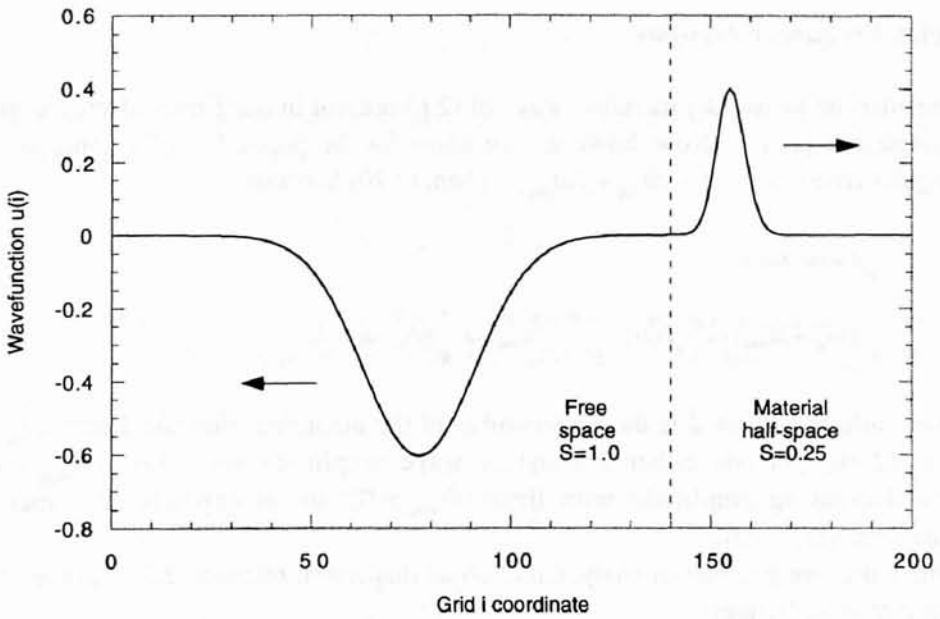


Fig. 2.5 Example of the calculated reflection and transmission of a Gaussian pulse at an interface between free space and a lossless material half-space having $v_p = c/4$.

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 for explicit numerical differential-equation solvers. It causes the computed results to spuriously increase without limit during time-marching.

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 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, $\tilde{\omega} = \tilde{\omega}_{\text{real}} + j\tilde{\omega}_{\text{imag}}$. Then, (2.20) becomes

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

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 ($\tilde{\omega}_{\text{imag}} = 0$), an exponentially decreasing amplitude with time ($\tilde{\omega}_{\text{imag}} > 0$), or an exponentially increasing amplitude with time ($\tilde{\omega}_{\text{imag}} < 0$).

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

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

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

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

where

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

We observe from (2.42b) that $1 - 2S^2 \le \xi \le 1$ 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 value of the cosine function is bounded by $-1 \le \cos(\tilde{k}\Delta x) \le 1$.) Now, consider dividing the $1 - 2S^2 \le \xi \le 1$ range into two subranges, as follows:

(a) $-1 \le \xi \le 1$. This subrange exists for $0 \le S \le 1$. Here, $\sin^{-1}(\xi)$ is real-valued, and hence, real values of $\tilde{\omega}$ are obtained in (2.42a). With $\tilde{\omega}_{\text{imag}} = 0$, (2.40) yields a constant wave amplitude with time.

(b) $1 - 2S^2 \le \xi < -1$. This subrange exists for $1 - 2S^2 < -1$, or equivalently $S > 1$. The most negative value in this subrange, $\xi_{\text{lower bound}}$, occurs for $\cos(\tilde{k}\Delta x) = -1$; that is, $\tilde{k}\Delta x = \pi$ (equivalent to a numerical wavelength $\tilde{\lambda} = 2\Delta x$). This lower bound is given by

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

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

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

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

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

Factoring out $j = \exp(j\pi/2)$ in the argument of the natural logarithm, we obtain

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

Upon taking the natural logarithm, we further obtain

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

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

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

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

$$u_i^n = e^{\left[-n \ln\left(-\xi - \sqrt{\xi^2 - 1}\right)\right]} e^{j[(\pi/\Delta t)(n\Delta t) - \tilde{k}\, i\Delta x]}$$

$$= \left(\frac{1}{-\xi - \sqrt{\xi^2 - 1}}\right)^{**n} e^{j[(\pi/\Delta t)(n\Delta t) - \tilde{k}\, i\Delta x]} \tag{2.47}$$

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:

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

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., $\xi_{\text{lowerbound}}$ defined in (2.43)]. As stated earlier in the context of (2.43), this occurs for $\tilde{k}\Delta x = \pi$ (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 $S \geq 1$:

$$q_{\text{growth}} = -\left(1 - 2S^2\right) + \sqrt{\left(1 - 2S^2\right)^2 - 1}$$

$$= 2S^2 + 2S\sqrt{S^2 - 1} - 1 = \left(S + \sqrt{S^2 - 1}\right)^2 \tag{2.49}$$

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 except Δt:

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

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

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

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

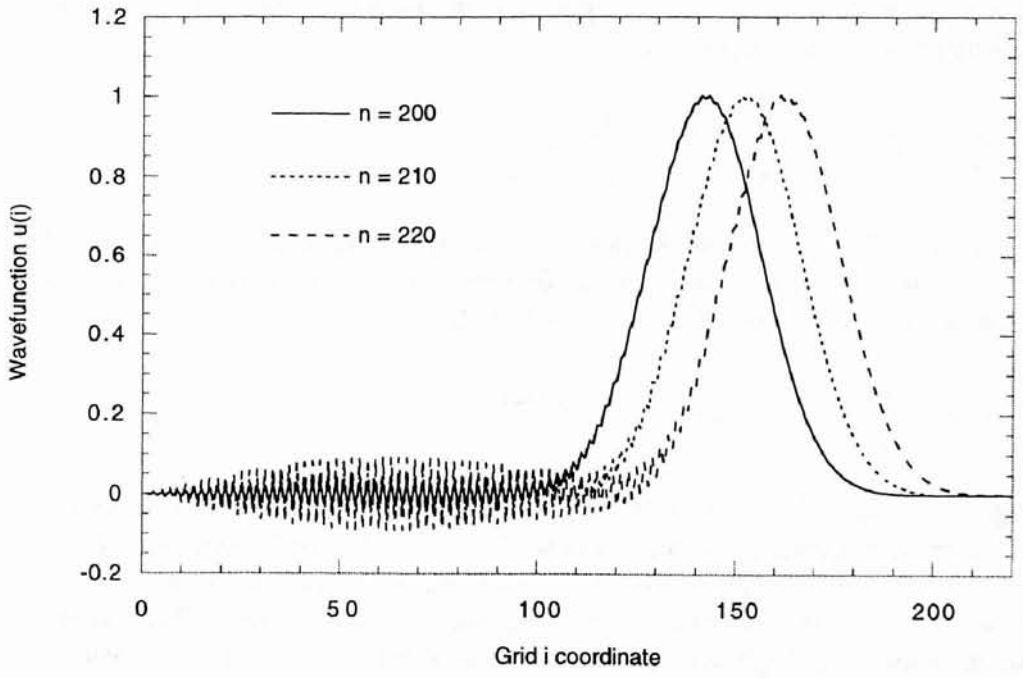
$$S \equiv \frac{c\,\Delta t}{\Delta x} \leq 1 \quad \longrightarrow \quad \Delta t \leq \frac{\Delta x}{c} \tag{2.52}$$

following the notation of (2.28a) and [1]. S is called the numerical stability factor or Courant number. In most modeling problems, Δx is the first 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, 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.

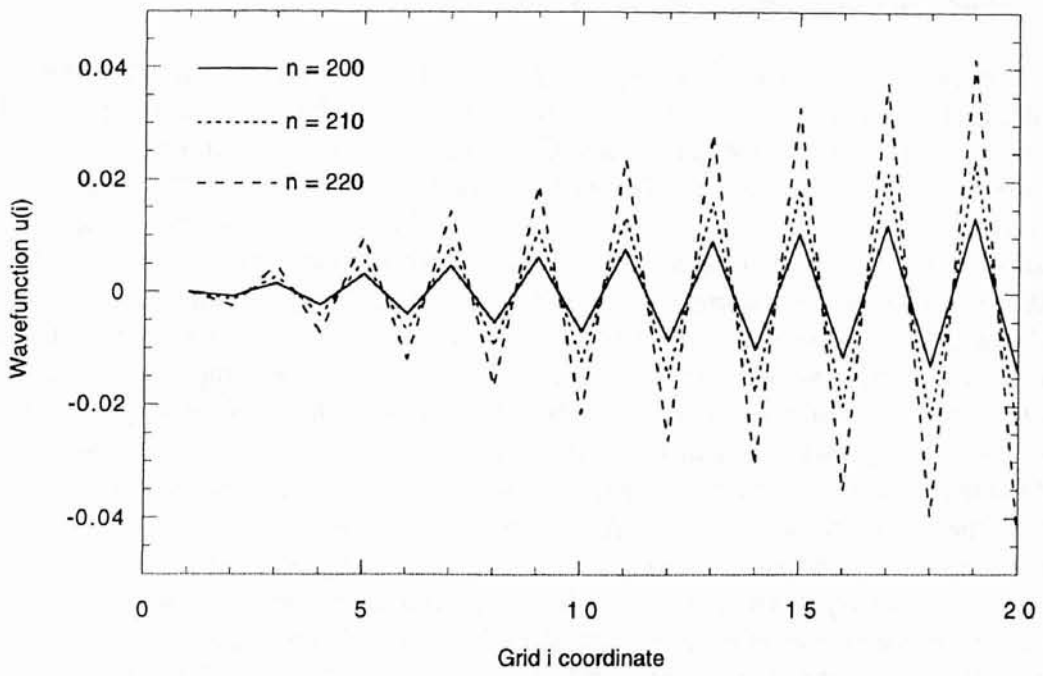
2.7.2 Examples of Calculations Involving Numerical Instability

We first consider an example of the beginning of a numerical instability arising because the Courant stability condition is violated equally at *every* gridpoint. 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 being greater than 1. 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 gridpoints $i = 1$ 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 to 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 gridpoint.

We next consider an example of the beginning of a numerical instability arising because the Courant stability condition is violated at only a *single* gridpoint. 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 $S = 1.075$.[1] 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 gridpoint 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.

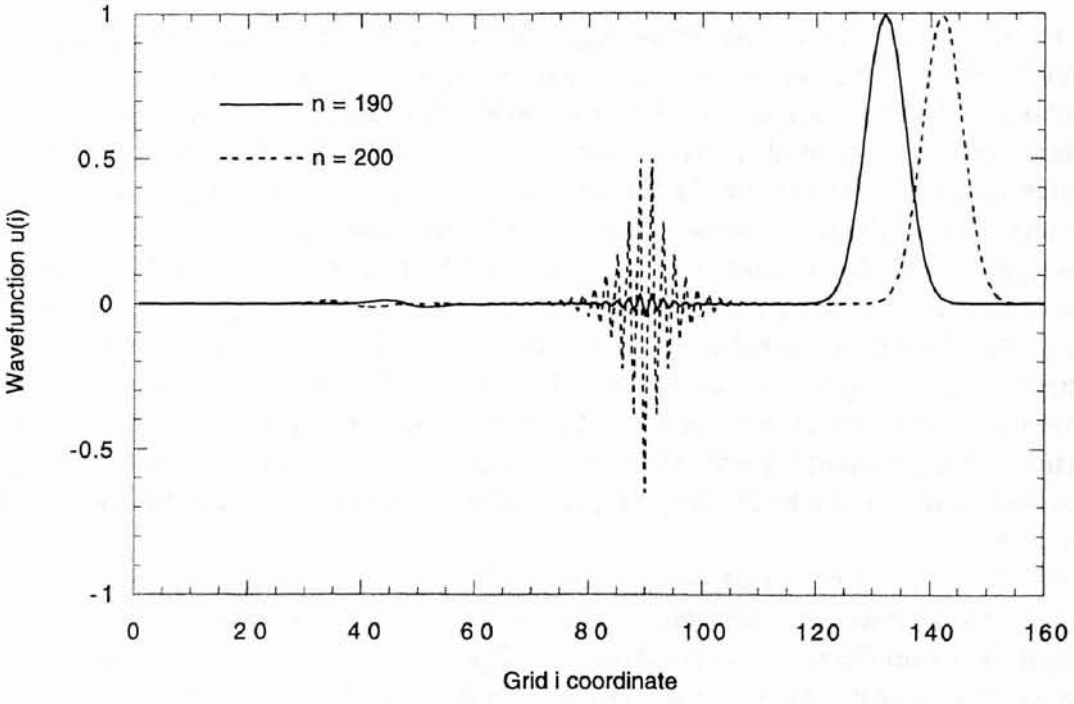
Fig. 2.7(b) is an expanded view of Fig. 2.7(a) between gridpoints $i = 70$ and $i = 110$, showing how the calculated noise due to the numerical instability originates at gridpoint $i = 90$. Again, the noise oscillates with a spatial period of 2 grid cells (i.e., $\tilde{\lambda} = 2\Delta x$). However, the rate of exponential growth here is much less than that predicted by the theory of Section 2.7.1, wherein *all* gridpoints were assumed to violate Courant stability. Upon analyzing the raw data underlying Fig. 2.7(b), the exponential growth factor $q \cong 15$ is observed for every 10-time-step advance of the algorithm.[1] This compares to $q \cong 2{,}200$ per 10 time-steps determined by substituting $S = 1.075$ into (2.49). Thus, it is clear that a grid having just a *single point* of numerical instability can "blow up." However, this process occurs much more slowly than a uniformly unstable grid having a comparable or even smaller Courant factor S.

2.8 SUMMARY

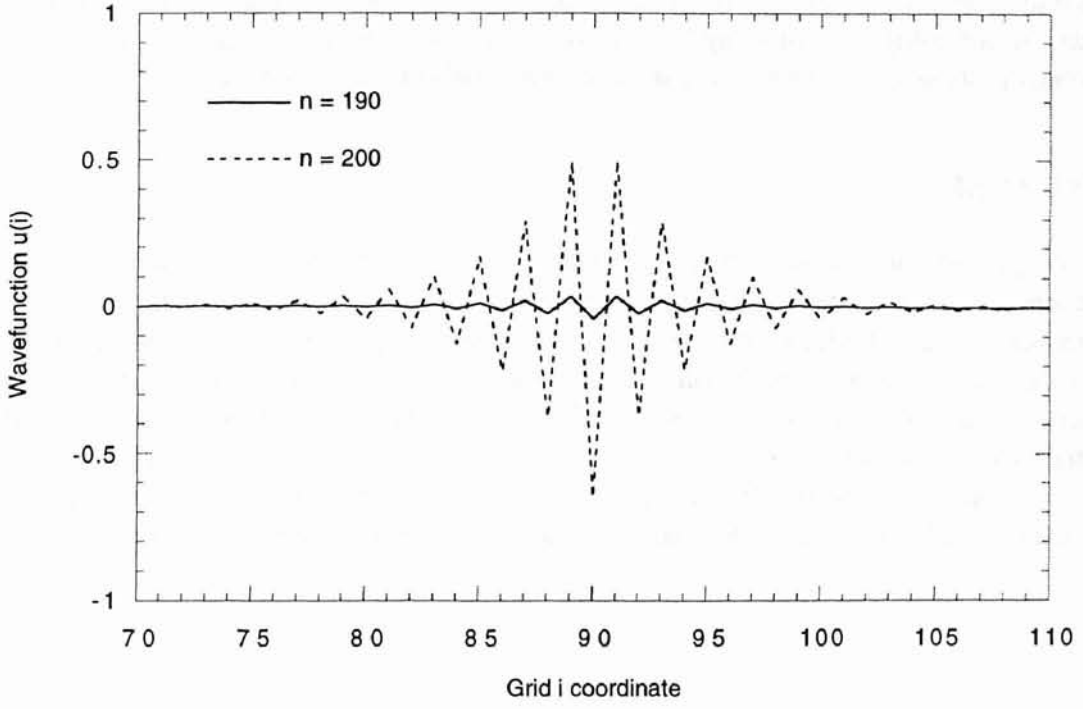
This chapter introduced the numerical FDTD solution of the one-dimensional scalar wave equation. First, the analytical propagating-wave solutions were obtained. Then, finite differences were introduced and applied to the wave equation, leading to introductory discussions of numerical dispersion, numerical 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.

---

[1]For this example, numerical experiments have shown that the observed rate of exponential growth of the numerical instability is a function of the precise details of the word length, arithmetic operations, and round-off processes implemented by a particular computer and its software. Therefore, readers who attempt to replicate this example may require values of S differing from that stated here.



(a) Comparison of calculated pulse propagation at $n = 190$ and $n = 200$ time-steps over grid coordinates $i = 1$ through $i = 160$.



(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 gridpoints but $i = 90$, where $S = 1.075$.

APPENDIX 2A: ORDER OF ACCURACY[2]

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

$$\frac{\partial u}{\partial t} = \mathcal{L} u \tag{2A.1}$$

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

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

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

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

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 a permittivity or permeability jump across a material interface.

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].

---

[2]This 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 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.

SELECTED 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, 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 Consider the numerical dispersion relation for the second-order-accurate finite-difference solution to the one-dimensional scalar wave equation for free-space.

a) Calculate and graph $\tilde{v}_p/c$, the numerical phase velocity normalized to the free-space speed of light, as a function of the grid sampling density for $N_\lambda|_{\text{transition}} \le N_\lambda \le 10$, assuming stability factors $S = 0.7$, 0.8, 0.9, and 0.99.

b) Calculate and graph the percent numerical phase-velocity error relative to the free-space speed of light as a function of the grid sampling density for $N_\lambda|_{\text{transition}} \le N_\lambda \le 100$, assuming stability factors $S = 0.7$, 0.8, 0.9, and 0.99. Use a logarithmic scale for the error axis.

c) What grid sampling density is needed to suppress the numerical phase-velocity errors to less than 0.1% for a stability factor $S = 0.9$?

2.8 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, b).

2.9 Use the scalar-wave-equation computer program developed in Problem 2.8 to replicate the graphical results of Figs. 2.4(a, b).

2.10 Use the scalar-wave-equation computer program developed in Problem 2.8 to replicate the graphical results of Fig. 2.5.

2.11 Use the scalar-wave-equation computer program developed in Problem 2.8 to replicate the graphical results of Figs. 2.6(a, b).

2.12 Use the scalar-wave-equation computer program developed in Problem 2.8 to replicate the graphical results of Figs. 2.7(a, b).

Chapter 3

Introduction to Maxwell's Equations and the Yee Algorithm

Allen Taflove and Jamesina Simpson

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 geometry for spatially sampling the electric and magnetic field vector components which 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 40 years since Yee's paper. However, none have had the seminal impact and longevity of his "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:

$$\frac{\partial \boldsymbol{B}}{\partial t} = -\nabla \times \boldsymbol{E} - \boldsymbol{M} \tag{3.1a}$$

$$\frac{\partial}{\partial t}\iint_A \boldsymbol{B} \cdot d\boldsymbol{A} = -\oint_L \boldsymbol{E} \cdot d\boldsymbol{L} - \iint_A \boldsymbol{M} \cdot d\boldsymbol{A} \tag{3.1b}$$

Ampere's law:

$$\frac{\partial \boldsymbol{D}}{\partial t} = \nabla \times \boldsymbol{H} - \boldsymbol{J} \tag{3.2a}$$

$$\frac{\partial}{\partial t}\iint_A \boldsymbol{D} \cdot d\boldsymbol{A} = \oint_L \boldsymbol{H} \cdot d\boldsymbol{L} - \iint_A \boldsymbol{J} \cdot d\boldsymbol{A} \tag{3.2b}$$

Gauss' law for the electric field:

$$\nabla \cdot \boldsymbol{D} = 0 \tag{3.3a}$$

$$\oiint_A \boldsymbol{D} \cdot d\boldsymbol{A} = 0 \tag{3.3b}$$

Gauss' law for the magnetic field:

$$\nabla \cdot \boldsymbol{B} = 0 \tag{3.4a}$$

$$\oiint_A \boldsymbol{B} \cdot d\boldsymbol{A} = 0 \tag{3.4b}$$

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

- $\boldsymbol{E}$: electric field (volts / meter)
- $\boldsymbol{D}$: electric flux density (coulombs / meter^2)
- $\boldsymbol{H}$: magnetic field (amperes / meter)
- $\boldsymbol{B}$: magnetic flux density (webers / meter^2)
- A : arbitrary three-dimensional surface
- $d\boldsymbol{A}$: differential normal vector that characterizes surface A (meter^2)
- L : closed contour that bounds surface A
- $d\boldsymbol{L}$: differential length vector that characterizes contour L (meters)
- $\boldsymbol{J}$: electric current density (amperes / meter^2)
- $\boldsymbol{M}$: equivalent magnetic current density (volts / meter^2)

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

$$\boldsymbol{D} = \varepsilon \boldsymbol{E} = \varepsilon_r \varepsilon_0 \boldsymbol{E} \; ; \qquad \boldsymbol{B} = \mu \boldsymbol{H} = \mu_r \mu_0 \boldsymbol{H} \tag{3.5}$$

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 $\boldsymbol{J}$ and $\boldsymbol{M}$ can act as *independent sources* of E- and H-field energy, $\boldsymbol{J}_{\text{source}}$ and $\boldsymbol{M}_{\text{source}}$. 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

$$\boldsymbol{J} = \boldsymbol{J}_{\text{source}} + \sigma \boldsymbol{E} \; ; \qquad \boldsymbol{M} = \boldsymbol{M}_{\text{source}} + \sigma^{*} \boldsymbol{H} \tag{3.6}$$

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:

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

$$\frac{\partial \boldsymbol{E}}{\partial t} = \frac{1}{\varepsilon} \nabla \times \boldsymbol{H} - \frac{1}{\varepsilon} \left(\boldsymbol{J}_{\text{source}} + \sigma \boldsymbol{E} \right) \tag{3.8}$$

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:

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

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

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

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

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

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

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 lattice must be structured so that the Gauss' law relations are *implicit* in the positions of the $\boldsymbol{E}$ and $\boldsymbol{H}$ components, 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 algorithm.

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

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

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

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

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

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

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

3.3.1 TM_z Mode

Consider grouping (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:

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

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

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

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:

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

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

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

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

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

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

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

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:

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

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

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

$$\frac{\partial E_x}{\partial t} = -\frac{1}{\varepsilon}\left(J_{\text{source}_x} + \sigma E_x\right) \tag{3.17a}$$

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

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

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:

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

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

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):

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

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

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

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

$$\frac{\partial^2 H_y}{\partial t^2} = \frac{1}{\mu}\cdot\frac{1}{\varepsilon}\cdot\frac{\partial^2 H_y}{\partial x^2} = c^2\frac{\partial^2 H_y}{\partial x^2} \tag{3.19c}$$

where $c = 1/\sqrt{\mu\varepsilon}$. 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 to be equal to 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 $\varepsilon = \varepsilon_0$, $c \equiv 3\times 10^8$ m/s, 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:

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

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

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

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

$$\frac{\partial^2 E_z}{\partial t^2} = \frac{1}{\varepsilon}\cdot\frac{1}{\mu}\cdot\frac{\partial^2 E_z}{\partial x^2} = c^2\frac{\partial^2 E_z}{\partial x^2} \tag{3.20c}$$

where again $c = 1/\sqrt{\mu\varepsilon}$. 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 reader 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, 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 $\boldsymbol{E}$ and $\boldsymbol{H}$ boundary conditions are enforced on the surface of a material structure.
- Using both $\boldsymbol{E}$ and $\boldsymbol{H}$ 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 $\boldsymbol{H}$ singularities near edges and corners, azimuthal (looping) $\boldsymbol{H}$ singularities near thin wires, and radial $\boldsymbol{E}$ 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 $\boldsymbol{E}$ and $\boldsymbol{H}$ components in three-dimensional space so that every $\boldsymbol{E}$ component is surrounded by four circulating $\boldsymbol{H}$ components, and every $\boldsymbol{H}$ component is surrounded by four circulating $\boldsymbol{E}$ components.

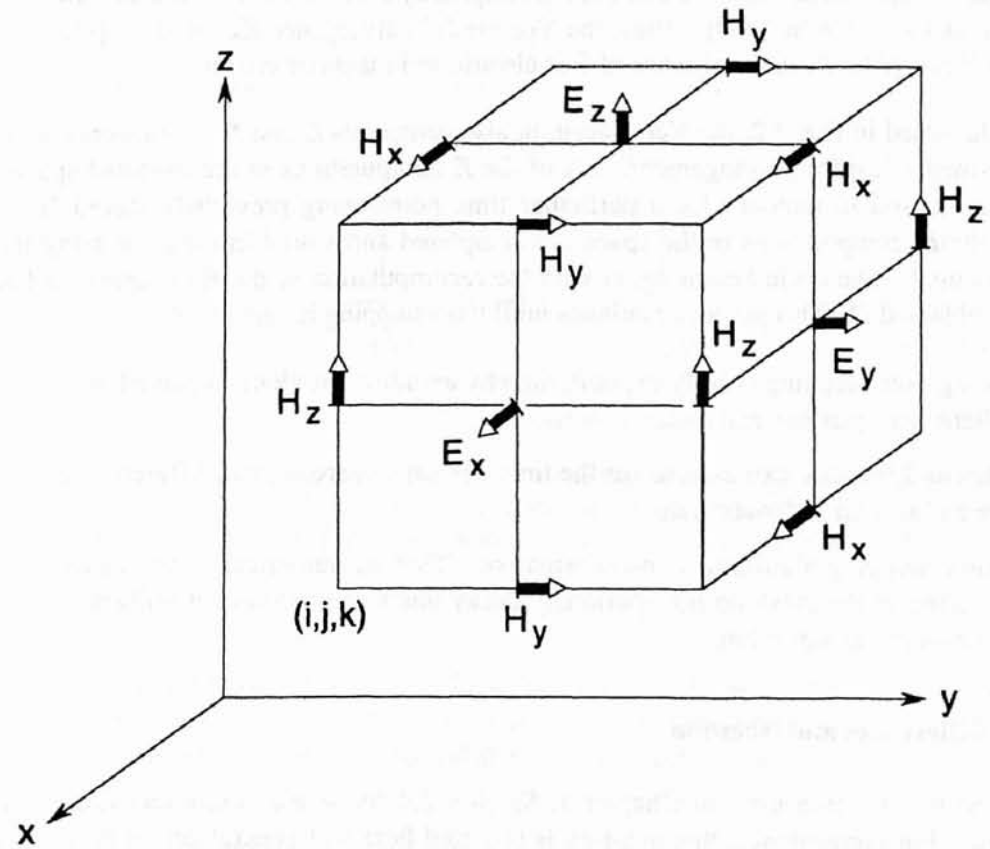


Fig. 3.1 Position of the electric and magnetic field vector components about a cubic unit cell of the Yee space lattice. *After:* K. S. Yee, *IEEE Trans. Antennas and Propagation*, Vol. 14, 1966, pp. 302–307, © 1966 IEEE.

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 $\boldsymbol{E}$ components associated with displacement current flux linking $\boldsymbol{H}$ loops, as well as $\boldsymbol{H}$ components associated with magnetic flux linking $\boldsymbol{E}$ 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 $\boldsymbol{E}$ and $\boldsymbol{H}$ 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 $\boldsymbol{E}$ and $\boldsymbol{H}$ 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 $\boldsymbol{E}$ and $\boldsymbol{H}$ fields in the absence of free electric and magnetic charge.

3. As illustrated in Fig. 3.2, the Yee algorithm also centers its $\boldsymbol{E}$ and $\boldsymbol{H}$ components in time, in what is termed a leapfrog arrangement. All of the $\boldsymbol{E}$ computations in the modeled space are completed and stored in memory for a particular time point using previously stored $\boldsymbol{H}$ data. Then, all of the $\boldsymbol{H}$ computations in the space are completed and stored in memory using the $\boldsymbol{E}$ data just computed. The cycle begins again with the recomputation of the $\boldsymbol{E}$ components based on the newly obtained $\boldsymbol{H}$. 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 Chapter 2, Section 2.4 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

$$(i, j, k) = (i\Delta x, j\Delta y, k\Delta z) \tag{3.21}$$

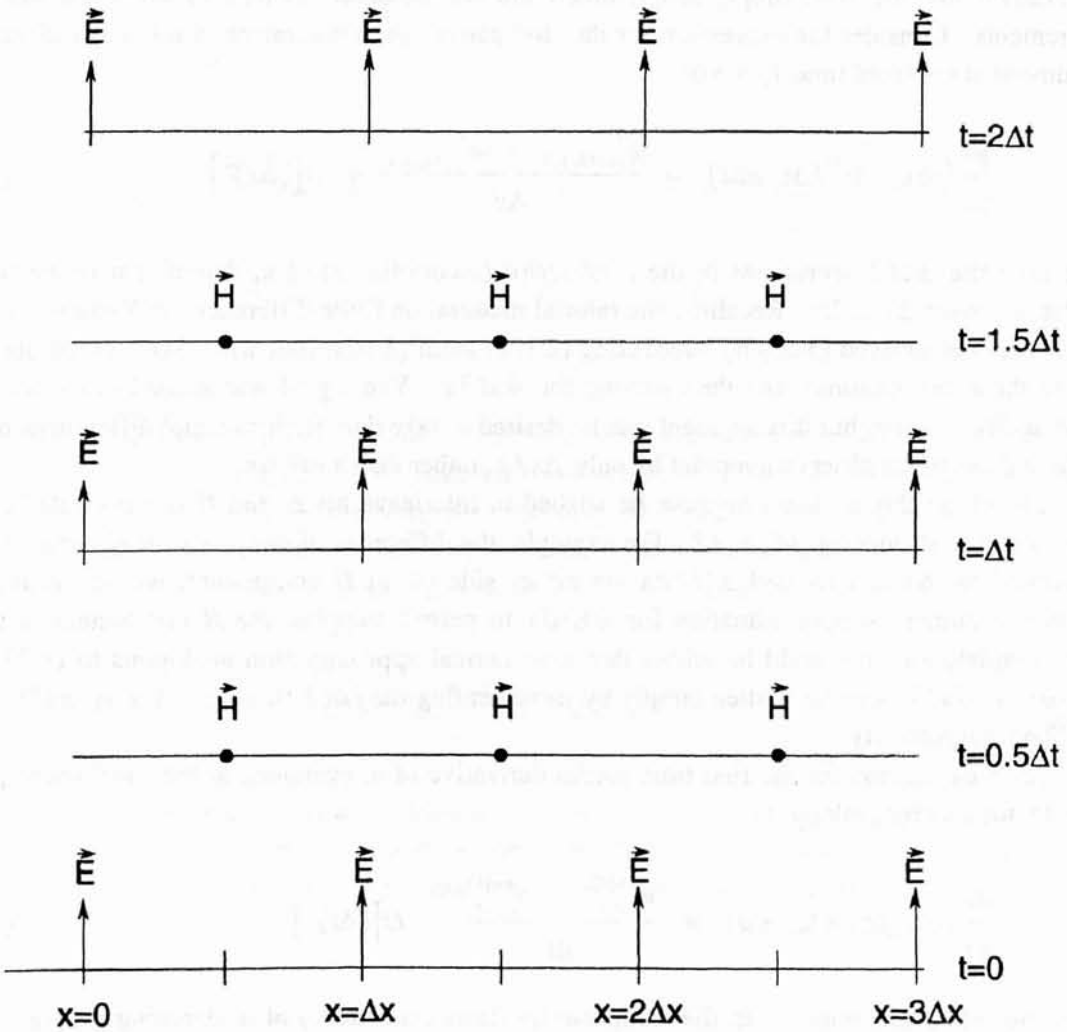


Fig. 3.2 Space-time chart of the Yee algorithm for a one-dimensional wave propagation example showing the use of central differences for the space derivatives and leapfrog for the time derivatives. Initial conditions for both electric and magnetic fields are zero everywhere in the grid.

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

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

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 $t_n = n\Delta t$:

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

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 $\boldsymbol{E}$ and $\boldsymbol{H}$ components in the space lattice at intervals of $\Delta x/2$. For example, the difference of two adjacent $\boldsymbol{E}$ components, separated by Δx and located $\pm 1/2\,\Delta x$ on either side of an $\boldsymbol{H}$ component, would be used to provide a numerical approximation for $\partial E/\partial x$ to permit stepping the $\boldsymbol{H}$ 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:

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

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 $\boldsymbol{E}$ and $\boldsymbol{H}$ 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 (3.9) and (3.10). We begin by considering the E_x field-component equation (3.10a), repeated here for convenience:

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

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

$$\frac{E_x\big|_{i,\,j+1/2,\,k+1/2}^{n+1/2} - E_x\big|_{i,\,j+1/2,\,k+1/2}^{n-1/2}}{\Delta t} =$$

$$\frac{1}{\varepsilon_{i,\,j+1/2,\,k+1/2}} \cdot \left(\begin{array}{l} \dfrac{H_z\big|_{i,\,j+1,\,k+1/2}^{n} - H_z\big|_{i,\,j,\,k+1/2}^{n}}{\Delta y} - \dfrac{H_y\big|_{i,\,j+1/2,\,k+1}^{n} - H_y\big|_{i,\,j+1/2,\,k}^{n}}{\Delta z} \\ - J_{\text{source}_x}\big|_{i,\,j+1/2,\,k+1/2}^{n} - \sigma_{i,\,j+1/2,\,k+1/2}\, E_x\big|_{i,\,j+1/2,\,k+1/2}^{n} \end{array} \right) \tag{3.25}$$

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*:

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

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

$$E_x\big|_{i,\,j+1/2,\,k+1/2}^{n+1/2} - E_x\big|_{i,\,j+1/2,\,k+1/2}^{n-1/2} =$$

$$\frac{\Delta t}{\varepsilon_{i,\,j+1/2,\,k+1/2}} \cdot \left[\begin{array}{l} \dfrac{H_z\big|_{i,\,j+1,\,k+1/2}^{n} - H_z\big|_{i,\,j,\,k+1/2}^{n}}{\Delta y} - \dfrac{H_y\big|_{i,\,j+1/2,\,k+1}^{n} - H_y\big|_{i,\,j+1/2,\,k}^{n}}{\Delta z} - \\ J_{\text{source}_x}\big|_{i,\,j+1/2,\,k+1/2}^{n} - \sigma_{i,\,j+1/2,\,k+1/2} \cdot \left(\dfrac{E_x\big|_{i,\,j+1/2,\,k+1/2}^{n+1/2} + E_x\big|_{i,\,j+1/2,\,k+1/2}^{n-1/2}}{2} \right) \end{array} \right] \tag{3.27}$$

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

$$\left(1 + \frac{\sigma_{i,j+1/2,k+1/2}\,\Delta t}{2\varepsilon_{i,j+1/2,k+1/2}}\right) E_x\Big|^{n+1/2}_{i,j+1/2,k+1/2} = \left(1 - \frac{\sigma_{i,j+1/2,k+1/2}\,\Delta t}{2\varepsilon_{i,j+1/2,k+1/2}}\right) E_x\Big|^{n-1/2}_{i,j+1/2,k+1/2}$$

$$+ \frac{\Delta t}{\varepsilon_{i,j+1/2,k+1/2}} \cdot \left(\frac{H_z\Big|^n_{i,j+1,k+1/2} - H_z\Big|^n_{i,j,k+1/2}}{\Delta y} - \frac{H_y\Big|^n_{i,j+1/2,k+1} - H_y\Big|^n_{i,j+1/2,k}}{\Delta z} - J_{\text{source}_x}\Big|^n_{i,j+1/2,\,k+1/2} \right) \tag{3.28}$$

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

$$E_x\Big|^{n+1/2}_{i,j+1/2,k+1/2} = \left(\frac{1 - \dfrac{\sigma_{i,j+1/2,k+1/2}\,\Delta t}{2\varepsilon_{i,j+1/2,k+1/2}}}{1 + \dfrac{\sigma_{i,j+1/2,k+1/2}\,\Delta t}{2\varepsilon_{i,j+1/2,k+1/2}}} \right) E_x\Big|^{n-1/2}_{i,j+1/2,k+1/2} \tag{3.29a}$$

$$+ \left(\frac{\dfrac{\Delta t}{\varepsilon_{i,j+1/2,k+1/2}}}{1 + \dfrac{\sigma_{i,j+1/2,k+1/2}\,\Delta t}{2\varepsilon_{i,j+1/2,k+1/2}}} \right) \cdot \left(\frac{H_z\Big|^n_{i,j+1,k+1/2} - H_z\Big|^n_{i,j,k+1/2}}{\Delta y} - \frac{H_y\Big|^n_{i,j+1/2,k+1} - H_y\Big|^n_{i,j+1/2,k}}{\Delta z} - J_{\text{source}_x}\Big|^n_{i,j+1/2,\,k+1/2} \right)$$

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 $E_x\Big|^{n+1/2}$. 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 $\boldsymbol{E}$ components normal to the remaining visible faces of the unit cell:

$$E_y\Big|_{i-1/2,\,j+1,\,k+1/2}^{n+1/2} = \left(\frac{1 - \dfrac{\sigma_{i-1/2,\,j+1,\,k+1/2}\,\Delta t}{2\,\varepsilon_{i-1/2,\,j+1,\,k+1/2}}}{1 + \dfrac{\sigma_{i-1/2,\,j+1,\,k+1/2}\,\Delta t}{2\,\varepsilon_{i-1/2,\,j+1,\,k+1/2}}}\right) E_y\Big|_{i-1/2,\,j+1,\,k+1/2}^{n-1/2}$$

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

$$E_z\Big|_{i-1/2,\,j+1/2,\,k+1}^{n+1/2} = \left(\frac{1 - \dfrac{\sigma_{i-1/2,\,j+1/2,\,k+1}\,\Delta t}{2\,\varepsilon_{i-1/2,\,j+1/2,\,k+1}}}{1 + \dfrac{\sigma_{i-1/2,\,j+1/2,\,k+1}\,\Delta t}{2\,\varepsilon_{i-1/2,\,j+1/2,\,k+1}}}\right) E_z\Big|_{i-1/2,\,j+1/2,\,k+1}^{n-1/2}$$

$$+ \left(\frac{\dfrac{\Delta t}{\varepsilon_{i-1/2,\,j+1/2,\,k+1}}}{1 + \dfrac{\sigma_{i-1/2,\,j+1/2,\,k+1}\,\Delta t}{2\varepsilon_{i-1/2,\,j+1/2,\,k+1}}}\right) \cdot \left(\begin{array}{l} \dfrac{H_y\Big|_{i,\,j+1/2,\,k+1}^{n} - H_y\Big|_{i-1,\,j+1/2,\,k+1}^{n}}{\Delta x} \\ - \dfrac{H_x\Big|_{i-1/2,\,j+1,\,k+1}^{n} - H_x\Big|_{i-1/2,\,j,\,k+1}^{n}}{\Delta y} \\ - J_{\text{source}_z}\Big|_{i-1/2,\,j+1/2,\,k+1}^{n} \end{array}\right) \tag{3.29c}$$

By analogy, we can derive finite-difference equations for (3.9a) to (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:

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

$$+\left(\frac{\dfrac{\Delta t}{\mu_{i-1/2,\,j+1,\,k+1}}}{1+\dfrac{\sigma^*_{i-1/2,\,j+1,\,k+1}\,\Delta t}{2\mu_{i-1/2,\,j+1,\,k+1}}}\right)\cdot\left(\begin{array}{l}\dfrac{E_y\Big|_{i-1/2,\,j+1,\,k+3/2}^{n+1/2}-E_y\Big|_{i-1/2,\,j+1,\,k+1/2}^{n+1/2}}{\Delta z}\\[2ex] -\dfrac{E_z\Big|_{i-1/2,\,j+3/2,\,k+1}^{n+1/2}-E_z\Big|_{i-1/2,\,j+1/2,\,k+1}^{n+1/2}}{\Delta y}\\[2ex] -M_{\text{source}_x}\Big|_{i-1/2,\,j+1,\,k+1}^{n+1/2}\end{array}\right) \tag{3.30a}$$

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

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

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

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

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

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

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):

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

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

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

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

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

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

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

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:

$$E_x\big|^{n+1/2}_{i,\,j+1/2,\,k+1/2} = C_{a,E_x}\big|_{i,\,j+1/2,\,k+1/2} E_x\big|^{n-1/2}_{i,\,j+1/2,\,k+1/2} + C_{b,E_x}\big|_{i,\,j+1/2,k+1/2} \cdot \left(\begin{array}{l} H_z\big|^n_{i,\,j+1,\,k+1/2} - H_z\big|^n_{i,\,j,\,k+1/2} + \\ H_y\big|^n_{i,\,j+1/2,\,k} - H_y\big|^n_{i,\,j+1/2,\,k+1} - J_{\text{source}_x}\big|^n_{i,\,j+1/2,\,k+1/2}\,\Delta \end{array} \right) \tag{3.33a}$$

$$E_y\big|^{n+1/2}_{i-1/2,\,j+1,\,k+1/2} = C_{a,E_y}\big|_{i-1/2,\,j+1,\,k+1/2} E_y\big|^{n-1/2}_{i-1/2,\,j+1,\,k+1/2} + C_{b,E_y}\big|_{i-1/2,\,j+1,\,k+1/2} \cdot \left(\begin{array}{l} H_x\big|^n_{i-1/2,\,j+1,\,k+1} - H_x\big|^n_{i-1/2,\,j+1,\,k} + \\ H_z\big|^n_{i-1,\,j+1,\,k+1/2} - H_z\big|^n_{i,\,j+1,\,k+1/2} - J_{\text{source}_y}\big|^n_{i-1/2,\,j+1,\,k+1/2}\,\Delta \end{array} \right) \tag{3.33b}$$

$$E_z\Big|_{i-1/2,\,j+1/2,\,k+1}^{n+1/2} = C_{a,E_z}\Big|_{i-1/2,\,j+1/2,\,k+1} E_z\Big|_{i-1/2,\,j+1/2,\,k+1}^{n-1/2}$$

$$+ C_{b,E_z}\Big|_{i-1/2,\,j+1/2,\,k+1} \cdot \left(\begin{array}{l} H_y\Big|_{i,\,j+1/2,\,k+1}^{n} - H_y\Big|_{i-1,\,j+1/2,\,k+1}^{n} + \\ H_x\Big|_{i-1/2,\,j,\,k+1}^{n} - H_x\Big|_{i-1/2,\,j+1,\,k+1}^{n} - J_{\text{source}_z}\Big|_{i-1/2,\,j+1/2,\,k+1}^{n} \Delta \end{array} \right) \tag{3.33c}$$

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

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

$$+ D_{b,H_x}\Big|_{i-1/2,\,j+1,\,k+1} \cdot \left(\begin{array}{l} E_y\Big|_{i-1/2,\,j+1,\,k+3/2}^{n+1/2} - E_y\Big|_{i-1/2,\,j+1,\,k+1/2}^{n+1/2} + \\ E_z\Big|_{i-1/2,\,j+1/2,\,k+1}^{n+1/2} - E_z\Big|_{i-1/2,\,j+3/2,\,k+1}^{n+1/2} - M_{\text{source}_x}\Big|_{i-1/2,\,j+1,\,k+1}^{n+1/2} \Delta \end{array} \right) \tag{3.34a}$$

$$H_y\Big|_{i,\,j+1/2,\,k+1}^{n+1} = D_{a,H_y}\Big|_{i,\,j+1/2,\,k+1} H_y\Big|_{i,\,j+1/2,\,k+1}^{n}$$

$$+ D_{b,H_y}\Big|_{i,\,j+1/2,\,k+1} \cdot \left(\begin{array}{l} E_z\Big|_{i+1/2,\,j+1/2,\,k+1}^{n+1/2} - E_z\Big|_{i-1/2,\,j+1/2,\,k+1}^{n+1/2} + \\ E_x\Big|_{i,\,j+1/2,\,k+1/2}^{n+1/2} - E_x\Big|_{i,\,j+1/2,\,k+3/2}^{n+1/2} - M_{\text{source}_y}\Big|_{i,\,j+1/2,\,k+1}^{n+1/2} \Delta \end{array} \right) \tag{3.34b}$$

$$H_z\Big|_{i,\,j+1,\,k+1/2}^{n+1} = D_{a,H_z}\Big|_{i,\,j+1,\,k+1/2} H_z\Big|_{i,\,j+1,\,k+1/2}^{n}$$

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

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, 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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 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, 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 were 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 $\boldsymbol{E}$ and $\boldsymbol{M}$ vectors:

$$\hat{\boldsymbol{E}} = (\Delta t / \mu_0 \Delta)\boldsymbol{E} \;; \qquad \hat{\boldsymbol{M}} = (\Delta t / \mu_0)\boldsymbol{M} \tag{3.37}$$

where $\Delta = \Delta x = \Delta y = \Delta z$ is the space cell size. 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

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

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

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

$$\hat{E}_x\Big|^{n+1/2}_{i,\, j+1/2,\, k+1/2} = C_a(m)\, \hat{E}_x\Big|^{n-1/2}_{i,\, j+1/2,\, k+1/2} + \hat{C}_b(m) \cdot \Big(H_z\Big|^{n}_{i,\, j+1,\, k+1/2} - H_z\Big|^{n}_{i,\, j,\, k+1/2} + H_y\Big|^{n}_{i,\, j+1/2,\, k} - H_y\Big|^{n}_{i,\, j+1/2,\, k+1} - J_{\text{source}_x}\Big|^{n}_{i,\, j+1/2,\, k+1/2} \Delta \Big)$$

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

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

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

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

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

$$H_x\Big|^{n+1}_{i-1/2,\, j+1,\, k+1} = H_x\Big|^{n}_{i-1/2,\, j+1,\, k+1} + \hat{E}_y\Big|^{n+1/2}_{i-1/2,\, j+1,\, k+3/2} - \hat{E}_y\Big|^{n+1/2}_{i-1/2,\, j+1,\, k+1/2} + \hat{E}_z\Big|^{n+1/2}_{i-1/2,\, j+1/2,\, k+1} - \hat{E}_z\Big|^{n+1/2}_{i-1/2,\, j+3/2,\, k+1} - \hat{M}_{\text{source}_x}\Big|^{n+1/2}_{i-1/2,\, j+1,\, k+1} \tag{3.40a}$$

$$H_y\Big|^{n+1}_{i,\, j+1/2,\, k+1} = H_y\Big|^{n}_{i,\, j+1/2,\, k+1} + \hat{E}_z\Big|^{n+1/2}_{i+1/2,\, j+1/2,\, k+1} - \hat{E}_z\Big|^{n+1/2}_{i-1/2,\, j+1/2,\, k+1} + \hat{E}_x\Big|^{n+1/2}_{i,\, j+1/2,\, k+1/2} - \hat{E}_x\Big|^{n+1/2}_{i,\, j+1/2,\, k+3/2} - \hat{M}_{\text{source}_y}\Big|^{n+1/2}_{i,\, j+1/2,\, k+1} \tag{3.40b}$$

$$H_z\Big|_{i,\,j+1,\,k+1/2}^{n+1} = H_z\Big|_{i,\,j+1,\,k+1/2}^{n} + \hat{E}_x\Big|_{i,\,j+3/2,\,k+1/2}^{n+1/2} - \hat{E}_x\Big|_{i,\,j+1/2,\,k+1/2}^{n+1/2} + \hat{E}_y\Big|_{i-1/2,\,j+1,\,k+1/2}^{n+1/2} - \hat{E}_y\Big|_{i+1/2,\,j+1,\,k+1/2}^{n+1/2} - \hat{M}_{\text{source}_z}\Big|_{i,\,j+1,\,k+1/2}^{n+1/2} \tag{3.40c}$$

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 $\mu_0 \Delta / \Delta t$.

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, $k+1$, and so forth, 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 $k+1/2$, $k+3/2$, and so forth, 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:

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

$$E_z\Big|_{i-1/2,\,j+1/2}^{n+1/2} = C_a(m)\, E_z\Big|_{i-1/2,\,j+1/2}^{n-1/2} + C_b(m) \cdot \left(H_y\Big|_{i,\,j+1/2}^{n} - H_y\Big|_{i-1,\,j+1/2}^{n} + H_x\Big|_{i-1/2,\,j}^{n} - H_x\Big|_{i-1/2,\,j+1}^{n} - J_{\text{source}_z}\Big|_{i-1/2,\,j+1/2}^{n} \Delta \right) \tag{3.41a}$$

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

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

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

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

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:

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

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

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

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

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

$$H_z\Big|_{i,\,j+1}^{n+1} = D_a(m)\, H_z\Big|_{i,\,j+1}^{n} + D_b(m)\cdot\Big(E_x\Big|_{i,\,j+3/2}^{n+1/2} - E_x\Big|_{i,\,j+1/2}^{n+1/2} + E_y\Big|_{i-1/2,\,j+1}^{n+1/2} - E_y\Big|_{i+1/2,\,j+1}^{n+1/2} - M_{\text{source}_z}\Big|_{i,\,j+1}^{n+1/2}\Delta\Big) \tag{3.42c}$$

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 simulates 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 and magnetomotive forces 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 remaining components is straightforward.

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

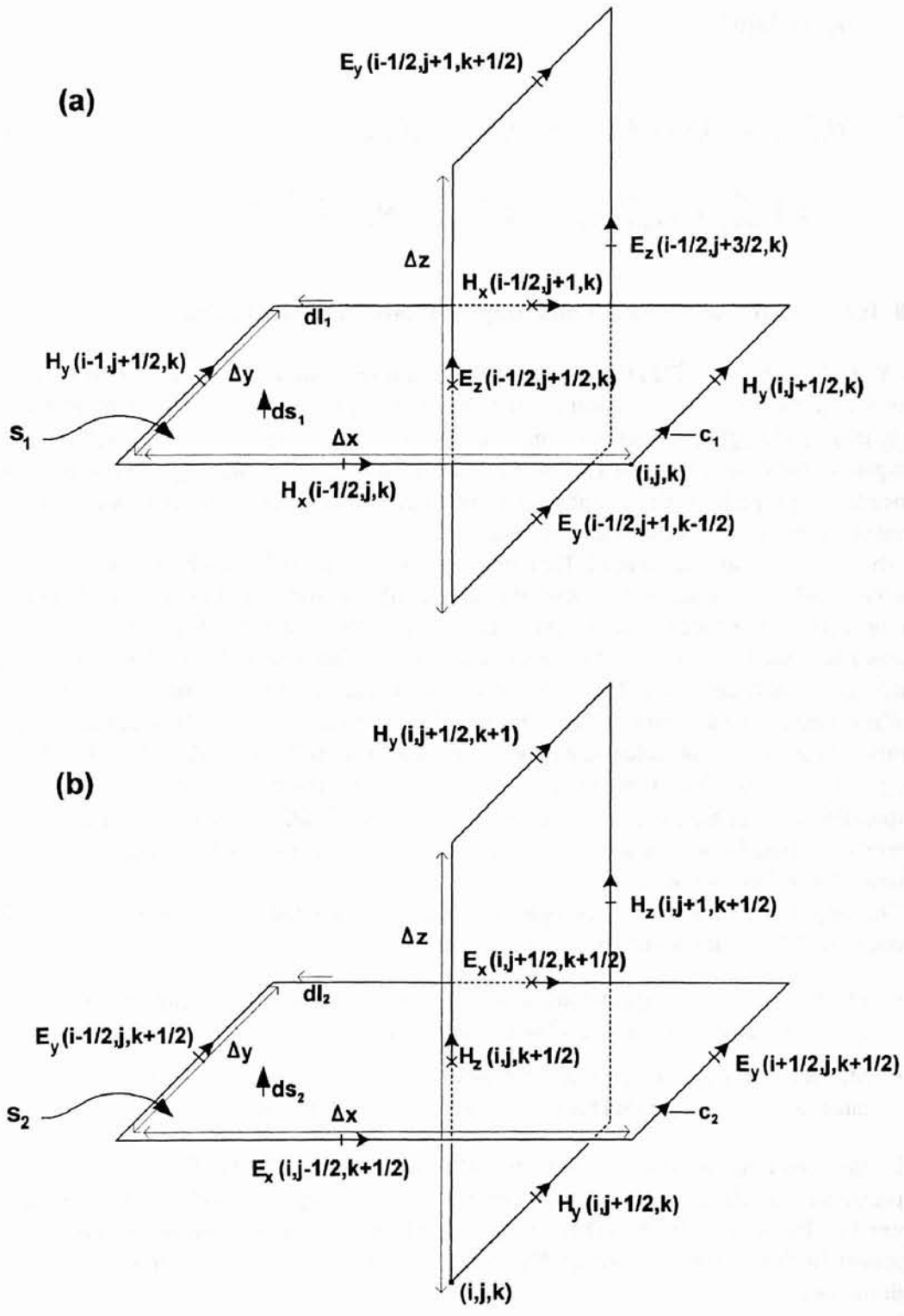


Fig. 3.3 Examples of chain-linked orthogonal contours in the free-space Yee mesh. (a) Ampere's law for time-stepping E_z; (b) Faraday's law for time-stepping H_z. *Adapted from*: Taflove et al., *IEEE Trans. Antennas and Propagation*, 1988, pp. 247–257, © 1988 IEEE.

$$\frac{\partial}{\partial t}\int_{S_1} \boldsymbol{D}\cdot d\boldsymbol{S}_1 \;=\; \oint_{C_1} \boldsymbol{H}\cdot d\boldsymbol{L}_1 \tag{3.43a}$$

$$\begin{aligned}\frac{\partial}{\partial t}\int_{S_1} \varepsilon_0 E_z\big|_{i-1/2,\,j+1/2,\,k}\, dS_1 \;\cong\; & H_x\big|_{i-1/2,\,j,\,k}\Delta x \;+\; H_y\big|_{i,\,j+1/2,\,k}\Delta y \\ & -\; H_x\big|_{i-1/2,\,j+1,\,k}\Delta x \;-\; H_y\big|_{i-1,\,j+1/2,\,k}\Delta y\end{aligned} \tag{3.43b}$$

Now, further assume that $E_z\big|_{i-1/2,\,j+1/2,\,k}$ 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

$$\begin{aligned}\varepsilon_0\,\Delta x\,\Delta y\left(\frac{E_z\big|_{i-1/2,\,j+1/2,\,k}^{n+1/2} \;-\; E_z\big|_{i-1/2,\,j+1/2,\,k}^{n-1/2}}{\Delta t}\right) \;=\; & \left(H_x\big|_{i-1/2,\,j,\,k}^{n} \;-\; H_x\big|_{i-1/2,\,j+1,\,k}^{n}\right)\Delta x \\ & +\; \left(H_y\big|_{i,\,j+1/2,\,k}^{n} \;-\; H_y\big|_{i-1,\,j+1/2,\,k}^{n}\right)\Delta y\end{aligned} \tag{3.43c}$$

Multiplying both sides by $\Delta t/(\varepsilon_0\,\Delta x\,\Delta y)$ and solving for $E_z\big|_{i-1/2,\,j+1/2,\,k}^{n+1/2}$ provides the following time-stepping relation:

$$\begin{aligned}E_z\big|_{i-1/2,\,j+1/2,\,k}^{n+1/2} \;=\; & E_z\big|_{i-1/2,\,j+1/2,\,k}^{n-1/2} \;+\; \left(H_x\big|_{i-1/2,\,j,\,k}^{n} \;-\; H_x\big|_{i-1/2,\,j+1,\,k}^{n}\right)\Delta t/(\varepsilon_0\,\Delta y) \\ & +\; \left(H_y\big|_{i,\,j+1/2,\,k}^{n} \;-\; H_y\big|_{i-1,\,j+1/2,\,k}^{n}\right)\Delta t/(\varepsilon_0\,\Delta x)\end{aligned} \tag{3.44}$$

Equation (3.44) is simply the free-space version of (3.29c), which is the Yee time-stepping equation for E_z that was obtained directly from implementing the curl $\boldsymbol{H}$ 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 apply Faraday's law along contour C_2 in Fig. 3.3(b) to obtain

$$\frac{\partial}{\partial t}\int_{S_2} \boldsymbol{B}\cdot d\boldsymbol{S}_2 \;=\; -\oint_{C_2} \boldsymbol{E}\cdot d\boldsymbol{L}_2 \tag{3.45a}$$

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

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

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

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

Equation (3.46) is simply the free-space version of (3.30c), which is the time-stepping expression for H_z that was obtained directly from implementing the curl $\boldsymbol{E}$ 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 reader.

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:

$$\frac{\partial}{\partial t} \oiint_{\text{Yee cell}} \boldsymbol{D} \cdot d\boldsymbol{S} = \underbrace{\varepsilon_0 \frac{\partial}{\partial t} \left(E_x\big|_{i,j+1/2,k+1/2} - E_x\big|_{i-1,j+1/2,k+1/2} \right) \Delta y \Delta z}_{\text{Term 1}} + \underbrace{\varepsilon_0 \frac{\partial}{\partial t} \left(E_y\big|_{i-1/2,j+1,k+1/2} - E_y\big|_{i-1/2,j,k+1/2} \right) \Delta x \Delta z}_{\text{Term 2}} + \underbrace{\varepsilon_0 \frac{\partial}{\partial t} \left(E_z\big|_{i-1/2,j+1/2,k+1} - E_z\big|_{i-1/2,j+1/2,k} \right) \Delta x \Delta y}_{\text{Term 3}} \tag{3.47}$$

Using the 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:

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

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

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

For all time-steps, this results in

$$\frac{\partial}{\partial t} \oiint_{\text{Yee cell}} \boldsymbol{D} \cdot \boldsymbol{dS} = (\text{Term 1})\Delta y \Delta z + (\text{Term 2})\Delta x \Delta z + (\text{Term 3})\Delta x \Delta y = 0 \tag{3.49}$$

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

$$\oiint_{\text{Yee cell}} \boldsymbol{D}(t) \cdot \boldsymbol{dS} = \oiint_{\text{Yee cell}} \boldsymbol{D}(t = 0) \cdot \boldsymbol{dS} = 0 \tag{3.50}$$

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 tetradecahedron / 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.

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]. Fig. 3.4(a) depicts the unstaggered, collocated grid, in which all E and H components are collocated at a single set of grid-cell vertices. Fig. 3.4(b) depicts 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.

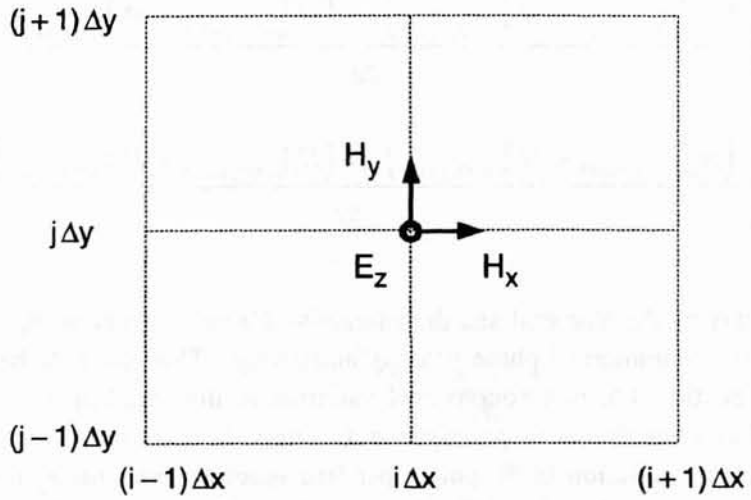
Upon applying second-order-accurate central space differences to the TM_z mode equations of (3.13) for the unstaggered, collocated grid of Fig. 3.4(a) (with a lossless material background assumed for simplicity), we obtain [4]:

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

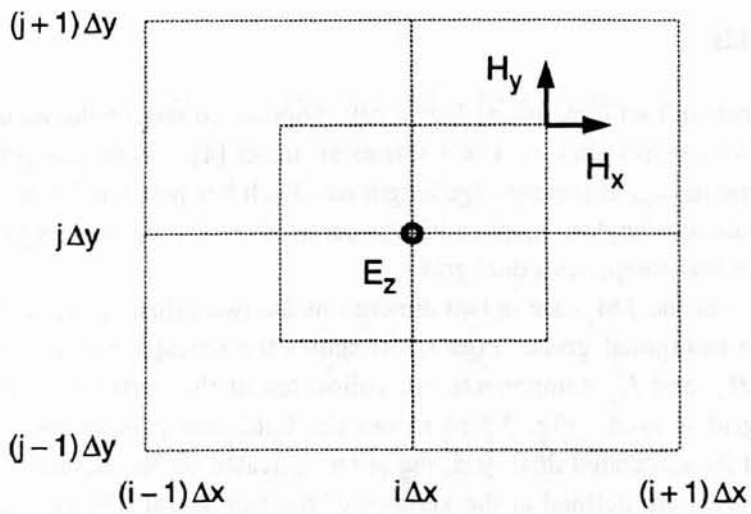
$$\frac{\partial H_y|_{i,j}}{\partial t} = \frac{1}{\mu_{i,j}} \cdot \left(\frac{E_z|_{i+1,j} - E_z|_{i-1,j}}{2\Delta x} \right) \tag{3.51b}$$

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

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



(a) Unstaggered, collocated grid



(b) Staggered, collocated grid

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.

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

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

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

Reference [4] analyzes the Yee grid and the alternative Cartesian grids of Figs. 3.4(a, 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.

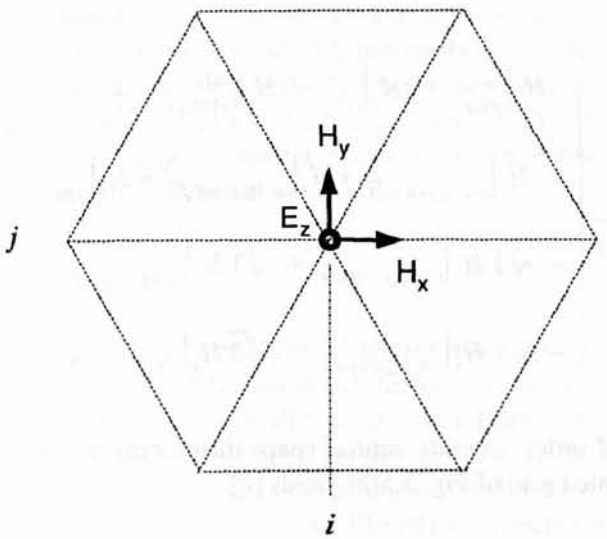
3.7.2 Hexagonal Grids

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.

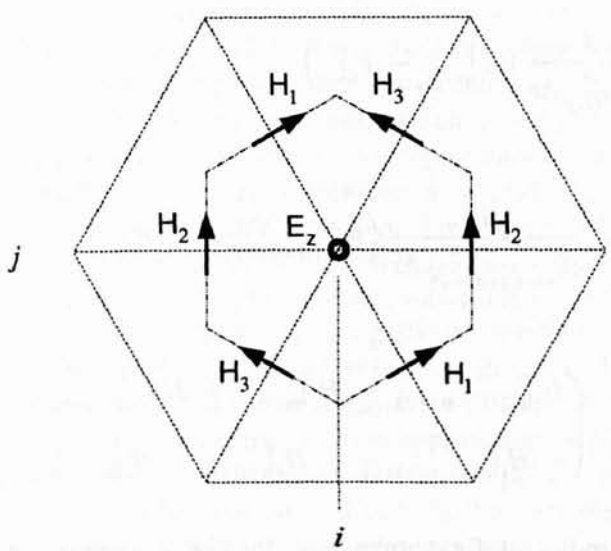
Fig. 3.5 illustrates for the TM_z case in two dimensions the two principal ways of arranging E and H components in hexagonal grids. Fig. 3.5(a) shows the unstaggered, collocated grid in which Cartesian E_z, H_x, and H_y components are collocated at the vertices of the equilateral triangles. No dual grid 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, and so forth, 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]:

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



(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.

$$\frac{\partial H_y\big|_{i,j}}{\partial t} = \frac{1}{\mu_{i,j}\,6\Delta s}\begin{pmatrix} 2E_z\big|_{i+1,j} - 2E_z\big|_{i-1,j} + E_z\big|_{i+1/2,\,j+0.5\sqrt{3}} \\ - E_z\big|_{i-1/2,\,j+0.5\sqrt{3}} + E_z\big|_{i+1/2,\,j-0.5\sqrt{3}} - E_z\big|_{i-1/2,\,j-0.5\sqrt{3}} \end{pmatrix} \tag{3.53b}$$

$$\frac{\partial E_z\big|_{i,j}}{\partial t} = \frac{1}{\varepsilon_{i,j}\,6\Delta s}\begin{pmatrix} 2H_y\big|_{i+1,j} - 2H_y\big|_{i-1,j} + H_y\big|_{i+1/2,\,j+0.5\sqrt{3}} \\ - H_y\big|_{i-1/2,\,j+0.5\sqrt{3}} + H_y\big|_{i+1/2,\,j-0.5\sqrt{3}} - H_y\big|_{i-1/2,\,j-0.5\sqrt{3}} \\ - \sqrt{3}\,H_x\big|_{i+1/2,\,j+0.5\sqrt{3}} + \sqrt{3}\,H_x\big|_{i+1/2,\,j-0.5\sqrt{3}} \\ - \sqrt{3}\,H_x\big|_{i-1/2,\,j+0.5\sqrt{3}} + \sqrt{3}\,H_x\big|_{i-1/2,\,j-0.5\sqrt{3}} \end{pmatrix} \tag{3.53c}$$

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]:

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

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

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

$$\frac{\partial E_z\big|_{i,j}}{\partial t} = \frac{2}{\varepsilon_{i,j}\,3\Delta s}\cdot\begin{pmatrix} H_1\big|_{i+1/4,\,j-0.25\sqrt{3}} + H_2\big|_{i+1/2,\,j} + H_3\big|_{i+1/4,\,j+0.25\sqrt{3}} \\ - H_1\big|_{i-1/4,\,j+0.25\sqrt{3}} - H_2\big|_{i-1/2,\,j} - H_3\big|_{i-1/4,\,j-0.25\sqrt{3}} \end{pmatrix} \tag{3.54d}$$

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 approximately 50%.

Analysis of the numerical dispersion of the hexagonal grids of Figs. 3.5(a, 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.

3.8 EMERGING APPLICATION: GRIDDING THE PLANET EARTH

This section discusses an emerging application of FDTD grid generation, which illustrates many of the concepts introduced earlier in this chapter: modeling of impulsive ELF propagation within the global Earth-ionosphere cavity. This permits for the first time a direct, three-dimensional, time-domain calculation of round-the-world ELF propagation, accounting for arbitrary horizontal as well as vertical geometrical and electrical inhomogeneities and anisotropies of the excitation, ionosphere, lithosphere, and oceans. Potential uses of this model include studies of electromagnetic precursors of major earthquakes and the development of novel means for remote sensing of underground ore and oil deposits.

3.8.1 Background

Propagation of ELF (3 Hz to 3 kHz) and VLF (3 to 30 kHz) electromagnetic waves in the Earth-ionosphere waveguide is a problem having a rich history of theoretical investigation extending over many years [5–10]. Currently, ELF/VLF propagation phenomena form the physics basis of important remote-sensing investigations of lightning and sprites [11], global temperature change [12], subsurface structures [13], and earthquake precursors [14].

Most theoretical techniques for modeling ELF/VLF propagation in the Earth-ionosphere waveguide are based upon frequency-domain waveguide mode theory [6, 9]. References [15–17] were the first to report the application of FDTD methods to this problem. These used two-dimensional cylindrical-coordinate FDTD grids to investigate VLF propagation over lossy ground paths due to either man-made sources or lightning discharges. Both reported very good agreement of their FDTD models with benchmark data.

Based upon the fundamental work in [18], [19–23] reported the initial application of FDTD to model the complete Earth-sphere at ELF. These used three-dimensional, spherical-coordinate, latitude-longitude space lattices with periodic boundary conditions. References [19, 20] reported no improvements relative to the grid of [18], which is subject to increasing space-cell eccentricity upon approaching the poles, due to converging lines of longitude. Since the Courant limit is set by the smallest space-cell dimension in a grid, this eccentricity mandates a corresponding reduction in the allowable time-step. For the grid of [18], the Courant limit is reduced to approximately d/r_E times the limit that would exist if all grid cells maintained the square configuration of those located at the Equator, where d is the cell dimension at the Equator and r_E is the radius of the Earth. This is a very significant mandatory time-step reduction factor for high-resolution grids having d smaller than 100 km (i.e., $d/r_E < 1/64$).

In contrast, [21–23] reported a key improvement in FDTD mesh generation to minimize the effects of increasing space-cell eccentricity upon approaching a pole. Here, the mesh was constructed to maintain approximately square cells in the polar regions by an adaptive cell-combination technique applied to adjacent grid cells in the East-West direction. This allows maintenance of the time-step at nearly the level permitted by the Courant stability condition for the square equatorial cells. Further, as shown in [21, 22], this approach achieves a high degree of isotropy for numerical wave propagation despite the mesh nonuniformity due to the adaptive East-West cell combinations. Overall, this technique permits a laboratory computer with 2 GB of random access memory to generate high-resolution (approximately $40 \times 40 \times 5$ km) modeling results for global, fully three-dimensional, impulsive ELF propagation within the entire Earth-ionosphere cavity, to a height of 100 km into the ionosphere and a depth of 100 km into the lithosphere. Section 3.8.2 provides the algorithmic details and sample results of this approach.

More recently, [24] reported an alternative means to grid the Earth that avoids the problem of geometrical singularities at the poles. Here, the FDTD grid is best described as resembling the surface of a soccer ball, being comprised entirely of hexagonal cells except for a small fixed number of pentagonal cells needed for grid completion. Grid-cell areas and locations are optimized to yield a smoothly varying area difference between adjacent cells, thereby maximizing numerical convergence. The new "geodesic" grid model is superior to the latitude-longitude grids of [21, 22] because it avoids geometrical singularities, executes approximately 14 times faster with only a 40% increase in the required memory, and provides even more isotropic numerical wave propagation. Section 3.8.3 describes this scheme and provides sample results.

3.8.2 The Latitude-Longitude Space Lattice

Fig. 3.6 illustrates the general layout of the latitude-longitude space lattice as seen along a TM surface at a constant radial coordinate. The lattice is a logically Cartesian $2M \times M \times K$ cell arrangement, where M is a power of 2. The lattice-cell position index in the West-to-East direction is $1 \le i \le 2M$; the lattice-cell position index in the South-to-North direction is $1 \le j \le M$; and the lattice-cell position index in the radial direction is $1 \le k \le K$.

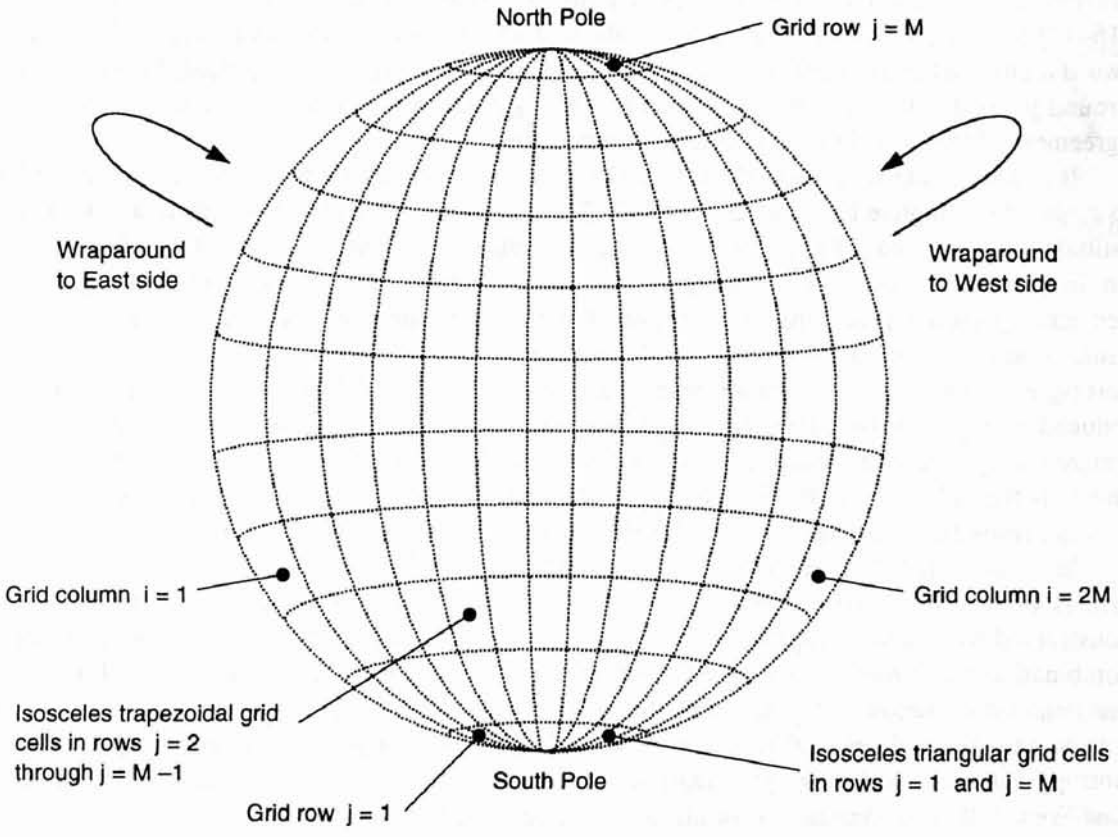
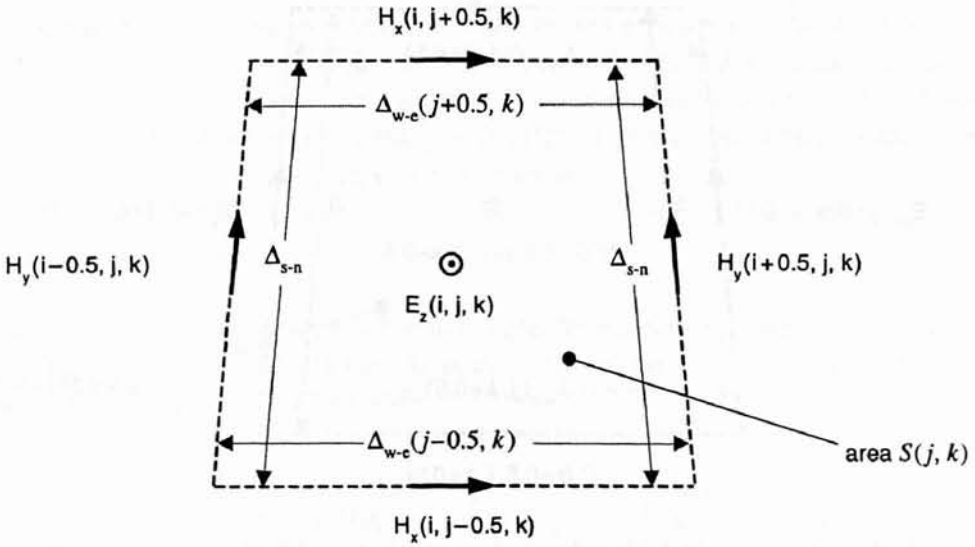
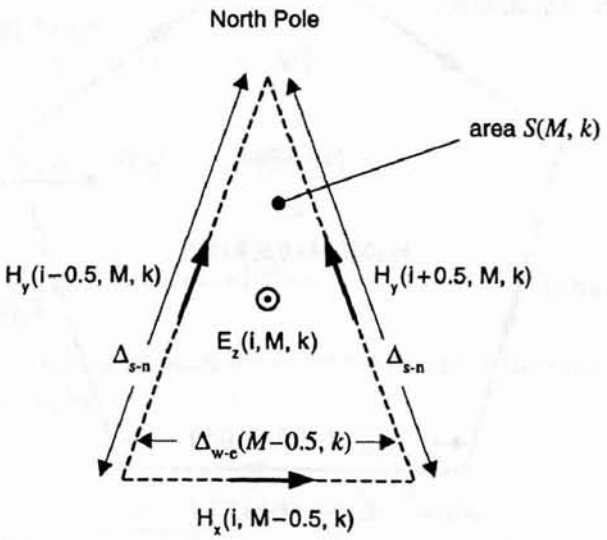


Fig. 3.6 General layout of the latitude-longitude FDTD space lattice covering the complete Earth-sphere as seen in a TM plane at a constant radial (height/depth) coordinate. *Source:* Simpson and Taflove, *IEEE Trans. Antennas and Propagation*, Vol. 52, 2004, pp. 443–451, © 2004 IEEE.

In Fig. 3.6, we see that the grid cells follow along lines of constant latitude $\theta = \text{constant}$, where θ is measured from the North Pole; and along lines of constant longitude $\phi = \text{constant}$, where ϕ is measured from a specified prime meridian. In this manner, each TM surface of the grid shown is comprised of isosceles trapezoidal cells away from the North and South Poles [Fig. 3.7(a)], and isosceles triangular cells at the poles [Fig. 3.7(b)].



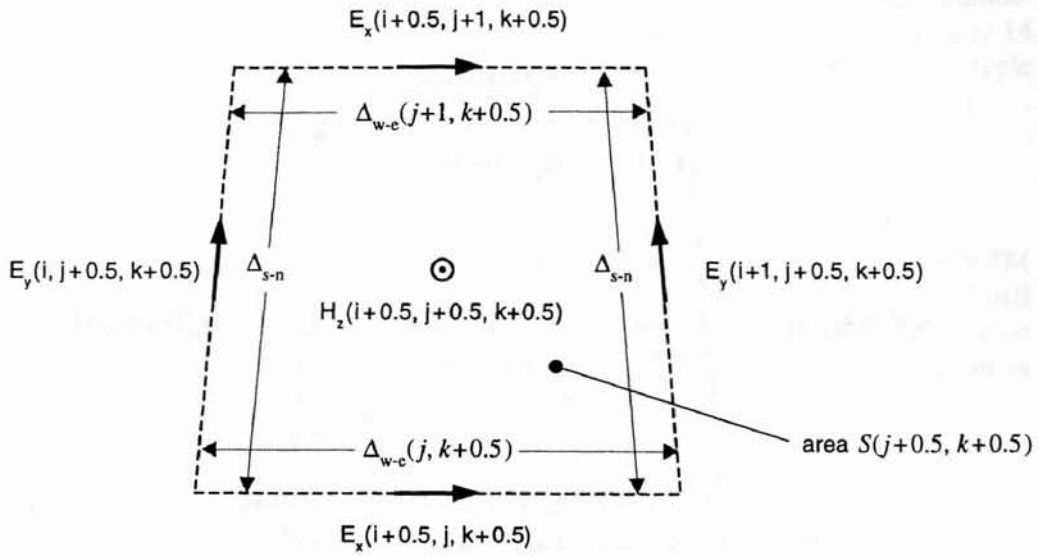
(a) Isosceles trapezoidal cell in the Northern Hemisphere away from the North Pole



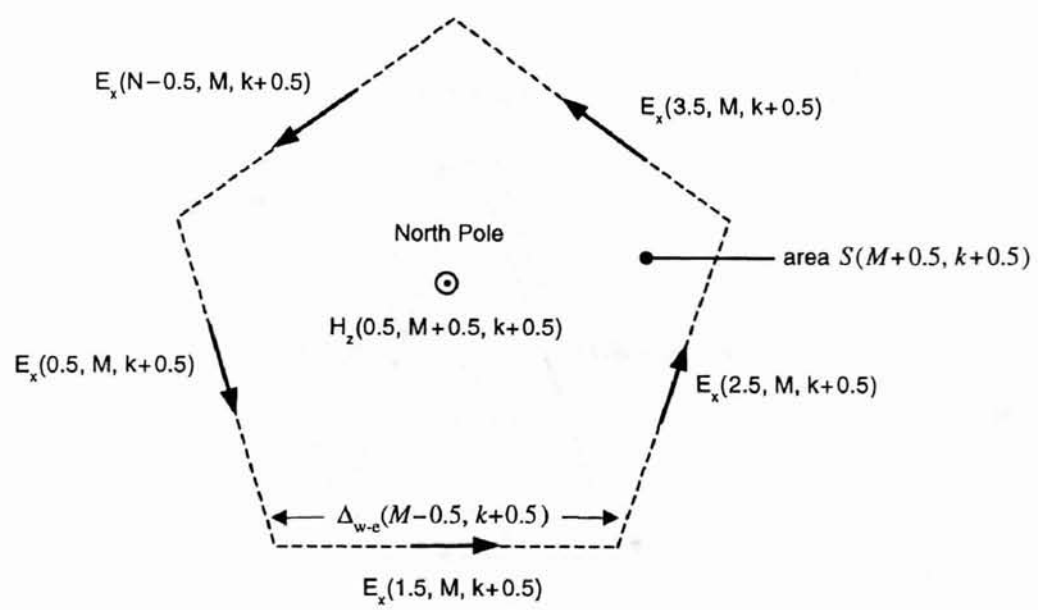
(b) Isosceles triangular cell at the North Pole

Fig. 3.7 Details of the TM-plane lattice-cell geometry: *Source:* Simpson and Taflove, *IEEE Trans. Antennas and Propagation*, Vol. 52, 2004, pp. 443–451, © 2004 IEEE.

Similarly, each TE surface at a constant radial coordinate is comprised of isosceles trapezoidal cells away from the North and South Poles [Fig. 3.8(a)], and a polygon at the poles [Fig. 3.8(b)].



(a) Isosceles trapezoidal cell in the Northern Hemisphere away from the North Pole



(b) Polygon cell directly at the North Pole

Fig. 3.8 Details of the TE-plane lattice-cell geometry: *Source:* Simpson and Taflove, *IEEE Trans. Antennas and Propagation*, Vol. 52, 2004, pp. 443–451, © 2004 IEEE.

We choose to have the same angular increment in latitude, $\Delta\theta = \pi/m$, for each cell in the grid. Thus, the South-North span of each trapezoidal or triangular grid cell is $\Delta_{s\text{-}n} = \pi R/m$, where R is the radial distance from the center of the Earth. To maintain square or nearly square grid cells near the Equator, we select the baseline value of the angular increment in longitude $\Delta\phi$ to equal $\Delta\theta$. However, this causes the West-East span of each cell $\Delta_{w\text{-}e} = R\Delta\phi \sin\theta$ to be a function of θ. This could be troublesome for cells near the North and South Poles where $\theta \to 0$ and $\theta \to \pi$, respectively. There, the geometrical eccentricity of each cell $\Delta_{s\text{-}n}/\Delta_{w\text{-}e} = \Delta\theta/(\Delta\phi \sin\theta)$ would become quite large, and the numerical stability and efficiency of the FDTD algorithm would be degraded. An algorithmic means to deal with this problem will be discussed below.

The wraparound or joining of the lattice is along a specific line of constant longitude, or meridian. As discussed below, this joining is, in effect, a periodic boundary condition applied at each j-row of lattice cells, whether trapezoids or triangles.

FDTD Algorithm, TM Components

Given the above assumptions, Ampere's law in integral form can be applied to develop an FDTD time-stepping relation for the electric field E_z at the center of the (i, j, k)'th trapezoidal grid cell. For example, referring to Fig. 3.7(a), we have

$$E_z\Big|_{i,j,k}^{n+1} = E_z\Big|_{i,j,k}^{n} + \frac{\Delta t}{\varepsilon_0 S|_{j,k}} \left[\begin{array}{l} H_x\Big|_{i,j-1/2,k}^{n+1/2} \Delta_{w-e}\Big|_{j-1/2,k} - H_x\Big|_{i,j+1/2,k}^{n+1/2} \Delta_{w-e}\Big|_{j+1/2,k} \\ + \left(H_y\Big|_{i+1/2,j,k}^{n+1/2} - H_y\Big|_{i-1/2,j,k}^{n+1/2} \right) \Delta_{s-n} \end{array} \right] \tag{3.55}$$

where Δt is the time-step, and

$$\Delta_{w-e}\Big|_{j+1/2,k} = R\Delta\phi \sin[(M-j)\pi/M] \tag{3.56a}$$

$$\Delta_{w-e}\Big|_{j-1/2,k} = R\Delta\phi \sin[(M-j+1)\pi/M] \tag{3.56b}$$

$$S|_{j,k} = \left[\Delta_{w-e}\Big|_{j-1/2,k} + \Delta_{w-e}\Big|_{j+1/2,k} \right] \Delta_{s-n}/2 \tag{3.56c}$$

Similarly, referring to Fig. 3.7(b), the update for E_z at the center of the i'th triangular grid cell at the North Pole ($j = M$) is given by

$$E_z\Big|_{i,M,k}^{n+1} = E_z\Big|_{i,M,k}^{n} + \frac{\Delta t}{\varepsilon_0 S|_{M,k}} \left[\begin{array}{l} H_x\Big|_{i,M-1/2,k}^{n+1/2} \Delta_{w-e}\Big|_{M-1/2,k} \\ + \left(H_y\Big|_{i+1/2,M,k}^{n+1/2} - H_y\Big|_{i-1/2,M,k}^{n+1/2} \right) \Delta_{s-n} \end{array} \right] \tag{3.57}$$

where $\Delta_{w\text{-}e}|_{M-1/2,k}$ is given by (3.56b) for the case $j = M$, and

$$S|_{M,k} = \frac{\Delta_{w-e}|_{M-1/2,k}\ \Delta_{s-n}}{2} \sin\left[\cos^{-1}\left(\frac{\Delta_{w-e}|_{M-1/2,k}}{2\Delta_{s-n}}\right)\right] \tag{3.58}$$

Analogous expressions can be derived for triangular grid cells at the South Pole ($j = 1$).

The basic FDTD time-stepping algorithm is completed by specifying the updates for the H_x and H_y fields. For example, referring to the trapezoidal grid cell shown in Fig. 3.7(a), we have

$$\begin{aligned} H_x|_{i,j-1/2,k}^{n+1.5} &= H_x|_{i,j-1/2,k}^{n+1/2} + \frac{\Delta t}{\mu_0 \Delta z}\left(E_y|_{i,j-1/2,k+1/2}^{n+1} - E_y|_{i,j-1/2,k-1/2}^{n+1}\right) \\ &+ \frac{\Delta t}{\mu_0 \Delta_{s-n}}\left(E_z|_{i,j-1,k}^{n+1} - E_z|_{i,j,k}^{n+1}\right) \end{aligned} \tag{3.59}$$

$$\begin{aligned} H_y|_{i+1/2,j,k}^{n+1.5} &= H_y|_{i+1/2,j,k}^{n+1/2} + \frac{\Delta t}{\mu_0 \Delta z}\left(E_x|_{i+1/2,j,k-1/2}^{n+1} - E_x|_{i+1/2,j,k+1/2}^{n+1}\right) \\ &+ \frac{\Delta t}{\mu_0 \Delta_{w-e}|_{j,k}}\left(E_z|_{i+1,j,k}^{n+1} - E_z|_{i,j,k}^{n+1}\right) \end{aligned} \tag{3.60}$$

where Δz is the cell span in the vertical (radial) direction.

For a triangular grid cell at the North Pole ($j = M$) shown in Fig. 3.7(b), we similarly have

$$\begin{aligned} H_x|_{i,M-1/2,k}^{n+1.5} &= H_x|_{i,M-1/2,k}^{n+1/2} + \frac{\Delta t}{\mu_0 \Delta z}\left(E_y|_{i,M-1/2,k+1/2}^{n+1} - E_y|_{i,M-1/2,k-1/2}^{n+1}\right) \\ &+ \frac{\Delta t}{\mu_0 \Delta_{s-n}}\left(E_z|_{i,M-1,k}^{n+1} - E_z|_{i,M,k}^{n+1}\right) \end{aligned} \tag{3.61}$$

$$\begin{aligned} H_y|_{i+1/2,M,k}^{n+1.5} &= H_y|_{i+1/2,M,k}^{n+1/2} + \frac{\Delta t}{\mu_0 \Delta z}\left(E_x|_{i+1/2,M,k-1/2}^{n+1} - E_x|_{i+1/2,M,k+1/2}^{n+1}\right) \\ &+ \frac{\Delta t}{\mu_0 \Delta_{w-e}|_{M,k}}\left(E_z|_{i+1,M,k}^{n+1} - E_z|_{i,M,k}^{n+1}\right) \end{aligned} \tag{3.62}$$

Analogous expressions can be derived for triangular grid cells at the South Pole ($j = 1$).

As stated earlier, the geometrical eccentricity of each trapezoidal cell can become quite large near the North and South Poles, thereby degrading the numerical stability and efficiency of the FDTD algorithm. Fig. 3.9 illustrates a means to mitigate this problem by merging pairs of adjacent cells of the TM plane in the West-East direction, effectively halving the cell eccentricity. This process can be repeated several times as the grid approaches a pole, allowing the user to specify a maximum allowable cell eccentricity.

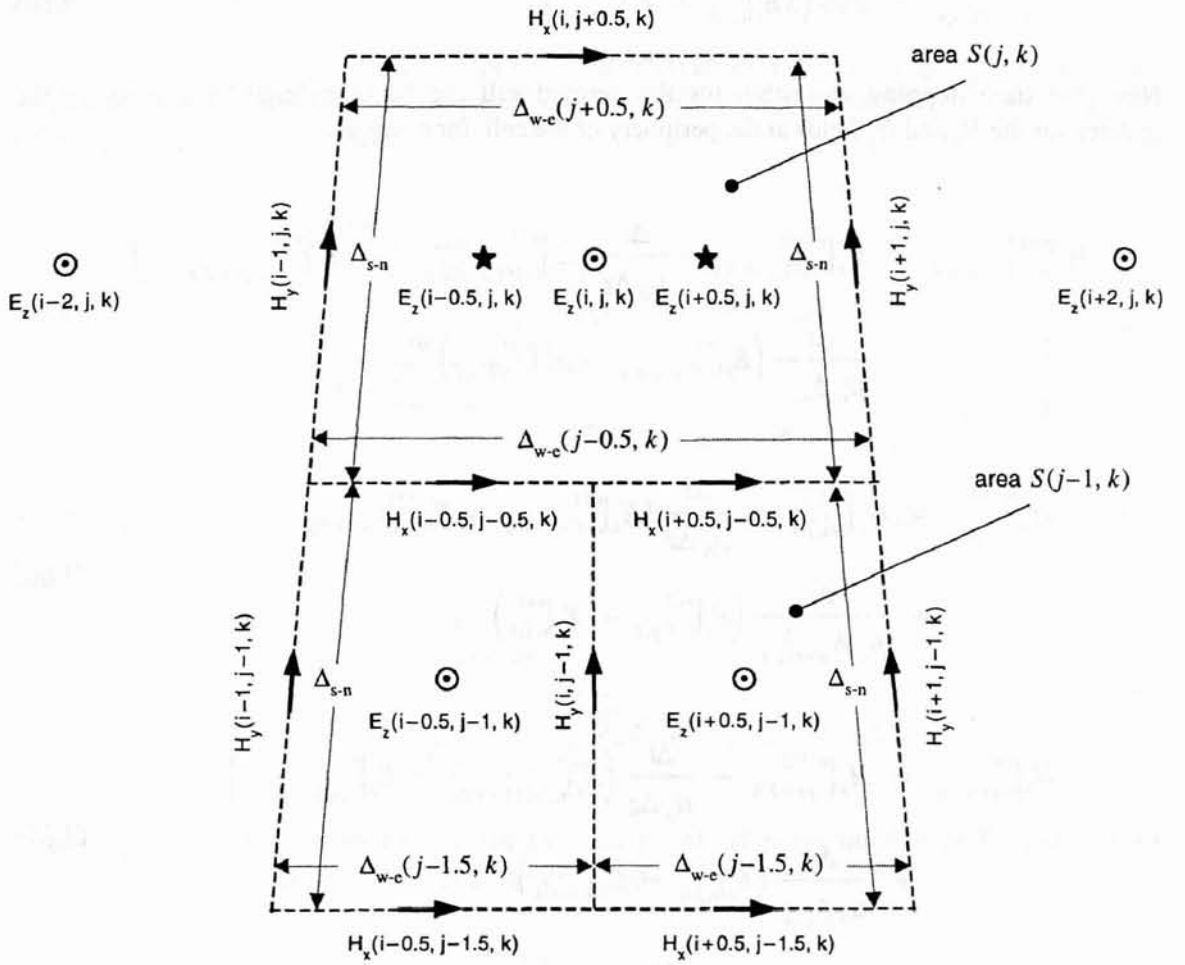


Fig. 3.9 Details of the TM-plane grid-cell geometry in the Northern Hemisphere at the transition between two adjacent regular cells and a single cell spanning twice the distance in the East-West direction. *Source:* Simpson and Taflove, *IEEE Trans. Antennas and Propagation*, Vol. 52, 2004, pp. 443–451, © 2004 IEEE.

The required algorithm is now presented. Again applying Ampere's law in integral form, the time-stepping relation for the E_z field at the center of the merged cell is given by

$$E_z\Big|_{i,j,k}^{n+1} = E_z\Big|_{i,j,k}^{n} + \frac{\Delta t}{\varepsilon_0 \, S\big|_{j,k}} \left[\begin{array}{l} 0.5 \left(H_x\Big|_{i-1/2,j-1/2,k}^{n+1/2} + H_x\Big|_{i+1/2,j-1/2,k}^{n+1/2} \right) \Delta_{\text{w-c}}\big|_{j-1/2,k} - \\ H_x\Big|_{i,j+1/2,k}^{n+1/2} \, \Delta_{\text{w-c}}\big|_{j+1/2,k} + \left(H_y\Big|_{i+1,j,k}^{n+1/2} - H_y\Big|_{i-1,j,k}^{n+1/2} \right) \Delta_{\text{s-n}} \end{array} \right] \tag{3.63}$$

The E_z fields indicated by the ★ symbols in Fig. 3.9 are required for the subsequent H-field updates. These are obtained by linearly interpolating the E_z fields calculated with (3.63), for example:

$$E_z\Big|_{i-1/2,j,k}^{n+1} = 0.25\left(3E_z\Big|_{i,j,k}^{n+1} + E_z\Big|_{i-2,j,k}^{n+1}\right) \tag{3.64}$$

Now, the time-stepping algorithm for the merged cell can be completed by specifying the updates for the H_x and H_y fields at the periphery of the cell, for example:

$$\begin{aligned} H_x\Big|_{i-1/2,j-1/2,k}^{n+1.5} &= H_x\Big|_{i-1/2,j-1/2,k}^{n+1/2} + \frac{\Delta t}{\mu_0\,\Delta z}\left(E_y\Big|_{i-1/2,j-1/2,k+1/2}^{n+1} - E_y\Big|_{i-1/2,j-1/2,k-1/2}^{n+1}\right) \\ &+ \frac{\Delta t}{\mu_0\,\Delta_{s-n}}\left(E_z\Big|_{i-1/2,j-1,k}^{n+1} - E_z\Big|_{i-1/2,j,k}^{n+1}\right) \end{aligned} \tag{3.65}$$

$$\begin{aligned} H_y\Big|_{i+1,j,k}^{n+1.5} &= H_y\Big|_{i+1,j,k}^{n+1/2} + \frac{\Delta t}{\mu_0\,\Delta z}\left(E_x\Big|_{i+1,j,k-1/2}^{n+1} - E_x\Big|_{i+1,j,k+1/2}^{n+1}\right) \\ &+ \frac{\Delta t}{\mu_0\,\Delta_{w-e}\big|_{j,k}}\left(E_z\Big|_{i+2,j,k}^{n+1} - E_z\Big|_{i,j,k}^{n+1}\right) \end{aligned} \tag{3.66}$$

$$\begin{aligned} H_x\Big|_{i,j+1/2,k}^{n+1.5} &= H_x\Big|_{i,j+1/2,k}^{n+1/2} + \frac{\Delta t}{\mu_0\,\Delta z}\left(E_y\Big|_{i,j+1/2,k+1/2}^{n+1} - E_y\Big|_{i,j+1/2,k-1/2}^{n+1}\right) \\ &+ \frac{\Delta t}{\mu_0\,\Delta_{s-n}}\left(E_z\Big|_{i,j,k}^{n+1} - E_z\Big|_{i,j+1,k}^{n+1}\right) \end{aligned} \tag{3.67}$$

The grid wraparound (periodic boundary condition) in the West-East direction completes the FDTD algorithm for the TM components. For each row j of grid column $i = 1$ in Fig. 3.6, this is implemented by the following special time-stepping relation for H_y at the West grid boundary:

$$\begin{aligned} H_y\Big|_{1/2,j,k}^{n+1.5} &= H_y\Big|_{1/2,j,k}^{n+1/2} + \frac{\Delta t}{\mu_0\,\Delta z}\left(E_x\Big|_{1/2,j,k-1/2}^{n+1} - E_x\Big|_{1/2,j,k+1/2}^{n+1}\right) \\ &+ \frac{\Delta t}{\mu_0\,\Delta_{w-e}\big|_{j,k}}\left(E_z\Big|_{1,j,k}^{n+1} - E_z\Big|_{2M,j,k}^{n+1}\right) \end{aligned} \tag{3.68a}$$

Upon obtaining each H_y value, the grid wraparound is completed by implementing the following special relation for each row j of grid column $i = 2M$ for H_y at the East grid boundary:

$$H_y\Big|_{2M+1/2,j,k}^{n+1.5} = H_y\Big|_{1/2,j,k}^{n+1.5} \tag{3.68b}$$

Steps (3.68a, b) apply to all grid rows, whether containing trapezoidal or triangular cells.

FDTD Algorithm, TE Components

Faraday's law in integral form can be applied to develop an FDTD time-stepping relation for the magnetic field H_z at the center of the $(i+1/2, j+1/2, k+1/2)$'th trapezoidal grid cell. For example, referring to Fig. 3.8(a), we have

$$H_z\Big|_{i+1/2,j+1/2,k+1/2}^{n+1.5} = H_z\Big|_{i+1/2,j+1/2,k+1/2}^{n+1/2} + \frac{\Delta t}{\mu_0 S\big|_{j+1/2,k+1/2}} \left[\begin{array}{l} E_x\Big|_{i+1/2,j+1,k+1/2}^{n} \Delta_{\text{w-c}}\big|_{j+1,k+1/2} \\ - E_x\Big|_{i+1/2,j,k+1/2}^{n} \Delta_{\text{w-c}}\big|_{j,k+1/2} \\ + \left(E_y\Big|_{i,j+1/2,k+1/2}^{n} - E_y\Big|_{i+1,j+1/2,k+1/2}^{n} \right) \Delta_{\text{s-n}} \end{array} \right] \tag{3.69}$$

Similarly, referring to Fig. 3.8(b), the update for H_z directly at the North Pole ($i = 1/2, j = M+1/2$) is given by

$$H_z\Big|_{1/2,M+1/2,k+1/2}^{n+1.5} = H_z\Big|_{1/2,M+1/2,k+1/2}^{n+1/2} - \frac{\Delta t}{\mu_0 S\big|_{M+1/2,k+1/2}} \left(\sum_{X=1}^{N} E_x\Big|_{X-1/2,M,k+1/2}^{n} \Delta_{\text{w-c}}\big|_{M,k+1/2} \right) \tag{3.70}$$

where $\Delta_{\text{w-c}}\big|_{M,k+1/2}$ is given by (3.56b) for the case $j = M$, N is the number of E_x components surrounding the H_z component directly at the pole, and

$$S\big|_{M+1/2,k+1/2} = N \left(\Delta_{\text{w-c}}\big|_{M,k+1/2} \right)^2 \Big/ 4 \tan(\pi/N) \tag{3.71}$$

Analogous expressions can be derived for H_z directly at the South Pole ($i = 1/2, j = 1/2$).

The basic FDTD time-stepping algorithm is completed by specifying the updates for the E_x and E_y fields. For example, referring to the trapezoidal grid cell shown in Fig. 3.8(a), we have

$$\begin{aligned} E_x\Big|_{i+1/2,j,k+1/2}^{n+1} &= E_x\Big|_{i+1/2,j,k+1/2}^{n} + \frac{\Delta t}{\varepsilon_0 \Delta z} \left(H_y\Big|_{i+1/2,j,k}^{n+1/2} - H_y\Big|_{i+1/2,j,k+1}^{n+1/2} \right) \\ &+ \frac{\Delta t}{\varepsilon_0 \Delta_{\text{s-n}}} \left(H_z\Big|_{i+1/2,j+1/2,k+1/2}^{n+1/2} - H_z\Big|_{i+1/2,j-1/2,k+1/2}^{n+1/2} \right) \end{aligned} \tag{3.72}$$

$$\begin{aligned} E_y\Big|_{i,j+1/2,k+1/2}^{n+1} &= E_y\Big|_{i,j+1/2,k+1/2}^{n} + \frac{\Delta t}{\varepsilon_0 \Delta z} \left(H_x\Big|_{i,j+1/2,k+1}^{n+1/2} - H_x\Big|_{i,j+1/2,k}^{n+1/2} \right) \\ &+ \frac{\Delta t}{\varepsilon_0 \Delta_{\text{w-c}}\big|_{j+1/2,k+1/2}} \left(H_z\Big|_{i-1/2,j+1/2,k+1/2}^{n+1/2} - H_z\Big|_{i+1/2,j+1/2,k+1/2}^{n+1/2} \right) \end{aligned} \tag{3.73}$$

As stated earlier, near the North and South Poles the geometrical eccentricity of each trapezoidal cell can become quite large, thereby degrading the numerical stability and efficiency of the FDTD algorithm. Fig. 3.10 illustrates a means to mitigate this problem by merging pairs of adjacent cells in the TE plane in the West-East direction, effectively halving the cell eccentricity. This process can be repeated several times as the grid approaches a pole, allowing the user to specify a maximum allowable cell eccentricity.

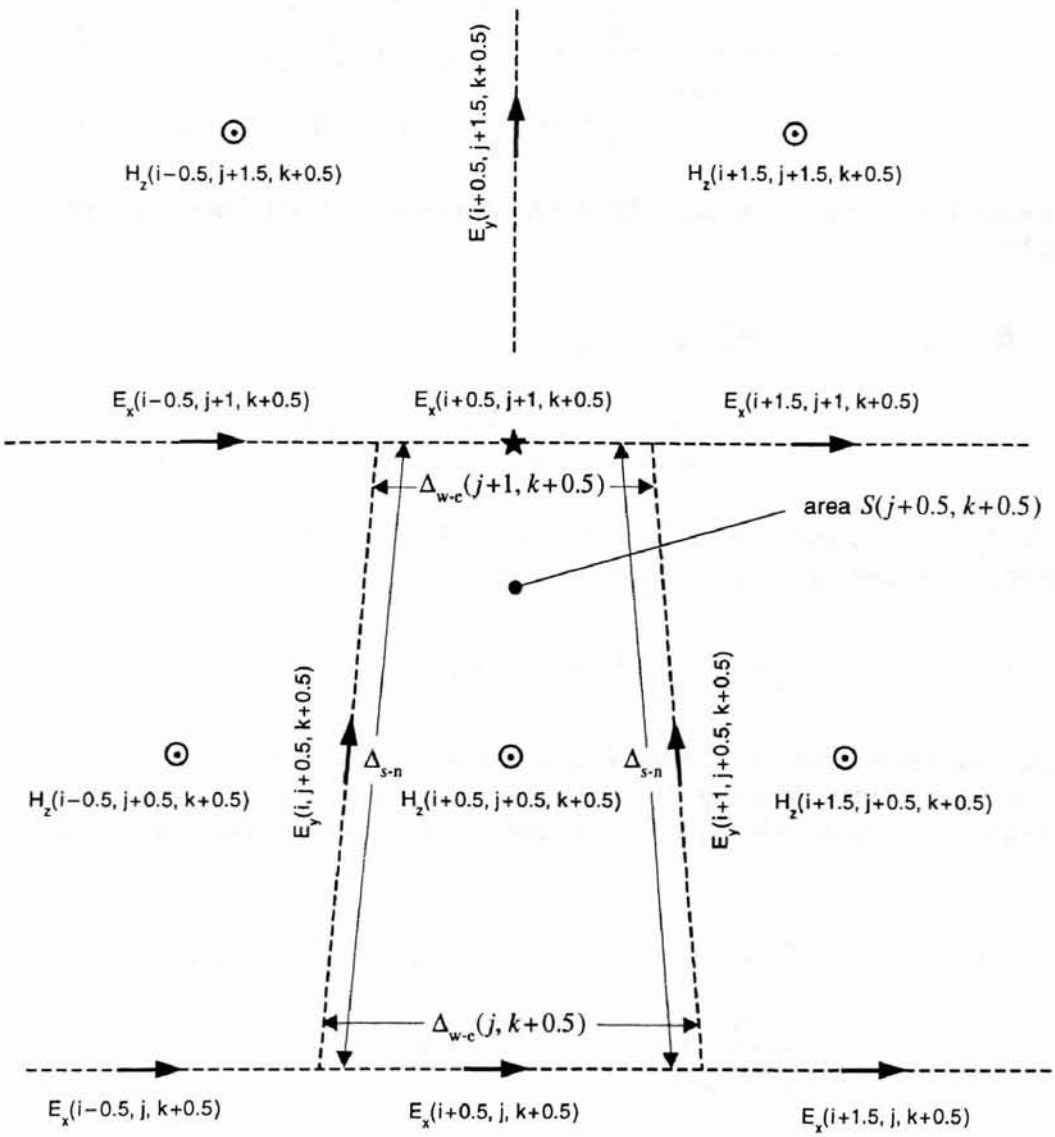


Fig. 3.10 Details of the TE-plane grid-cell geometry in the Northern Hemisphere at the transition between three adjacent regular cells and two cells each spanning twice the distance in the East-West direction. *Source:* Simpson and Taflove, *IEEE Trans. Antennas and Propagation*, Vol. 52, 2004, pp. 443–451, © 2004 IEEE.

The required algorithm is now presented. First, we obtain the E_x field indicated by the ★ symbol in Fig. 3.10. This field component is obtained by linear interpolation:

$$E_x\Big|^n_{i+1/2,\,j+1,\,k+1/2} = 0.5\left(E_x\Big|^n_{i-1/2,\,j+1,\,k+1/2} + E_x\Big|^n_{i+1.5,\,j+1,\,k+1/2}\right) \tag{3.74}$$

Next, applying Faraday's law in integral form, we use the E_x component obtained in (3.74) in the time-stepping relation for an H_z component located below a row of merged cells:

$$\begin{aligned} H_z\Big|^{n+1/2}_{i+1/2,\,j+1/2,\,k+1/2} &= H_z\Big|^{n-1/2}_{i+1/2,\,j+1/2,\,k+1/2} \\ &+ \frac{\Delta t}{\mu_0\, S\big|_{j+1/2,\,k+1/2}} \left[\begin{array}{l} E_x\Big|^n_{i+1/2,\,j+1,\,k+1/2}\ \Delta_{\text{w-e}}\big|_{j+1,\,k+1/2} - \\ E_x\Big|^n_{i+1/2,\,j,\,k+1/2}\ \Delta_{\text{w-e}}\big|_{j,\,k+1/2} + \\ \left(E_y\Big|^n_{i,\,j+1/2,\,k+1/2} - E_y\Big|^n_{i+1,\,j+1/2,\,k+1/2}\right)\Delta_{\text{s-n}} \end{array}\right] \end{aligned} \tag{3.75}$$

Again applying Faraday's law in integral form, the time-stepping relation for an H_z component located in a row of merged cells is given by

$$\begin{aligned} H_z\Big|^{n+1/2}_{i+1.5,\,j+1.5,\,k+1/2} &= H_z\Big|^{n-1/2}_{i+1.5,\,j+1.5,\,k+1/2} \\ &+ \frac{\Delta t}{\mu_0\, S\big|_{j+1.5,\,k+1/2}} \left[\begin{array}{l} E_x\Big|^n_{i+1.5,\,j+2,\,k+1/2}\ \Delta_{\text{w-e}}\big|_{j+2,\,k+1/2} - \\ E_x\Big|^n_{i+1.5,\,j+1,\,k+1/2}\ \Delta_{\text{w-e}}\big|_{j+1,\,k+1/2} + \\ \left(E_y\Big|^n_{i+1/2,\,j+1.5,\,k+1/2} - E_y\Big|^n_{i+2.5,\,j+1.5,\,k+1/2}\right)\Delta_{\text{s-n}} \end{array}\right] \end{aligned} \tag{3.76}$$

Now, the time-stepping algorithm for a merged cell can be completed by specifying the updates for the E_x and E_y fields at the periphery of the cell, for example:

$$\begin{aligned} E_x\Big|^{n+1}_{i+1.5,\,j+1,\,k+1/2} &= E_x\Big|^n_{i+1.5,\,j+1,\,k+1/2} + \frac{\Delta t}{\varepsilon_0\,\Delta z}\left(H_y\Big|^{n+1/2}_{i+1.5,\,j+1,\,k} - H_y\Big|^{n+1/2}_{i+1.5,\,j+1,\,k+1}\right) \\ &+ \frac{\Delta t}{\varepsilon_0\,\Delta_{\text{s-n}}}\left(H_z\Big|^{n+1/2}_{i+1.5,\,j+1.5,\,k+1/2} - H_z\Big|^{n+1/2}_{i+1.5,\,j+1/2,\,k+1/2}\right) \end{aligned} \tag{3.77}$$

$$\begin{aligned} E_y\Big|^{n+1}_{i+1/2,\,j+1.5,\,k+1/2} &= E_y\Big|^n_{i+1/2,\,j+1.5,\,k+1/2} + \frac{\Delta t}{\varepsilon_0\,\Delta z}\left(H_x\Big|^{n+1/2}_{i+1/2,\,j+1.5,\,k+1} - H_x\Big|^{n+1/2}_{i+1/2,\,j+1.5,\,k}\right) \\ &+ \frac{\Delta t}{\varepsilon_0\,\Delta_{\text{w-e}}\big|_{j+1.5,\,k+1/2}}\left(H_z\Big|^{n+1/2}_{i-1/2,\,j+1.5,\,k+1/2} - H_z\Big|^{n+1/2}_{i+1.5,\,j+1.5,\,k+1/2}\right) \end{aligned} \tag{3.78}$$

The grid wraparound (periodic boundary condition) in the West-East direction completes the FDTD algorithm for the TE components. For each row j of grid column $i = 1/2$, this is implemented by the following special time-stepping relation for H_z at the West grid boundary:

$$H_z\Big|^{n+1/2}_{1/2,\,j+1/2,\,k+1/2} = H_z\Big|^{n-1/2}_{1/2,\,j+1/2,\,k+1/2} + \frac{\Delta t}{\mu_0\, S\big|_{j+1/2,\,k+1/2}} \left[\begin{array}{l} E_x\Big|^{n}_{1/2,\,j+1,\,k+1/2}\, \Delta_{w-e}\big|_{j+1,\,k+1/2} - \\ E_x\Big|^{n}_{1/2,\,j,\,k+1/2}\, \Delta_{w-e}\big|_{j,\,k+1/2} + \\ \left(E_y\Big|^{n}_{2M,\,j+1/2,\,k+1/2} - E_y\Big|^{n}_{1,\,j+1/2,\,k+1/2} \right) \Delta_{s-n} \end{array} \right] \tag{3.79a}$$

Upon obtaining each H_z value, the grid wraparound is completed by implementing the following special relation for each row j of grid column $i = 2M + 1/2$ at the East grid boundary:

$$H_z\Big|^{n+1/2}_{2M+1/2,\,j+1/2,\,k+1/2} = H_z\Big|^{n+1/2}_{1/2,\,j+1/2,\,k+1/2} \tag{3.79b}$$

Validation Study

Reference [23] reported a validation of the above algorithm by comparing its predictions for round-the-world ELF propagation with data reported in [25]. Topographic and bathymetric data from the NOAA-NGDC "Global Relief CD-ROM" [26] were mapped onto the space lattice with an assumed resolution of $40 \times 40 \times 5$ km at the Equator. This cell size was a compromise which: (1) limited the space-lattice size to permit its residence in memory on an available workstation with 2 GB of memory, (2) resolved ELF wavelengths of 800 km or greater (frequencies of 375 Hz or less) with 20 or more cells per wavelength, and (3) provided one-to-two cell resolution for the radial extent (height and depth) of the Earth's oceans and mountain ranges.

For the lithosphere, conductivity values were assigned according to [27], depending upon the location of the $\boldsymbol{E}$ component (i.e., below an ocean or within a continent). For the atmosphere, the exponential conductivity profile used in [28] was assumed. This permitted the most straightforward comparison of the FDTD modeling results with the data reported in [25] since ELF propagation is crucially affected by the ionosphere characteristics. Note that, however, the FDTD model is capable of significantly greater ionospheric detail (i.e., day-to-night transitions, anisotropy, and so forth).

The model was excited with a vertical, 5-km-long, Gaussian current pulse located just above the Earth's surface on the Equator at longitude 47°W. The pulse had a $1/e$ full-width of $480\Delta t$ (where $\Delta t = 3.0$ μs) and its temporal center at $960\Delta t$.

Fig. 3.11 provides snapshot visualizations of the global propagation of the electromagnetic pulse resulting from the assumed excitation. (The corresponding color snapshot visualizations are shown in Fig. 1.1 of Chapter 1; a color video of this phenomenon can be downloaded from http://www.ece.northwestern.edu/ecefaculty/taflove/3Dmovietext@gif.avi.) Slight departures from circular symmetry of the propagating pulse are observed. Analysis reveals that this lack of symmetry is caused to a high degree by the inhomogeneities of the lithosphere conductivity assumed in the model, rather than by numerical distortions resulting from the gridding technique.

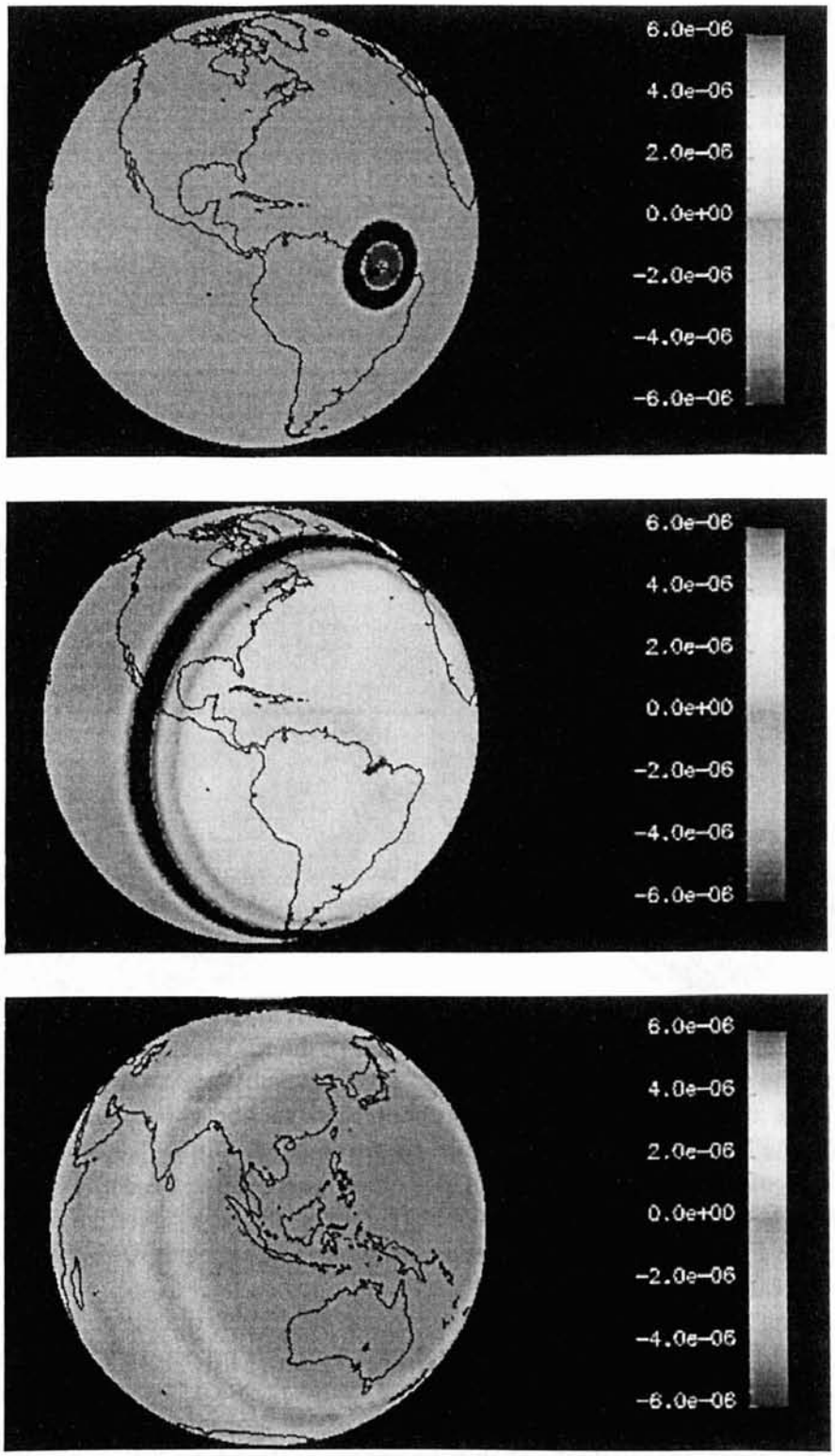


Fig. 3.11 Snapshot visualizations of the FDTD-computed global propagation of an ELF electromagnetic pulse generated by a vertical 5-km-long Gaussian current pulse off the coast of South America at the Equator. All features of the lithosphere and atmosphere located within ±100 km of sea level are modeled in three dimensions with a resolution of approximately $40 \times 40 \times 5$ km.

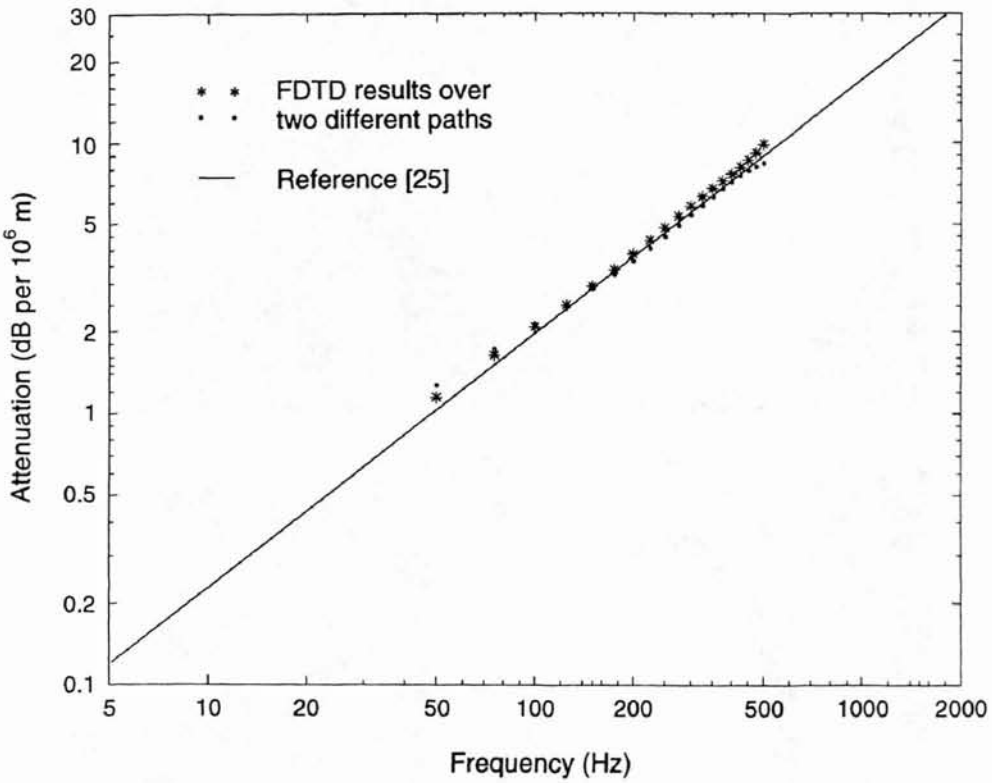


Fig. 3.12 Comparison between FDTD-calculated ELF propagation attenuation versus frequency with data reported in [25]. *Source:* Simpson and Taflove, *IEEE Trans. Antennas and Propagation*, Vol. 52, 2004, pp. 443–451, © 2004 IEEE.

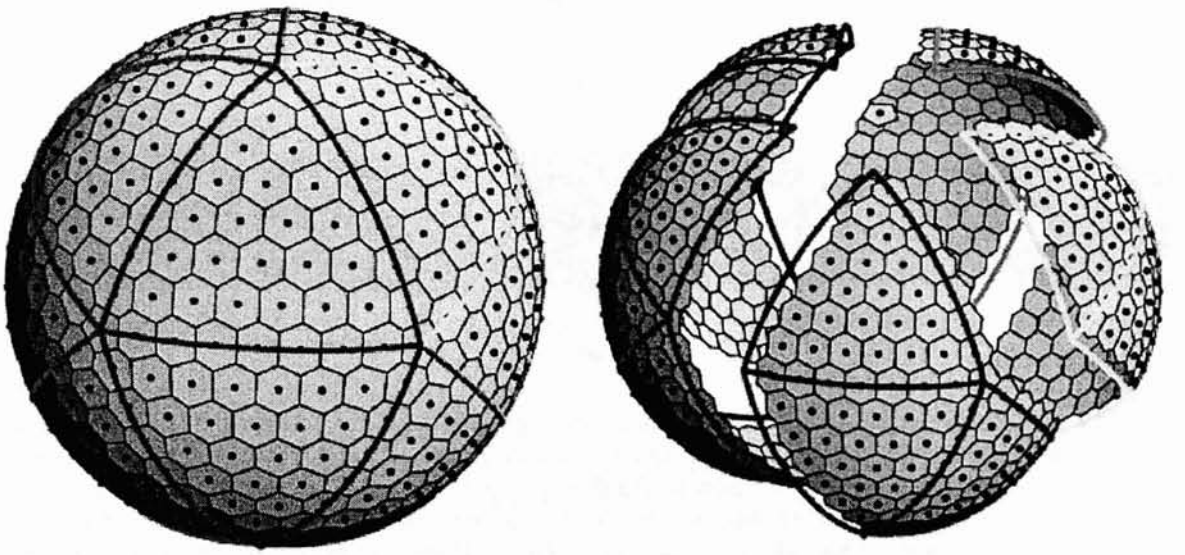
Fig. 3.12 compares, relative to benchmark data [25], the FDTD-calculated ELF propagation attenuation in the example visualized in Fig. 3.11. Results are shown for two separate propagation paths along the Equator in decibels per megameter versus frequency. Here, for each path, broadband FDTD results were obtained by forming the ratio of the *discrete Fourier transform* (DFT) of the vertical E-field time-waveform observed at the end of the path to the DFT of the E-field waveform observed at the beginning of the path. Over the frequency range 50 to 500 Hz, the FDTD-computed propagation attenuation values agree with the results of [25] to within about ±1.0 dB per megameter.

In developing Fig. 3.12, temporal-windowing of the FDTD-calculated field-versus-time waveforms was required to eliminate the "reception" of signals propagating completely around the model Earth in the opposite direction along the Equator. Such "long-path" signals would eventually add to, and corrupt, each desired time-waveform propagated along the "short path" directly from the source, especially the slow-tail response analyzed in [29]. In combination with the high-frequency wavelength-resolution limitations implied by the usage of a 40-km grid cell, the time-windowing process resulted in a frequency window of approximately 50 to 500 Hz within which the FDTD results could be validly compared with the benchmark data of [25]. Nevertheless, the FDTD results should remain valid well below 50 Hz.

3.8.3 The Geodesic (Hexagon-Pentagon) Grid

The geodesic Earth grid is comprised entirely of hexagonal cells except for a small fixed number (12) of pentagonal cells needed for grid completion [30]. Grid-cell areas and locations are optimized to yield a smoothly varying area difference between adjacent cells, thereby maximizing numerical convergence [31].

Fig. 3.13 illustrates an example of the layout of a spherical geodesic grid along a TM surface at a constant radial coordinate. This sample grid has a total of 642 planar cells, of which 630 are hexagons and 12 are pentagons [30]. For purposes of efficient mapping into the computer memory, this grid can be divided into five equal panels of size $im5 \times jm$ cells, where $im5 = 18$ and $jm = 10$.



(a) (b)

Fig. 3.13 642-cell, two-dimensional geodesic grid covering the Earth-sphere: (a) layout of the entire grid, showing the arrangement of the planar grid cells; (b) division of the grid into five equal panels with $im5 = 18$ and $jm = 10$. *Source:* Randall et al., *Computing in Science and Engineering*, Vol. 4(5), 2002, pp. 32–41, © 2002 IEEE.

Fig. 3.14(a) shows the five grid panels of Fig. 3.13 after unwrapping and stretching them flat. The cells at the North and South Poles can be arbitrarily connected to any of the five panels. Fig. 3.14(b) illustrates how the five grid panels of Fig. 3.13 can be assigned logically Cartesian coordinates, and how all five panels can then be laid side-by-side to constitute an overall ($im = 90$, $jm = 10$) logically Cartesian grid. This powerful interpretation of a spherical geodesic grid as a single logically Cartesian grid for purposes of efficient computer processing can be implemented for a variety of meshing densities over the sphere, as shown in Table 3.1.

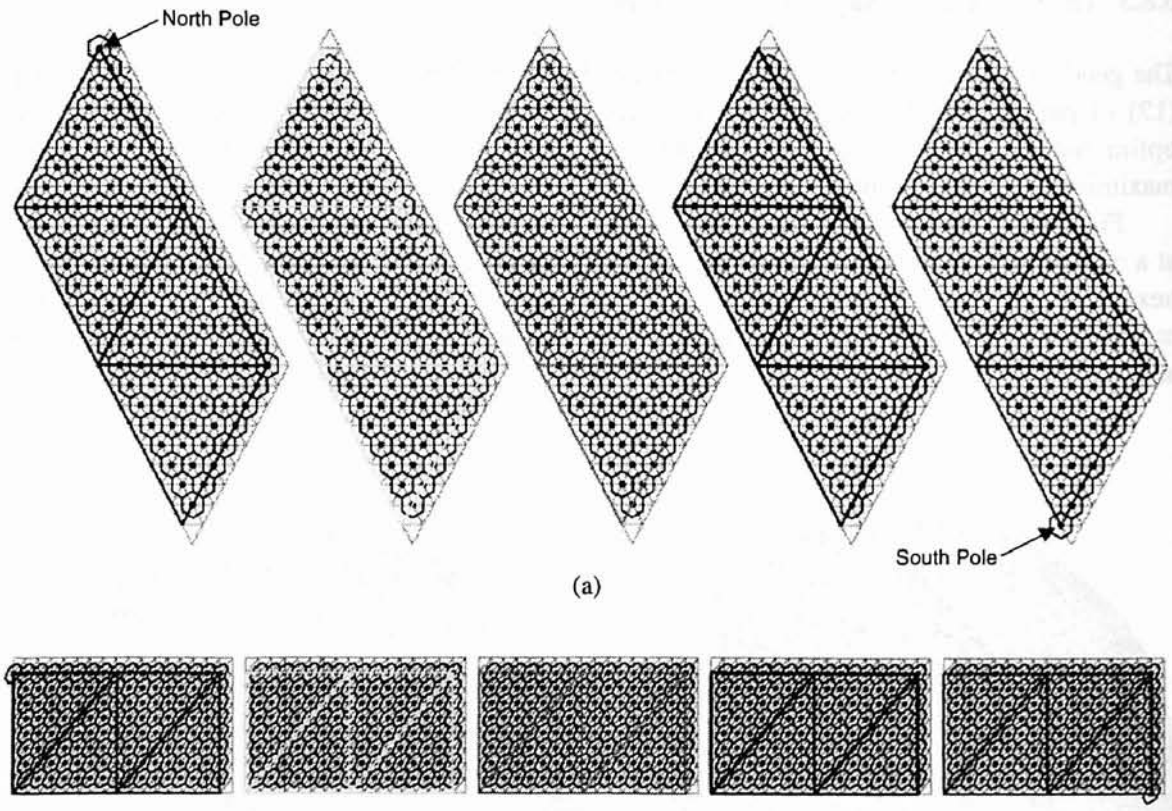


Fig. 3.14 (a) The five grid panels of Fig. 3.13 after unwrapping and stretching them flat. The grid cells at the North Pole and South Pole can be arbitrarily connected to any of the five panels. (b) Illustration of how the cells in the five grid panels can be assigned logically Cartesian coordinates, and how all five panels can then be laid side-by-side to constitute an overall ($im = 90$, $jm = 10$) logically Cartesian grid. *Source:* Randall et al., *Computing in Science and Engineering*, Vol. 4(5), 2002, pp. 32–41, © 2002 IEEE.

TABLE 3.1

Horizontal (*im5*) and Vertical (*jm*) Number of Cells for the Five Panels at Different Resolutions. *Source:* Randall et al., *Computing in Science and Engineering*, Vol. 4(5), 2002, pp. 32–41, © 2002 IEEE.

Number of Cells	*im5*	*jm*
642	18	10
2,562	34	18
10,242	66	34
40,962	130	66
163,842	258	130
655,362	514	258

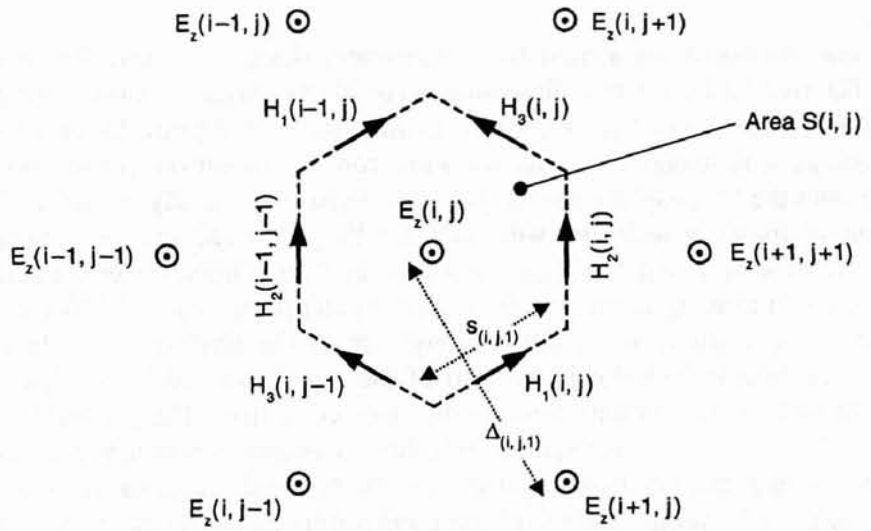


Fig. 3.15 Field components in the vicinity of the (i, j)'th hexagonal cell of the grid of Fig. 3.14.

Algorithm

Fig. 3.15 illustrates the field components in the vicinity of the (i, j)'th hexagonal cell of the grid of Fig. 3.14. As discussed in [24], sample field updates are given by the following expressions:

$$H_1\big|_{i,j}^{n+1/2} = H_1\big|_{i,j}^{n-1/2} + \frac{\Delta t}{\mu_0\, \Delta\big|_{i,j,1}} \left(E_z\big|_{i+1,j}^{n} - E_z\big|_{i,j}^{n} \right) \tag{3.80}$$

$$E_z\big|_{i,j}^{n+1} = E_z\big|_{i,j}^{n} + \frac{\Delta t}{\varepsilon_0\, S\big|_{i,j}} \left(\begin{array}{l} H_1\big|_{i,j}^{n+1/2} s\big|_{i,j,1} + H_2\big|_{i,j}^{n+1/2} s\big|_{i,j,2} + H_3\big|_{i,j}^{n+1/2} s\big|_{i,j,3} - \\ H_1\big|_{i-1,j}^{n+1/2} s\big|_{i-1,j,1} - H_2\big|_{i-1,j-1}^{n+1/2} s\big|_{i-1,j-1,2} - H_3\big|_{i,j-1}^{n+1/2} s\big|_{i,j-1,3} \end{array} \right) \tag{3.81}$$

Field updates for the pentagonal cells are by analogy, noting of course that only five $\boldsymbol{H}$-components circulate about each E_z in these cells.

The periodic boundary condition to be discussed next completes the algorithm. For each of the five panels in Fig. 3.14, the algorithm represented in (3.81) is used to update only the E_z fields for which $2 < i < (im5 - 1)$ and $2 < j < (jm-1)$, along with the E_z fields in the pentagons at the North and South Poles. The set of "ghost" E_z fields located at $(i = 1,\ 1 \le j \le jm-2)$, $(j = 1,\ 1 \le i \le im5-2)$, $(i = im5,\ 2 \le j \le jm-1)$, and $(j = jm,\ 2 \le i \le im5)$, along with the set of E_z fields located at $(i = 1,\ j = jm-1)$ and $(i = im5-1,\ j = 1)$ in the panels not updating the North and South Poles, is then specified after each time-step by setting each such field equal to the corresponding E_z field in the neighboring panel.

Results

We now review the results of a numerical experiment designed to test the isotropy of the geodesic FDTD grid model for two-dimensional TM electromagnetic wave propagation about the lossless Earth-sphere [24]. The idea here is to track an impulsive circular cylindrical wave in time and space as it propagates radially outward from a filamentary current source, travels completely around the Earth-sphere model, and then propagates radially inward to the antipode. Results are shown for a geodesic grid with 40,962 cells ($im5 = 130$ and $jm = 66$) spanning the Earth-sphere, with the time step $\Delta t = 25$ μs. For this simulation, implemented in Fortran 90 on a Dell 530 workstation running Linux, the required computer memory is 3.3 MB and the running time is 2.5 min for a single wave circumnavigation of the Earth-sphere. In comparison, a comparably resolved latitude-longitude grid of the type discussed in Section 3.8.2 would require 2.4 MB and 34 min running time on the same computer. The geodesic grid executes approximately 14 times faster while requiring only 40% more data storage in two dimensions.

Fig. 3.16 is a snapshot visualization of the electric field calculated using the geodesic grid FDTD model projected onto the surface of the Earth-sphere as the radiated wave converges to the antipode. Superimposed on this visualization is: (1) a sample computed contour of equal-amplitude electric field, and (2) a circle centered at the antipode that is drawn to have a radius matching as well as possible that of the equal-amplitude electric field contour.

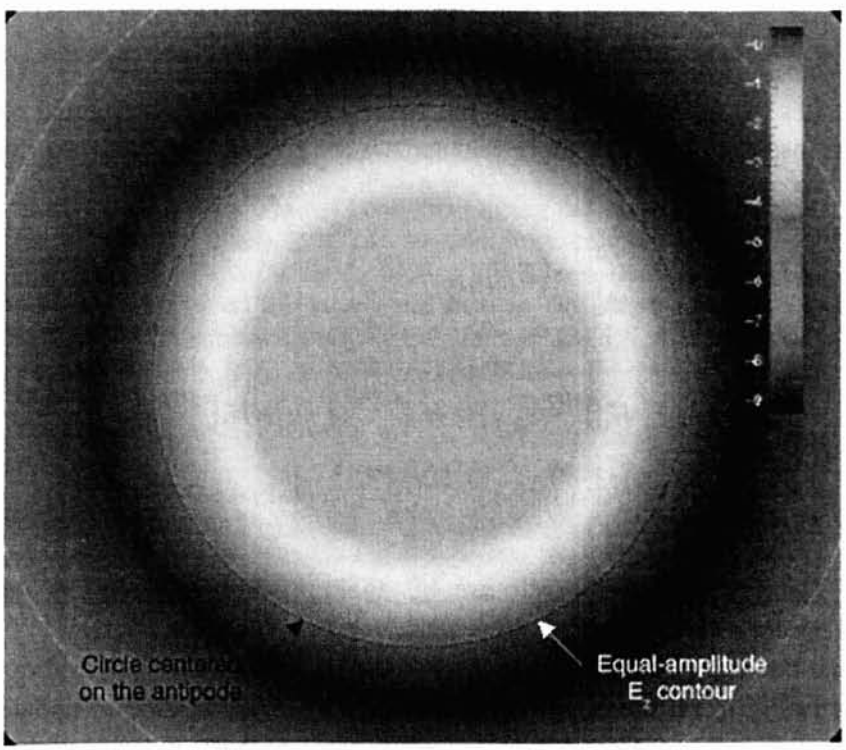


Fig. 3.16 Visualization of the electric field projected onto the surface of the Earth-sphere as the radiated wave converges to the antipode. Note that the white contour of calculated equal-amplitude electric field coincides almost exactly with a black dotted circle centered on the antipode. *Adapted from:* Simpson and Taflove, *IEEE Antennas and Wireless Propagation Lett.*, Vol. 3, 2004, pp. 215–218, © 2004 IEEE.

From Fig. 3.16, we see that there is negligible deviation of the computed equal-amplitude electric field contour from a circle centered at the antipode. This means that the impulsive numerical wave has propagated from its source point almost completely to the antipode in an azimuthally isotropic manner, despite passing through regions of hexagonal and pentagonal grid cells of varying areas. By way of comparison, using this same test, the azimuthal isotropy of numerical wave propagation within the latitude-longitude grid is poorer by approximately one order-of-magnitude for grids of comparable spatial resolution. Combined with the 14:1 speed advantage of the geodesic grid relative to the latitude-longitude grid, it appears likely that the geodesic grid will become the choice for FDTD modeling of the global electromagnetic environment of the Earth in the ELF spectral band.

3.9 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;
- An emerging illustrative application: gridding the Earth to model global electromagnetic wave propagation in the ELF spectral band.

Since the space lattice is the foundation of all FDTD theory and applications discussed in the remaining chapters of this book, the reader is advised to develop a thorough understanding of each topic discussed in this chapter.

REFERENCES

[1] Yee, K. S., "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, 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 Propagat.*, 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 Propagat.*, 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.

[5] Wait, J. R., "Terrestrial propagation of very-low-frequency radio waves, a theoretical investigation," *J. Res. Nat. Bureau Stand.*, Vol. 64D, 1960, p. 153.

[6] Budden, K. G., "The influence of the Earth's magnetic field on radio propagation by waveguide modes," *Proc. Royal Society A*, Vol. 265, 1962, p. 538.

[7] Wait, J. R., "Earth-ionosphere cavity resonances and the propagation of ELF radio waves," *Radio Science*, Vol. 69D, 1965, p. 1057.

[8] Wait, J. R., "Distortion of ELF pulses after propagation through an antipode," *J. Geophys. Res.*, Vol. 74, 1969, pp. 2982–2986.

[9] Wait, J. R., *Electromagnetic Waves in Stratified Media*. New York: Pergamon Press, 1970.

[10] Wait, J. R., "Reflection of VLF radio waves at a junction in the Earth-ionosphere waveguide," *IEEE Trans. Electromagn. Compat.*, Vol. 34, 1992, p. 4.

[11] Cummer, S. A., U. S. Inan, T. F. Bell, and C. P. Barrington-Leigh, "ELF radiation produced by electrical currents in sprites," *Geophys. Res. Lett.*, Vol. 25, 1998, p. 1281.

[12] Williams, E. R., "The Schumann resonance—A global tropical thermometer," *Science*, Vol. 256, 1992, p. 1184.

[13] Strangway, D. W., C. M. Swift, and R. C. Holmer, "The application of audio-frequency magnetotellurics to mineral exploration," *Geophysics*, Vol. 38, 1975, p. 1159.

[14] Parrot, M., "Statistical study of ELF/VLF emissions recorded by a low-altitude satellite during seismic events," *J. Geophys. Res.*, Vol. 99, 1994, pp. 323–339.

[15] Berenger, J.-P., "A FDTD scheme for the computation of VLF-LF propagation in the anisotropic Earth-ionosphere waveguide," *Ann Telecommunications*, Vol. 54, 1999, pp. 297–310.

[16] Berenger, J.-P., "FDTD computation of VLF-LF propagation in the Earth-ionosphere waveguide," *Ann Telecommunications*, Vol. 57, 2002, pp. 1059–1090.

[17] Cummer, S. A., "Modeling electromagnetic propagation in the Earth-ionosphere waveguide," *IEEE Trans. Antennas Propagat.*, Vol. 48, 2000, p. 1420.

[18] Holland, R., "THREDS: A finite-difference time-domain EMP code in 3D spherical coordinates," *IEEE Trans. Nuclear Science*, Vol. 30, 1983, pp. 4592–4595.

[19] Hayakawa, M., and T. Otsuyama, "FDTD analysis of ELF wave propagation in inhomogeneous subionospheric waveguide models," *ACES J.*, Vol. 17, 2002, pp. 239–244.

[20] Otsuyama, T., D. Sakuma, and M. Hayakawa, "FDTD analysis of ELF wave propagation and Schumann resonances for a subionospheric waveguide model," *Radio Science*, Vol. 38(6), 2003, 1103.

[21] Simpson, J. J., and A. Taflove, "Two-dimensional FDTD modeling of impulsive ELF antipodal propagation about the Earth-sphere," *Proc. IEEE AP-S Intl. Symp.*, Vol. 3, San Antonio, TX, June 2002, pp. 678–681.

[22] Simpson, J. J., and A. Taflove, "Two-dimensional FDTD model of antipodal ELF propagation and Schumann resonance of the Earth," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 1, 2002, pp. 53–56.

[23] Simpson, J. J., and A. Taflove, "Three-dimensional FDTD modeling of impulsive ELF propagation about the Earth-sphere," *IEEE Trans. Antennas Propagat.*, Vol. 52, 2004, pp. 443–451.

[24] Simpson, J. J., and A. Taflove, "Efficient modeling of impulsive ELF antipodal propagation about the Earth-sphere using an optimized two-dimensional geodesic FDTD grid," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 3, 2004, pp. 215–218.

[25] Bannister, P., "ELF propagation update," *IEEE J. Oceanic Engineering*, Vol. 9, 1984, pp. 179–188.

[26] NOAA-NGDC "Global Relief CD-ROM"

[27] Hermance, J., "Electrical conductivity of the crust and mantle," in *Global Earth Physics: A Handbook of Physical Constants*, American Geophysical Union, 1995.

[28] Bannister, P., "The determination of representative ionospheric conductivity parameters for ELF propagation in the Earth-ionosphere waveguide," *Radio Science*, Vol. 20, 1985, pp. 977–984.

[29] Wait, J. R., "On the theory of the slow tail portion of atmospheric wave forms," *J. Geophysics Research*, Vol. 65, 1960, pp. 1939–1946.

[30] Randall, D. A., T. D. Ringler, and R. P. Heikes, "Climate modeling with spherical geodesic grids," *Computing in Science and Engineering*, Vol. 4(5), 2002, pp. 32–41.

[31] Heikes, R., and D. A. Randall, "Numerical integration of the shallow-water equations on a twisted icosahedral grid. Part II: A detailed description of the grid and an analysis of numerical accuracy," *Monthly Weather Review*, Vol. 123, 1995, pp. 1881–1887.

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. You only need to prove this for the differential forms of the equations. Note that "source-free" means that there are no imposed electric or magnetic current densities in the space region (i.e., $\boldsymbol{J}_{\text{source}} = 0$ and $\boldsymbol{M}_{\text{source}} = 0$) in (3.6), as well as no free charges, as already indicated in (3.3) and (3.4). Assume a linear, isotropic, nondispersive, and locally homogeneous medium and zero electromagnetic fields at time $t = 0$.

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 with E_z and H_y field components. Use (3.41a) and (3.41c) as the 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, a Gaussian pulse, or a 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. 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 perfect electric conductor 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. Compare the wave-reflection properties of the E_z and H_y components of the wave due to the presence of the perfect magnetic conductor.

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 perfect electric conductors. 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, a Gaussian pulse, or a 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 the time-stepping algorithm. Terminate the grid in $E_x = 0$ and $E_y = 0$ components at its outer boundaries, thereby simulating the presence of perfect electric conductors. 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 perfect magnetic conductor 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 ether" having properties very close to vacuum, but not quite. This "ether" 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 taken into account in order 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-c), 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

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

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu}\frac{\partial E_z}{\partial x} \tag{4.1b}$$

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

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:

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

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

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

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:

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

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

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

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:

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

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

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

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

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

where $c = 1/\sqrt{\mu\varepsilon}$ 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 $\Delta x = \Delta y \equiv \Delta$. 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:

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

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

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

or equivalently

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

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 $\mu = 1$, $\varepsilon = 1$, $\sigma = 0$, $\sigma^* = 0$, and $c = 1$. Letting $j = \sqrt{-1}$, we rewrite Maxwell's equations in compact form as

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

or more simply as

$$j\nabla \times \boldsymbol{V} = \frac{\partial \boldsymbol{V}}{\partial t} \tag{4.8b}$$

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

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

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

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

where $\hat{x}$, $\hat{y}$, and $\hat{z}$ are unit vectors in the x-, y-, and z-coordinate directions. After performing the cross product in (4.10) and writing out the x, y, and z 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

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

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:

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

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

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

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 desired level if we use sufficiently fine 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 results for a numerical plane wave propagating along a diagonal of a three-dimensional cubic lattice ($\tilde{k}_x = \tilde{k}_y = \tilde{k}_z = \tilde{k}/\sqrt{3}$) if $S = 1/\sqrt{3}$. Similarly, ideal dispersion results for a numerical plane wave propagating along a diagonal of a two-dimensional square grid ($\tilde{k}_x = \tilde{k}_y = \tilde{k}/\sqrt{2}$) if $S = 1/\sqrt{2}$. Finally, ideal dispersion results for any numerical wave in a uniform one-dimensional linear grid if $S = 1$, the magic time-step. These reductions to the ideal case have little practical value for two-dimensional and three-dimensional simulations, occurring only for diagonal wave propagation. 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-dimensional 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

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: $\phi = 0°$, $90°$, $180°$, and $270°$. For this case, the solution for $\tilde{k}$ is given by (4.7b), which is repeated here for convenience:

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

The corresponding numerical phase velocity is given by

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

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

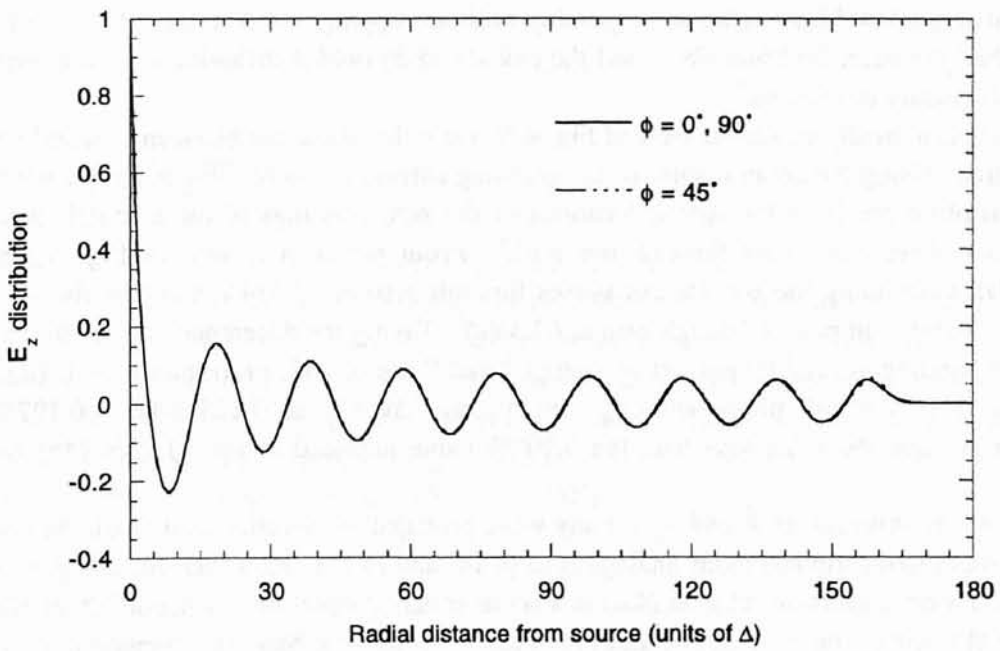
$$\tilde{k} = \frac{2\sqrt{2}}{\Delta}\sin^{-1}\left[\frac{1}{S\sqrt{2}}\sin\left(\frac{\pi S}{N_\lambda}\right)\right] \quad : \text{propagation along grid diagonals} \tag{4.15a}$$

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

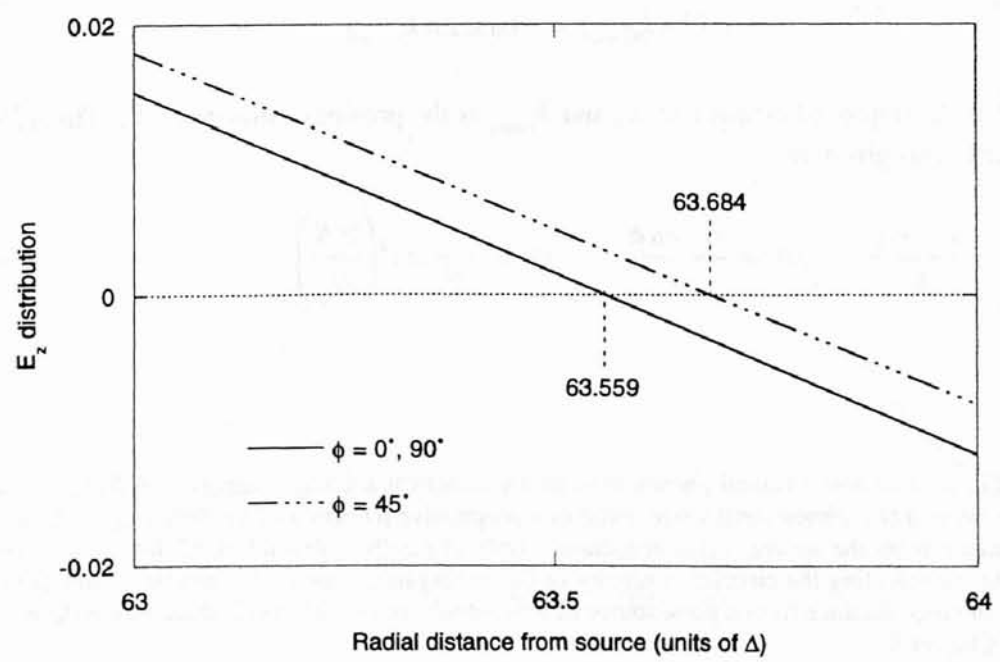
As an example of values arising from these expressions, assume a grid having $S = 0.5$ and $N_\lambda = 20$. Then (4.14b) and (4.15b) yield numerical phase velocities of $0.996892c$ and $0.998968c$, respectively. Both of these velocities are less than c, indicating a phase lag relative to a physical 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 $\Delta x = \Delta y \equiv \Delta = 1.0$. 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 immediately above.

In Fig. 4.1(a), we graph snapshots of the E_z field distribution versus 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 $\phi = 0°$, 90°, or parallel to the grid diagonal $\phi = 45°$. (We note that, by the 90° rotational symmetry of the Cartesian grid geometry, identical field distributions are obtained along $\phi = 0°$ and 90°.)



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



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

Fig. 4.1 Effect of numerical dispersion upon a radially propagating cylindrical wave in a two-dimensional 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 $N_\lambda = 20$.

The snapshots are taken $328\Delta t$ after the beginning of time-stepping. At this time, the wave has not yet reached the outer grid boundary, and the calculated E_z field distribution is free of error due to outer-boundary reflections.[1]

Fig. 4.1(b) is a greatly expanded view of Fig. 4.1(a) at radial distances between 63Δ and 64Δ from the source. Using the zoom feature of the graphing software, it is feasible to evaluate with three-decimal-place precision the spatial locations of the zero-crossings of the E_z distributions along the two observation cuts through the grid.[2] From the data shown in Fig. 4.1(b), the sinusoidal wave along the $\phi = 45^\circ$ cut passes through zero at 63.684Δ, whereas the wave along the $\phi = 0^\circ, 90^\circ$ cut passes through zero at 63.559Δ. Taking the difference, we see that the obliquely propagating sinusoidal numerical energy "leads" the on-axis-propagation by 0.125Δ. This yields a numerical phase-velocity anisotropy $\Delta\tilde{v}_p / \tilde{v}_p \cong 0.125/63.6 = 0.197\%$. This number is only about 5% less than the 0.208% value obtained above from (4.14b) and (4.15b).

To permit determination of $\tilde{k}$ and $\tilde{v}_p$ for any wave-propagation direction ϕ, it would be very useful to derive closed-form equations analogous to (4.14) and (4.15). However, for this general case, the underlying dispersion relation (4.6) is a transcendental equation. A useful alternative approach for obtaining sample values of $\tilde{v}_p$ is to apply the following Newton's method iterative procedure to (4.6):

$$\tilde{k}_{icount+1} = \tilde{k}_{icount} - \frac{\sin^2(A\tilde{k}_{icount}) + \sin^2(B\tilde{k}_{icount}) - C}{A\sin(2A\tilde{k}_{icount}) + B\sin(2B\tilde{k}_{icount})} \tag{4.16a}$$

Here, $\tilde{k}_{icount+1}$ is the improved estimate of $\tilde{k}$, and $\tilde{k}_{icount}$ is the previous estimate of $\tilde{k}$. The A, B, and C are coefficients given by

$$A = \frac{\Delta \cdot \cos\phi}{2}, \qquad B = \frac{\Delta \cdot \sin\phi}{2}, \qquad C = \frac{1}{S^2}\sin^2\left(\frac{\pi S}{N_\lambda}\right) \tag{4.16b}$$

---

[1] From Fig. 4.1(a), we also note a natural phenomenon (*not* a numerical artifact). Namely, radial cylindrical wave propagation in a two-dimensional space involves a progressive reduction of the field amplitude with the radial distance from the source. This reduction profile is ideally independent of the angle of the observation cut, representing the circular symmetry of the propagating wave. The precise nature of the field falloff with radial distance from a point source in a two-dimensional Yee grid is discussed in detail in Section 5.3 of Chapter 5.

[2] The primary cautionary note in applying this procedure is to position the measurement of the zero-crossings sufficiently behind the wavefront so that field jitter generated by the startup of the numerical wave (see Section 4.6) is diminished to a negligible level. Zero-crossings located near the wavefront are contaminated by the leading-edge jitter and cannot be reliably used to measure numerical phase velocity. By positioning the zero-crossing measurement sufficiently behind the wavefront, the transient (homogeneous) solution to the numerical model of Maxwell's equations, that is, the leading-edge jitter, decays to nearly zero. Here, only the driven (particular) solution remains. This is the desired sinusoidal steady-state response.

Additional simplicity results if Δ is normalized to the free-space wavelength, λ_0 (equivalent to setting $\lambda_0 = 1$). Then, a very good guess for $\tilde{k}_{icount=0}$ is simply 2π. For this case, $\tilde{v}_p$ is given by

$$\frac{\tilde{v}_p}{c} = \frac{2\pi}{\tilde{k}_{final\ icount}} \tag{4.17}$$

Usually, only two or three iterations are required for convergence.

Fig. 4.2 graphs results obtained using this procedure which illustrate the variation of $\tilde{v}_p$ with propagation direction ϕ. Here, three different grid-sampling densities N_λ are examined: $N_\lambda = 5$ points per λ_0, $N_\lambda = 10$, and $N_\lambda = 20$. For each case, the Courant stability factor is $S = 0.5$. We see that $\tilde{v}_p < c$ for the three cases studied, and is clearly a function of both ϕ and N_λ. $\tilde{v}_p$ is maximum for waves propagating obliquely within the grid ($\phi = 45°$), and is minimum for waves propagating along either grid axis ($\phi = 0°, 90°$).

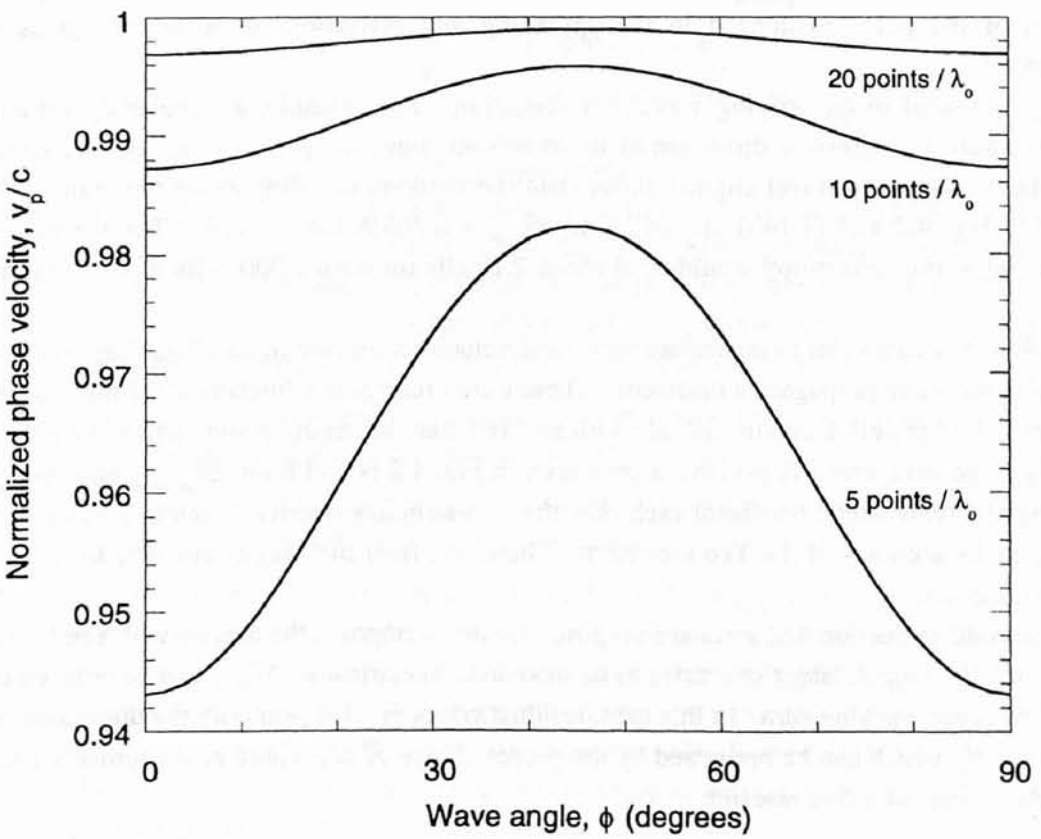


Fig. 4.2 Variation of numerical phase velocity with wave-propagation angle in a two-dimensional FDTD grid for three sampling densities of the square unit cells. $S = c\Delta t/\Delta = 0.5$ for all of the cases.

It is useful to summarize the algorithmic dispersive-error performance by defining two normalized error measures: (1) the physical phase-velocity error $\Delta\tilde{v}_{physical}$, and (2) the velocity-anisotropy error $\Delta\tilde{v}_{aniso}$. These are given by

$$\Delta\tilde{v}_{\text{physical}}\Big|_{N_\lambda} = \frac{\min\left[\tilde{v}_p(\phi)\right] - c}{c} \times 100\% \tag{4.18a}$$

$$\Delta\tilde{v}_{\text{aniso}}\Big|_{N_\lambda} = \frac{\max\left[\tilde{v}_p(\phi)\right] - \min\left[\tilde{v}_p(\phi)\right]}{\min\left[\tilde{v}_p(\phi)\right]} \times 100\% \tag{4.18b}$$

$\Delta\tilde{v}_{\text{physical}}$ is useful in quantifying the phase lead or lag that numerical modes suffer relative to physical modes propagating at c. For example, from Fig. 4.2 and (4.14b), $\Delta\tilde{v}_{\text{physical}} = -0.31\%$ for $N_\lambda = 20$. This means that a sinusoidal numerical wave traveling over a $10\lambda_0$ distance in the grid (200 cells) could develop a lagging phase error up to 11°. We note that $\Delta\tilde{v}_{\text{physical}}$ is a function of N_λ. Since the grid cell size Δ is fixed, for an impulsive wave-propagation problem there exists a spread of effective N_λ values for the spectral components comprising the pulse. This causes a spread of $\Delta\tilde{v}_{\text{physical}}$ over the pulse spectrum, which in turn yields a temporal dispersion of the pulse evidenced in the spreading and distortion of its waveform as it propagates.

$\Delta\tilde{v}_{\text{aniso}}$ is useful in quantifying wavefront distortion. For example, a circular cylindrical wave would suffer progressive distortion of its wavefront, since the portions propagating along the grid diagonals would travel slightly faster than the portions traveling along the major grid axes. From Fig. 4.2 and (4.14b) and (4.15b), $\Delta\tilde{v}_{\text{aniso}} = 0.208\%$ for $N_\lambda = 20$. The wavefront distortion due to this anisotropy would total about 2.1 cells for each 1,000 cells of propagation distance.

It is clear that errors due to inaccurate numerical velocities are cumulative (i.e., they increase linearly with the wave-propagation distance). These errors represent a fundamental limitation of *all* grid-based Maxwell's equations' algorithms, and can be troublesome when modeling electrically large structures. A positive aspect seen in Fig. 4.2 is that both $\Delta\tilde{v}_{\text{physical}}$ and $\Delta\tilde{v}_{\text{aniso}}$ decrease by approximately a 4:1 factor each time the grid-sampling density doubles, indicative of the second-order accuracy of the Yee algorithm. Therefore, finer meshing is one way to control the dispersion error.

As discussed in Section 4.9, there are proposed means to improve the accuracy of Yee-based algorithms to allow much larger structures to be modeled. Specifically, $\Delta\tilde{v}_{\text{aniso}}$ can be reduced to very low levels approaching zero. In this case, residual errors involve primarily the dispersion of $\Delta\tilde{v}_{\text{physical}}$ with N_λ, which can be optimized by the proper choice of Δt. These new approaches are presently the subject of active research.

4.5.2 Intrinsic Grid Velocity Anisotropy

This section provides a deeper discussion of the numerical phase-velocity errors of the Yee algorithm. We show that the nature of the grid discretization determines the velocity anisotropy $\Delta\tilde{v}_{\text{aniso}}$ in a manner that is virtually independent of the time-stepping scheme.

Relation of the Time and Space Discretizations in Generating Numerical Velocity Error

In Section 4.5.1, we determined that $\Delta\tilde{v}_{\text{aniso}} = 0.208\%$ for a two-dimensional Yee algorithm having $N_\lambda = 20$ and $S = 0.5$. An important and revealing question is: How is $\Delta\tilde{v}_{\text{aniso}}$ affected by the choice of S, assuming that N_λ is fixed at 20?

To begin to answer this question, we first choose (what will later be shown to be) the largest possible value of S for numerical stability in two dimensions, $S = 1/\sqrt{2}$. Substituting this value of S into (4.14b) and (4.15b) yields

$$\left.\begin{aligned} \tilde{v}_p(\phi = 0^\circ) &= 0.997926\, c \\ \tilde{v}_p(\phi = 45^\circ) &= c \end{aligned}\right\} \quad \Delta\tilde{v}_{\text{aniso}} = \frac{c - 0.997926\, c}{0.997926\, c} \times 100\% = 0.208\%$$

To three decimal places, there is no change in $\Delta\tilde{v}_{\text{aniso}}$ from the previous value, $S = 0.5$. We next choose a very small value $S = 0.01$ for substitution into (4.14b) and (4.15b):

$$\left.\begin{aligned} \tilde{v}_p(\phi = 0^\circ) &= 0.995859\, c \\ \tilde{v}_p(\phi = 45^\circ) &= 0.997937\, c \end{aligned}\right\} \quad \Delta\tilde{v}_{\text{aniso}} = \frac{0.997937\, c - 0.995859\, c}{0.995859\, c} \times 100\% = 0.208\%$$

Again, there is no change in $\Delta\tilde{v}_{\text{aniso}}$ to three decimal places.

We now suspect that, for a given N_λ, $\Delta\tilde{v}_{\text{aniso}}$ is at most a weak function of S, and therefore is only weakly dependent on Δt. In fact, this is the case. More generally, it can be shown that $\Delta\tilde{v}_{\text{aniso}}$ is only weakly dependent on the type of time-marching scheme used, whether leapfrog, Runge-Kutta, and so forth [2]. Thus, we can say that $\Delta\tilde{v}_{\text{aniso}}$ is virtually an intrinsic characteristic of the space-lattice discretization. Three key points should be made in this regard [2]:

- Numerical-dispersion errors associated with the time discretization are isotropic relative to the propagation direction of the wave.
- The choice of time discretization has little effect upon the phase-velocity anisotropy $\Delta\tilde{v}_{\text{aniso}}$ for $N_\lambda > 10$.
- The choice of time discretization does influence $\Delta\tilde{v}_{\text{physical}}$. However, it is not always true that higher-order time-marching schemes, such as Runge-Kutta, yield less $\Delta\tilde{v}_{\text{physical}}$ than simple Yee leapfrogging. Errors in $\Delta\tilde{v}_{\text{physical}}$ are caused separately by the space and time discretizations, and can either partially reinforce or cancel each other. Thus, the use of fourth-order Runge-Kutta may actually shift the $\tilde{v}_p(\phi)$ profile away from c, representing an increased $\Delta\tilde{v}_{\text{physical}}$ relative to ordinary leapfrogging.

The Associated Eigenvalue Problem

To determine the relative velocity-anisotropy characteristic intrinsic to a space grid, it is useful to set up an eigenvalue problem for the matrix that delineates the spatial derivatives used in the numerical algorithm [2]. Consider as an example the finite-difference system (4.2) for the case of two-dimensional TM_z electromagnetic wave propagation. The associated eigenvalue problem is written as:

$$\Delta\tilde{v}_{\text{physical}}\Big|_{N_\lambda} = \frac{\min\left[\tilde{v}_p(\phi)\right] - c}{c} \times 100\% \tag{4.18a}$$

$$\Delta\tilde{v}_{\text{aniso}}\Big|_{N_\lambda} = \frac{\max\left[\tilde{v}_p(\phi)\right] - \min\left[\tilde{v}_p(\phi)\right]}{\min\left[\tilde{v}_p(\phi)\right]} \times 100\% \tag{4.18b}$$

$\Delta\tilde{v}_{\text{physical}}$ is useful in quantifying the phase lead or lag that numerical modes suffer relative to physical modes propagating at c. For example, from Fig. 4.2 and (4.14b), $\Delta\tilde{v}_{\text{physical}} = -0.31\%$ for $N_\lambda = 20$. This means that a sinusoidal numerical wave traveling over a $10\lambda_0$ distance in the grid (200 cells) could develop a lagging phase error up to 11°. We note that $\Delta\tilde{v}_{\text{physical}}$ is a function of N_λ. Since the grid cell size Δ is fixed, for an impulsive wave-propagation problem there exists a spread of effective N_λ values for the spectral components comprising the pulse. This causes a spread of $\Delta\tilde{v}_{\text{physical}}$ over the pulse spectrum, which in turn yields a temporal dispersion of the pulse evidenced in the spreading and distortion of its waveform as it propagates.

$\Delta\tilde{v}_{\text{aniso}}$ is useful in quantifying wavefront distortion. For example, a circular cylindrical wave would suffer progressive distortion of its wavefront, since the portions propagating along the grid diagonals would travel slightly faster than the portions traveling along the major grid axes. From Fig. 4.2 and (4.14b) and (4.15b), $\Delta\tilde{v}_{\text{aniso}} = 0.208\%$ for $N_\lambda = 20$. The wavefront distortion due to this anisotropy would total about 2.1 cells for each 1,000 cells of propagation distance.

It is clear that errors due to inaccurate numerical velocities are cumulative (i.e., they increase linearly with the wave-propagation distance). These errors represent a fundamental limitation of *all* grid-based Maxwell's equations' algorithms, and can be troublesome when modeling electrically large structures. A positive aspect seen in Fig. 4.2 is that both $\Delta\tilde{v}_{\text{physical}}$ and $\Delta\tilde{v}_{\text{aniso}}$ decrease by approximately a 4:1 factor each time the grid-sampling density doubles, indicative of the second-order accuracy of the Yee algorithm. Therefore, finer meshing is one way to control the dispersion error.

As discussed in Section 4.9, there are proposed means to improve the accuracy of Yee-based algorithms to allow much larger structures to be modeled. Specifically, $\Delta\tilde{v}_{\text{aniso}}$ can be reduced to very low levels approaching zero. In this case, residual errors involve primarily the dispersion of $\Delta\tilde{v}_{\text{physical}}$ with N_λ, which can be optimized by the proper choice of Δt. These new approaches are presently the subject of active research.

4.5.2 Intrinsic Grid Velocity Anisotropy

This section provides a deeper discussion of the numerical phase-velocity errors of the Yee algorithm. We show that the nature of the grid discretization determines the velocity anisotropy $\Delta\tilde{v}_{\text{aniso}}$ in a manner that is virtually independent of the time-stepping scheme.

Relation of the Time and Space Discretizations in Generating Numerical Velocity Error

In Section 4.5.1, we determined that $\Delta\tilde{v}_{\text{aniso}} = 0.208\%$ for a two-dimensional Yee algorithm having $N_\lambda = 20$ and $S = 0.5$. An important and revealing question is: How is $\Delta\tilde{v}_{\text{aniso}}$ affected by the choice of S, assuming that N_λ is fixed at 20?

To begin to answer this question, we first choose (what will later be shown to be) the largest possible value of S for numerical stability in two dimensions, $S = 1/\sqrt{2}$. Substituting this value of S into (4.14b) and (4.15b) yields

$$\left.\begin{aligned} \tilde{v}_p(\phi = 0^\circ) &= 0.997926\ c \\ \tilde{v}_p(\phi = 45^\circ) &= c \end{aligned}\right\} \quad \Delta\tilde{v}_{\text{aniso}} = \frac{c - 0.997926\ c}{0.997926\ c} \times 100\% = 0.208\%$$

To three decimal places, there is no change in $\Delta\tilde{v}_{\text{aniso}}$ from the previous value, $S = 0.5$. We next choose a very small value $S = 0.01$ for substitution into (4.14b) and (4.15b):

$$\left.\begin{aligned} \tilde{v}_p(\phi = 0^\circ) &= 0.995859\ c \\ \tilde{v}_p(\phi = 45^\circ) &= 0.997937\ c \end{aligned}\right\} \quad \Delta\tilde{v}_{\text{aniso}} = \frac{0.997937\ c - 0.995859\ c}{0.995859\ c} \times 100\% = 0.208\%$$

Again, there is no change in $\Delta\tilde{v}_{\text{aniso}}$ to three decimal places.

We now suspect that, for a given N_λ, $\Delta\tilde{v}_{\text{aniso}}$ is at most a weak function of S, and therefore is only weakly dependent on Δt. In fact, this is the case. More generally, it can be shown that $\Delta\tilde{v}_{\text{aniso}}$ is only weakly dependent on the type of time-marching scheme used, whether leapfrog, Runge-Kutta, and so forth [2]. Thus, we can say that $\Delta\tilde{v}_{\text{aniso}}$ is virtually an intrinsic characteristic of the space-lattice discretization. Three key points should be made in this regard [2]:

- Numerical-dispersion errors associated with the time discretization are isotropic relative to the propagation direction of the wave.
- The choice of time discretization has little effect upon the phase-velocity anisotropy $\Delta\tilde{v}_{\text{aniso}}$ for $N_\lambda > 10$.
- The choice of time discretization does influence $\Delta\tilde{v}_{\text{physical}}$. However, it is not always true that higher-order time-marching schemes, such as Runge-Kutta, yield less $\Delta\tilde{v}_{\text{physical}}$ than simple Yee leapfrogging. Errors in $\Delta\tilde{v}_{\text{physical}}$ are caused separately by the space and time discretizations, and can either partially reinforce or cancel each other. Thus, the use of fourth-order Runge-Kutta may actually shift the $\tilde{v}_p(\phi)$ profile away from c, representing an increased $\Delta\tilde{v}_{\text{physical}}$ relative to ordinary leapfrogging.

The Associated Eigenvalue Problem

To determine the relative velocity-anisotropy characteristic intrinsic to a space grid, it is useful to set up an eigenvalue problem for the matrix that delineates the spatial derivatives used in the numerical algorithm [2]. Consider as an example the finite-difference system (4.2) for the case of two-dimensional TM_z electromagnetic wave propagation. The associated eigenvalue problem is written as:

$$-\frac{1}{\mu}\left(\frac{E_z\big|_{i,\,j+1/2} - E_z\big|_{i,\,j-1/2}}{\Delta y}\right) = \Lambda H_x\big|_{i,\,j} \tag{4.19a}$$

$$\frac{1}{\mu}\left(\frac{E_z\big|_{i+1/2,\,j} - E_z\big|_{i-1/2,\,j}}{\Delta x}\right) = \Lambda H_y\big|_{i,\,j} \tag{4.19b}$$

$$\frac{1}{\varepsilon}\left(\frac{H_y\big|_{i+1/2,\,j} - H_y\big|_{i-1/2,\,j}}{\Delta x} - \frac{H_x\big|_{i,\,j+1/2} - H_x\big|_{i,\,j-1/2}}{\Delta y}\right) = \Lambda E_z\big|_{i,\,j} \tag{4.19c}$$

We note that, at any time-step n, the instantaneous values of the E and H fields distributed in space across the grid can be Fourier-transformed with respect to the i and j grid coordinates to provide a spectrum of sinusoidal modes. The result is often called the two-dimensional spatial-frequency spectrum, or the plane-wave eigenmodes of the grid. Let the following specify a typical mode of this spectrum having $\tilde{k}_x$ and $\tilde{k}_y$ as, respectively, the x- and y-components of its numerical wavevector:

$$\begin{aligned} E_z\big|_{I,\,J} &= E_{z_0}\, e^{j(\tilde{k}_x I\Delta x + \tilde{k}_y J\Delta y)} \\ H_x\big|_{I,\,J} &= H_{x_0}\, e^{j(\tilde{k}_x I\Delta x + \tilde{k}_y J\Delta y)} \\ H_y\big|_{I,\,J} &= H_{y_0}\, e^{j(\tilde{k}_x I\Delta x + \tilde{k}_y J\Delta y)} \end{aligned} \tag{4.20}$$

Upon substituting the eigenmode expressions of (4.20) into (4.19a), we obtain

$$-\frac{1}{\mu}\left(\frac{E_{z_0}\, e^{j[\tilde{k}_x I\Delta x + \tilde{k}_y (J+1/2)\Delta y]} - E_{z_0}\, e^{j[\tilde{k}_x I\Delta x + \tilde{k}_y (J-1/2)\Delta y]}}{\Delta y}\right) = \Lambda H_{x_0}\, e^{j(\tilde{k}_x I\Delta x + \tilde{k}_y J\Delta y)} \tag{4.21a}$$

Factoring out the $e^{j(\tilde{k}_x I\Delta x + \tilde{k}_y J\Delta y)}$ term that is common to both sides, we simplify (4.21a):

$$-\frac{E_{z_0}}{\mu\,\Delta y}\left(e^{j[\tilde{k}_y(\Delta y/2)]} - e^{-j[\tilde{k}_y(\Delta y/2)]}\right) = \Lambda H_{x_0} \tag{4.21b}$$

Applying Euler's identity to the complex exponentials yields a sine relation:

$$-\frac{2jE_{z_0}}{\mu\Delta y}\sin\left(\frac{\tilde{k}_y\Delta y}{2}\right) = \Lambda H_{x_0} \quad \rightarrow \quad H_{x_0} = -\frac{2jE_{z_0}}{\Lambda\mu\Delta y}\sin\left(\frac{\tilde{k}_y\Delta y}{2}\right) \tag{4.22a}$$

In a similar manner, substituting the eigenmode expressions of (4.20) into (4.19b) and (4.19c) yields, after the same type of simplifications:

$$H_{y_0} = \frac{2jE_{z_0}}{\Lambda\mu\Delta x}\sin\left(\frac{\tilde{k}_x\Delta x}{2}\right) \tag{4.22b}$$

$$E_{z_0} = \frac{2j}{\Lambda\varepsilon}\left[\frac{H_{y_0}}{\Delta x}\sin\left(\frac{\tilde{k}_x\Delta x}{2}\right) - \frac{H_{x_0}}{\Delta y}\sin\left(\frac{\tilde{k}_y\Delta y}{2}\right)\right] \tag{4.22c}$$

Substituting H_{x_0} of (4.22a) and H_{y_0} of (4.22b) into (4.22c) yields

$$E_{z_0} = \frac{2j}{\Lambda\varepsilon}\left[\begin{array}{l} \dfrac{1}{\Delta x}\cdot\dfrac{2jE_{z_0}}{\Lambda\mu\Delta x}\cdot\sin\left(\dfrac{\tilde{k}_x\Delta x}{2}\right)\cdot\sin\left(\dfrac{\tilde{k}_x\Delta x}{2}\right) \\ \\ -\dfrac{1}{\Delta y}\cdot\dfrac{-2jE_{z_0}}{\Lambda\mu\Delta y}\cdot\sin\left(\dfrac{\tilde{k}_y\Delta y}{2}\right)\cdot\sin\left(\dfrac{\tilde{k}_y\Delta y}{2}\right) \end{array}\right] \tag{4.23}$$

Factoring out the common E_{z_0} term, simplifying, and solving for Λ^2, we obtain

$$\Lambda^2 = -\frac{4}{\mu\varepsilon}\left[\frac{1}{(\Delta x)^2}\sin^2\left(\frac{\tilde{k}_x\Delta x}{2}\right) + \frac{1}{(\Delta y)^2}\sin^2\left(\frac{\tilde{k}_y\Delta y}{2}\right)\right] \tag{4.24}$$

From the elementary properties of the sine function (assuming that $\tilde{k}_x$ and $\tilde{k}_y$ are real numbers for propagating numerical waves), the right-hand side of (4.24) is negative. Hence, Λ is a pure imaginary number given by

$$\Lambda = j2c\left[\frac{1}{(\Delta x)^2}\sin^2\left(\frac{\tilde{k}_x\Delta x}{2}\right) + \frac{1}{(\Delta y)^2}\sin^2\left(\frac{\tilde{k}_y\Delta y}{2}\right)\right]^{1/2} \tag{4.25}$$

where $c = 1/\sqrt{\mu\varepsilon}$ is the speed of light in the homogeneous material being modeled. Following the definition provided in [2], we finally obtain a "normalized numerical phase speed" c^*/c intrinsic to the grid discretization. This is given by

$$\frac{c^*}{c} = \frac{\Lambda_{imag}}{c\tilde{k}} = \frac{2}{\tilde{k}}\left[\frac{1}{(\Delta x)^2}\sin^2\left(\frac{\tilde{k}_x\Delta x}{2}\right) + \frac{1}{(\Delta y)^2}\sin^2\left(\frac{\tilde{k}_y\Delta y}{2}\right)\right]^{1/2} \tag{4.26}$$

Following [2], a closed-form expression for c^*/c can be written by using the approximation $\tilde{k} \cong k$. For additional convenience, we assume a uniform square-cell grid. This yields

$$\frac{c^*}{c} \cong \frac{N_\lambda}{\pi}\left[\sin^2\left(\frac{\pi\cos\phi}{N_\lambda}\right) + \sin^2\left(\frac{\pi\sin\phi}{N_\lambda}\right)\right]^{1/2} \quad : N_\lambda > 10 \tag{4.27}$$

The Meaning of c^*/c

The reader is cautioned that c^*/c is *not* the same as $\tilde{v}_p/c$. This is because the derivation of c^*/c utilizes no information regarding the time-stepping process. Thus, c^*/c cannot be used directly to determine $\Delta\tilde{v}_{\text{physical}}$ defined in (4.18a). However, c^*/c does provide direct information regarding $\Delta\tilde{v}_{\text{aniso}}$ defined in (4.18b). Following [2], we can expand (4.27) to isolate the leading-order velocity-anisotropy term. This yields a simple expression for $\Delta\tilde{v}_{\text{aniso}}$ that is useful for $N_\lambda > 10$:

$$\Delta\tilde{v}_{\text{aniso}}\Big|_{\text{Yee}} \cong \frac{\max\left[\frac{c^*(\phi)}{c}\right] - \min\left[\frac{c^*(\phi)}{c}\right]}{\min\left[\frac{c^*(\phi)}{c}\right]} \times 100\% \cong \frac{\pi^2}{12\left(N_\lambda\right)^2} \times 100\% \tag{4.28}$$

For example, (4.28) provides $\Delta\tilde{v}_{\text{aniso}} \cong 0.206\%$ for $N_\lambda = 20$. This is very close to the 0.208% value previously obtained using (4.14b) and (4.15b), which are the exact solutions of the full numerical dispersion relation for $\phi = 0°$ and 45°, respectively.

In summary, we can use (4.28) to estimate the numerical phase-velocity anisotropy $\Delta\tilde{v}_{\text{aniso}}$ of the basic Yee algorithm applied to a square-cell grid without having to resort to the Newton's method solution (4.16). This approach provides a convenient means to compare the relative anisotropy of alternative space-gridding techniques, including the higher-order methods and non-Cartesian meshes to be discussed in Section 4.9.

4.6 COMPLEX-VALUED NUMERICAL WAVENUMBERS

Section 2.6.3 of Chapter 2 demonstrated that the central-difference solution of the one-dimensional scalar wave equation has a low-sampling-density regime that allows complex-valued numerical wavenumbers. In this sampling regime, a spatially decaying numerical wave can propagate faster than light. This can cause weak, nonphysical, superluminal energy to appear in the numerical simulation at the leading edges of sharply defined pulses.

In an analysis similar to that of Section 2.6.3 and originally presented in [3], this section investigates the possibility of complex-valued numerical wavenumbers existing for the FDTD solution of Maxwell's equations. It is shown that the Yee algorithm also has a low-sampling-density regime that allows complex-valued numerical wavenumbers, and consequently, weak superluminal propagation.

4.6.1 Case 1: Numerical Wave Propagation Along the Principal Lattice Axes

Consider again numerical wave propagation along the major axes of a Yee space grid. For convenience, we rewrite (4.14a), the corresponding numerical dispersion relation:

$$\tilde{k} = \frac{2}{\Delta}\sin^{-1}\left[\frac{1}{S}\sin\left(\frac{\pi S}{N_\lambda}\right)\right] \equiv \frac{2}{\Delta}\sin^{-1}(\zeta) \tag{4.29a}$$

where

$$\zeta = \frac{1}{S}\sin\left(\frac{\pi S}{N_\lambda}\right) \tag{4.29b}$$

As discussed in [3] and Section 2.6.3, caution must be exercised in evaluating numerical dispersion relations such as (4.29), since it is possible to choose S and N_λ such that $\tilde{k}$ is complex. In the specific case of (4.29), 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

$$N_\lambda|_{\text{transition}} = \frac{\pi S}{\sin^{-1}(S)} \tag{4.30}$$

For a grid-sampling density greater 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, $\tilde{v}_p < c$. For a coarser grid-sampling density $N_\lambda < N_\lambda|_{\text{transition}}$, $\tilde{k}$ is a complex number and the numerical wave undergoes a nonphysical exponential decay while propagating [3]. Further, in this coarse-resolution regime, $\tilde{v}_p$ can exceed c [3].

We now discuss how $\tilde{k}$ and $\tilde{v}_p$ vary with grid sampling N_λ both above and below the transition between real and complex numerical wavenumbers.

Real-Numerical-Wavenumber Regime

For $N_\lambda > N_\lambda|_{\text{transition}}$, we have from (4.29a)

$$\tilde{k}_{real} = \frac{2}{\Delta}\sin^{-1}\left[\frac{1}{S}\sin\left(\frac{\pi S}{N_\lambda}\right)\right] \quad ; \qquad \tilde{k}_{imag} = 0 \tag{4.31}$$

The numerical phase velocity is given by (4.14b) as

$$\tilde{v}_p = \frac{\omega}{\tilde{k}_{real}} = \frac{\pi}{N_\lambda \sin^{-1}\left[\frac{1}{S}\sin\left(\frac{\pi S}{N_\lambda}\right)\right]}\, c \tag{4.32}$$

From (2.21), the wave-amplitude multiplier per grid cell of propagation is

$$e^{\tilde{k}_{imag}\,\Delta} \equiv e^{-\alpha\Delta} = e^{0} = 1 \tag{4.33}$$

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

Complex-Numerical-Wavenumber Regime

For $N_\lambda < N_\lambda|_{\text{transition}}$, we observe that $\zeta > 1$ in (4.29a). The following relation for the complex-valued arc-sine function [3, 4] becomes useful:

$$\sin^{-1}(\zeta) = -j \ln\left(j\zeta + \sqrt{1 - \zeta^2}\right) \tag{4.34}$$

Substituting (4.34) into (4.29a) yields

$$\tilde{k} = \frac{2}{\Delta}\left[-j \ln\left(j\zeta + \sqrt{1 - \zeta^2}\right)\right] = -\frac{j2}{\Delta} \ln\left(j\zeta + j\sqrt{\zeta^2 - 1}\right) \tag{4.35a}$$

Expressing the argument of the natural logarithm in polar form, we obtain

$$\tilde{k} = -\frac{j2}{\Delta} \ln\left[\left(\zeta + \sqrt{\zeta^2 - 1}\right) e^{+j\pi/2}\right] = \frac{\pi}{\Delta} - j\frac{2}{\Delta} \ln\left(\zeta + \sqrt{\zeta^2 - 1}\right) \tag{4.35b}$$

This yields

$$\tilde{k}_{real} = \frac{\pi}{\Delta} \; ; \qquad \tilde{k}_{imag} = -\frac{2}{\Delta} \ln\left(\zeta + \sqrt{\zeta^2 - 1}\right) \tag{4.36}$$

From [3], the numerical phase velocity is given by

$$\tilde{v}_p = \frac{\omega}{\tilde{k}_{real}} = \frac{\omega}{(\pi / \Delta)} = \frac{2\pi f \Delta}{\pi} = \frac{2 f \lambda_0}{N_\lambda} = \frac{2}{N_\lambda} c \tag{4.37a}$$

and from (2.21), the wave-amplitude multiplier per grid cell of propagation is

$$e^{\tilde{k}_{imag}\,\Delta} \equiv e^{-\alpha\Delta} = e^{-2 \ln\left(\zeta + \sqrt{\zeta^2 - 1}\right)} = \frac{1}{\left(\zeta + \sqrt{\zeta^2 - 1}\right)^2} \tag{4.37b}$$

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 [3]. 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

$$\lambda_{0,\min} = c/f_{\max} = 2c\Delta t \tag{4.38a}$$

The corresponding minimum spatial-sampling density is

$$N_{\lambda,\min} = \lambda_{0,\min}/\Delta = 2c\Delta t/\Delta = 2S \tag{4.38b}$$

Then from (4.37a), the maximum numerical phase velocity is given by

$$\tilde{v}_{p,\max} = \frac{2}{N_{\lambda,\min}} c = \frac{2}{2S} c = \frac{c}{S} \tag{4.39a}$$

From the definition of S, this maximum phase velocity can also be expressed as

$$\tilde{v}_{p,\max} = \frac{1}{S} c = \left(\frac{\Delta}{c\Delta t}\right) c = \frac{\Delta}{\Delta t} \tag{4.39b}$$

This relation tells us that in one time-step, a numerical value can propagate at most one space cell. This is intuitively correct, given the local nature of the spatial difference used in the Yee algorithm. That is, a field point more than one space cell away from a source point that undergoes a sudden change cannot possibly "feel" the effect of that change during the next time-step. 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. These results are analogous to those derived in Section 2.6.3 for the numerical implementation of the one-dimensional scalar wave equation.

4.6.2 Case 2: Numerical Wave Propagation Along a Grid Diagonal

We next explore the possibility of complex-valued wavenumbers arising for oblique numerical wave propagation in a square-cell grid. For convenience, we rewrite (4.15a), the corresponding numerical dispersion relation:

$$\tilde{k} = \frac{2\sqrt{2}}{\Delta} \sin^{-1}\left[\frac{1}{S\sqrt{2}} \sin\left(\frac{\pi S}{N_\lambda}\right)\right] \equiv \frac{2\sqrt{2}}{\Delta} \sin^{-1}(\zeta) \tag{4.40a}$$

where

$$\zeta = \frac{1}{S\sqrt{2}} \sin\left(\frac{\pi S}{N_\lambda}\right) \tag{4.40b}$$

Similar to the previous case of numerical wave propagation along the principal lattice axes, it is possible to choose S and N_λ such that $\tilde{k}$ is complex. In the specific case of (4.40a), the transition between real and complex values of $\tilde{k}$ occurs when $\zeta = 1$. Solving for N_λ at this transition results in

$$N_\lambda\big|_{\text{transition}} = \frac{\pi S}{\sin^{-1}\left(S\sqrt{2}\right)} \tag{4.41}$$

We now discuss how $\tilde{k}$ and $\tilde{v}_p$ vary with grid sampling N_λ both above and below the transition between real and complex numerical wavenumbers.

Real-Numerical-Wavenumber Regime

For $N_\lambda > N_\lambda|_{\text{transition}}$, we have from (4.40a)

$$\tilde{k}_{real} = \frac{2\sqrt{2}}{\Delta}\sin^{-1}\left[\frac{1}{S\sqrt{2}}\sin\left(\frac{\pi S}{N_\lambda}\right)\right] \; ; \qquad \tilde{k}_{imag} = 0 \tag{4.42}$$

The numerical phase velocity is given by (4.15b) as

$$\tilde{v}_p = \frac{\omega}{\tilde{k}_{real}} = \frac{\pi}{N_\lambda\sqrt{2}\,\sin^{-1}\left[\frac{1}{S\sqrt{2}}\sin\left(\frac{\pi S}{N_\lambda}\right)\right]}\,c \tag{4.43}$$

From (2.21), the wave-amplitude multiplier per grid cell of propagation is

$$e^{\tilde{k}_{imag}\,\Delta} \equiv e^{-\alpha\Delta} = e^{0} = 1 \tag{4.44}$$

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

Complex-Numerical-Wavenumber Regime

For $N_\lambda < N_\lambda|_{\text{transition}}$, we observe that $\zeta > 1$ in (4.40a). Substituting the complex-valued arc-sine function of (4.34) into (4.40a) yields

$$\tilde{k} = \frac{2\sqrt{2}}{\Delta}\left[-j\,\ln\left(j\zeta + \sqrt{1-\zeta^2}\right)\right] \tag{4.45}$$

By direct analogy with the development in (4.35a) and (4.35b), we obtain

$$\tilde{k}_{real} = \frac{\pi\sqrt{2}}{\Delta} \; ; \qquad \tilde{k}_{imag} = -\frac{2\sqrt{2}}{\Delta} \ln\left(\zeta + \sqrt{\zeta^2 - 1}\right) \tag{4.46}$$

The numerical phase velocity for this case is

$$\tilde{v}_p = \frac{\omega}{\tilde{k}_{real}} = \frac{\omega}{\left(\pi\sqrt{2}/\Delta\right)} = \frac{\sqrt{2} f \lambda_0}{N_\lambda} = \frac{\sqrt{2}}{N_\lambda} c \tag{4.47a}$$

and the wave-amplitude multiplier per grid cell of propagation is

$$e^{\tilde{k}_{imag}\Delta} \equiv e^{-\alpha\Delta} = e^{-2\sqrt{2}\ln\left(\zeta + \sqrt{\zeta^2 - 1}\right)} = \frac{1}{\left(\zeta + \sqrt{\zeta^2 - 1}\right)^{2\sqrt{2}}} \tag{4.47b}$$

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

We again consider the possibility of $\tilde{v}_p$ exceeding c. From our previous discussion of (4.38a) and (4.38b), the minimum free-space wavelength that can be sampled without aliasing is $\lambda_{0,\min} = c/f_{\max} = 2c\Delta t$, and the corresponding minimum spatial-sampling density is $N_{\lambda,\min} = \lambda_{0,\min}/\Delta = 2S$. Then, from (4.47a), the maximum numerical phase velocity is given by

$$\tilde{v}_{p,\max} = \frac{\sqrt{2}}{N_{\lambda,\min}} c = \frac{\sqrt{2}}{2S} c \tag{4.48a}$$

From the definition of S, this maximum phase velocity can also be expressed as

$$\tilde{v}_{p,\max} = \frac{\sqrt{2}}{2}\left(\frac{\Delta}{c\Delta t}\right) c = \frac{\sqrt{2}\Delta}{2\Delta t} \tag{4.48b}$$

This relation tells us that in two time-steps, a numerical value can propagate at most $\sqrt{2}\Delta$ along the grid diagonal. We can show that this upper bound on $\tilde{v}_p$ is intuitively correct, given the local nature of the spatial differences used in the Yee algorithm. Consider two nearest-neighbor field points $P_{i,j}$ and $P_{i+1,j+1}$ along a grid diagonal, and how a sudden change at $P_{i,j}$ could be communicated to $P_{i+1,j+1}$. Now, a basic principle is that the Yee algorithm can communicate field data only along Cartesian (x and y) grid lines, and not along grid diagonals. Thus, at the minimum, a single time-step would be needed to transfer any part of the field perturbation at $P_{i,j}$ over a distance of one space cell in the x-direction to $P_{i+1,j}$. Then, a second time-step would be needed, at the minimum, to transfer any part of the resulting field perturbation at $P_{i+1,j}$ over a distance of one space cell in the y-direction to reach $P_{i+1,j+1}$. Because the distance between $P_{i,j}$ and $P_{i+1,j+1}$ is $\sqrt{2}\Delta$, the maximum effective velocity of signal transmission between the two points is $\sqrt{2}\Delta/2\Delta t$. By this reasoning, we see that $\tilde{v}_{p,\max}$ is independent of the material parameters modeled in the grid. It is an inherent property of the FDTD grid and its method of obtaining space derivatives.

4.6.3 Example of Calculation of Numerical Phase Velocity and Attenuation

This section provides sample calculations of values of the numerical phase velocity and the exponential attenuation constant for the case of a two-dimensional square-cell grid. These calculations are based upon the dispersion analyses of Sections 4.6.1 and 4.6.2.

Fig. 4.3 graphs the normalized numerical phase velocity and the exponential attenuation constant per grid cell as a function of grid-sampling density N_λ. A Courant factor $S = 0.5$ is assumed. For propagation along the principal grid axes $\phi = 0°, 90°$, a minimum value of $\tilde{v}_p = (2/3)c$ is reached at $N_\lambda = 3$. This sampling density is also the onset of attenuation. 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 $N_\lambda \to 1$. In this limit, as well, the attenuation constant approaches a value of 2.634 nepers/cell.

For propagation along the grid diagonal at $\phi = 45°$, a minimum value of $\tilde{v}_p = (\sqrt{2}/2)c$ is reached at $N_\lambda = 2$. This point is also the onset of exponential attenuation. As N_λ is reduced below 2, $\tilde{v}_p$ increases inversely with N_λ. Eventually, $\tilde{v}_p$ exceeds c for $N_\lambda < \sqrt{2}$, and reaches a limiting velocity of $\sqrt{2}c$ as $N_\lambda \to 1$. In this limit, as well, the attenuation constant approaches a value of 2.493 nepers/cell.

For both the on-axis and oblique cases of numerical wave propagation, we see that very coarsely resolved wave modes in the grid can propagate at superluminal speeds, but are rapidly attenuated [3]. This is similar to the scalar-wave-equation results obtained in Section 2.6.3.

Fig. 4.4 graphs the percent error in the numerical phase velocity relative to c for lossless wave propagation along the principal grid axes $\phi = 0°, 90°$. In the present example, wherein $S = 0.5$, this lossless propagation regime exists for $N_\lambda \geq 3$. Fig. 4.4 also graphs the percent velocity error for lossless wave propagation along the grid diagonal $\phi = 45°$. This lossless regime exists for $N_\lambda \geq 2$ for $S = 0.5$. As $N_\lambda >> 10$, we see that the numerical phase-velocity error at each wave-propagation angle diminishes as the inverse square of N_λ. This is indicative of the second-order-accurate nature of the Yee algorithm.

4.6.4 Example of Calculation of Wave Propagation

Fig. 4.5 presents an example of the calculation of a radially outward-propagating cylindrical wave in a two-dimensional TM_z Yee grid. A 360×360-cell square grid with $\Delta x = \Delta y = \Delta = 1.0$ is used in this example. The grid is numerically excited at its center point by applying a unit-step time-function to a single E_z field component. We assume the Courant factor $S = \sqrt{2}/2$, which yields dispersion-free propagation for numerical plane-wave modes propagating along the grid diagonals $\phi = 45°, 135°, 225°$, and $315°$. In Fig. 4.5(a), we graph snapshots of the E_z distribution versus radial distance from the source. Here, field observations are made along cuts through the grid passing through the source and either parallel to the principal grid axes $\phi = 0°, 90°$ or parallel to the grid diagonal $\phi = 45°$. The snapshots are taken $232\Delta t$ after the beginning of time-stepping. At this time, the wave has not yet reached the outer grid boundary, and the calculated E_z distribution is free of error due to outer-boundary reflections.

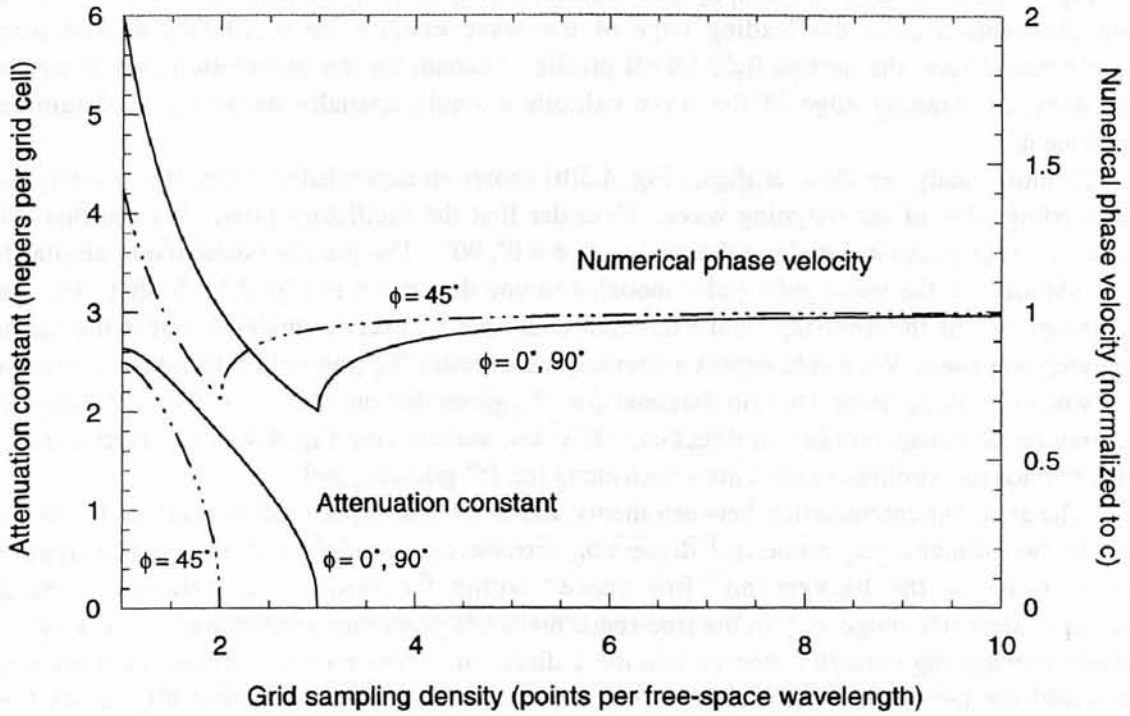


Fig. 4.3 Normalized numerical phase velocity and exponential attenuation constant per grid cell versus the grid-sampling density for on-axis and oblique wave propagation. $S = 0.5$ is assumed.

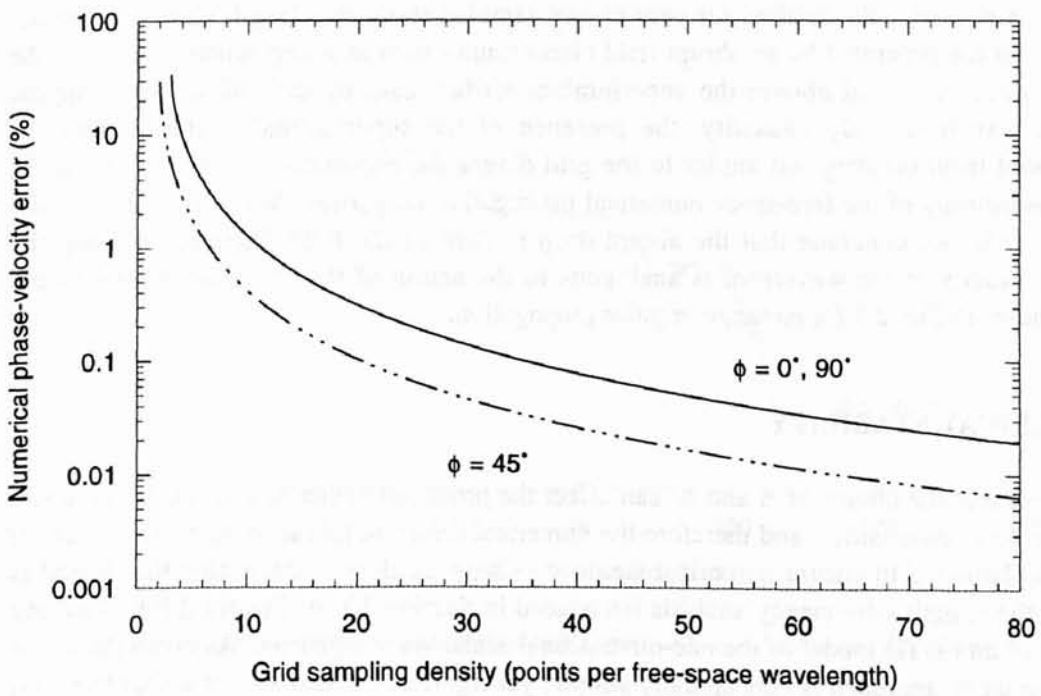


Fig. 4.4 Percent numerical phase-velocity error relative to the free-space speed of light as a function of the grid-sampling density for on-axis and oblique wave propagation. $S = 0.5$ is assumed.

Fig. 4.5(a) illustrates two nonphysical artifacts arising from numerical dispersion. First, for both observation cuts, the leading edge of the wave exhibits an oscillatory spatial jitter superimposed upon the normal field falloff profile. Second, for the observation cuts along the grid axes, the leading edge of the wave exhibits a small, spatially decaying, superluminal component.

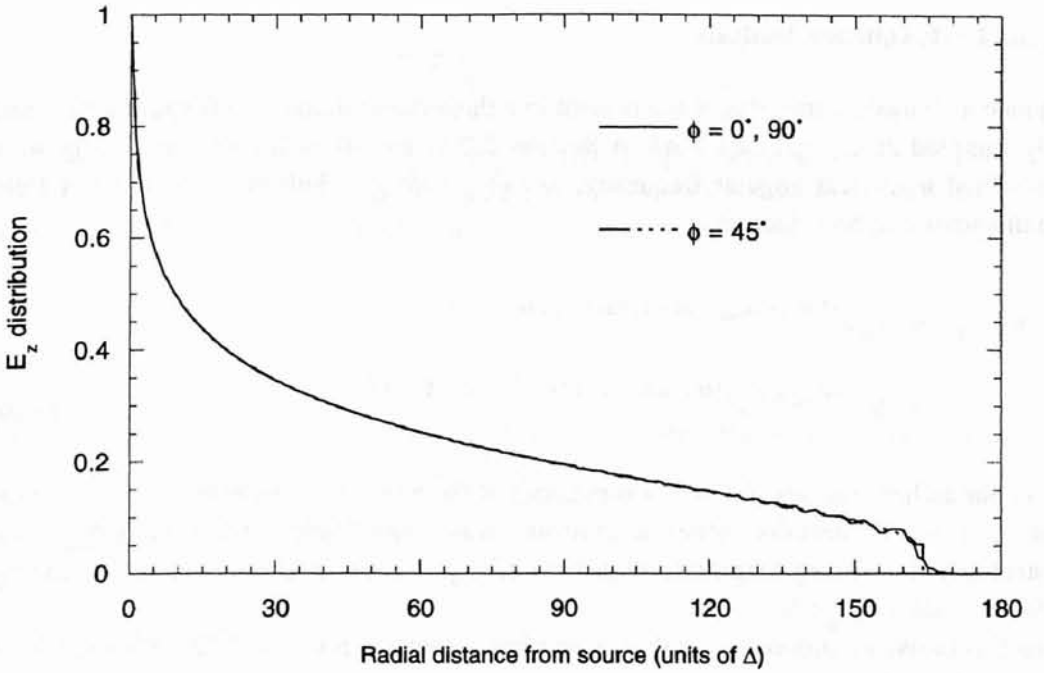
To more easily see these artifacts, Fig. 4.5(b) shows an expanded view in the vicinity of the leading edge of the outgoing wave. Consider first the oscillatory jitter. We see that this artifact is most pronounced along the grid axes $\phi = 0°, 90°$. The jitter is qualitatively similar to that exhibited by the rectangular pulse modeled in one dimension in Fig. 2.3. In that example, the abrupt rise of the unit-step field excitation could not be exactly modeled unless the magic time-step was used. We might expect a similarly exact result (i.e., no field jitter), for the present case when observing along the grid diagonal $\phi = 45°$, given that our choice of $S = \sqrt{2}/2$ implies dispersionless propagation in that direction. However, we see from Fig. 4.4(b) that there exists a small but nonzero oscillatory field jitter even along the 45° grid diagonal.

The apparent contradiction between theory and numerical experiment is resolved by noting that, in two dimensions, numerical dispersion introduces a slightly anisotropic propagation characteristic of the background "free space" within the grid versus azimuth angle ϕ. This equivalent inhomogeneity in the free-space dielectric properties scatters part of the original radially propagating numerical energy into the ϕ direction. Thus, no point behind the wavefront can avoid the radial field oscillation — this artifact is communicated azimuthally as well as radially in space. An implication of this result is that we must be cautious in using plane-wave dispersion analyses to probe the behavior of propagating numerical waves having curved fronts.

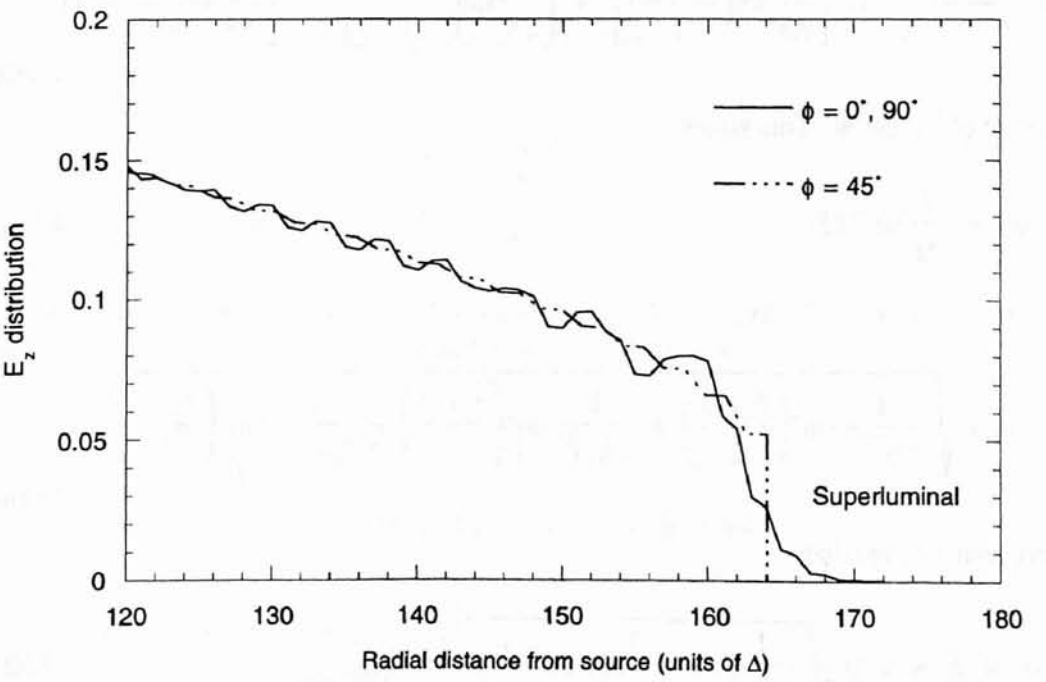
We last consider the superluminal artifact present at the leading edge of the wave shown in Fig. 4.5(b) for $\phi = 0°, 90°$; but not for $\phi = 45°$. The theory developed in Section 4.6.1 points toward the existence of this artifact for very poorly sampled short-wavelength numerical modes in the grid that are generated by an abrupt field discontinuity such as a step function. Unlike the oscillatory jitter discussed above, the superluminal artifact can, by definition, never appear behind the wavefront. By causality, the presence of the superluminal artifact cannot be communicated from off-diagonal angles to the grid diagonals, regardless of the strength of the azimuthal anisotropy of the free-space numerical propagation properties. With $S = \sqrt{2}/2$ in the present example, we conclude that the abrupt drop to zero of the field distribution along the $\phi = 45°$ cut exactly at the wavefront is analogous to the action of the one-dimensional magic time-step shown in Fig. 2.3 for rectangular pulse propagation.

4.7 NUMERICAL STABILITY

We have seen that the choice of Δ and Δt can affect the propagation characteristics of numerical waves in the Yee space lattice, and therefore the numerical error. In this section, we show that Δt must also be bounded to ensure numerical stability. Our procedure to determine this bound is based upon the complex-frequency analysis introduced in Section 2.7 of Chapter 2 for analyzing the stability of an FDTD model of the one-dimensional scalar wave equation. As noted there, the complex-frequency approach is conceptually simple, yet rigorous. It also allows straightforward estimates of the growth-rate of unstable solutions. The method is based upon the analysis of numerical dispersion relation (4.5) for the two-dimensional case, and (4.12) for the three-dimensional case.



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



(b) Expanded view of (a) at distances between 120Δ and 180Δ from the source.

Fig. 4.5 Effect of numerical dispersion upon a radially propagating cylindrical wave in a two-dimensional TM Yee grid. The grid is excited at its center point by applying a unit-step time-function to a single E_z field component. The Courant stability factor is $S = \sqrt{2}/2$.

4.7.1 Complex-Frequency Analysis

We postulate a sinusoidal traveling wave present in a three-dimensional FDTD space lattice and discretely sampled at (x_I, y_J, z_K, t_n). As in Section 2.7.1, we allow for the possibility of a complex-valued numerical angular frequency, $\tilde{\omega} = \tilde{\omega}_{\text{real}} + j\tilde{\omega}_{\text{imag}}$. Following (4.3a-c), a field vector in this wave can be written as

$$\begin{aligned} V|_{I,J,K}^{n} &= V_0 e^{j\left[(\tilde{\omega}_{\text{real}} + j\tilde{\omega}_{\text{imag}})n\Delta t - \tilde{k}_x I\Delta x - \tilde{k}_y J\Delta y - \tilde{k}_z K\Delta z\right]} \\ &= V_0 e^{-\tilde{\omega}_{\text{imag}} n\Delta t} e^{j\left(\tilde{\omega}_{\text{real}} n\Delta t - \tilde{k}_x I\Delta x - \tilde{k}_y J\Delta y - \tilde{k}_z K\Delta z\right)} \end{aligned} \tag{4.49}$$

Here, as in our earlier analyses, $\tilde{k}$ is the wavenumber of the numerical sinusoidal traveling wave. We note that (4.49) permits either a constant wave amplitude with time ($\tilde{\omega}_{\text{imag}} = 0$), an exponentially decreasing amplitude with time ($\tilde{\omega}_{\text{imag}} > 0$), or an exponentially increasing amplitude with time ($\tilde{\omega}_{\text{imag}} < 0$).

Given this basis, we proceed to analyze numerical dispersion relation (4.12), allowing for a complex-valued angular frequency:

$$\left[\frac{1}{c\Delta t}\sin\left(\frac{\tilde{\omega}\Delta t}{2}\right)\right]^2 = \left[\frac{1}{\Delta x}\sin\left(\frac{\tilde{k}_x\Delta x}{2}\right)\right]^2 + \left[\frac{1}{\Delta y}\sin\left(\frac{\tilde{k}_y\Delta y}{2}\right)\right]^2 + \left[\frac{1}{\Delta z}\sin\left(\frac{\tilde{k}_z\Delta z}{2}\right)\right]^2 \tag{4.50}$$

We first solve (4.50) for $\tilde{\omega}$. This yields

$$\tilde{\omega} = \frac{2}{\Delta t}\sin^{-1}(\xi) \tag{4.51a}$$

where

$$\xi = c\Delta t\sqrt{\frac{1}{(\Delta x)^2}\sin^2\left(\frac{\tilde{k}_x\Delta x}{2}\right) + \frac{1}{(\Delta y)^2}\sin^2\left(\frac{\tilde{k}_y\Delta y}{2}\right) + \frac{1}{(\Delta z)^2}\sin^2\left(\frac{\tilde{k}_z\Delta z}{2}\right)} \tag{4.51b}$$

We observe from (4.51b) that

$$0 \le \xi \le c\Delta t\sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2} + \frac{1}{(\Delta z)^2}} \equiv \xi_{\text{upper bound}} \tag{4.52}$$

for all possible real values of $\tilde{k}$, that is, those numerical waves having zero exponential attenuation per grid space cell. $\xi_{\text{upper bound}}$ is obtained when each $\sin^2$ term under the square root of (4.51b) simultaneously reaches a value of 1. This occurs for the propagating numerical wave having the wavevector components

$$\tilde{k}_x = \pm\frac{\pi}{\Delta x} \; ; \qquad \tilde{k}_y = \pm\frac{\pi}{\Delta y} \; ; \qquad \tilde{k}_z = \pm\frac{\pi}{\Delta z} \tag{4.53}$$

It is clear that $\xi_{\text{upper bound}}$ can exceed 1 depending upon the choice of Δt. This can yield complex values of $\sin^{-1}(\xi)$ in (4.51a), and therefore complex values for $\tilde{\omega}$ which give rise to numerical instability. To investigate further, we divide the range of ξ given in (4.52) into two subranges, as follows.

Stable Range: $0 \le \xi \le 1$

Here, $\sin^{-1}(\xi)$ is real-valued and hence, real values of $\tilde{\omega}$ are obtained in (4.51a). With $\tilde{\omega}_{\text{imag}} = 0$, (4.49) yields a constant wave amplitude with time.

Unstable Range: $1 < \xi < \xi_{\text{upper bound}}$

This subrange exists only if

$$\xi_{\text{upper bound}} = c\,\Delta t \sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2} + \frac{1}{(\Delta z)^2}} > 1 \tag{4.54a}$$

The unstable range is defined in an equivalent manner by

$$\Delta t > \frac{1}{c\sqrt{\dfrac{1}{(\Delta x)^2} + \dfrac{1}{(\Delta y)^2} + \dfrac{1}{(\Delta z)^2}}} \equiv \Delta t_{\substack{\text{stable} \\ \text{limit-3D}}} \tag{4.54b}$$

To prove the claim of instability for the range $\xi > 1$, we apply the complex-valued $\sin^{-1}(\xi)$ function given by (2.34), which is repeated here for convenience:

$$\sin^{-1}(\xi) = -j \ln\left(j\xi + \sqrt{1 - \xi^2}\right) \tag{4.55}$$

We now substitute (4.55) into (4.51a) and solve for $\tilde{\omega}$. This yields

$$\tilde{\omega} = \frac{-j2}{\Delta t} \ln\left(j\xi + \sqrt{1 - \xi^2}\right) \tag{4.56a}$$

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

$$\tilde{\omega} = \frac{-j2}{\Delta t} \ln\left[\left(\xi + \sqrt{\xi^2 - 1}\right) e^{+j\pi/2}\right] \tag{4.56b}$$

Upon taking the natural logarithm, we further obtain

$$\tilde{\omega} = \frac{-j2}{\Delta t}\left[\ln\left(\xi + \sqrt{\xi^2 - 1}\right) + j\frac{\pi}{2}\right] = \frac{\pi}{\Delta t} - \frac{j2}{\Delta t}\ln\left(\xi + \sqrt{\xi^2 - 1}\right) \tag{4.56c}$$

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

$$\tilde{\omega}_{\text{real}} = \frac{\pi}{\Delta t}\ ; \qquad \tilde{\omega}_{\text{imag}} = -\frac{2}{\Delta t}\ln\left(\xi + \sqrt{\xi^2 - 1}\right) \tag{4.57}$$

Finally, substituting (4.57) into (4.49), we obtain

$$\begin{aligned} V|_{I,J,K}^{n} &= V_0 e^{2n\ln\left(\xi+\sqrt{\xi^2-1}\right)} e^{j\left[(\pi/\Delta t)(n\Delta t) - \tilde{k}_x I\Delta x - \tilde{k}_y J\Delta y - \tilde{k}_z K\Delta z\right]} \\ &= V_0\left(\xi + \sqrt{\xi^2 - 1}\right)^{**2n} e^{j\left[(\pi/\Delta t)(n\Delta t) - \tilde{k}_x I\Delta x - \tilde{k}_y J\Delta y - \tilde{k}_z K\Delta z\right]} \end{aligned} \tag{4.58}$$

where $**2n$ denotes the $2n$'th power. From (4.58), we define the following multiplicative factor greater than 1 that amplifies the numerical wave every time-step:

$$q_{\text{growth}} \equiv \left(\xi + \sqrt{\xi^2 - 1}\right)^2 \tag{4.59}$$

Equations (4.58) and (4.59) define an exponential growth of the numerical wave with time-step number n. We see that the dominant exponential growth occurs for the most positive possible value of ξ; that is, $\xi_{\text{upper bound}}$ defined in (4.52) for the numerical wavevector components defined in (4.53). This is the origin of numerical instability.

Example of Calculating a Stability Bound

Consider the practical case of a three-dimensional cubic-cell space lattice with $\Delta x = \Delta y = \Delta z = \Delta$. From (4.54b), numerical instability arises when

$$\Delta t > \frac{1}{c\sqrt{\dfrac{1}{(\Delta)^2} + \dfrac{1}{(\Delta)^2} + \dfrac{1}{(\Delta)^2}}} = \frac{1}{c\sqrt{\dfrac{3}{(\Delta)^2}}} = \frac{\Delta}{c\sqrt{3}} \tag{4.60a}$$

We define an equivalent Courant stability limit for the cubic-cell-lattice case:

$$S_{\substack{\text{stability}\\ \text{limit-3D}}} = \frac{1}{\sqrt{3}} \tag{4.60b}$$

From (4.53), the dominant exponential growth is seen to occur for numerical waves propagating along the lattice diagonals. The relevant wavevectors are

$$\tilde{k} = \frac{\pi}{\Delta}(\pm\hat{x} \pm \hat{y} \pm \hat{z}) \quad \rightarrow \quad |\tilde{k}| = \frac{\pi\sqrt{3}}{\Delta} \quad \rightarrow \quad \tilde{\lambda} = \left(\frac{2\sqrt{3}}{3}\right)\Delta \tag{4.61}$$

where $\hat{x}$, $\hat{y}$, and $\hat{z}$ are unit vectors defining the major lattice axes. Further, (4.52) yields

$$\xi_{\substack{\text{upper}\\\text{bound}}} = c\,\Delta t \sqrt{\frac{1}{(\Delta)^2} + \frac{1}{(\Delta)^2} + \frac{1}{(\Delta)^2}} = \left(\frac{c\,\Delta t}{\Delta}\right)\sqrt{3} = S\sqrt{3} \tag{4.62}$$

From (4.59), this implies the following maximum possible growth-factor per time-step under conditions of numerical instability:

$$q_{\text{growth}} \equiv \left[\, S\sqrt{3} + \sqrt{\left(S\sqrt{3}\right)^2 - 1}\,\right]^2 \quad \Bigg\} \text{ for } S \geq \frac{1}{\sqrt{3}} \tag{4.63}$$

Courant Factor Normalization and Extension to Two Dimensions and One Dimension

It is instructive to use the results of (4.60b) to normalize the Courant factor S in (4.63). This will permit us to generalize the three-dimensional results for the maximum growth-factor q to two-dimensional and one-dimensional Yee grids. In this spirit, we define

$$S_{\text{norm-3D}} \equiv \frac{S}{S_{\substack{\text{stability}\\\text{limit-3D}}}} = \frac{S}{\left(1/\sqrt{3}\right)} = S\sqrt{3} \tag{4.64}$$

Then, (4.63) can be written as

$$q_{\text{growth}} = \left[\, S_{\text{norm-3D}} + \sqrt{\left(S_{\text{norm-3D}}\right)^2 - 1}\,\right]^2 \quad \Bigg\} \text{ for } S_{\text{norm-3D}} \geq 1 \tag{4.65}$$

Given this notation, it can be shown that analogous expressions for the Courant stability limit and the growth-factor under conditions of numerical instability are given by:

Two-dimensional square Yee grid:

$$S_{\substack{\text{stability}\\\text{limit-2D}}} = \frac{1}{\sqrt{2}} \tag{4.66a}$$

$$S_{\text{norm-2D}} \equiv \frac{S}{S_{\substack{\text{stability}\\\text{limit-2D}}}} = \frac{S}{\left(1/\sqrt{2}\right)} = S\sqrt{2} \tag{4.66b}$$

Dominant exponential growth occurs for numerical waves propagating along the grid diagonals. The relevant wavevectors are

$$\tilde{k} = \frac{\pi}{\Delta}(\pm\hat{x} \pm \hat{y}) \quad \rightarrow \quad |\tilde{k}| = \frac{\pi\sqrt{2}}{\Delta} \quad \rightarrow \quad \tilde{\lambda} = \sqrt{2}\Delta \tag{4.66c}$$

This yields the following solution growth-factor per time-step:

$$\left.\begin{array}{l} q_{\text{growth}} = \left[S_{\text{norm-2D}} + \sqrt{(S_{\text{norm-2D}})^2 - 1} \right]^2 \\ \end{array}\right\} \text{ for } S_{\text{norm-2D}} \geq 1 \tag{4.66d}$$

One-dimensional uniform Yee grid:

$$S_{\substack{\text{stability} \\ \text{limit-1D}}} = 1 \tag{4.67a}$$

$$S_{\text{norm-1D}} \equiv \frac{S}{S_{\substack{\text{stability} \\ \text{limit-1D}}}} = \frac{S}{1} = S \tag{4.67b}$$

Dominant exponential growth occurs for numerical waves with the wavevectors

$$\tilde{k} = \pm\frac{\pi}{\Delta}\hat{x} \quad \rightarrow \quad |\tilde{k}| = \frac{\pi}{\Delta} \quad \rightarrow \quad \tilde{\lambda} = 2\Delta \tag{4.67c}$$

This yields the following solution growth-factor per time-step:

$$\left.\begin{array}{l} q_{\text{growth}} = \left(S + \sqrt{S^2 - 1} \right)^2 \\ \end{array}\right\} \text{ for } S \geq 1 \tag{4.67d}$$

Interestingly, (4.67d) is *exactly* (2.49), the growth-factor under conditions of numerical instability for the scalar one-dimensional wave equation.

We see from the above discussion that the solution growth-factor q under conditions of numerical instability is the same, regardless of the dimensionality of the FDTD space lattice, if the same normalized Courant number is used. For example, q is identical for the following three cases:

1. $S = 1.0005$ for a uniform, one-dimensional grid;
2. $S = 1.0005 \times (1/\sqrt{2}) = 0.707460$ for a uniform, square, two-dimensional grid;
3. $S = 1.0005 \times (1/\sqrt{3}) = 0.577639$ for a uniform, cubic, three-dimensional grid.

This is illustrated below by way of a sample FDTD calculation.

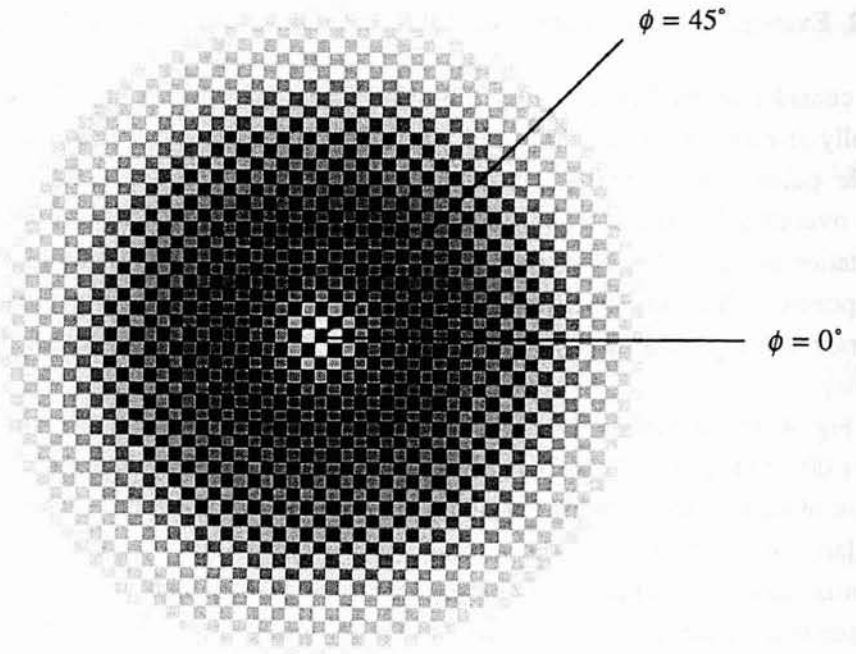
4.7.2 Example of a Numerically Unstable Two-Dimensional FDTD Model

We consider an FDTD modeling example where the Courant stability condition is violated equally at every point in a TM_z grid. To allow direct comparison with a previous example of stable pulse propagation, the same grid discussed in Section 4.6.4 and Fig. 4.5 is used. The overall grid size is again 360×360 square cells, with $\Delta x = \Delta y = 1.0 = \Delta$. Numerical excitation to the grid is again provided by specifying a unit-step time-function for the center E_z component. The only condition that differs from those assumed in Section 4.6.4 is that the Courant stability factor S is increased just above the threshold for numerical instability given by (4.66a).

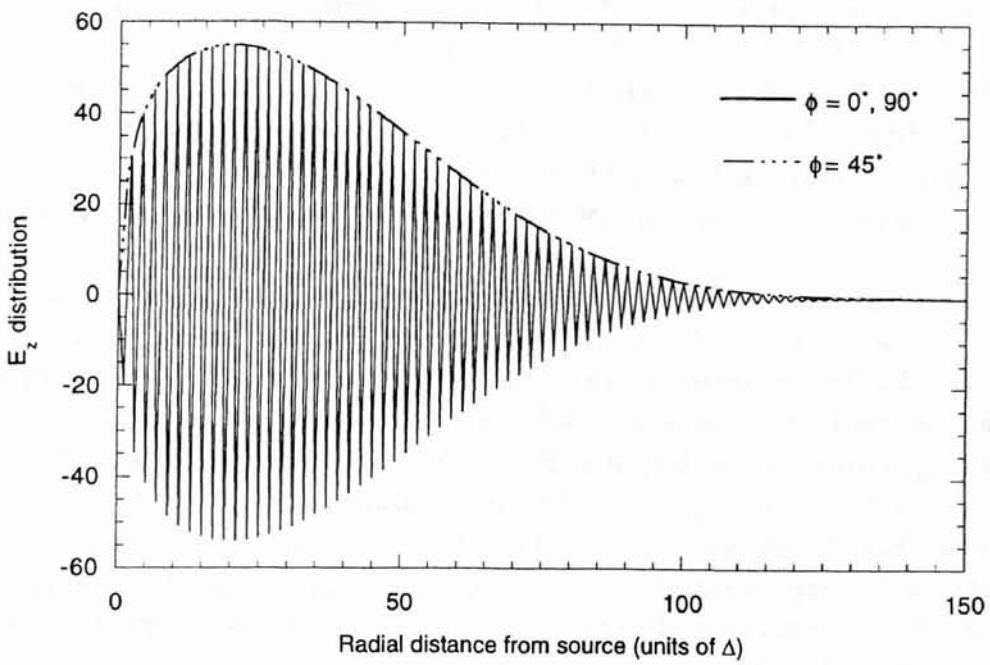
Fig. 4.6(a) visualizes the two-dimensional E_z distribution in space at $n = 40$ time-steps for $S = 1.005 \times (1/\sqrt{2})$. S is sufficiently greater than the instability threshold to quickly generate a region of numerical instability spreading out radially from the source, where the field amplitudes are large enough to mask the normal wave propagation. This permits a high-resolution visualization of individual E_z components in the grid, which are depicted as square pixels. We see that the unstable field pattern has the form of a checkerboard wherein the dark and gray pixels denote positive and negative E_z field values, respectively. Here, the pixel saturation denotes the relative amplitude of its positive or negative value.

Fig. 4.6(b) graphs the variation of E_z versus radial distance from the source at $n = 200$ time-steps for $S = 1.0005 \times (1/\sqrt{2})$. Two distinct plots are shown. The solid line graph exhibits a rapid spatial oscillation with the period 2Δ. This is the E_z behavior along the $\phi = 0°$, $90°$ (and similar on-axis) cuts through the grid. The smooth dashed-dotted curve with no spatial oscillation represents the E_z behavior along the $\phi = 45°$ (and similar oblique) cuts through the grid. Analysis of the underlying data reveals growth factors in the range 1.060 to 1.069 per time-step along the leading edge of the instability region. This agrees very well with $q_{\text{growth}} = 1.0653$ calculated using (4.66d), and is an excellent validation of the Courant-factor-normalization theory.

An interesting observation in Fig. 4.6(b) is that the smooth E_z variation along $\phi = 45°$ forms the envelope of the oscillatory E_z distribution observed along the grid's major axes. This difference in behavior is confirmed in Fig. 4.6(a), which shows that the $\phi = 45°$ cut lies entirely within a diagonal string of dark (positive) pixels, whereas the $\phi = 0°$ cut passes through alternating dark (positive) and gray (negative) pixels. We attribute this behavior to (4.66c), which states that the exponential growth along the grid diagonal has $\tilde{\lambda} = \sqrt{2}\Delta$. That is, the numerical wavelength along the 45° observation cut for the unstable mode is exactly the diagonal length across a single $\Delta \times \Delta$ grid cell. Thus, there exists 2π (or equivalently, 0) phase shift of the unstable mode between adjacent observation points along the $\phi = 45°$ cut. This means that adjacent E_z values along $\phi = 45°$ cannot change sign. In contrast, (4.66c) reduces to $\tilde{k} = \pi/\Delta$ (i.e., $\tilde{\lambda} = 2\Delta$), for the unstable mode along the $\phi = 0°$, $90°$ cuts. Therefore, there is π phase shift of the unstable mode between adjacent observation points along $\phi = 0°$, $90°$; yielding the point-by-point sign reversals (rapid spatial oscillations) seen in Fig. 4.6(b).



(a) Visualization of the two-dimensional E_z distribution at $n = 40$ for $S = 1.005 \times (1/\sqrt{2})$.



(b) Comparison of E_z distributions along the grid axes and along a grid diagonal at $n = 200$ for $S = 1.0005 \times (1/\sqrt{2})$. The theoretical and measured growth-factor is $q_{growth} \equiv 1.065$ per time-step.

Fig. 4.6 Effect of numerical instability upon the two-dimensional pulse-propagation model of Fig. 4.5.

4.7.3 Linear Growth Mode When the Normalized Courant Factor Equals 1

Equation (4.49) postulates a numerical wave amplitude that is either constant, exponentially growing, or exponentially decaying as time-stepping progresses. As shown in [4], an additional *linear* growth mode is possible when the normalized Courant factor equals exactly 1.0. While the amplitude of this mode increases much more slowly than the exponential instability discussed previously, the analyst should proceed with caution when using $S_{\text{norm}} = 1$.

4.8 GENERALIZED STABILITY PROBLEM

The stability of the overall FDTD solution procedure for Maxwell's equations depends upon more than the stability of the Yee algorithm. In fact, a *generalized stability problem* arises due to interactions between the Yee algorithm and any augmenting algorithms used to model boundary conditions, variable meshing, and complex materials. Factors involved in the generalized stability problem are now discussed.

4.8.1 Absorbing and Impedance Boundary Conditions

Numerical realizations of electromagnetic field boundary conditions that require the processing of field data located nonlocally in space or time can lead to instability of the overall time-stepping algorithm. An important example of this possibility arises when implementing ABCs at the outermost space-lattice planes to model scattering or radiation phenomena in unbounded regions. As discussed in detail in Chapters 6 and 7, ABCs have been the subject of much research, with several distinct physics modeling approaches and numerical implementations emphasized by the FDTD research community.

The nature of the numerical stability problem here is exemplified by one of the most popular ABCs of the early 1990s, the Liao ABC (Section 6.5). In augmentation of the Yee algorithm, Liao et al. postulated a polynomial extrapolation of field data at interior gridpoints and past time-steps to the desired outer-boundary gridpoint at the latest time-step. However, the Liao ABC was found by later workers to be marginally stable. Similar issues had previously arisen with regard to other ABCs, including those of Engquist-Majda (Section 6.3) and Higdon (Section 6.4). More recently, PML ABCs (Chapter 7) have come under scrutiny for potential numerical instability due to their novel formulations.

Overall, the experience of workers in this area is that ABC numerical stability is maintained for many thousands of iterations, if not indefinitely, with the proper choice of time-step. This has permitted the ABCs of Chapters 6 and 7 to be used successfully in many different engineering simulations. A similar experience base has been established for the numerical stability of dispersive impedance boundary conditions (Section 10.8).

4.8.2 Variable and Unstructured Meshing

The analysis of numerical instability can become complicated when the FDTD space lattice is generated to conformally fit a specific structure by varying the size, position, and shape of the lattice cells, rather than using the uniform "bricks" postulated by Yee. Three varieties of such meshes are discussed in detail in Chapters 10, 11, and 12. Groups working in this area have

found that even if the mesh construction is so complex that an exact stability criterion cannot be derived, a part-analytical / part-empirical upper bound on the time-step can be derived for each gridding approach so that numerical stability is maintained for many thousands of time-steps, if not indefinitely. This has permitted numerous successful engineering applications for non-Cartesian and unstructured FDTD meshes. See Chapters 11 and 12 for specific numerical stability conditions for such meshes.

4.8.3 Lossy, Dispersive, Nonlinear, and Gain Materials

Much literature has emerged concerning FDTD modeling of dispersive and nonlinear materials, and the reader is referred to Chapter 9 for a detailed discussion of the primary algorithms in this area. For linear-dispersion algorithms, it is usually possible to derive precise bounds on numerical stability. However, stability analysis may not be feasible for models of nonlinear dispersive materials. Fortunately, substantial modeling experience has shown that numerical stability can be maintained for thousands of time-steps, if not indefinitely, for linear, nonlinear, and gain materials with a properly chosen time-step. Again, this has permitted numerous successful engineering applications.

4.9 MODIFIED YEE-BASED ALGORITHMS FOR MITIGATING NUMERICAL DISPERSION

The numerical algorithm for Maxwell's equations introduced by Yee in 1966 is very robust. Evidence of this claim is provided by the existence of thousands of successful electromagnetic engineering applications, refereed journal papers, and conference presentations derived from the basic Yee algorithm over the past 40 years. However, it is clear from our previous discussions that Yee's approach is not perfect. For certain modeling problems, numerical dispersion can cause significant errors to arise in the calculated field.

This section reviews a small set of representative strategies aimed at mitigating the effects of numerical dispersion in Yee-based algorithms for Maxwell's equations. No attempt is made to provide a comprehensive summary because such an effort would require several chapters. The intent here is to provide the flavor of what may be possible in this area. Clearly, this area will evolve as our understanding of the underlying physical and numerical issues improves.

4.9.1 Strategy 1: Center a Specific Numerical Phase-Velocity Curve About *c*

We have seen from Fig. 4.2 that the Yee algorithm yields a family of numerical phase-velocity curves contained entirely in the range $\tilde{v}_p < c$. We also observe that each velocity curve is centered about the value

$$\tilde{v}_{\text{avg}} = \frac{\tilde{v}_p(\phi = 0^\circ) + \tilde{v}_p(\phi = 45^\circ)}{2} \tag{4.68}$$

where $\tilde{v}_p(\phi = 0^\circ)$ and $\tilde{v}_p(\phi = 45^\circ)$ are given by (4.14b) and (4.15b), respectively. This symmetry can be exploited if a narrowband grid excitation is used, such that a specific phase-velocity curve accurately characterizes the propagation of most of the numerical modes in the grid. Then, it is

possible to shift the phase-velocity curve of interest so that it is centered about the free-space speed of light c, thereby cutting $\Delta\tilde{v}_{\text{physical}}$ by almost 3:1. Centering is implemented by simply scaling the free-space values of ε_0 and μ_0 used in the finite-difference system of (4.2):

$$\varepsilon_0' = \left(\frac{\tilde{v}_{\text{avg}}}{c}\right)\varepsilon_0 \quad ; \qquad \mu_0' = \left(\frac{\tilde{v}_{\text{avg}}}{c}\right)\mu_0 \tag{4.69}$$

This scaling increases the baseline value of the model's "free-space" speed of light to compensate for the too-slow value of $\tilde{v}_{\text{avg}}$. By scaling both ε_0 and μ_0, the required shift in $\tilde{v}_{\text{avg}}$ is achieved without introducing any changes in wave impedance.

There are four difficulties with this approach: (1) The phase-velocity anisotropy error $\Delta\tilde{v}_{\text{aniso}}$ *remains unmitigated.* (2) The velocity compensation is only in the average sense over all possible directions in the grid. Hence, important numerical modes can still have phase velocities not equal to c. (3) The increased free-space speed of light requires Δt to be proportionately reduced below the Courant limit to ensure numerical stability. (4) Propagating wave pulses having broad spectral content cannot be compensated over their entire frequency range. Nevertheless, this approach is so easy to implement that its use can be almost routine.

4.9.2 Strategy 2: Use Fourth-Order-Accurate Explicit Spatial Differences

It is possible to substantially reduce the phase-velocity anisotropy error $\Delta\tilde{v}_{\text{aniso}}$ for the Yee algorithm by incorporating a fourth-order-accurate finite-difference scheme for the spatial first-derivatives needed to implement the curl operator. This section reviews an explicit method wherein a fourth-order-accurate spatial central-difference is calculated from a stencil of four field values located at distances of $\Delta/2$ and $3\Delta/2$ on each side of the observation point [5].

Basic Method

The basic fourth-order-accurate spatial-difference scheme [5] can be derived either by: (1) using a Taylor's series approach similar to that of Chapter 2, Section 2.4; or (2) fitting a cubic polynomial to the nearest four field values centered about the observation point, and then analytically evaluating the first derivative of this polynomial at the observation point. Assuming that standard Yee leapfrog time-stepping is used, either approach results in the following set of finite-difference expressions for the TM_z mode analogous to (4.2a-c):

$$\frac{H_x\big|_{i,\,j+1/2}^{n+1/2} - H_x\big|_{i,\,j+1/2}^{n-1/2}}{\Delta t} = -\frac{1}{\mu_{i,\,j+1/2}}\left(\frac{-E_z\big|_{i,\,j+2}^{n} + 27E_z\big|_{i,\,j+1}^{n} - 27E_z\big|_{i,\,j}^{n} + E_z\big|_{i,\,j-1}^{n}}{24\Delta y}\right) \tag{4.70a}$$

$$\frac{H_y\big|_{i+1/2,\,j}^{n+1/2} - H_y\big|_{i+1/2,\,j}^{n-1/2}}{\Delta t} = \frac{1}{\mu_{i+1/2,\,j}}\left(\frac{-E_z\big|_{i+2,\,j}^{n} + 27E_z\big|_{i+1,\,j}^{n} - 27E_z\big|_{i,\,j}^{n} + E_z\big|_{i-1,\,j}^{n}}{24\Delta x}\right) \tag{4.70b}$$

$$\frac{E_z\big|_{i,j}^{n+1} - E_z\big|_{i,j}^{n}}{\Delta t} = \frac{1}{\varepsilon_{i,j}} \left(\begin{array}{l} \dfrac{-H_y\big|_{i+3/2,j}^{n+1/2} + 27H_y\big|_{i+1/2,j}^{n+1/2} - 27H_y\big|_{i-1/2,j}^{n+1/2} + H_y\big|_{i-3/2,j}^{n+1/2}}{24\Delta x} - \\ \dfrac{-H_x\big|_{i,j+3/2}^{n+1/2} + 27H_x\big|_{i,j+1/2}^{n+1/2} - 27H_x\big|_{i,j-1/2}^{n+1/2} + H_x\big|_{i,j-3/2}^{n+1/2}}{24\Delta y} \end{array} \right) \tag{4.70c}$$

The numerical dispersion relation for this algorithm analogous to (4.5) is given by

$$\begin{aligned} \left[\frac{1}{c\Delta t}\sin\left(\frac{\omega\Delta t}{2}\right)\right]^2 = {} & \frac{1}{(\Delta x)^2}\left[\frac{27}{24}\sin\left(\frac{\tilde{k}_x\Delta x}{2}\right) - \frac{1}{24}\sin\left(\frac{3\tilde{k}_x\Delta x}{2}\right)\right]^2 \\ & + \frac{1}{(\Delta y)^2}\left[\frac{27}{24}\sin\left(\frac{\tilde{k}_y\Delta y}{2}\right) - \frac{1}{24}\sin\left(\frac{3\tilde{k}_y\Delta y}{2}\right)\right]^2 \end{aligned} \tag{4.71}$$

By analogy with the development in Section 4.5.2 leading to (4.27), it can be shown that the intrinsic numerical phase-velocity anisotropy for a square-cell grid of this type is given by

$$\frac{c^*}{c} \cong \frac{N_\lambda}{\pi}\sqrt{\begin{array}{l} \left[\dfrac{27}{24}\sin\left(\dfrac{\pi\cos\phi}{N_\lambda}\right) - \dfrac{1}{24}\sin\left(\dfrac{3\pi\cos\phi}{N_\lambda}\right)\right]^2 + \\ \left[\dfrac{27}{24}\sin\left(\dfrac{\pi\sin\phi}{N_\lambda}\right) - \dfrac{1}{24}\sin\left(\dfrac{3\pi\sin\phi}{N_\lambda}\right)\right]^2 \end{array}} \tag{4.72}$$

and the numerical phase-velocity anisotropy error [by analogy with (4.28)] is given by

$$\Delta\tilde{v}_{\text{aniso}}\big|_{\text{explicit 4th-order}} \cong \frac{\pi^4}{18(N_\lambda)^4} \times 100\% \tag{4.73}$$

Finally, we can apply a complex-frequency analysis analogous to that discussed in Section 4.7.1 to numerical dispersion relation (4.71). Solving (4.71) for $\tilde{\omega}$ yields an expression comparable to (4.51a) and (4.51b):

$$\tilde{\omega} = \frac{2}{\Delta t}\sin^{-1}(\xi) \tag{4.74a}$$

where

$$\xi = c\Delta t \sqrt{\frac{1}{(\Delta x)^2}\left[\frac{27}{24}\sin\left(\frac{\tilde{k}_x \Delta x}{2}\right) - \frac{1}{24}\sin\left(\frac{3\tilde{k}_x \Delta x}{2}\right)\right]^2 + \frac{1}{(\Delta y)^2}\left[\frac{27}{24}\sin\left(\frac{\tilde{k}_y \Delta y}{2}\right) - \frac{1}{24}\sin\left(\frac{3\tilde{k}_y \Delta y}{2}\right)\right]^2} \tag{4.74b}$$

ξ reaches its maximum possible value when all of the sin terms in the square root reinforce, yielding an expression analogous to (4.52):

$$\xi_{\text{upper bound}} = c\Delta t \sqrt{\frac{1}{(\Delta x)^2}\left(\frac{27}{24} + \frac{1}{24}\right)^2 + \frac{1}{(\Delta y)^2}\left(\frac{27}{24} + \frac{1}{24}\right)^2}$$

$$= \left(\frac{28}{24}\right) c\Delta t \sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2}} \tag{4.75}$$

The $\sin^{-1}$ function, and thus $\tilde{\omega}$, becomes complex-valued if $\xi_{\text{upper bound}} > 1$. As discussed in Section 4.7.1, this causes numerical instability through exponential growth of the solution. Hence, the numerically stable range is defined by requiring that $\xi_{\text{upper bound}} \leq 1$:

$$\left(\frac{7}{6}\right) c\Delta t \sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2}} \leq 1 \tag{4.76}$$

yielding immediately

$$\Delta t\Big|_{\substack{\text{explicit} \\ \text{4th-order}}} \leq \frac{(6/7)}{c\sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2}}} \quad : \text{two-dimensions} \tag{4.77}$$

as the maximum time-step for numerical stability. The stability bounds for the one- and three-dimensional versions of the algorithm can now be written by inspection from (4.74) and (4.75):

$$\Delta t\Big|_{\substack{\text{explicit} \\ \text{4th-order}}} \leq \left(\frac{6}{7}\right)\frac{\Delta x}{c} \quad : \text{one-dimension} \tag{4.78}$$

$$\Delta t\Big|_{\substack{\text{explicit} \\ \text{4th-order}}} \leq \frac{(6/7)}{c\sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2} + \frac{1}{(\Delta z)^2}}} \quad : \text{three-dimensions} \tag{4.79}$$

In each case, the maximum Δt permitted for algorithm stability is 6/7 of the corresponding limit for the classic Yee algorithm using second-order space differences.

For a uniform square-cell space lattice, the Courant factor $S = c\Delta t/\Delta$ can be defined. This permits the following Courant factor normalization analogous to (4.64), (4.66b), and (4.67b):

$$S_{\text{norm-3D}} = \frac{S}{S_{\substack{\text{stability} \\ \text{limit-3D}}}} = \frac{S}{\left(\frac{6}{7} \cdot \frac{1}{\sqrt{3}}\right)} = \frac{7}{6} S\sqrt{3} \tag{4.80a}$$

$$S_{\text{norm-2D}} = \frac{S}{S_{\substack{\text{stability} \\ \text{limit-2D}}}} = \frac{S}{\left(\frac{6}{7} \cdot \frac{1}{\sqrt{2}}\right)} = \frac{7}{6} S\sqrt{2} \tag{4.80b}$$

$$S_{\text{norm-1D}} = \frac{S}{S_{\substack{\text{stability} \\ \text{limit-1D}}}} = \frac{S}{\left(\frac{6}{7} \cdot 1\right)} = \frac{7}{6} S \tag{4.80c}$$

This permits the per-time-step solution growth-factor under conditions of numerical instability to be defined exactly as in (4.65), (4.66d), and (4.67d):

$$q_{\text{growth}} = \left[S_{\text{norm}} + \sqrt{(S_{\text{norm}})^2 - 1} \right]^2 \quad \Big\} \text{ for } S_{\text{norm}} \geq 1 \tag{4.81}$$

Equations (4.80a-c) and (4.81) comprise a unified treatment of the exponential growth under unstable conditions for the fourth-order space-derivative algorithm applied to one-, two-, and three-dimensional space lattices.

Optimized Method

Reference [6] reported a means to optimize the spatial differences of (4.70) to reduce the numerical dispersion error of the fourth-order accurate scheme discussed above by a factor of better than 4:1 for $N_\lambda > 10$. Here, a four-point central-difference operator for the first derivative of a field quantity f at grid point i is given by the following general form:

$$\left.\frac{\partial f}{\partial w}\right|_i = \frac{C_{w1}\left(f_{i+1/2} - f_{i-1/2}\right) + C_{w2}\left(f_{i+3/2} - f_{i-3/2}\right)}{\Delta w} \tag{4.82}$$

where w denotes either x or y for two-dimensional problems, and Δw is the corresponding space increment in the x or y directions. Assuming usage of standard Yee leapfrog time-stepping, a dispersion analysis of the type considered in Section 4.2 is implemented. The resulting dispersion relation is expanded in a Taylor series in both the space and time increments (retaining terms up to order 2), and the ideal condition $\tilde{\boldsymbol{k}} = \boldsymbol{k}$ is applied to this expansion. Coefficients C_{w1} and C_{w2} are then selected to satisfy the resulting system of equations:

$$C_{x1} + 3C_{x2} = 1 \tag{4.83a}$$

$$C_{y1} + 3C_{y2} = 1 \tag{4.83b}$$

$$C_{x1} + 27C_{x2} + \left(\frac{\Delta y}{\Delta x}\right)^2 \left(C_{y1} + 27C_{y2}\right) = \frac{8}{3}\left(\frac{c\Delta t}{\Delta x}\right)^2 \tag{4.83c}$$

$$C_{x1} + 27C_{x2} - \left(\frac{\Delta y}{\Delta x}\right)^2 \left(C_{y1} + 27C_{y2}\right) = 0 \tag{4.83d}$$

For the special case of square grid cells ($\Delta x = \Delta y = \Delta$), the system of equations given by (4.83) reduces to

$$\begin{aligned} C_1 + 3C_2 &= 1 \\ C_1 + 27C_2 &= \frac{4}{3}\left(\frac{c\Delta t}{\Delta x}\right)^2 \end{aligned} \tag{4.84}$$

where the x and y subscripts have been dropped. For a three-dimensional space lattice having cubic cells, [6] reported the following system of equations for optimally selecting C_1 and C_2:

$$\begin{aligned} C_1 + 3C_2 &= 1 \\ C_1 + 27C_2 &= \frac{5}{3}\left(\frac{c\Delta t}{\Delta x}\right)^2 \end{aligned} \tag{4.85}$$

We note that the significant reduction in numerical dispersion achieved by implementing (4.83), (4.84), or (4.85) has no impact upon the required computer storage or number of arithmetic operations relative to the basic fourth-order accurate algorithm of [5] discussed previously. Further, there is little perturbation of the numerical instability threshold.

Example of Improved Accuracy

Fig. 4.7 compares the accuracy of the fourth-order-accurate spatial-differencing method reported in [7] with that of the classic Yee algorithm for a generic two-dimensional wave-propagation problem — a sinusoidal line source radiating in free space after being switched on at $t = 0$. The fourth-order spatial-differencing technique is optimized for a square-cell grid of sampling density $N_\lambda = 5$, whereas the Yee grid has a much finer spatial resolution of $N_\lambda = 40$. We see that, despite its much coarser grid, the fourth-order technique achieves accuracy (in the L_2-normed sense relative to the exact solution) comparable to that of the Yee algorithm. This results in a measured savings of $(40/5)^2$:1 = 8^2:1 = 64:1 in computer storage, and 23:1 in running-time. The advantage in computer storage for the fourth-order technique is expected to scale to the order of 8^3:1 = 512:1 in three dimensions.

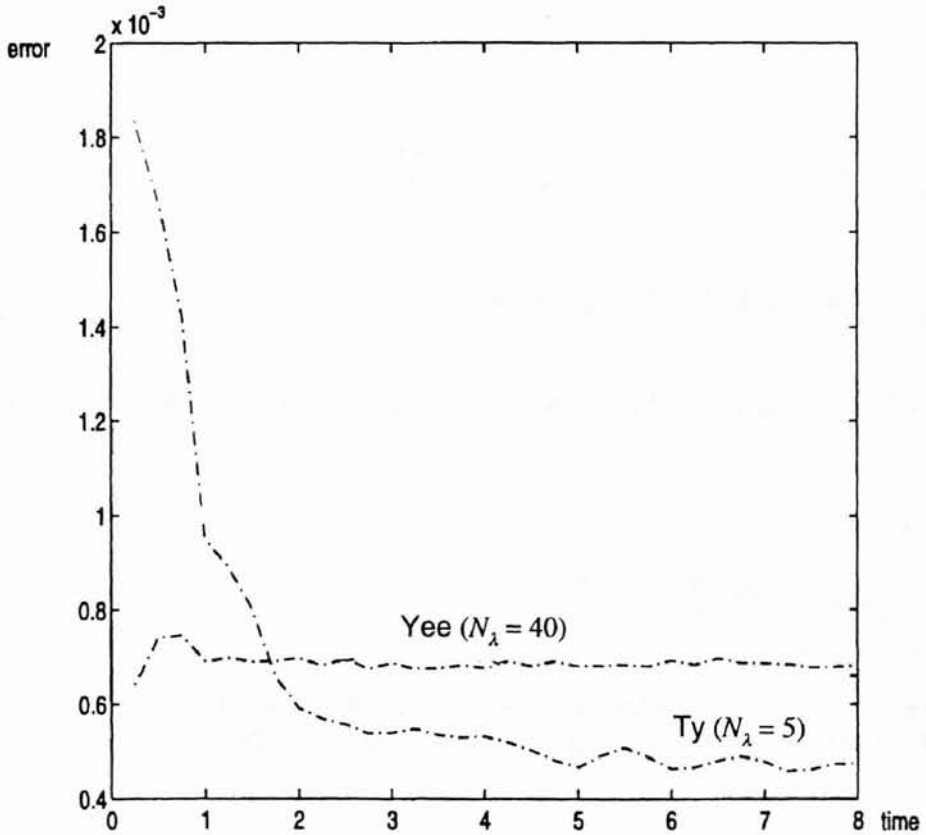


Fig. 4.7 Comparison of the L_2-normed errors in modeling a radially propagating sinusoidal wave (as a function of the simulated time) for a high-resolution ($N_\lambda = 40$) classic Yee grid and a low-resolution ($N_\lambda = 5$) grid employing Ty fourth-order-accurate spatial differences. *Source:* E. Turkel, Chapter 2 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, A. Taflove, (ed.), © 1998 Artech House, Inc.

Discussion

From a growing set of published results similar to those of Fig. 4.7, it appears that fourth-order-accurate spatial schemes allow modeling electromagnetic wave-propagation and interaction problems that are about eight times the electrical size of those permitted by the classic Yee algorithm. This is a very worthwhile increase in efficiency and modeling capability.

However, this improvement is not without cost. Although easy to set up in homogeneous-material regions, the larger stencil needed to calculate fourth-order spatial differences is troublesome when dealing with material interfaces. Metal boundaries are especially challenging, since they effectively cause field discontinuities in the grid. Special boundary conditions required for such interfaces significantly complicate the computer software used to render structures in the grid.

Fourth-Order Smooth Approximation of Jumps in Material Parameters at Interfaces

As stated above, special boundary conditions must be derived and programmed to deal with material discontinuities (especially metal boundaries) when implementing high-order-accuracy finite-difference approximations of spatial derivatives. This is because the nonlocal nature of the numerical space-differentiation process (i.e., the enlarged required stencil of field components) may convey electromagnetic field data across such discontinuities in a nonphysical manner.

Reference [7] reported a means to markedly reduce error due to abrupt dielectric interfaces. This approach replaces the discontinuous permittivity function ε by a fourth-order-accurate smooth implicit approximation. (A similar strategy can be applied to jumps in μ.) Relative to the use of a polynomial approximation to ε, this strategy avoids the overshoot artifact. We note that, with an implicit interpolation, ε varies in the entire domain and not just near the interface. However, far from the interface, the variation is small.

Consider a dielectric interface separating two regions defined along the x-axis of the space lattice. Following (2.82) of [7], a fourth-order-accurate interpolation of the permittivity distribution with grid position i can be achieved using

$$\frac{1}{8}\begin{bmatrix} 10 & -5 & 4 & -1 & . & . & 0 \\ 1 & 6 & 1 & 0 & . & . & 0 \\ 0 & 1 & 6 & 1 & 0 & . & 0 \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & 1 & 6 & 1 \\ 0 & . & . & -1 & 4 & -5 & 10 \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_{p-1} \end{bmatrix} = \frac{1}{2}\left(\begin{bmatrix} \varepsilon_{3/2} \\ \varepsilon_{5/2} \\ . \\ \varepsilon_{p-3/2} \\ \varepsilon_{p-1/2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1/2} \\ \varepsilon_{3/2} \\ . \\ \varepsilon_{p-5/2} \\ \varepsilon_{p-3/2} \end{bmatrix} \right) \tag{4.86}$$

Here, $[\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{p-1}]$ is the initially unknown set of values of the smooth approximation to the abrupt dielectric interface; and $[\varepsilon_{1/2}, \varepsilon_{3/2}, \ldots, \varepsilon_{p-1/2}]$ is the known set of permittivities for the original dielectric interface geometry. Inversion of the linear system of (4.86) yields the desired smooth approximation, $[\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{p-1}]$.

We now consider an example of the use of this dielectric interface smoothing technique [7]. This example involves the error in calculating the sinusoidal standing wave within a two-dimensional rectangular cavity comprised of a block of lossless dielectric ($\varepsilon_r = 4$) surrounded by free space. Here, both the Yee and the fourth-order-accurate spatial-differencing grids use square unit cells, wherein $N_\lambda = 30$ within the dielectric. An available analytical (exact) solution is used to specify the L_2-normed errors for both the Yee and fourth-order-accurate spatial-differencing simulations.

Fig. 4.8 shows that the error generated by the fourth-order-accurate spatial-differencing technique in combination with the fourth-order dielectric interface smoothing provided by (4.86) is more than one order-of-magnitude below the error of the Yee algorithm used with the same smoothing. In additional studies in [7], it is demonstrated that the Yee error can be reduced to that of the fourth-order-accurate spatial differencing / interface smoothing by increasing the Yee grid sampling density eightfold, just what was observed for the free-space propagation example discussed previously in the context of Fig. 4.7. Apparently, fourth-order-accurate smoothing of abrupt permittivity jumps succeeds in preserving the accuracy advantage of the fourth-order-accurate spatial-differencing scheme that is observed for the homogeneous-permittivity case.

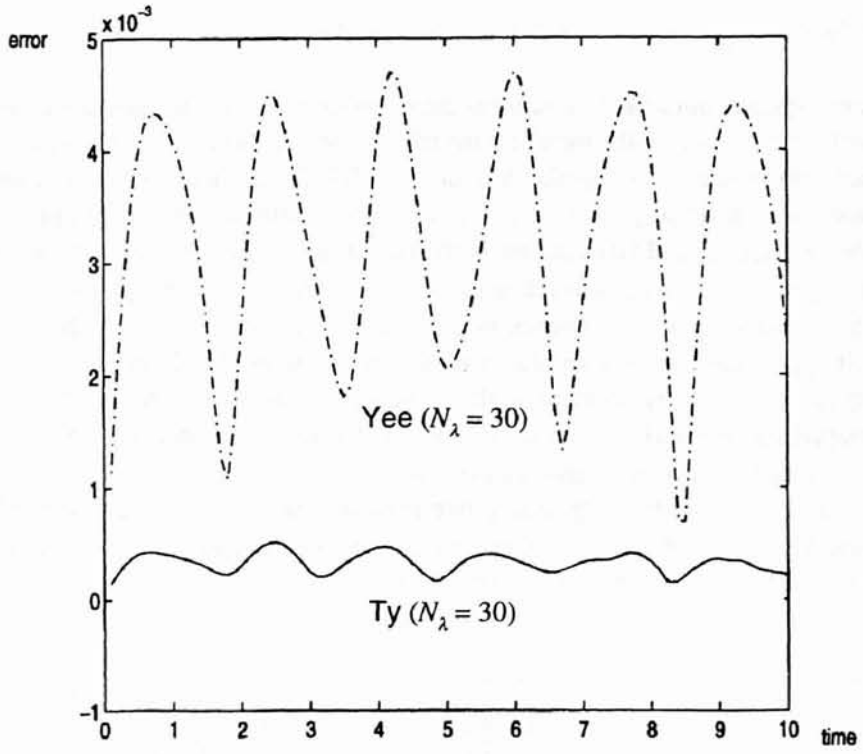


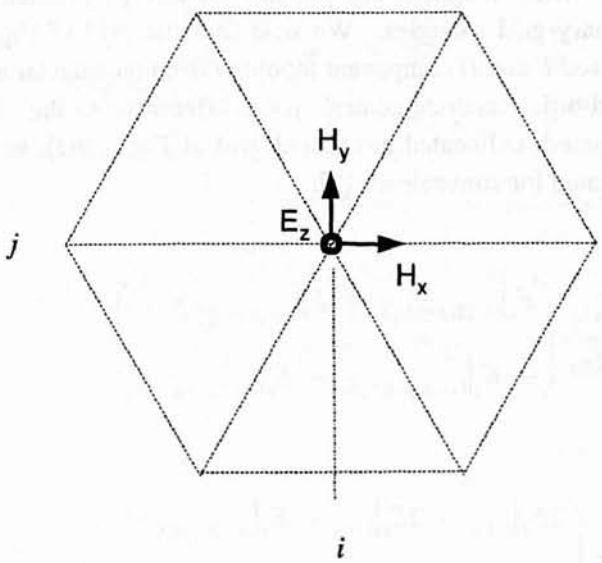
Fig. 4.8 Comparison of the L_2-normed errors in modeling the sinusoidal standing-wave fields within a rectangular dielectric cavity (as a function of the simulated time) for a high-resolution ($N_\lambda = 30$) classic Yee grid and a high-resolution ($N_\lambda = 30$) grid employing Ty fourth-order-accurate spatial differences. Both models use fourth-order-accurate smoothing of the permittivity at the dielectric interfaces as per (4.86). *Source:* E. Turkel, Chapter 2 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, A. Taflove, (ed.), © 1998 Artech House, Inc.

4.9.3 Strategy 3: Use a Hexagonal Grid, If Possible

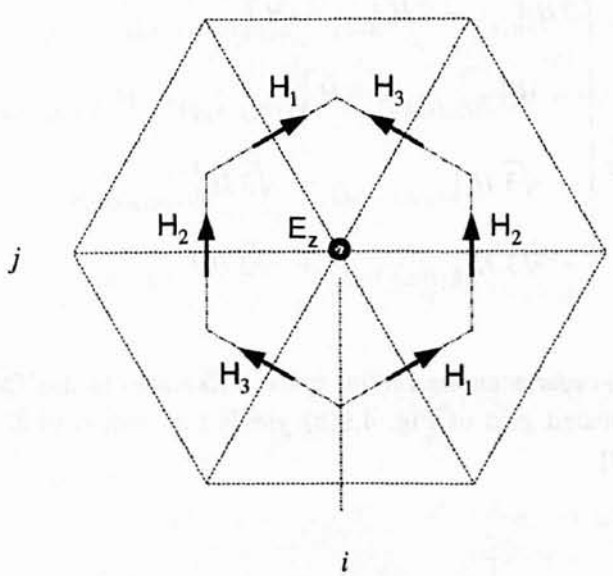
Regular hexagonal grids in two dimensions have been proposed to reduce the numerical phase-velocity anisotropy well below that of Yee's Cartesian mesh. Here, the primary grid is composed of equilateral hexagons having edge length Δs. Each hexagon can be considered to be the union of six equilateral triangles.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

Basic Approaches

Fig. 4.9, which is Fig. 3.5 repeated for convenience, illustrates for the TM_z case in two dimensions the two principal ways of arranging $\boldsymbol{E}$ and $\boldsymbol{H}$ vector components about hexagonal grids [2]. Fig. 4.9(a) shows the unstaggered, collocated grid in which Cartesian E_z, H_x, and H_y field components are collocated at the vertices of the equilateral triangles. No dual-grid is used. Fig. 4.9(b) shows the field arrangement for the staggered, uncollocated grid and its associated dual-grid, the latter indicated by the dashed line segments.



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



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

Fig. 4.9 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.

In Fig. 4.9(b), only E_z components are defined at the vertices of the equilateral triangles, which are exactly the centroids of the hexagonal faces of the dual-grid. Magnetic field components H_1, H_2, H_3, and so forth, 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. 4.9(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 (4.1a-c) for the unstaggered, collocated hexagonal grid of Fig. 4.9(a), we obtain the system (4.87), which is (3.53) repeated for convenience [2]:

$$\frac{\partial H_x|_{i,j}}{\partial t} = -\frac{\sqrt{3}}{\mu_{i,j}\,6\Delta s}\begin{pmatrix} E_z|_{i-1/2,\,j+0.5\sqrt{3}} + E_z|_{i+1/2,\,j+0.5\sqrt{3}} \\ - E_z|_{i-1/2,\,j-0.5\sqrt{3}} - E_z|_{i+1/2,\,j-0.5\sqrt{3}} \end{pmatrix} \tag{4.87a}$$

$$\frac{\partial H_y|_{i,j}}{\partial t} = \frac{1}{\mu_{i,j}\,6\Delta s}\begin{pmatrix} 2E_z|_{i+1,\,j} - 2E_z|_{i-1,\,j} + E_z|_{i+1/2,\,j+0.5\sqrt{3}} \\ - E_z|_{i-1/2,\,j+0.5\sqrt{3}} + E_z|_{i+1/2,\,j-0.5\sqrt{3}} - E_z|_{i-1/2,\,j-0.5\sqrt{3}} \end{pmatrix} \tag{4.87b}$$

$$\frac{\partial E_z|_{i,j}}{\partial t} = \frac{1}{\varepsilon_{i,j}\,6\Delta s}\begin{pmatrix} 2H_y|_{i+1,\,j} - 2H_y|_{i-1,\,j} + H_y|_{i+1/2,\,j+0.5\sqrt{3}} \\ - H_y|_{i-1/2,\,j+0.5\sqrt{3}} + H_y|_{i+1/2,\,j-0.5\sqrt{3}} - H_y|_{i-1/2,\,j-0.5\sqrt{3}} \\ - \sqrt{3}\,H_x|_{i+1/2,\,j+0.5\sqrt{3}} + \sqrt{3}\,H_x|_{i+1/2,\,j-0.5\sqrt{3}} \\ - \sqrt{3}\,H_x|_{i-1/2,\,j+0.5\sqrt{3}} + \sqrt{3}\,H_x|_{i-1/2,\,j-0.5\sqrt{3}} \end{pmatrix} \tag{4.87c}$$

Similarly, applying second-order-accurate central space differences to the TM_z mode equations for the staggered, uncollocated grid of Fig. 4.9(b) yields the system (4.88), which is (3.54) repeated for convenience [2]:

$$\frac{\partial H_1|_{i+1/4,\,j-0.25\sqrt{3}}}{\partial t} = \frac{1}{\mu_{i+1/4,\,j-0.25\sqrt{3}}\,\Delta s}\left(E_z|_{i+1/2,\,j-0.5\sqrt{3}} - E_z|_{i,j}\right) \tag{4.88a}$$

$$\frac{\partial H_2|_{i+1/2,\,j}}{\partial t} = \frac{1}{\mu_{i+1/2,\,j}\,\Delta s}\left(E_z|_{i+1,\,j} - E_z|_{i,j}\right) \tag{4.88b}$$

$$\frac{\partial H_3\big|_{i+1/4,\,j+0.25\sqrt{3}}}{\partial t} = \frac{1}{\mu_{i+1/4,\,j+0.25\sqrt{3}}\,\Delta s}\left(E_z\big|_{i+1/2,\,j+0.5\sqrt{3}} - E_z\big|_{i,\,j}\right) \tag{4.88c}$$

$$\frac{\partial E_z\big|_{i,\,j}}{\partial t} = \frac{2}{\varepsilon_{i,\,j}\,3\Delta s}\cdot\begin{pmatrix} H_1\big|_{i+1/4,\,j-0.25\sqrt{3}} + H_2\big|_{i+1/2,\,j} + H_3\big|_{i+1/4,\,j+0.25\sqrt{3}} - \\ H_1\big|_{i-1/4,\,j+0.25\sqrt{3}} - H_2\big|_{i-1/2,\,j} - H_3\big|_{i-1/4,\,j-0.25\sqrt{3}} \end{pmatrix} \tag{4.88d}$$

We note that the total number of field unknowns for the staggered, uncollocated grid of Fig. 4.9(b) is 33% more than that for the unstaggered, collocated grid of Fig. 4.9(a). However, the discretization of Fig. 4.9(b) is simpler, and the number of total operations is less by about 50%.

Using the analysis method of Section 4.5.2, the following expressions are obtained for the numerical phase-velocity anisotropy error for the gridding cases of Fig. 4.9 [2]:

$$\Delta\tilde{v}_{\text{aniso}}\big|_{\substack{\text{hex. grid,}\\ \text{Fig. 4.9(a)}}} \cong \frac{1\cdot 2\cdot\pi^4}{120\left(N_\lambda\right)^4}\times 100\% = \frac{\pi^4}{60\left(N_\lambda\right)^4}\times 100\% \tag{4.89}$$

$$\Delta\tilde{v}_{\text{aniso}}\big|_{\substack{\text{hex. grid,}\\ \text{Fig. 4.9(b)}}} \cong \frac{1\cdot 2\cdot\pi^4}{720\left(N_\lambda\right)^4}\times 100\% = \frac{\pi^4}{360\left(N_\lambda\right)^4}\times 100\% \tag{4.90}$$

Interestingly, we note that $\Delta\tilde{v}_{\text{aniso}}$ for both hexagonal grids exhibits a *fourth-order* dependence on the grid-sampling density N_λ despite the second-order accuracy of each spatial difference used. This is because the leading second-order error term becomes isotropic for the hexagonal gridding case, with a value exactly equal to the average of its ϕ-dependent Cartesian counterpart [2].

Discussion

Comparison of $\Delta\tilde{v}_{\text{aniso}}$ of the Yee algorithm given by (4.28) with $\Delta\tilde{v}_{\text{aniso}}$ of the hexagonal gridding given by (4.89) and (4.90) yields the following error ratios:

$$\frac{\Delta\tilde{v}_{\text{aniso}}\big|_{\text{hex. grid, Fig. 4.9(a)}}}{\Delta\tilde{v}_{\text{aniso}}\big|_{\text{Yee}}}(N_\lambda) \cong \frac{\pi^2}{5\left(N_\lambda\right)^2}\ ;\quad \frac{\Delta\tilde{v}_{\text{aniso}}\big|_{\text{hex. grid, Fig. 4.9(b)}}}{\Delta\tilde{v}_{\text{aniso}}\big|_{\text{Yee}}}(N_\lambda) \cong \frac{\pi^2}{30\left(N_\lambda\right)^2} \tag{4.91}$$

We note that these are approximate ratios that have increasing validity for $N_\lambda > 10$. We can also compare the velocity-anisotropy errors of the hexagonal gridding with that of the basic explicit fourth-order gridding method given by (4.73):

$$\frac{\Delta\tilde{v}_{\text{aniso}}\big|_{\text{hex. grid, Fig. 4.9(a)}}}{\Delta\tilde{v}_{\text{aniso}}\big|_{\text{explicit 4th-order}}}(N_\lambda) \cong \frac{3}{10}\ ;\quad \frac{\Delta\tilde{v}_{\text{aniso}}\big|_{\text{hex. grid, Fig. 4.9(b)}}}{\Delta\tilde{v}_{\text{aniso}}\big|_{\text{explicit 4th-order}}}(N_\lambda) \cong \frac{1}{20} \tag{4.92}$$

Note that the ratios of (4.92) are independent of the grid-sampling density N_λ.

It appears that the hexagonal gridding of Fig. 4.9(b) can attain an intrinsic grid dispersion that is significantly better than that achieved by the fourth-order-accurate Cartesian spatial-differencing algorithms discussed in Section 4.9.2. Thus, based upon the previous studies, we expect the hexagonal gridding of Fig. 4.9(b) to provide at least an 8^2:1 reduction in computer storage relative to the Yee algorithm for electrically large two-dimensional modeling problems. In addition, hexagonal gridding uses only nearest-neighbor field data, thereby allowing material discontinuities (including metal boundaries) to be modeled as easily as with the classic Yee algorithm.

What, if any, are the limitations in using hexagonal grid algorithms relative to Yee's original method? The answer is: none very significant in two dimensions. This is because it is only moderately more complicated to generate (even manually) uniform hexagonal grids than it is to generate uniform Cartesian grids. The difficulty arises in attempting to extend hexagonal gridding to three dimensions. Such an extension involves filling space with tetradecahedron and dual-tetrahedron unit cells [2]. This greatly increases the complexity of generating and visualizing the computational mesh to the point where manual mesh generation is no longer possible, and an automated computer-based approach is required.

4.9.4 Strategy 4: Use Discrete Fourier Transforms to Calculate the Spatial Derivatives

The fourth and final approach reviewed here for reduction of the numerical dispersion artifact is called the *pseudospectral time-domain* (PSTD) method [8, 9]. This technique uses a *discrete Fourier transform* (DFT) algorithm to represent the spatial derivatives in the Maxwell's equations' computational lattice. The *fast Fourier transform* (FFT) can also be applied to increase numerical efficiency. (See Chapter 17 for a comprehensive discussion.)

Formulation

The PSTD method works on unstaggered, collocated Cartesian space lattices wherein all field components are located at the same points. An example of such an arrangement is the two-dimensional TM_z grid of Fig. 4.10. To see how PSTD works, consider an x-directed cut through this grid. At every sample point along this cut, we wish to compute the x-derivatives of the general field component V. Let $\{V_i\}$ denote the set of initially known values of V at all gridpoints along the observation cut, and let $\{(\partial V/\partial x)_i\}$ denote the set of initially unknown x-derivatives of V at the same gridpoints. Then, using the differentiation theorem for Fourier transforms, we can write:

$$\left\{ \left. \frac{\partial V}{\partial x} \right|_i \right\} = -\mathcal{F}^{-1}\left(j\tilde{k}_x \, \mathcal{F}\{V_i\}\right) \tag{4.93}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote respectively the forward and inverse DFTs, and $\tilde{k}_x$ is the Fourier transform variable representing the x-component of the numerical wavevector. In this manner, the entire set of spatial derivatives of V along the observation cut can be calculated in one step. In multiple dimensions, this process is repeated for each observation cut parallel to the major axes of the space lattice.

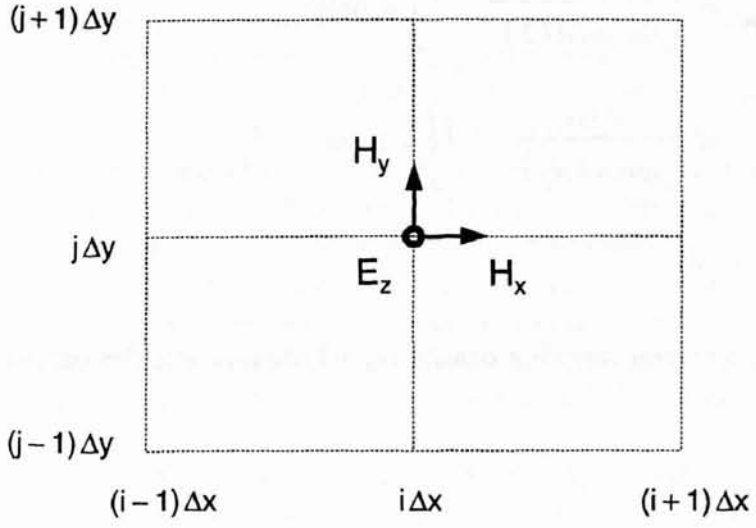


Fig. 4.10 Unstaggered, collocated Cartesian grid (with no dual-grid) used for the PSTD method, illustrated in two dimensions for the TM_z case.

According to the Nyquist sampling theorem, (4.93) is *exact* for $|\tilde{k}_x| \leq \pi/\Delta x$; that is, $\Delta x \leq \tilde{\lambda}/2$. Thus, the spatial-differencing process here can be said to be of "infinite order" for grid-sampling densities of two or more points per wavelength. The wraparound effect, a potentially major limitation caused by the periodicity assumed in the FFT, is eliminated by using the perfectly matched layer absorbing boundary condition [9]. Finally, the time-differencing for PSTD as reported in [8, 9] uses conventional second-order-accurate Yee leapfrogging.

For the PSTD method, the following expressions are obtained for the wavenumber and phase velocity of a sinusoidal numerical wave of temporal period $T = 2\pi/\omega$ propagating in an arbitrary direction within a three-dimensional space lattice [8, 9]:

$$\tilde{k} = |\tilde{\boldsymbol{k}}| = \frac{2}{c\Delta t}\sin\left(\frac{\omega\Delta t}{2}\right) \tag{4.94}$$

$$\tilde{v}_p = \frac{\omega}{\tilde{k}} = \frac{\omega}{\dfrac{2}{c\Delta t}\sin\left(\dfrac{\omega\Delta t}{2}\right)} = \frac{\omega\Delta t/2}{\sin(\omega\Delta t/2)}\,c \tag{4.95}$$

We see from (4.95) that the numerical phase velocity is *independent* of the propagation direction of the wave, unlike any of the methods considered previously in this chapter. Applying our definitions of numerical phase-velocity error, we therefore have the following figures of merit for the PSTD method:

$$\Delta\tilde{v}_{\text{physical}}\Big|_{\text{PSTD}} = \left[\frac{\omega\Delta t/2}{\sin(\omega\Delta t/2)} - 1\right] \times 100\%$$

$$= \left[\frac{\pi/N_T}{\sin(\pi/N_T)} - 1\right] \times 100\% \tag{4.96}$$

$$\Delta\tilde{v}_{\text{aniso}}\Big|_{\text{PSTD}} = 0 \tag{4.97}$$

where we define the temporal sampling density $N_T = T/\Delta t$ time samples per wave-oscillation period.

Discussion

Remarkably, $\Delta\tilde{v}_{\text{aniso}} = 0$ for the PSTD method for *all* propagating sinusoidal waves sampled at $N_\lambda \geq 2$. Therefore, to specify the gridding density of the PSTD simulation, we need only a reliable estimate of λ_{min}, the fastest oscillating spectral component of significance. This estimate is based upon the wavelength spectrum of the exciting pulse and the size of significant structural details such as material inhomogeneities. Then, the space-cell dimension is set at $\Delta = \lambda_{\text{min}}/2$, regardless of the problem's overall electrical size. This is because our choice of Δ assures zero $\Delta\tilde{v}_{\text{aniso}}$ error, and thus, zero accumulation of this error, even if the number of space cells increases without bound. Consequently, we conclude that the density of the PSTD mesh-sampling is *independent* of the electrical size of the modeling problem.

However, the fact that $\Delta\tilde{v}_{\text{aniso}} = 0$ does *not* mean that PSTD yields perfect results analogous to the magic-time-step case of the one-dimensional scalar wave equation. In fact, (4.96) shows that there remains a numerical phase-velocity error relative to c. This residual velocity error is not a function of the wave-propagation direction ϕ, and is therefore isotropic within the space grid. The residual velocity error arises from the Yee-type leapfrog time-stepping used in the algorithm, and is a function only of N_T. Table 4.1 provides representative values of this residual velocity error.

TABLE 4.1

Residual Numerical Phase-Velocity Error of the PSTD Method Versus Its Time-Sampling

N_T	$\Delta\tilde{v}_{\text{physical}}$	N_T	$\Delta\tilde{v}_{\text{physical}}$
2	+57%	15	+0.73%
4	+11%	20	+0.41%
8	+2.6%	25	+0.26%
10	+1.7%	30	+0.18%

The key point from Table 4.1 is that N_T limits the accuracy of the PSTD technique when modeling impulsive wave-propagation problems. This is not an issue for a monochromatic wave where there is only a single value of N_T, and $\Delta\tilde{v}_{\text{physical}}$ can be nulled in the manner of Strategy 1. For a pulse, however, with Δt a fixed parameter in the algorithm, there exists a spread of equivalent N_T values for the spectral components of the pulse, which possess a range of temporal periods T. This causes a spread of $\tilde{v}_p$ over the pulse spectrum, which in turn results in an isotropic progressive broadening and distortion of the pulse waveform as it propagates in the grid. To bound such dispersion, it is important to choose Δt small enough so that it adequately resolves the period $T_{\min}$ of the fastest oscillating spectral component $\lambda_{\min}$. Because this dispersion is cumulative with the wave-propagation distance, we have a second key point: the density of the PSTD time sampling must *increase* with the electrical size of the modeling problem if we apply a fixed upper bound on the maximum total phase error of propagating waves within the mesh.

Despite this need for a small Δt, PSTD can provide a large reduction in computer resources relative to the classic Yee algorithm for electrically large problems not having spatial details or material inhomogeneities smaller than $\lambda_{\min}/2$. Increased efficiency is expected, even relative to the fourth-order-accurate spatial algorithms discussed previously. References [8, 9] report that, within the range of problem sizes from 16 to 64 wavelengths, the use of PSTD permits up to an 8^D:1 reduction in computer storage and running-time relative to the Yee algorithm to produce results with comparable accuracy, where D is the problem dimensionality. We expect the PSTD advantage to increase for even larger problems. In fact, the computational benefit of PSTD theoretically increases without limit as the electrical size of the modeling problem expands.

Similar to techniques employing fourth-order approximations to the space derivatives, PSTD's global calculation of space derivatives along observation cuts through the lattice may have difficulties at material interfaces. A concern is that PSTD might yield nonphysical results for problems having abrupt jumps in permittivity and/or permeability unless a parameter-smoothing technique is used [see (4.86) and Fig. 4.8]. This concern is especially acute for computation spaces containing high-contrast material objects or metal structures. For the latter, depending upon the orientation and thickness of the metal surfaces, tangential-field discontinuities could appear for two reasons:

1. A space-cell boundary lies at a metal surface or within a metal layer. The tangential H-field within the space cell drops abruptly to zero at the metal surface, and remains at zero for the remainder of the space cell.

2. A metal sheet splits the space lattice so there exist distinct illuminated and shadow regions within the lattice. Here, the tangential H-field on the far (shadowed or shielded) side of the metal sheet may be physically isolated from the field immediately across the metal sheet on its near (illuminated) side. Gross error is caused by the global nature of PSTD's spatial-derivative calculation which nonphysically transports field information directly across the shielding metal barrier from the illuminated to the shadow sides.

In principle, these problems can be solved by the prescription of special boundary conditions, or by dividing the modeled space into zones that limit the spatial extent of each field stencil used to implement a DFT or FFT. The latter technique is discussed in detail in Chapter 17.

We note that the PSTD usage of collocated field components simplifies rendering material structures within the mesh. It is also ideal for modeling nonlinear optical problems where the local index of refraction is dependent upon a power of the magnitude of the local $\boldsymbol{E}$. Here, collocation of the vector components of $\boldsymbol{E}$ avoids the need for error-causing spatial interpolations of nearby electric field vector components staggered in space.

4.10 ALTERNATING-DIRECTION-IMPLICIT TIME-STEPPING ALGORITHM FOR OPERATION BEYOND THE COURANT LIMIT

Section 4.7 showed that numerical stability of the Yee algorithm requires a bounding of the time-step Δt relative to the space increments Δx, Δy, and Δz. Rewriting (4.54b) for convenience, this Courant stability bound is given in three dimensions by

$$\Delta t \leq \Delta t_{\max} = \frac{1}{c\sqrt{\dfrac{1}{(\Delta x)^2} + \dfrac{1}{(\Delta y)^2} + \dfrac{1}{(\Delta z)^2}}} = \left.\frac{\Delta}{c\sqrt{3}}\right|_{\substack{\text{cubic-cell} \\ \text{space lattice}}} \tag{4.98}$$

The limit on Δt set by (4.98) has allowed application of FDTD methods to a wide variety of electromagnetic wave interaction problems of moderate electrical size and quality factor. Typically, such problems require 10^3 to 10^4 time-steps to complete a single simulation.

However, there are important potential applications of FDTD modeling where the Courant stability bound of (4.98) is much too restrictive. Modeling applications that fall into this difficult regime have the following characteristics:

- The cell size Δ needed to resolve the fine-scale geometric detail of the electromagnetic wave interaction structure is much less than the shortest wavelength $\lambda_{\min}$ of a significant spectral component of the source.
- The simulated time T_{sim} needed to evolve the electromagnetic wave physics to the desired endpoint is related to the cycle time T of $\lambda_{\min}$.

With Δ fixed by the need to resolve the problem geometry, (4.98) in turn fixes $\Delta t_{\max}$ to avoid numerical instability. This, in turn, fixes the total number of time-steps needed to complete the simulation, $N_{\text{sim}} = T_{\text{sim}}/\Delta t_{\max}$. Table 4.2 lists parameters of two important classes of problems where this decision process results in values of N_{sim} that are so large that standard FDTD modeling in three dimensions is difficult, or even impossible.

If these classes of electromagnetics problems are to be explored using FDTD modeling, we need an advancement of FDTD techniques that permits accurate and numerically stable operation for values of Δt exceeding the Courant limit of (4.98) by much more than 10:1. Chapter 18 provides a state-of-the-art review of such *unconditionally stable* techniques. In this section, we summarize one candidate approach: the *alternating-direction-implicit* (ADI) time-stepping algorithm. In fact, work with ADI FDTD methods in the early 1980s [10, 11] achieved promising results for two-dimensional models. However, using the ADI techniques of [10, 11], it proved difficult to demonstrate the required numerical stability for the general three-dimensional case, and research in this area was largely discontinued.

TABLE 4.2

Two Important Classes of Three-Dimensional FDTD Modeling Problems Made Difficult or Impossible by the Courant Limit on Δt

Problem Class	Δ	T_{sim}	Δt_{max}	N_{sim}
Low-frequency bioelectromagnetics	~1 mm	~100 ms	~2 ps	$\sim 5 \times 10^{10}$
Operation of VLSI digital logic	~0.25 μm	~1 ns	~0.5 fs	$\sim 2 \times 10^{6}$

Since the late 1990s, there has been a revival of interest in the use of ADI time-stepping algorithms to obtain unconditional numerical stability for FDTD [12–16]. This research has focused on rigorous mathematical derivation of the numerical stability and dispersion properties of proposed ADI algorithms. In this section, we focus on the work reported in [15, 16], wherein for the first time unconditional numerical stability is derived for the full three-dimensional case.

Note that, with any unconditionally stable ADI FDTD algorithm, the upper bound on Δt is only that implied by the required numerical accuracy in implementing the time derivatives of the electromagnetic field. Thus, Δt need only be small enough to provide about 20 or more field samples during the cycle time T of the fastest oscillating significant spectral component of the exciting source. For example, in Table 4.2, Δt could be 0.1 ms rather than 2 ps for bioelectromagnetics problems, yielding $N_{sim} = 1{,}000$ rather than $N_{sim} = 5 \times 10^{10}$.

4.10.1 Numerical Formulation of the Zheng / Chen / Zhang Algorithm

Zheng, Chen, and Zhang (ZCZ) reported a new ADI time-stepping algorithm for FDTD that has theoretical unconditional numerical stability for the general three-dimensional case [15, 16]. While the ZCZ technique uses the same Yee space lattice as conventional FDTD, the six field-vector components are collocated rather than staggered in time. In discussing the numerical formulation of this algorithm, we assume that all of the field components are known everywhere in the lattice at time-step n and stored in the computer memory.

Unsimplified System of Time-Stepping Equations

The ADI nature of the ZCZ algorithm can be best understood by first considering its unsimplified form, and then proceeding to obtain the final simplified system of field update equations. To advance a single time-step from n to $n+1$, we perform two subiterations: the first from n to $n+1/2$, and the second from $n+1/2$ to $n+1$. These subiterations are as follows.

Subiteration 1: Advance the six field components from time-step n to time-step $n+1/2$

$$E_x\Big|_{i+1/2,j,k}^{n+1/2} = E_x\Big|_{i+1/2,j,k}^{n} + \frac{\Delta t}{2\varepsilon\,\Delta y}\left(H_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} - H_z\Big|_{i+1/2,j-1/2,k}^{n+1/2}\right) - \frac{\Delta t}{2\varepsilon\,\Delta z}\left(H_y\Big|_{i+1/2,j,k+1/2}^{n} - H_y\Big|_{i+1/2,j,k-1/2}^{n}\right) \tag{4.99a}$$

$$E_y\Big|_{i,j+1/2,k}^{n+1/2} = E_y\Big|_{i,j+1/2,k}^{n} + \frac{\Delta t}{2\varepsilon\,\Delta z}\left(H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} - H_x\Big|_{i,j+1/2,k-1/2}^{n+1/2}\right) - \frac{\Delta t}{2\varepsilon\,\Delta x}\left(H_z\Big|_{i+1/2,j+1/2,k}^{n} - H_z\Big|_{i-1/2,j+1/2,k}^{n}\right) \tag{4.99b}$$

$$E_z\Big|_{i,j,k+1/2}^{n+1/2} = E_z\Big|_{i,j,k+1/2}^{n} + \frac{\Delta t}{2\varepsilon\,\Delta x}\left(H_y\Big|_{i+1/2,j,k+1/2}^{n+1/2} - H_y\Big|_{i-1/2,j,k+1/2}^{n+1/2}\right) - \frac{\Delta t}{2\varepsilon\,\Delta y}\left(H_x\Big|_{i,j+1/2,k+1/2}^{n} - H_x\Big|_{i,j-1/2,k+1/2}^{n}\right) \tag{4.99c}$$

$$H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} = H_x\Big|_{i,j+1/2,k+1/2}^{n} + \frac{\Delta t}{2\mu\,\Delta z}\left(E_y\Big|_{i,j+1/2,k+1}^{n+1/2} - E_y\Big|_{i,j+1/2,k}^{n+1/2}\right) - \frac{\Delta t}{2\mu\,\Delta y}\left(E_z\Big|_{i,j+1,k+1/2}^{n} - E_z\Big|_{i,j,k+1/2}^{n}\right) \tag{4.100a}$$

$$H_y\Big|_{i+1/2,j,k+1/2}^{n+1/2} = H_y\Big|_{i+1/2,j,k+1/2}^{n} + \frac{\Delta t}{2\mu\,\Delta x}\left(E_z\Big|_{i+1,j,k+1/2}^{n+1/2} - E_z\Big|_{i,j,k+1/2}^{n+1/2}\right) - \frac{\Delta t}{2\mu\,\Delta z}\left(E_x\Big|_{i+1/2,j,k+1}^{n} - E_x\Big|_{i+1/2,j,k}^{n}\right) \tag{4.100b}$$

$$H_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} = H_z\Big|_{i+1/2,j+1/2,k}^{n} + \frac{\Delta t}{2\mu\,\Delta y}\left(E_x\Big|_{i+1/2,j+1,k}^{n+1/2} - E_x\Big|_{i+1/2,j,k}^{n+1/2}\right) - \frac{\Delta t}{2\mu\,\Delta x}\left(E_y\Big|_{i+1,j+1/2,k}^{n} - E_y\Big|_{i,j+1/2,k}^{n}\right) \tag{4.100c}$$

In each of the above equations, the first finite-difference on the right-hand side is set up to be evaluated implicitly from as yet unknown field data at time-step $n+1/2$, while the second finite-difference on the right-hand side is evaluated explicitly from known field data at time-step n.

Subiteration 2: Advance the six field components from time-step $n+1/2$ to $n+1$

$$E_x\Big|_{i+1/2,j,k}^{n+1} = E_x\Big|_{i+1/2,j,k}^{n+1/2} + \frac{\Delta t}{2\varepsilon\,\Delta y}\left(H_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} - H_z\Big|_{i+1/2,j-1/2,k}^{n+1/2}\right) - \frac{\Delta t}{2\varepsilon\,\Delta z}\left(H_y\Big|_{i+1/2,j,k+1/2}^{n+1} - H_y\Big|_{i+1/2,j,k-1/2}^{n+1}\right) \tag{4.101a}$$

$$E_y\Big|_{i,j+1/2,k}^{n+1} = E_y\Big|_{i,j+1/2,k}^{n+1/2} + \frac{\Delta t}{2\varepsilon\,\Delta z}\left(H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} - H_x\Big|_{i,j+1/2,k-1/2}^{n+1/2}\right) - \frac{\Delta t}{2\varepsilon\,\Delta x}\left(H_z\Big|_{i+1/2,j+1/2,k}^{n+1} - H_z\Big|_{i-1/2,j+1/2,k}^{n+1}\right) \tag{4.101b}$$

$$E_z\Big|_{i,j,k+1/2}^{n+1} = E_z\Big|_{i,j,k+1/2}^{n+1/2} + \frac{\Delta t}{2\varepsilon\,\Delta x}\left(H_y\Big|_{i+1/2,j,k+1/2}^{n+1/2} - H_y\Big|_{i-1/2,j,k+1/2}^{n+1/2}\right) - \frac{\Delta t}{2\varepsilon\,\Delta y}\left(H_x\Big|_{i,j+1/2,k+1/2}^{n+1} - H_x\Big|_{i,j-1/2,k+1/2}^{n+1}\right) \tag{4.101c}$$

$$H_x\Big|_{i,j+1/2,k+1/2}^{n+1} = H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} + \frac{\Delta t}{2\mu\,\Delta z}\left(E_y\Big|_{i,j+1/2,k+1}^{n+1/2} - E_y\Big|_{i,j+1/2,k}^{n+1/2}\right) - \frac{\Delta t}{2\mu\,\Delta y}\left(E_z\Big|_{i,j+1,k+1/2}^{n+1} - E_z\Big|_{i,j,k+1/2}^{n+1}\right) \tag{4.102a}$$

$$H_y\Big|_{i+1/2,j,k+1/2}^{n+1} = H_y\Big|_{i+1/2,j,k+1/2}^{n+1/2} + \frac{\Delta t}{2\mu\,\Delta x}\left(E_z\Big|_{i+1,j,k+1/2}^{n+1/2} - E_z\Big|_{i,j,k+1/2}^{n+1/2}\right) - \frac{\Delta t}{2\mu\,\Delta z}\left(E_x\Big|_{i+1/2,j,k+1}^{n+1} - E_x\Big|_{i+1/2,j,k}^{n+1}\right) \tag{4.102b}$$

$$H_z\Big|_{i+1/2,j+1/2,k}^{n+1} = H_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} + \frac{\Delta t}{2\mu\,\Delta y}\left(E_x\Big|_{i+1/2,j+1,k}^{n+1/2} - E_x\Big|_{i+1/2,j,k}^{n+1/2}\right) - \frac{\Delta t}{2\mu\,\Delta x}\left(E_y\Big|_{i+1,j+1/2,k}^{n+1} - E_y\Big|_{i,j+1/2,k}^{n+1}\right) \tag{4.102c}$$

In each of the above equations, the second finite-difference on the right-hand side is set up to be evaluated implicitly from as yet unknown field data at time-step $n+1$, while the first finite-difference on the right-hand side is evaluated explicitly from known field data at time-step $n+1/2$ previously computed using (4.99) and (4.100).

Simplified System of Time-Stepping Equations

The system of equations summarized above for each subiteration can be greatly simplified. For Subiteration 1, this is done by substituting the expressions of (4.100) for the H-field components evaluated at time-step $n+1/2$ into the E-field updates of (4.99). Similarly, for Subiteration 2, this is done by substituting the expressions of (4.102) for the H-field components evaluated at time-step $n+1$ into the E-field updates of (4.101). This yields the following simplified system of time-stepping equations for the ZCZ algorithm:

Subiteration 1: Advance the six field components from time-step n to time-step $n+1/2$

$$\left[1+\frac{(\Delta t)^2}{2\mu\varepsilon(\Delta y)^2}\right]E_x\Big|^{n+1/2}_{i+1/2,j,k}-\left[\frac{(\Delta t)^2}{4\mu\varepsilon(\Delta y)^2}\right]\left(E_x\Big|^{n+1/2}_{i+1/2,j-1,k}+E_x\Big|^{n+1/2}_{i+1/2,j+1,k}\right)$$

$$=E_x\Big|^{n}_{i+1/2,j,k}+\frac{\Delta t}{2\varepsilon\Delta y}\left(H_z\Big|^{n}_{i+1/2,j+1/2,k}-H_z\Big|^{n}_{i+1/2,j-1/2,k}\right)$$

$$-\frac{\Delta t}{2\varepsilon\Delta z}\left(H_y\Big|^{n}_{i+1/2,j,k+1/2}-H_y\Big|^{n}_{i+1/2,j,k-1/2}\right) \tag{4.103a}$$

$$-\left[\frac{(\Delta t)^2}{4\mu\varepsilon\Delta x\Delta y}\right]\left(E_y\Big|^{n}_{i+1,j+1/2,k}-E_y\Big|^{n}_{i,j+1/2,k}-E_y\Big|^{n}_{i+1,j-1/2,k}+E_y\Big|^{n}_{i,j-1/2,k}\right)$$

$$\left[1+\frac{(\Delta t)^2}{2\mu\varepsilon(\Delta z)^2}\right]E_y\Big|^{n+1/2}_{i,j+1/2,k}-\left[\frac{(\Delta t)^2}{4\mu\varepsilon(\Delta z)^2}\right]\left(E_y\Big|^{n+1/2}_{i,j+1/2,k-1}+E_y\Big|^{n+1/2}_{i,j+1/2,k+1}\right)$$

$$=E_y\Big|^{n}_{i,j+1/2,k}+\frac{\Delta t}{2\varepsilon\Delta z}\left(H_x\Big|^{n}_{i,j+1/2,k+1/2}-H_x\Big|^{n}_{i,j+1/2,k-1/2}\right)$$

$$-\frac{\Delta t}{2\varepsilon\Delta x}\left(H_z\Big|^{n}_{i+1/2,j+1/2,k}-H_z\Big|^{n}_{i-1/2,j+1/2,k}\right) \tag{4.103b}$$

$$-\left[\frac{(\Delta t)^2}{4\mu\varepsilon\Delta y\Delta z}\right]\left(E_z\Big|^{n}_{i,j+1,k+1/2}-E_z\Big|^{n}_{i,j,k+1/2}-E_z\Big|^{n}_{i,j+1,k-1/2}+E_z\Big|^{n}_{i,j,k-1/2}\right)$$

$$\left[1+\frac{(\Delta t)^2}{2\mu\varepsilon(\Delta x)^2}\right]E_z\Big|^{n+1/2}_{i,j,k+1/2}-\left[\frac{(\Delta t)^2}{4\mu\varepsilon(\Delta x)^2}\right]\left(E_z\Big|^{n+1/2}_{i-1,j,k+1/2}+E_z\Big|^{n+1/2}_{i+1,j,k+1/2}\right)$$

$$=E_z\Big|^{n}_{i,j,k+1/2}+\frac{\Delta t}{2\varepsilon\Delta x}\left(H_y\Big|^{n}_{i+1/2,j,k+1/2}-H_y\Big|^{n}_{i-1/2,j,k+1/2}\right)$$

$$-\frac{\Delta t}{2\varepsilon\Delta y}\left(H_x\Big|^{n}_{i,j+1/2,k+1/2}-H_x\Big|^{n}_{i,j-1/2,k+1/2}\right) \tag{4.103c}$$

$$-\left[\frac{(\Delta t)^2}{4\mu\varepsilon\Delta x\Delta z}\right]\left(E_x\Big|^{n}_{i+1/2,j,k+1}-E_x\Big|^{n}_{i+1/2,j,k}-E_x\Big|^{n}_{i-1/2,j,k+1}+E_x\Big|^{n}_{i-1/2,j,k}\right)$$

$$H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} = H_x\Big|_{i,j+1/2,k+1/2}^{n} + \frac{\Delta t}{2\mu\,\Delta z}\left(E_y\Big|_{i,j+1/2,k+1}^{n+1/2} - E_y\Big|_{i,j+1/2,k}^{n+1/2}\right) - \frac{\Delta t}{2\mu\,\Delta y}\left(E_z\Big|_{i,j+1,k+1/2}^{n} - E_z\Big|_{i,j,k+1/2}^{n}\right) \tag{4.104a}$$

$$H_y\Big|_{i+1/2,j,k+1/2}^{n+1/2} = H_y\Big|_{i+1/2,j,k+1/2}^{n} + \frac{\Delta t}{2\mu\,\Delta x}\left(E_z\Big|_{i+1,j,k+1/2}^{n+1/2} - E_z\Big|_{i,j,k+1/2}^{n+1/2}\right) - \frac{\Delta t}{2\mu\,\Delta z}\left(E_x\Big|_{i+1/2,j,k+1}^{n} - E_x\Big|_{i+1/2,j,k}^{n}\right) \tag{4.104b}$$

$$H_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} = H_z\Big|_{i+1/2,j+1/2,k}^{n} + \frac{\Delta t}{2\mu\,\Delta y}\left(E_x\Big|_{i+1/2,j+1,k}^{n+1/2} - E_x\Big|_{i+1/2,j,k}^{n+1/2}\right) - \frac{\Delta t}{2\mu\,\Delta x}\left(E_y\Big|_{i+1,j+1/2,k}^{n} - E_y\Big|_{i,j+1/2,k}^{n}\right) \tag{4.104c}$$

We see that (4.103a) yields a set of simultaneous equations for $E_x^{n+1/2}$ when written for each j coordinate along a y-directed line through the space lattice. The matrix associated with this system is tridiagonal, and hence, easily solved. This process is repeated for each y-cut through the grid where E_x components are located. Similarly, (4.103b) yields a tridiagonal matrix system for each z-cut through the lattice to obtain $E_y^{n+1/2}$, and (4.103c) yields a tridiagonal matrix system for each x-cut through the lattice to obtain $E_z^{n+1/2}$.

We also note that (4.104a-c) are exactly (4.100a-c), repeated for convenience. These H-field updating equations are now fully explicit because all of their required E-field component data at time-step $n+1/2$ are available upon solving (4.103a-c) in the manner described above.

Subiteration 2: Advance the six field components from time-step $n+1/2$ to $n+1$

$$\left[1 + \frac{(\Delta t)^2}{2\mu\varepsilon(\Delta z)^2}\right] E_x\Big|_{i+1/2,j,k}^{n+1} - \left[\frac{(\Delta t)^2}{4\mu\varepsilon(\Delta z)^2}\right]\left(E_x\Big|_{i+1/2,j,k-1}^{n+1} + E_x\Big|_{i+1/2,j,k+1}^{n+1}\right)$$

$$= E_x\Big|_{i+1/2,j,k}^{n+1/2} + \frac{\Delta t}{2\varepsilon\,\Delta y}\left(H_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} - H_z\Big|_{i+1/2,j-1/2,k}^{n+1/2}\right) - \frac{\Delta t}{2\varepsilon\,\Delta z}\left(H_y\Big|_{i+1/2,j,k+1/2}^{n+1/2} - H_y\Big|_{i+1/2,j,k-1/2}^{n+1/2}\right) - \left[\frac{(\Delta t)^2}{4\mu\varepsilon\,\Delta x\,\Delta z}\right]\left(E_z\Big|_{i+1,j,k+1/2}^{n+1/2} - E_z\Big|_{i,j,k+1/2}^{n+1/2} - E_z\Big|_{i+1,j,k-1/2}^{n+1/2} + E_z\Big|_{i,j,k-1/2}^{n+1/2}\right) \tag{4.105a}$$

$$\left[1 + \frac{(\Delta t)^2}{2\mu\varepsilon(\Delta x)^2}\right] E_y\Big|_{i,j+1/2,k}^{n+1} - \left[\frac{(\Delta t)^2}{4\mu\varepsilon(\Delta x)^2}\right]\left(E_y\Big|_{i-1,j+1/2,k}^{n+1} + E_y\Big|_{i+1,j+1/2,k}^{n+1}\right)$$

$$= E_y\Big|_{i,j+1/2,k}^{n+1/2} + \frac{\Delta t}{2\varepsilon\Delta z}\left(H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} - H_x\Big|_{i,j+1/2,k-1/2}^{n+1/2}\right)$$

$$- \frac{\Delta t}{2\varepsilon\Delta x}\left(H_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} - H_z\Big|_{i-1/2,j+1/2,k}^{n+1/2}\right)$$

$$- \left[\frac{(\Delta t)^2}{4\mu\varepsilon\Delta x\Delta y}\right]\left(E_x\Big|_{i+1/2,j+1,k}^{n+1/2} - E_x\Big|_{i+1/2,j,k}^{n+1/2} - E_x\Big|_{i-1/2,j+1,k}^{n+1/2} + E_x\Big|_{i-1/2,j,k}^{n+1/2}\right) \tag{4.105b}$$

$$\left[1 + \frac{(\Delta t)^2}{2\mu\varepsilon(\Delta y)^2}\right] E_z\Big|_{i,j,k+1/2}^{n+1} - \left[\frac{(\Delta t)^2}{4\mu\varepsilon(\Delta y)^2}\right]\left(E_z\Big|_{i,j-1,k+1/2}^{n+1} + E_z\Big|_{i,j+1,k+1/2}^{n+1}\right)$$

$$= E_z\Big|_{i,j,k+1/2}^{n+1/2} + \frac{\Delta t}{2\varepsilon\Delta x}\left(H_y\Big|_{i+1/2,j,k+1/2}^{n+1/2} - H_y\Big|_{i-1/2,j,k+1/2}^{n+1/2}\right)$$

$$- \frac{\Delta t}{2\varepsilon\Delta y}\left(H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} - H_x\Big|_{i,j-1/2,k+1/2}^{n+1/2}\right)$$

$$- \left[\frac{(\Delta t)^2}{4\mu\varepsilon\Delta y\Delta z}\right]\left(E_y\Big|_{i,j+1/2,k+1}^{n+1/2} - E_y\Big|_{i,j+1/2,k}^{n+1/2} - E_y\Big|_{i,j-1/2,k+1}^{n+1/2} + E_y\Big|_{i,j-1/2,k}^{n+1/2}\right) \tag{4.105c}$$

$$H_x\Big|_{i,j+1/2,k+1/2}^{n+1} = H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} + \frac{\Delta t}{2\mu\Delta z}\left(E_y\Big|_{i,j+1/2,k+1}^{n+1/2} - E_y\Big|_{i,j+1/2,k}^{n+1/2}\right)$$

$$- \frac{\Delta t}{2\mu\Delta y}\left(E_z\Big|_{i,j+1,k+1/2}^{n+1} - E_z\Big|_{i,j,k+1/2}^{n+1}\right) \tag{4.106a}$$

$$H_y\Big|_{i+1/2,j,k+1/2}^{n+1} = H_y\Big|_{i+1/2,j,k+1/2}^{n+1/2} + \frac{\Delta t}{2\mu\Delta x}\left(E_z\Big|_{i+1,j,k+1/2}^{n+1/2} - E_z\Big|_{i,j,k+1/2}^{n+1/2}\right)$$

$$- \frac{\Delta t}{2\mu\Delta z}\left(E_x\Big|_{i+1/2,j,k+1}^{n+1} - E_x\Big|_{i+1/2,j,k}^{n+1}\right) \tag{4.106b}$$

$$H_z\Big|_{i+1/2,j+1/2,k}^{n+1} = H_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} + \frac{\Delta t}{2\mu\Delta y}\left(E_x\Big|_{i+1/2,j+1,k}^{n+1/2} - E_x\Big|_{i+1/2,j,k}^{n+1/2}\right)$$

$$- \frac{\Delta t}{2\mu\Delta x}\left(E_y\Big|_{i+1,j+1/2,k}^{n+1} - E_y\Big|_{i,j+1/2,k}^{n+1}\right) \tag{4.106c}$$

We see that (4.105a) yields a tridiagonal matrix system for each z-cut through the lattice to allow calculation of E_x^{n+1}. Further, (4.105b) yields a tridiagonal matrix system for each x-cut through the lattice to allow calculation of E_y^{n+1}, and (4.105c) yields a tridiagonal matrix system for each y-cut through the lattice to allow calculation of E_z^{n+1}. We also note that (4.106a–c) are exactly (4.102a–c), repeated for convenience. These H-field updating equations are now fully explicit because all of their required E-component data at time-step $n+1$ are available upon solving (4.105a–c) in the manner described above. This completes the algorithm.

4.10.2 Sources

The ADI-FDTD algorithm given by (4.103) to (4.106) was derived for the source-free Maxwell's curl equations. Extending this algorithm to include an electric or magnetic current source must be done carefully. The use of explicit source conditions wherein specific electric and/or magnetic field components at source points in the grid are updated separately from the implicit electric field or explicit magnetic field ADI-FDTD updating equations leads to localized errors [17]. If, for example, the ADI-FDTD algorithm is formulated so that the electric fields are updated implicitly, then the electric current source excitation function should be embedded in the known column vector on the right-hand side of the tridiagonal matrix system for the x-, y-, or z-directed lines that pass through the location of the current source.

By considering the ADI-FDTD algorithm as a second-order-in-time perturbation of a Crank-Nicolson FDTD scheme, [18] rigorously determined the discrete time values that should be used to evaluate current source terms in each of the two subiterations of the ADI-FDTD scheme. In the first subiteration, the currents should be sampled at $t = (n + 1/2)\Delta t$, even though the time derivatives and the fields are evaluated at $t = (n + 1/4)\Delta t$. Likewise, in the second subiteration, the currents should be sampled at $t = (n + 1/2)\Delta t$, while the time derivatives and the fields are evaluated at $t = (n + 3/4)\Delta t$. This suggests that the update equations in each subiteration are not consistent with Maxwell's curl equations; that is, they do not approximate the curl equations when all of the space and time increments tend to zero. However, the overall scheme is still consistent with Maxwell's curl equations up to the second order in both time and space. This can be shown by obtaining the truncation error in the same manner used for the source-free lossless case in [19]. It is interesting to note that each subiteration in the time-stepping algorithm can lose consistency without impacting the overall consistency of the total scheme. An alternative method for the temporal sampling of the current sources is found in [17]. In this alternative formulation, the currents are evaluated at $t = (n + 1/4)\Delta t$ in the first subiteration and at $t = (n + 3/4)\Delta t$ in the second subiteration, thereby maintaining consistency within each subiteration. However, this alternative method yields less accurate results [18].

4.10.3 Numerical Stability

In [15], Zheng, Chen, and Zhang provided a detailed theoretical stability analysis of their ZCZ algorithm summarized in Section 4.10.1. The key results of their analysis are now summarized. We assume that for each time-step n, the instantaneous values of the E- and H-fields are Fourier-transformed into the spatial spectral domain with wavenumbers $\tilde{k}_x$, $\tilde{k}_y$, and $\tilde{k}_z$ along the x-, y-, and z-directions, respectively. Denoting the composite field vector in the spatial spectral domain at time-step n as

$$\{F^n\} = \left\{E_x^n \quad E_y^n \quad E_z^n \quad H_x^n \quad H_y^n \quad H_z^n\right\}^T \tag{4.107}$$

it can be shown that Subiteration 1 [consisting of the systems of equations (4.103) and (4.104)] can be written in the spatial spectral domain in matrix form as

$$\{F^{n+1/2}\} = [M_1]\{F^n\} \tag{4.108}$$

where

$$[M_1] = \begin{bmatrix} \frac{1}{Q_y} & \frac{W_x W_y}{\mu\varepsilon Q_y} & 0 & 0 & \frac{jW_z}{\varepsilon Q_y} & \frac{-jW_y}{\varepsilon Q_y} \\ 0 & \frac{1}{Q_z} & \frac{W_z W_y}{\mu\varepsilon Q_z} & \frac{-jW_z}{\varepsilon Q_z} & 0 & \frac{jW_x}{\varepsilon Q_z} \\ \frac{W_x W_z}{\mu\varepsilon Q_x} & 0 & \frac{1}{Q_x} & \frac{jW_y}{\varepsilon Q_x} & \frac{-jW_x}{\varepsilon Q_x} & 0 \\ 0 & \frac{-jW_z}{\mu Q_z} & \frac{jW_z}{\mu Q_z} & \frac{1}{Q_z} & 0 & \frac{W_x W_z}{\mu\varepsilon Q_z} \\ \frac{jW_z}{\mu Q_x} & 0 & \frac{-jW_x}{\mu Q_x} & \frac{W_x W_y}{\mu\varepsilon Q_x} & \frac{1}{Q_x} & 0 \\ \frac{-jW_y}{\mu Q_y} & \frac{jW_x}{\mu Q_y} & 0 & 0 & \frac{W_z W_y}{\mu\varepsilon Q_y} & \frac{1}{Q_y} \end{bmatrix} \tag{4.109}$$

and

$$W_x = \frac{\Delta t}{\Delta x}\sin\left(\frac{\tilde{k}_x \Delta x}{2}\right); \qquad W_y = \frac{\Delta t}{\Delta y}\sin\left(\frac{\tilde{k}_y \Delta y}{2}\right); \qquad W_z = \frac{\Delta t}{\Delta z}\sin\left(\frac{\tilde{k}_z \Delta z}{2}\right) \tag{4.110a}$$

$$Q_x = 1 + \frac{(W_x)^2}{\mu\varepsilon}; \qquad Q_y = 1 + \frac{(W_y)^2}{\mu\varepsilon}; \qquad Q_z = 1 + \frac{(W_z)^2}{\mu\varepsilon} \tag{4.110b}$$

Similarly, it can be shown that Subiteration 2 [consisting of the systems of equations (4.105) and (4.106)] can be written in the spatial spectral domain in matrix form as

$$\{F^{n+1}\} = [M_2]\{F^{n+1/2}\} \tag{4.111}$$

where

$$[M_2] = \begin{bmatrix} \frac{1}{Q_z} & 0 & \frac{W_z W_x}{\mu\varepsilon Q_z} & 0 & \frac{jW_z}{\varepsilon Q_z} & \frac{-jW_y}{\varepsilon Q_z} \\ \frac{W_x W_y}{\mu\varepsilon Q_x} & \frac{1}{Q_x} & 0 & \frac{-jW_z}{\varepsilon Q_x} & 0 & \frac{jW_x}{\varepsilon Q_x} \\ 0 & \frac{W_y W_z}{\mu\varepsilon Q_y} & \frac{1}{Q_y} & \frac{jW_y}{\varepsilon Q_y} & \frac{-jW_x}{\varepsilon Q_y} & 0 \\ 0 & \frac{-jW_z}{\mu Q_y} & \frac{jW_y}{\mu Q_y} & \frac{1}{Q_y} & \frac{W_z W_y}{\mu\varepsilon Q_y} & 0 \\ \frac{jW_z}{\mu Q_z} & 0 & \frac{-jW_x}{\mu Q_z} & 0 & \frac{1}{Q_z} & \frac{W_z W_y}{\mu\varepsilon Q_z} \\ \frac{-jW_y}{\mu Q_x} & \frac{jW_x}{\mu Q_x} & 0 & 0 & \frac{W_x W_z}{\mu\varepsilon Q_x} & \frac{1}{Q_x} \end{bmatrix} \tag{4.112}$$

Now, substituting (4.108) into (4.111) yields in matrix form the complete single time-step update expression in the spatial spectral domain:

$$\{F^{n+1}\} = [M_2][M_1]\{F^n\} \tag{4.113}$$

Using the software package MAPLE™, Zheng, Chen, and Zhang found that the magnitudes of all of the eigenvalues of the composite matrix $[M] = [M_2][M_1]$ equal unity, regardless of the time-step Δt. Therefore, they concluded that their ZCZ algorithm is *unconditionally stable* for all Δt, and the Courant stability condition is removed.

4.10.4 Numerical Dispersion

In [16], Zheng and Chen provided a detailed analysis of the numerical dispersion of their ZCZ algorithm summarized in Section 4.10.1. They derived the following relation:

$$\sin^2(\omega t) = \frac{4\mu\varepsilon\left[\begin{matrix}\mu\varepsilon(W_x)^2 + \mu\varepsilon(W_y)^2 + \mu\varepsilon(W_z)^2 + \\ (W_x)^2(W_y)^2 + (W_y)^2(W_z)^2 + (W_z)^2(W_x)^2\end{matrix}\right]\left[(\mu\varepsilon)^3 + (W_x)^2(W_y)^2(W_z)^2\right]}{\left[\mu\varepsilon + (W_x)^2\right]^2\left[\mu\varepsilon + (W_y)^2\right]^2\left[\mu\varepsilon + (W_z)^2\right]^2} \tag{4.114}$$

For Δt below the Courant limit of the classic Yee leapfrog algorithm, it was found that the numerical dispersion given by (4.114) is quite close to that of the Yee algorithm. For Δt above the usual Courant limit, the dispersive error given by (4.114) increases steadily.

4.10.5 Additional Accuracy Limitations and Their Implications

González García, Lee, and Hagness [19] demonstrated additional accuracy limitations of ADI-FDTD not revealed by previously published numerical dispersion analyses such as that given in [16]. They showed that some terms of its truncation error grow with $(\Delta t)^2$ multiplied by the spatial derivatives of the fields. These error terms, which are not present in a fully implicit time-stepping method such as the Crank-Nicolson scheme, give rise to potentially large numerical errors as Δt is increased. Excessive error can occur even if Δt is still small enough to highly resolve key temporal features of the modeled electromagnetic field waveform.

As a result, the primary usage of existing ADI-FDTD techniques appears to be for problems involving a fine mesh needed to model a small geometric feature in an overall much larger structure that is discretized using a coarse mesh; and where, for computational efficiency, it is desirable to use a large time-step satisfying Courant stability for the coarse mesh. While this limits the impact of the excess error introduced locally within the fine mesh, this also limits the usefulness of ADI-FDTD when considering how to model the key problem areas outlined in Table 4.2.

4.11 SUMMARY

This chapter derived the key relations for numerical dispersion and stability applicable to FDTD modeling of Maxwell's equations in multiple dimensions. These included:

- Derivation of the numerical dispersion relation for two-dimensional wave propagation;
- Extension to three-dimensions;
- Comparison with the ideal dispersion case;
- Anisotropy of the numerical phase velocity;
- Complex-valued numerical wavenumbers;
- Numerical stability derived using a complex-frequency analysis;
- The generalized numerical stability problem;
- Modified Yee-based algorithms for improved numerical dispersion, including fourth-order spatial differencing, use of hexagonal grids, and the pseudospectral time-domain method;
- Alternating-direction-implicit time-stepping algorithm for operation beyond the Courant limit.

Because this area is rapidly evolving, the reader is encouraged to keep up with technology developments by monthly reading of relevant technical journals such as *IEEE Transactions on Antennas and Propagation* and *IEEE Transactions on Microwave Theory and Techniques.*

REFERENCES

[1] 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 Tech.*, Vol. 23, 1975, pp. 623–630.

[2] Liu, Y., "Fourier analysis of numerical algorithms for the Maxwell's equations," *J. Computational Physics*, Vol. 124, 1996, pp. 396–416.

[3] Schneider, J. B., and C. L. Wagner, "FDTD dispersion revisited: Faster-than-light propagation," *IEEE Microwave Guided Wave Lett.*, Vol. 9, 1999, pp. 54–56.

[4] Min, M. S., and C. H. Teng, "The instability of the Yee scheme for the magic time step," *J. Computational Physics*, Vol. 166, 2001, pp. 418–424.

[5] Fang, J., *Time-Domain Finite Difference Computations for Maxwell's Equations*, Ph.D. dissertation, EECS Dept., Univ. of California, Berkeley, CA, 1989.

[6] Zygiridis, T., and T. D. Tsiboukis, "Low-dispersion algorithms based on the higher order (2, 4) FDTD method," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, 2004, pp. 1321–1327.

[7] Turkel, E., "High-order methods," Chap. 2 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Taflove, A., (ed.), Boston, MA: Artech House, 1998.

[8] Liu, Q. H., "The pseudospectral time-domain (PSTD) method: A new algorithm for solutions of Maxwell's equations," *Proc. 1997 IEEE Antennas Propagat. Soc. Intl. Symp.*, Vol. 1, 1997, pp. 122–125.

[9] Liu, Q. H., *The PSTD Algorithm: A Time-Domain Method Requiring Only Two Grids Per Wavelength*, New Mexico State Univ., Las Cruces, NM, Tech. Report NMSU-ECE96-013, Dec. 1996.

[10] Holland, R., "Implicit three-dimensional finite differencing of Maxwell's equations," *IEEE Trans. Nuclear Science*, Vol. NS-31, 1984, pp. 1322–1326.

[11] Holland, R., and K. S. Cho, *Alternating-Direction Implicit Differencing of Maxwell's Equations: 3D Results*, Computer Sciences Corp., Albuquerque, NM, Tech. Report, Contract DAAL02-85-C-0200, Harry Diamond Labs., Adelphi, MD, June 1, 1986.

[12] Namiki, T., "A new FDTD algorithm based on alternating-direction implicit method," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, 1999, pp. 2003–2007.

[13] Namiki, T., "3-D ADI-FDTD method—Unconditionally stable time-domain algorithm for solving full vector Maxwell's equations," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 1743–1748.

[14] Namiki, T., and K. Ito, "Investigation of numerical errors of the two-dimensional ADI-FDTD method," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 1950–1956.

[15] Zhen, F., Z. Chen, and J. Zhang, "Toward the development of a three-dimensional unconditionally stable finite-difference time-domain method," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 1550–1558.

[16] Zheng, F., and Z. Chen, "Numerical dispersion analysis of the unconditionally stable 3-D ADI-FDTD method," *IEEE Trans. Microwave Theory Tech.*, Vol. 49, 2001, pp. 1006–1009.

[17] Lee, T. W., and S. C. Hagness, "Wave source conditions for the unconditionally stable ADI-FDTD method," *Proc. 2001 IEEE Antennas Propagat. Soc. Intl. Symp.*, Vol. 4, 2001, pp. 142–145.

[18] González García, S., A. Rubio Bretones, R. Gomez Martin, and S. C. Hagness, "Accurate implementation of current sources in the ADI–FDTD scheme," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 3, 2004, pp. 141–144.

[19] González García, S., T.-W. Lee, and S. C. Hagness, "On the accuracy of the ADI-FDTD method," *IEEE Antennas Wireless Propagation Lett.*, Vol. 1, 2002, pp. 31–34.

PROBLEMS

4.1 Derive numerical dispersion relation (4.5) by completing the steps of substituting traveling waves (4.3) into finite-difference equations (4.2).

4.2 Show that the numerical dispersion relations associated with the second-order-accurate FDTD solution to the one-dimensional scalar wave equation given by (2.29a) and the one-dimensional Maxwell's curl equations given by (4.7b) are identical.

4.3 Derive numerical dispersion relation (4.11) by completing the steps of substituting traveling wave (4.9) into the Yee finite-differencing of (4.8b).

4.4 Derive analytical dispersion relation (4.13) for an arbitrarily directed plane wave in three dimensions. Show that numerical dispersion relation (4.12) reduces to (4.13) in the limit of infinitesimal time and space increments.

4.5 Derive numerical wavenumber and numerical phase-velocity expressions (4.14a), (4.14b), (4.15a), and (4.15b).

4.6 Using the two-dimensional TM_z FDTD code that was constructed in Chapter 3, replicate the results shown in Figs. 4.1(a) and 4.1(b).

4.7 Write a computer program to implement the Newton's method iterative procedure of (4.16a) and (4.16b), and replicate the results shown in Fig. 4.2.

4.8 Confirm the derivations of (4.36), (4.37a), and (4.37b).

4.9 Confirm the derivations of (4.46), (4.47a), and (4.47b).

4.10 Replicate the results shown in Figs. 4.3 and 4.4.

4.11 Using the two-dimensional TM_z FDTD code that was constructed in Chapter 3, replicate the results shown in Figs. 4.5(a) and 4.5(b).

4.12 Confirm the derivations of (4.57), (4.58), and (4.59).

4.13 Confirm the Courant factor normalization given by (4.66d) and (4.67d).

4.14 Using the two-dimensional TM_z FDTD code that was constructed in Chapter 3, replicate the results shown in Figs. 4.6(a) and 4.6(b).

4.15 Using the differentiation property of Fourier transforms given by (4.93), write and test a computer program that performs forward and inverse FFTs on a function $V(x)$ to obtain its numerical derivative. For the numerical experiment, let $V(x) = \exp[-(x - x_0)^2/a^2]$ where $x_0 = 45\text{m}$, $a = 5\text{m}$, and the spatial domain spans 100m. Verify the accuracy of the program by comparing the numerical and analytical derivatives of $V(x)$.

PROJECTS

P4.1 Modify the one-dimensional FDTD code that was constructed in Chapter 3 to use the fourth-order-accurate spatial-differencing of (4.70). In the manner of Fig. 4.7, compare the L_2-normed errors versus time for the two codes when modeling free-space, non-magic-time-step propagation of a Gaussian pulse of full width equal to 6 cells at its $1/e$ points.

P4.2 Modify the two-dimensional TM_z FDTD code that was constructed in Chapter 3 to use the explicit fourth-order-accurate spatial differencing system of (4.70). Then, repeat Problem 4.5 using the new code, and compare with the previous results.

P4.3 Model an air-dielectric half-space problem using the following one-dimensional FDTD codes: (a) classic Yee algorithm; and (b) basic fourth-order-explicit spatial differencing. Do this first with no special treatment of the permittivity jump between air and the dielectric half-space. Then, implement the permittivity-smoothing tactic of (4.86) in both codes. Compare the L_2-normed errors versus time for all cases in the manner of Fig. 4.8.

P4.4 Construct a new two-dimensional TM_z FDTD code using the space-time algorithm system of (4.88) applied on the hexagonal grid shown in Fig. 4.9(b). Then, repeat Problem 4.5 using the new code, and compare with the previous results in Problem 4.5 and Project P4.2.

P4.5 Construct a new one-dimensional code that implements the PSTD scheme of (4.93) on an unstaggered collocated grid. In the manner of Fig. 4.7, compare the L_2-normed errors versus time for the new PSTD code with the results of Projects P4.1 and P4.3, when modeling non-magic-time-step propagation of a Gaussian pulse of full width equal to 6 cells at its $1/e$ points.

Chapter 5

Incident Wave Source Conditions

Allen Taflove, Geoff Waldschmidt, Christopher Wagner, John Schneider, and Susan Hagness

5.1 INTRODUCTION

In the two previous chapters, we considered basic aspects of the discretization, stability, and dispersion of Maxwell's equations when numerically approximated in a uniform Cartesian FDTD space lattice. In this chapter, we consider another fundamental point: how to introduce electromagnetic wave excitations that are useful for modeling engineering problems.

A generic issue in FDTD modeling has been to accurately realize the physics of an electromagnetic wave source in as spatially compact a manner as possible. This means using relatively few $\boldsymbol{E}$ and $\boldsymbol{H}$ components to realize the wave source compared to the total number of field components in the lattice. In this manner, the computer storage and running time needed to simulate the source is small compared to that needed for ordinary time-stepping of the fields. Maximum algorithm efficiency is achieved.

This chapter reviews four very useful, general classes of compact electromagnetic wave sources to spotlight their key features: (1) hard-sourced $\boldsymbol{E}$ and $\boldsymbol{H}$ fields in one- and two-dimensional grids; (2) $\boldsymbol{J}$ and $\boldsymbol{M}$ current sources in three-dimensional lattices; (3) the total-field / scattered-field formulation for plane-wave excitation in one, two, and three dimensions; and (4) waveguide sources. While not compact, a fifth class of wave source, the pure scattered-field formulation, is also reviewed because of its utility. Chapter 15 will discuss in detail an emerging sixth class of FDTD wave sources: linear and nonlinear electronic circuits. Sources considered there include resistors, capacitors, inductors, diodes, transistors, and logic gates. These are useful for three-dimensional modeling of high-speed digital circuits and microwave amplifiers. Research is ongoing in developing such hybrid FDTD / electronic-circuit models.

5.2 POINTWISE *E* AND *H* HARD SOURCES IN ONE DIMENSION

A *hard source* is set up simply by assigning a desired time function to specific components of $\boldsymbol{E}$ or $\boldsymbol{H}$ in the FDTD space lattice. This time function is independent of anything else in the model. For example, in a one-dimensional, x-directed Yee grid, the following E_z hard source could be established at grid-point i_s to generate a continuous sinusoidal wave of frequency f_0 that is switched on at $n = 0$:

$$E_z\big|_{i_s}^n = E_0 \sin(2\pi f_0 n\Delta t) \tag{5.1}$$

The resulting wave would propagate bilaterally in both the $+x$ and $-x$ directions from the source point. To provide the proper direction of the Poynting power flow away from the source, the FDTD-calculated spatial distribution of the E_z component of the wave would assume even symmetry about the source point, while the H_y component would assume odd symmetry.

A second common pointwise hard source provides a lowpass Gaussian pulse with finite dc content. The time waveform of the pulse is centered at time-step n_0 and has a $1/e$ characteristic decay of n_{decay} time-steps:

$$E_z\big|_{i_s}^n = E_0 e^{-\left[(n-n_0)/n_{\text{decay}}\right]^2} \tag{5.2}$$

Note that (5.2) has a nonzero value at $n = 0$, so that if a smooth transition from zero into the Gaussian pulse is required, n_0 should be taken as at least $3n_{\text{decay}}$.

A third common pointwise hard source provides a bandpass Gaussian pulse with zero dc content. The Fourier spectrum of this pulse has even symmetry about f_0. The pulse is again centered at time-step n_0 and has a $1/e$ characteristic decay of n_{decay} time-steps:

$$E_z\big|_{i_s}^n = E_0 e^{-\left[(n-n_0)/n_{\text{decay}}\right]^2} \sin\left[2\pi f_0 (n-n_0)\Delta t\right] \tag{5.3}$$

Each hard source in (5.1) to (5.3) radiates a numerical wave having a time waveform corresponding to the source function. If a material structure is specified in the FDTD grid at some distance from the source point, the radiated numerical wave eventually propagates to this structure and undergoes partial transmission and partial reflection. In principle, time-stepping can be continued until all transients decay. For the source of (5.1), this would mean the attainment of the sinusoidal steady state for the transmitted and reflected fields. For the sources of (5.2) and (5.3), this would mean the evolution of the complete time histories of the transmitted and reflected waves. Discrete Fourier analysis of these time histories obtained in a single FDTD run would provide the magnitude and phase of the transmission and reflection coefficients over a potentially wide frequency band starting at dc.

However, these optimistic scenarios are problematic. As time-stepping continues to obtain either the sinusoidal steady state or the late-time impulse response, the reflected numerical wave eventually returns to the grid source location i_s. Because the total tangential E-field is specified at i_s without regard to any possible reflected waves in the grid, the hard source causes a spurious, nonphysical retroreflection of these waves at i_s back toward the material structure of interest. In effect, the hard source prevents the movement of reflected wave energy through the source position toward infinity, and thereby fails to properly simulate a physical incident wave.

This can be easily seen for the hard sources of (5.2) and (5.3), where for $n - n_0 >> n_{\text{decay}}$ the E-field at i_s decays to zero. This gridpoint then simulates the behavior of a perfect electric conductor (PEC) having zero tangential E-field. It can be demonstrated that the same retroreflecting effect occurs for the sinusoidal hard source of (5.1) or any other hard source, even though the tangential fields at the source point may not be pinned at zero. In fact, total reflection of an impinging wave always occurs whenever the total tangential field at a surface is specified without regard to the values of the incident field.

One way to eliminate the reflective nature of a pulsed hard source is simply to remove it from the algorithm after its pulse time waveform has decayed essentially to zero. Subsequently, we apply instead the normal Yee field update at i_s. In the context of the Gaussian pulse source of (5.2), we would program the equivalent of the following update relation for the E-field at i_s:

$$\text{IF}(\ (n-n_0)/n_{\text{decay}} \le 3.0\ ) \quad E_z\Big|_{i_s}^{n} = E_0 e^{-\left[(n-n_0)/n_{\text{decay}}\right]^2}$$
$$\text{ELSE} \quad E_z\Big|_{i_s}^{n} = C_a(m)E_z\Big|_{i_s}^{n-1} + C_b(m)\left(H_y\Big|_{i_s+1/2}^{n-1/2} - H_y\Big|_{i_s-1/2}^{n-1/2}\right) \tag{5.4}$$

However, this approach cannot be used for source waveforms such as continuous sinusoids, which have extended durations such that the source remains active after reflections from the material structure propagate back to it.

In general, using a hard source in a one-dimensional FDTD simulation limits the maximum number of time-steps that can be run without spurious retroreflections contaminating the computed fields in the vicinity of the material structure being modeled. However, pointwise hard sources in two dimensions or collinear arrays of hard-sourced field-vector components in three dimensions can be useful for exciting numerical models of waveguides and microstrips. Here, the hard sources simulate a metal probe extending from a feeding coaxial line into the waveguide. (Note that a zero generator source impedance is being modeled.) It is observed that much less error in the calculated field occurs for pointwise hard sources in two and three dimensions than in one dimension. This is because pointwise hard sources intercept and retroreflect much smaller fractions of the total energy in a multidimensional FDTD space lattice.

5.3 POINTWISE *E* AND *H* HARD SOURCES IN TWO DIMENSIONS

A pointwise ***E*** (***H***) hard source located within a two-dimensional FDTD TM_z (TE_z) grid excites a radially propagating cylindrical wave centered on the source point. Due to numerical dispersion, the propagating wavefront experiences a fractional-percent lagging velocity error relative to the free-space speed of light *c*. The wavefront also undergoes a fractional-percent distortion from its ideal circular shape because of the variation of the numerical phase velocity with the local propagation direction along the wavefront. Means to theoretically and numerically quantify these errors were presented in Sections 4.2, 4.5, and 4.6 of Chapter 4.

This section discusses a revealing alternative means to test the accuracy of cylindrical wave generation by a pointwise hard source: comparison of the FDTD calculations with the exact analytical solution [1]. Following conventional terminology, we say that this section focuses on the frequency-domain Green function for the two-dimensional Yee grid. Our analysis leads to the interesting and useful conclusion that a "pointwise" hard source is really not a point at all. In fact, a pointwise hard source has a finite effective radius of approximately 0.2 grid cell in two-dimensional FDTD models.

5.3.1 Green Function for the Scalar Wave Equation in Two Dimensions

For simplicity and clarity in our discussion, we consider Maxwell's equations for the two-dimensional TE_z mode given by the system (3.14). After some manipulation, it can be shown that the *H*-field in free space is governed by the following scalar wave equation:

$$\nabla_{xy}^2 H_z = \frac{\partial^2 H_z}{\partial x^2} + \frac{\partial^2 H_z}{\partial y^2} = \frac{1}{c^2}\frac{\partial^2 H_z}{\partial t^2} \tag{5.5}$$

where $c = 1/\sqrt{\mu_0 \varepsilon_0}$ is the speed of light in free space. Noting that the same type of wave equation is obtained for E_z for the TM_z mode given by (3.13), we can generalize (5.5) for both the TM_z and TE_z modes by identifying either E_z or H_z as the scalar function $u(x, y, t)$. Now, we assume that such a physical system is excited by a Dirac delta function occurring at $t = 0$ at the coordinate origin:

$$u(x, y, t) = \delta(x)\,\delta(y)\,\delta(t) \tag{5.6}$$

It can be shown that this generates a circular cylindrical wave radiating outward from the coordinate origin (the z-axis in our TM_z and TE_z field models) at the speed of light, as given exactly by the following two-dimensional Green function G_{2D}:

$$u(r, t) \equiv G_{2D}(r, t) = \frac{c\,U(ct - r)}{2\pi\sqrt{c^2t^2 - r^2}} \tag{5.7}$$

where $r^2 = x^2 + y^2$ and the unit step function U is given by

$$U(ct-r) = \begin{cases} 0 & \text{for } r > ct \\ 1 & \text{for } r < ct \end{cases} \tag{5.8}$$

The radial coordinate of the outgoing wavefront is given by $r = ct$, an intuitive result. Here, the denominator in (5.7) equals zero and the amplitude of the outgoing wave is infinite.

As discussed in Chapter 4, numerical dispersion prevents the Yee algorithm from accurately following the step discontinuity of u at the wavefront. However, much can be learned by taking the Fourier transform of (5.7) to obtain the two-dimensional frequency-domain Green function:

$$\breve{u}(r, \omega) \equiv G_{2D}(r, \omega) = \frac{j}{4} H_0^{(2)}(kr) \tag{5.9}$$

Here, $\breve{u}$ is the phasor E_z or H_z field of angular frequency $\omega = 2\pi f$, $H_0^{(2)}(kr)$ is a Hankel function of the second kind, and k is the free-space wavenumber at ω. The magnitude of $G_{2D}(r, \omega)$ is a continuous function of r. It is exactly the envelope of the oscillating u field that radiates symmetrically outward from a time-harmonic source of angular frequency ω at the origin.

5.3.2 Obtaining Comparative FDTD Data

Because of the spatially continuous nature of $G_{2D}(r, \omega)$, it is possible to obtain a direct comparison with sinusoidal steady-state FDTD results at any selected set of points away from the origin. The required FDTD data can be obtained in one of two ways. For example, if a smooth, essentially time-limited impulse function such as a Gaussian pulse is used to excite the FDTD grid, the time-harmonic response $\breve{u}(r, \omega)$ can be obtained by running a discrete Fourier transformation on $u(r, t)$ concurrently with the time-stepping (taking care to continue this process until all transients die down to a negligible level). Alternatively, if a sinusoidal source is turned on at $t = 0$ and then kept on sufficiently long, eventually only the time-harmonic response remains at any point. Thereafter, we need only to activate a simple peak-detection function working on $u(r, t)$ to directly sense the envelope of the oscillations.

Two related problems remain in comparing FDTD data with $G_{2D}(r, \omega)$:

- The theoretical magnitude of $G_{2D}(r, \omega)$ is infinite at the origin ($r = 0$) because of the logarithmic singularity of the Hankel function of (5.9):

$$\lim_{kr \to 0} H_0^{(2)}(kr) \sim \ln(kr) \tag{5.10}$$

- The amplitude of the waveform $u(t)$ used to excite a single E_z or H_z field component at the center of the FDTD grid cannot be infinite.

Despite these problems, the procedure now discussed permits accurate calculation of $G_{2D}(r, \omega)$ using a square-cell FDTD grid having an appropriate cell size Δ. The procedure is based upon the assumption that exciting a single H_z (E_z) component in a two-dimensional TE_z (TM_z) grid is equivalent to exciting a cylindrical virtual surface of radius $r_{eq} = f^* \Delta$ centered about that component, where f^* is a fraction that is virtually independent of Δ. We call f^* the effective action radius of the excited field component. The following steps are taken [1]:

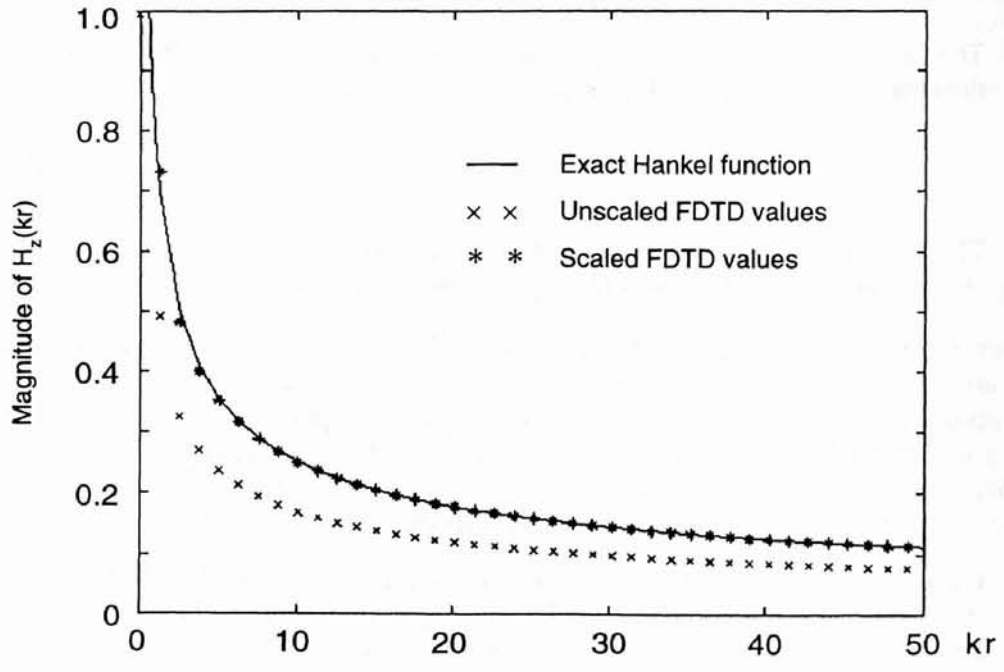
1. Hard source a single H_z (E_z) component at the center of a large TE_z (TM_z) free-space grid. Use a sinusoid of period T and amplitude 1 A/m (1 V/m), and let the grid cell size be $\Delta = \lambda_0 / N_\lambda = (cT)/N_\lambda$. Run the FDTD simulation for a sufficient number of T to reach the sinusoidal steady state in grid region *SSS* centered on the hard source. Make sure that the radiated wavefront does not reach the outer boundary of the grid to avoid computational artifacts due to wave reflections.
2. Process the FDTD data obtained in Step 1 to yield the sinusoidal steady-state values of all of the H_z (E_z) field components in *SSS*. Graph these values versus their radial distance from the hard-sourced H_z (E_z).
3. Determine a scaling factor C that multiplies the entire set of FDTD data obtained in Step 2 to yield the best fit (in the L_2 or energy-normed sense) between the FDTD and exact Green function results for the field-magnitude variation with radial distance from the source.
4. Then, by our basic assumption concerning the effective action radius of an excited field component in a two-dimensional grid, we equate

$$C = \left| G_{2D}(f^* \Delta, \omega) \right| = \left| \frac{j}{4} H_0^{(2)}(k f^* \Delta) \right| \tag{5.11}$$

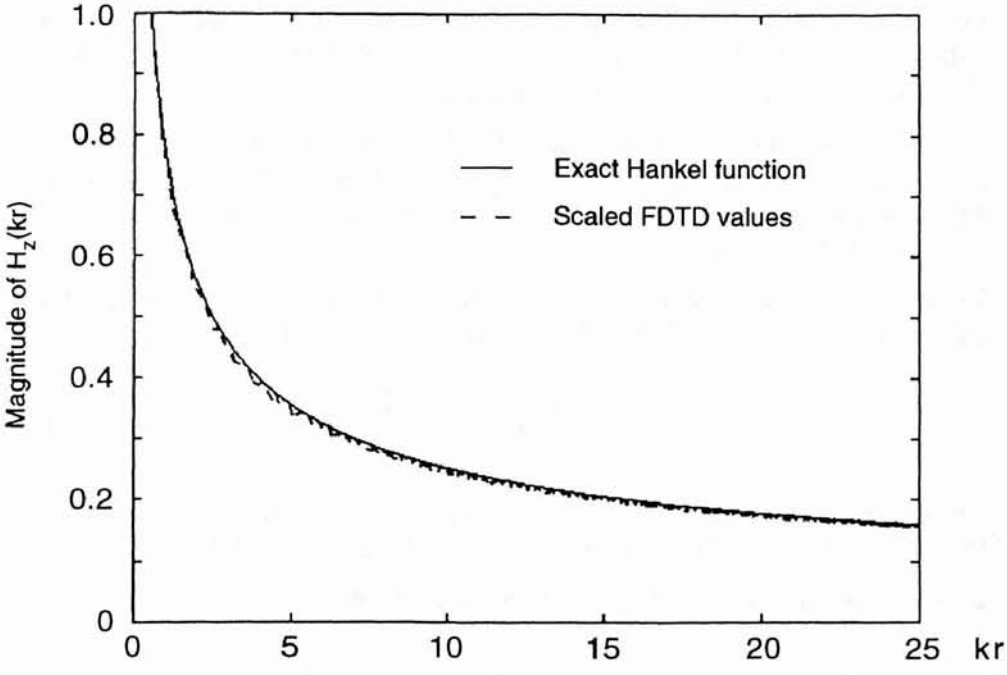
   where $k = 2\pi / \lambda_0$. Upon substituting C, k, and Δ into (5.11), we obtain f^* by using a table of Hankel function values generated by MATLAB.
5. Repeat Steps 1 to 4 for other N_λ and check the convergence of f^*.

5.3.3 Results for Effective Action Radius of a Hard-Sourced Field Component

Fig. 5.1(a) compares the radial variation of the magnitude of the exact Green function solution (5.9) with both raw and best-fit scaled FDTD data for $|H_z(kr)|$ along the x-axis of a coarse grid having $\Delta = \lambda_0 / 5$. From (5.11), this yields $f^* = 0.247$. Fig. 5.1(b) is similar except that it compares scaled FDTD data throughout the *entire grid* for $\Delta = \lambda_0 / 10$.



(a) Comparative field values along the x-axis for $\lambda_0/5$ grid resolution



(b) Comparative field values throughout the entire grid for $\lambda_0/10$ grid resolution

Fig. 5.1 Comparison of the radial variation of the exact frequency-domain Green function with scaled FDTD sinusoidal steady-state H_z data in a uniform, square-cell TE_z grid. *Source:* Waldschmidt and Taflove, *IEEE Microwave and Guided Wave Lett.*, 2000, pp. 217–219, © 2000 IEEE.

Relative to Fig. 5.1(b), our methodology is to scale the FDTD results twice so that the Green function forms either the upper or lower envelope of the scatter of FDTD values. [For clarity, Fig. 5.1(b) shows only the best fit to the scaled upper envelope.] This yields two results for the optimal scaling factor, which by (5.11) provides two corresponding values of f^*. These bound f^* for the particular grid resolution used.

TABLE 5.1

Range of Effective Action Radius of a Hard-Sourced Field Component in a 2D FDTD Grid.
Source: Waldschmidt and Taflove, *IEEE Microwave and Guided Wave Lett.*, 2000, pp. 217–219, © 2000 IEEE.

Grid Cell Size Δ	*Range of Effective Radius* f^*
$\lambda_0/5$	0.149 – 0.247
$\lambda_0/10$	0.187 – 0.211
$\lambda_0/20$	0.200 – 0.207
$\lambda_0/40$	0.206 – 0.209

Table 5.1 lists the range of the effective action radius f^* of the hard-sourced H_z field component, as determined using the above procedure. For a grid cell size Δ finer than $\lambda_0/10$, $f^* \cong 0.2$ and is only a weak function of Δ, as originally assumed. In the limit as $\Delta << \lambda_0/10$, the effective action radius converges to the value $f^* \cong 0.21$. We conclude that, over the commonly used range of FDTD grid resolutions $\lambda_0/10$ to $\lambda_0/40$, a hard-sourced H_z (E_z) field component in a two-dimensional TE_z (TM_z) grid exhibits an effective action radius of approximately 0.2 space cell.

5.4 *J* AND *M* CURRENT SOURCES IN THREE DIMENSIONS

A pointwise source located within a three-dimensional FDTD space lattice excites a radially propagating wave centered on the source point. Similar to the two-dimensional case, numerical dispersion causes the propagating wavefront to experience a fractional-percent lagging velocity error relative to the free-space speed of light c. The wavefront also undergoes a fractional-percent distortion because of the variation of the numerical phase velocity with the local propagation direction along the wavefront. Eventually, the radiated wave exits the computational space, assuming that a suitable absorbing boundary condition (discussed in Chapters 6 and 7) is applied at the outer-boundary planes of the space lattice.

In addition to radiating fields to infinity, $\boldsymbol{J}$ and $\boldsymbol{M}$ current sources in three-dimensional FDTD space lattices can deposit charge and generate charge-associated fields [2]. Such charge can exist even though there is no explicit storage location for it. Indeed, the charge exists only insofar as diverging fields exist. These diverging fields, which are required to satisfy the continuity equation, can persist *indefinitely*, and hence remain in the computational domain even after all of the radiated fields have exited.

As shown in Chapter 3 (Section 3.6.9), the Yee space lattice is divergence-free in free-space regions in the absence of a source. However, substantial field divergence (and thus, equivalent charge deposition) properly results from current sources embedded in the lattice that have specific geometrical and temporal properties. An example of a current-source geometry that deposits charge is a contiguous, open-ended chain of $\boldsymbol{J}$ components forming a filamentary radiator. Here, electric charge is deposited at the two ends of the filament where the current diverges. An example of a temporal property that influences charge deposition is the presence (or absence) of a dc component in the time waveform used to excite the current source. A time waveform having a dc component permits deposition of persistent charge, while a waveform having no dc component can produce only temporary charging.

This section examines how $\boldsymbol{J}$ and $\boldsymbol{M}$ current sources in the three-dimensional Yee space lattice cause equivalent electric and magnetic charge deposition. This results in either temporary or persistent nonpropagating E- and H-fields to be generated within the lattice, in addition to the desired radiated, outward-propagating wave. Since we show that the Yee lattice stores electric and magnetic charge, we can define an effective lattice capacitance and inductance. This must be accounted for in models of embedded electronic components, as discussed in Chapter 15.

5.4.1 Sources and Charging

In the following discussion, the relationship between the electric field $\boldsymbol{E}$ and the electric charge density ρ in free space is given by Gauss' law[1]:

$$\varepsilon_0 \nabla \cdot \boldsymbol{E} = \rho \tag{5.12}$$

When the E-field diverges from a point, (5.12) states that the charge density is nonzero. The Yee space lattice allows the divergence to be computed with central differences, thus preserving the second-order accuracy of the charge-density computations.

Further, the relationship between the electric current density $\boldsymbol{J}$ and the electric charge density ρ is given by the continuity equation[2]:

$$\nabla \cdot \boldsymbol{J} = -\frac{\partial \rho}{\partial t} \tag{5.13}$$

We recall that, in the Yee lattice, the components of $\boldsymbol{J}$ are spatially collocated with corresponding collinear components of $\boldsymbol{E}$, but are offset by one-half time-step.

Consider an electric-current filament that consists of one or more contiguous components of $\boldsymbol{J}$ forming a continuous path. In accordance with (5.13), electric charge exists at the ends of this source. The amount of charge at one end of the current source (i.e., the amount of charge enclosed within a surface surrounding the end of the filament), is given by the space and time integration of (5.13):

---

[1] An equation analogous to (5.12) relates the divergence of the magnetic field $\boldsymbol{H}$ to the presence of nonphysical magnetic charge density ρ_m.

[2] An equation analogous to (5.13) relates the nonphysical magnetic current density $\boldsymbol{M}$ to the time-rate of change of the nonphysical magnetic charge density ρ_m.

$$-\int_{-\infty}^{t}\left(\oiint_{S} \boldsymbol{J}\cdot d\boldsymbol{s}\right)dt' \;=\; \int_{-\infty}^{t} I(t')dt' \;=\; Q_{\text{enclosed}} \tag{5.14}$$

where $I(t)$ is the total current entering the volume and Q_{enclosed} is the enclosed charge. We see that Q_{enclosed} can be determined by integrating $I(t)$ over time. Alternately, Q_{enclosed} can be determined numerically in an FDTD simulation by a discrete volume integration of Gauss' law (5.12). If the volume containing the charge is a single space lattice cell of edge length Δ, this yields

$$\varepsilon_0 \oiint_{S} \boldsymbol{E}\cdot d\boldsymbol{s} \;=\; \varepsilon_0(\Delta)^2 \sum_{\text{six faces}} E_{\text{face}} \;=\; Q_{\text{enclosed}} \tag{5.15}$$

where E_{face} is the total E-field on a cube face.

We now illustrate the deposition of charge and confirm the correspondence between: (1) the "expected" charge given by the time integral of $I(t)$ in (5.14), and (2) the "measured" or calculated charge in the space lattice given by (5.15), the flux integral of the FDTD-computed $\boldsymbol{E}$. To this end, we source a single J_z component (realizing a Hertzian dipole) located at the center of an $81 \times 81 \times 81$ cubic-cell space lattice having $\Delta = 1\text{m}$. The J_z component is driven by a Gaussian-pulse time waveform.

Fig. 5.2 shows the expected and measured charge as a function of time at the two ends of the current source [2]. We see that the measured FDTD results correspond precisely to those predicted by the time integral of the current. Thus, although the FDTD method is not considered to be a dc analysis technique, it does, nevertheless, properly predict the fields associated with the rearrangement of fixed (dc) electric charge.

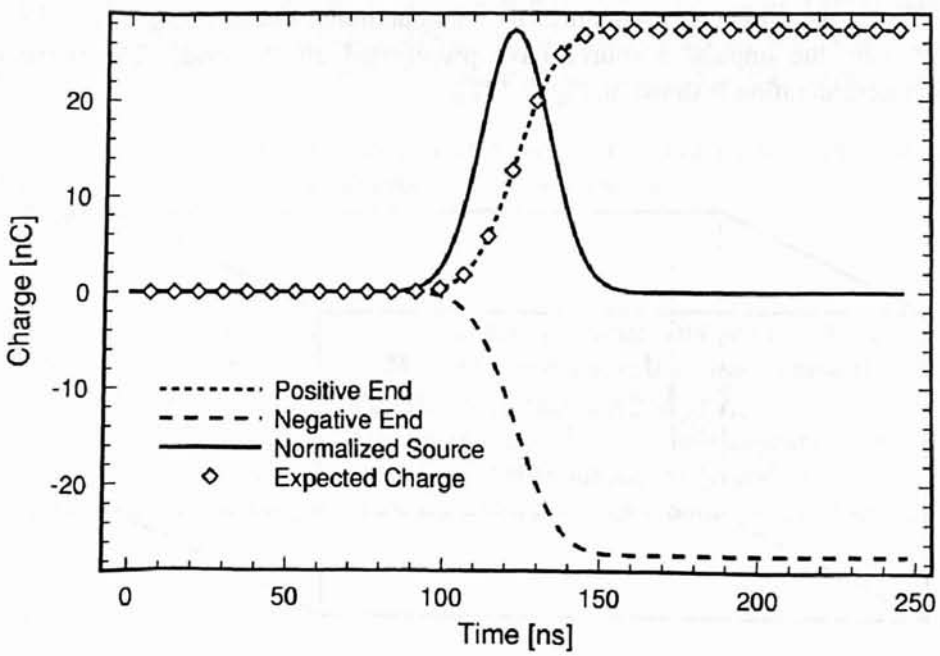


Fig. 5.2 Agreement of the time dependence of the expected and measured charge deposited at the ends of an impulsively excited Hertzian dipole located at the center of a three-dimensional Yee space lattice. *Source:* Wagner and Schneider, *IEEE Trans. Microwave Theory and Techniques*, 1998, pp. 2131–2136, © 1998 IEEE.

5.4.2 Sinusoidal Sources

FDTD simulations sometimes employ sinusoidal sources. However, because such a source is turned on at some initial time and the total modeling time is finite, we should not expect the resulting model to be completely valid for the sinusoidal steady-state (time-harmonic) case. One important aspect of this problem is that the time integral of the current-source function can have a nonzero dc value. A specific example is a sinusoidal current turned on at $t = 0$, which deposits the following charge into the computational domain:

$$Q(t) = \int_0^t I(t')\,dt' = \int_0^t \sin t'\,dt' = 1 - \cos t \tag{5.16}$$

Here, for simplicity, unit frequency and unit amplitude are used. We see that the charge oscillates between 0 and 2 and has a time-average value of 1. This average charge might be deposited at the ends of a dipole source, thereby producing nonzero dc fields throughout unshielded portions of the computational domain.

We note that cosine currents turned on at $t = 0$ deposit a zero time-average charge. However, the large turn-on discontinuity of the cosine source function at $t = 0$ contains significant high-frequency spectral components which suffer large numerical dispersion. Therefore, the use of cosine source functions is not advised.

5.4.3 Transient (Pulse) Sources

This section illustrates how the static fields generated due to charge deposition within the three-dimensional Yee space lattice can remain in the computational domain long after all radiated fields generated by the impulsive source have propagated off the grid. The space lattice geometry under consideration is shown in Fig. 5.3 [2].

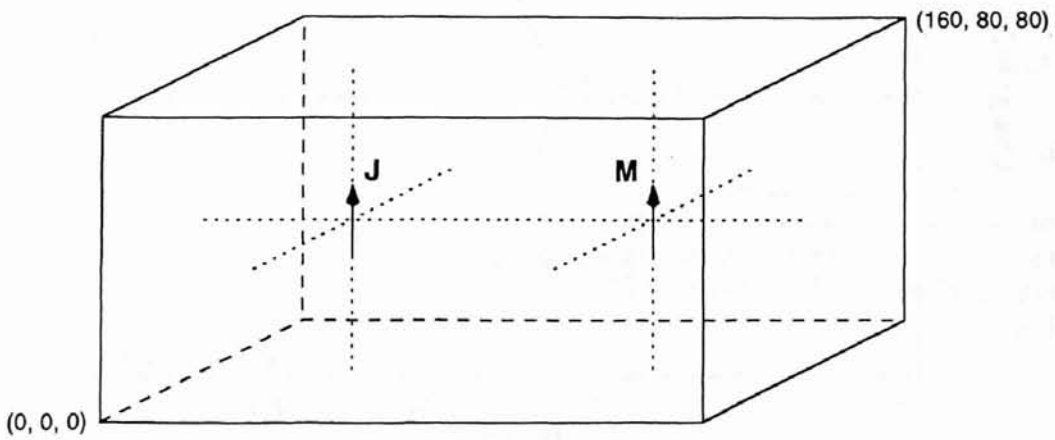


Fig. 5.3 FDTD space-lattice geometry for the illustration of the persistence of static fields generated due to charge deposition. *Source:* Wagner and Schneider, *IEEE Trans. Microwave Theory and Techniques*, 1998, pp. 2131–2136, © 1998 IEEE.

Fig. 5.3 depicts a $161 \times 81 \times 81$ cubic-cell Yee space lattice that contains two z-directed filamentary current sources, each 20 collinear elements long. The first source is an electric-current filament $\boldsymbol{J}$ whose center is located at cell (40, 40, 40), and the second source is a magnetic-current filament $\boldsymbol{M}$ whose center is located at cell (120, 40, 40). The time-waveform of each filament's excitation is a Gaussian pulse having a width of 16 time-steps between its $1/e$ points.

Figs. 5.4(a) and 5.4(b) are logarithmic-grayscale visualizations[3] of the magnitude of $\boldsymbol{E}$ and $\boldsymbol{H}$, respectively, at $n = 100$ time-steps within the z-directed lattice cut plane that contains both the electric-current and magnetic-current filaments [2]. At this early time in the FDTD simulation, both current sources are radiating outward-propagating dynamic fields. Note that Figs. 5.4(a) and 5.4(b) are duals.

Figs. 5.5(a) and 5.5(b) repeat these visualizations at $n = 300$ time-steps [2]. Now, the dynamic fields have propagated off the space lattice, which is terminated by an absorbing boundary condition. Only static fields remain. These diverge from electric or magnetic charges deposited at the ends of each source filament. For example, Fig. 5.5(a) shows that the $\boldsymbol{J}$ filament deposits electric charge at its ends and generates a dc E-field that is strongest near this charge. However, the $\boldsymbol{M}$ filament deposits no electric charge and therefore generates zero dc E-field. The dual situation is given in Fig. 5.5(b), which shows that the $\boldsymbol{M}$ filament deposits magnetic charge at its ends and generates a dc H-field that is strongest near this charge. Because the $\boldsymbol{J}$ filament deposits no magnetic charge, it generates zero dc H-field.

5.4.4 Intrinsic Lattice Capacitance

In the physical world, electric charges exert on each other an attractive or repulsive Coulomb force, leading to charge motion under the influence of this force. In an FDTD simulation that does not specifically provide for charge dynamics, electric charges do not move. Consequently, since adjacent space lattice cells can store charge, it is natural to define a capacitance between cells of the space lattice.

The capacitance of adjacent cells in free space in the Yee lattice can be derived in the following manner [2]. The standard definition of capacitance is

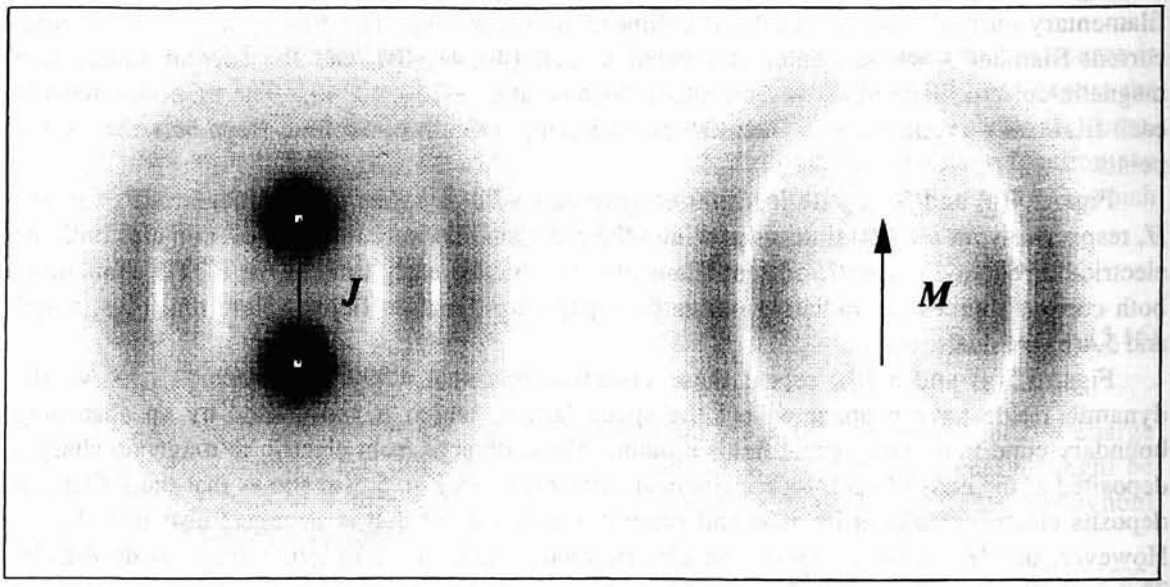
$$C = Q / V \tag{5.17}$$

where Q is the stored charge and V is the electric potential between the charges. The charge stored in the grid can be found either from the time integral of the current that deposited the charge, as in (5.14), or from the finite-difference flux integral of $\boldsymbol{E}$ (Gauss' law), as in (5.15). By expressing the charge in terms of the flux integral, an E-field appears in the numerator of (5.17), which can be cancelled with the E-field that subsequently appears in the denominator. When computing the flux integral, the total E-field at any point on the flux surface can be decomposed into two parts:

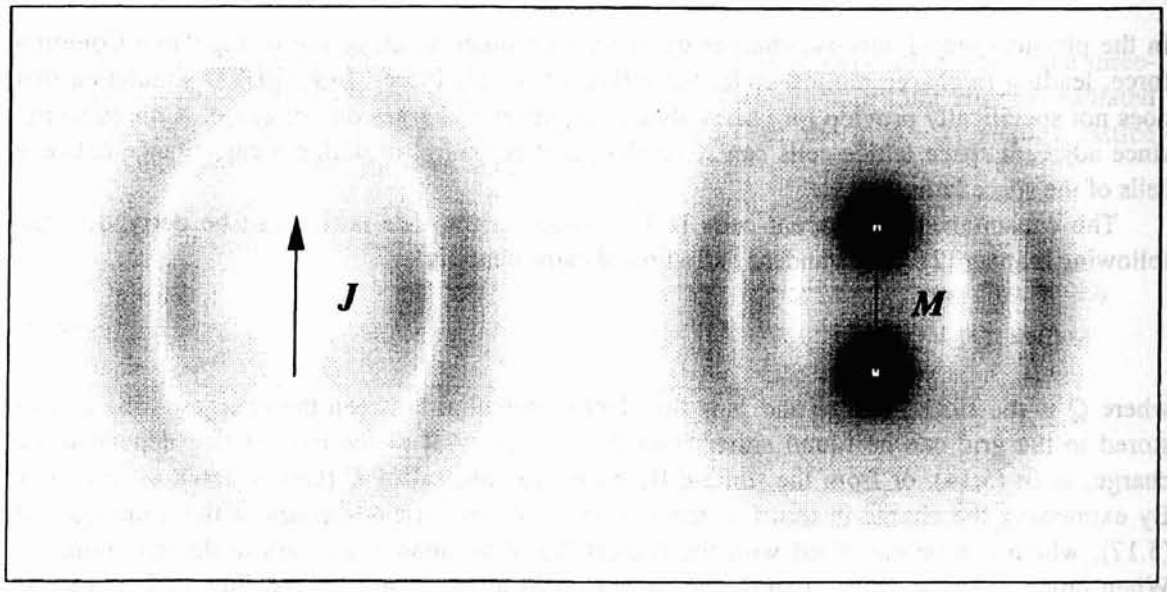
$$\boldsymbol{E}_{\text{face}} = \boldsymbol{E}_{\text{enclosed}} + \boldsymbol{E}_{\text{distant}} \tag{5.18}$$

---

[3]For convenience, all field data in Figs. 5.4 and 5.5 are normalized to permit a direct visual comparison.

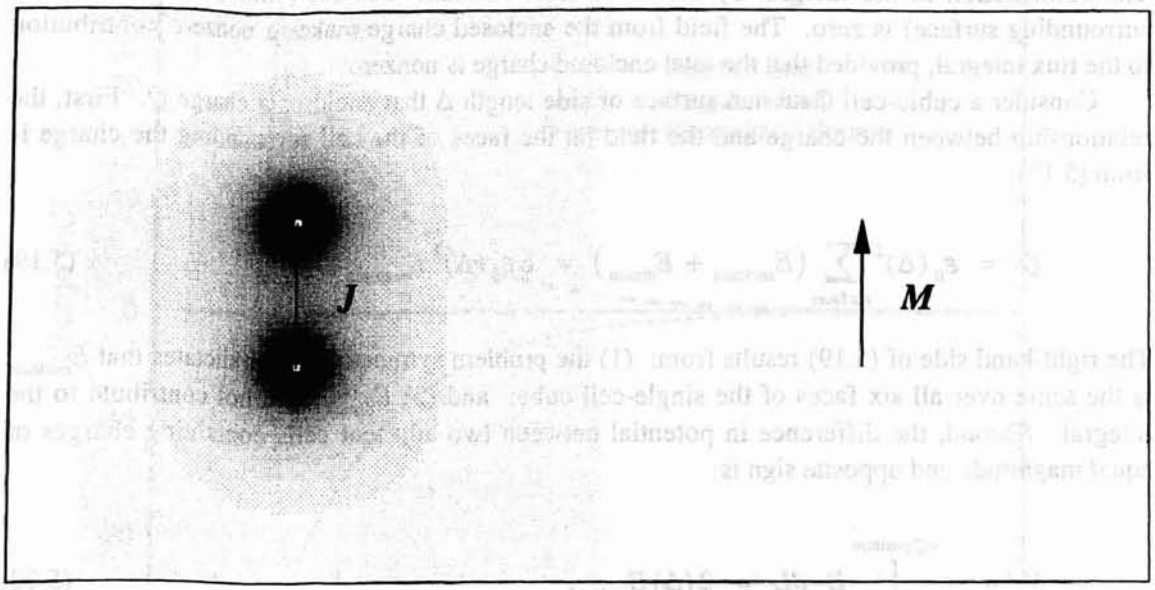


(a) Electric field magnitude.

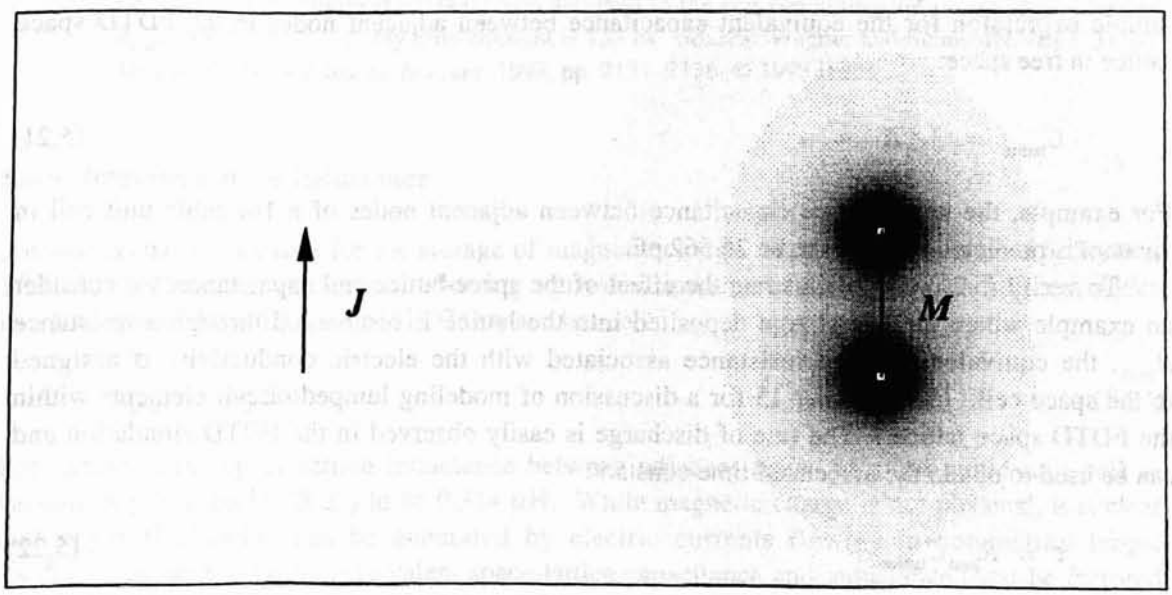


(b) Magnetic field magnitude.

Fig. 5.4 Logarithmic-grayscale visualization of the magnitude of (a) $\boldsymbol{E}$ and (b) $\boldsymbol{H}$ in Fig. 5.3 at $n = 100$ time-steps within the z-directed lattice cut plane that contains both the electric-current and magnetic-current filaments. *Source:* Wagner and Schneider, *IEEE Trans. Microwave Theory and Techniques*, 1998, pp. 2131–2136, © 1998 IEEE.



(a) Electric field magnitude.



(b) Magnetic field magnitude.

Fig. 5.5 The same as Fig. 5.4, but at $n = 300$ time-steps. Comparing with Fig. 5.4, note the departure of the radiated fields from the space lattice, leaving only static fields due to the effective transfer of electric and magnetic charge. *Source:* Wagner and Schneider, *IEEE Trans. Microwave Theory and Techniques*, 1998, pp. 2131–2136, © 1998 IEEE.

The contribution to the integral by the fields from "distant" sources (sources external to the surrounding surface) is zero. The field from the enclosed charge makes a nonzero contribution to the flux integral, provided that the total enclosed charge is nonzero.

Consider a cubic-cell Gaussian surface of side length Δ that encloses a charge Q. First, the relationship between the charge and the field on the faces of the cell surrounding the charge is from (5.15):

$$Q = \varepsilon_0 (\Delta)^2 \sum_{\text{six faces}} \left(E_{\text{enclosed}} + E_{\text{distant}} \right) = 6 \varepsilon_0 (\Delta)^2 E_{\text{enclosed}} \tag{5.19}$$

The right-hand side of (5.19) results from: (1) the problem symmetry, which dictates that E_{enclosed} is the same over all six faces of the single-cell cube; and (2) E_{distant} does not contribute to the integral. Second, the difference in potential between two adjacent cells containing charges of equal magnitude and opposite sign is:

$$V = - \int_{-Q \text{ position}}^{+Q \text{ position}} \boldsymbol{E} \cdot \boldsymbol{dL} = 2(\Delta) E_{\text{enclosed}} \tag{5.20}$$

Here, the factor of 2 on the right-hand side is a consequence of the opposite charges doubling the total E-field over the face common to both cells.

Now, combining these results, we substitute (5.19) and (5.20) into (5.17). This yields a simple expression for the equivalent capacitance between adjacent nodes in the FDTD space lattice in free space:

$$C_{\text{lattice}} = 3 \varepsilon_0 \Delta \tag{5.21}$$

For example, the space-lattice capacitance between adjacent nodes of a 1m cubic unit cell in vacuum is predicted by (5.21) to be 26.562 pF.

To verify (5.21) and to illustrate the effect of the space-lattice cell capacitance, we consider an example where electric charge deposited into the lattice is discharged through a resistance R_{load}, the equivalent lumped resistance associated with the electric conductivity σ assigned to the space cell. (See Chapter 15 for a discussion of modeling lumped circuit elements within the FDTD space lattice.) The rate of discharge is easily observed in the FDTD simulation and can be used to obtain the associated time constant:

$$\tau = R_{\text{load}} C_{\text{lattice}} \tag{5.22}$$

This yields the cell capacitance upon dividing by R_{load}.

Fig. 5.6 shows the results of one numerical experiment of this type involving the same computational domain and impulsively excited Hertzian dipole discussed in the context of Fig. 5.2 [2]. Here, the Δ = 1m cubic lattice cell associated with the source has a cell capacitance predicted by (5.21) to be 26.562 pF. When this cell is assigned the conductivity $\sigma = 0.0002$ S/m (equivalent to $R_{\text{load}} = 5{,}000\,\Omega$), a decay time constant of 133 ns results from (5.22). In fact, this agrees almost exactly with the measured time constant observed from the FDTD-calculated charge decay in Fig. 5.6, thereby verifying the cell capacitance predicted by (5.21).

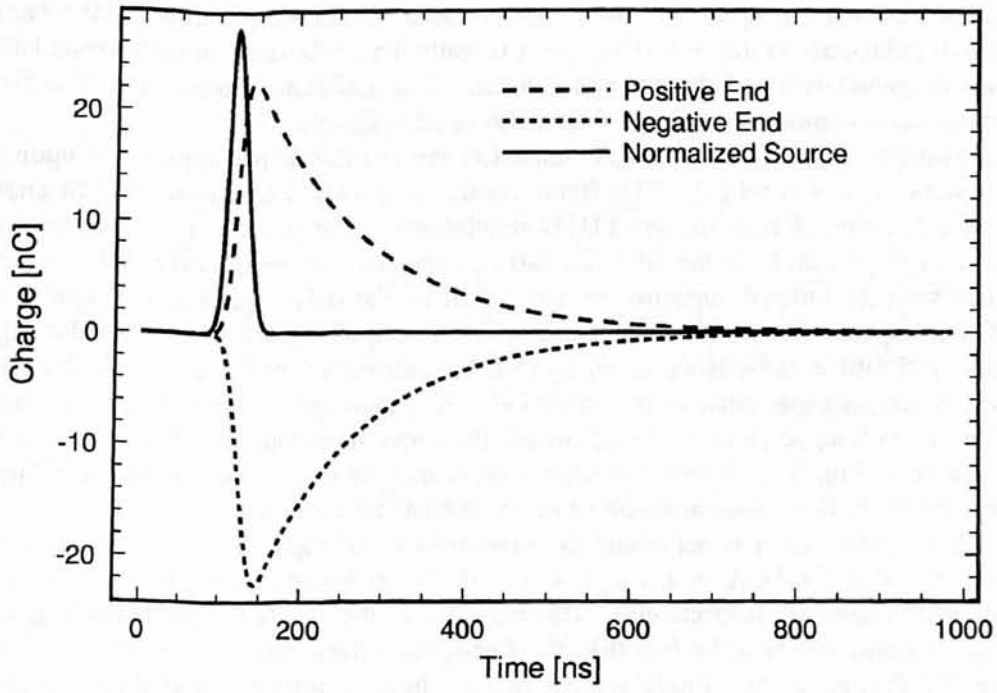


Fig. 5.6 Computational demonstration of the charge-decay effect of the intrinsic lattice capacitance, showing the charge versus time for the Hertzian-dipole electric current source used in Fig. 5.2, but with a conductivity $\sigma = 0.0002$ S/m assigned to the cell containing the source (equivalent to $R_{load} = 5{,}000\,\Omega$). The decay time constant is 133 ns. *Source:* Wagner and Schneider, *IEEE Trans. Microwave Theory and Techniques*, 1998, pp. 2131–2136, © 1998 IEEE.

5.4.5 Intrinsic Lattice Inductance

An analogous effect exists for the storage of magnetic charge. Referring to Chapter 15, Section 15.10.3, it can be shown that the following equivalent inductance exists between adjacent nodes in a free-space, three-dimensional FDTD space lattice:

$$L_{lattice} = \mu_0 \Delta / 4 \tag{5.23}$$

For example, the space-lattice inductance between adjacent nodes of a 1m cubic unit cell in vacuum is predicted by (5.23) to be 0.314 μH. While magnetic charge is not physical, it is clear that physical H-fields can be generated by electric currents flowing in conducting loops. As discussed next, both the equivalent space-lattice capacitance and inductance must be factored into FDTD simulations dealing with discrete electronic circuit elements posed as sources.

5.4.6 Impact upon FDTD Simulations of Lumped-Element Capacitors and Inductors

As discussed in Chapter 15, it is often desirable to embed a lumped-element capacitor or a lumped-element inductor at a specific point within the FDTD space lattice, especially if a combined electronic-circuit / electromagnetic-field model is useful. Chapter 15 shows that the required field-update equation at the embedding point of the circuit element is simply a modified

version of the standard Yee update. However, the modeler should realize that the total effective capacitance or inductance at the embedding point is really the combination of the intrinsic lattice capacitance or inductance and the embedded value. The additional modification of a field-update expression to compensate for this effect is discussed in Chapter 15.

As an example, Fig. 5.7 illustrates the impact of the intrinsic lattice capacitance upon the overall capacitance at a point [2]. This figure shows the potential measured between charge locations as a function of time for two FDTD simulations. (The potential is indicative of the amount of charge present.) In the first simulation, deposited charge is discharged through a conductance with no lumped capacitor present; that is, the only capacitance present is the intrinsic lattice capacitance. In the second simulation, a lumped capacitor having a value equal to the theoretical lattice capacitance given by (5.21) is introduced in the manner described in Chapter 15. Since the capacitance of the lumped element equals the intrinsic lattice capacitance, the addition of the lumped element should double the decay time constant. This is indeed the case, as seen from Fig. 5.7. Viewed another way, if one were to ignore the inherent lattice capacitance, the decay time constant would be incorrect by a factor of two.

It should be clear that it is not trivial to model the overall capacitance or inductance at a lattice point such that $C < 3\varepsilon_0\Delta$ or $L < \mu_0\Delta/4$ (i.e., the lower bounds set by the intrinsic lattice capacitance and inductance, respectively). This may require the locally defined lattice cell size, permittivity, or permeability to be less than the free-space values existing only one space cell away from the desired point. Implementing one of these options could lead to numerical instability, and must be approached with caution.

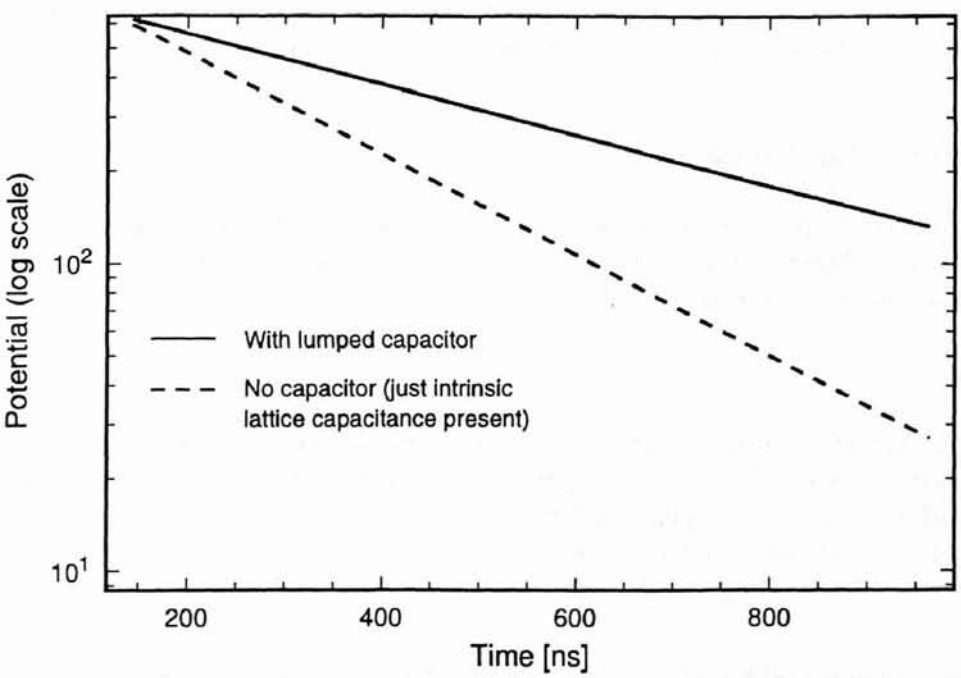


Fig. 5.7 Demonstration for the case of Fig. 5.6 of the impact of intrinsic lattice capacitance upon overall capacitance at a point. The solid line shows the charge decay versus time for an embedded lumped capacitor equal to the theoretical intrinsic lattice capacitance. The time constant for this decay is double that for the case where the lattice is empty (dashed line). *Source:* Wagner and Schneider, *IEEE Trans. Microwave Theory and Techniques*, 1998, pp. 2131–2136, © 1998 IEEE.

5.5 THE PLANE-WAVE SOURCE CONDITION

The development of the plane-wave source condition allowed the earliest engineering applications of FDTD computational electromagnetics modeling. These applications were in the defense and bioelectromagnetics areas, and involved the interaction of complex-shaped, inhomogeneous material structures with impinging pulsed or continuous-wave electromagnetic fields. Initially, many problems of these types located the material structure of interest far from the radiating antenna, where the incident illumination could be approximated by a plane wave.

In the context of FDTD solvers for Maxwell's equations, sourcing an incident plane wave in the Yee space lattice poses a number of challenges. The sourced incident wave must have:

- An arbitrary and easily specified propagation direction, polarization, time waveform, and duration;
- A planar wavefront that is perpendicular to the direction of propagation;
- A constant amplitude along any plane parallel to the wavefront.

In addition, the incident wave source must be invisible to any scattered numerical waves, allowing them to pass through the wave source without any hindrance or interaction to eventually reach the exterior region.

Yee's 1966 paper provided the original plane-wave source [3]. This approach involved inserting the incident wave as an initial condition at each $\boldsymbol{E}$ and $\boldsymbol{H}$ component location in the space lattice. With this method, all values of $E_x|^0$, $E_y|^0$, $E_z|^0$, $H_x|^{1/2}$, $H_y|^{1/2}$, and $H_z|^{1/2}$ of the incident wave throughout the lattice are prefilled by the modeler. The sign and magnitude of each initial field component is selected to give the desired wave polarization and propagation direction.

We note that, due to the Yee leapfrog time-stepping, this approach requires a $0.5\Delta t$ delay between the insertion of the initial conditions $\boldsymbol{E}|^0$ and $\boldsymbol{H}|^{1/2}$. Thus, when implementing such grid prefilling, we must account for the shift in the wave location due to its propagation during this delay. The shift depends upon the time-step used in the simulation. For example, using the magic time-step $c\Delta t = \Delta$ in a one-dimensional FDTD code implies a 0.5Δ wave motion in free space between insertion of the $\boldsymbol{E}$ and $\boldsymbol{H}$ initial conditions. Similarly, the use of $c\Delta t = \Delta/2$ in a two- or three-dimensional code implies a 0.25Δ free-space wave motion between the $\boldsymbol{E}$ and $\boldsymbol{H}$ initial conditions. These positional shifts are reduced proportionately if the grid is filled with a nondispersive dielectric medium that reduces the speed of wave motion below c. However, if the dielectric filling of the grid is dispersive (i.e., has a frequency-dependent dielectric constant), it is no longer straightforward to calculate the wave positional shift.

The initial-condition approach of [3] has two profound problems:

1. The space lattice must be enlarged to physically contain long-duration pulses or continuous sinusoids as initial conditions. This wastes computer resources, and is a classic example of a noncompact wave source.
2. A wave sourced by this method at an oblique angle in a two- or three-dimensional space lattice undergoes distortion of its wavefront as it drags against the lattice outer boundaries.

As a result of these difficulties, the initial-condition wave source currently finds only limited, specialized usage.

5.6 THE TOTAL-FIELD / SCATTERED-FIELD TECHNIQUE: IDEAS AND ONE-DIMENSIONAL FORMULATION

The *total-field / scattered-field* (TF/SF) formulation [4–6] resulted from attempts to realize a plane-wave source that avoids the difficulties caused by using either hard sources or the initial-condition approach, as discussed in the previous sections. The TF/SF technique succeeded in all respects, permitting FDTD modeling of long-duration pulsed or sinusoidal illuminations for arbitrary plane-wave propagation directions. It remains in use today for popular FDTD software.

5.6.1 Ideas

The TF/SF formulation is based on the linearity of Maxwell's equations. It assumes that the physical total electric and magnetic fields $\boldsymbol{E}_{\text{total}}$ and $\boldsymbol{H}_{\text{total}}$ can be decomposed as follows:

$$\boldsymbol{E}_{\text{total}} = \boldsymbol{E}_{\text{inc}} + \boldsymbol{E}_{\text{scat}} \qquad \boldsymbol{H}_{\text{total}} = \boldsymbol{H}_{\text{inc}} + \boldsymbol{H}_{\text{scat}} \tag{5.24a, b}$$

Here, $\boldsymbol{E}_{\text{inc}}$ and $\boldsymbol{H}_{\text{inc}}$ are the values of the incident-wave fields, assumed to be known at all points of the space lattice at all time-steps. These are the fields that would exist in vacuum, that is, if there were no materials of any sort in the modeling space. $\boldsymbol{E}_{\text{scat}}$ and $\boldsymbol{H}_{\text{scat}}$ are the values of the scattered-wave fields, which are initially unknown. These result from the interaction of the incident wave with any materials in the space lattice.

We note that the finite-difference operations of the Yee algorithm can be applied with equal validity to the incident field, the scattered field, and the total field. As shown in Fig. 5.8(a), this property permits zoning the Yee space lattice into two distinct regions: Region 1, where total fields are assumed to be stored in the computer memory; and Region 2 (surrounding Region 1), where scattered fields are assumed stored in the computer memory. Regions 1 and 2 are separated by a nonphysical virtual surface that serves to connect the fields in each region, and thereby generates the incident wave.

In Region 1, the inner zone of the space lattice, we assume that the Yee algorithm operates on total-field vector components, necessarily including the fields of the incident wave as well as those of the scattered wave. The interacting structure of interest is embedded within this region.

In Region 2, the outer zone of the space lattice, we assume that the Yee algorithm operates only on scattered-field vector components. This implies that there is no incident wave in Region 2. The outer lattice planes bounding Region 2 truncate the computation space and serve to implement an ABC. As discussed in Chapters 6 and 7, an ABC simulates the lattice extending to infinity. An ideal ABC would permit all outward-propagating numerical waves to exit the computation space with zero reflection.

TF/SF lattice zoning provides a number of key features that enhance the computational flexibility and dynamic range of the FDTD method. These features are summarized below.

1. *Arbitrary incident wave.* The connecting condition provided at the interface of the TF and SF regions, which ensures consistency of the numerical space-derivative operations across the interface, simultaneously generates an arbitrary incident plane wave in Region 1 having a user-specified time waveform and duration, angle of incidence, and angle of polarization. This connecting condition almost completely confines the numerical incident wave to Region 1, and yet is transparent to outgoing numerical scattered-wave modes, which are free to enter Region 2. Therefore, Region 2 is a well-defined scattered-field zone.

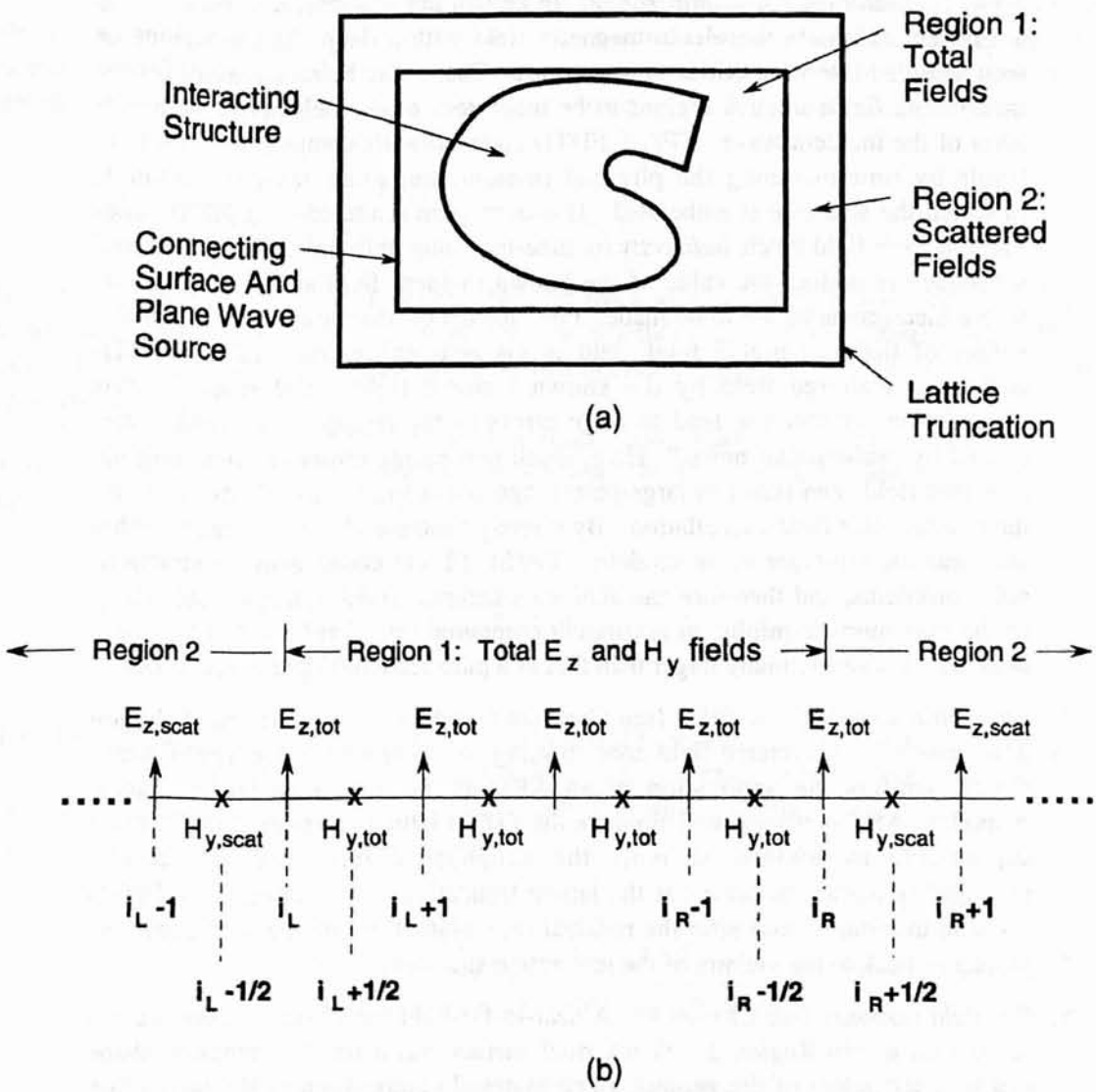


Fig. 5.8 Total-field / scattered-field zoning of the FDTD space lattice: (a) Total-field and scattered-field regions, connecting virtual surface (plane-wave source), and lattice truncation (absorbing boundary condition); (b) detail of field component locations in a one-dimensional x-directed cut through the space lattice of (a).

2. *Relatively simple programming of interaction structures.* The required continuity of total tangential $\boldsymbol{E}$ and $\boldsymbol{H}$ across the interface of dissimilar materials is automatically provided by the Yee algorithm because all materials are located in Region 1, where the total fields are time-marched. Unlike pure *scattered-field* FDTD codes, to be discussed in Section 5.10, the incident field needs to be calculated only along the simple connecting virtual surface between Regions 1 and 2. This connecting surface has a fixed rectangular shape that is independent of the geometry and material composition of the interaction structure in Region 1.

3. *Wide computational dynamic range.* In certain applications, it is important to accurately calculate the electromagnetic field within deep shadow regions or well-shielded internal cavities of a structure. There may be requirements for the measurable fields in such regions to be many tens of decibels below the power level of the incident wave. TF/SF FDTD codes *directly* compute such low field levels by time-marching the physical (measurable) total fields in Region 1, in which the structure is embedded. However, pure scattered-field FDTD codes calculate low field levels *indirectly* by time-marching only the scattered field, and subsequently adding the value of the known incident field at the lattice points where measurements are to be made. For such codes, the only way to obtain low values of the measurable total field is via near cancellation of the FDTD-computed scattered field by the known incident field. Unfortunately, this cancellation process can lead to large errors in the required total-field values caused by "subtraction noise." Here, small-percentage errors in calculating the scattered fields can result in large-percentage errors in the total fields which are the residue after field cancellation. By directly time-marching total fields within and near the structure to be modeled, TF/SF FDTD codes avoid subtraction-noise problems, and therefore can achieve a computational dynamic range (ratio of the maximum-to-minimum accurately computed total-field level in the space lattice) that is substantially larger than that of a pure scattered-field FDTD code.

4. *Absorbing boundary condition* (see Chapters 6 and 7). The provision of Region 2, a well-defined scattered-field zone forming the outer part of the TF/SF space lattice, permits the application of an ABC at the outermost lattice planes. A modern ABC can very well simulate the FDTD lattice extending to infinity by suppressing by 99.99% or more the nonphysical reflection of outward-propagating numerical waves at the lattice truncations. This permits the FDTD solution to remain valid after the residual reflections from the lattice truncations propagate back to the vicinity of the interaction structure.

5. *Far-field response* (see Chapter 8). A near-to-far-field transformation surface can be located within Region 2. This virtual surface has a fixed rectangular shape that is independent of the geometry and material composition of the interaction structure modeled in Region 1. FDTD-computed tangential scattered-field data along this surface are weighted by the free-space Green function and then summed to obtain the complete far-field radiation or bistatic scattering pattern of the structure being modeled.

5.6.2 One-Dimensional Formulation

The virtual surface constituting the interface of Regions 1 and 2 contains $\boldsymbol{E}$ and $\boldsymbol{H}$ components, which, according to the Yee algorithm, require the formulation of various field component spatial differences to advance one time-step. When the required spatial difference is taken across the interface plane, there arises a problem of consistency. That is, on the Region-1 side of the interface, the field to be used in the difference expression is assumed to be a total field, whereas on the Region-2 side of the interface, the field to be used in the difference expression is assumed to be a scattered field. It would be inconsistent to perform an arithmetic difference between scattered- and total-field values.

To precisely illustrate this problem, consider the one-dimensional linear grid of E_z and H_y field components shown in Fig. 5.8(b). This also represents an x-directed cut through the two-dimensional TM_z space grid of Fig. 5.8(a). Assume that all points in the one-dimensional grid are in free space. Then, the Yee algorithm to time-step the E-field is given at any gridpoint by

$$E_z\Big|_i^{n+1} = E_z\Big|_i^n + \frac{\Delta t}{\varepsilon_0 \Delta x}\left(H_y\Big|_{i+1/2}^{n+1/2} - H_y\Big|_{i-1/2}^{n+1/2}\right) \tag{5.25}$$

Note that (5.25) is consistent and completely valid when all three field components on the right-hand side are contained within a single grid region, whether total-field or scattered-field. This is due to the linearity of Maxwell's equations, which operate equally well on the total field and the scattered field.

The validity of applying (5.25) within a single grid region is now shown. For example, if points $i - 1/2$, i, and $i + 1/2$ are each in Region 1 of Fig. 5.8(b), the application of (5.25) can be written in an equivalent manner as

$$E_{z,\,\text{total}}\Big|_i^{n+1} = \underbrace{E_{z,\,\text{total}}\Big|_i^n}_{\substack{\text{assumed stored in} \\ \text{computer memory}}} + \frac{\Delta t}{\varepsilon_0 \Delta x}\underbrace{\left(H_{y,\,\text{total}}\Big|_{i+1/2}^{n+1/2} - H_{y,\,\text{total}}\Big|_{i-1/2}^{n+1/2}\right)}_{\text{assumed stored in computer memory}} \tag{5.26a}$$

Here, total-field E_z and H_y components that are assumed to be stored in the computer memory are manipulated to time-step a total E_z component, which is then returned to computer memory. Similarly, if points $i - 1/2$, i, and $i + 1/2$ are each in Region 2 of Fig. 5.8(b), the application of (5.25) can be written in an equivalent manner as

$$E_{z,\,\text{scat}}\Big|_i^{n+1} = \underbrace{E_{z,\,\text{scat}}\Big|_i^n}_{\substack{\text{assumed stored in} \\ \text{computer memory}}} + \frac{\Delta t}{\varepsilon_0 \Delta x}\underbrace{\left(H_{y,\,\text{scat}}\Big|_{i+1/2}^{n+1/2} - H_{y,\,\text{scat}}\Big|_{i-1/2}^{n+1/2}\right)}_{\text{assumed stored in computer memory}} \tag{5.26b}$$

Here, scattered-field E_z and H_y components that are assumed to be stored in the computer memory are manipulated to time-step a scattered E_z component, which is then returned to computer memory.

Now, consider applying (5.25) to time-step the E_z component located at gridpoint i_L, the left interface between Regions 1 and 2. Assume that this E_z component is stored in computer memory as a total field. Direct application of (5.25) yields

$$E_{z,\,\text{total}}\Big|_{i_L}^{n+1} = \underbrace{E_{z,\,\text{total}}\Big|_{i_L}^n}_{\substack{\text{assumed stored in} \\ \text{computer memory}}} + \frac{\Delta t}{\varepsilon_0 \Delta x}\underbrace{\left(H_{y,\,\text{total}}\Big|_{i_L+1/2}^{n+1/2} - H_{y,\,\text{scat}}\Big|_{i_L-1/2}^{n+1/2}\right)}_{\text{assumed stored in computer memory}} \tag{5.27}$$

In (5.27), we note that the H_y component at $i_L + 1/2$ (just to the right of E_z) is assumed to be stored in the computer memory as a total field, whereas the H_y component at $i_L - 1/2$ (just to the left of E_z) is assumed to be stored in computer memory as a scattered field. Thus, (5.27) is inconsistent and incorrect, since it subtracts unlike H_y quantities on the right-hand side. For consistency, we must subtract total-H_y from total-H_y to advance total-E_z in time, as in (5.26a).

However, (5.27) can be made consistent and correct simply by adding another term that is a function of the assumed-known incident wave:

$$E_{z,\,\text{total}}\Big|_{i_L}^{n+1} = \underbrace{E_{z,\,\text{total}}\Big|_{i_L}^{n}}_{\substack{\text{assumed stored in}\\ \text{computer memory}}} + \frac{\Delta t}{\varepsilon_0 \Delta x}\underbrace{\left(H_{y,\,\text{total}}\Big|_{i_L+1/2}^{n+1/2} - H_{y,\,\text{scat}}\Big|_{i_L-1/2}^{n+1/2}\right)}_{\text{assumed stored in computer memory}} - \frac{\Delta t}{\varepsilon_0 \Delta x}\underbrace{H_{y,\,\text{inc}}\Big|_{i_L-1/2}^{n+1/2}}_{\text{assumed known}} \tag{5.28}$$

since

$$-H_{y,\,\text{scat}}\Big|_{i_L-1/2}^{n+1/2} - H_{y,\,\text{inc}}\Big|_{i_L-1/2}^{n+1/2} = -H_{y,\,\text{total}}\Big|_{i_L-1/2}^{n+1/2} \tag{5.29}$$

This can be done as a correction implemented only at gridpoint i_L after first applying the generic Yee algorithm of (5.25) at *all* E_z locations in the grid. The correction "fixes" the only gridpoint (i_L), where inconsistency and error arises during the application of (5.25), and avoids breaking up the E_z time-stepping loop to achieve maximum speed on a vectorizing machine.

However, we are not quite finished dealing with consistency problems at the left interface between Regions 1 and 2. Consider the generic Yee algorithm to time-step the H_y field in free space at any gridpoint in Fig. 5.8(b):

$$H_y\Big|_{i-1/2}^{n+1/2} = H_y\Big|_{i-1/2}^{n-1/2} + \frac{\Delta t}{\mu_0 \Delta x}\left(E_z\Big|_i^n - E_z\Big|_{i-1}^n\right) \tag{5.30}$$

In the manner of (5.26a, b), it can be easily shown that (5.30) is consistent and valid when all three field components on the right-hand side are contained within a single grid region, whether Region 1 or Region 2. However, upon applying (5.30) to time-step the H_y component at gridpoint $i_L - 1/2$, an inconsistency arises. Assume that this H_y component is stored in the computer memory as a scattered field. Direct application of (5.30) yields

$$H_{y,\,\text{scat}}\Big|_{i_L-1/2}^{n+1/2} = \underbrace{H_{y,\text{scat}}\Big|_{i_L-1/2}^{n-1/2}}_{\substack{\text{assumed stored in}\\ \text{computer memory}}} + \frac{\Delta t}{\mu_0 \Delta x}\underbrace{\left(E_{z,\,\text{total}}\Big|_{i_L}^{n} - E_{z,\,\text{scat}}\Big|_{i_L-1}^{n}\right)}_{\text{assumed stored in computer memory}} \tag{5.31}$$

In (5.31), we note that the E_z component at i_L (just to the right of H_y) is assumed to be stored in the computer memory as a total field, whereas the E_z component at $i_L - 1$ (just to the left of H_y) is assumed to be stored in computer memory as a scattered field. Thus, (5.31) is inconsistent and incorrect, since it subtracts unlike E_z quantities on the right-hand side. For consistency, we must subtract scattered-E_z from scattered-E_z to advance scattered-H_y in time.

However, (5.31) can be made consistent and correct simply by adding another term that is a function of the assumed-known incident wave:

$$H_{y,\,\text{scat}}\Big|_{i_L-1/2}^{n+1/2} = \underbrace{H_{y,\text{scat}}\Big|_{i_L-1/2}^{n-1/2}}_{\substack{\text{assumed stored in}\\ \text{computer memory}}} + \frac{\Delta t}{\mu_0 \Delta x}\underbrace{\left(E_{z,\,\text{total}}\Big|_{i_L}^{n} - E_{z,\,\text{scat}}\Big|_{i_L-1}^{n}\right)}_{\text{assumed stored in computer memory}} - \frac{\Delta t}{\mu_0 \Delta x}\underbrace{E_{z,\,\text{inc}}\Big|_{i_L}^{n}}_{\text{assumed known}} \tag{5.32}$$

since

$$E_{z,\,\text{total}}\Big|_{i_L}^{n} - E_{z,\,\text{inc}}\Big|_{i_L}^{n} = E_{z,\,\text{scat}}\Big|_{i_L}^{n} \tag{5.33}$$

This can be done as a correction implemented only at gridpoint $i_L - 1/2$ after first applying the generic Yee algorithm of (5.30) at *all* H_y locations in the grid. This correction "fixes" the only gridpoint $(i_L - 1/2)$ where inconsistency and error arises during the application of (5.30), and avoids needlessly breaking up the H_y time-stepping loop.

Similar procedures are implemented to preserve consistency of the numerical differentiation at the right interface between Regions 1 and 2. As shown in Fig. 5.8(b), we assume that this interface is located at the E_z component at gridpoint i_R, and further assume that this E_z is a total field. The consistency condition analogous to (5.28) is given in a shorthand notation by

$$E_z\Big|_{i_R}^{n+1} = \left\{E_z\Big|_{i_R}^{n+1}\right\}_{(5.25)} + \frac{\Delta t}{\varepsilon_0 \Delta x} \underbrace{H_{y,\,\text{inc}}\Big|_{i_R+1/2}^{n+1/2}}_{\text{assumed known}} \tag{5.34}$$

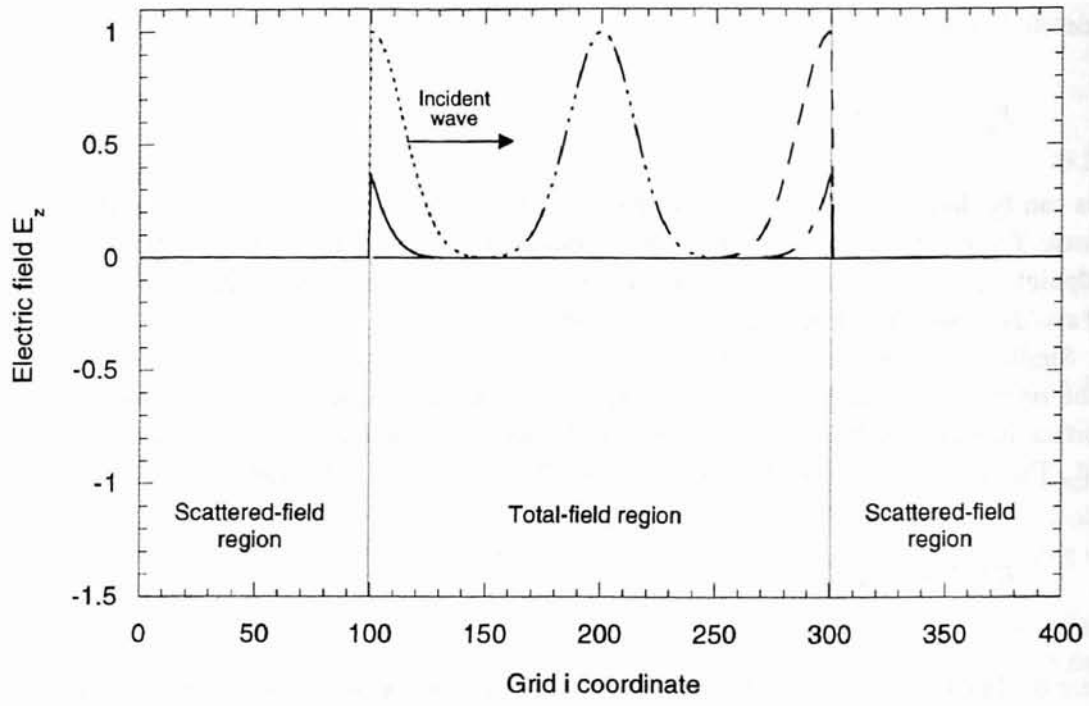
where the bracket denotes that the generic E_z time-stepping operations of (5.25) are implemented before adding the incident-wave correction term. The consistency condition analogous to (5.32) is given by

$$H_y\Big|_{i_R+1/2}^{n+1/2} = \left\{H_y\Big|_{i_R+1/2}^{n+1/2}\right\}_{(5.30)} + \frac{\Delta t}{\mu_0 \Delta x} \underbrace{E_{z,\,\text{inc}}\Big|_{i_R}^{n}}_{\text{assumed known}} \tag{5.35}$$

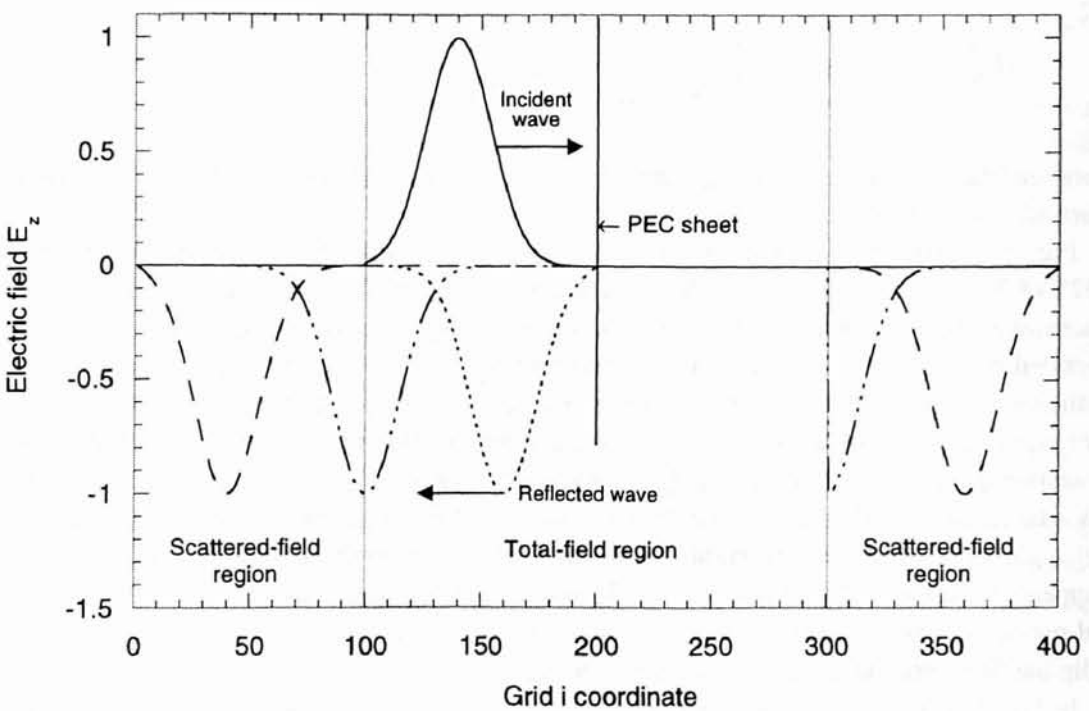
where the bracket denotes that the generic H_y time-stepping operations of (5.30) are implemented before adding the incident-wave correction term.

Fig. 5.9 illustrates the overall effect of implementing the TF/SF formulation of (5.28), (5.32), (5.34), and (5.35). In Fig. 5.9(a), we assume that free space exists everywhere in a one-dimensional linear grid with $c\Delta t = \Delta x$. A wave having a Gaussian-pulse time behavior is generated at the left TF/SF interface at gridpoint $i_L = 100$. This wave propagates virtually unidirectionally within the grid. Specifically, to the right of i_L in the total-field region, it propagates in the $+x$-direction with full amplitude and no distortion; whereas to the left of i_L in the scattered-field region, there is only a very small leakage wave that is suppressed by more than a factor of 10^5. Eventually, the desired $+x$-directed wave passes entirely through the total-field region and arrives at the right TF/SF interface at gridpoint $i_R = 300$. Here, the wave disappears from the grid with negligible reflection and negligible transmission into the scattered-field region beyond i_R. Overall, in the absence of a scatterer in the total-field region, there are negligible fields present in surrounding scattered-field regions.

In Fig. 5.9(b), we assume that a PEC sheet enforcing $E_z = 0$ is centered within the total-field region at gridpoint $i = 200$. This sheet generates a $-x$-directed reflected wave that eventually passes through the left TF/SF interface at $i_L = 100$. This passage is accomplished with negligible retroreflection and essentially complete transmission — the left TF/SF interface is transparent to the wave reflected by the PEC sheet. To the right of the PEC sheet in the total-field region, there exists zero field for all time-steps, as expected.



(a)



(b)

Fig. 5.9 Snapshots of a $+x$-propagating Gaussian pulse in a one-dimensional total-field / scattered-field Yee grid for $c\Delta t = \Delta x$, showing the action of TF/SF grid zoning. (a) Free space everywhere in the grid; and (b) PEC mirror placed within the total-field region.

However, to the right of the PEC sheet in the scattered-field region, there eventually appears a $+x$-directed wave that is identical to the incident wave, but negatively signed. This is *not* an error. In fact, a scattered wave properly exists within the shadow region of the PEC sheet by the assumptions of (5.24). Where the total field is identically zero, as in the shadow region of the PEC sheet, the scattered and incident fields must be equal and opposite so that their sum (i.e., the total field) is zero. Thus, a negative-amplitude scattered wave must appear in the scattered-field region to the right of $i_R = 300$ after sufficient time has passed to permit the incident wave to propagate this far.

5.7 TWO-DIMENSIONAL FORMULATION OF THE TF/SF TECHNIQUE

The TF/SF technique can be extended in a straightforward manner to model two-dimensional TM and TE problems. Here, this technique permits generation of an incident plane wave having an arbitrary time waveform, duration, and propagation direction in the Yee space lattice. As an example of the extension of the TF/SF formulation to FDTD modeling in two dimensions, this section discusses in detail the algorithmic modifications required for the TM_z case.

5.7.1 Consistency Conditions

Fig. 5.10 illustrates the extension of the TF/SF technique to a TM_z grid having the zoning shown in Fig. 5.8(a). Referring to Fig. 5.10(a), consider first the Region-1 / Region-2 interface located at $j = j_0$ where the total-field components $E_{z,\,\text{total}}$ and $H_{y,\,\text{total}}$ are located. To time-step the $E_{z,\,\text{total}}|_{i,\,j_0}$ components (indicated by circled dots), we must know $H_{x,\,\text{total}}|_{i,\,j_0+1/2}$ and $H_{x,\,\text{total}}|_{i,\,j_0-1/2}$. Clearly, the first $H_{x,\,\text{total}}$ is known and stored in the computer memory, since gridpoint $(i, j_0+1/2)$ is in the total-field zone. However, the second $H_{x,\,\text{total}}$ is not stored in computer memory, since $(i, j_0-1/2)$ is in the scattered-field zone. Only $H_{x,\,\text{scat}}|_{i,\,j_0-1/2}$ is available in memory. Yet, since

$$H_{x,\,\text{total}}\Big|_{i,\,j_0-1/2} = H_{x,\,\text{scat}}\Big|_{i,\,j_0-1/2} + H_{x,\,\text{inc}}\Big|_{i,\,j_0-1/2} \tag{5.36}$$

for all time-steps, we can modify the normal Yee time-stepping relation for $E_{z,\,\text{total}}|_{i,\,j_0}$ to achieve consistency. For simplicity, we assume a square-cell grid. Using the shorthand notation established in (5.34) and (5.35), we have the following consistency condition to be applied to all of the E_z components (circled dots) located at the front face of Region 1 of Fig. 5.10(a):

Front Face of Region 1 $(j = j_0;\ \ i = i_0, \ldots, i_1)$

$$E_z\Big|_{i,\,j_0}^{n+1} = \left\{E_z\Big|_{i,\,j_0}^{n+1}\right\}_{(3.41\text{a})} + \underbrace{\frac{\Delta t}{\varepsilon_0 \Delta} H_{x,\,\text{inc}}\Big|_{i,\,j_0-1/2}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.37a}$$

Here, the bracket denotes that the generic E_z time-stepping operations of (3.41a) are implemented before adding the incident-wave correction term. Implicit in (5.37a) is the assumption that the TF / SF interface is located in source-free vacuum.

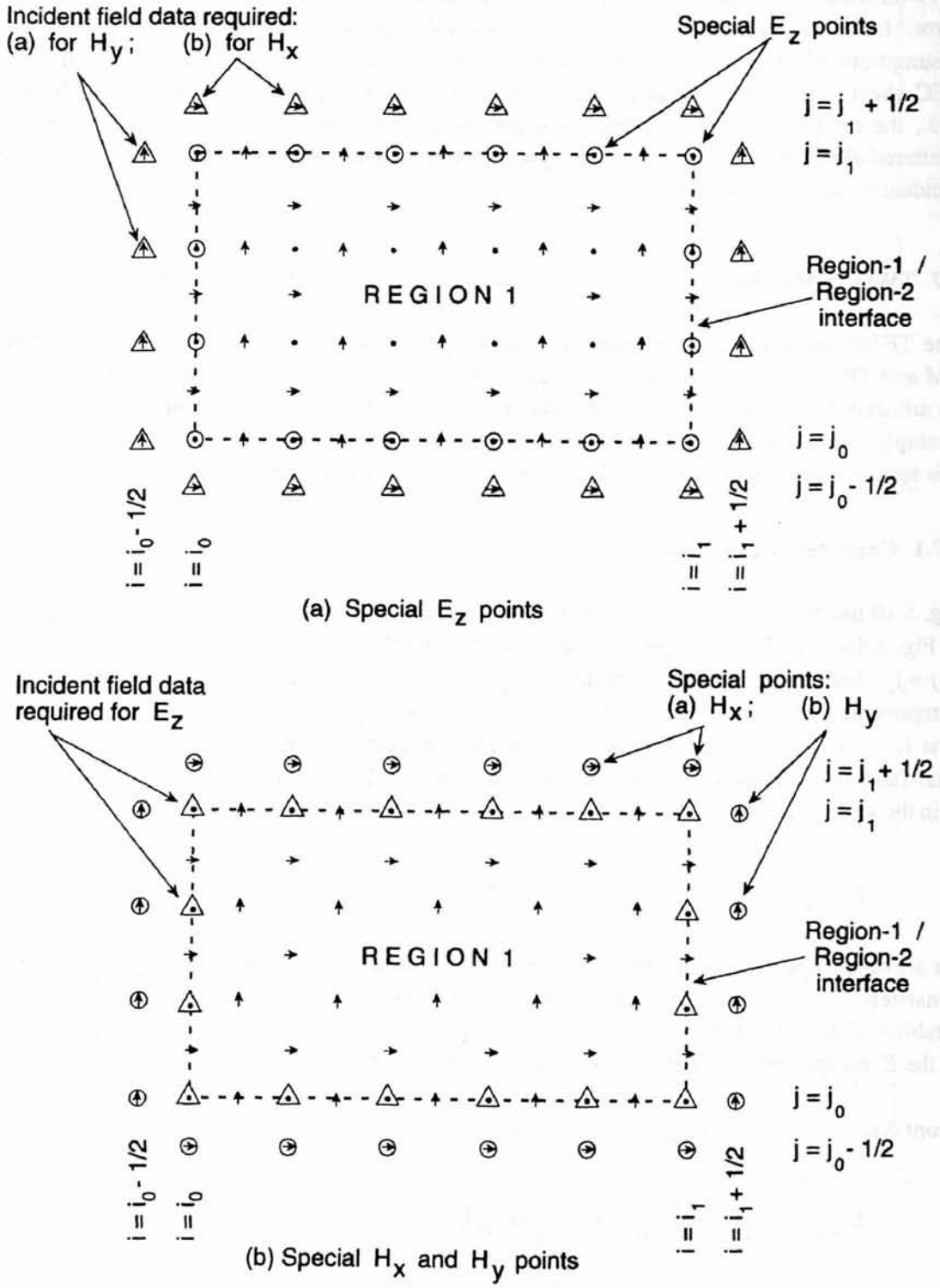


Fig. 5.10 Field components in the two-dimensional TM_z grid at the interface of the total-field and scattered-field regions used for the incident-wave source condition.

The remaining E_z components located on the Region-1 / Region-2 interface, also indicated by circled dots in Fig. 5.10(a), can be treated in an analogous manner to achieve consistency. The following are the required operations:

Back Face of Region 1 $(j = j_1 \,;\; i = i_0, \ldots, i_1)$

$$E_z\Big|_{i,j_1}^{n+1} = \left\{E_z\Big|_{i,j_1}^{n+1}\right\}_{(3.41a)} \underbrace{- \frac{\Delta t}{\varepsilon_0 \Delta} H_{x,\,\mathrm{inc}}\Big|_{i,\,j_1+1/2}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.37b}$$

Left Face of Region 1 $(i = i_0 \,;\; j = j_0, \ldots, j_1)$

$$E_z\Big|_{i_0,j}^{n+1} = \left\{E_z\Big|_{i_0,j}^{n+1}\right\}_{(3.41a)} \underbrace{- \frac{\Delta t}{\varepsilon_0 \Delta} H_{y,\,\mathrm{inc}}\Big|_{i_0-1/2,\,j}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.37c}$$

Right Face of Region 1 $(i = i_1 \,;\; j = j_0, \ldots, j_1)$

$$E_z\Big|_{i_1,j}^{n+1} = \left\{E_z\Big|_{i_1,j}^{n+1}\right\}_{(3.41a)} \underbrace{+ \frac{\Delta t}{\varepsilon_0 \Delta} H_{y,\,\mathrm{inc}}\Big|_{i_1+1/2,\,j}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.37d}$$

To implement the consistency conditions of (5.37), data are required for the incident H_x and H_y field components at grid locations 0.5Δ outside of the interface. These are indicated by triangles in Fig. 5.10(a). We note that proper treatment of the four interface corner points at $E_z|_{i_0,\,j_0}$, $E_z|_{i_0,\,j_1}$, $E_z|_{i_1,\,j_0}$, and $E_z|_{i_1,\,j_1}$ (where two adjacent H-components are in Region 2) is provided by the overlapping operations implied by (5.37a–d).

Referring to Fig. 5.10(b), we see that a consistency problem also exists for the $H_{x,\,\mathrm{scat}}$ and $H_{y,\,\mathrm{scat}}$ components (indicated by circled arrows) that are located 0.5Δ outside of the Region-1 / Region-2 interface. To properly time-step each of these components, we must know $E_{z,\,\mathrm{scat}}$ located 0.5Δ to each side. Yet, on one side of each H-component of interest is $E_{z,\,\mathrm{total}}$ located on the Region-1 / Region-2 interface.

The solution of the consistency problem for $H_{x,\,\mathrm{scat}}$ and $H_{y,\,\mathrm{scat}}$ is analogous to the development leading to (5.37) for $E_{z,\,\mathrm{total}}$. Knowing that

$$E_{z,\,\mathrm{total}}\Big|_{i,j} = E_{z,\,\mathrm{scat}}\Big|_{i,j} + E_{z,\,\mathrm{inc}}\Big|_{i,j} \tag{5.38}$$

for all time-steps, we can slightly modify the normal Yee time-stepping relation for the affected H-components to achieve consistency. For example, consider the calculation of the scattered H_x components (circled rightward arrows) that are located just outside the front face of Region 1. Again assuming a square-cell grid, we have

Outside Front Face of Region 1 $(j = j_0 - 1/2;\ \ i = i_0, \ldots, i_1)$

$$H_x\Big|_{i,j_0-1/2}^{n+1/2} = \left\{H_x\Big|_{i,j_0-1/2}^{n+1/2}\right\}_{(3.41b)} + \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{z,\,\mathrm{inc}}\Big|_{i,\,j_0}^{n}}_{\text{assumed known correction term}} \tag{5.39a}$$

Similar to our previous shorthand notation, the bracket denotes that the generic H_x time-stepping operations of (3.41b) are implemented before adding the incident-wave correction term. Implicit in (5.39a) is the assumption that the TF/SF interface is located in source-free vacuum.

The remaining H_x and H_y components located just outside of the Region-1 / Region-2 interface, also indicated by circled arrows in Fig. 5.10(b), can be treated in an analogous manner to achieve consistency. The following are the required operations:

Outside Back Face of Region 1 $(j = j_1 + 1/2;\ \ i = i_0, \ldots, i_1)$

$$H_x\Big|_{i,j_1+1/2}^{n+1/2} = \left\{H_x\Big|_{i,j_1+1/2}^{n+1/2}\right\}_{(3.41b)} - \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{z,\,\mathrm{inc}}\Big|_{i,\,j_1}^{n}}_{\text{assumed known correction term}} \tag{5.39b}$$

Outside Left Face of Region 1 $(i = i_0 - 1/2;\ \ j = j_0, \ldots, j_1)$

$$H_y\Big|_{i_0-1/2,j}^{n+1/2} = \left\{H_y\Big|_{i_0-1/2,j}^{n+1/2}\right\}_{(3.41c)} - \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{z,\,\mathrm{inc}}\Big|_{i_0,\,j}^{n}}_{\text{assumed known correction term}} \tag{5.39c}$$

Outside Right Face of Region 1 $(i = i_1 + 1/2;\ \ j = j_0, \ldots, j_1)$

$$H_y\Big|_{i_1+1/2,j}^{n+1/2} = \left\{H_y\Big|_{i_1+1/2,j}^{n+1/2}\right\}_{(3.41c)} + \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{z,\,\mathrm{inc}}\Big|_{i_1,\,j}^{n}}_{\text{assumed known correction term}} \tag{5.39d}$$

To implement the consistency conditions of (5.39), data are required for the incident E_z components at grid locations exactly on the interface. These are indicated by triangles in Fig. 5.10(b).

Together, the consistency conditions of (5.37) and (5.39) properly connect the assumed total-field and scattered-field grid regions for the two-dimensional TM_z mode. A *direct result* of implementing these conditions is the generation of a high-quality numerical plane wave in the total-field zone, Region 1, with little field leakage appearing in the scattered-field zone, Region 2. The wave generated in Region 1 has all of the properties specified for $\boldsymbol{E}_{\mathrm{inc}}$ and $\boldsymbol{H}_{\mathrm{inc}}$ in the various correction terms used in (5.37) and (5.39): time waveform, duration, and angle of incidence. Further, the interface between Regions 1 and 2 is transparent to all outgoing numerical scattered waves, permitting them to pass without reflection or attenuation into Region 2. Overall, this TF/SF technique meets all of the requirements for the plane-wave source condition.

5.7.2 Calculation of the Incident Field

Referring to Fig. 5.10 and the previous discussion, we see that data for the incident E_z, H_x, and H_y components at or near the interface of Regions 1 and 2 are required to implement the TF/SF consistency conditions for the TM_z mode. A simple approach to calculate these data for arbitrary wave angles of incidence is now described. This approach is based upon a table look-up procedure, which avoids the need to compute large numbers of mathematical functions (sinusoids or exponentials) to describe the space-time behavior of the incident wave.

Coordinate Origins

Consider first the two-dimensional TM_z simulation of an incident plane wave. The wave is assumed to propagate with a wavevector $\boldsymbol{k}_{inc}$ that is oriented at the angle ϕ relative to the $+x$-axis of the FDTD grid, where $0^\circ \le \phi \le 90^\circ$.

Fig. 5.11(a) is a schematic diagram that is useful for analysis of this case. A key point is that connecting conditions (5.37) and (5.39) generate a numerical analog of the desired plane wave that appears to originate from the front-left corner of Region 1. We designate this point as Origin O_1 with the grid coordinates (i_0, j_0). In effect, O_1 is the first total-field gridpoint that is "contacted" by the incident wavefront. For a square-cell Cartesian grid, the other field components in Region 1 are contacted by the wavefront after n_{delay} time-steps, where

$$n_{delay} = d\Delta \Big/ \left[\tilde{v}_p(\phi)\Delta t\right] \tag{5.40}$$

Here, $\tilde{v}_p(\phi)$ is the numerical phase velocity of the incident wave at the propagation angle ϕ, as discussed in Chapter 4, Section 4.5. Also, d is the distance in grid cells along the wavevector from Origin O_1 to a perpendicular dropped to the wavevector from the location of the field component in question. Distance d can be expressed conveniently as

$$d = \hat{\boldsymbol{k}}_{inc} \cdot \boldsymbol{r}_{comp} \tag{5.41}$$

where $\hat{\boldsymbol{k}}_{inc}$ is the unit incident wavevector given by

$$\hat{\boldsymbol{k}}_{inc} = \hat{\boldsymbol{x}}\cos\phi + \hat{\boldsymbol{y}}\sin\phi \tag{5.42}$$

and $\boldsymbol{r}_{comp}$ is the position vector from O_1 to the location of the field vector component of interest:

$$\boldsymbol{r}_{comp} = \left(i_{comp} - i_0\right)\hat{\boldsymbol{x}} + \left(j_{comp} - j_0\right)\hat{\boldsymbol{y}} \tag{5.43a}$$

Note that i_{comp} and j_{comp} can be either integers or integers $\pm 1/2$, since a field component can be positioned at a half-cell location.

The delay distance d specified by (5.41) is necessary to calculate the incident field at the desired grid location. However, for $\phi > 90^\circ$, it is clear that the incident wavefront no longer makes an initial contact with Region 1 at Origin O_1. In fact, as the wavevector angle is rotated through $+360^\circ$, three new points of initial contact are defined, representing the remaining three corners of Region 1: $90^\circ < \phi \le 180^\circ$ – Origin O_2 at (i_1, j_0); $180^\circ < \phi \le 270^\circ$ – Origin O_3 at (i_1, j_1); and $270^\circ < \phi \le 360^\circ$ – Origin O_4 at (i_0, j_1).

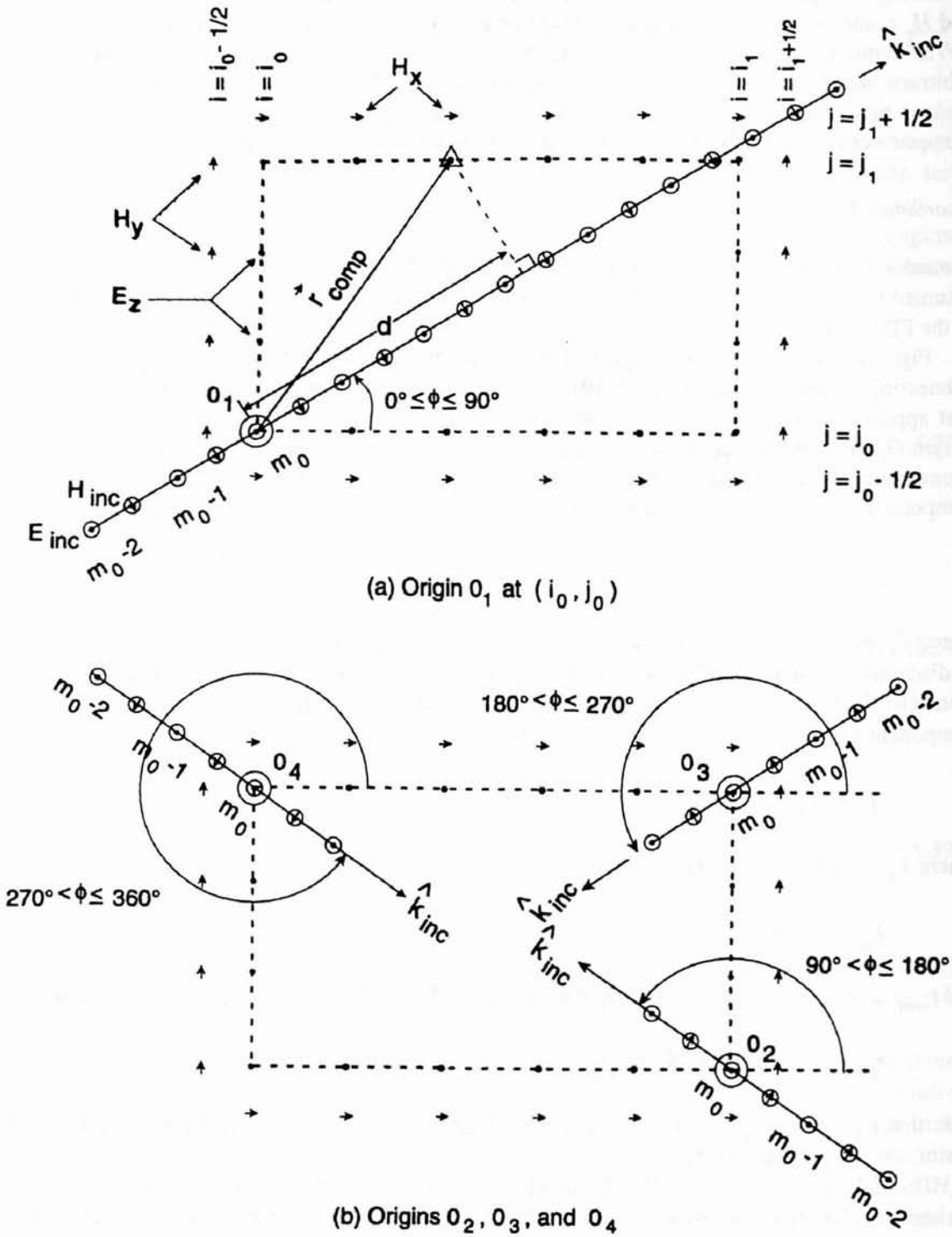


Fig. 5.11 Coordinate origins for efficiently calculating the incident field in a two-dimensional TM_z FDTD grid zoned into total-field and scattered-field regions.

These points are shown in Fig. 5.11(b). For these cases, the position vector $\boldsymbol{r}_{\text{comp}}$ of (5.43a) must be modified. We have

$$90^\circ < \phi \le 180^\circ \qquad \boldsymbol{r}_{\text{comp}} = \left(i_{\text{comp}} - i_1\right)\hat{x} + \left(j_{\text{comp}} - j_0\right)\hat{y} \tag{5.43b}$$

$$180^\circ < \phi \le 270^\circ \qquad \boldsymbol{r}_{\text{comp}} = \left(i_{\text{comp}} - i_1\right)\hat{x} + \left(j_{\text{comp}} - j_1\right)\hat{y} \tag{5.43c}$$

$$270^\circ < \phi \le 360^\circ \qquad \boldsymbol{r}_{\text{comp}} = \left(i_{\text{comp}} - i_0\right)\hat{x} + \left(j_{\text{comp}} - j_1\right)\hat{y} \tag{5.43d}$$

Selection of one of these four cases of coordinate origins and corresponding position vectors can be performed automatically in a computer program upon specification of ϕ by the user.

Generation of Look-Up Table

It is clear that $\boldsymbol{E}_{\text{inc}}$ and $\boldsymbol{H}_{\text{inc}}$ can be calculated at any grid location using an analytical expression for the space-time behavior of the incident wave (e.g., a sine function or an exponential), once n_{delay} and d are determined for that location from (5.40) and (5.41). However, a large amount of computer arithmetic is needed to calculate the incident wavefunction to implement the TF/SF connecting condition. For example, approximately 240,000 computations of a sine or exponential function are needed each time-step for a six-sided Region-1 / Region-2 interface in a $100 \times 100 \times 100$-cell space lattice. (This is still much less than the *12 million* computations of the same type that would be required each time-step for a pure scattered-field formulation, as discussed in Section 5.10.) To reduce the burden of computing many sines or exponentials each time-step, an approach based upon a table look-up procedure has been used successfully in FDTD software. This approach is now summarized.

Referring to Fig. 5.11(a), we assume that an auxiliary one-dimensional FDTD grid is positioned so that one of its E-field components, $E_{\text{inc}}|_{m_0}$, coincides with origin O_1 of the main two-dimensional TM_z grid. Further, the auxiliary grid is assumed to be oriented parallel to the incident wavevector in the main grid. This allows us to use the auxiliary grid to calculate the space-time variation of the incident fields in the main grid. With the delay distance d known for a point in the main grid, the incident field at this point can be obtained by interpolating similar field values adjacent to point $m_0 + d$ in the auxiliary grid. As Fig. 5.11(a) shows, this has the geometrical interpretation of dropping a perpendicular from the point of interest in the main two-dimensional TM_z grid to the auxiliary one-dimensional grid, and then interpolating about the foot of the dropped perpendicular.

Note that this procedure requires computation of the incident-wave time dependence at only a single point in the auxiliary grid, a hard source at $m_0 - 2$:

$$E_{\text{inc}}\big|_{m_0-2}^{n} = E_0\, g(n\Delta t) \tag{5.44}$$

where g is an arbitrary time function. Thus, *only one* sine or exponential function must be computed per time-step to excite the auxiliary grid, thereby exciting the complete two-dimensional main TM_z grid.

The system of finite-difference equations for the auxiliary grid is easily derived from (3.41a) and (3.41c). Assuming positive directions for E_{inc} and H_{inc} as shown in Fig. 5.11(a), and the same Δx and Δt in both the auxiliary and main grids, we have

$$E_{\text{inc}}\Big|_m^{n+1} = E_{\text{inc}}\Big|_m^n + \frac{\Delta t}{\left[\dfrac{\tilde{v}_p(\phi = 0^\circ)}{\tilde{v}_p(\phi)}\right]\varepsilon_0 \Delta} \cdot \left(H_{\text{inc}}\Big|_{m-1/2}^{n+1/2} - H_{\text{inc}}\Big|_{m+1/2}^{n+1/2}\right) \tag{5.45a}$$

$$H_{\text{inc}}\Big|_{m+1/2}^{n+1/2} = H_{\text{inc}}\Big|_{m+1/2}^{n-1/2} + \frac{\Delta t}{\left[\dfrac{\tilde{v}_p(\phi = 0^\circ)}{\tilde{v}_p(\phi)}\right]\mu_0 \Delta} \cdot \left(E_{\text{inc}}\Big|_m^n - E_{\text{inc}}\Big|_{m+1}^n\right) \tag{5.45b}$$

Note that the factor $\tilde{v}_p(\phi = 0^\circ)/\tilde{v}_p(\phi)$, a ratio of numerical phase velocities in the main TM_z grid that is slightly less than one, is introduced as a multiplier of both μ_0 and ε_0 to slightly *speed up* the numerical wave in the auxiliary grid. This speed-up is needed to equalize the numerical phase velocities of the incident wave in the main grid (propagating at the angle ϕ) and the wave in the auxiliary grid, which would otherwise propagate more slowly because it behaves as if it were on-axis in the main grid (i.e., $\phi = 0^\circ$). Using (4.14b), (4.16), and (4.17), the required ratio of numerical phase velocities can readily be calculated at the beginning of the FDTD run. (Note that this velocity compensation is valid only at a single sinusoidal frequency. See Section 5.9 for more sophisticated compensation techniques appropriate for broadband pulses.)

Although the above discussion considered the case of Fig. 5.11(a) as an example, the use of an auxiliary one-dimensional FDTD grid to source incident-field values applies equally well to the three cases of Fig. 5.11(b). Changes of the incident-wave angle that cause a shift of the coordinate origin simply shift the assumed position of the auxiliary grid, as shown in this figure. In all cases, $E_{\text{inc}}|_{m_0}$ of the auxiliary grid coincides with the appropriate origin of the main grid.

Interpolation of Look-Up Table Data

Linear interpolation can be used to obtain the incident fields given knowledge of the delay distance d and the auxiliary grid field values. Let IFIX(r) denote the largest integer in the real number r. Then, with d known from (5.41) to (5.43), we obtain for an E-field located at d:

$$\begin{aligned} d' &= d - \text{IFIX}(d) \\ E_{\text{inc}}\Big|_d^n &= (1 - d') \cdot E_{\text{inc}}\Big|_{m_0 + \text{IFIX}(d)}^n + d' \cdot E_{\text{inc}}\Big|_{m_0 + \text{IFIX}(d)+1}^n \end{aligned} \tag{5.46a}$$

For a magnetic field located at d, we have similarly

$$\begin{aligned} d'' &= d + 1/2 \\ d' &= d'' - \text{IFIX}(d'') \\ H_{\text{inc}}\Big|_d^{n+1/2} &= (1 - d') \cdot H_{\text{inc}}\Big|_{m_0 - 1/2 + \text{IFIX}(d'')}^{n+1/2} + d' \cdot H_{\text{inc}}\Big|_{m_0 + 1/2 + \text{IFIX}(d'')}^{n+1/2} \end{aligned} \tag{5.46b}$$

Incident-Field Components

The last step in calculating the incident field needed to implement the TF/SF connecting condition is to compute the vector components of the field in the x-, y-, and z-coordinate directions. For the TM_z case under discussion, we have simply

$$E_{z,\,\mathrm{inc}}\Big|_d^n = E_{\mathrm{inc}}\Big|_d^n \tag{5.47a}$$

$$H_{x,\,\mathrm{inc}}\Big|_d^{n+1/2} = H_{\mathrm{inc}}\Big|_d^{n+1/2} \sin\phi \tag{5.47b}$$

$$H_{y,\,\mathrm{inc}}\Big|_d^{n+1/2} = -H_{\mathrm{inc}}\Big|_d^{n+1/2} \cos\phi \tag{5.47c}$$

where $E_{\mathrm{inc}}\Big|_d^n$ and $H_{\mathrm{inc}}\Big|_d^{n+1/2}$ are provided from (5.46a) and (5.46b). The incident-field components calculated using (5.47a–c) can now be substituted into the Region-1 / Region-2 connecting condition provided by systems (5.37) and (5.39).

5.7.3 Illustrative Example

We now consider an example of the action of the TF/SF grid zoning provided by systems (5.37) and (5.39) for a pulsed plane wave propagating in free space at $\phi = 45^\circ$ within a two-dimensional TE_z grid. The grid has uniform square cells of size $\Delta = 1.5$ cm and is time-stepped with $\Delta t = \Delta/2c$. Its total-field zone spans 100×100 cells, and is surrounded on all sides by a 20-cell scattered-field zone. A Gaussian pulse having a $1/e$ spectral bandwidth of 1 GHz is assumed for the incident wave.

Fig. 5.12 is a grayscale visualization that depicts four snapshots of the H_z field distribution within the grid. In this figure, the dotted lines show the location of the TF/SF interface. Above and to the right of O_1 in the total-field region, the incident wave propagates exactly at $\phi = 45^\circ$ with full amplitude and no wavefront distortion. Below and to the left of O_1 in the scattered-field region, there is only a small leakage wave that is suppressed by a factor of more than 10^3. Eventually, the plane wave propagates entirely through the total-field region and arrives at the upper-right TF/SF interface adjacent to gridpoint O_3. Here, the wave disappears from the grid with only a small leakage into the scattered-field region that is suppressed by approximately 10^2.

Fig. 5.13 repeats the visualization of Fig. 5.12, but here a 20×20-cell PEC cylinder is centered within the total-field region. This obstacle generates a complex scattered wave that eventually passes through all four sides of the TF/SF interface. This passage occurs with zero retroreflection and complete transmission. Thus, the TF/SF interface is completely transparent to outgoing scattered waves, regardless of their nature.

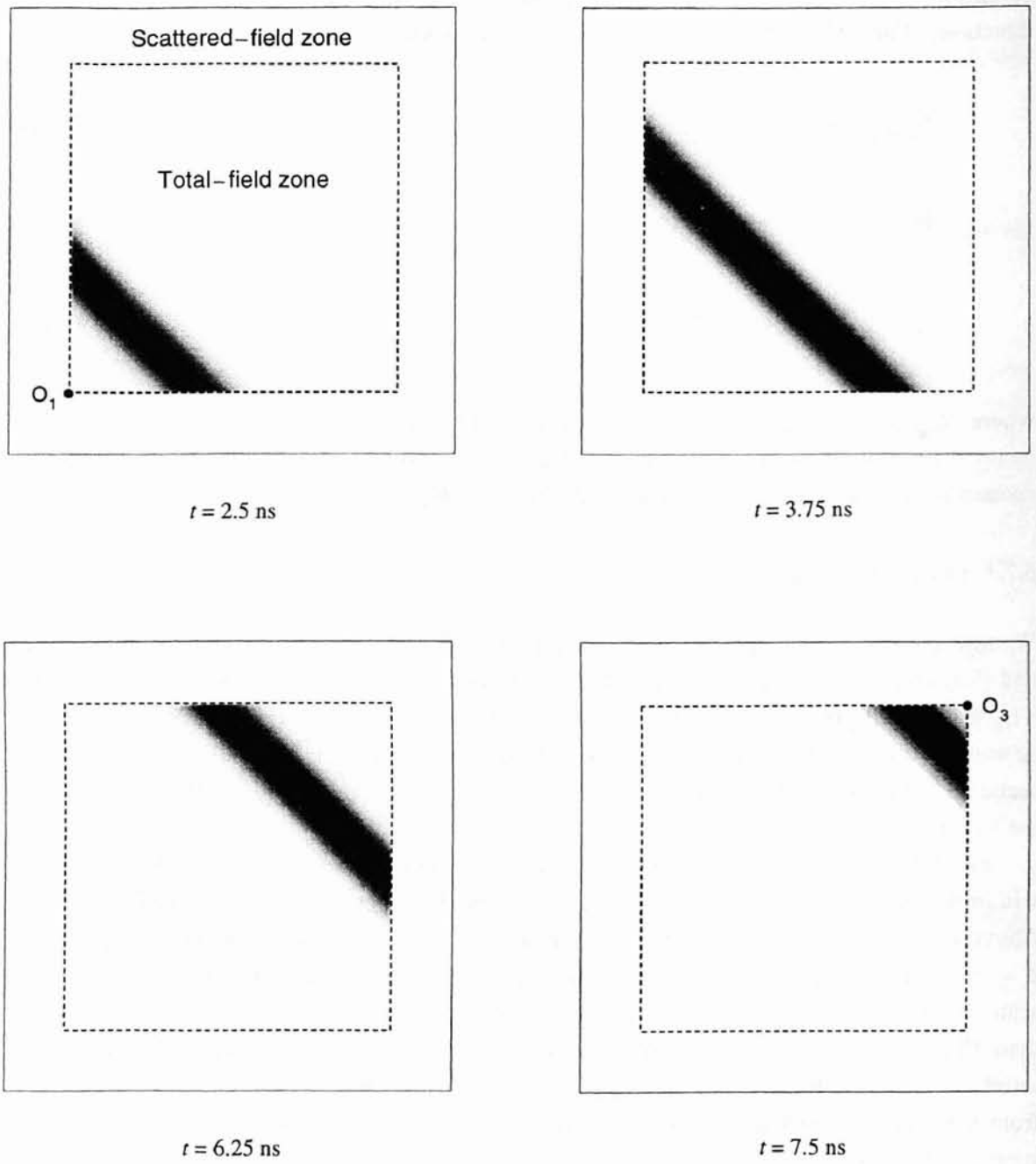


Fig. 5.12 Action of total-field / scattered-field grid zoning for a plane wave propagating in free space at $\phi = 45^\circ$ within a square-cell two-dimensional TE_z grid for $\Delta t = \Delta/2c$. The H_z field is visualized in four snapshots as the wave propagates through the total-field zone.

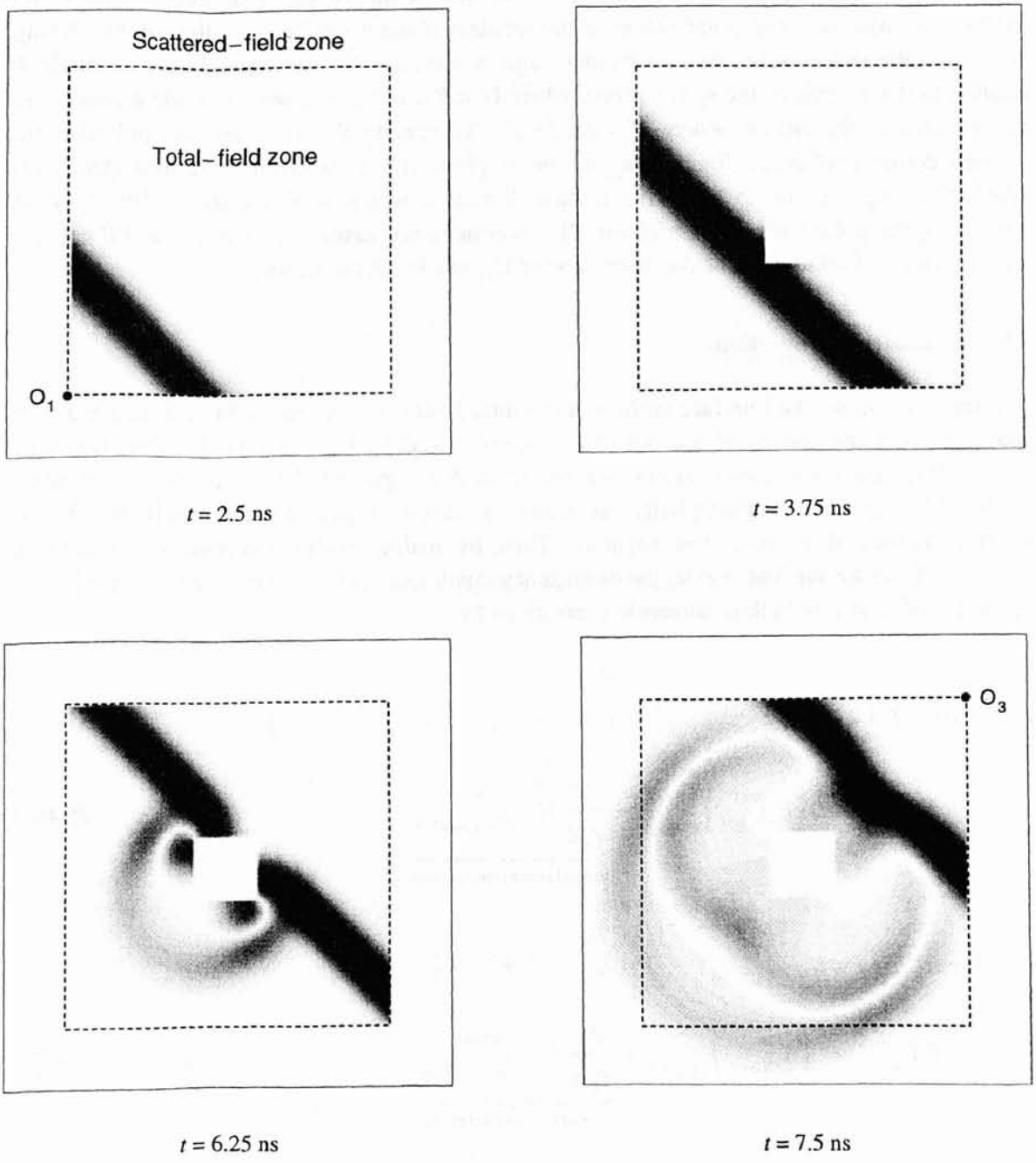


Fig. 5.13 Action of total-field / scattered-field grid zoning for a plane wave impinging upon a square PEC object at $\phi = 45^\circ$ within a square-cell two-dimensional TE_z grid for $\Delta t = \Delta/2c$. The H_z field is visualized in four snapshots as the wave propagates through the total-field zone.

5.8 THREE-DIMENSIONAL FORMULATION OF THE TF/SF TECHNIQUE

The total-field / scattered-field grid zoning discussed above can be readily extended to three-dimensional FDTD space lattices. Fig. 5.14(a) shows the coordinates used to define the propagation direction and polarization of the incident plane wave in three dimensions. Using standard spherical coordinates, the incident unit wavevector $\hat{\boldsymbol{k}}_{\text{inc}}$ is oriented with an angle θ relative to the $+z$-axis of the space lattice, where $0^\circ < \theta < 180^\circ$; and with an angle ϕ relative to the $+x$-axis of the lattice, where $0^\circ \le \phi < 360^\circ$. To specify the incident-wave polarization, we first define a reference direction $\hat{\boldsymbol{k}}_{\text{inc}} \times \hat{\boldsymbol{z}}$ in the plane of the wavefront. We then specify an orientation angle ψ for $\boldsymbol{E}_{\text{inc}}$ relative to this direction, where $0^\circ \le \psi < 360^\circ$. This way of specifying the polarization is useful for all wave-incidence cases except $\theta = 0^\circ$ and $\theta = 180^\circ$, where ϕ can be used to describe the orientation of $\boldsymbol{E}_{\text{inc}}$ relative to the $+x$-axis.

5.8.1 Consistency Conditions

In three dimensions, the interface surface of the total-field and scattered-field regions in the Yee space lattice is composed of six flat planes forming a closed rectangular box, as shown in Fig. 5.14(b). Each box face contains two tangential $\boldsymbol{E}$ components having a type and location specified in Fig. 5.15. For simplicity, we assume a cubic-cell space lattice wherein the TF/SF interface is located in source-free vacuum. Then, by analogy with the consistency conditions given by (5.37) for the TM_z mode, the consistency conditions for the $\boldsymbol{E}$ components at each face of the TF/SF interface in three dimensions are given by

$$j = j_0 \text{ Face: } E_x\left(i = i_0 + \tfrac{1}{2}, \ldots, i_1 - \tfrac{1}{2};\quad j = j_0;\quad k = k_0, \ldots, k_1\right)$$

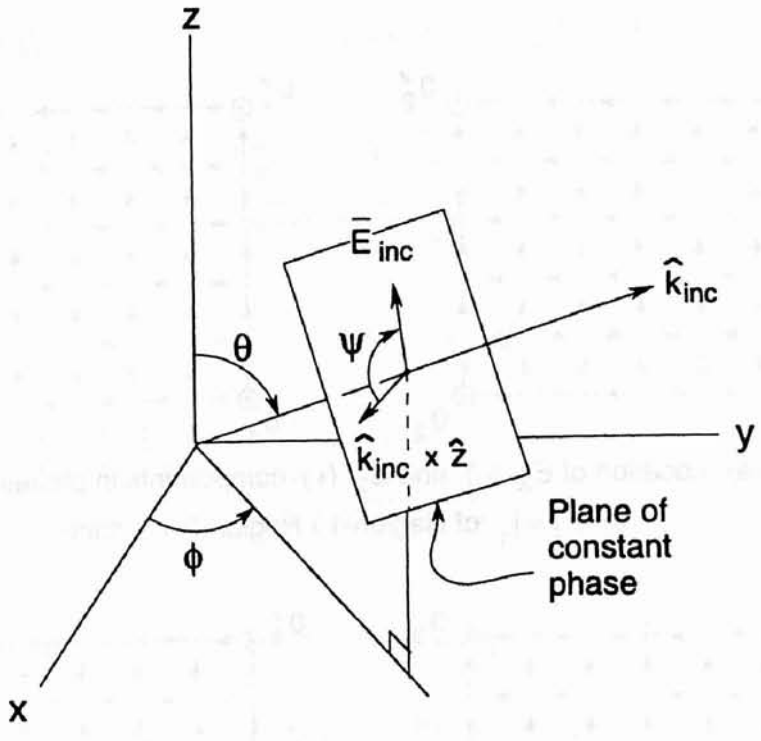
$$E_x\Big|_{i,j_0,k}^{n+1} = \left\{E_x\Big|_{i,j_0,k}^{n+1}\right\}_{(3.35a)} \underbrace{- \frac{\Delta t}{\varepsilon_0 \Delta} H_{z,\text{inc}}\Big|_{i,j_0-1/2,k}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.48a}$$

$$j = j_0 \text{ Face: } E_z\left(i = i_0, \ldots, i_1;\quad j = j_0;\quad k = k_0 + \tfrac{1}{2}, \ldots, k_1 - \tfrac{1}{2}\right)$$

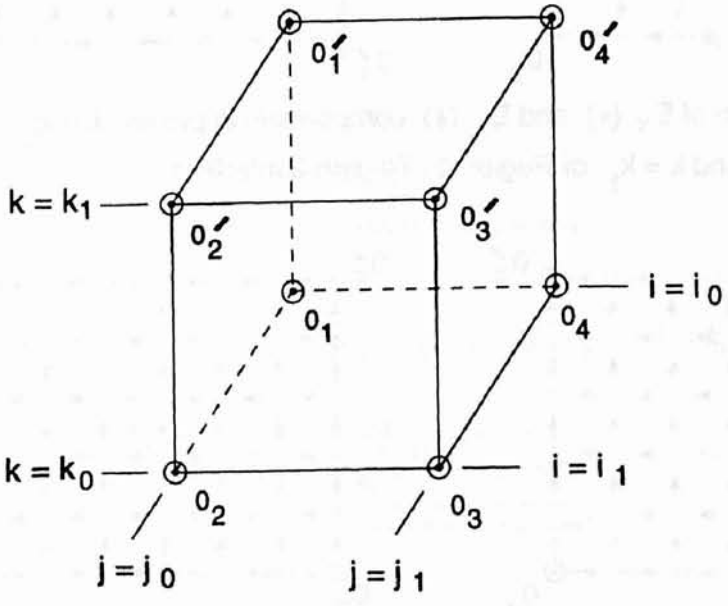
$$E_z\Big|_{i,j_0,k}^{n+1} = \left\{E_z\Big|_{i,j_0,k}^{n+1}\right\}_{(3.35c)} \underbrace{+ \frac{\Delta t}{\varepsilon_0 \Delta} H_{x,\text{inc}}\Big|_{i,j_0-1/2,k}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.48b}$$

$$j = j_1 \text{ Face: } E_x\left(i = i_0 + \tfrac{1}{2}, \ldots, i_1 - \tfrac{1}{2};\quad j = j_1;\quad k = k_0, \ldots, k_1\right)$$

$$E_x\Big|_{i,j_1,k}^{n+1} = \left\{E_x\Big|_{i,j_1,k}^{n+1}\right\}_{(3.35a)} \underbrace{+ \frac{\Delta t}{\varepsilon_0 \Delta} H_{z,\text{inc}}\Big|_{i,j_1+1/2,k}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.49a}$$

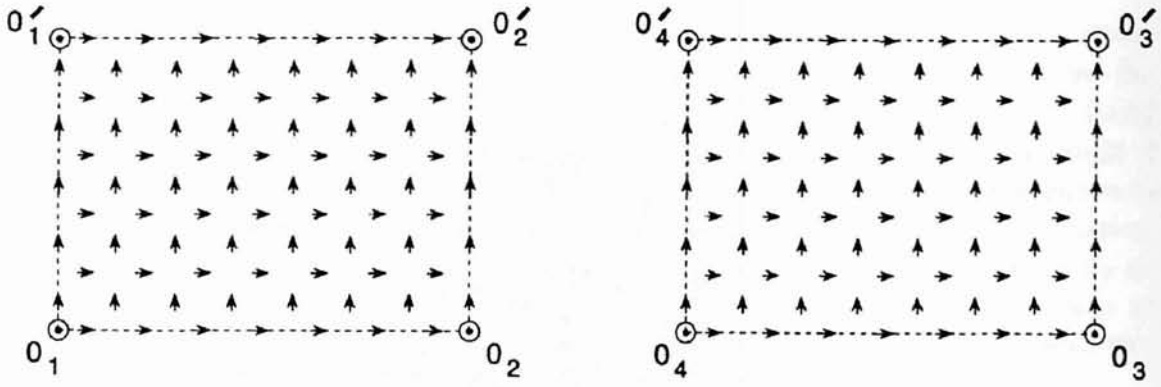


(a) Incident wave propagation direction and polarization

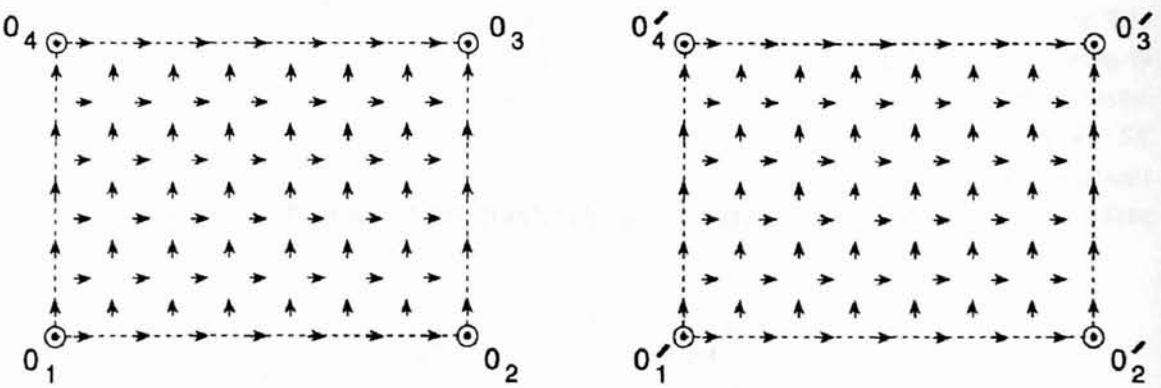


(b) Coordinate origins for calculation of incident field

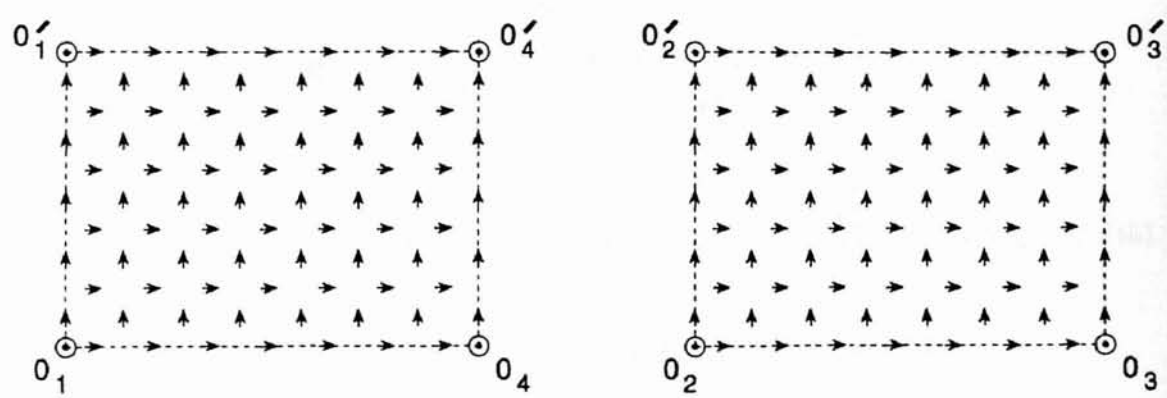
Fig. 5.14 Definitions of the incident-wave propagation direction and polarization and the six-sided total-field / scattered-field interface surface for the three-dimensional FDTD space lattice.



(a) Location of E_x (→) and E_z (↑) components in planes $j = j_0$ and $j = j_1$ of Region-1 / Region-2 interface



(b) Location of E_x (→) and E_y (↑) components in planes $k = k_0$ and $k = k_1$ of Region-1 / Region-2 interface



(c) Location of E_y (→) and E_z (↑) components in planes $i = i_0$ and $i = i_1$ of Region-1 / Region-2 interface

Fig. 5.15 Type and location of the tangential $\boldsymbol{E}$ components in the six faces of the total-field / scattered-field interface surface of Fig. 5.14(b).

$j = j_1$ *Face:* $E_z\left(i = i_0, \ldots, i_1 ; \quad j = j_1 ; \quad k = k_0 + \frac{1}{2}, \ldots, k_1 - \frac{1}{2}\right)$

$$E_z\Big|_{i,j_1,k}^{n+1} = \left\{E_z\Big|_{i,j_1,k}^{n+1}\right\}_{(3.35c)} \underbrace{- \frac{\Delta t}{\varepsilon_0 \Delta} H_{x,\,\mathrm{inc}}\Big|_{i,\,j_1+1/2,\,k}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.49b}$$

$k = k_0$ *Face:* $E_x\left(i = i_0 + \frac{1}{2}, \ldots, i_1 - \frac{1}{2}; \quad j = j_0, \ldots, j_1 ; \quad k = k_0\right)$

$$E_x\Big|_{i,j,k_0}^{n+1} = \left\{E_x\Big|_{i,j,k_0}^{n+1}\right\}_{(3.35a)} \underbrace{+ \frac{\Delta t}{\varepsilon_0 \Delta} H_{y,\,\mathrm{inc}}\Big|_{i,\,j,\,k_0-1/2}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.50a}$$

$k = k_0$ *Face:* $E_y\left(i = i_0, \ldots, i_1 ; \quad j = j_0 + \frac{1}{2}, \ldots, j_1 - \frac{1}{2}; \quad k = k_0\right)$

$$E_y\Big|_{i,j,k_0}^{n+1} = \left\{E_y\Big|_{i,j,k_0}^{n+1}\right\}_{(3.35b)} \underbrace{- \frac{\Delta t}{\varepsilon_0 \Delta} H_{x,\,\mathrm{inc}}\Big|_{i,\,j,\,k_0-1/2}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.50b}$$

$k = k_1$ *Face:* $E_x\left(i = i_0 + \frac{1}{2}, \ldots, i_1 - \frac{1}{2}; \quad j = j_0, \ldots, j_1 ; \quad k = k_1\right)$

$$E_x\Big|_{i,j,k_1}^{n+1} = \left\{E_x\Big|_{i,j,k_1}^{n+1}\right\}_{(3.35a)} \underbrace{- \frac{\Delta t}{\varepsilon_0 \Delta} H_{y,\,\mathrm{inc}}\Big|_{i,\,j,\,k_1+1/2}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.51a}$$

$k = k_1$ *Face:* $E_y\left(i = i_0, \ldots, i_1 ; \quad j = j_0 + \frac{1}{2}, \ldots, j_1 - \frac{1}{2}; \quad k = k_1\right)$

$$E_y\Big|_{i,j,k_1}^{n+1} = \left\{E_y\Big|_{i,j,k_1}^{n+1}\right\}_{(3.35b)} \underbrace{+ \frac{\Delta t}{\varepsilon_0 \Delta} H_{x,\,\mathrm{inc}}\Big|_{i,\,j,\,k_1+1/2}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.51b}$$

$i = i_0$ *Face:* $E_y\left(i = i_0 ; \quad j = j_0 + \frac{1}{2}, \ldots, j_1 - \frac{1}{2}; \quad k = k_0, \ldots, k_1\right)$

$$E_y\Big|_{i_0,j,k}^{n+1} = \left\{E_y\Big|_{i_0,j,k}^{n+1}\right\}_{(3.35b)} \underbrace{+ \frac{\Delta t}{\varepsilon_0 \Delta} H_{z,\,\mathrm{inc}}\Big|_{i_0-1/2,\,j,\,k}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.52a}$$

$i = i_0$ *Face:* $E_z\left(i = i_0;\quad j = j_0, \ldots, j_1;\quad k = k_0 + \frac{1}{2}, \ldots, k_1 - \frac{1}{2}\right)$

$$E_z\Big|_{i_0,j,k}^{n+1} = \left\{E_z\Big|_{i_0,j,k}^{n+1}\right\}_{(3.35c)} \underbrace{- \frac{\Delta t}{\varepsilon_0 \Delta} H_{y,\,\mathrm{inc}}\Big|_{i_0-1/2,\,j,k}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.52b}$$

$i = i_1$ *Face:* $E_y\left(i = i_1;\quad j = j_0 + \frac{1}{2}, \ldots, j_1 - \frac{1}{2};\quad k = k_0, \ldots, k_1\right)$

$$E_y\Big|_{i_1,j,k}^{n+1} = \left\{E_y\Big|_{i_1,j,k}^{n+1}\right\}_{(3.35b)} \underbrace{- \frac{\Delta t}{\varepsilon_0 \Delta} H_{z,\,\mathrm{inc}}\Big|_{i_1+1/2,\,j,k}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.53a}$$

$i = i_1$ *Face:* $E_z\left(i = i_1;\quad j = j_0, \ldots, j_1;\quad k = k_0 + \frac{1}{2}, \ldots, k_1 - \frac{1}{2}\right)$

$$E_z\Big|_{i_1,j,k}^{n+1} = \left\{E_z\Big|_{i_1,j,k}^{n+1}\right\}_{(3.35c)} \underbrace{+ \frac{\Delta t}{\varepsilon_0 \Delta} H_{y,\,\mathrm{inc}}\Big|_{i_1+1/2,\,j,k}^{n+1/2}}_{\text{assumed known correction term}} \tag{5.53b}$$

By analogy with the consistency conditions given by (5.39) for the TM_z mode, the consistency conditions for the $\boldsymbol{H}$ components located 0.5Δ outside of each face of the TF/SF interface in three dimensions are given by

$j = j_0 - \frac{1}{2}$ *Face:* $H_z\left(i = i_0 + \frac{1}{2}, \ldots, i_1 - \frac{1}{2};\quad j = j_0 - \frac{1}{2};\quad k = k_0, \ldots, k_1\right)$

$$H_z\Big|_{i,j_0-1/2,k}^{n+1/2} = \left\{H_z\Big|_{i,j_0-1/2,k}^{n+1/2}\right\}_{(3.36c)} \underbrace{- \frac{\Delta t}{\mu_0 \Delta} E_{x,\,\mathrm{inc}}\Big|_{i,\,j_0,k}^{n}}_{\text{assumed known correction term}} \tag{5.54a}$$

$j = j_0 - \frac{1}{2}$ *Face:* $H_x\left(i = i_0, \ldots, i_1;\quad j = j_0 - \frac{1}{2};\quad k = k_0 + \frac{1}{2}, \ldots, k_1 - \frac{1}{2}\right)$

$$H_x\Big|_{i,j_0-1/2,k}^{n+1/2} = \left\{H_x\Big|_{i,j_0-1/2,k}^{n+1/2}\right\}_{(3.36a)} \underbrace{+ \frac{\Delta t}{\mu_0 \Delta} E_{z,\,\mathrm{inc}}\Big|_{i,\,j_0,k}^{n}}_{\text{assumed known correction term}} \tag{5.54b}$$

$j = j_1 + \frac{1}{2}$ *Face:* $H_z\left(i = i_0 + \frac{1}{2}, \ldots, i_1 - \frac{1}{2};\quad j = j_1 + \frac{1}{2};\quad k = k_0, \ldots, k_1\right)$

$$H_z\Big|_{i,j_1+1/2,k}^{n+1/2} = \left\{H_z\Big|_{i,j_1+1/2,k}^{n+1/2}\right\}_{(3.36c)} \underbrace{+ \frac{\Delta t}{\mu_0 \Delta} E_{x,\,\mathrm{inc}}\Big|_{i,\,j_1,k}^{n}}_{\text{assumed known correction term}} \tag{5.55a}$$

$j = j_1 + \frac{1}{2}$ *Face:* $H_x\left(i = i_0, \ldots, i_1;\quad j = j_1 + \frac{1}{2};\quad k = k_0 + \frac{1}{2}, \ldots, k_1 - \frac{1}{2}\right)$

$$H_x\Big|^{n+1/2}_{i,j_1+1/2,k} = \left\{H_x\Big|^{n+1/2}_{i,j_1+1/2,k}\right\}_{(3.36a)} - \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{z,\,\mathrm{inc}}\Big|^{n}_{i,\,j_1,\,k}}_{\text{assumed known correction term}} \tag{5.55b}$$

$k = k_0 - \frac{1}{2}$ *Face:* $H_y\left(i = i_0 + \frac{1}{2}, \ldots, i_1 - \frac{1}{2};\quad j = j_0, \ldots, j_1;\quad k = k_0 - \frac{1}{2}\right)$

$$H_y\Big|^{n+1/2}_{i,j,k_0-1/2} = \left\{H_y\Big|^{n+1/2}_{i,j,k_0-1/2}\right\}_{(3.36b)} + \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{x,\,\mathrm{inc}}\Big|^{n}_{i,\,j,\,k_0}}_{\text{assumed known correction term}} \tag{5.56a}$$

$k = k_0 - \frac{1}{2}$ *Face:* $H_x\left(i = i_0, \ldots, i_1;\quad j = j_0 + \frac{1}{2}, \ldots, j_1 - \frac{1}{2};\quad k = k_0 - \frac{1}{2}\right)$

$$H_x\Big|^{n+1/2}_{i,j,k_0-1/2} = \left\{H_x\Big|^{n+1/2}_{i,j,k_0-1/2}\right\}_{(3.36a)} - \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{y,\,\mathrm{inc}}\Big|^{n}_{i,\,j,\,k_0}}_{\text{assumed known correction term}} \tag{5.56b}$$

$k = k_1 + \frac{1}{2}$ *Face:* $H_y\left(i = i_0 + \frac{1}{2}, \ldots, i_1 - \frac{1}{2};\quad j = j_0, \ldots, j_1;\quad k = k_1 + \frac{1}{2}\right)$

$$H_y\Big|^{n+1/2}_{i,j,k_1+1/2} = \left\{H_y\Big|^{n+1/2}_{i,j,k_1+1/2}\right\}_{(3.36b)} - \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{x,\,\mathrm{inc}}\Big|^{n}_{i,\,j,\,k_1}}_{\text{assumed known correction term}} \tag{5.57a}$$

$k = k_1 + \frac{1}{2}$ *Face:* $H_x\left(i = i_0, \ldots, i_1;\quad j = j_0 + \frac{1}{2}, \ldots, j_1 - \frac{1}{2};\quad k = k_1 + \frac{1}{2}\right)$

$$H_x\Big|^{n+1/2}_{i,j,k_1+1/2} = \left\{H_x\Big|^{n+1/2}_{i,j,k_1+1/2}\right\}_{(3.36a)} + \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{y,\,\mathrm{inc}}\Big|^{n}_{i,\,j,\,k_1}}_{\text{assumed known correction term}} \tag{5.57b}$$

$i = i_0 - \frac{1}{2}$ *Face:* $H_z\left(i = i_0 - \frac{1}{2};\quad j = j_0 + \frac{1}{2}, \ldots, j_1 - \frac{1}{2};\quad k = k_0, \ldots, k_1\right)$

$$H_z\Big|^{n+1/2}_{i_0-1/2,j,k} = \left\{H_z\Big|^{n+1/2}_{i_0-1/2,j,k}\right\}_{(3.36c)} + \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{y,\,\mathrm{inc}}\Big|^{n}_{i_0,\,j,\,k}}_{\text{assumed known correction term}} \tag{5.58a}$$

$i = i_0 - \frac{1}{2}$ *Face:* $H_y\left(i = i_0 - \frac{1}{2};\quad j = j_0, \ldots, j_1;\quad k = k_0 + \frac{1}{2}, \ldots, k_1 - \frac{1}{2}\right)$

$$H_y\Big|_{i_0-1/2,j,k}^{n+1/2} = \left\{H_y\Big|_{i_0-1/2,j,k}^{n+1/2}\right\}_{(3.36b)} - \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{z,\text{inc}}\Big|_{i_0,j,k}^{n}}_{\text{assumed known correction term}} \tag{5.58b}$$

$i = i_1 + \frac{1}{2}$ *Face:* $H_z\left(i = i_1 + \frac{1}{2};\quad j = j_0 + \frac{1}{2}, \ldots, j_1 - \frac{1}{2};\quad k = k_0, \ldots, k_1\right)$

$$H_z\Big|_{i_1+1/2,j,k}^{n+1/2} = \left\{H_z\Big|_{i_1+1/2,j,k}^{n+1/2}\right\}_{(3.36c)} - \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{y,\text{inc}}\Big|_{i_1,j,k}^{n}}_{\text{assumed known correction term}} \tag{5.59a}$$

$i = i_1 + \frac{1}{2}$ *Face:* $H_y\left(i = i_1 + \frac{1}{2};\quad j = j_0, \ldots, j_1;\quad k = k_0 + \frac{1}{2}, \ldots, k_1 - \frac{1}{2}\right)$

$$H_y\Big|_{i_1+1/2,j,k}^{n+1/2} = \left\{H_y\Big|_{i_1+1/2,j,k}^{n+1/2}\right\}_{(3.36b)} + \underbrace{\frac{\Delta t}{\mu_0 \Delta} E_{z,\text{inc}}\Big|_{i_1,j,k}^{n}}_{\text{assumed known correction term}} \tag{5.59b}$$

5.8.2 Calculation of the Incident Field

Coordinate Origins

As shown in Fig. 5.14(b), the TF/SF interface surface in three dimensions provides eight possible points of initial contact with the incident wavefront as the wavevector angles θ and ϕ are varied through their ranges. Analogous to the two-dimensional TM_z case discussed earlier, these points of initial contact are, in fact, coordinate origins for calculation of the delay distance d needed to obtain the incident-wave fields for the consistency conditions of (5.48) to (5.59). Following the notation established for the TM_z case, distance d for the three-dimensional case is again given by (5.41), which is repeated here for convenience:

$$d = \hat{\boldsymbol{k}}_{\text{inc}} \cdot \boldsymbol{r}_{\text{comp}}$$

Here, $\hat{\boldsymbol{k}}_{\text{inc}}$ is the unit incident wavevector given by

$$\hat{\boldsymbol{k}}_{\text{inc}} = \hat{x} \sin\theta \cos\phi + \hat{y} \sin\theta \sin\phi + \hat{z} \cos\theta \tag{5.60}$$

and $\boldsymbol{r}_{\text{comp}}$ is the position vector from the appropriate origin to the field vector component of interest.

For $0^\circ < \theta \le 90^\circ$, Origins O_1, O_2, O_3, and O_4 of Fig. 5.14(b) are the points of the TF/SF interface that can make initial contact with the incident wavefront. We have the following four possibilities for r_{comp} for this range of θ:

$$0^\circ \le \phi \le 90^\circ \quad \rightarrow \quad \text{Origin } O_1 \text{ at } (i_0, j_0, k_0)$$

$$r_{comp} = (i_{comp} - i_0)\hat{x} + (j_{comp} - j_0)\hat{y} + (k_{comp} - k_0)\hat{z} \tag{5.61a}$$

$$90^\circ < \phi \le 180^\circ \quad \rightarrow \quad \text{Origin } O_2 \text{ at } (i_1, j_0, k_0)$$

$$r_{comp} = (i_{comp} - i_1)\hat{x} + (j_{comp} - j_0)\hat{y} + (k_{comp} - k_0)\hat{z} \tag{5.61b}$$

$$180^\circ < \phi \le 270^\circ \quad \rightarrow \quad \text{Origin } O_3 \text{ at } (i_1, j_1, k_0)$$

$$r_{comp} = (i_{comp} - i_1)\hat{x} + (j_{comp} - j_1)\hat{y} + (k_{comp} - k_0)\hat{z} \tag{5.61c}$$

$$270^\circ < \phi < 360^\circ \quad \rightarrow \quad \text{Origin } O_4 \text{ at } (i_0, j_1, k_0)$$

$$r_{comp} = (i_{comp} - i_0)\hat{x} + (j_{comp} - j_1)\hat{y} + (k_{comp} - k_0)\hat{z} \tag{5.61d}$$

For $90^\circ < \theta < 180^\circ$, Origins O_1', O_2', O_3', and O_4' of Fig. 5.14(b) are the points of the TF/SF interface that can make initial contact with the incident wavefront. We have the following four additional possibilities for r_{comp} for this range of θ:

$$0^\circ \le \phi \le 90^\circ \quad \rightarrow \quad \text{Origin } O_1' \text{ at } (i_0, j_0, k_1)$$

$$r_{comp} = (i_{comp} - i_0)\hat{x} + (j_{comp} - j_0)\hat{y} + (k_{comp} - k_1)\hat{z} \tag{5.62a}$$

$$90^\circ < \phi \le 180^\circ \quad \rightarrow \quad \text{Origin } O_2' \text{ at } (i_1, j_0, k_1)$$

$$r_{comp} = (i_{comp} - i_1)\hat{x} + (j_{comp} - j_0)\hat{y} + (k_{comp} - k_1)\hat{z} \tag{5.62b}$$

$$180^\circ < \phi \le 270^\circ \quad \rightarrow \quad \text{Origin } O_3' \text{ at } (i_1, j_1, k_1)$$

$$r_{comp} = (i_{comp} - i_1)\hat{x} + (j_{comp} - j_1)\hat{y} + (k_{comp} - k_1)\hat{z} \tag{5.62c}$$

$$270^\circ < \phi < 360^\circ \quad \rightarrow \quad \text{Origin } O_4' \text{ at } (i_0, j_1, k_1)$$

$$r_{comp} = (i_{comp} - i_0)\hat{x} + (j_{comp} - j_1)\hat{y} + (k_{comp} - k_1)\hat{z} \tag{5.62d}$$

Selection of one of these eight cases of coordinate origins and corresponding position vectors can be performed automatically in a computer program upon specification of the wavevector angles θ and ϕ by the user. Then the delay distance d can be calculated using (5.41).

Generation of Look-Up Table

By analogy with the TM_z case discussed in Section 5.7, an auxiliary linear FDTD grid can be used to generate a look-up table for the space-time variation of the incident fields in the three-dimensional Yee lattice. The source grid is again assumed to be placed along the incident wavevector so that the appropriate origin $(O_1 - O_4,\ O_1' - O_4')$ of the TF/SF interface in the lattice coincides with one of the E-components of the source grid, $E_{\text{inc}}|_{m_0}$.

Both the main three-dimensional space lattice and the auxiliary one-dimensional source grid use the same Δ, Δt, and time-step number n. The source grid is again time-stepped using (5.45a, b). Here, the only modification is that the equalization factor for the numerical phase velocities is written as $\tilde{v}_p(\theta = 0^\circ, \phi = 0^\circ) \,/\, \tilde{v}_p(\theta, \phi)$, reflecting the additional angular possibility for wavevector direction in three dimensions. In a manner analogous to (4.16) and (4.17), this ratio of numerical phase velocities can be calculated via a Newton's method solution of (4.12), the numerical dispersion relation for the three-dimensional Yee algorithm. (See Section 5.9 for more sophisticated dispersion-compensation techniques suitable for broadband pulses.)

Interpolation of Look-Up Table Data

Given the delay distance d and the source-grid FDTD values, the linear interpolation method of (5.46) can again be used to obtain the incident-field values $E_{\text{inc}}|_d$ and $H_{\text{inc}}|_d$.

Incident-Field Components

The last step is to compute the required vector components of the incident field:

$$H_{x,\,\text{inc}}\Big|_d^{n+1/2} = H_{\text{inc}}\Big|_d^{n+1/2}(\sin\psi\,\sin\phi + \cos\psi\,\cos\theta\,\cos\phi) \tag{5.63a}$$

$$H_{y,\,\text{inc}}\Big|_d^{n+1/2} = H_{\text{inc}}\Big|_d^{n+1/2}(-\sin\psi\,\cos\phi + \cos\psi\,\cos\theta\,\sin\phi) \tag{5.63b}$$

$$H_{z,\,\text{inc}}\Big|_d^{n+1/2} = H_{\text{inc}}\Big|_d^{n+1/2}(-\cos\psi\,\sin\theta) \tag{5.63c}$$

$$E_{x,\,\text{inc}}\Big|_d^{n} = E_{\text{inc}}\Big|_d^{n}(\cos\psi\,\sin\phi - \sin\psi\,\cos\theta\,\cos\phi) \tag{5.63d}$$

$$E_{y,\,\text{inc}}\Big|_d^{n} = E_{\text{inc}}\Big|_d^{n}(-\cos\psi\,\cos\phi - \sin\psi\,\cos\theta\,\sin\phi) \tag{5.63e}$$

$$E_{z,\,\text{inc}}\Big|_d^{n} = E_{\text{inc}}\Big|_d^{n}(\sin\psi\,\sin\theta) \tag{5.63f}$$

The incident-field vector components of (5.63) can now be substituted into the consistency conditions provided by (5.48) to (5.59). This completes the process of generating the incident plane wave in three dimensions via total-field / scattered-field zoning of the Yee space lattice.

5.9 ADVANCED DISPERSION COMPENSATION IN THE TF/SF TECHNIQUE

As discussed in Sections 5.7.2 and 5.8.2, implementation of the TF/SF technique requires calculation of the incident E and H components at FDTD space-lattice points on or adjacent to the interface of total-field Region 1 and scattered-field Region 2. A simple procedure was described for this calculation, wherein an auxiliary one-dimensional FDTD grid is assumed to be oriented along the incident wavevector in the main two- or three-dimensional space lattice. The idea is to use the linear source grid, which has the same Δ, Δt, and time-step number n as the main grid, to calculate the free-space propagation of the incident wave. In this manner, the source grid generates a look-up table for the space-time variation of the incident fields. The required incident field at any point in the main grid can be obtained by interpolating corresponding field values along the source grid.

While this strategy is conceptually straightforward, there is the nuance that, in general, the numerical dispersion characteristics of the incident and main grids differ. This leads to imperfect confinement of the incident wave within Region 1. Field leakage into Region 2 is most pronounced at the corner of the TF/SF interface that is opposite to the point of first contact of the incident wave with Region 1. While small, this leakage can lead to significant error in calculating scattering behavior, especially for those objects that exhibit pronounced suppression of scattering in certain directions.

Equation (5.45) represents an attempt to compensate for the difference in the numerical phase velocities of waves in the incident and main grids by adjusting the effective speed of light in the incident grid. However, (5.45) is exact only in the limit of prolonged time-stepping of a single-frequency sinusoidal incident wave. While helpful in reducing field leakage even during nonsinusoidal-steady-state conditions, (5.45) is simply not formulated to address the issue of matching the numerical dispersion characteristics of the incident and main grids over the broad range of Fourier components of the incident wave that are injected during impulsive excitations.

This section describes two enhancements to the TF/SF formulation that address the problem of minimizing field leakage into Region 2. The first, the *matched numerical dispersion* (MND) technique of Guiffaut and Mahdjoubi [7], requires only a slight modification of the method presented in Sections 5.7 and 5.8. As will be shown, one merely uses a different space-cell size Δ in the one-dimensional wave-source grid than in the main grid to ensure that the dispersion properties in both grids are nearly the same. As opposed to the velocity adjustment of (5.45), wherein the dispersion correction is at only a single frequency, this modification matches the dispersions of the two grids over a broad spectrum. As a result, leakage into the scattered-field region is suppressed by 30 to 40 dB relative to the technique of (5.45) for broadband pulse excitations.

The second enhancement, called the *analytical field-propagation* (AFP) technique, takes an entirely different approach. Instead of using an auxiliary grid to calculate the incident field, an analytical expression is obtained for the incident field at each required space-lattice point. This expression accounts for all of the numerical artifacts inherent in grid-based propagation. In addition, unlike one-dimensional auxiliary grid schemes, it accounts for the fact that for non-grid-aligned propagation of a homogeneous plane wave, E, H, and wavevector $\bar{k}$ do not exactly form a mutually orthogonal set, as they do in the continuous world. The AFP technique realizes orders-of-magnitude reduction in leakage into the scattered-field region, at the expense of additional calculations to perform Fourier and inverse-Fourier transformations at each point where an incident field component is required.

The use of a one-dimensional auxiliary incident-wave grid requires interpolation of the fields within this grid to points corresponding to projected locations in the higher-dimensional main grid. This interpolation, which the MND scheme must employ, comes with its own set of artifacts, which has been addressed in [8, 9]. The AFP technique does not require interpolation, and hence avoids any interpolating errors. It should also be noted that multidimensional auxiliary-grid schemes have been developed [10, 11]. Such schemes avoid interpolation, but have both certain advantages and disadvantages compared to the methods described here. The interested reader is referred to [10, 11].

5.9.1 Matched Numerical Dispersion Technique

Guiffaut and Mahdjoubi [7] described a slight modification to the one-dimensional auxiliary grid technique, which makes a significant difference in the ability of the wave-source grid to match the dispersion characteristics for nongrid-aligned propagation in two- or three-dimensional main grids. Following [7], our goal is to have the numeric wavenumber in one dimension, $\tilde{k}_{1D}$, match the numeric wavenumber in three dimensions, $\tilde{k}_{3D}$. (The two-dimensional case is a subset, wherein propagation is perpendicular to one of the axes of the three-dimensional space lattice, and need not be separately considered.) These wavenumbers are related to the true wavenumber k by a scalar factor ζ which is yet to be determined. Hence,

$$\tilde{k}_{1D} = \tilde{k}_{3D} = \zeta k = \zeta\left(\frac{2\pi}{\lambda}\right) = \zeta\left(\frac{2\pi}{N_\lambda \Delta_{3D}}\right) \tag{5.64}$$

where the number of cells per wavelength N_λ is defined in terms of the space-cell size Δ_{3D} in the three-dimensional main grid. For a plane wave propagating at spherical angles θ and ϕ in the main grid, the factor ζ is obtained from the three-dimensional dispersion relation, which can be written as

$$\sin^2\left(\frac{\pi S_{3D}}{N_\lambda}\right) = S_{3D}^2\left[\sin^2\left(\frac{\pi\zeta\cos\phi\sin\theta}{N_\lambda}\right) + \sin^2\left(\frac{\pi\zeta\sin\phi\sin\theta}{N_\lambda}\right) + \sin^2\left(\frac{\pi\zeta\cos\theta}{N_\lambda}\right)\right] \tag{5.65}$$

Here, the equality $\omega\Delta t/2 = \pi S_{3D}/N_\lambda$ has been used in the argument for the sine function on the left-hand side. In (5.65), one specifies the number of cells per wavelength N_λ, the Courant number S_{3D}, and the direction of propagation. There is no closed-form solution for ζ, and hence it must be obtained numerically by such techniques as bisection or Newton's method.

Having obtained ζ, we turn our attention to the one-dimensional dispersion relation. In general this can be written as

$$\sin^2\left(\frac{\omega\Delta t}{2}\right) = S_{1D}^2 \sin^2\left(\frac{\tilde{k}_{1D}\Delta_{1D}}{2}\right) \tag{5.66}$$

where Δ_{1D} is the space-cell size in the one-dimensional wave-source grid, and S_{1D} is the Courant number in that grid. Since the same time-step is used in both the one- and three-dimensional grids, the argument on the left-hand side can written as in (5.65). The argument on the right can be modified by substituting the last term shown in (5.64) for $\tilde{k}_{1D}$. Then, after taking the square root of both sides, this yields

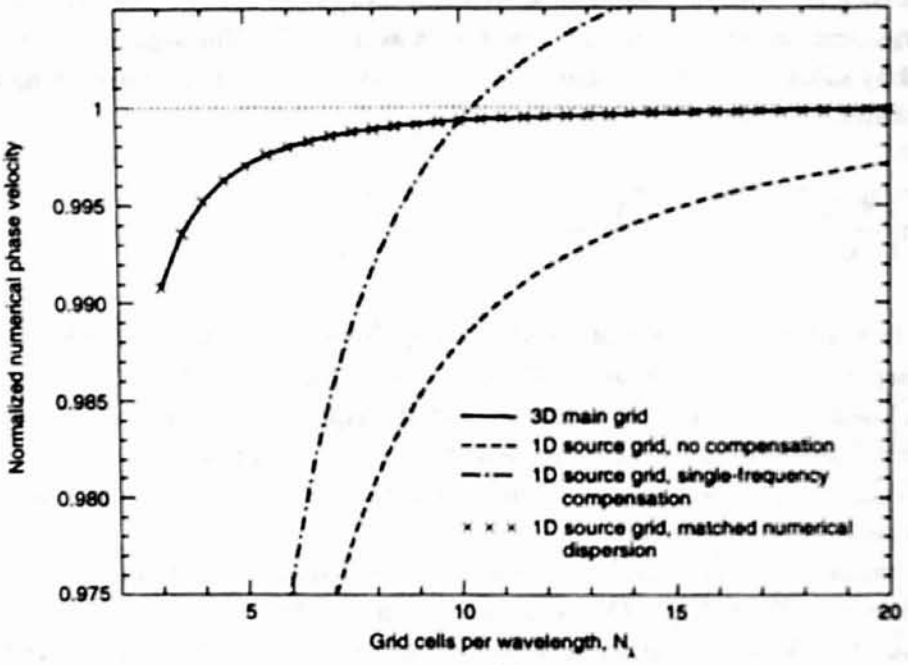
$$\sin\left(\frac{\pi S_{3D}}{N_\lambda}\right) = S_{1D}\sin\left(\frac{\pi\zeta\Delta_{1D}}{N_\lambda\Delta_{3D}}\right) = S_{1D}\sin\left(\frac{\pi\zeta S_{3D}}{N_\lambda S_{1D}}\right) \tag{5.67}$$

The second form of the equation is obtained by recognizing that $\Delta_{1D}/\Delta_{3D} = S_{3D}/S_{1D}$ because the same time-step Δt is used for both grids. Equation (5.67) is now solved for S_{1D}, and hence, Δ_{1D}. Again, no closed-form solution is available, and the equation must be solved numerically. The one-dimensional auxiliary wave-source grid can now be constructed using this value of Δ_{1D}. This grid matches the dispersion characteristics of the three-dimensional main grid nearly exactly over a broad range of discretizations.

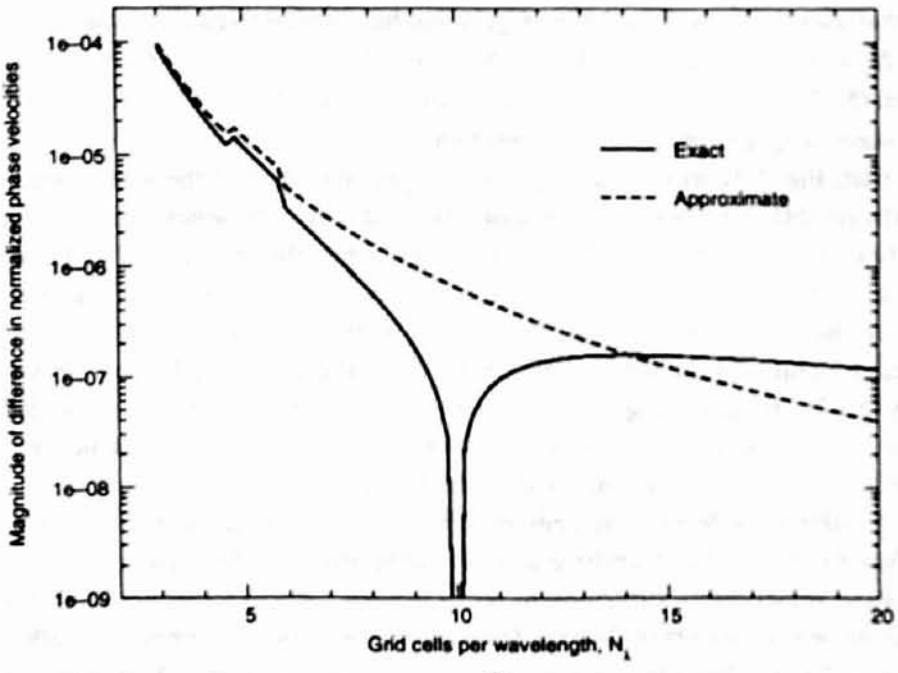
To demonstrate this, we consider an example of propagation in a three-dimensional main grid in the direction ($\theta = 60^\circ$, $\phi = 45^\circ$) with the Courant number $S_{3D} = 0.95/\sqrt{3}$. Fig. 5.16(a) graphs (versus N_λ) the numerical phase velocity normalized to the free-space speed of light. Four separate cases are shown: (1) in the main grid; (2) in a one-dimensional wave-source grid that uses the same Δ (and hence the same Courant number) as the main grid; (3) in a wave-source grid that uses the same Δ as the main grid, but has its phase velocity adjusted to match the main grid at $N_\lambda = 10$ using the approach of (5.45); and (4) in a wave-source grid that uses the Δ_{1D} obtained from (5.67) for $N_\lambda = 10$. In the last case, the ratio Δ_{3D}/Δ_{1D} is found to be 1.70570174 for the assumed wave-propagation direction in the main grid.

We see from Fig. 5.16(a) that the numerical dispersion curve of the uncompensated wave-source grid always differs from that of the main grid, and only approaches a match with the main grid at very fine grid discretizations. We also see that, while the approach of (5.45) succeeds in locally matching the dispersions in the main and wave-source grids (as indicated by the intersection of the curves at $N_\lambda = 10$), this match degrades rapidly away from the single-frequency match point. *By far* the best dispersion matching is achieved using the value of Δ_{1D} prescribed by (5.67). Here, the dispersion in the wave-source grid agrees so well with that of the main grid that the two curves are *indistinguishable* at the scale of 5.16(a). Hence, the usage of the term "matched numerical dispersion" to describe this technique becomes clear.

To better quantify the level of agreement of the dispersion characteristics of the main and MND wave-source grids, Fig. 5.16(b) graphs the magnitude of the difference between their normalized phase velocities as a function of N_λ. Observing the exact values of this difference (the solid curve), we see numbers that are four to seven orders-of-magnitude smaller than the free-space speed of light. For propagating modes of primary interest in FDTD modeling, that is, modes sampled at $N_\lambda \geq 10$, this dispersion match approaches perfection. The null at $N_\lambda = 10$ corresponds to the grid discretization at which the dispersion in the two grids was chosen to match. The dashed curve in Fig. 5.16(b) will be explained in the context of the discussion to follow, regarding the essential independence of the solution for Δ_{1D} relative to N_λ.



(a)



(b)

Fig. 5.16 Attributes of the matched numerical dispersion technique for a one-dimensional wave-source grid: (a) Superior matching of the numerical dispersion characteristic of wave propagation at ($\theta = 60^\circ$, $\phi = 45^\circ$) in a 3D grid relative to uncompensated or single-frequency compensated source grids; (b) corresponding miniscule normalized phase-velocity mismatch of MND grid.

We note that (5.67) requires specification of N_λ, S_{3D}, and the factor ζ previously obtained by solving (5.65). It is clear that N_λ is a fixed (single-frequency) parameter that cannot be uniquely defined for a broadband impulsive excitation. However, we can show that the precise value used for N_λ is not particularly important, and one can safely use any "reasonable" number. To show this, consider (5.65) and (5.66). Essentially, S_{1D} was obtained by equating the right-hand sides of these two equations [although first (5.65) had to be solved for ζ]. Here, we directly equate these right-hand sides after using the first two terms of the Taylor series expansion of the squared sine functions (i.e., $\sin^2(x) \equiv x^2 - x^4/3$). The result is

$$S_{3D}^2 \left\{ \left(\frac{\pi\zeta}{N_\lambda} \right)^2 - \frac{1}{3}\left(\frac{\pi\zeta}{N_\lambda} \right)^4 \left[\sin^4\theta\left(\cos^4\phi + \sin^4\phi\right) + \cos^4\theta \right] \right\} \equiv$$

$$S_{1D}^2 \left[\left(\frac{\pi\zeta S_{3D}}{N_\lambda S_{1D}} \right)^2 - \frac{1}{3}\left(\frac{\pi\zeta S_{3D}}{N_\lambda S_{1D}} \right)^4 \right] \tag{5.68}$$

Upon simplifying (5.68), we obtain

$$\frac{S_{1D}}{S_{3D}} = \frac{\Delta_{3D}}{\Delta_{1D}} \equiv \frac{1}{\sqrt{\sin^4\theta\left(\cos^4\phi + \sin^4\phi\right) + \cos^4\theta}} \tag{5.69}$$

This shows, rather surprisingly, that at least approximately, the optimum ratio of the space-cell sizes in the two grids is independent of N_λ and depends only on the direction of propagation. For the example of Fig. 5.16 where the direction of propagation is ($\theta = 60^\circ$, $\phi = 45^\circ$), (5.69) yields a ratio of 1.7056057. Comparing this to the nominally exact value of 1.70570174, we see that the two agree through the first four digits. Hence, if one wants to avoid the root-finding needed to solve for this ratio rigorously, one is relatively safe merely using (5.69). The dashed line in Fig. 5.16(b) shows the residual normalized phase-velocity mismatch using this approximate approach. While the null is no longer present, the mismatch is in the same order of magnitude as before (i.e., exceedingly small).

The final step in implementing the MND technique is to slightly modify the TF/SF procedure described previously in this chapter. First, we simply use Δ_{1D} [determined either rigorously with (5.65) and (5.67) or approximately with (5.69)] in the updating coefficients for E and H in the wave-source grid. Second, and just as simple, when doing the interpolations outlined in Sections 5.7.2 and 5.8.2, we replace the distance d given in (5.41) with $(\Delta_{3D}/\Delta_{1D})d$. All other implementation details remain the same.

For broadband pulses launched into the main grid, numerical experiments show that MND optimization of the TF/SF wave source reduces field leakage into the scattered-field region by 30:1 to 100:1 (i.e., 30 to 40 dB) relative to the single-frequency compensation technique of (5.45). While remarkable given the simplicity of implementing MND, we may ask: "Why is there not even more leakage reduction, given the microscopic residual phase-velocity mismatch shown in Fig. 5.16(b)?" The answer is that error due to the required interpolation of field values in the wave-source grid becomes the dominant cause of leakage, once the numerical dispersion characteristics of the wave-source and main grids are exactly matched. The next section discusses a means to achieve dispersion-matching and avoid interpolation, thereby suppressing leakage into the scattered-field region by more than 90 dB for wideband pulses.

5.9.2 Analytical Field Propagation

Having a complete quantification of the way in which plane waves propagate in the FDTD grid, the field at any point in the grid can be determined, given an incident-field time-series at a reference point and the direction of propagation. For notational simplicity, we will describe two-dimensional TM_z plane-wave propagation, but extension to three dimensions is straightforward. For those three-dimensional simulations in which the incident wave propagates in a plane that is perpendicular to a grid axis, the derivation presented here pertains directly.

Assume that we designate the time series for the E_z field component of the incident wave to be $g(n\Delta t)$ at grid reference point (i_{ref}, j_{ref}). In the frequency domain, the phasor E_z field at any other node (i, j) in the grid is simply

$$\breve{E}_z\big|_{i,j} = P(i - i_{ref}, j - j_{ref})\, \mathcal{F}(g) \tag{5.70}$$

where $\mathcal{F}$ is the discrete Fourier transform, and the complex-valued propagator P is given by

$$P(i - i_{ref}, j - j_{ref}) = \exp\left\{-j\left[\tilde{k}_x(i - i_{ref})\Delta + \tilde{k}_y(j - j_{ref})\Delta\right]\right\} \tag{5.71}$$

In the time domain, the field is given by

$$E_z\big|_{i,j} = \mathcal{F}^{-1}\left[P(i - i_{ref}, j - j_{ref})\, \mathcal{F}(g)\right] \tag{5.72}$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform. In two dimensions, one can write the x- and y-components of the H-field directly in terms of E_z. It can be shown that, in the time domain, these components are given by

$$H_x\big|_{i,j} = \mathcal{F}^{-1}\left[\frac{S\sin\left(\tilde{k}_y \Delta/2\right)}{\eta\sin\left(\omega\Delta t/2\right)} P(i - i_{ref}, j - j_{ref})\, \mathcal{F}(g)\right] \tag{5.73}$$

$$H_y\big|_{i,j} = -\mathcal{F}^{-1}\left[\frac{S\sin\left(\tilde{k}_x \Delta/2\right)}{\eta\sin\left(\omega\Delta t/2\right)} P(i - i_{ref}, j - j_{ref})\, \mathcal{F}(g)\right] \tag{5.74}$$

where $\eta = \sqrt{\mu/\varepsilon}$. The staggering of the field components is inherently accounted for by the values of i and j, which are not restricted to integers (i.e., an offset of one-half is used when appropriate).

Equations (5.71) to (5.74) comprise the heart of the analytical field propagation TF/SF technique. These expressions give the incident field at any point in the grid, and can be used for the incident field in the TF/SF update equations, as given in Sections 5.7 and 5.8. For arbitrary three-dimensional propagation, we add the third wavevector component term to the argument of the exponential in (5.71), and find six field components rather than three. In addition, we must account for the polarization, although there is no unique orientation for the fields since the orientation is frequency dependent.

We now consider an example that illustrates just how effective the AFP TF/SF technique is in suppressing field leakage into the scattered-field region relative to the use of an auxiliary one-dimensional wave-source grid, even one which has been compensated using the matched numerical dispersion method. This example involves a 141×141 square-cell TM_z FDTD grid, which is excited by a plane wave propagating at 30° relative to the x-axis with the temporal dependence of a Ricker wavelet (a twice-differentiated Gaussian pulse). The grid discretization is 20 cells per wavelength at the most energetic frequency of the Ricker wavelet, and the Courant number is set to $1/\sqrt{3}$ so that the fields associated with this example could pertain to a three-dimensional simulation. Leakage is recorded five space cells to the right and five space cells above the upper-right corner of the total-field region. The reference point for the incident wave is the lower-left corner of the total-field region.

Fig. 5.17 compares the absolute value of the calculated field leakage waveform versus time-step number for the single-frequency compensated source grid, the MND source grid, and the AFP technique. We see that the MND results are approximately 40 dB better than the traditional single-frequency compensated incident-wave grid. However, the AFP technique is much better yet, suppressing field leakage by approximately 120 dB relative to the MND method. We note that the AFP technique exhibits nonphysical, acausal energy that arrives at the recording point essentially at the beginning of time-stepping. However, this energy is microscopic compared with the energy contained in the incident pulse and should not pose a concern in practice.

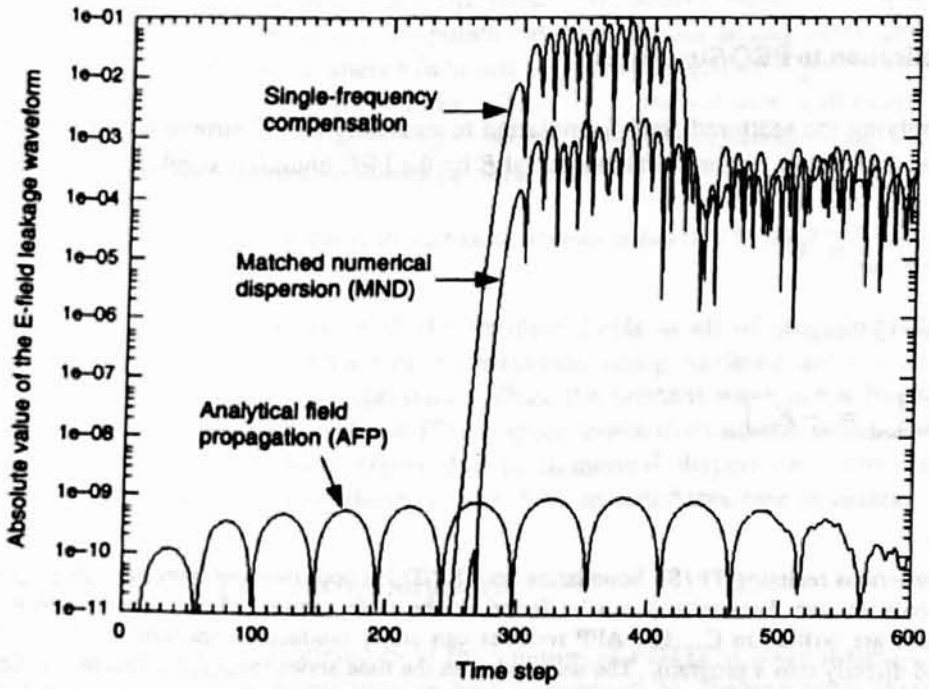


Fig. 5.17 Comparison of the field leakage versus time-step number into the scattered-field region of a two-dimensional TM_z TF/SF FDTD grid for three wave-source techniques: the single-frequency compensated source grid, the matched numerical dispersion source grid, and analytical field propagation. The injected plane wave has the time dependence of a Ricker wavelet (a twice-differentiated Gaussian pulse) that is selected such that the FDTD grid discretization is 20 cells per wavelength at the most energetic frequency of the excitation.

Fig. 5.17 shows time out to 600 time-steps. However, in this example, the TF/SF code was required only to calculate the incident field to 500 time-steps, at which point the incident field was effectively declared zero at all points along the TF/SF interface. Subsequently, the FDTD simulation could be run to any number of time-steps. Since the incident field was only needed for 500 time-steps, 512-point FFTs were employed in the AFP algorithm. In general, the length of the FFT must be sufficient to contain the entire incident time series, but need not be longer (in this case, after performing the inverse FFT, 12 points were discarded).[4]

5.10 SCATTERED-FIELD FORMULATION

The scattered-field formulation [12] borrows from a method popular with the frequency-domain integral equation (method of moments) community. Again, the concept evolves from the linearity of Maxwell's equations and the decomposition of total $\boldsymbol{E}$ and total $\boldsymbol{H}$ into known incident fields and unknown scattered fields, according to (5.24).

Here, however, the FDTD method is used to time-step *only* scattered $\boldsymbol{E}$ and scattered $\boldsymbol{H}$. That is, the FDTD grid is not segmented into total-field and scattered-field regions, but instead assumes scattered-field quantities everywhere. No incident wave propagates within the grid. If the total $\boldsymbol{E}$ or total $\boldsymbol{H}$ time waveform at a gridpoint is needed, it is obtained in a postprocessing step by adding the FDTD-computed scattered-field time waveform to the known incident-wave time dependence at that point.

5.10.1 Application to PEC Structures

Consider applying the scattered-field formulation to modeling a PEC structure. At the surface of this structure, there must be zero total tangential $\boldsymbol{E}$ by the PEC boundary condition:

$$\boldsymbol{E}_{\tan}\big|_{\text{total}} = 0 \tag{5.75}$$

Therefore, by (5.24a):

$$\boldsymbol{E}_{\tan}\big|_{\text{scat}} = -\boldsymbol{E}_{\tan}\big|_{\text{inc}} \tag{5.76}$$

---

[4]Computer programs realizing TF/SF boundaries both MND (in both two and three dimensions) and AFP (currently only in two dimensions) can be found in the code section of the website www.fdtd.org. The programs are written in C. The AFP routines can either produce an incident-field file or can be incorporated directly into a program. The user specifies the time series (actually a function which can be called to obtain the time series), a reference point, the Courant number, the impedance, the direction of propagation, and the desired number of time-steps. The incident field is calculated, and a function is supplied which applies these fields to nodes adjacent to the TF/SF boundary. All calculations are in double precision. The FFTs are performed using the FFTw routines, which are available from www.fftw.org. For the example presented here, which required the calculation of approximately 800 FFTs of 512 points, the computation time was approximately 0.11 seconds on a 2-GHz Apple G5 running OSX 10.3.

must hold at the structure surface at all time-steps. This causes a scattered wave to be locally generated at the surface. The scattered wave has surface tangential $\boldsymbol{E}$ components equal and opposite to those of the incident wave, which are assumed to be known everywhere in space and time. (Recall that the incident wave is assumed to propagate in free space, that is, in the absence of any of the structures being modeled, so that it can be expressed as a simple analytical wavefunction of space and time.) Thus, at each lattice space-cell edge that defines a small part of the structure surface, the FDTD code need only specify as a *hard source* the local $\boldsymbol{E}$ component as being equal and opposite to the incident $\boldsymbol{E}$ at that point in space and time.

There are two principal disadvantages of the scattered-field formulation relative to the total-field / scattered-field technique when used to model PEC structures:

- To permit enforcement of the PEC boundary condition, the scattered-field method requires that components of the incident-wave tangential $\boldsymbol{E}$ be calculated at all space-lattice points comprising the surface of the structure being modeled. Because the location of these points varies from structure to structure, the programming of the wave-source condition in the scattered-field method is structure-dependent and potentially complicated if a complex surface shape is involved. Further, complex surface shapes can require a large amount of computer arithmetic to evaluate the required incident $\boldsymbol{E}$ components at all of the surface points.

- In space regions where total $\boldsymbol{E}$ and total $\boldsymbol{H}$ are small, such as inside well-shielded cavities, the incident and scattered fields are nearly equal and opposite (canceling). Here, the normal computational uncertainties arising from applying the Yee algorithm to the scattered field can be greatly magnified if postprocessing is needed to obtain the total field. This is a well-known computational phenomenon called *subtraction noise*, which limits the dynamic range of the scattered-field method in calculating the total field.

There is one major advantage of the scattered-field approach relative to the total-field / scattered-field technique:

- With the scattered-field method, the incident fields at all of the space-lattice points comprising the surface of the structure being modeled are calculated *exactly* using an analytical expression. Thus, the incident wave is not forced to numerically propagate through the FDTD space lattice to excite the structure, and therefore suffers no phase errors due to numerical dispersion. This is of increasing importance as the electrical size of the modeled structure increases.

5.10.2 Application to Lossy Dielectric Structures

The initial application of the scattered-field formulation of FDTD to lossy dielectric structures was reported in [13]. This formulation is best understood by starting with Maxwell's curl equations (3.7) and (3.8) and writing them first for total-field quantities in a region free of current sources:

$$\mu \frac{\partial \boldsymbol{H}_{\text{total}}}{\partial t} = -\nabla \times \boldsymbol{E}_{\text{total}} - \sigma^* \boldsymbol{H}_{\text{total}} \tag{5.77}$$

$$\varepsilon \frac{\partial \boldsymbol{E}_{\text{total}}}{\partial t} = \nabla \times \boldsymbol{H}_{\text{total}} - \sigma \boldsymbol{E}_{\text{total}} \tag{5.78}$$

Note that the material parameters μ, ε, σ^*, and σ are permitted to vary in an arbitrary manner throughout the space of interest for this total-field formulation. Next, write (3.7) and (3.8) for the incident field:

$$\mu_0 \frac{\partial \boldsymbol{H}_{\text{inc}}}{\partial t} = -\nabla \times \boldsymbol{E}_{\text{inc}} \tag{5.79}$$

$$\varepsilon_0 \frac{\partial \boldsymbol{E}_{\text{inc}}}{\partial t} = \nabla \times \boldsymbol{H}_{\text{inc}} \tag{5.80}$$

Note again that the incident field is defined as propagating in a free-space region having the electric parameters μ_0, ε_0, $\sigma^* = 0$, and $\sigma = 0$.

Now, working under the assumption of the linearity of Maxwell's equations, we can subtract (5.79) from (5.77), and (5.80) from (5.78) to obtain

$$\mu \frac{\partial \boldsymbol{H}_{\text{scat}}}{\partial t} + \sigma^* \boldsymbol{H}_{\text{scat}} = -\nabla \times \boldsymbol{E}_{\text{scat}} - \sigma^* \boldsymbol{H}_{\text{inc}} - (\mu - \mu_0) \frac{\partial \boldsymbol{H}_{\text{inc}}}{\partial t} \tag{5.81}$$

$$\varepsilon \frac{\partial \boldsymbol{E}_{\text{scat}}}{\partial t} + \sigma \boldsymbol{E}_{\text{scat}} = \nabla \times \boldsymbol{H}_{\text{scat}} - \sigma \boldsymbol{E}_{\text{inc}} - (\varepsilon - \varepsilon_0) \frac{\partial \boldsymbol{E}_{\text{inc}}}{\partial t} \tag{5.82}$$

Equations (5.81) and (5.82) can be realized numerically using ordinary Yee central differences to implement the space and time derivatives. These equations differ from analogous total-field expressions derived in Chapter 3 *only* in that:

1. The scattered field is subject to time-differentiation, and therefore time-stepping;
2. There are simple functions of the known incident E- and H-fields on the right-hand sides.

However, the latter fact implies that the incident fields must be calculated at *all* space-lattice locations where $\mu \neq \mu_0$ and $\varepsilon \neq \varepsilon_0$. Despite the fact that the incident fields have a known functional form in space and time, merely implementing this functional form at possibly tens of millions of space-points and thousands of time-steps can pose a formidable computational burden.

Relative to the total-field / scattered-field approach, this need to compute the incident field components *volumetrically* within the FDTD space lattice is the primary disadvantage of the scattered-field formulation for lossy dielectric structures. However, this burden can be mitigated somewhat by calculating the incident wave via a table look-up procedure using an auxiliary one-dimensional FDTD source grid, as discussed previously in this chapter in the context of total-field / scattered-field zoning. Here, the source grid could use the magic time-step to yield exact incident-field data (before the interpolation process).

5.10.3 Choice of Incident Plane-Wave Formulation

Over the years, substantial high-quality FDTD modeling experience has been obtained with both the total-field / scattered-field and scattered-field formulations. The total-field / scattered-field technique seems particularly useful for guided-wave simulations, including transmission lines and microwave circuits. Both approaches have done well with free-space scattering problems. The reader should be aware of the benefits and limitations of each approach, as discussed above, and in the optimum case be able to construct working codes for each.

5.11 WAVEGUIDE SOURCE CONDITIONS

Sourcing an incident numerical wave in an FDTD model of a PEC or dielectric waveguide presents additional challenges relative to the ones discussed above for plane-wave sources in free space. The characteristics of a waveguide that provide these challenges are as follows:

- The waveguide usually supports a number of distinct propagating modes that have substantially different spatial distributions of $\boldsymbol{E}$ and $\boldsymbol{H}$. It may be difficult to simulate the numerical excitation of one particular mode without inadvertently exciting some or all of the rest.
- Some waveguiding systems such as open dielectric channels and microstrip structures can have energy that is not strictly bounded by conducting outer walls, but instead decays with distance from the center of the system. The transverse distribution of the modal field may be initially unknown for such systems.
- A pulsed wideband excitation of a waveguide introduces spectral energy that travels at possibly widely varying phase and group velocities due to the dispersive nature of the guide's frequency-wavenumber characteristic. This can cause difficulty in specifying any numerical source condition that is not completely localized in space.
- Waveguides exhibiting cutoff phenomena can be subject to reactive fields loitering in the vicinity of a source. The distance between a numerical source and the interaction structure of interest within the waveguide must be selected to allow substantially complete decay of these reactive fields.
- The use of a wideband pulsed source in a waveguide exhibiting cutoff phenomena may require substantially prolonging the time-stepping of the FDTD simulation to permit the very slowly propagating spectral energy just above the waveguide cutoff frequency to reach the interaction structure of interest, if these spectral components are of interest in the simulation.

5.11.1 Pulsed Electric Field Modal Hard Source

A simple way to excite an incident wave in an FDTD model of a waveguide or microstrip is to specify a pulsed $\boldsymbol{E}$ hard-source distribution in a transverse cross section. That is, some or all of the $\boldsymbol{E}$ components located in and tangential to the transverse source plane are provided with a space-time variation that coincides with that of the desired propagating mode, and are independent of the presence of any other numerical waves in the grid.

In certain cases, the complete transverse field distribution of the desired propagating mode is known analytically, allowing the pulsed numerical wave that is launched to immediately represent the desired mode without unduly generating undesired propagating modes and below-cutoff reactive fields. In other cases, the transverse distribution of the desired mode is initially unknown or can only be approximated. There are two options in this situation:

1. *Implement an approximation of the true mode.* The impact of specifying a subset of the complete transverse distribution is to possibly generate undesired numerical waves that represent multimoding and reactive fields. A substantial buffer length of waveguiding structure may then be required in the model to permit the reactive fields to decay. An example of an approximate waveguide modal source is the pulsing of a collinear array of $\boldsymbol{E}$ components in the FDTD mesh to simulate a probe extending from a PEC waveguide wall and dead-ending in air at a point halfway to the opposing wall. An example of an approximate microstrip source is the pulsing of a similar collinear array of $\boldsymbol{E}$ components extending from a ground plane to the metal trace of the microstrip.
2. *Use "bootstrapping."* By this we mean running a preliminary FDTD model of the waveguide that is sufficiently electrically long to decay all undesired reactive and multimode fields. The transverse field distribution at the far end of this preliminary model is stored in a data file and is subsequently used in all production runs as a compact hard source. Effectively, the preliminary FDTD run is used to solve for the correct transverse field distribution.

The principal concern with this (or for that matter, *any*) hard-source condition is that its effective source impedance is zero. As a result, numerical waves properly reflected from a simulated discontinuity further down the waveguide will nonphysically retroreflect from the hard source if it is still operating. This nonphysical retroreflection is total or partial, depending on whether the hard source completely or partially specifies the entire transverse source plane.

A good way to eliminate this problem is to turn off the hard source before any reflecting numerical waves reach its position. Following the strategy discussed in Section 5.2, the pulsed source waveform should be smoothly decayed to zero, and a program flag set so that execution diverts around the hard source statements. In effect, the source plane is removed and replaced by standard free-space Yee cells by the time the reflecting numerical waves reach its original position. The reflecting numerical waves can then freely propagate through what was the source plane to reach an ABC representing the extension of the waveguiding system to infinity.

This procedure requires elongation of the simulated waveguiding system between the source plane and the first reflective discontinuity to permit the complete evolution and return to zero of the pulsed source waveform before arrival of the initial reflection at the source plane. For certain classes of models, this may not be possible because of the constraints of source pulse width and dimensions of the waveguiding system being modeled.

A second problem occurs when the bandwidth of the pulsed source is so wide that there is significant spectral energy below cutoff of the waveguiding system. The nonpropagating reactive fields that are generated by the source *never leave* its vicinity. Upon turning off the hard source and replacing it with normal Yee updating, a spurious transient is generated in the fields loitering about the source. This transient can have sufficient spectral content above cutoff to propagate down the waveguide an indefinite distance, thereby contaminating field data along the entire guide. To avoid this problem, the bandwidth of a hard source that is turned off should be carefully controlled to minimize spectral content below cutoff.

5.11.2 Total-Field / Reflected-Field Modal Formulation

The total-field / reflected-field modal formulation is simply the application of the total-field / scattered-field technique discussed earlier in this chapter to guided-wave systems. Here, the FDTD space lattice modeling a microwave structure is zoned into two distinct regions separated by a nonphysical virtual surface that serves to connect the fields in each region.

The interacting structure of interest is embedded within total-field Region 1. In reflected-field Region 2, located outside of Region 1, the FDTD algorithm operates only on reflected-field components with no incident wave present. A consistency condition is provided at the interface of the two regions. This generates an incident wave in Region 1 having a user-specified transverse distribution and time waveform, while allowing little leakage into Region 2. The Region-1 / Region-2 interface is transparent to outgoing numerical reflected-wave modes, which freely enter Region 2 and then disappear from the grid due to the action of the ABC.

Unlike the pulsed E-field modal hard source discussed in Section 5.11.1, there is no need to decay the total-field / reflected-field source to zero and remove it from the lattice, and no need to elongate the simulated waveguide system to causally isolate the source from adjacent waveguide discontinuities that generate reflections. Here, the source can be continuously operated in close proximity to adjacent discontinuities.

The programming of the total-field / reflected-field source for a waveguide can be easily placed in the context of this chapter's previous detailed discussions. For example, assume that we wish to generate a $+x$-directed incident waveguide mode. Let the interface between Regions 1 and 2 be the y-z plane at grid coordinate $i = i_s$ containing total-field E_y and E_z components. We immediately identify this interface with the $i = i_0$ face of the total-field / scattered-field surface of Figs. 5.14(b) and 5.15(c). The consistency conditions for this face are provided by (5.52) and (5.58). Thus, all that is needed to implement the wave source is knowledge of the incident $\boldsymbol{E}$ values at the total-field component locations in plane $i = i_s$, and knowledge of the incident $\boldsymbol{H}$ values at the reflected-field component locations in plane $i = i_s - 1/2$.

Overall, this technique is well suited to sourcing an accurate waveguide mode for an arbitrary PEC or dielectric system if the incident-field values are known over the complete transverse cross section of the waveguide. Bootstrapping, as discussed earlier, may be advised to obtain the necessary field component data, especially for open dielectric and microstrip guides. Again, caution should be exercised in using this approach for wideband pulse excitations. Here, it is possible that the *physical* dispersion of the propagating spectral energy between planes $i = i_s - 1/2$ and $i = i_s$ is sufficient to introduce error in the assumed incident-field distributions at these planes. In addition, the inadvertent sourcing of reactive below-cutoff fields can muddy the assumed incident distributions.

5.11.3 Resistive Source and Load Conditions

In the laboratory, a commonly used means to excite microstrips or waveguides is an insulated metal probe extending from the metal ground plane or waveguide wall. For microstrips, the probe connects to the signal trace. For waveguides, the probe dead-ends in free space at approximately the center of the waveguide cross section at the location of an E-field maximum of the waveguide mode. The metal probe is an extension of the center conductor of a coaxial cable that is connected to a microwave generator. Usually, the output impedance of the microwave generator is matched to the characteristic impedance of the coaxial cable, which in turn is matched to the characteristic impedance of the excited microstrip or waveguide.

Three-dimensional FDTD modeling of such structures is clearly desirable. Here, it is useful to compactly simulate a microwave generator having a finite resistive source impedance. A simple means to accomplish this goal is to *physically model* the length and the location of the exciting metal probe. For all time-steps, we merely set to zero a daisy chain of $\boldsymbol{E}$ components in the FDTD space lattice that is located at the metal probe. The base of this probe model is located at a distance of two lattice cells from the local reference ground plane or waveguide wall. We excite the probe model by providing special time-stepping relations for the two daisy-chained $\boldsymbol{E}$-components, E_{gap1} and E_{gap2}, located in the gap between the base of the probe and ground. For convenience, assume a cubic Yee lattice of cell-size Δ. Then, we have:

- E_{gap1} designated as a hard source of $V_{Thev}(t)/\Delta$ V/m.
  - This yields an FDTD model of the zero-impedance Thevenin-equivalent voltage source $V_{Thev}(t)$ that characterizes the microwave generator.
- E_{gap2} assigned the electrical conductivity $\sigma = 1/(R_{Thev}\Delta)$ S/m.
  - This yields an FDTD model of a physical resistor having the value R_{Thev} Ω and the dimensions of a single space-lattice cell. R_{Thev} is the Thevenin-equivalent source resistance that characterizes the microwave generator.

Specifying E_{gap1} and E_{gap2} in this manner models the connection of the base of the metal probe to a general microwave source. In effect, a Thevenin-equivalent circuit representing the microwave source is placed in the FDTD space lattice between the base of the simulated metal probe and the local reference ground plane or waveguide wall. We note that, by simply setting $E_{gap1} = 0$, this technique models a passive resistive load connected at the base of the metal probe. (See Section 15.9 of Chapter 15 for a much more general discussion of modeling electronic circuit elements and sources embedded within the FDTD lattice.)

The modeling technique discussed above is a realistic means to use FDTD to simulate the connection of common microwave signal sources to transmission-line and waveguide structures. Interestingly, most artifacts resulting from this excitation technique also occur physically in the laboratory. For example, below-cutoff, high-order, or retroreflected numerical modes can be generated in the FDTD space lattice when the probe model is used to excite a simulated waveguide. However, these modes also appear in the laboratory as *physical* artifacts when an actual microwave generator is connected to a waveguide probe via a coaxial cable. Therefore, to ensure proper results, whether using FDTD modeling or working in the laboratory, care must be taken to account for the electromagnetic wave phenomena due to probe excitation.

5.12 SUMMARY

This chapter reviewed four useful, general classes of compact electromagnetic wave sources for FDTD models to spotlight their key features: (1) hard-sourced $\boldsymbol{E}$ and $\boldsymbol{H}$ fields in one- and two-dimensional grids; (2) $\boldsymbol{J}$ and $\boldsymbol{M}$ current sources in three-dimensional lattices; (3) the total-field / scattered-field formulation for plane-wave excitation in one, two, and three dimensions; and (4) waveguide sources. While not compact, a fifth class of wave source, the pure scattered-field formulation, was also reviewed because of its utility. Chapter 15 will discuss in detail a sixth class of potential FDTD wave sources, electronic circuits. Circuit components and sources to be considered there include resistors, capacitors, inductors, diodes, transistors, and logic gates. These are useful for modeling high-speed digital circuits and microwave amplifiers.

REFERENCES

[1] Waldschmidt, G., and A. Taflove, "The determination of the effective radius of a filamentary source in the FDTD mesh," *IEEE Microwave Guided Wave Lett.*, Vol. 10, 2000, pp. 217–219.

[2] Wagner, C. L., and J. B. Schneider, "Divergent fields, charge, and capacitance in FDTD simulations," *IEEE Trans. Microwave Theory Tech.*, Vol. 46, 1998, pp. 2131–2136.

[3] Yee, K. S., "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, Vol. 14, 1966, pp. 302–307.

[4] Merewether, D. E., R. Fisher, and F. W. Smith, "On implementing a numeric Huygen's source scheme in a finite difference program to illuminate scattering bodies," *IEEE Trans. Nuclear Science*, Vol. 27, 1980, pp. 1829–1833.

[5] Mur, G., "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations," *IEEE Trans. Electromagn. Compat.*, Vol. 23, 1981, pp. 377–382.

[6] Umashankar, K. R., and A. Taflove, "A novel method to analyze electromagnetic scattering of complex objects," *IEEE Trans. Electromagn. Compat.*, Vol. 24, 1982, pp. 397–405.

[7] Guiffaut, C., and K. Mahdjoubi, "Perfect wideband plane wave injector for FDTD method," *Proc. IEEE Antennas Propagation Soc. Intl. Symp.*, Salt Lake City, UT, Vol. 1, 2000, pp. 236–239.

[8] Oguz, U., and L. Gürel, "Interpolation techniques to improve the accuracy of the plane wave excitations in the finite difference time domain method," *Radio Science*, Vol. 32, 1997, pp. 2189–2199.

[9] Oguz, U., and L. Gürel, "An efficient and accurate technique for the incident-wave excitations in the FDTD method," *IEEE Trans. Microwave Theory Tech.*, Vol. 46, 1998, pp. 869–882.

[10] Watts, M. E., and R. E. Díaz, "Perfect plane-wave injection into a finite FDTD domain through teleportation of fields," *Electromagnetics*, Vol. 23, 2003, pp. 187–201.

[11] Díaz, R. E., and I. Scherbatko, "A simple stackable re-radiating boundary condition (rRBC) for FDTD," *IEEE Antennas Propagat. Magazine*, Vol. 46, No. 1, 2004, pp. 124-130.

[12] Holland, R., "Threde: a free-field EMP coupling and scattering code," *IEEE Trans. Nuclear Science*, Vol. 24, 1977, pp. 2416–2421.

[13] Holland, R., L. Simpson, and K. S. Kunz, "Finite-difference analysis of EMP coupling to lossy dielectric structures," *IEEE Trans. Electromagn. Compat.*, Vol. 22, 1980, pp. 203–209.

PROBLEMS

5.1 Using the magic time-step in a one-dimensional FDTD grid, implement the hard source of (5.1), assuming $+x$-directed propagation. Use a resolution of 20 grid cells per wavelength. Set up a reflecting barrier in the grid, and demonstrate that the hard source causes a retroreflection artifact.

5.2 Using the magic time-step in a one-dimensional FDTD grid, implement the hard source of (5.2), assuming $+x$-directed propagation. Use a resolution of 20 or more grid cells between the $1/e$ points of the pulse.

5.3 Using the magic time-step in a one-dimensional FDTD grid, implement the logically removed hard source of (5.4). Set up a reflecting barrier in the grid and demonstrate that the pulse width and distance to the reflecting barrier can be adjusted to permit zero retroreflection artifact.

5.4 Using the magic time-step in a one-dimensional FDTD grid, implement the total-field / scattered-field wave-source condition of (5.28), (5.32), (5.34), and (5.35). Replicate the results of Fig. 5.9.

PROJECTS

P5.1 Develop a working two-dimensional square-cell Yee TE_z grid with simple outer boundary conditions (PEC sheets). Then, replicate the results of Fig. 5.1 and Table 5.1. Be sure to adjust the grid size so that numerical wave reflections from the outer grid boundary do not contaminate the results in the field-observation region. Use the time-step relation $c\Delta t = \Delta/\sqrt{2}$.

P5.2 Develop a working three-dimensional cubic-cell Yee space lattice with simple outer boundary conditions (PEC sheets). Replicate the results of Figs. 5.3 to 5.5, using a smaller space lattice as needed to keep the problem in the DRAM of your computer. Use the time-step relation $c\Delta t = \Delta/\sqrt{3}$.

P5.3 Use the FDTD code developed in Project P5.2 to replicate the results of Figs. 5.6 and 5.7.

P5.4 Implement total-field / scattered-field zoning for a two-dimensional, square-cell TE_z grid over the incident-angle range $0^\circ \le \phi \le 90^\circ$. Use both single-frequency compensated and MND incident-wave grids. Replicate the results of Figs. 5.12 and 5.13.

P5.5 Modify the FDTD code of Project P5.4 to allow the incident-angle range $90^\circ < \phi \le 180^\circ$. Check the programming by obtaining the mirror-image results of Figs. 5.12 and 5.13 for a plane wave incident at $\phi = 135^\circ$.

P5.6 Replicate the results of Project P5.4, but use instead the scattered-field formulation of Section 5.10.1.

P5.7 Develop a working two-dimensional square-cell Yee TM_z grid with simple outer boundary conditions (PEC sheets). After adjusting the grid size so that numerical wave reflections from the outer grid boundary do not contaminate the results in the field-observation region, simulate the lowest order (TE_1) waveguide mode propagating between two parallel PEC plates. Use the pulsed E-field modal hard source of Section 5.11.1 for a time waveform similar to (5.3). Adjust the carrier frequency and pulse decay to implement first a narrowband excitation well above cutoff, and then a broadband excitation with spectral components near or even below cutoff. Graph or visualize the fields in the waveguide. Justify the dispersive nature of the propagating pulse.

P5.8 Adapt the TM_z FDTD code developed in Project P5.7 to use the total-field / reflected-field modal source of Section 5.10.2. Contrast the operation of this source with that of the modal hard source of Project P5.7, considering both the accuracy of the generated waveguide mode and suppression of the source-retroreflection artifact.

P5.9 Adapt the TM_z FDTD code developed in Project P5.7 to implement a single-component, hard-sourced E_z excitation at the center of the waveguide. Observe the electrical length needed for the generated fields to settle into the fundamental waveguide mode. Justify this distance. Contrast the operation of this source with that of the modal hard source of Project P5.7, considering both the accuracy of the generated waveguide mode and suppression of the source-retroreflection artifact.

Chapter 6

Analytical Absorbing Boundary Conditions

6.1 INTRODUCTION

A basic consideration with FDTD modeling of electromagnetic wave interaction problems is that many are defined in "open" regions where the spatial domain of the computed field is unbounded in one or more coordinate directions. Clearly, no computer can store an unlimited amount of data, and therefore, the field computation domain must be limited in size. The computation domain must be large enough to enclose the structure of interest, and a suitable boundary condition on the outer perimeter of the domain must be used to simulate its extension to infinity.

Fig. 6.1 shows schematically such a finite computational domain Ω. In the interior of Ω, we apply one of the FDTD algorithms of Chapter 3 that models wave propagation in all directions. On $\partial\Omega$, the outer boundary of Ω, only numerical wave motion outward from Ω is desired. Here, we need a boundary condition that permits all outward-propagating numerical waves to exit Ω as if the simulation were performed on a computational domain of infinite extent. This boundary condition must suppress spurious reflections of outgoing numerical waves to an acceptable level, permitting the FDTD solution to remain valid for all time-steps, especially after the reflected numerical waves return to the vicinity of the modeled structure. Depending upon their theoretical basis, outer-boundary conditions of this type have been called either *radiation boundary conditions* (RBCs) or *absorbing boundary conditions* (ABCs). For simplicity, the notation ABC will be used in the remainder of this book to describe all outer-boundary conditions used to simulate the extension of an FDTD computational domain to infinity.

ABCs cannot be directly obtained from the numerical algorithms for Maxwell's curl equations defined by the finite-difference systems reviewed in Chapter 3. Principally, this is because these systems employ a central spatial-difference scheme that requires knowledge of the field which is one-half space cell to each side of an observation point. Central differences cannot be implemented at the outermost planes of the space lattice, since by definition there is no information concerning the fields at points one-half space cell outside of these planes. Although backward finite differences could conceivably be used here, these are generally of lower accuracy for a given space discretization and have not been used in major FDTD software.

Modern ABCs have capabilities for virtually reflection-free truncation of two-dimensional and three-dimensional FDTD space lattices in free space, in lossy or dispersive materials, or in dispersive metal or dielectric waveguides. Extremely small local reflection coefficients in the order of 10^{-4} to 10^{-6} can be attained with an acceptable computational burden. These ABCs allow the possibility of achieving FDTD simulations having a wide dynamic range of 70 dB or more. Currently, FDTD computational dynamic range is limited more by numerical dispersion and by the precision in defining the shapes of the structures being modeled than by artifacts due to reflection of numerical waves from the outer boundary of the computation space.

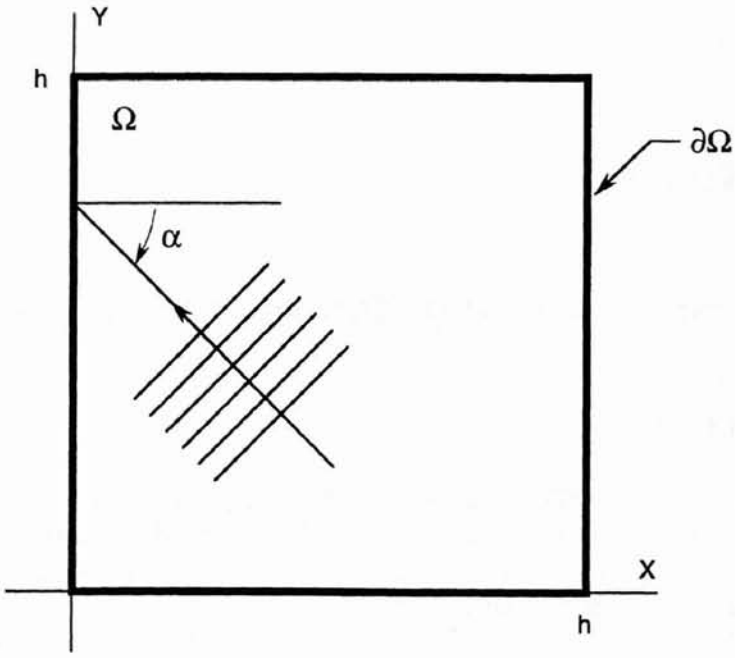


Fig. 6.1 Two-dimensional Cartesian computational domain Ω showing the outer domain boundary $\partial\Omega$ at which an ABC is implemented to simulate the extension of the grid to infinity. *Source:* Moore et al., *IEEE Trans. Antennas and Propagation*, 1988, pp. 1797–1812, © 1988 IEEE.

This chapter discusses in detail several important, approximate analytical ABCs that indicate the progression of research in this area.[1] The discussion is presented in the approximate chronological order that these ABCs were examined in the literature. The next chapter (Chapter 7) will provide a similarly detailed treatment of an alternative approach to realizing ABCs, the use of the *perfectly matched layer* (PML) to literally absorb outgoing waves at the outer boundary of the computation space.

6.2 BAYLISS-TURKEL RADIATION OPERATORS

Historically, the theory and application of radiation operators represents one of four major achievements in ABC theory during the 1970s and 1980s. These operators constitute a class of ABCs based upon the expansion of outward-propagating solutions of the wave equation in spherical or cylindrical coordinates.

---

[1]Using Green's theorem, it is possible to derive exact analytical expressions for the fields at the outer domain boundary in terms of known interior fields. However, such expressions are "global" in nature in that the field data required are located at a contour or surface completely enclosing the computation space. Storing and processing these data involves a large computer-resource burden. While local ABCs are only approximate, they require interior field data only in the vicinity of the outer boundary point of interest. The computer-resource advantage of local ABCs relative to global ABCs has been decisive, with almost all reported FDTD software using local ABCs.

As presented by Bayliss and Turkel [1, 2], the idea here is to construct a linear partial-differential operator from a weighted sum of three types of derivatives of the field: (1) spatial derivatives in the direction of outgoing wave propagation, (2) spatial derivatives in directions transverse to the direction of outgoing wave propagation, and (3) time derivatives. Properly constructed, this differential operator systematically "annihilates" terms of the outgoing-wave expansion, leaving a remainder term that diminishes to zero as an inverse power of R, the distance from the observation point to the origin. At any point on the outer boundary of the computation space, applying this differential operator to the local field via finite differences permits the derivative of the field in the direction of outgoing wave propagation to be accurately estimated in terms of known transverse spatial derivatives and time derivatives. Such an estimate permits the computation space to be closed, with little nonphysical wave reflection.

6.2.1 Spherical Coordinates

Consider solutions $U(R, \theta, \phi, t)$ to the three-dimensional scalar wave equation

$$\frac{\partial^2 U}{\partial t^2} = c^2 \nabla^2 U \tag{6.1}$$

The radiating solutions of the scalar wave equation (i.e., solutions propagating in directions that are outward from the origin of a spherical coordinate system) can be expanded in a convergent series of the form [3, 4]

$$\begin{aligned} U(R, \theta, \phi, t) &= \sum_{i=1}^{\infty} \frac{u_i(ct - R, \theta, \phi)}{R^i} \\ &= \frac{u_1(ct - R, \theta, \phi)}{R} + \frac{u_2(ct - R, \theta, \phi)}{R^2} + \cdots \end{aligned} \tag{6.2}$$

where the u_i are unknown functions of angular position, but are all propagating in the radial direction at the free-space phase velocity c. For a very large distance R from the origin, we note that the leading terms of (6.2) dominate. Forming the partial-derivative operator

$$L \equiv \frac{1}{c}\frac{\partial}{\partial t} + \frac{\partial}{\partial R} \tag{6.3}$$

and applying to the expansion of (6.2) yields for the limiting case of very large R:

$$\begin{aligned} \lim_{R\to\infty} LU &= \left[\frac{1}{c}\cdot\frac{u_1' \cdot c}{R} + \frac{u_1' \cdot (-1)}{R} + \frac{u_1 \cdot (-1)}{R^2}\right] \\ &\quad + \left[\frac{1}{c}\cdot\frac{u_2' \cdot c}{R^2} + \frac{u_2' \cdot (-1)}{R^2} + \frac{u_2 \cdot (-2)}{R^3}\right] \end{aligned} \tag{6.4a}$$

where the prime denotes differentiation with respect to the argument. After performing the term cancellations indicated in (6.4a), we obtain

$$\lim_{R\to\infty} LU = \frac{1}{c}\frac{\partial U}{\partial t} + \frac{\partial U}{\partial R} = \frac{u_1\cdot(-1)}{R^2} + \frac{u_2\cdot(-2)}{R^3} = O\left(R^{-2}\right) \tag{6.4b}$$

where $O(R^{-2})$ denotes the leading-order remainder term which varies as R^{-2}. We call operator L the Sommerfeld radiation condition [5]. Solving (6.4b) for $\partial U/\partial R$ yields

$$\frac{\partial U}{\partial R} = -\frac{1}{c}\frac{\partial U}{\partial t} + O\left(R^{-2}\right) \tag{6.5}$$

In principle, (6.5) can be used at the outer boundary of the computation space to estimate $\partial U/\partial R$ from known interior field values if the remainder term is neglected (i.e., assumed to be zero). As discussed above, this permits closure of the computation space. However, the remainder term in (6.5) would not be negligible, in general, unless the outer boundary were moved quite far from the origin. This would be wasteful of computer resources. Realizing this, Bayliss and Turkel sought to devise a partial differential operator similar to L of (6.3) that would have a remainder term diminishing more quickly to zero than $O(R^{-2})$.

To follow the logic of the Bayliss-Turkel procedure, let us consider first a slightly augmented version of the Sommerfeld radiation condition operator L:

$$B_1 = L + 1/R \tag{6.6}$$

which we call the Bayliss-Turkel radiation operator of order 1. Applying B_1 to outgoing wave expansion (6.2) yields

$$\begin{aligned} B_1 U = &\left[\frac{1}{c}\cdot\frac{u_1'\cdot c}{R} + \frac{u_1'\cdot(-1)}{R} + \frac{u_1\cdot(-1)}{R^2} + \frac{1}{R}\cdot\frac{u_1}{R}\right] \\ &+ \left[\frac{1}{c}\cdot\frac{u_2'\cdot c}{R^2} + \frac{u_2'\cdot(-1)}{R^2} + \frac{u_2\cdot(-2)}{R^3} + \frac{1}{R}\cdot\frac{u_2}{R^2}\right] \\ &+ \left[\frac{1}{c}\cdot\frac{u_3'\cdot c}{R^3} + \frac{u_3'\cdot(-1)}{R^3} + \frac{u_3\cdot(-3)}{R^4} + \frac{1}{R}\cdot\frac{u_3}{R^3}\right] + \cdots \end{aligned} \tag{6.7a}$$

After performing the term cancellations indicated in (6.7a), we obtain

$$\lim_{R\to\infty} B_1 U = -\frac{u_2}{R^3} - \frac{2u_3}{R^4} - \frac{3u_4}{R^5} = O\left(R^{-3}\right) \tag{6.7b}$$

Comparing (6.7b) with (6.4b), we see that B_1 is superior to the Sommerfeld radiation condition L in that the leading-order remainder term of B_1 diminishes to zero more quickly than that of L by one full inverse power of R.

Next, pursuing even better performance, we apply the transitional operator $B_{12} = L + (3/R)$ to what remains in (6.7b), yielding the composite operator:

$$B_2 = B_{12}B_1 = (L + 3/R)(L + 1/R) \tag{6.8}$$

which we call the Bayliss-Turkel radiation operator of order 2. Applying B_2 to (6.2) results in

$$\begin{aligned} B_2 U &= -\left[\frac{1}{c}\cdot\frac{u_2'\cdot c}{R^3} + \frac{u_2'\cdot(-1)}{R^3} + \frac{u_2\cdot(-3)}{R^4} + \frac{3}{R}\cdot\frac{u_2}{R^3}\right] \\ &\quad - 2\left[\frac{1}{c}\cdot\frac{u_3'\cdot c}{R^4} + \frac{u_3'\cdot(-1)}{R^4} + \frac{u_3\cdot(-4)}{R^5} + \frac{3}{R}\cdot\frac{u_3}{R^4}\right] \\ &\quad - 3\left[\frac{1}{c}\cdot\frac{u_4'\cdot c}{R^5} + \frac{u_4'\cdot(-1)}{R^5} + \frac{u_4\cdot(-5)}{R^6} + \frac{3}{R}\cdot\frac{u_4}{R^5}\right] - \cdots \end{aligned} \tag{6.9a}$$

After performing the term cancellations indicated in (6.9a), we obtain

$$\lim_{R\to\infty} B_2 U = \frac{2u_3}{R^5} + \frac{6u_4}{R^6} + \frac{12u_5}{R^7} = O\left(R^{-5}\right) \tag{6.9b}$$

Comparing (6.9b) with (6.7b) and (6.4b), we see that B_2 is superior to both B_1 and L. The leading-order remainder term of B_2 diminishes to zero more quickly than that of B_1 by two full inverse powers of R. Relative to L, the leading-order remainder term of B_2 diminishes to zero more quickly by three full inverse powers of R.

It can be shown that the procedure discussed above can continue indefinitely via a recursion relation to construct B_n, the Bayliss-Turkel radiation operator of order n:

$$\begin{aligned} B_n &= \prod_{k=1}^{n}\left(L + \frac{2k-1}{R}\right) = \left(L + \frac{2n-1}{R}\right) B_{n-1} \\ &= \left(L + \frac{2n-1}{R}\right)\cdots\left(L + \frac{5}{R}\right)\left(L + \frac{3}{R}\right)\left(L + \frac{1}{R}\right) \end{aligned} \tag{6.10}$$

B_n annihilates the first n terms of outgoing wave expansion (6.2). The leading-order remainder term of B_n is given by

$$\lim_{R\to\infty} B_n U = O\left(R^{-2n-1}\right) \tag{6.11}$$

To emphasize a remarkable and key point that might otherwise be lost in the mathematics, the Bayliss-Turkel radiation operators are constructed to achieve their annihilation property *without our having any knowledge whatsoever* of the angular dependence of the partial wavefunctions u_i that make up outgoing wave expansion (6.2). Further, the user can select the order of B_n as necessary to annihilate the dominant expected outgoing waves. Of course, the complexity of the operator and its numerical realization increases significantly with its order. In the research literature, B_2 has been the most popular radiation operator for providing an ABC for spherical meshes, or for more general meshes having locally defined curvatures that permit exploitation of the Bayliss-Turkel formulation. Apparently, B_2 represents a good compromise between accuracy and complexity of implementation. Depending on the radial distance at which the outer computational boundary is located, B_2 yields reflection coefficients on the order of 1%.

6.2.2 Cylindrical Coordinates

The derivation of B_n for use with cylindrical wavefunctions $U(r, \phi, t)$ in two space dimensions proceeds in a manner analogous to the above. The radiating solutions of the scalar wave equation (i.e., solutions propagating in directions that are outward from the origin of a cylindrical coordinate system) can be expanded in a convergent series of the form [6]

$$\begin{aligned} U(r, \phi, t) &= \sum_{i=0}^{\infty} \frac{u_i(ct - r, \phi)}{r^{i+1/2}} \\ &= \frac{u_1(ct - r, \phi)}{r^{1/2}} + \frac{u_2(ct - r, \phi)}{r^{3/2}} + \cdots \end{aligned} \tag{6.12}$$

where the u_i are unknown functions of angular position, but are all propagating in the radial direction at the free-space phase velocity c. For a very large distance r from the origin, we note that the leading terms of (6.12) dominate. Again forming the Sommerfeld radiation operator L of (6.3) and applying to the expansion of (6.12) yields in the limit of very large r:

$$\begin{aligned} \lim_{r\to\infty} LU &= \left[\frac{1}{c}\cdot\frac{u_1'\cdot c}{r^{1/2}} + \frac{u_1'\cdot(-1)}{r^{1/2}} + \frac{u_1\cdot(-1/2)}{r^{3/2}}\right] \\ &+ \left[\frac{1}{c}\cdot\frac{u_2'\cdot c}{r^{3/2}} + \frac{u_2'\cdot(-1)}{r^{3/2}} + \frac{u_2\cdot(-3/2)}{r^{5/2}}\right] \end{aligned} \tag{6.13a}$$

After performing the term cancellations indicated in (6.13a), we obtain

$$\lim_{r\to\infty} LU = \frac{u_1\cdot(-1/2)}{r^{3/2}} + \frac{u_2\cdot(-3/2)}{r^{5/2}} = O\left(r^{-3/2}\right) \tag{6.13b}$$

By analogy with the procedure discussed in Section 6.2.1, we consider first a slightly augmented version of L. We call this the Bayliss-Turkel radiation operator of order 1 in cylindrical coordinates:

$$B_1 = L + 1/2r \tag{6.14}$$

Applying B_1 of (6.14) to outgoing wave expansion (6.12) yields

$$\begin{aligned} B_1U &= \left[\frac{1}{c}\cdot\frac{u_1'\cdot c}{r^{1/2}} + \frac{u_1'\cdot(-1)}{r^{1/2}} + \frac{u_1\cdot(-1/2)}{r^{3/2}} + \frac{1}{2r}\cdot\frac{u_1}{r^{1/2}}\right] \\ &+ \left[\frac{1}{c}\cdot\frac{u_2'\cdot c}{r^{3/2}} + \frac{u_2'\cdot(-1)}{r^{3/2}} + \frac{u_2\cdot(-3/2)}{r^{5/2}} + \frac{1}{2r}\cdot\frac{u_2}{r^{3/2}}\right] \\ &+ \left[\frac{1}{c}\cdot\frac{u_3'\cdot c}{r^{5/2}} + \frac{u_3'\cdot(-1)}{r^{5/2}} + \frac{u_3\cdot(-5/2)}{r^{7/2}} + \frac{1}{2r}\cdot\frac{u_3}{r^{5/2}}\right] + \cdots \end{aligned} \tag{6.15a}$$

After performing the term cancellations indicated in (6.15a), we obtain

$$\lim_{r\to\infty} B_1 U = -\frac{u_2}{r^{5/2}} - \frac{2u_3}{r^{7/2}} = O\left(r^{-5/2}\right) \tag{6.15b}$$

Comparing (6.15b) with (6.13b), we see that B_1 is superior to the Sommerfeld radiation operator L, in that the leading-order remainder term of B_1 diminishes to zero more quickly than that of L by one full inverse power of r. This is the same degree of error reduction noted for B_1 versus L in the three-dimensional case.

Again pursuing even better performance, we apply the transitional operator $B_{12} = L + (5/2r)$ to what remains in (6.15b). This yields the composite operator B_2, which we call the Bayliss-Turkel radiation operator of order 2 in cylindrical coordinates:

$$B_2 = B_{12} B_1 = (L + 5/2r)(L + 1/2r) \tag{6.16}$$

Applying B_2 of (6.16) to outgoing wave expansion (6.12) results in

$$\begin{aligned} B_2 U = &-\left[\frac{1}{c}\cdot\frac{u_2'\cdot c}{r^{5/2}} + \frac{u_2'\cdot(-1)}{r^{5/2}} + \frac{u_2\cdot(-5/2)}{r^{7/2}} + \frac{5}{2r}\cdot\frac{u_2}{r^{5/2}}\right] \\ &- 2\left[\frac{1}{c}\cdot\frac{u_3'\cdot c}{r^{7/2}} + \frac{u_3'\cdot(-1)}{r^{7/2}} + \frac{u_3\cdot(-7/2)}{r^{9/2}} + \frac{5}{2r}\cdot\frac{u_3}{r^{7/2}}\right] \\ &- 3\left[\frac{1}{c}\cdot\frac{u_4'\cdot c}{r^{9/2}} + \frac{u_4'\cdot(-1)}{r^{9/2}} + \frac{u_4\cdot(-9/2)}{r^{11/2}} + \frac{5}{2r}\cdot\frac{u_4}{r^{9/2}}\right] - \cdots \end{aligned} \tag{6.17a}$$

After performing the term cancellations indicated in (6.17a), we obtain

$$\lim_{r\to\infty} B_2 U = \frac{2u_3}{r^{9/2}} + \frac{6u_4}{r^{11/2}} = O\left(r^{-9/2}\right) \tag{6.17b}$$

Comparing (6.17b) with (6.15b) and (6.13b), we see that B_2 is superior to both B_1 and L. The leading-order remainder term of B_2 diminishes to zero more quickly than that of B_1 by two full inverse powers of r. Relative to L, the leading-order remainder term of B_2 diminishes to zero more quickly by three full inverse powers of r. These accuracy advantages of B_2 versus B_1 and L are the same as for the three-dimensional case discussed earlier.

It can be shown that this procedure can continue indefinitely via a recursion relation to construct B_n, the Bayliss-Turkel radiation operator of order n in cylindrical coordinates:

$$\begin{aligned} B_n &= \prod_{k=1}^{n}\left(L + \frac{4k-3}{2r}\right) = \left(L + \frac{4n-3}{2r}\right) B_{n-1} \\ &= \left(L + \frac{4n-3}{2r}\right)\cdots\left(L + \frac{9}{2r}\right)\left(L + \frac{5}{2r}\right)\left(L + \frac{1}{2r}\right) \end{aligned} \tag{6.18}$$

B_n annihilates the first n terms of outgoing wave expansion (6.12). The leading-order remainder term of B_n is given by

$$\lim_{r\to\infty} B_n U = O\left(r^{-2n-1/2}\right) \tag{6.19}$$

B_2 has been the most popular radiation operator for providing an ABC for cylindrical grids, or for more general grids having locally defined curvatures that permit exploitation of the Bayliss-Turkel formulation. This is because B_2 represents a good compromise between accuracy and complexity of implementation. Depending on the radial distance at which the outer computational boundary is located, B_2 yields reflection coefficients on the order of 1%.

Although Bayliss-Turkel operators can be adapted to provide ABCs for two- and three-dimensional Cartesian FDTD grids, the adaptation is really a forced one. The outer grid boundary in Cartesian space is no longer at a fixed radial distance from the center of the grid, and finite-difference realizations of the necessary azimuthal spatial derivatives no longer lie completely within the computation space. Here, a better approach is to define a local coordinate system at the outer boundary that follows natural grid planes, rather than cutting across them. This is the approach of the next group of ABCs to be considered.

6.3 ENGQUIST-MAJDA ONE-WAY WAVE EQUATIONS

A partial-differential equation that permits wave propagation only in certain directions is called a *one-way wave equation.* When applied at the outer boundary of an FDTD computation space, a one-way wave equation numerically absorbs impinging outgoing waves. The theory and numerical application of such equations constitute the second of the principal thrusts in ABC technology during the 1970s and 1980s.

Engquist and Majda derived a theory of one-way wave equations suitable for ABCs in Cartesian FDTD grids [7]. Their theory can be explained in terms of factoring partial-derivative operators. For example, consider the two-dimensional wave equation in Cartesian coordinates:

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} - \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} = 0 \tag{6.20}$$

where U is a scalar field component and c is the wave phase velocity. We can define the partial-differential operator

$$G \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \equiv D_x^2 + D_y^2 - \frac{1}{c^2}D_t^2 \tag{6.21}$$

The wave equation is then compactly written as

$$GU = 0 \tag{6.22}$$

It can be shown that the wave operator G can be factored in the following manner:

$$GU = G^+ G^- U = 0 \tag{6.23a}$$

where G^- and G^+ are defined as

$$G^- \equiv D_x - \frac{D_t}{c}\sqrt{1-s^2} \;\; ; \;\; G^+ \equiv D_x + \frac{D_t}{c}\sqrt{1-s^2} \tag{6.23b, c}$$

with

$$s \equiv \frac{D_y}{(D_t/c)} \tag{6.23d}$$

Engquist and Majda showed that at the left and right grid boundaries in Fig. 6.1, applying G^- and G^+, respectively, to the wavefunction U exactly absorbs a plane wave propagating toward the boundary at *any* angle α. Thus, the conditions

$$G^-U\big|_{x=0} = 0 \; ; \qquad G^+U\big|_{x=h} = 0 \tag{6.24a, b}$$

function as exact analytical ABCs for general leftward and rightward wave motion originating from the interior of the spatial domain Ω.

6.3.1 One-Term and Two-Term Taylor Series Approximations

The presence of the square-root functions in (6.23b, c) classifies G^- and G^+ as pseudodifferential operators that are nonlocal in both space and time. This is an unfortunate characteristic, in that it prohibits the direct numerical implementation of (6.24) as an ABC. However, we can approximate the square-root function to produce a family of normal partial-differential equations that *can* be implemented numerically and are useful in FDTD simulations. These approximations are not exact, in that a small amount of reflection does develop as numerical waves pass through the outer boundary of the spatial domain. However, it is possible to design an ABC that minimizes the reflection over a meaningful range of outgoing wave angles.

Consider the use of a Taylor series to approximate the square-root functions in (6.23b, c). A one-term series of this type suitable for very small values of s is simply

$$\sqrt{1-s^2} \cong 1 \tag{6.25a}$$

Note that a very small value of s means that the y partial derivative of the outgoing wave is negligible compared to the time derivative of the wave scaled by c. Equivalently, the outgoing wave impinges upon the $x = 0$ or $x = h$ grid boundary at an angle very close to broadside. For example, consider the $x = 0$ grid boundary. Substituting (6.25a) into (6.23b), we obtain

$$G^- \cong D_x - D_t/c \tag{6.25b}$$

Now, substituting (6.25b) into (6.24a) and identifying the differential operators as partial derivatives, we obtain the corresponding partial-differential equation that can be numerically implemented as a first-order-accurate ABC at the $x = 0$ grid boundary:

$$\frac{\partial U}{\partial x} - \frac{1}{c}\frac{\partial U}{\partial t} = 0 \tag{6.26}$$

Next, we consider the use of a two-term Taylor series to approximate the square-root functions in (6.23b, c):

$$\sqrt{1 - s^2} \cong 1 - s^2/2 \tag{6.27a}$$

Clearly, this approximation remains useful for larger values of s than that of (6.25a), or equivalently, for waves impinging upon the left or right grid boundary at angles somewhat further from broadside. For example, substituting (6.27a) into (6.23b), we obtain at $x = 0$:

$$\begin{aligned} G^- &\cong D_x - \frac{D_t}{c}\left(1 - \frac{1}{2}s^2\right) \\ &\cong D_x - \frac{D_t}{c}\left[1 - \frac{1}{2}\left(\frac{c\,D_y}{D_t}\right)^2\right] \\ &\cong D_x - \frac{D_t}{c} + \frac{c\,D_y^{\,2}}{2D_t} \end{aligned} \tag{6.27b}$$

After substituting (6.27b) into (6.24a), multiplying through by D_t, and identifying the differential operators as partial derivatives, we obtain the corresponding partial-differential equation that can be numerically implemented as a second-order-accurate ABC at the $x = 0$ grid boundary:

$$\frac{\partial^2 U}{\partial x\,\partial t} - \frac{1}{c}\frac{\partial^2 U}{\partial t^2} + \frac{c}{2}\frac{\partial^2 U}{\partial y^2} = 0 \tag{6.28a}$$

Analogous approximate analytical ABCs can be derived for the other grid boundaries of Fig. 6.1:

$$\frac{\partial^2 U}{\partial x\,\partial t} + \frac{1}{c}\frac{\partial^2 U}{\partial t^2} - \frac{c}{2}\frac{\partial^2 U}{\partial y^2} = 0, \quad x = h \text{ boundary} \tag{6.28b}$$

$$\frac{\partial^2 U}{\partial y\,\partial t} - \frac{1}{c}\frac{\partial^2 U}{\partial t^2} + \frac{c}{2}\frac{\partial^2 U}{\partial x^2} = 0, \quad y = 0 \text{ boundary} \tag{6.28c}$$

$$\frac{\partial^2 U}{\partial y\,\partial t} + \frac{1}{c}\frac{\partial^2 U}{\partial t^2} - \frac{c}{2}\frac{\partial^2 U}{\partial x^2} = 0, \quad y = h \text{ boundary} \tag{6.28d}$$

For the FDTD simulation of Maxwell's equations, the ABCs of (6.28) are applied to individual Cartesian components of $\boldsymbol{E}$ and $\boldsymbol{H}$ that are located at, and tangential to, the outer boundaries of the space lattice.

The derivation of ABCs for the three-dimensional case follows the above development closely. The wave equation, given by

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} = 0 \tag{6.29a}$$

has the associated partial-differential operator

$$G \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \equiv D_x^2 + D_y^2 + D_z^2 - \frac{1}{c^2}D_t^2 \tag{6.29b}$$

G can be factored in the manner of (6.23a) to provide an exact absorbing boundary operator G^- at $x = 0$ having the same form as that of (6.23b), but with s given by

$$s \equiv \sqrt{\left(\frac{D_y}{D_t / c}\right)^2 + \left(\frac{D_z}{D_t / c}\right)^2} \tag{6.29c}$$

Again, (6.24a) holds, where G^- applied to scalar wavefunction U at the $x = 0$ boundary of the space lattice exactly absorbs a plane wave propagating toward this boundary at an arbitrary angle.

The first-order ABC in three dimensions, which uses a one-term Taylor series expansion of (6.23b), remains the same as (6.26) obtained for the two-dimensional case. However, the second-order ABC in three dimensions, which uses a two-term Taylor series expansion, has a z partial-derivative operator added to (6.27b):

$$\left(D_x - \frac{D_t}{c} + \frac{cD_y^2}{2D_t} + \frac{cD_z^2}{2D_t}\right) U = 0 \tag{6.30}$$

Multiplying (6.30) by D_t, and identifying the differential operators as partial derivatives, yields the desired ABC at $x = 0$:

$$\frac{\partial^2 U}{\partial x\, \partial t} - \frac{1}{c}\frac{\partial^2 U}{\partial t^2} + \frac{c}{2}\frac{\partial^2 U}{\partial y^2} + \frac{c}{2}\frac{\partial^2 U}{\partial z^2} = 0 \tag{6.31a}$$

Analogous second-order ABCs can be derived for the other outer boundaries of the space lattice:

$$\frac{\partial^2 U}{\partial x\, \partial t} + \frac{1}{c}\frac{\partial^2 U}{\partial t^2} - \frac{c}{2}\frac{\partial^2 U}{\partial y^2} - \frac{c}{2}\frac{\partial^2 U}{\partial z^2} = 0, \quad x = h \text{ boundary} \tag{6.31b}$$

$$\frac{\partial^2 U}{\partial y\, \partial t} - \frac{1}{c}\frac{\partial^2 U}{\partial t^2} + \frac{c}{2}\frac{\partial^2 U}{\partial x^2} + \frac{c}{2}\frac{\partial^2 U}{\partial z^2} = 0, \quad y = 0 \text{ boundary} \tag{6.31c}$$

$$\frac{\partial^2 U}{\partial y\, \partial t} + \frac{1}{c}\frac{\partial^2 U}{\partial t^2} - \frac{c}{2}\frac{\partial^2 U}{\partial x^2} - \frac{c}{2}\frac{\partial^2 U}{\partial z^2} = 0, \quad y = h \text{ boundary} \tag{6.31d}$$

$$\frac{\partial^2 U}{\partial z\,\partial t} - \frac{1}{c}\frac{\partial^2 U}{\partial t^2} + \frac{c}{2}\frac{\partial^2 U}{\partial x^2} + \frac{c}{2}\frac{\partial^2 U}{\partial y^2} = 0, \quad z = 0 \text{ boundary} \tag{6.31e}$$

$$\frac{\partial^2 U}{\partial z\,\partial t} + \frac{1}{c}\frac{\partial^2 U}{\partial t^2} - \frac{c}{2}\frac{\partial^2 U}{\partial x^2} - \frac{c}{2}\frac{\partial^2 U}{\partial y^2} = 0, \quad z = h \text{ boundary} \tag{6.31f}$$

For the FDTD simulation of Maxwell's equations, the ABCs of (6.31) are applied to individual Cartesian components of $\boldsymbol{E}$ and $\boldsymbol{H}$ that are located at, and tangential to, each of the outer boundaries of the space lattice.

The approximate second-order ABCs derived above have been found to be useful in FDTD codes when implemented using the finite-difference scheme reported by Mur [8], discussed next. These ABCs truncate FDTD space lattices with an overall level of spurious reflections in the range of 1% to 5%, sufficiently low to provide simulations having engineering value.

6.3.2 Mur Finite-Difference Scheme

Mur introduced a simple and successful finite-difference scheme for the ABCs of (6.28a–d) [8]. For clarity, this scheme is illustrated at the $x = 0$ grid boundary. Let $W|_{0,j}$ represent a Cartesian component of $\boldsymbol{E}$ or $\boldsymbol{H}$ located in the Yee grid at $x = 0$ and tangential to this boundary. Mur implemented the partial derivatives of (6.28a) as numerical central differences expanded about an auxiliary gridpoint $(1/2, j)$. In the first step, the mixed x and t derivative is written as

$$\begin{aligned}\left.\frac{\partial^2 W}{\partial x\,\partial t}\right|^n_{1/2,j} &= \frac{1}{2\Delta t}\left(\left.\frac{\partial W}{\partial x}\right|^{n+1}_{1/2,j} - \left.\frac{\partial W}{\partial x}\right|^{n-1}_{1/2,j}\right) \\ &= \frac{1}{2\Delta t}\left[\left(\frac{W|^{n+1}_{1,j} - W|^{n+1}_{0,j}}{\Delta x}\right) - \left(\frac{W|^{n-1}_{1,j} - W|^{n-1}_{0,j}}{\Delta x}\right)\right]\end{aligned} \tag{6.32a}$$

Next, the second time derivative is written as the average of the second time derivatives at the adjacent points $(0, j)$ and $(1, j)$:

$$\begin{aligned}\left.\frac{\partial^2 W}{\partial t^2}\right|^n_{1/2,j} &= \frac{1}{2}\left(\left.\frac{\partial^2 W}{\partial t^2}\right|^n_{0,j} + \left.\frac{\partial^2 W}{\partial t^2}\right|^n_{1,j}\right) \\ &= \frac{1}{2}\left\{\left[\frac{W|^{n+1}_{0,j} - 2W|^n_{0,j} + W|^{n-1}_{0,j}}{(\Delta t)^2}\right] + \left[\frac{W|^{n+1}_{1,j} - 2W|^n_{1,j} + W|^{n-1}_{1,j}}{(\Delta t)^2}\right]\right\}\end{aligned} \tag{6.32b}$$

The second y derivative is written as the average of the second y derivatives at the adjacent points $(0, j)$ and $(1, j)$:

$$\left.\frac{\partial^2 W}{\partial y^2}\right|_{1/2,j}^{n} = \frac{1}{2}\left(\left.\frac{\partial^2 W}{\partial y^2}\right|_{0,j}^{n} + \left.\frac{\partial^2 W}{\partial y^2}\right|_{1,j}^{n}\right)$$

$$= \frac{1}{2}\left\{\left[\frac{W|_{0,j+1}^{n} - 2W|_{0,j}^{n} + W|_{0,j-1}^{n}}{(\Delta y)^2}\right] + \left[\frac{W|_{1,j+1}^{n} - 2W|_{1,j}^{n} + W|_{1,j-1}^{n}}{(\Delta y)^2}\right]\right\} \tag{6.32c}$$

Substituting the finite-difference expressions of (6.32a–c) into (6.28a) and then solving for $W|_{0,j}^{n+1}$, we obtain the following time-stepping algorithm along the $x = 0$ grid boundary:

$$\begin{aligned} W|_{0,j}^{n+1} &= -W|_{1,j}^{n-1} + \frac{c\Delta t - \Delta x}{c\Delta t + \Delta x}\left(W|_{1,j}^{n+1} + W|_{0,j}^{n-1}\right) \\ &+ \frac{2\Delta x}{c\Delta t + \Delta x}\left(W|_{0,j}^{n} + W|_{1,j}^{n}\right) \\ &+ \frac{(c\Delta t)^2 \Delta x}{2(\Delta y)^2 (c\Delta t + \Delta x)}\begin{pmatrix} W|_{0,j+1}^{n} - 2W|_{0,j}^{n} + W|_{0,j-1}^{n} + \\ W|_{1,j+1}^{n} - 2W|_{1,j}^{n} + W|_{1,j-1}^{n} \end{pmatrix} \end{aligned} \tag{6.33}$$

For a square grid, $\Delta x = \Delta y = \Delta$, and the Mur ABC at $x = 0$ can be written as

$$\begin{aligned} W|_{0,j}^{n+1} &= -W|_{1,j}^{n-1} + \frac{c\Delta t - \Delta}{c\Delta t + \Delta}\left(W|_{1,j}^{n+1} + W|_{0,j}^{n-1}\right) + \frac{2\Delta}{c\Delta t + \Delta}\left(W|_{0,j}^{n} + W|_{1,j}^{n}\right) \\ &+ \frac{(c\Delta t)^2}{2\Delta(c\Delta t + \Delta)}\begin{pmatrix} W|_{0,j+1}^{n} - 2W|_{0,j}^{n} + W|_{0,j-1}^{n} + \\ W|_{1,j+1}^{n} - 2W|_{1,j}^{n} + W|_{1,j-1}^{n} \end{pmatrix} \end{aligned} \tag{6.34}$$

In the same manner, we can derive analogous finite-difference expressions for the Mur ABC at each of the other grid boundaries $x = h$, $y = 0$, and $y = h$ by substituting into (6.28b), (6.28c), and (6.28d), respectively. More simply, these expressions can be obtained by inspection from (6.33) and (6.34), using coordinate-symmetry arguments to permute the subscripts of the W's. Further, it is easy to show that the corresponding first-order Mur ABC is obtained by removing the y-derivative terms, yielding at $x = 0$:

$$W|_{0,j}^{n+1} = -W|_{1,j}^{n-1} + \frac{c\Delta t - \Delta}{c\Delta t + \Delta}\left(W|_{1,j}^{n+1} + W|_{0,j}^{n-1}\right) + \frac{2\Delta}{c\Delta t + \Delta}\left(W|_{0,j}^{n} + W|_{1,j}^{n}\right) \tag{6.35}$$

In three dimensions, the derivation of the Mur finite-difference expression for the ABC at the $x = 0$ boundary of the space lattice follows the above development closely. Here, the Mur scheme involves implementing the partial derivatives of (6.31a) as numerical central differences expanded about the auxiliary gridpoint $(1/2, j, k)$. The second derivatives, $\partial^2 W/\partial x\,\partial t$, $\partial^2 W/\partial t^2$, and $\partial^2 W/\partial y^2$ are identical in form to (6.32a), (6.32b), and (6.32c), respectively, and are evaluated in lattice plane $z = k\Delta z$. The second z-derivative, $\partial^2 W/\partial z^2$, is expressed as the average of the second z-derivatives evaluated at the adjacent points $(0, j, k)$ and $(1, j, k)$. Substituting these finite-difference expressions into (6.31a) and solving for $W|_{0,j,k}^{n+1}$, we obtain the following time-stepping algorithm along the $x = 0$ lattice boundary:

$$
\begin{aligned}
W|_{0,j,k}^{n+1} &= -W|_{1,j,k}^{n-1} + \frac{c\Delta t - \Delta x}{c\Delta t + \Delta x}\left(W|_{1,j,k}^{n+1} + W|_{0,j,k}^{n-1}\right) \\
&+ \frac{2\Delta x}{c\Delta t + \Delta x}\left(W|_{0,j,k}^{n} + W|_{1,j,k}^{n}\right) \\
&+ \frac{(c\Delta t)^2\,\Delta x}{2\Delta y^2(c\Delta t + \Delta x)}\begin{pmatrix} W|_{0,j+1,k}^{n} - 2W|_{0,j,k}^{n} + W|_{0,j-1,k}^{n} + \\ W|_{1,j+1,k}^{n} - 2W|_{1,j,k}^{n} + W|_{1,j-1,k}^{n} \end{pmatrix} \\
&+ \frac{(c\Delta t)^2\,\Delta x}{2\Delta z^2(c\Delta t + \Delta x)}\begin{pmatrix} W|_{0,j,k+1}^{n} - 2W|_{0,j,k}^{n} + W|_{0,j,k-1}^{n} + \\ W|_{1,j,k+1}^{n} - 2W|_{1,j,k}^{n} + W|_{1,j,k-1}^{n} \end{pmatrix}
\end{aligned}
\tag{6.36}
$$

For a cubic lattice, $\Delta x = \Delta y = \Delta z = \Delta$, and the Mur ABC at $x = 0$ can be written as

$$
\begin{aligned}
W|_{0,j,k}^{n+1} &= -W|_{1,j,k}^{n-1} + \frac{c\Delta t - \Delta}{c\Delta t + \Delta}\left(W|_{1,j,k}^{n+1} + W|_{0,j,k}^{n-1}\right) \\
&+ \frac{2\Delta}{c\Delta t + \Delta}\left(W|_{0,j,k}^{n} + W|_{1,j,k}^{n}\right) \\
&+ \frac{(c\Delta t)^2}{2\Delta(c\Delta t + \Delta)}\begin{pmatrix} W|_{0,j+1,k}^{n} - 4W|_{0,j,k}^{n} + W|_{0,j-1,k}^{n} + \\ W|_{1,j+1,k}^{n} - 4W|_{1,j,k}^{n} + W|_{1,j-1,k}^{n} + \\ W|_{0,j,k+1}^{n} + W|_{0,j,k-1}^{n} + W|_{1,j,k+1}^{n} + W|_{1,j,k-1}^{n} \end{pmatrix}
\end{aligned}
\tag{6.37}
$$

In the same manner, we can derive analogous finite-difference expressions for the Mur ABC at each of the other lattice boundaries $x = h$, $y = 0$, $y = h$, $z = 0$, and $z = h$ by substituting into (6.31b–f), respectively. More simply, these expressions can be obtained from (6.36) and (6.37) using coordinate-symmetry arguments to permute the subscripts of the W's.

6.3.3 Trefethen-Halpern Generalized and Higher-Order ABCs

Trefethen and Halpern [9, 10] proposed a systematic means to improve the accuracy of ABCs derived from the one-way wave equation. Their idea was to achieve a better approximation of $\sqrt{1-s^2}$ in (6.23b, c) than is possible using a Taylor's series. A general interpolant $r(s)$ for this purpose was defined over the range $-1 \le s \le +1$. This interpolant was postulated to be the ratio of a numerator polynomial $p_m(s)$ of degree m, and a denominator polynomial $q_n(s)$ of degree n:

$$\sqrt{1-s^2} \equiv r(s) = \frac{p_m(s)}{q_n(s)} \tag{6.38}$$

$r(s)$ is called a rational-function interpolant of type (m, n) over the s interval $[-1, 1]$. Since $s \equiv cD_y/D_t$, implementing (6.38) is equivalent to approximating, for example, the exact one-way wave equation of (6.24a) at the $x = 0$ lattice boundary for outgoing waves ranging from $-y$-directed grazing incidence along the lattice boundary ($s = -1$) through broadside incidence ($s = 0$) to $+y$-directed grazing incidence ($s = +1$).

In perhaps the most simple case, we can specify $r(s)$ as a general (2, 0) interpolant. Here, the square-root function is approximated by a polynomial of the form

$$\sqrt{1-s^2} \equiv p_0 + p_2 s^2 \tag{6.39a}$$

This results in the general second-order approximate analytical ABC:

$$\frac{\partial^2 U}{\partial x\,\partial t} - \frac{p_0}{c}\frac{\partial^2 U}{\partial t^2} - p_2 c \frac{\partial^2 U}{\partial y^2} = 0 \tag{6.39b}$$

A more flexible interpolation is provided by the general (2, 2) rational function:

$$\sqrt{1-s^2} \equiv \frac{p_0 + p_2 s^2}{q_0 + q_2 s^2} \tag{6.40a}$$

This results in the general third-order approximate analytical ABC:

$$q_0 \frac{\partial^3 U}{\partial x\,\partial t^2} + q_2 c^2 \frac{\partial^3 U}{\partial x\,\partial y^2} - \frac{p_0}{c}\frac{\partial^3 U}{\partial t^3} - p_2 c \frac{\partial^3 U}{\partial t\,\partial y^2} = 0 \tag{6.40b}$$

Appropriate selection of the p_i and q_i in (6.39a) and (6.40a) produces various families of ABCs. Standard Padé, least-square, or Chebyshev interpolations are applied to achieve a "best fit" to $\sqrt{1-s^2}$ over all or part of the range $-1 \le s \le +1$. This produces an approximate ABC whose performance is good over all or part of the range of outgoing wave directions. The two-term Taylor's series of (6.27a) is now seen in a more general sense to be a Padé (2, 0) interpolant with the coefficients $p_0 = +1$ and $p_2 = -1/2$ in (6.39a). This provides a double zero of the wave reflection at $s = 0$ (i.e., for waves at normal incidence to the $x = 0$ boundary of the space lattice).

A more accurate interpolation involves choosing $q_0 = p_0 = 1$, $p_2 = -3/4$, and $q_2 = -1/4$ in (6.40a). This yields a Padé (2, 2) approximation that provides a triple zero of the wave reflection at normal incidence, $s = 0$. In fact, this is the third-order ABC originally proposed by Engquist and Majda in [7]. Other types of approximating polynomials "tune" the ABC to absorb numerical waves incident at specified angles other than broadside to the lattice boundary. These were considered by Trefethen and Halpern to be a means to improve the wide-angle performance of the ABC. Tables 6.1 and 6.2 list the p and q coefficients for seven techniques of approximation:

1. Padé
2. Chebyshev on a subinterval (L_a^∞ norm)
3. Interpolation in Chebyshev points
4. Least squares (L^2 norm)
5. Chebyshev-Padé (or C-P)
6. Interpolation in Newman points
7. Chebyshev (L^∞ norm)

Tables 6.1 and 6.2 also show the angles of incidence $\pm\alpha$ defined in Fig. 6.1, at which the rational-function interpolant exactly equals $\sqrt{1-s^2}$. Theoretically, the corresponding ABC exactly absorbs numerical plane waves propagating at these angles.

TABLE 6.1

Coefficients for Second-Order ABCs

Type of Approximation	p_0	p_2	*±Angles of Exact Absorption*	
Padé (reported by Mur)	1.00000	–0.50000	0°	0°
Chebyshev on a subinterval	1.00023	–0.51555	7.6°	18.7°
Chebyshev points	1.03597	–0.76537	22.5°	67.5°
Least-squares	1.03084	–0.73631	22.1°	64.4°
Chebyshev-Padé	1.06103	–0.84883	25.8°	73.9°
Newman points	1.00000	–1.00000	0°	90°
Chebyshev	1.12500	–1.00000	31.4°	81.6°

TABLE 6.2

Coefficients for Third-Order ABCs ($q_0 = 1.0$ for each technique)

Type of Approximation	p_0	p_2	q_2	*±Angles of Exact Absorption*		
Padé	1.00000	–0.75000	–0.25000	0°	0°	0°
Chebyshev on a subinterval	0.99973	–0.80864	–0.31657	11.7°	31.9°	43.5°
Chebyshev points	0.99650	–0.91296	–0.47258	15°	45°	75°
Least-squares	0.99250	–0.92233	–0.51084	18.4°	51.3°	76.6°
Chebyshev-Padé	0.99030	–0.94314	–0.5556	18.4°	53.1°	81.2°
Newman points	1.00000	–1.00000	–0.66976	0°	60.5°	90°
Chebyshev	0.95651	–0.94354	–0.70385	26.9°	66.6°	87.0°

In implementing any of the third-order ABCs of Table 6.2, the user must consider the tradeoff between improved performance relative to the Padé (2, 0) condition reported by Mur and increased complexity in the finite-difference realization. The following two sections will provide information regarding the performance side of this tradeoff.

6.3.4 Theoretical Reflection Coefficient Analysis

We can think of a radiated or scattered wave propagating outward from an arbitrary structure as being a superposition of plane waves having a range of incident angles α relative to an outer-boundary plane of the space lattice, where α is defined in Fig. 6.1. This allows the performance of a proposed ABC to be assessed theoretically by deriving the plane-wave reflection coefficient $R_{ABC}(\alpha)$ for that lattice outer-boundary plane. Clearly, an effective ABC would have a small value of R_{ABC} over a wide range of α, thereby permitting most of the impinging numerical energy to exit the lattice.

Consider the outgoing numerical plane wave in Fig. 6.1. Assuming that it is sinusoidal and propagates with the speed $c = \omega/k$, the wave has the analytical form

$$U_{inc} = e^{j(\omega t + kx\cos\alpha - ky\sin\alpha)} \tag{6.41}$$

The total field at the $x = 0$ boundary plane of the computational domain must satisfy the specific ABC in effect there. Postulating the existence of a reflected numerical wave launched from this plane, the total field has the form

$$U = e^{j(\omega t + kx\cos\alpha - ky\sin\alpha)} + R_{ABC}\, e^{j(\omega t - kx\cos\alpha - ky\sin\alpha)} \tag{6.42}$$

R_{ABC} can now be determined by substituting U of (6.42) directly into (6.39b) or (6.40b), the partial differential equations for the general second- or third-order analytical ABCs, respectively, and then setting $x = 0$ while requiring y to be arbitrary. In this manner, we obtain for the general second-order ABC [11]:

$$R_{ABC} = \frac{\cos\alpha - p_0 - p_2 \sin^2\alpha}{\cos\alpha + p_0 + p_2 \sin^2\alpha} \tag{6.43}$$

and for the general third-order ABC:

$$R_{ABC} = \frac{q_0\cos\alpha + q_2\cos\alpha\sin^2\alpha - p_0 - p_2\sin^2\alpha}{q_0\cos\alpha + q_2\cos\alpha\sin^2\alpha + p_0 + p_2\sin^2\alpha} \tag{6.44}$$

where the coefficients p and q correspond to those of the approximating function used in the derivation of the ABC (see Tables 6.1 and 6.2).

Fig. 6.2 shows the variation of the theoretical reflection coefficient R_{ABC} for second- and third-order Padé and Chebyshev ABCs as a function of α [11]. In Fig. 6.2(a), we see that $R_{ABC} < 0.1\%$ over $0 < \alpha < 35°$ for the Padé (2, 2) ABC. In Fig. 6.2(b), we see that $R_{ABC} < 0.02\%$ over $0 < \alpha < 45°$ for the type (2, 2) Chebyshev-on-a-subinterval ABC. Here, the dips in R_{ABC} near 12°, 32°, and 44° represent the angles of exact absorption for this ABC as listed in Table 6.2.

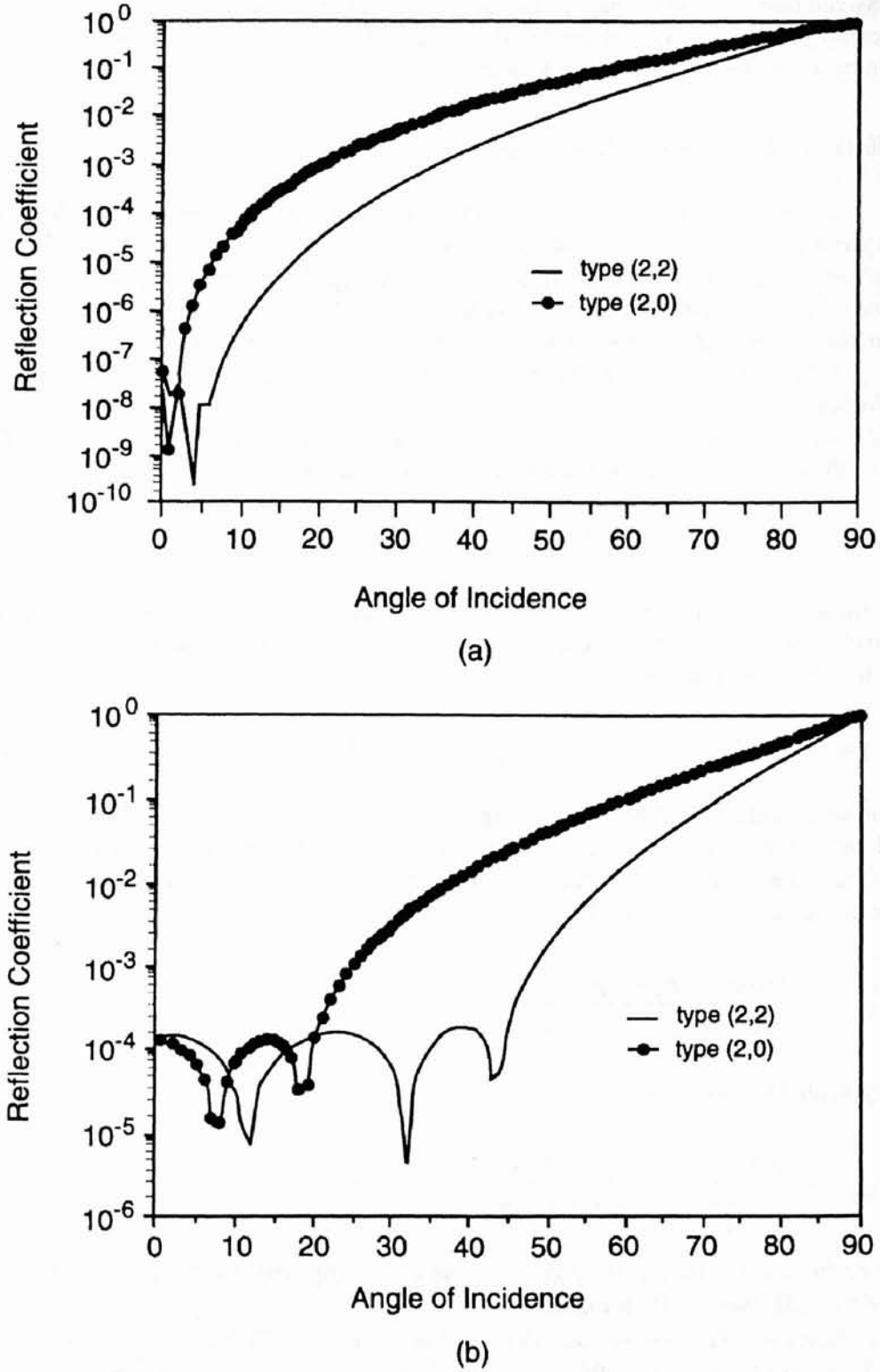


Fig. 6.2 Theoretical reflection coefficient versus incident wave angle for: (a) second- and third-order Padé ABCs; (b) second- and third-order Chebyshev-on-subinterval ABCs. *Source:* Moore et al., *IEEE Trans. Antennas and Propagation*, 1988, pp. 1797–1812, © 1988 IEEE.

On the basis of these excellent theoretical results for R_{ABC}, one would think that the solution to the ABC problem is well in hand: use a third-order condition, most likely the Chebyshev-on-a-subinterval. However, as shown in the following section, numerical experiments implementing third-order ABCs *do not* indicate reflection coefficients much below 1%. Further, these experiments *do not* indicate any performance advantage for the Chebyshev-on-a-subinterval ABC relative to the Padé ABC. We shall discuss a probable reason for this anomaly at the conclusion of the next section, after reviewing the results of the numerical experiments.

6.3.5 Numerical Experiments

We now discuss numerical experiments [11] which measure the amount of nonphysical reflection that a given ABC produces as a pulsed wave propagates through a grid outer boundary. Fig. 6.3(a) shows the two FDTD domains used in the experiments: the 100×50-cell test domain Ω_T, and the much larger benchmark domain Ω_B. On each domain, the TM_z algorithm is time-stepped. Ω_T is centered within Ω_B and has exactly the same grid-cell size, so that each field component in Ω_T corresponds to a field component in Ω_B. The size of Ω_B is selected so that its outer boundaries are so remote from Ω_T as to be causally isolated during the time-stepping. In this manner, Ω_B exactly simulates the ideal of an infinitely large FDTD grid during tests of the ABC on Ω_T.

On the outer boundary of Ω_T, a test ABC is applied. This ABC is excited by an outward-propagating cylindrical wave generated by a hard source located at the center of Ω_T, and, by design, the center of Ω_B as well. The outgoing wave is therefore spatially coincident in both domains until time-stepping advances to the point when the wave interacts with the test ABC at the outer boundary of Ω_T. Any reflection from this boundary causes the FDTD-computed field values at points within Ω_T to differ from the effective infinite-grid FDTD field values obtained at corresponding points within Ω_B. By calculating the difference between the FDTD solutions in the two domains at each gridpoint at each time-step, a measure of the spurious reflection caused by the test ABC is obtained.

We define at time-step n the local error at any point (i, j) within Ω_T due to its ABC as

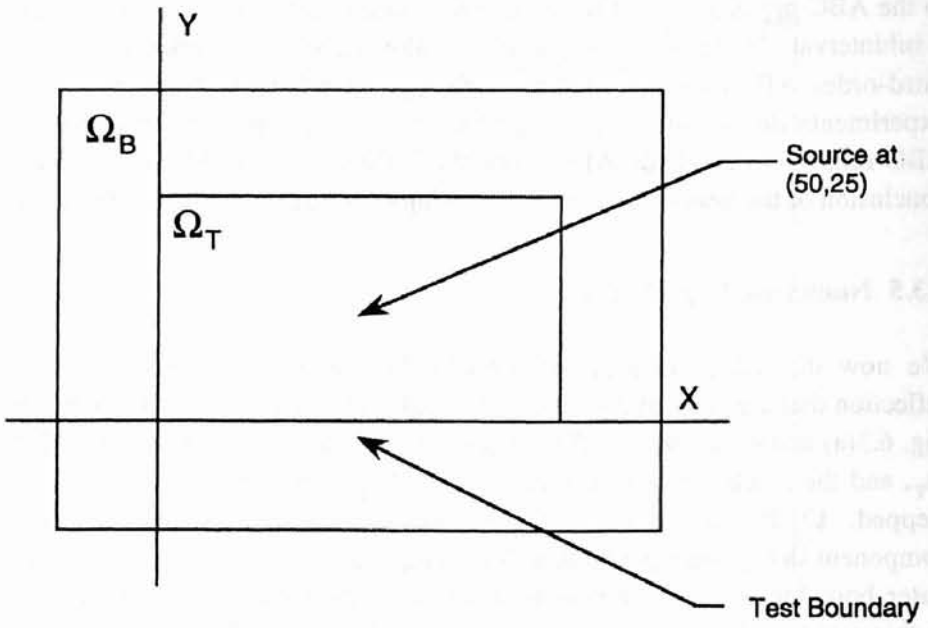
$$e_{\text{local}}\Big|_{i,j}^{n} = E_{z,T}\Big|_{i,j}^{n} - E_{z,B}\Big|_{i,j}^{n} \tag{6.45a}$$

where $E_{z,T}$ and $E_{z,B}$ are, respectively, the FDTD-computed E-fields within the test and benchmark domains. We also define the global error measure

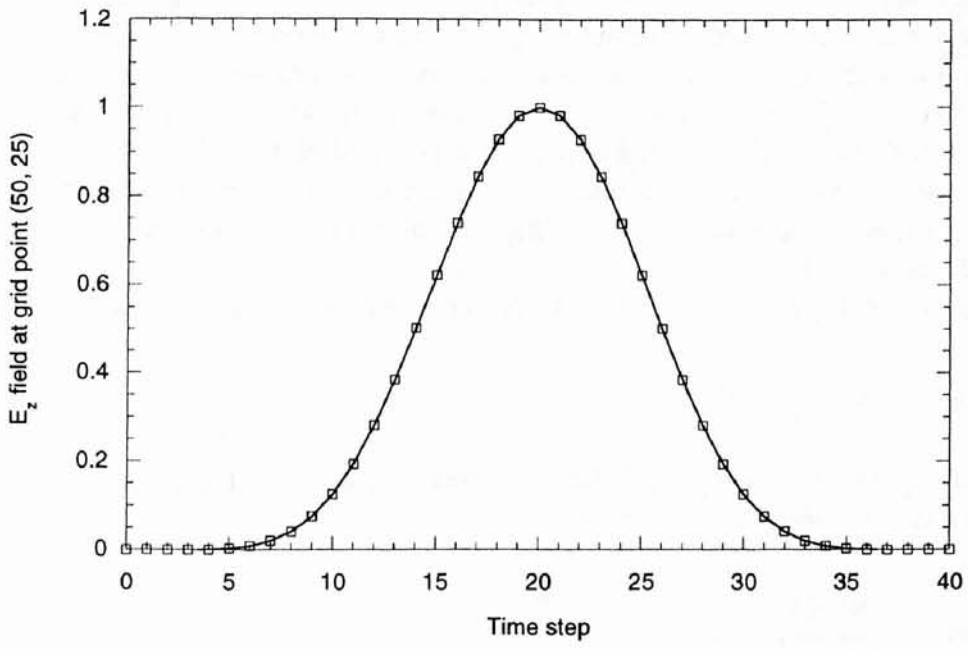
$$e_{\text{global}}\Big|^{n} = \sum_{i} \sum_{j} \left| E_{z,T}\Big|_{i,j}^{n} - E_{z,B}\Big|_{i,j}^{n} \right|^2 \tag{6.45b}$$

where the double sum is taken over all points (i, j) within Ω_T. This error measure has the units of power, and if summed over all time-steps, would provide the total error energy within Ω_T.

As shown in Fig. 6.3(b), the time waveform of the line source used in the numerical experiments is a *smooth compact pulse* peaking at $n = 20$ time-steps with a value of 1.0. The pulse excitation is provided by the following hard source located at point (50, 25) in the test grid:



(a)



(b)

Fig. 6.3 Dual-grid setup for numerical experiments testing the effectiveness of proposed ABCs implemented at the outer boundary of test grid Ω_T: (a) test and benchmark computational domains; (b) time waveform of the cylindrical- or spherical-wave source. *Source:* Moore et al., *IEEE Trans. Antennas and Propagation*, 1988, pp. 1797–1812, © 1988 IEEE.

$$E_{z,\mathrm{T}}\Big|_{50,25}^{n} = \begin{cases} \dfrac{1}{32}\left[\begin{array}{l} 10 - 15\cos(\pi n/20) + \\ 6\cos(2\pi n/20) - \cos(3\pi n/20) \end{array}\right] & n \le 40 \\ 0 & n > 40 \end{cases} \tag{6.46}$$

The analytical form of this source function, introduced in [12], causes it to have an extremely smooth transition to zero, with the first five time derivatives vanishing at $n = 0$ and at $n = 40$. As a result, this pulse has very little high-frequency content.

A square-cell grid is used in all of the numerical experiments with $c\Delta t = \Delta/2$. With this specification, the leading edge of the cylindrical wave generated by the source requires $50\Delta t$ to propagate over the 25Δ distance to the boundary of the test grid at $j = 0$. Therefore, at time-step $n = 70$, the peak of the radially propagating pulse reaches the ABC. We choose to observe the local error along $j = 1$, the first row of gridpoints away from the boundary, at time-step $n = 100$. This permits the bulk of the outgoing pulse to pass through the boundary and excite the largest possible reflection.

Fig. 6.4 graphs the local and global reflection errors observed for the second-order (2, 0) and third-order (2, 2) Padé ABCs. Fig. 6.5 compares the reflection errors of the third-order Padé and Chebyshev-on-subinterval ABCs. In each of the local-error graphs, the data are normalized to the peak value of the incident pulse striking the grid boundary at time-step $n = 70$ and grid position (50, 0). While the third-order ABCs have less error than the second-order ABCs, the local reflections are still on the order of 1%. This is a much poorer performance than the order 0.1% or less theoretical plane-wave reflection coefficients derived earlier. Further, inspection of Fig. 6.5 reveals no performance advantage for the third-order Chebyshev-on-subinterval ABC versus the third-order Padé ABC. The improved wide-angle performance suggested by Trefethen and Halpern [9, 10] is not evidenced in these numerical experiments.

These contradictory theoretical and numerical-experimental results call into question the formulation of ABCs based upon one-way wave equations. A possible explanation for the discrepancy is the effect of numerical dispersion in the grid (see Chapter 4 and [13, 14]). *All* ABC theory developed so far in this chapter assumes that numerical waves propagate in the FDTD grid at the speed c regardless of their spectral content and propagation angle. Effectively, these ABCs enforce $\tilde{v}_p = c$ at the outer boundary of the grid. However, unless extremely fine meshing is used, it is impossible to avoid the order 0.1% to 1% variations in $\tilde{v}_p$ introduced by the Yee algorithm. Without doubt, these artifacts were present in the numerical experiments providing data for Figs. 6.4 and 6.5. The mismatch between the dispersive $\tilde{v}_p \neq c$ conditions within the interior of the grid and the imposed $\tilde{v}_p = c$ condition at the outer grid boundary *must* generate wave reflections, as does any interface between media of differing wave speeds. The percentage reflection is easily estimated as the percentage that $\tilde{v}_p$ deviates from c, namely in the order of 0.1% to 1%. It is therefore futile to apply a (2, 2) Chebyshev-on-subinterval ABC and expect to realize its theoretical reflection coefficient of less than 0.02%, unless somehow this ABC is modified to account for the variability of the speed of the numerical waves impinging upon it.

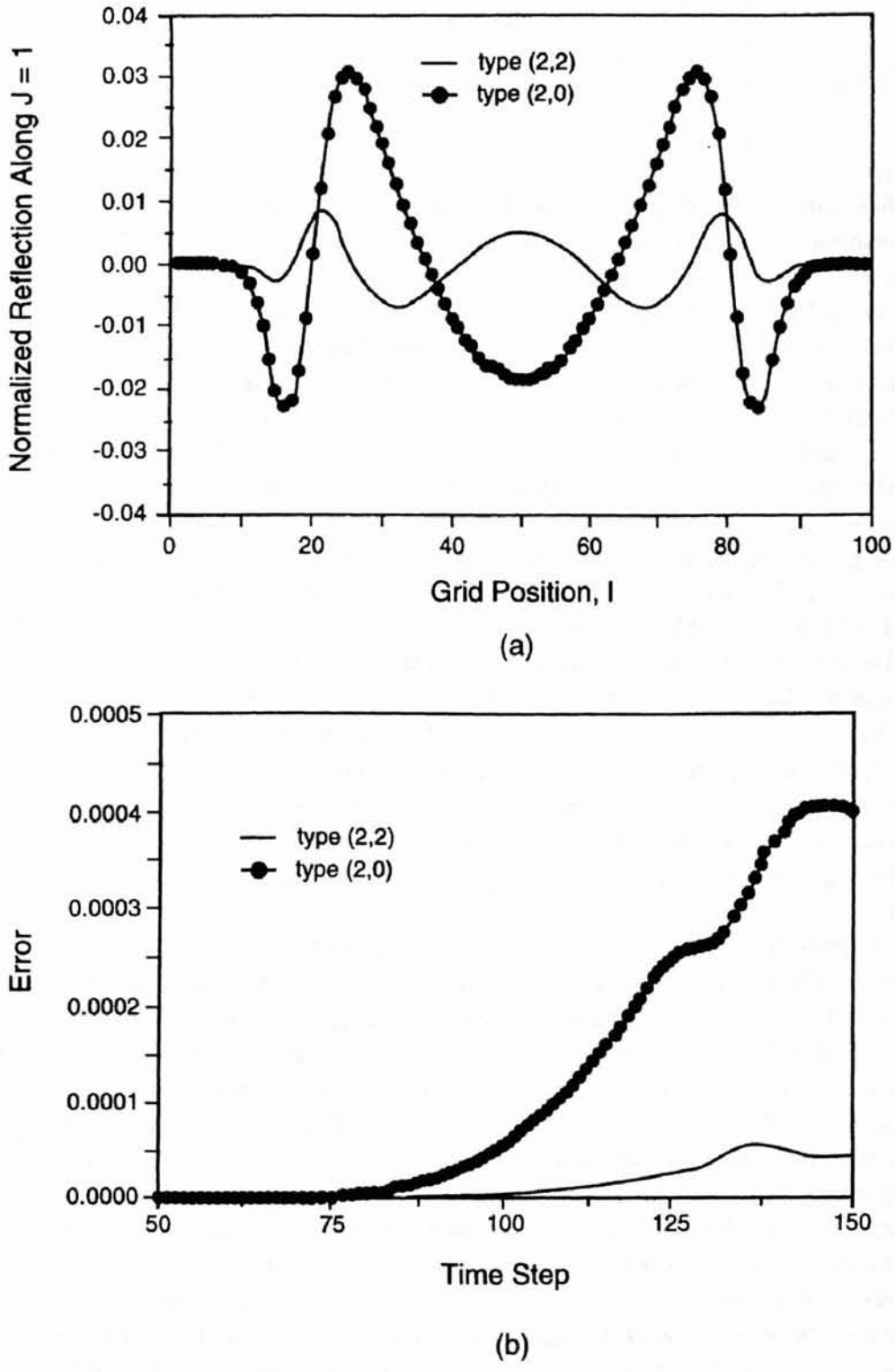
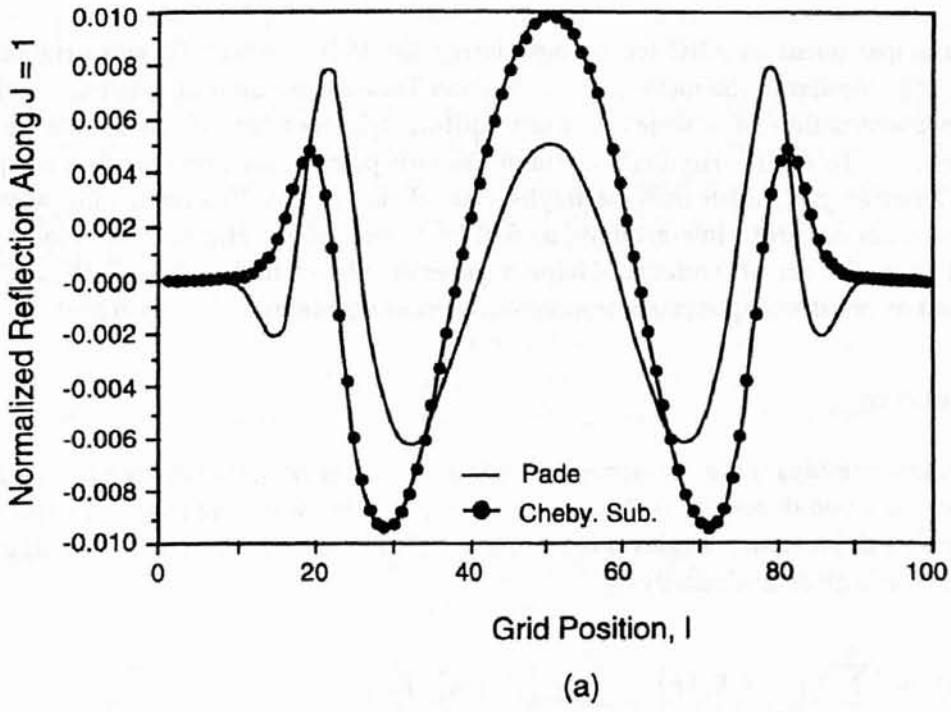


Fig. 6.4 Comparative error measures for second- and third-order Padé ABCs from the numerical experiments of Fig. 6.3: (a) local error along the test-grid outer boundary at time-step $n = 100$; (b) global error within the test grid observed over the first 150 time-steps. *Source:* Moore et al., *IEEE Trans. Antennas and Propagation*, 1988, pp. 1797–1812, © 1988 IEEE.



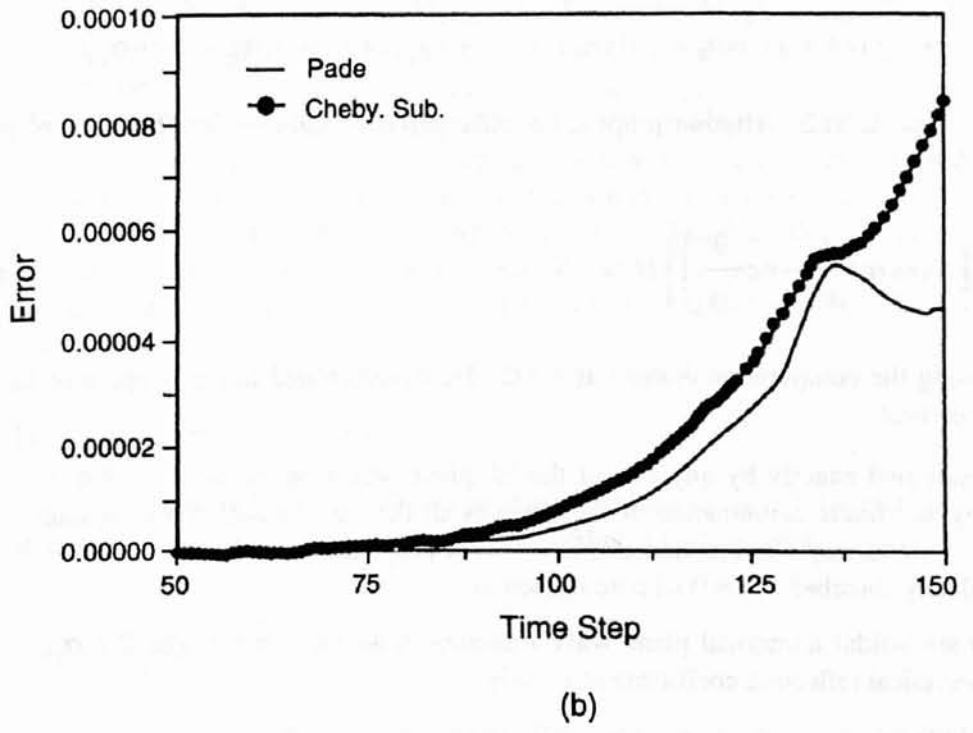


Fig. 6.5 Comparison of the error of the third-order Padé and Chebyshev-on-subinterval ABCs obtained in the numerical experiments of Fig. 6.3: (a) local error along the test-grid outer boundary at time-step $n = 100$; (b) global error within the test grid observed over the first 150 time-steps. *Source:* Moore et al., *IEEE Trans. Antennas and Propagation*, 1988, pp. 1797–1812, © 1988 IEEE.

6.4 HIGDON RADIATION OPERATORS

The third principal thrust in ABC technology during the 1970s and 1980s was originated by Higdon [15, 16]. Similar to the method of Bayliss and Turkel (Section 6.2), Higdon's technique involves the construction of a series of linear differential operators to annihilate outgoing numerical waves. However, Higdon's operators absorb plane waves propagating at specific angles in a Cartesian grid, rather than the Bayliss-Turkel sum of radially propagating waves in a cylindrical or spherical grid. Interestingly, as will be shown below, Higdon's approach yields, as a special case, the set of Trefethen-Halpern generalized and higher-order ABCs (Section 6.3.3) obtained by constructing approximate one-wave wave equations in Cartesian grids.

6.4.1 Formulation

Consider a linear combination of numerical plane waves each traveling with speed c toward the $x = 0$ boundary of a two-dimensional Cartesian FDTD grid. The waves are assumed to propagate at the symmetrical incidence angles $\pm\alpha_1, \ldots, \pm\alpha_L$ relative to the $-x$-axis. The aggregate propagating mode is given analytically by

$$\begin{aligned} U(x, y, t) &= \sum_{\ell=1}^{L} f_\ell\left(ct + \hat{\boldsymbol{k}}_\ell \cdot \hat{\boldsymbol{r}}\right) + \sum_{\ell=1}^{L} g_\ell\left(ct + \hat{\boldsymbol{k}}_\ell^* \cdot \hat{\boldsymbol{r}}\right) \\ &= f_1\left(ct + x\cos\alpha_1 + y\sin\alpha_1\right) + \ldots + f_L\left(ct + x\cos\alpha_L + y\sin\alpha_L\right) \\ &+ g_1\left(ct + x\cos\alpha_1 - y\sin\alpha_1\right) + \ldots + g_L\left(ct + x\cos\alpha_L - y\sin\alpha_L\right) \end{aligned} \tag{6.47}$$

where $-\pi/2 \le \alpha_\ell \le \pi/2$. Higdon proposed a differential annihilator for this sum of plane waves of the form

$$\left[\prod_{\ell=1}^{L}\left(\cos\alpha_\ell \frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right)\right] U = 0 \tag{6.48}$$

to permit closing the computation domain at $x = 0$. He demonstrated that this operator has the following properties:

1. It is satisfied exactly by any one of the $2L$ plane waves in the sum of (6.47), and by any linear combination of these waves all the way through the complete sum. That is, any combination of plane waves propagating at the angles α_ℓ is completely absorbed at $x = 0$ with no reflection.
2. For a sinusoidal numerical plane wave traveling at an incidence angle $\theta \ne \alpha_\ell$, the theoretical reflection coefficient at $x = 0$ is

$$R = -\prod_{\ell=1}^{L}\left(\frac{\cos\alpha_\ell - \cos\theta}{\cos\alpha_\ell + \cos\theta}\right) \tag{6.49}$$

where each factor in (6.49) has absolute value less than 1 when $|\theta| < \pi/2$.

3. For any given order L of the annihilator and any given problem, the exact wave-absorption angles α_ℓ can be chosen to optimize the overall transmission characteristics of the grid outer boundary.

4. Finite-differencing (6.48) requires no space derivatives parallel to the grid outer boundary. Thus, the stencil of gridpoints needed to implement the Higdon operator is one-dimensional in the direction perpendicular to the grid boundary. This simplifies the numerical implementation, especially near the corners of rectangular grids.

5. The operator of (6.48) provides a general representation of ABCs in the following sense. If an ABC is based on a symmetric rational approximation to the outgoing-wave portion of the factored wave equation (equivalently, the portion of the dispersion relation corresponding to outgoing waves), then it is either:

   (a) Equivalent to (6.48) for a suitable choice of the α_ℓ satisfying $|\alpha_\ell| < \pi/2$;

   (b) Unstable;

   (c) Not optimal in that the coefficients in the ABC can be modified to reduce the reflection coefficient for all Fourier modes.

The fifth point proved by Higdon is particularly interesting, since it implies that the ABCs proposed by Engquist-Majda and Trefethen-Halpern are *special cases* of the more general Higdon operator of (6.48). As stated by Givoli [17]:

> "In other words, any stable [ABC] that is derived by using a symmetric rational approximation and that cannot be improved by a simple modification of its coefficients, is characterized completely by its angles of perfect absorption. Higdon's theorem seems to imply that there is not much value in using rational approximations to derive [ABCs] for the scalar wave equation, since equivalent boundary conditions can be obtained in a simpler way by using [the Higdon operator]. The fact that [the Higdon operator] requires the choice of the parameters α_ℓ may seem a disadvantage, but one may argue that those [ABCs] which are special cases of [the Higdon operator] *implicitly* make this kind of choice for the analyst."

6.4.2 First Two Higdon Operators

The simplicity of the Higdon approach can be illustrated by writing out the first two operators of the sequence defined by (6.48). Letting $L = 1$, we obtain the first operator:

$$\cos\alpha_1 \frac{\partial U}{\partial t} - c\frac{\partial U}{\partial x} = 0 \tag{6.50}$$

This operator completely absorbs plane waves propagating with speed c at the angles $\pm\alpha_1$ with respect to the $-x$-axis. If $\alpha_1 = 0°$, (6.50) reduces to (6.26), the first-order ABC at $x = 0$ resulting from the one-term Taylor approximation of $\sqrt{1-s^2}$. We recall that this approximation implies normal incidence to the $x = 0$ boundary. The first-order Higdon and the one-term Taylor approximation ABCs are therefore equivalent.

Now consider the second operator ($L = 2$) in the Higdon sequence (6.48). We obtain

$$\left(\cos\alpha_2 \frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right)\left(\cos\alpha_1 \frac{\partial U}{\partial t} - c\frac{\partial U}{\partial x}\right) = 0 \tag{6.51}$$

Carrying out the indicated operations and grouping terms yields

$$\cos\alpha_1 \cos\alpha_2 \frac{\partial^2 U}{\partial t^2} - c(\cos\alpha_1 + \cos\alpha_2)\frac{\partial^2 U}{\partial x\,\partial t} + c^2 \frac{\partial^2 U}{\partial x^2} = 0 \tag{6.52a}$$

At the $x = 0$ grid boundary, the ABC of (6.52a) completely absorbs plane waves propagating with speed c at the angles $\pm\alpha_1$ and $\pm\alpha_2$ with respect to the $-x$-axis. As stated by Higdon, (6.52a) can be directly implemented with finite differences on a straight-line stencil of grid-points perpendicular to the boundary. However, we have the option of using the wave equation $\partial^2 U/\partial t^2 = c^2\,\partial^2 U/\partial x^2 + c^2\,\partial^2 U/\partial y^2$ to eliminate the second x partial derivative, and thereby simplify the x finite-differencing at the grid boundary. This also permits a direct comparison between the partial derivative terms of the second-order Higdon ABC and the Trefethen-Halpern general second-order ABC of (6.39b). After algebraic manipulation, (6.52a) is written as

$$\frac{\partial^2 U}{\partial x\,\partial t} - \frac{1}{c}\left(\frac{1 + \cos\alpha_1 \cos\alpha_2}{\cos\alpha_1 + \cos\alpha_2}\right)\frac{\partial^2 U}{\partial t^2} + \left(\frac{c}{\cos\alpha_1 + \cos\alpha_2}\right)\frac{\partial^2 U}{\partial y^2} = 0 \tag{6.52b}$$

Comparison of (6.52b) with (6.39b) reveals that the two can be made identical with the following correspondences:

$$p_0 = \frac{1 + \cos\alpha_1 \cos\alpha_2}{\cos\alpha_1 + \cos\alpha_2}; \qquad p_2 = -\frac{1}{\cos\alpha_1 + \cos\alpha_2} \tag{6.52c}$$

To illustrate the identity, we substitute into (6.52c) the pair of angles of exact absorption from Table 6.1 corresponding to any one of the seven approximations. We can easily show that (6.52c) then yields the corresponding p_0 and p_2 coefficients listed in the table. This indicates that the Higdon operator of (6.52b) can be configured to be identical to each of the approximate second-order ABCs in Table 6.1, simply by substituting the appropriate angles of exact absorption. In a similar manner, the third operator ($L = 3$) in the Higdon sequence (6.48) can be configured to be identical to each of the third-order ABCs in Table 6.2.

6.4.3 Discussion

In general, the Higdon operator provides for a simple and explicit setting of the exact absorption angles desired by the modeler. Using Trefethen-Halpern operators is more cumbersome, requiring the modeler to go through the intermediate step of setting up an interpolating polynomial.

For a sample computation modeling the outward propagation of a pulsed cylindrical wave sourced in the center of the grid, Higdon reported a reflection of about 7% at the grid outer boundary when implementing his first-order operator with $\alpha = 30°$. For the second-order Higdon operator, the outer-boundary reflection was reduced to the range 2.46% to 3.80%, depending upon the choice of α_1 and α_2, with the minimum reflection obtained when $\alpha_1 = \alpha_2 = 50°$. These results are consistent with the magnitude of wave reflections obtained using Engquist-Majda and Trefethen-Halpern ABCs of the corresponding order.

Despite the greater generality and ease of implementation of Higdon ABCs, there is still the fundamental problem of the assumption of a uniform wave speed c in its formulation. As discussed earlier, this likely limits the suppression of reflections to the order of the numerical phase-velocity variations within the mesh, 0.1% to 1%, regardless of the order of the Higdon ABC implemented. One possible way to mitigate this effect is to incorporate the theoretical numerical phase velocity $\tilde{v}_p$ at each of the exact annihilation angles α_ℓ in the Higdon annihilator of (6.48). That is, $\tilde{v}_p(\alpha_\ell)$ takes the place of the free-space velocity c in the operator:

$$\left\{ \prod_{\ell=1}^{L} \left[\cos\alpha_\ell \frac{\partial}{\partial t} - \tilde{v}_p(\alpha_\ell) \frac{\partial}{\partial x} \right] \right\} U = 0 \tag{6.53}$$

Unlike the Trefethen-Halpern approach, the straightforward formulation of the Higdon operator allows this insight to be readily implemented. However, it is not apparent from the literature that anyone has pursued this improvement.

6.5 LIAO EXTRAPOLATION IN SPACE AND TIME

The fourth principal thrust in ABC technology during the 1970s and 1980s was originated by Liao et al. [18], and based upon what the authors called the "multitransmitting theory." Numerical experiments showed that this ABC exhibits 10 to 20 dB less reflection than the second-order Mur condition (Section 6.3.2), with little sensitivity to the wave-propagation angle or to numerical phase-velocity variations.

To assist in reader understanding, this section provides an alternate derivation of the Liao ABC relative to that presented in [18]. Our derivation involves a straightforward extrapolation of the numerical wave fields in space and time using a Newton backward-difference polynomial.

6.5.1 Formulation

Consider an outer grid boundary located at x_{max}. To close the grid and absorb outgoing waves at this boundary, we need to appropriately calculate the updated field $u(x_{max}, t+\Delta t)$ using field values that are already available in the computer memory. For this purpose, we consider a set of L known fields u_ℓ that are positioned along a straight-line stencil perpendicular to the grid boundary with a uniform spacing $h = \alpha c \Delta t$, where $0 \le \alpha \le 2$. Given the normal propagation delays for outgoing waves approaching this boundary, our strategy also involves using field data that are progressively retarded in time by Δt as the distance along the stencil from x_{max} increases by h. We write:

$$
\begin{array}{lll}
\ell = 0 & u_0 = u(x_{\max},\ t+\Delta t) & \} \text{ unknown field at grid boundary} \\
\ell = 1 & u_1 = u(x_{\max}-h,\ t) & \\
\ell = 2 & u_2 = u(x_{\max}-2h,\ t-\Delta t) & \} \ L \text{ known fields along space-time stencil} \\
\ell = 3 & u_3 = u(x_{\max}-3h,\ t-2\Delta t) & \\
\vdots & \vdots & \\
\ell = L & u_L = u(x_{\max}-Lh,\ t-(L-1)\Delta t) &
\end{array}
\tag{6.54}
$$

Before trying to evaluate u_0, consider how we might estimate u at an arbitrary point $\bar{\ell}$ located *within* the space-time stencil of (6.54). Call this field $u_{\bar{\ell}} \equiv u[x_{\max}-\bar{\ell}h,\ t-(\bar{\ell}-1)\Delta t]$, where $\bar{\ell}$ is a real number in the range $1 \le \bar{\ell} \le L$. A straightforward means to estimate $u_{\bar{\ell}}$ is to set up a backward-difference interpolating polynomial. For this purpose, we define a sequence of backward differences $\{\Delta^m\}$ originating at u_1 (see [19, 20] for example):

$$
\begin{array}{llll}
\Delta^1 u_1 = u_1 - u_2 & \} \ \Delta^2 u_1 = \Delta^1 u_1 - \Delta^1 u_2 & \} \ \Delta^3 u_1 = \Delta^2 u_1 - \Delta^2 u_2 & \} \ \cdots \\
\Delta^1 u_2 = u_2 - u_3 & \} \ \Delta^2 u_2 = \Delta^1 u_2 - \Delta^1 u_3 & \} \ \Delta^3 u_2 = \Delta^2 u_2 - \Delta^2 u_3 & \\
\Delta^1 u_3 = u_3 - u_4 & \} \ \Delta^2 u_3 = \Delta^1 u_3 - \Delta^1 u_4 & & \\
\Delta^1 u_4 = u_4 - u_5 & & & \\
\vdots & \vdots & \vdots &
\end{array}
\tag{6.55}
$$

In general, the mth backward difference at u_1 can be written as

$$
\Delta^m u_1 = \sum_{\ell=1}^{m+1} (-1)^{\ell+1} C_{\ell-1}^m \, u\left[x_{\max} - \ell h,\ \ t-(\ell-1)\Delta t\right] \tag{6.56a}
$$

where C_ℓ^m is the binomial coefficient defined by

$$
C_\ell^m \equiv \frac{m!}{(m-\ell)!\ \ell!} \tag{6.56b}
$$

With the backward differences of (6.55) and (6.56), Newton's backward-difference polynomial can be applied to calculate an interpolated value for $u_{\bar{\ell}}$. Letting $\beta = 1 - \bar{\ell}$, we write

$$
\begin{aligned}
u_{\bar{\ell}} \equiv\ & u_1 + \beta \Delta^1 u_1 + \frac{\beta(\beta+1)}{2!}\Delta^2 u_1 + \frac{\beta(\beta+1)(\beta+2)}{3!}\Delta^3 u_1 \\
& + \cdots + \frac{\beta(\beta+1)(\beta+2)\cdots(\beta+L-2)}{(L-1)!}\Delta^{L-1} u_1
\end{aligned}
\tag{6.57}
$$

Liao's idea can now be interpreted very simply: Use the polynomial derived in (6.57), but instead of interpolating in the range $1 \le \bar{\ell} \le L$, we *extrapolate* to $\bar{\ell} = 0$, the desired outer-boundary field u_0. We immediately have $\beta = 1$, and (6.57) yields

$$u_0 \equiv u(x_{\max}, t+\Delta t) \equiv u_1 + \Delta^1 u_1 + \Delta^2 u_1 + \Delta^3 u_1 + \cdots + \Delta^{L-1} u_1 \tag{6.58}$$

This is simply equation (9) of [18] with the symbol "L" replacing "N." Following Liao's nomenclature, we call this the "generalized L-order transmitting formula." More commonly, (6.58) is termed the Liao ABC of order L implemented at $x_{\max}$.

6.5.2 Discussion

It is well known that using an interpolating polynomial such as (6.57) for extrapolation can yield substantially larger errors than for the original interpolation case. However, Liao's ABC is viable, providing that enough terms in the Newton's difference polynomial are retained.

Accuracy

Consider a unit-amplitude sinusoidal plane wave of wavelength λ incident upon the ABC specified by (6.58). For this case, Liao showed that the maximum amplitude error is given by

$$\left|\Delta^L u\right|_{\max} = 2^L \sin^L(\pi c \Delta t / \lambda) \tag{6.59}$$

For a realistic simulation in three dimensions, we could have $c\Delta t = \Delta/2$ and $\Delta = \lambda/20$. Then, (6.59) yields a maximum error of about 0.4% for $L = 3$. Note that this error is 2 to 10 times less than that of the ABCs considered earlier. In fact, actual FDTD modeling runs have demonstrated this advantage. Unlike the ABCs discussed earlier, the Liao condition is capable of realizing its theoretical level of error. This is because its underlying Newton extrapolation process is not founded on any assumptions about the angle of incidence or velocity of the outgoing numerical waves. It is robust.

Choice of h

The choice of $h = \alpha c \Delta t$ in (6.54) determines the spatial position of the fields along the stencil used in the extrapolation process. Liao reported considerable flexibility in choosing α in the range $0.5 \le \alpha \le 2$, stating that

> "It can be seen that the variance of α influences the accuracy of numerical results; however, all theoretical [waveforms] converge to exact solutions for all the different values of α."

This is additional evidence of the robust nature of the underlying extrapolation process.

However, depending upon the choice of α (and thus h), the needed data may not coincide with field-calculation points in the Yee grid. For example, if $\alpha = 1$ and $c\Delta t = \Delta/2$, we have $h = \Delta/2$. Then, the following field data are required for $L = 3$:

$$\begin{aligned} \ell = 1 \qquad & u_1 = u(x_{\max} - \Delta/2,\ t) \\ \ell = 2 \qquad & u_2 = u(x_{\max} - \Delta,\ t - \Delta t) \\ \ell = 3 \qquad & u_3 = u(x_{\max} - 3\Delta/2,\ t - 2\Delta t) \end{aligned} \tag{6.60}$$

Here, u_2 is available on the Yee grid as an explicitly computed field value. However, u_1 and u_3 are not available because they are located at half-cell points occupied by vector components of the dual field. To obtain such field values, Liao reported a quadratic-interpolation procedure and suggested study of the spline function method.

However, given that the Liao ABC itself results from a Newton interpolating polynomial, the additional interpolation suggested by Liao may provide no additional information. To avoid any interpolation within the Yee grid, it is sufficient to set $\alpha = 2$, the maximum value permitted by Liao. For $c\Delta t = \Delta/2$, this yields $h = \Delta$. Then, the following field data are required for $L = 3$:

$$\begin{aligned} \ell = 1 \qquad & u_1 = u(x_{\max} - \Delta,\ t) \\ \ell = 2 \qquad & u_2 = u(x_{\max} - 2\Delta,\ t - \Delta t) \\ \ell = 3 \qquad & u_3 = u(x_{\max} - 3\Delta,\ t - 2\Delta t) \end{aligned} \tag{6.61}$$

All of the fields specified in (6.61) are explicitly computed in the Yee grid, and Liao's ABC simplifies very nicely.

Potential for Numerical Instability

It was reported in [21] that the use of single-precision computer arithmetic causes FDTD algorithms employing Liao's ABC to become unstable with time-stepping, whereas double-precision arithmetic maintains stability. A z-transform analysis was presented, which shows that Liao's field updating at the grid outer boundary has roots in the z-plane that are all inside the unit circle, with the exception of one root, which lies on the unit circle. Reference [21] states that

> "If there exist some numerical errors, for example, from the use of single-precision arithmetic, this pole could be driven outside the unit circle, rendering the system unstable. To cure this problem, we subtract adiabatic loss terms from [the interpolation coefficients] which are large enough compared to the machine accuracy to keep the pole inside the unit circle, yet small enough to preserve the desired accuracy of the final results."

Reference [21] goes on to suggest that loss terms of order (0.5%) or less be subtracted from each of the interpolation coefficients. This introduces a perturbation that is comparable to the error of the original Liao ABC (assuming $\Delta = \lambda/20$ grid resolution), and results in the overall error increased to about 1% or less. Numerical experiments have shown that the suggested loss terms eliminate instability, and at the same time maintain a solution accuracy superior to that available using the second-order Mur ABC. However, this tactic reduces the accuracy of Liao's ABC, relative to the alternative of using double-precision arithmetic.

6.6 RAMAHI COMPLEMENTARY OPERATORS

In the late 1990s, Ramahi introduced the *complementary operators method* (COM) [22, 23] as an effective means of canceling the residual outer-boundary numerical wave reflections that result from applying standard analytical ABCs. This section summarizes the basis of COM and provides representative FDTD modeling results that indicate its utility.

6.6.1 Basic Idea

The basic premise of the COM is that first-order reflections of numerical waves from the outer boundary of the FDTD space lattice can be cancelled. This cancellation is made possible by averaging two independent solutions to the modeling problem. These two solutions are obtained by imposing radiation boundary operators that are complementary to each other. That is, their reflection coefficients are equal in magnitude, but 180° out of phase.

To place this concept on a more precise footing, we consider the action of an analytical ABC applied at the rightmost outer-boundary plane $x = a$ of a Cartesian FDTD space lattice. Impinging upon this boundary is the $+x$-directed sinusoidal numerical plane wave given by:[2]

$$U_{\text{inc}} = e^{j(\omega t - \tilde{k}_x x - \tilde{k}_y y - \tilde{k}_z z)} \tag{6.62}$$

The total field at $x = a$ must satisfy the specific ABC in effect there. Upon postulating the existence of a wave reflected from plane $x = a$, the total field in the region $x \le a$ has the form

$$U = e^{j(\omega t - \tilde{k}_x x - \tilde{k}_y y - \tilde{k}_z z)} + R_{\text{ABC}}\, e^{j(\omega t + \tilde{k}_x x - \tilde{k}_y y - \tilde{k}_z z)} \tag{6.63}$$

If the ABC is ideal (completely reflectionless), then $R_{\text{ABC}} = 0$, and the second term in (6.63) equals zero. This has been the goal in formulating all of the analytical ABCs reviewed in this chapter. However, Ramahi reported a strategy to compensate for the reality that the ABC is not ideal (and thus $R_{\text{ABC}} \neq 0$). To this end, he defined $\overline{\text{ABC}}$, a *complementary* boundary condition having the property

$$R_{\overline{\text{ABC}}} = -R_{\text{ABC}} \tag{6.64}$$

In the region $x \le a$, this new condition generates the field

$$\overline{U} = e^{j(\omega t - \tilde{k}_x x - \tilde{k}_y y - \tilde{k}_z z)} - R_{\text{ABC}}\, e^{j(\omega t + \tilde{k}_x x - \tilde{k}_y y - \tilde{k}_z z)} \tag{6.65}$$

Averaging the fields in (6.63) and (6.65) yields the desired result, a solution that has *zero reflection* from the outer boundary of the space lattice at $x = a$:

$$U_{\text{exact}} = (U + \overline{U})/2 \tag{6.66}$$

---

[2]This analytical form is similar to (6.41). However, note that (6.41) described $-x$-directed wave propagation toward the ABC acting at the $x = 0$ lattice boundary.

The primary strength of COM is that the cancellation of first-order outer-boundary reflections indicated by (6.66) takes place for any field independent of its numerical wavenumber $\tilde{k}$. This implies that reflections at the outer boundary of the space lattice are suppressed, regardless of whether the fields are composed of evanescent or purely traveling waves.

6.6.2 Complementary Operators

To implement the COM, Ramahi derived two independent, complementary radiation boundary operators from Higdon's ABC, discussed in Section 6.4:

$$B_L U = \frac{\partial}{\partial x} \prod_{\ell=1}^{L} \left(\frac{\partial}{\partial x} + \frac{\cos\phi}{c} \frac{\partial}{\partial t} + \alpha_\ell \right) U = 0 \tag{6.67a}$$

$$\overline{B_L} U = \frac{\partial}{\partial t} \prod_{\ell=1}^{L} \left(\frac{\partial}{\partial x} + \frac{\cos\phi}{c} \frac{\partial}{\partial t} + \alpha_\ell \right) U = 0 \tag{6.67b}$$

Here, the parameter α_ℓ ensures the stability of the numerical simulation. This parameter has been shown empirically to depend on the space-cell size Δ and the problem dimensionality, and is typically $\alpha_\ell = 0.02/\Delta$ and $\alpha_\ell = 0.01/\Delta$ for two- and three-dimensional computation spaces, respectively. The angle ϕ can be optimized to minimize reflection in a specific propagation direction. In a general application, however, where no advanced knowledge of the propagation is assumed, ϕ is set to zero.

For a sinusoidal plane wave, the reflection coefficients for the radiation operators of (6.67a) and (6.67b) are given respectively by

$$R_{B_L} = \prod_{\ell=1}^{L} \left(\frac{-j\tilde{k}_x + j\tilde{k}\cos\phi + \alpha_\ell}{j\tilde{k}_x + j\tilde{k}\cos\phi + \alpha_\ell} \right) ; \qquad R_{\overline{B_L}} = -R_{B_L} \tag{6.68a, b}$$

Given the equal but opposite nature of the outer-boundary reflection coefficients in (6.68a) and (6.68b), it suffices to implement separate FDTD modeling runs of the electromagnetic wave problem of interest, using the corresponding ABCs B_L and $\overline{B_L}$ in (6.67a) and (6.67b). Then, as a postprocessing step, the field values obtained in these two runs are averaged according to (6.66) at each space-time point of interest. This yields a composite set of field data devoid of first-order reflections from the outer boundary of the space lattice. Because this procedure requires two independent FDTD modeling runs, the total floating-point operation count is exactly doubled in comparison with the traditional implementation of the ABC.

6.6.3 Effect of Multiple Wave Reflections

In practical FDTD models, the ideal scenario for the COM technique outlined above is complicated by the fact that multiple reflections of the outgoing numerical wave can occur. As a result, complete cancellation of reflections from the outer lattice boundary may not occur when averaging the two solutions as in (6.66).

Fig. 6.6(a) illustrates the nature of this problem for a two-dimensional grid, showing three categories of uncancelled multiple reflections. Here, we assume that the original outgoing wave has unity magnitude. We denote the initial reflections at the outer grid boundary due to B_L and $\overline{B_L}$ as $[R]$ and $(-R)$, respectively; and a retroreflection at the structure being modeled as a. A double bounce between two B_L ABCs yields the reflection $[R^2] = [R]\cdot[R]$, which is equal to the double-bounce reflection between two $\overline{B_L}$ ABCs, $(R^2) = (-R)\cdot(-R)$. At any point in Fig. 6.6(a), the residual field after implementing the COM averaging process of (6.66) is obtained by adding $[R] + (-R) = 0$, yielding the desired cancellation of the initial reflections, or by adding $[R^2] + (R^2) = 2R^2$, yielding undesired reinforcement. The categories of uncancelled multiple reflections in Fig. 6.6(a) are as follows:

1. *Double-bounce reflections at the corners of the grid.* While such reflections are of order (R^2), that is, second order, they can be a significant source of error. This is because the outgoing waves impinge upon the grid corners at highly oblique angles where the ABC generally has poor performance. Numerical experiments with the COM, using B_3 and $\overline{B_3}$ defined in (6.67a) and (6.67b), have shown that a wave incident upon a grid corner at 70° comes back into the computational domain with approximately a 1% reflection after implementation of (6.66).

2. *Double-bounce reflections between two parallel grid outer-boundary planes.* While also of order (R^2), these are less significant than those of Category 1 because they occur near normal incidence at the grid outer boundary, where the ABC is most effective.

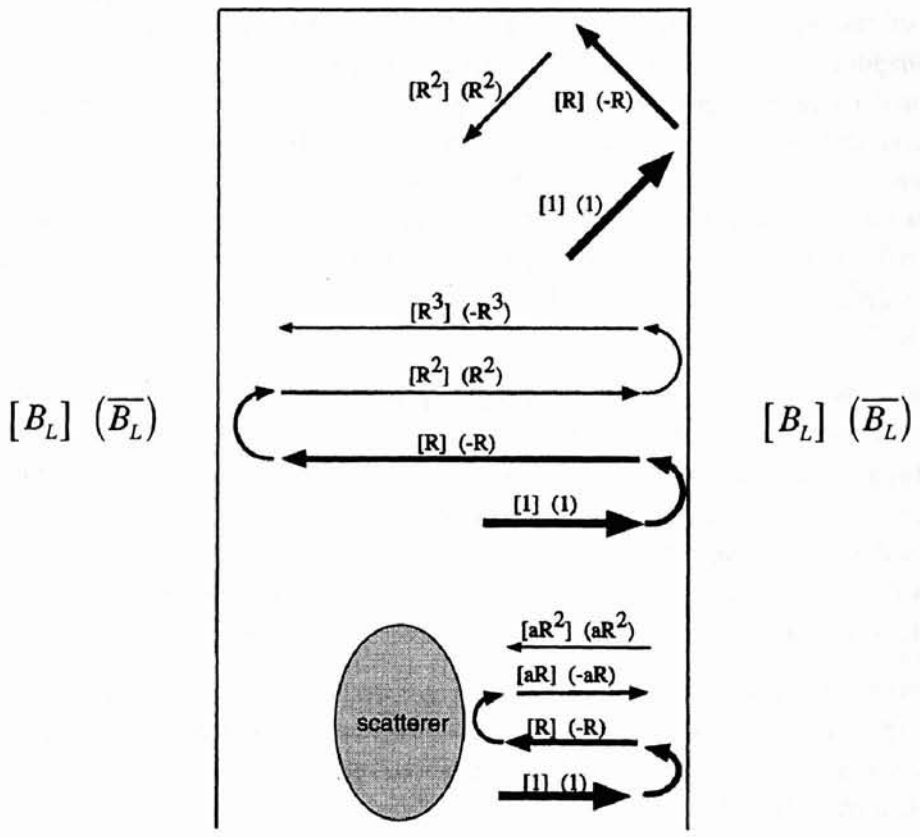
3. *Triple-bounce reflections between a grid-outer boundary plane and the structure being modeled.* Similar physical mechanisms apply here as in Category 2, resulting in weaker reflections than those of Category 1.

Ramahi reported a variation of COM to cancel the double-bounce corner reflections in a two-dimensional grid. Here, four independent FDTD runs are required. For each run, one imposes a unique combination of B_L and $\overline{B_L}$ over the four planes of the grid outer boundary. This is shown in Fig. 6.6(b), where for conciseness, we use "+" and "–" symbols to denote the presence of the B_L or $\overline{B_L}$ ABC, respectively, at each outer-boundary plane. By analogy to (6.66), postprocessing of the fields is implemented by taking the arithmetic average of the fields obtained in the four runs at each space-time point of interest. Using this procedure, the net reflection from a grid corner is easily seen to be $0.25 \times \{ [R]\cdot[R] + (-R)\cdot(-R) + [R]\cdot(-R) + (-R)\cdot[R] \} = 0$.

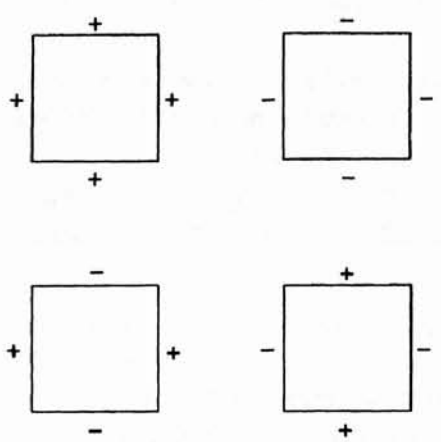
6.6.4 Basis of the Concurrent Complementary Operator Method

The computational penalty involved in implementing the COM in two dimensions is a doubling or quadrupling (depending upon whether corner reflections are cancelled) of the total number of floating-point operations relative to a conventional FDTD model of the same grid size. This discourages the use of the COM for large-scale simulations.

Recognizing this difficulty, Ramahi formulated an improved version of the original COM, which he called the *concurrent COM* (C-COM). This technique, suitable for use in either two- or three-dimensional FDTD space lattices, is intended to achieve two objectives:



(a) Multiple reflections due to the outer lattice boundary and the scatterer.



(b) Permutations of boundary operators needed to cancel corner reflections in a two-dimensional grid.

Fig. 6.6 Effect of multiple wave reflections in the complementary operators method. *Source:* Ramahi, *IEEE Trans. Antennas and Propagation*, 1998, pp. 1475–1482, © 1998 IEEE.

- Implement the complementary radiation operators within a single FDTD modeling run to cancel the reflections from the outer planes of the space lattice;
- At the user's option during the same FDTD run, cancel the reflections from the corners formed by the outer planes of the space lattice.

The ability of C-COM to cancel outer-boundary reflections with only a single FDTD run rather than multiple runs greatly reduces the required computer time to implement Ramahi's original concept, thereby enlarging the scope of its potential applications.

Cancellation of Reflections from the Outer Side Planes of the Space Lattice

We implement C-COM by dividing the FDTD space lattice into an interior region and a surrounding boundary layer. Components of $\boldsymbol{E}$ are assumed to lie tangentially on the interface between these regions, and on the lattice outer-boundary planes. The interior region includes the structure being modeled as well as any sources. Here, each of the components of $\boldsymbol{E}$ and $\boldsymbol{H}$ is assigned a single computer storage location in the usual manner. However, in the boundary layer, two storage locations are assigned to each component of $\boldsymbol{E}$ and $\boldsymbol{H}$, which we designate as $\boldsymbol{E}^{(1)}$, $\boldsymbol{E}^{(2)}$, $\boldsymbol{H}^{(1)}$, and $\boldsymbol{H}^{(2)}$. This double allocation of field storage in the boundary layer permits the *concurrent and independent* implementation of the complementary radiation operators in this region.

A good way to understand C-COM is to consider it an inductive procedure founded upon the following assumptions concerning the field values stored in computer memory at the beginning of a time-step:

- $\boldsymbol{E}$ and $\boldsymbol{H}$ within the interior region are devoid of any error due to first-order reflections from the lattice outer-boundary planes.
- $\boldsymbol{E}^{(1)}$ and $\boldsymbol{H}^{(1)}$ within the boundary layer are corrupted by reflection error due to the imperfect action of the ABC operator B_L.
- $\boldsymbol{E}^{(2)}$ and $\boldsymbol{H}^{(2)}$ within the boundary layer are corrupted by reflection error due to the imperfect action of the ABC operator $\overline{B_L}$.

Given these assumptions, time-stepping proceeds as follows:

<u>$\boldsymbol{E}$ Updates</u>

1. Update $\boldsymbol{E}$ within the interior region using the standard Yee algorithm applied to the stored $\boldsymbol{H}$ values.
2. With the exception of points along the outer boundary of the lattice, update $\boldsymbol{E}^{(1)}$ in the boundary layer using the standard Yee algorithm applied to the stored $\boldsymbol{H}^{(1)}$ values. Similarly, update $\boldsymbol{E}^{(2)}$ in the boundary layer using the stored $\boldsymbol{H}^{(2)}$ values.
3. Using B_L of (6.67a), update each tangential component of $\boldsymbol{E}^{(1)}$ at its location P on the lattice outer boundary. The required data are the L nearest parallel components of $\boldsymbol{E}^{(1)}$ located on a straight-line stencil perpendicular to the lattice boundary at P. Similarly, update each tangential component of $\boldsymbol{E}^{(2)}$ at P by applying $\overline{B_L}$ of (6.67b) to the L nearest parallel components of $\boldsymbol{E}^{(2)}$ located on the same stencil.

4. Set each tangential component of $\boldsymbol{E}$ at its location P' on the interior-region / boundary-layer interface to the average of the values of the tangential components of $\boldsymbol{E}^{(1)}$ and $\boldsymbol{E}^{(2)}$ at P'. Set each tangential component of $\boldsymbol{E}^{(1)}$ and $\boldsymbol{E}^{(2)}$ at P' to this same average value. This connects the solutions in the interior region and the boundary layer.

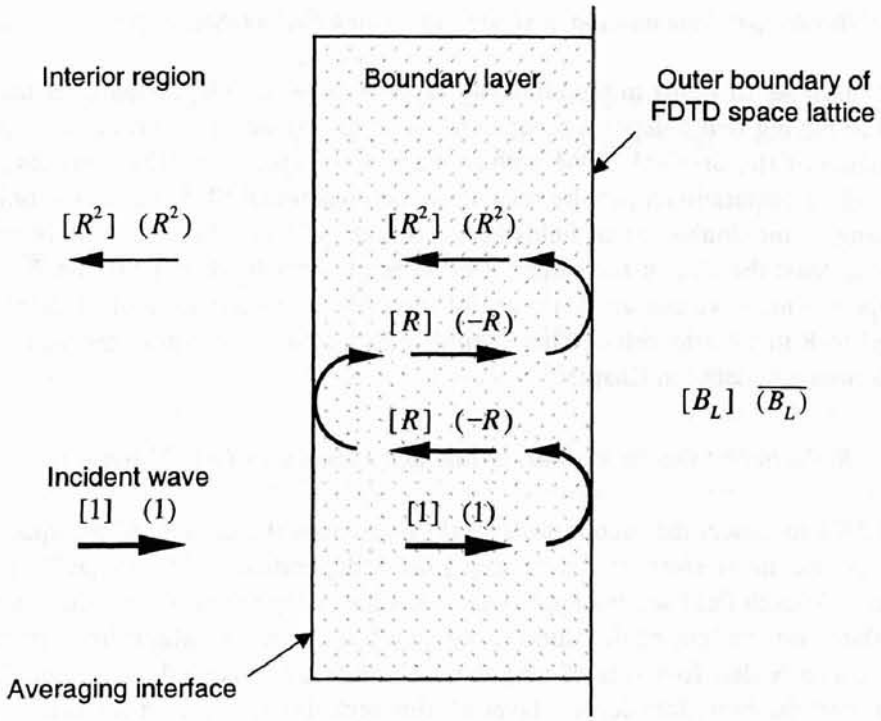
<u>$\boldsymbol{H}$ Updates</u>

1. Advance time by one-half time-step.
2. Update $\boldsymbol{H}$ within the interior region using the standard Yee algorithm applied to the $\boldsymbol{E}$ values just calculated in the previous half time-step.
3. Update $\boldsymbol{H}^{(1)}$ within the boundary layer using the standard Yee algorithm applied to the $\boldsymbol{E}^{(1)}$ values just calculated in the previous half time-step. Similarly, update $\boldsymbol{H}^{(2)}$ in the boundary layer using the just-calculated $\boldsymbol{E}^{(2)}$ values.

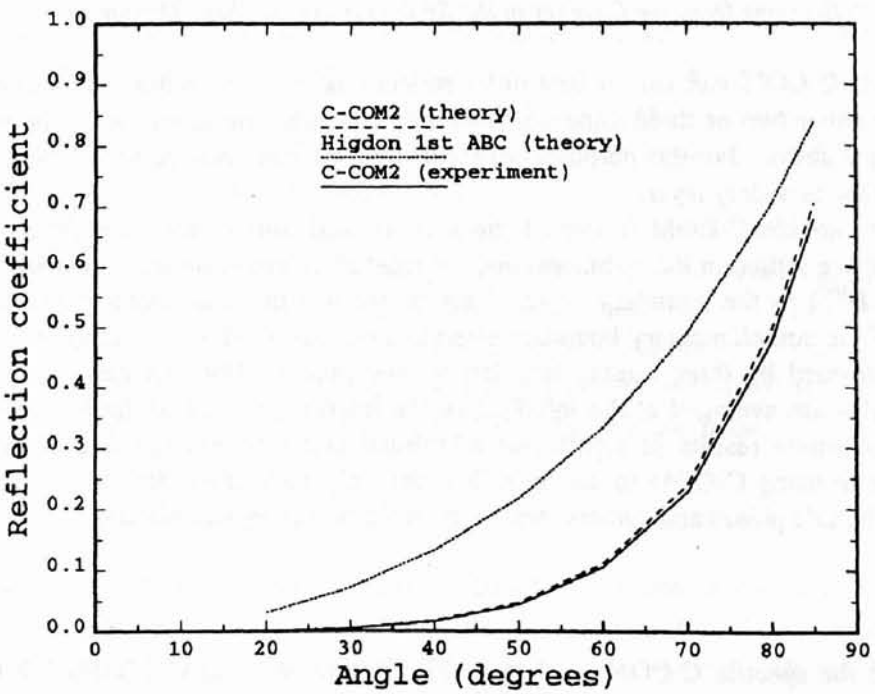
Fig. 6.7(a) shows that the net effect of the seven-step procedure summarized above is to create a virtual (nonphysical) discontinuity at the interface of the interior region and the boundary layer. The incident wave, assumed to have a unity amplitude, passes through the interface with no reflection. This excites two independent waves $\{\boldsymbol{E}^{(1)}, \boldsymbol{H}^{(1)}\}$ and $\{\boldsymbol{E}^{(2)}, \boldsymbol{H}^{(2)}\}$ having amplitudes [1] and (1), respectively, within the boundary layer. The initial (first-order) reflection of these waves from the lattice outer boundary yields two waves propagating back toward the interface with amplitudes $[R]$ and $(-R)$, due to the action of B_L and $\overline{B_L}$, respectively. However, Step 4 of the time-stepping algorithm sets the tangential E-field at the interface to be the average of the wave fields in the boundary layer. To satisfy this constraint and to satisfy the required continuity of the tangential E-field across the interface, complete retroreflection of the first-order reflection errors must take place at the interface. These retroreflected error signals propagate back toward the outer space lattice boundary with amplitudes $[R] = [R]\cdot 1$ and $(-R) = (-R)\cdot 1$, respectively.

Subsequently, these retroreflected error signals reach the outer space lattice boundary. There, they are once again reflected back toward the interface with the amplitudes $[R^2] = [R]\cdot 1\cdot[R]$ and $(R^2) = (-R)\cdot 1\cdot(-R)$, due to the action of B_L and $\overline{B_L}$, respectively. It is clear that these second-order reflections from the outer boundary are in-phase and have equal amplitudes. Hence, when these reflections are averaged at the interface, no cancellation takes place, and they penetrate into the interior region carrying the reflection error R^2. Thus, the overall effect of implementing C-COM is to yield a composite reflection coefficient equal to the square of the reflection coefficient generated by the ABC operators used at the sides of the outer lattice boundary.

Fig. 6.7(b) is a validation study relating to the above discussion. This figure shows the results of FDTD numerical experiments where the theoretical plane-wave reflection coefficient for Higdon's first-order ABC is compared with the theoretical and measured reflection coefficients for the C-COM using this same ABC and its complement. (Ramahi designates this implementation as C-COM2.) We see that the measured reflectivity of C-COM2 is very close to the square of the reflectivity of the Higdon ABC over incident wave angles from 0° to 85° relative to the lattice outer boundary.



(a) Illustration showing the cancellation mechanism of the C-COM.



(b) Comparison of reflection coefficients for Higdon's first-order ABC and C-COM2.

Fig. 6.7 Basis of the concurrent complementary operator method. *Source:* Ramahi, *IEEE Trans. Antennas and Propagation*, 1998, pp. 1475–1482, © 1998 IEEE.

Computational Burden for Cancellation of Outer Side-Plane Reflections

With C-COM, each set of fields in the boundary layer is updated independently of the other set. This amounts to having two independent simulations in the boundary layer, thereby meeting the data requirements of the original COM without having to repeat the FDTD processing of the entire lattice. The computational penalty relative to a conventional FDTD model is only that due to the processing of the double set of fields in the boundary layer. The width W of this layer is required to be at least the size of the stencil used to implement B_L of (6.67a) and $\overline{B_L}$ of (6.67b). Numerical experiments have shown that a useful increase in the accuracy of C-COM occurs as W is increased to 8 to 12 grid cells. This is similar to the thickness requirement for the PML ABC, to be discussed in detail in Chapter 7.

Cancellation of Reflections from the Corners of the Space Lattice in Two Dimensions

To enable C-COM to cancel the second-order reflections from the corners of the space lattice in two dimensions, we must store and time-step *four* independent field sets $[\boldsymbol{E}^{(m)}, \boldsymbol{H}^{(m)}]$ in the boundary layer. To each field set, we apply one of the boundary-operator permutations shown in Fig. 6.6(b). Then, we implement the same seven-step time-marching algorithm outlined above. The only difference is that four sets of tangential E fields are averaged at the interface of the interior region and the boundary-layer. Overall, this technique results in additional computer storage and running-time burdens relative to using C-COM to cancel reflections only from the outer side planes of the lattice.

Cancellation of Reflections from the Corners of the Space Lattice in Three Dimensions

As stated earlier, C-COM can cancel first-order reflections from the outer side planes of the space lattice in either two or three dimensions by implementing the seven-step time-marching algorithm outlined above. For this purpose, only two storage locations need be reserved for each E and H field in the boundary layer.

However, to enable C-COM to cancel the second- and third-order reflections from the corners of the space lattice in three dimensions, we must store and time-step *eight* independent field sets $[\boldsymbol{E}^{(m)}, \boldsymbol{H}^{(m)}]$ in the boundary layer. This is because there are eight possible unique permutations of the complementary boundary operators occurring when considering reflections from a corner formed by three intersecting lattice side planes. For this case, eight sets of tangential E fields are averaged at the interface of the interior region and the boundary layer. Overall, this technique results in significant additional computer-storage and running-time burdens relative to using C-COM to cancel reflections only from the outer side planes of the lattice, or from the side planes and corners formed by two intersecting side planes.

Notation

Ramahi denoted the specific C-COM used in an FDTD simulation as C-COM$L(S, W)$. Here, parameter L specifies the order of accuracy of the radiation operator. Parameter $S = 2, 4, 8$ specifies the need to store either two, four, or eight independent field sets in the boundary layer to cancel reflections, respectively, from the outer side planes of the space lattice, from corners formed by two intersecting side planes, or from corners formed by three intersecting side planes. Parameter W specifies the width of the boundary layer in lattice cells.

6.6.5 Illustrative FDTD Modeling Results Obtained Using the C-COM

This section reviews two results of FDTD numerical experiments published by Ramahi [23] that illustrate both the level of effectiveness of C-COM and the impact of varying its governing parameters.

Radiation from a Line Current Source in Two Dimensions

This numerical experiment evaluates how C-COM performs in two dimensions for a radiating source located close to the outer boundary of the FDTD grid. In this example, the grid spans only 21×21 square cells, and the boundary layer needed to implement C-COM is added to this space. A line current source is located at point (11, 11) and is excited with a time waveform similar to the compact pulse specified in (6.46). A normalized error at the observation point (16, 16) is defined as

$$Error(t) = \frac{\left| y(t) - y_{\text{reference}}(t) \right|}{\max\left[\left| y_{\text{reference}}(t) \right| \right]} \tag{6.69}$$

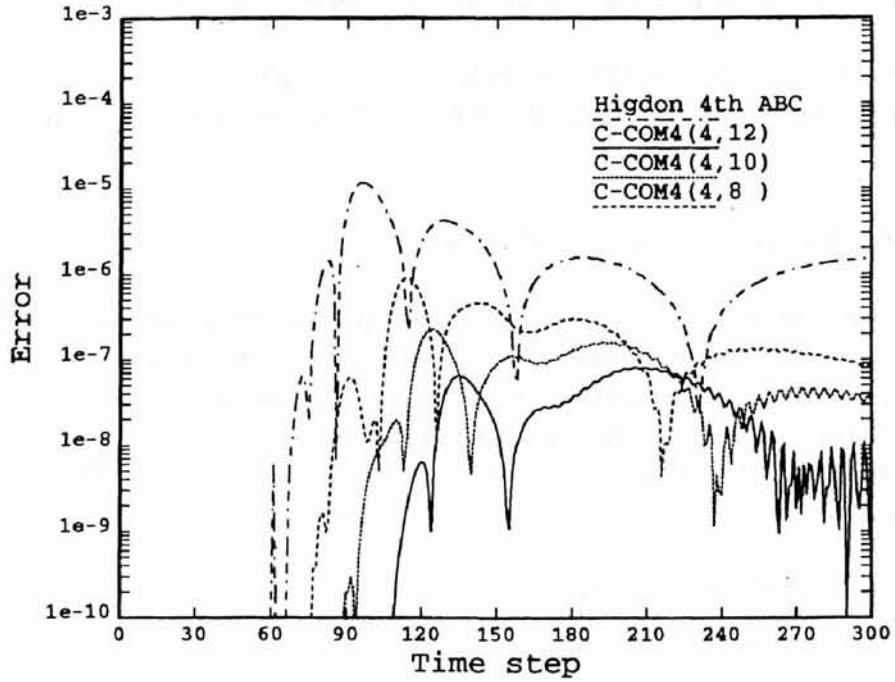
where $y(t)$ is the FDTD-calculated field at point (16, 16) corresponding to the usage of either Higdon's ABC or C-COM at the outer boundary of the 21×21-cell test grid, and $y_{\text{reference}}(t)$ is the FDTD-calculated field obtained in a much larger reference grid that has zero outer-boundary reflections during the time-stepping span considered.

Fig. 6.8(a) shows the effect of varying the width W of the C-COM4(4, W) boundary layer from 8 to 10 to 12 cells. We see that each application of C-COM considered here provides less error than Higdon's fourth-order ABC. Further, the C-COM error decreases as W increases. For a 12-cell boundary-layer width, the error in the C-COM calculation is generally one to two orders of magnitude below that of Higdon's ABC.

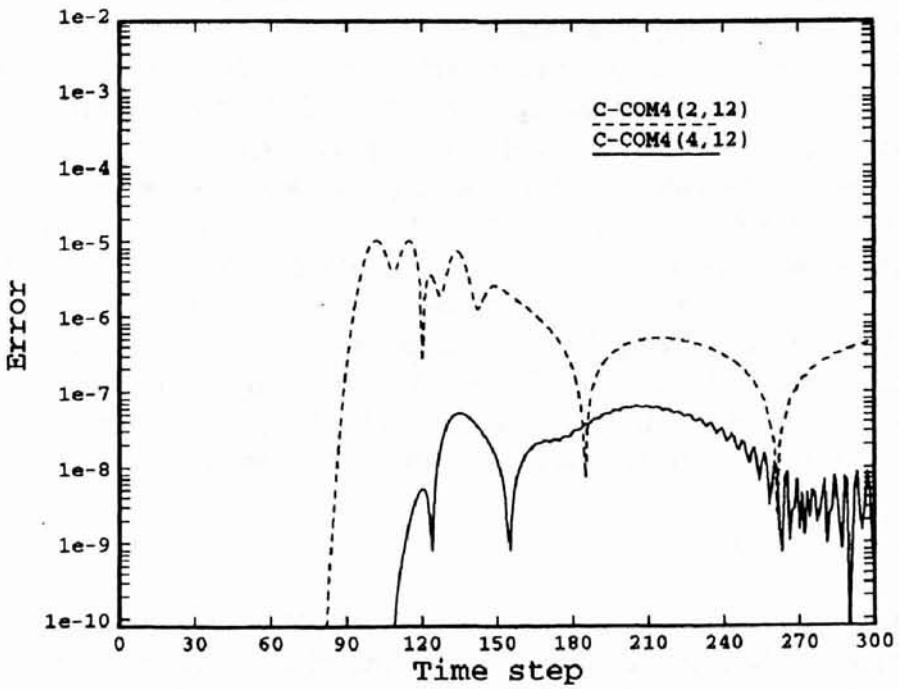
Fig. 6.8(b) compares the error of the fourth-order, 12-cell-wide C-COM4(4, 12), which cancels both side-boundary and corner reflections, with the fourth-order, 12-cell-wide C-COM4(2, 12), which cancels only the side-boundary reflections. We see that suppressing corner reflections reduces the error by approximately one order of magnitude. This is significant, and in accordance with our expectation that relatively large reflection coefficients arise from waves impinging on the corner regions at highly oblique angles.

Radiation from a Hertzian Dipole in Three Dimensions

This numerical experiment evaluates the performance of C-COM in three dimensions relative to Higdon's fourth-order ABC. It also demonstrates the importance of canceling both the side-boundary and corner reflections of the space lattice. In this example, the FDTD lattice spans only $21 \times 21 \times 21$ cubic cells, inclusive of the C-COM boundary layer. A z-polarized Hertzian dipole is located at point (11, 11, 11) and is excited with a Gaussian pulse modulating a sinusoidal signal. The normalized error in the E_z field, as defined by (6.69), is measured at the observation point (11, 11, 19).



(a) Reduction of the error of fourth-order C-COM as the boundary layer widens.



(b) Reduction of the error of fourth-order C-COM by canceling corner reflections.

Fig. 6.8 Measurement of the outer-boundary reflection error due to C-COM for a two-dimensional FDTD model of a radiating line current source. *Source:* Ramahi, *IEEE Trans. Antennas and Propagation*, 1998, pp. 1475–1482, © 1998 IEEE.

Fig. 6.9 compares the error of the fourth-order, 8-cell-wide C-COM4(8, 8), which cancels both side-boundary and corner reflections, with the fourth-order, 8-cell-wide C-COM4(2, 8), which cancels only the side-boundary reflections, and with Higdon's fourth-order ABC. We see that each application of C-COM considered here provides much less error than Higdon's ABC. Namely, the C-COM4(2, 8) error is almost two orders of magnitude below Higdon, while the C-COM4(8, 8) error is three orders of magnitude below Higdon. The primary caveat is that implementing the latter C-COM technique involves a substantial burden on computer memory and running time, and thus is recommended only for FDTD simulations requiring a wide dynamic range of computational accuracy.

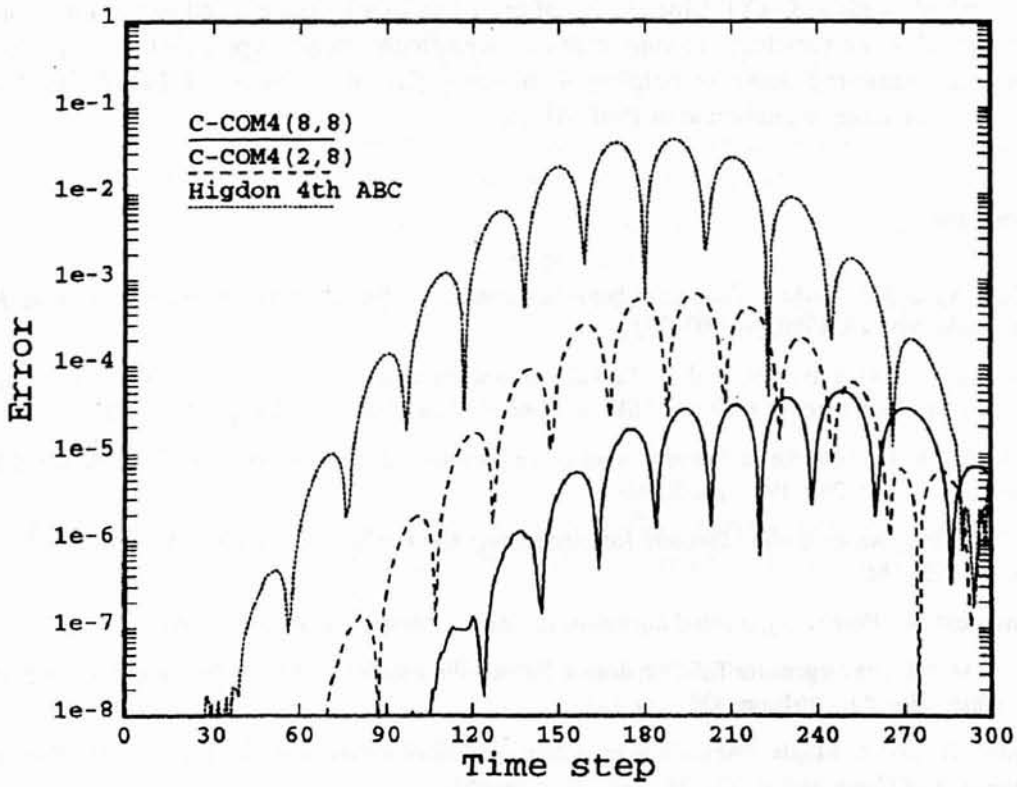


Fig. 6.9 Measurement of the outer-boundary reflection error due to C-COM for a three-dimensional FDTD model of a radiating Hertzian dipole. *Source:* Ramahi, *IEEE Trans. Antennas and Propagation*, 1998, pp. 1475–1482, © 1998 IEEE.

What These Results Mean

Ramahi has convincingly demonstrated that his systematic C-COM technique can achieve an outer-boundary reflectivity as little as 1/1,000th that of the best previous analytical ABCs, such as Higdon's fourth-order operator. This translates to net reflection coefficients as low as 10^{-5} for three-dimensional open-region FDTD models, or equivalently, –100 dB. These are extremely low levels that permit, in principle, the accurate simulation of very weakly scattering objects of interest in defense, geophysical, optical, and biomedical applications.

6.7 SUMMARY

Three decades of work in formulating analytical ABCs for open-region electromagnetic wave interaction problems has culminated in systematic means, such as Ramahi's C-COM technique, to reduce reflection coefficients in open-region FDTD space lattices to levels in the order of –100 dB. As a result, current FDTD computational dynamic range for such problems is limited more by numerical dispersion and by the precision in defining the shapes of the structures being modeled, than by artifacts due to reflection of numerical waves from the outer boundary of the computation space.

The next chapter, Chapter 7, will provide a completely different treatment of ABCs for FDTD simulations involving the use of PML absorbers. PML ABCs may have more utility than analytical methods such as C-COM for classes of problems involving the need to terminate space lattices in guided-wave structures having multiple, significant corner-type reflections and other reverberations. Ongoing work is helping to better define the classes of FDTD modeling problems most amenable to analytical or PML ABCs.

REFERENCES

[1] Bayliss, A., and E. Turkel, "Radiation boundary conditions for wave-like equations," *Comm. Pure Appl. Math.*, Vol. 23, 1980, pp. 707–725.

[2] Bayliss, A., M. Gunzburger, and E. Turkel, "Boundary conditions for the numerical solution of elliptic equations in exterior regions," *SIAM J. Applied Math.*, Vol. 42, 1982, pp. 430–451.

[3] Friedlander, F. G., "On the radiation field of pulse solutions of the wave equation," *Proc. Royal Soc. London Ser. A*, Vol. 269, 1962, pp. 53–69.

[4] Wilcox, C. H., "An expansion theorem for electromagnetic fields," *Comm. Pure Appl. Math.*, Vol. 9, 1956, pp. 115–134.

[5] Sommerfeld, A., *Partial Differential Equations in Physics*, New York: Academic Press, 1949.

[6] Karp, S. N., "A convergent far-field expansion for two-dimensional radiation functions," *Comm. Pure Appl. Math.*, Vol. 14, 1961, pp. 427–434.

[7] Engquist, B., and A. Majda, "Absorbing boundary conditions for the numerical simulation of waves," *Mathematics of Computation*, Vol. 31, 1977, pp. 629–651.

[8] Mur, G., "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations," *IEEE Trans. Electromagn. Compat.*, Vol. 23, 1981, pp. 377–382.

[9] Halpern, L., and L. N. Trefethen, "Wide-angle one-way wave equations," Numerical Analysis Report 86-5, Dept. of Mathematics, Massachusetts Institute of Technology, 1986.

[10] Trefethen, L. N., and L. Halpern, "Well-posedness of one-way wave equations and absorbing boundary conditions," *Mathematics of Computation*, Vol. 47, 1986, pp. 421–435.

[11] Moore, T. G., J. G. Blaschak, A. Taflove, and G. A. Kriegsmann, "Theory and application of radiation boundary operators," *IEEE Trans. Antennas Propagat.*, Vol. 36, 1988, pp. 1797–1812.

[12] Kriegsmann, G. A., A. N. Norris, and E. L. Reiss, "Acoustic pulse scattering by baffled membranes," *J. Acoustic Society America*, Vol. 79, 1986.

[13] Trefethen, L. N., "Group velocity in finite-difference schemes," *SIAM Review*, No. 24, 1982, pp. 114–136.

[14] Trefethen, L. N., "Group velocity interpretation of stability theory of Gustafson, Kreiss, and Sundstrom," *J. Computational Physics*, No. 49, 1983, pp. 199–217.

[15] Higdon, R. L., "Absorbing boundary conditions for difference approximations to the multi-dimensional wave equation," *Mathematics of Computation*, Vol. 47, 1986, pp. 437–459.

[16] Higdon, R. L., "Numerical absorbing boundary conditions for the wave equation," *Mathematics of Computation*, Vol. 49, 1987, pp. 65–90.

[17] Givoli, D., "Non-reflecting boundary conditions," *J. Computational Physics*, Vol. 94, 1991, pp. 1–29.

[18] Liao, Z. P., H. L. Wong, B. P. Yang, and Y. F. Yuan, "A transmitting boundary for transient wave analyses," *Scientia Sinica* (series A), Vol. XXVII, 1984, pp. 1063–1076.

[19] Carnahan, B., H. A. Luther, and J. O. Wilkes, *Applied Numerical Methods*, New York: Wiley, 1969, Chapter 1.

[20] Al-Khafaji, A. W., and J. R. Tooley, *Numerical Methods in Engineering Practice*, New York: Holt, Rinehart, and Winston, 1986, Chap. 8, Sect. 8.3.2.

[21] Moghaddam, M., *Modeling of the Subsurface Interface Radar*, Ph.D. Dissertation, Dept. of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 1990.

[22] Ramahi, O. M., "The complementary operators method in FDTD simulations," *IEEE Antennas Propagat. Magazine*, Vol. 39, No. 6, Dec. 1997, pp. 33–45.

[23] Ramahi, O. M., "The concurrent complementary operators method for FDTD mesh truncation," *IEEE Trans. Antennas Propagat.*, Vol. 46, 1998, pp. 1475–1482.

PROBLEMS

6.1 Apply the third-order Bayliss-Turkel operator to the outgoing wave expansion of (6.2) in spherical coordinates and determine the first three terms in the remainder. Show that the order of the leading term satisfies (6.11).

6.2 Apply the third-order Bayliss-Turkel operator to the outgoing wave expansion of (6.12) in cylindrical coordinates and determine the first three terms in the remainder. Show that the order of the leading term satisfies (6.19).

6.3 Derive the partial differential equation corresponding to the three-term Taylor series approximation of the Engquist-Majda one-way wave equation of (6.24) at the $x = 0$ grid boundary.

6.4 Using appropriate finite differences, implement the first-order Engquist-Majda ABC of (6.26) in an FDTD computer program for the two-dimensional TM_z grid. Obtain field values at the corners of the grid outer boundary by finite-differencing a first-order Engquist-Majda ABC along the diagonal line running from the center of the grid to the corner point. Then, follow the method illustrated by Fig. 6.3 to conduct numerical experiments to obtain the local and global errors of your ABC system.

6.5 Repeat Problem 6.4, but for a two-dimensional TE_z grid.

6.6 Repeat Problem 6.4, but now implement the second-order Mur ABC of (6.34). At the corners of the grid outer boundary, retain the first-order Engquist-Majda ABC finite-differenced along the diagonal line running from the center of the grid to the corner point.

6.7 Repeat Problem 6.4, but now implement the second-order Higdon ABC of (6.48). Select the incidence angles α_ℓ to optimize the local and global error. Use a straight-line stencil to implement the needed spatial finite differences.

6.8 Repeat Problem 6.4, but now implement the third-order ($L = 3$) Liao ABC of (6.58) at all boundary points, including the corners, using the space-time stencil of (6.61). Use double-precision computer arithmetic to avoid numerical instability problems.

6.9 Repeat Problem 6.4, but now implement Ramahi's COM ABC using (6.67a) and (6.67b) evaluated for $L = 1$.

Chapter 7

Perfectly Matched Layer Absorbing Boundary Conditions

Stephen Gedney

7.1 INTRODUCTION

One of the greatest challenges of the FDTD method has been the efficient and accurate solution of electromagnetic wave interaction problems in unbounded regions. For such problems, an *absorbing boundary condition* (ABC) must be introduced at the outer lattice boundary to simulate the extension of the lattice to infinity. As reviewed in Chapter 6, a number of analytical techniques have been introduced to achieve this goal.

An alternate approach to realize an ABC is to terminate the outer boundary of the space lattice in an absorbing material medium. This is analogous to the physical treatment of the walls of an anechoic chamber. Ideally, the absorbing medium is only a few lattice cells thick, reflectionless to all impinging waves over their full frequency spectrum, highly absorbing, and effective in the near field of a source or a scatterer. An early attempt at implementing such an absorbing material boundary condition was reported by Holland [1], who utilized a conventional lossy dispersionless absorbing medium. The difficulty with this tactic is that such an absorbing layer is matched only to normally incident plane waves. As a result, the entire category of lossy-material ABCs had only limited application in electromagnetics.

Since 1994, a new fervor in this area has been created by Berenger's introduction of a highly effective absorbing-material ABC designated the *perfectly matched layer* (PML) [2]. The innovation of Berenger's PML is that plane waves of arbitrary incidence, polarization, and frequency are matched at the boundary. Perhaps of equal importance is that the PML can be used as an absorbing boundary to terminate domains comprised of inhomogeneous, dispersive, anisotropic, and even nonlinear media, which was previously not possible with analytically derived ABCs.

In his pioneering work, Berenger derived a novel split-field formulation of Maxwell's equations where each vector field component is split into two orthogonal components [2]. Maxwell's curl equations were also appropriately split, leading to a set of 12 coupled first-order partial differential equations. Then, by choosing loss parameters consistent with a dispersionless medium, a perfectly matched planar interface is derived.

In a continuous space, the PML absorber and the host medium are perfectly matched. However, in the discrete FDTD lattice, the electric and magnetic material parameters are represented in a piecewise-constant manner and are spatially staggered. This results in discretization errors that can degrade the ideal behavior of the PML. Berenger proposed a spatial

scaling of the PML parameters to reduce the discretization errors at material interfaces [2]. He also reported means to reduce the discretization errors for the second-order FDTD scheme, and to reduce the required PML thickness [3–7].

Following Berenger's seminal work, many papers appeared validating his technique [8–17] as well as applying FDTD with the PML medium [18–28]. Several modifications of the PML were also proposed to enhance its performance [3, 6, 14, 29–31]. The original split-field PML concept was also restated in a stretched-coordinate form [9, 32]. This form extends the use of the PML to other orthogonal coordinate systems [33–35] through a coordinate mapping of the metrics. The split-field PML can also be applied directly within an FDTD discretization using a very effective scheme referred to as the *convolutional PML* (CPML), introduced by Roden and Gedney [36]. The CPML accommodates more general metric tensor coefficients that can lead to improved absorption of slowly varying evanescent waves [36, 37].

The split-field PML was also posed as a uniaxial anisotropic medium having both magnetic and electric permittivity tensors [38, 39]. Introducing loss into the tensors results in a perfectly matched absorbing medium. This *uniaxial PML* (UPML) was first applied within the FDTD method by Gedney [24, 39]. The UPML is intriguing because it is based on a Maxwellian formulation rather than a mathematical model. In fact, attempts to realize a physical UPML medium have been reported [40, 41].

This chapter reviews the theoretical foundations of Berenger's split-field PML, its stretched-coordinate representation, the UPML, and the CPML. The relationships between these methods are highlighted. Efficient implementations of the UPML and CPML within the Cartesian FDTD space lattice are presented. Additional topics include the numerical discretization error of PML and the optimal choice of its constitutive parameters and spatial grading functions. We show that PML is applicable to FDTD models having regions of inhomogeneous, conductive, and dispersive media that exit normally from the space lattice.

7.2 PLANE WAVE INCIDENT UPON A LOSSY HALF-SPACE

To establish a foundation for our discussion of PML absorbers, this section reviews the physics of a sinusoidal plane wave impinging upon a conventional lossy material at the arbitrary angle θ relative to the x-axis. Specifically, we consider a TE_z-polarized wave incident from Region 1, the lossless material half-space $x < 0$, onto Region 2, the material half-space $x > 0$ having the electric conductivity σ and magnetic loss σ^*. Here, the incident-wave magnetic field is given by

$$\breve{\boldsymbol{H}}^{\text{inc}} = \hat{z}\, H_0\, e^{-j\beta_{1_x}x - j\beta_{1_y}y} \tag{7.1}$$

where the hat denotes a phasor quantity. Then, the total fields in Region 1 are given by

$$\breve{\boldsymbol{H}}_1 = \hat{z}\, H_0 \left(1 + \Gamma e^{2j\beta_{1_x}x}\right) e^{-j\beta_{1_x}x - j\beta_{1_y}y} \tag{7.2a}$$

$$\breve{\boldsymbol{E}}_1 = \left[-\hat{x}\,\frac{\beta_{1_y}}{\omega\varepsilon_1}\left(1 + \Gamma e^{2j\beta_{1_x}x}\right) + \hat{y}\,\frac{\beta_{1_x}}{\omega\varepsilon_1}\left(1 - \Gamma e^{2j\beta_{1_x}x}\right)\right] H_0\, e^{-j\beta_{1_x}x - j\beta_{1_y}y} \tag{7.2b}$$

The fields transmitted into Region 2 are given by

$$\breve{H}_2 = \hat{z}\, H_0 \tau\, e^{-j\beta_{2_x} x - j\beta_{2_y} y} \tag{7.3a}$$

$$\breve{E}_2 = \left[-\hat{x}\, \frac{\beta_{2_y}}{\omega\varepsilon_2\left(1+\dfrac{\sigma}{j\omega\varepsilon_2}\right)} + \hat{y}\, \frac{\beta_{2_x}}{\omega\varepsilon_2\left(1+\dfrac{\sigma}{j\omega\varepsilon_2}\right)} \right] H_0 \tau\, e^{-j\beta_{2_x} x - j\beta_{2_y} y} \tag{7.3b}$$

Γ and τ are the H-field reflection and transmission coefficients, respectively, and

$$\left.\begin{aligned} \beta_{1_x} &= k_1 \cos\theta \;; \qquad \beta_{1_y} = k_1 \sin\theta \end{aligned}\right\} \quad x < 0 \tag{7.4a}$$

$$\left.\begin{aligned} \beta_{2_x} &= \sqrt{(k_2)^2\left(1+\frac{\sigma}{j\omega\varepsilon_2}\right)\left(1+\frac{\sigma^*}{j\omega\mu_2}\right) - \left(\beta_{2_y}\right)^2} \end{aligned}\right\} \quad x > 0 \tag{7.4b}$$

where $k_i = \omega\sqrt{\mu_i \varepsilon_i}$ for $i = 1, 2$. Enforcing the continuity of the tangential fields across the Region-1 / Region-2 interface at $x = 0$ leads to $\beta_{2_y} = \beta_{1_y} = k_1 \sin\theta$ and

$$\Gamma = \frac{\dfrac{\beta_{1_x}}{\omega\varepsilon_1} - \dfrac{\beta_{2_x}}{\omega\varepsilon_2\left(1+\sigma/j\omega\varepsilon_2\right)}}{\dfrac{\beta_{1_x}}{\omega\varepsilon_1} + \dfrac{\beta_{2_x}}{\omega\varepsilon_2\left(1+\sigma/j\omega\varepsilon_2\right)}} \;; \qquad \tau = 1 + \Gamma \tag{7.5a, b}$$

In general, $\Gamma \neq 0$ for the arbitrary incidence angle θ. However, for the special case of normal incidence ($\theta = 0$) of the impinging wave, we have

$$\Gamma = \frac{\eta_1 - \eta_2}{\eta_1 + \eta_2} \tag{7.6}$$

where the wave impedances in Regions 1 and 2 are given by

$$\eta_1 = \sqrt{\frac{\mu_1}{\varepsilon_1}} \;; \qquad \eta_2 = \sqrt{\frac{\mu_2\left(1+\sigma^*/j\omega\mu_2\right)}{\varepsilon_2\left(1+\sigma/j\omega\varepsilon_2\right)}} \tag{7.7a, b}$$

Then, if we set $\varepsilon_1 = \varepsilon_2$ and $\mu_1 = \mu_2$ and further enforce the condition

$$\sigma^*/\mu_1 = \sigma/\varepsilon_1 \quad \rightarrow \quad \sigma^* = \sigma\mu_1/\varepsilon_1 = \sigma(\eta_1)^2 \tag{7.8}$$

then $k_1 = k_2$ and $\eta_1 = \eta_2$. This yields $\Gamma = 0$, a reflectionless Region-1 / Region-2 interface for the normally impinging wave. From (7.4b), we also find for this case that

$$\beta_{2_x} = \left(1 + \frac{\sigma}{j\omega\varepsilon_1}\right) k_1 = k_1 - j\sigma\eta_1 \tag{7.9}$$

yielding the following transmitted fields in Region 2:

$$\breve{\boldsymbol{E}}_2 = \hat{y}\,\eta_1 H_0 e^{-jk_1 x} e^{-\sigma\eta_1 x} \quad ; \qquad \breve{\boldsymbol{H}}_2 = \hat{z}\, H_0 e^{-jk_1 x} e^{-\sigma\eta_1 x} \tag{7.10a, b}$$

We observe that, for the normal-incidence case, the transmitted wave in Region 2 attenuates exponentially along the normal direction. Further, despite propagating in a lossy medium, this wave is dispersionless. That is, its wave speed is independent of frequency. Therefore, the material half-space Region 2, having magnetic and electric losses defined by (7.8), is perfectly matched to Region 1 for normally incident waves.

7.3 PLANE WAVE INCIDENT UPON BERENGER'S PML MEDIUM

The lossy medium discussed in Section 7.2 has been used with limited success as an absorbing layer to terminate the outer boundary of the FDTD space lattice [1, 42]. A major difficulty is that this medium is matched to the interior of the lattice only for normally incident waves. Therefore, oblique waves partially reflect back into the computation region and corrupt the solution. As a result, this absorber is not useful for most problems of practical interest.

Berenger provided the seminal insight that a nonphysical absorber can be postulated that is matched independent of the frequency *and* the angle of incidence of the impinging wave by exploiting additional degrees of freedom arising from a novel field-splitting strategy. In effect, this allows the construction of what he called a "perfectly matched layer" adjacent to the outer boundary of the FDTD space lattice for absorption of all impinging waves. Berenger reported local reflection coefficients for his PML 1/3,000th that of the Mur ABC discussed in Chapter 6. The overall noise power in the lattice due to reflections from the outer boundary was reduced by 70 or more decibels. This section reviews the theoretical basis of Berenger's PML for the case of plane-wave incidence upon a half-space comprised of this medium.

7.3.1 Two-Dimensional TE_z Case

Field-Splitting Modification of Maxwell's Equations

Consider again a TE_z plane wave impinging on the planar interface $x = 0$ of the material half-space Region 2. Within Region 2, Maxwell's curl equations (3.14a–c) as modified by Berenger are expressed in their time-dependent form as

$$\varepsilon_2 \frac{\partial E_x}{\partial t} + \sigma_y E_x = \frac{\partial H_z}{\partial y} \tag{7.11a}$$

$$\varepsilon_2 \frac{\partial E_y}{\partial t} + \sigma_x E_y = -\frac{\partial H_z}{\partial x} \tag{7.11b}$$

$$\mu_2 \frac{\partial H_{zx}}{\partial t} + \sigma_x^* H_{zx} = -\frac{\partial E_y}{\partial x} \tag{7.11c}$$

$$\mu_2 \frac{\partial H_{zy}}{\partial t} + \sigma_y^* H_{zy} = \frac{\partial E_x}{\partial y} \tag{7.11d}$$

Here, H_z is assumed to be split into two additive subcomponents

$$H_z = H_{zx} + H_{zy} \tag{7.12}$$

Further, the parameters σ_x and σ_y denote electric conductivities, and the parameters σ_x^* and σ_y^* denote magnetic losses.

We see that Berenger's formulation represents a generalization of normally modeled physical media. If $\sigma_x = \sigma_y = 0$ and $\sigma_x^* = \sigma_y^* = 0$, (7.11a–d) reduce to Maxwell's equations in a lossless medium. If $\sigma_x = \sigma_y = \sigma$ and $\sigma_x^* = \sigma_y^* = 0$, (7.11a–d) describe an electrically conductive medium. If $\varepsilon_1 = \varepsilon_2$, $\mu_1 = \mu_2$, $\sigma_x = \sigma_y = \sigma$, $\sigma_x^* = \sigma_y^* = \sigma^*$, and (7.8) is satisfied, then (7.11a–d) describe an absorbing medium that is impedance-matched to the Region-1 half-space $x < 0$ for normally incident plane waves, as discussed in Section 7.2.

Additional possibilities present themselves, however. If $\sigma_y = \sigma_y^* = 0$, then the medium can absorb a plane wave having field components (E_y, H_{zx}) propagating along x, but does not absorb a wave having field components (E_x, H_{zy}) propagating along y, since in the first case propagation is governed by (7.11b, c), and in the second case by (7.11a, d). The converse situation is true for waves (E_y, H_{zx}) and (E_x, H_{zy}) if $\sigma_x = \sigma_x^* = 0$. These properties of particular Berenger media characterized by the pairwise parameter sets $(\sigma_x, \sigma_x^*, 0, 0)$ and $(0, 0, \sigma_y, \sigma_y^*)$ are closely related to the fundamental premise of this novel ABC, proved later. That is, if the pairwise electric and magnetic losses satisfy (7.8), then at interfaces normal to x and y, respectively, the Berenger media have zero reflection of electromagnetic waves.

Now consider (7.11a–d) expressed in their time-harmonic form in the Berenger medium. Again letting the hat symbol denote a phasor quantity, we write

$$j\omega\varepsilon_2\left(1 + \frac{\sigma_y}{j\omega\varepsilon_2}\right)\breve{E}_x = \frac{\partial}{\partial y}\left(\breve{H}_{zx} + \breve{H}_{zy}\right) \tag{7.13a}$$

$$j\omega\varepsilon_2\left(1 + \frac{\sigma_x}{j\omega\varepsilon_2}\right)\breve{E}_y = -\frac{\partial}{\partial x}\left(\breve{H}_{zx} + \breve{H}_{zy}\right) \tag{7.13b}$$

$$j\omega\mu_2\left(1 + \frac{\sigma_x^*}{j\omega\mu_2}\right)\breve{H}_{zx} = -\frac{\partial \breve{E}_y}{\partial x} \tag{7.13c}$$

$$j\omega\mu_2\left(1 + \frac{\sigma_y^*}{j\omega\mu_2}\right)\breve{H}_{zy} = \frac{\partial \breve{E}_x}{\partial y} \tag{7.13d}$$

The notation is simplified by introducing the variables

$$s_w = \left(1 + \frac{\sigma_w}{j\omega\varepsilon_2}\right) \quad ; \quad s_w^* = \left(1 + \frac{\sigma_w^*}{j\omega\mu_2}\right) \quad : \; w = x, \, y \tag{7.14a, b}$$

Then, (7.13a, b) are rewritten as

$$j\omega\varepsilon_2 s_y \breve{E}_x = \frac{\partial}{\partial y}\left(\breve{H}_{zx} + \breve{H}_{zy}\right) ; \quad j\omega\varepsilon_2 s_x \breve{E}_y = -\frac{\partial}{\partial x}\left(\breve{H}_{zx} + \breve{H}_{zy}\right) \tag{7.15a, b}$$

Plane-Wave Solution Within the Berenger Medium

The next step is to derive the plane-wave solution within the Berenger medium. To this end, (7.15a) is differentiated with respect to y, and (7.15b) with respect to x. Substituting the expressions for $\partial E_y / \partial x$ and $\partial \breve{E}_x / \partial y$ from (7.13c, d) leads to

$$-\omega^2 \mu_2 \varepsilon_2 \breve{H}_{zx} = -\frac{1}{s_x^*}\frac{\partial}{\partial x}\frac{1}{s_x}\frac{\partial}{\partial x}\left(\breve{H}_{zx} + \breve{H}_{zy}\right) \tag{7.16a}$$

$$-\omega^2 \mu_2 \varepsilon_2 \breve{H}_{zy} = -\frac{1}{s_y^*}\frac{\partial}{\partial y}\frac{1}{s_y}\frac{\partial}{\partial y}\left(\breve{H}_{zx} + \breve{H}_{zy}\right) \tag{7.16b}$$

Adding these together and using (7.12), we obtain the representative wave equation

$$\frac{1}{s_x^*}\frac{\partial}{\partial x}\frac{1}{s_x}\frac{\partial}{\partial x}\breve{H}_z + \frac{1}{s_y^*}\frac{\partial}{\partial y}\frac{1}{s_y}\frac{\partial}{\partial y}\breve{H}_z + \omega^2 \mu_2 \varepsilon_2 \breve{H}_z = 0 \tag{7.17}$$

This wave equation supports the solutions

$$\breve{H}_z = H_0 \tau \, e^{-j\sqrt{s_x s_x^*}\,\beta_{2_x} x \, - \, j\sqrt{s_y s_y^*}\,\beta_{2_y} y} \tag{7.18}$$

with the dispersion relationship

$$\left(\beta_{2_x}\right)^2 + \left(\beta_{2_y}\right)^2 = \left(k_2\right)^2 \quad \rightarrow \quad \beta_{2_x} = \left[\left(k_2\right)^2 - \left(\beta_{2_y}\right)^2\right]^{1/2} \tag{7.19}$$

Then, from (7.15a, b) and (7.12), we have

$$\breve{E}_x = -H_0 \tau \frac{\beta_{2_y}}{\omega\varepsilon_2}\sqrt{\frac{s_y^*}{s_y}}\, e^{-j\sqrt{s_x s_x^*}\,\beta_{2_x} x \, - \, j\sqrt{s_y s_y^*}\,\beta_{2_y} y} \tag{7.20a}$$

$$\breve{E}_y = H_0 \tau \frac{\beta_{2_x}}{\omega \varepsilon_2} \sqrt{\frac{s_x^*}{s_x}}\, e^{-j\sqrt{s_x s_x^*}\,\beta_{2_x} x \;-\; j\sqrt{s_y s_y^*}\,\beta_{2_y} y} \tag{7.20b}$$

Despite field splitting, continuity of the tangential E and H fields must be preserved across the $x = 0$ interface for the fields given by (7.2), (7.18), and (7.20). To enforce this continuity, we have $s_y = s_y^* = 1$, or equivalently, $\sigma_y = 0 = \sigma_y^*$. This yields the phase-matching condition $\beta_{2_y} = \beta_{1_y} = k_1 \sin\theta$. Further, we derive the H-field reflection and transmission coefficients

$$\Gamma = \left(\frac{\beta_{1_x}}{\omega \varepsilon_1} - \frac{\beta_{2_x}}{\omega \varepsilon_2} \sqrt{\frac{s_x^*}{s_x}} \right) \cdot \left(\frac{\beta_{1_x}}{\omega \varepsilon_1} + \frac{\beta_{2_x}}{\omega \varepsilon_2} \sqrt{\frac{s_x^*}{s_x}} \right)^{-1} ; \qquad \tau = 1 + \Gamma \tag{7.21a, b}$$

Reflectionless Matching Condition

Now, assume $\varepsilon_1 = \varepsilon_2$, $\mu_1 = \mu_2$, and $s_x = s_x^*$. This is equivalent to $k_1 = k_2$, $\eta_1 = \sqrt{\mu_1 / \varepsilon_1} = \sqrt{\mu_2 / \varepsilon_2}$, and $\sigma_x / \varepsilon_1 = \sigma_x^* / \mu_1$ [i.e., σ_x and σ_x^* satisfying (7.8) in a pairwise manner]. With $\beta_{2_y} = \beta_{1_y}$, (7.19) now yields $\beta_{2_x} = \beta_{1_x}$. Substituting into (7.21) gives the reflectionless condition $\Gamma = 0$ for *all* incident angles θ. For this case, (7.18) and (7.20) specify the following transmitted fields within the Berenger medium $x > 0$:

$$\breve{H}_z = H_0 e^{-j s_x \beta_{1_x} x \;-\; j\beta_{1_y} y} = H_0 e^{-j\beta_{1_x} x \;-\; j\beta_{1_y} y} e^{-\sigma_x x \eta_1 \cos\theta} \tag{7.22a}$$

$$\breve{E}_x = -H_0 \eta_1 \sin\theta\, e^{-j\beta_{1_x} x \;-\; j\beta_{1_y} y} e^{-\sigma_x x \eta_1 \cos\theta} \tag{7.22b}$$

$$\breve{E}_y = H_0 \eta_1 \cos\theta\, e^{-j\beta_{1_x} x \;-\; j\beta_{1_y} y} e^{-\sigma_x x \eta_1 \cos\theta} \tag{7.22c}$$

Within the matched Berenger medium, the transmitted wave propagates with the same speed and direction as the impinging wave, while simultaneously undergoing exponential decay along the x-axis normal to the interface between Regions 1 and 2. The attenuation factor $\sigma_x \eta_1 \cos\theta$ is independent of frequency. Unlike the conventional lossy material considered in Section 7.2, these desirable properties apply to all angles of incidence. Hence, Berenger's coining of the term "perfectly matched layer" makes excellent sense.

Structure of an FDTD Grid Employing Berenger's PML ABC

The above analysis can be repeated for PMLs that are normal to the y-direction. This permitted Berenger to propose the two-dimensional TE_z FDTD grid shown in Fig. 7.1, which uses PMLs to greatly reduce outer-boundary reflections. Here, a free-space computation zone is surrounded on all sides by PML backed by PEC walls.

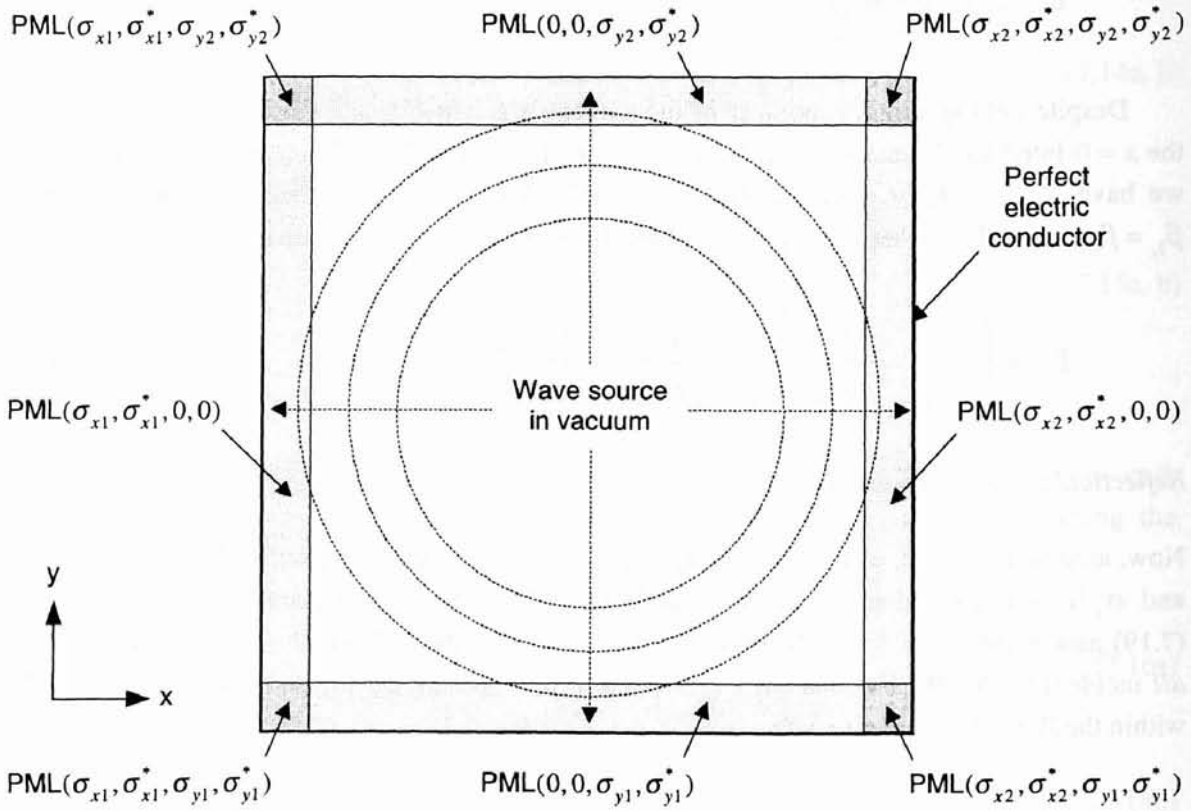


Fig. 7.1 Structure of a two-dimensional TE_z FDTD grid employing the Berenger PML ABC. *After:* Berenger, *J. Computational Physics*, 1994, pp. 185–200.

Referring to Fig. 7.1, at the left and right sides of the grid (x_1 and x_2), each PML has σ_x and σ_x^* matched according to (7.8), along with $\sigma_y = 0 = \sigma_y^*$ to permit reflectionless transmission across the vacuum-PML interface. At the lower and upper sides of the grid (y_1 and y_2), each PML has σ_y and σ_y^* matched according to (7.8), along with $\sigma_x = 0 = \sigma_x^*$. At the four corners of the grid where there is overlap of two PMLs, all four losses (σ_x, σ_x^*, σ_y, and σ_y^*) are present and set equal to those of the adjacent PMLs.

Impact of Numerical Dispersion

A reasonable question to ask at this point is whether or not numerical phase-velocity dispersion limits the ultimate effectiveness of the Berenger PML ABC as it does the analytical ABCs of Chapter 6. Based upon the experience with PML to date, the answer is that measured PML reflection coefficients are orders of magnitude smaller than measured numerical phase-velocity variations with incident wave angle and frequency. Therefore, the absorptive action of the PML is robust relative to numerical dispersion artifacts in the FDTD space lattice.

7.3.2 Two-Dimensional TM_z Case

The analysis of Section 7.3.1 can be repeated for the case of a TM_z-polarized incident wave, wherein we implement the field splitting $E_z = E_{zx} + E_{zy}$. Analogous to (7.11), the PML-modified Maxwell's equations for the TM_z case are as follows:

$$\mu_2 \frac{\partial H_x}{\partial t} + \sigma_y^* H_x = -\frac{\partial E_z}{\partial y} \tag{7.23a}$$

$$\mu_2 \frac{\partial H_y}{\partial t} + \sigma_x^* H_y = \frac{\partial E_z}{\partial x} \tag{7.23b}$$

$$\varepsilon_2 \frac{\partial E_{zx}}{\partial t} + \sigma_x E_{zx} = \frac{\partial H_y}{\partial x} \tag{7.23c}$$

$$\varepsilon_2 \frac{\partial E_{zy}}{\partial t} + \sigma_y E_{zy} = -\frac{\partial H_x}{\partial y} \tag{7.23d}$$

A derivation of the PML properties analogous to that of the TE_z case yields slightly changed results. In most of the equations, the change is only a permutation of ε_2 with μ_2, and of σ with σ^*. However, the PML matching conditions are unchanged. This permits an absorbing reflectionless layer to be constructed adjacent to the outer grid boundary, as in the TE_z case.

7.3.3 Three-Dimensional Case

Analogous to the two-dimensional cases of (7.11a–d) and (7.23a–d), three-dimensional time-domain Maxwell's equations for Berenger's split-field PML have been developed [5, 8, 12]. In three dimensions, all six Cartesian field vector components are split. For example, the modified Ampere's law is given by

$$\left(\varepsilon \frac{\partial}{\partial t} + \sigma_y\right) E_{xy} = \frac{\partial}{\partial y}\left(H_{zx} + H_{zy}\right) \tag{7.24a}$$

$$\left(\varepsilon \frac{\partial}{\partial t} + \sigma_z\right) E_{xz} = -\frac{\partial}{\partial z}\left(H_{yx} + H_{yz}\right) \tag{7.24b}$$

$$\left(\varepsilon \frac{\partial}{\partial t} + \sigma_z\right) E_{yz} = \frac{\partial}{\partial z}\left(H_{xy} + H_{xz}\right) \tag{7.24c}$$

$$\left(\varepsilon \frac{\partial}{\partial t} + \sigma_x\right) E_{yx} = -\frac{\partial}{\partial x}\left(H_{zx} + H_{zy}\right) \tag{7.24d}$$

$$\left(\varepsilon\frac{\partial}{\partial t} + \sigma_x\right) E_{zx} = \frac{\partial}{\partial x}\left(H_{yx} + H_{yz}\right) \tag{7.24e}$$

$$\left(\varepsilon\frac{\partial}{\partial t} + \sigma_y\right) E_{zy} = -\frac{\partial}{\partial y}\left(H_{xy} + H_{xz}\right) \tag{7.24f}$$

Similarly, the modified Faraday's law is given by

$$\left(\mu\frac{\partial}{\partial t} + \sigma_y^*\right) H_{xy} = -\frac{\partial}{\partial y}\left(E_{zx} + E_{zy}\right) \tag{7.25a}$$

$$\left(\mu\frac{\partial}{\partial t} + \sigma_z^*\right) H_{xz} = \frac{\partial}{\partial z}\left(E_{yx} + E_{yz}\right) \tag{7.25b}$$

$$\left(\mu\frac{\partial}{\partial t} + \sigma_z^*\right) H_{yz} = -\frac{\partial}{\partial z}\left(E_{xy} + E_{xz}\right) \tag{7.25c}$$

$$\left(\mu\frac{\partial}{\partial t} + \sigma_x^*\right) H_{yx} = \frac{\partial}{\partial x}\left(E_{zx} + E_{zy}\right) \tag{7.25d}$$

$$\left(\mu\frac{\partial}{\partial t} + \sigma_x^*\right) H_{zx} = -\frac{\partial}{\partial x}\left(E_{yx} + E_{yz}\right) \tag{7.25e}$$

$$\left(\mu\frac{\partial}{\partial t} + \sigma_y^*\right) H_{zy} = \frac{\partial}{\partial y}\left(E_{xy} + E_{xz}\right) \tag{7.25f}$$

PML matching conditions analogous to the two-dimensional cases discussed previously are used. Specifically, if we denote $w = x, y, z$, the matching condition at a normal-to-w PML interface in the FDTD lattice has the parameter pair (σ_w, σ_w^*) satisfy (7.8). This causes the transmitted wave within the PML to undergo exponential decay in the $\pm w$-directions. All other (σ_w, σ_w^*) pairs within this PML are zero. In a corner region, the PML is provided with each matched (σ_w, σ_w^*) pair that is assigned to the overlapping PMLs forming the corner. Thus, PML media located in dihedral-corner overlapping regions have two nonzero and one zero (σ_w, σ_w^*) pairs, and in trihedral-corner overlapping regions have three nonzero (σ_w, σ_w^*) pairs.

7.4 STRETCHED-COORDINATE FORMULATION OF BERENGER'S PML

A more compact form of the split-field equations of (7.24) and (7.25) was introduced by Chew and Weedon [9], and independently by Rappaport [32]. Here, the split-field equations are reposed in a nonsplit form that maps Maxwell's equations into a complex coordinate space. To this end, the following coordinate mapping is introduced:

$$\tilde{x} \rightarrow \int_0^x s_x(x')dx' \; ; \quad \tilde{y} \rightarrow \int_0^y s_y(y')dy' \; ; \quad \tilde{z} \rightarrow \int_0^z s_z(z')dz' \tag{7.26}$$

In (7.26), we assume that the PML parameters s_w are continuous functions along the axial directions. The partial derivatives in the stretched coordinate space are then

$$\frac{\partial}{\partial \tilde{x}} = \frac{1}{s_x}\frac{\partial}{\partial x} \; ; \qquad \frac{\partial}{\partial \tilde{y}} = \frac{1}{s_y}\frac{\partial}{\partial y} \; ; \qquad \frac{\partial}{\partial \tilde{z}} = \frac{1}{s_z}\frac{\partial}{\partial z} \tag{7.27}$$

Thus, the ∇ operator in the mapped space is defined as

$$\tilde{\nabla} = \hat{x}\frac{\partial}{\partial \tilde{x}} + \hat{y}\frac{\partial}{\partial \tilde{y}} + \hat{z}\frac{\partial}{\partial \tilde{z}} = \hat{x}\frac{1}{s_x}\frac{\partial}{\partial x} + \hat{y}\frac{1}{s_y}\frac{\partial}{\partial y} + \hat{z}\frac{1}{s_z}\frac{\partial}{\partial z} \tag{7.28}$$

The time-harmonic Maxwell's equations in the complex-coordinate stretched space are then expressed as

$$\begin{aligned} j\omega\varepsilon\breve{\boldsymbol{E}} = \tilde{\nabla}\times\breve{\boldsymbol{H}} = {} & \hat{x}\left(\frac{1}{s_y}\frac{\partial}{\partial y}\breve{H}_z - \frac{1}{s_z}\frac{\partial}{\partial z}\breve{H}_y\right) + \hat{y}\left(\frac{1}{s_z}\frac{\partial}{\partial z}\breve{H}_x - \frac{1}{s_x}\frac{\partial}{\partial x}\breve{H}_z\right) \\ & + \hat{z}\left(\frac{1}{s_x}\frac{\partial}{\partial x}\breve{H}_y - \frac{1}{s_y}\frac{\partial}{\partial y}\breve{H}_x\right) \end{aligned} \tag{7.29}$$

$$\begin{aligned} -j\omega\mu\breve{\boldsymbol{H}} = \tilde{\nabla}\times\breve{\boldsymbol{E}} = {} & \hat{x}\left(\frac{1}{s_y}\frac{\partial}{\partial y}\breve{E}_z - \frac{1}{s_z}\frac{\partial}{\partial z}\breve{E}_y\right) + \hat{y}\left(\frac{1}{s_z}\frac{\partial}{\partial z}\breve{E}_x - \frac{1}{s_x}\frac{\partial}{\partial x}\breve{E}_z\right) \\ & + \hat{z}\left(\frac{1}{s_x}\frac{\partial}{\partial x}\breve{E}_y - \frac{1}{s_y}\frac{\partial}{\partial y}\breve{E}_x\right) \end{aligned} \tag{7.30}$$

A direct relationship can now be shown between the stretched-coordinate form of Maxwell's equations and Berenger's split-field PML. To demonstrate this, we first rewrite the split-field equations of (7.24) for the time-harmonic case:

$$j\omega\varepsilon s_y\breve{E}_{xy} = \frac{\partial}{\partial y}\left(\breve{H}_{zx} + \breve{H}_{zy}\right) \tag{7.31a}$$

$$j\omega\varepsilon s_z\breve{E}_{xz} = -\frac{\partial}{\partial z}\left(\breve{H}_{yx} + \breve{H}_{yz}\right) \tag{7.31b}$$

$$j\omega\varepsilon s_z\breve{E}_{yz} = \frac{\partial}{\partial z}\left(\breve{H}_{xy} + \breve{H}_{xz}\right) \tag{7.31c}$$

$$j\omega\varepsilon s_x \breve{E}_{yx} = -\frac{\partial}{\partial x}\left(\breve{H}_{zx} + \breve{H}_{zy}\right) \tag{7.31d}$$

$$j\omega\varepsilon s_x \breve{E}_{zx} = \frac{\partial}{\partial x}\left(\breve{H}_{yx} + \breve{H}_{yz}\right) \tag{7.31e}$$

$$j\omega\varepsilon s_y \breve{E}_{zy} = -\frac{\partial}{\partial y}\left(\breve{H}_{xy} + \breve{H}_{xz}\right) \tag{7.31f}$$

Then, add (7.31a) + (7.31b), (7.31c) + (7.31d), and (7.31e) + (7.31f), and use the relationships $E_x = E_{xy} + E_{xz}$, $E_y = E_{yx} + E_{yz}$, and $E_z = E_{zx} + E_{zy}$. This yields

$$j\omega\varepsilon \breve{E}_x = \frac{1}{s_y}\frac{\partial}{\partial y}\breve{H}_z - \frac{1}{s_z}\frac{\partial}{\partial z}\breve{H}_y \tag{7.32a}$$

$$j\omega\varepsilon \breve{E}_y = \frac{1}{s_z}\frac{\partial}{\partial z}\breve{H}_x - \frac{1}{s_x}\frac{\partial}{\partial x}\breve{H}_z \tag{7.32b}$$

$$j\omega\varepsilon \breve{E}_z = \frac{1}{s_x}\frac{\partial}{\partial x}\breve{H}_y - \frac{1}{s_y}\frac{\partial}{\partial y}\breve{H}_x \tag{7.32c}$$

which is identical to (7.29). This procedure is repeated for the split-field equations of (7.25) rewritten for the time-harmonic case, leading exactly to (7.30). Specifically, we see that the stretched-coordinate form of the PML is equivalent to the split-field PML; however, it reposes it in a nonsplit form.

In the time-domain, the stretched coordinate form of the PML is expressed as

$$\frac{\partial \boldsymbol{D}}{\partial t} = \hat{x}\left(\bar{s}_y * \frac{\partial}{\partial y}H_z - \bar{s}_z * \frac{\partial}{\partial z}H_y\right) + \hat{y}\left(\bar{s}_z * \frac{\partial}{\partial z}H_x - \bar{s}_x * \frac{\partial}{\partial x}H_z\right) + \hat{z}\left(\bar{s}_x * \frac{\partial}{\partial x}H_y - \bar{s}_y * \frac{\partial}{\partial y}H_x\right) \tag{7.33a}$$

$$-\frac{\partial \boldsymbol{B}}{\partial t} = \hat{x}\left(\bar{s}_y * \frac{\partial}{\partial y}E_z - \bar{s}_z * \frac{\partial}{\partial z}E_y\right) + \hat{y}\left(\bar{s}_z * \frac{\partial}{\partial z}E_x - \bar{s}_x * \frac{\partial}{\partial x}E_z\right) + \hat{z}\left(\bar{s}_x * \frac{\partial}{\partial x}E_y - \bar{s}_y * \frac{\partial}{\partial y}E_x\right) \tag{7.33b}$$

where * denotes a convolution,

$$\bar{s}_w = \mathcal{F}^{-1}\left(\frac{1}{s_w}\right) \qquad w = x, y, z \tag{7.34}$$

and $\mathcal{F}^{-1}$ denotes the inverse Fourier transform.

The complex stretched-coordinate formulation allows straightforward mathematical manipulation of the PML equations, which in turn allows extension of our understanding of PML behavior. For example, the stretched-coordinate formulation has provided a pathway to mapping the PML into other orthogonal coordinate systems or into general curvilinear coordinate systems [33–35]. It also can be shown that if the metric coefficients s_w are independent of the material parameters, then the PML matches waves propagating in a general host medium. As shown in Section 7.8, a highly effective discrete implementation of the stretched-coordinate formulation can be applied within the FDTD method.

7.5 AN ANISOTROPIC PML ABSORBING MEDIUM

The split-field PML introduced by Berenger is a hypothetical medium based on a mathematical model. Neither it nor its stretched-coordinate analog is a physical medium. Due to the coordinate dependence of the loss terms, if such a physical medium exists, it must be anisotropic.

Indeed, a physical model based on an anisotropic, perfectly matched medium can be formulated. This was first discussed by Sacks et al. [38]. For a single interface, the anisotropic medium is uniaxial, and is composed of both electric and magnetic permittivity tensors. The uniaxial material performs as well as Berenger's PML while avoiding the nonphysical field splitting. This section introduces the theoretical basis of the uniaxial PML, and compares its formulation with Berenger's PML and the stretched-coordinate PML.

7.5.1 Perfectly Matched Uniaxial Medium

Consider an arbitrarily polarized time-harmonic plane wave $\breve{\boldsymbol{H}}^{\text{inc}} = \boldsymbol{H}_0 e^{-j\beta_{1_x} x - j\beta_{1_y} y}$ propagating in Region 1, the isotropic half-space $x < 0$. This wave is assumed to be incident on Region 2, the half-space $x > 0$ comprised of a uniaxial anisotropic medium having the permittivity and permeability tensors

$$\bar{\bar{\varepsilon}}_2 = \varepsilon_2 \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & b \end{bmatrix}, \qquad \bar{\bar{\mu}}_2 = \mu_2 \begin{bmatrix} c & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d \end{bmatrix} \tag{7.35}$$

Here, $\varepsilon_{yy} = \varepsilon_{zz}$ and $\mu_{yy} = \mu_{zz}$, since the medium is assumed to be rotationally symmetric about the x-axis.

The fields excited within Region 2 are also plane-wave in nature and satisfy Maxwell's curl equations. We obtain

$$\boldsymbol{\beta}_2 \times \breve{\boldsymbol{E}} = \omega \bar{\bar{\mu}}_2 \breve{\boldsymbol{H}} ; \qquad \boldsymbol{\beta}_2 \times \breve{\boldsymbol{H}} = -\omega \bar{\bar{\varepsilon}}_2 \breve{\boldsymbol{E}} \tag{7.36}$$

where $\boldsymbol{\beta}_2 = \hat{x}\beta_{2_x} + \hat{y}\beta_{2_y}$ is the wavevector in anisotropic Region 2. In turn, this permits the derivation of the wave equation

$$\boldsymbol{\beta}_2 \times \left(\bar{\bar{\varepsilon}}_2^{-1}\boldsymbol{\beta}_2\right) \times \breve{\boldsymbol{H}} + \omega^2 \bar{\bar{\mu}}_2 \breve{\boldsymbol{H}} = \mathbf{0} \tag{7.37}$$

Expressing the cross products as matrix operators, this wave equation can be expressed in matrix form as

$$\begin{bmatrix} k_2^2 c - \left(\beta_{2_y}\right)^2 b^{-1} & \beta_{2_x}\beta_{2_y} b^{-1} & 0 \\ \beta_{2_x}\beta_{2_y} b^{-1} & k_2^2 d - \left(\beta_{2_x}\right)^2 b^{-1} & 0 \\ 0 & 0 & k_2^2 d - \left(\beta_{2_x}\right)^2 b^{-1} - \left(\beta_{2_y}\right)^2 a^{-1} \end{bmatrix} \begin{bmatrix} \breve{H}_x \\ \breve{H}_y \\ \breve{H}_z \end{bmatrix} = \mathbf{0} \tag{7.38}$$

where $k_2^2 = \omega^2 \mu_2 \varepsilon_2$. The dispersion relation for the uniaxial medium in Region 2 is derived from the determinant of the matrix operator. Solving for β_{2_x}, we find that there are four eigenmode solutions. Conveniently, these solutions can be decoupled into forward and backward TE_z and TM_z modes, which satisfy the dispersion relations

$$k_2^2 - \left(\beta_{2_x}\right)^2 b^{-1} d^{-1} - \left(\beta_{2_y}\right)^2 a^{-1} d^{-1} = 0 \qquad : \mathrm{TE}_z \ (\breve{H}_x, \breve{H}_y = 0) \tag{7.39}$$

$$k_2^2 - \left(\beta_{2_x}\right)^2 b^{-1} d^{-1} - \left(\beta_{2_y}\right)^2 b^{-1} c^{-1} = 0 \qquad : \mathrm{TM}_z \ (\breve{H}_z = 0) \tag{7.40}$$

The reflection coefficient at the interface $x = 0$ of Regions 1 and 2 can now be derived. Let us assume a TE_z incident wave in Region 1. Then, in isotropic Region 1, the fields are expressed as a superposition of the incident and reflected fields as

$$\breve{\boldsymbol{H}}_1 = \hat{z} H_0 \left(1 + \Gamma e^{2j\beta_{1_x} x}\right) e^{-j\beta_{1_x} x - j\beta_{1_y} y} \tag{7.41a}$$

$$\breve{\boldsymbol{E}}_1 = \left[-\hat{x}\frac{\beta_{1_y}}{\omega\varepsilon_1}\left(1 + \Gamma e^{2j\beta_{1_x} x}\right) + \hat{y}\frac{\beta_{1_x}}{\omega\varepsilon_1}\left(1 - \Gamma e^{2j\beta_{1_x} x}\right)\right] H_0 e^{-j\beta_{1_x} x - j\beta_{1_y} y} \tag{7.41b}$$

The wave transmitted into Region 2 is also TE_z with propagation characteristics governed by (7.39). These fields are expressed as

$$\breve{\boldsymbol{H}}_2 = \hat{z}\, H_0\, \tau\, e^{-j\beta_{2_x} x - j\beta_{2_y} y} \tag{7.42a}$$

$$\breve{\boldsymbol{E}}_2 = \left(-\hat{x}\, \frac{\beta_{2_y}}{\omega \varepsilon_2 a} + \hat{y}\, \frac{\beta_{2_x}}{\omega \varepsilon_2 b} \right) H_0\, \tau\, e^{-j\beta_{2_x} x - j\beta_{2_y} y} \tag{7.42b}$$

where Γ and τ are the H-field reflection and transmission coefficients, respectively. These are derived by enforcing continuity of the tangential E and H fields across $x = 0$, and are given by

$$\Gamma = \frac{\beta_{1_x} - \beta_{2_x} b^{-1}}{\beta_{1_x} + \beta_{2_x} b^{-1}} \; ; \qquad \tau = 1 + \Gamma = \frac{2\beta_{1_x}}{\beta_{1_x} + \beta_{2_x} b^{-1}} \tag{7.43a, b}$$

Further, for all angles of wave incidence we have

$$\beta_{2_y} = \beta_{1_y} \tag{7.44}$$

due to phase-matching across the $x = 0$ interface. Substituting (7.44) into (7.39) and solving for β_{2_x} yields

$$\beta_{2_x} = \sqrt{k_2^2 b d - \left(\beta_{1_y}\right)^2 a^{-1} b} \tag{7.45}$$

Then, if we set $\varepsilon_1 = \varepsilon_2$, $\mu_1 = \mu_2$, $d = b$, and $a^{-1} = b$, we have $k_2 = k_1$ and

$$\beta_{2_x} = \sqrt{k_1^2 b^2 - \left(\beta_{1_y}\right)^2 b^2} = b\sqrt{k_1^2 - \left(\beta_{1_y}\right)^2} \equiv b\beta_{1_x} \tag{7.46}$$

Substituting (7.46) into (7.43) yields $\Gamma = 0$ for all β_{1_x}. Thus, the interface between Regions 1 and 2 is reflectionless for all angles of wave incidence.

The above exercise can be repeated for the TM_z-polarization case. Here, the E-field reflection coefficient is the dual of (7.43), and is found by replacing b with d (and vice versa), and a with c. For this case, the reflectionless condition holds if $b = d$ and $c^{-1} = d$.

Combining the results for the TE_z and TM_z cases, we see that reflectionless wave transmission into Region 2 occurs when it is composed of a uniaxial medium having the ε and μ tensors

$$\bar{\bar{\varepsilon}}_2 = \varepsilon_1 \bar{\bar{s}} \; ; \qquad \bar{\bar{\mu}}_2 = \mu_1 \bar{\bar{s}} \; ; \qquad \bar{\bar{s}} = \begin{bmatrix} s_x^{-1} & 0 & 0 \\ 0 & s_x & 0 \\ 0 & 0 & s_x \end{bmatrix} \tag{7.47}$$

This reflectionless property is completely independent of the angle of incidence, polarization, and frequency of the incident wave. Further, from (7.39) and (7.40), the propagation characteristics of the TE- and TM-polarized waves are identical. We call this medium a "uniaxial PML" in recognition of its uniaxial anisotropy and perfect matching.

Similar to Berenger's PML, the reflectionless property of the UPML in Region 2 is valid for any s_x. For example, choose $s_x = 1 + \sigma_x / j\omega\varepsilon_1 = 1 - j\sigma_x / \omega\varepsilon_1$. Then, from (7.46), we have

$$\beta_{2_x} = s_x \beta_{1_x} = \left(1 - j\sigma_x / \omega\varepsilon_1\right) \beta_{1_x} \tag{7.48}$$

We note that the real part of β_{2_x} is identical to β_{1_x}. Combined with (7.42), this implies that the phase velocities of the impinging and transmitted waves are identical for all incident angles. The characteristic wave impedance in Region 2 is also identical to that in Region 1, which is a consequence of the fact that the media are perfectly matched.

Finally, substituting (7.44) and (7.48) into (7.42) yields the fields transmitted into the Region-2 UPML for a TE_z incident wave:

$$\breve{\boldsymbol{H}}_2 = \hat{z} H_0 e^{-j\beta_{1_x} x - j\beta_{1_y} y} e^{-\sigma_x x \eta_1 \cos\theta} \tag{7.49a}$$

$$\breve{\boldsymbol{E}}_2 = \left(-\hat{x} s_x \eta_1 \sin\theta + \hat{y} \eta_1 \cos\theta\right) H_0 e^{-j\beta_{1_x} x - j\beta_{1_y} y} e^{-\sigma_x x \eta_1 \cos\theta} \tag{7.49b}$$

Here, $\eta_1 = \sqrt{\mu_1 / \varepsilon_1}$ and θ is the angle of incidence relative to the x-axis. Thus, the transmitted wave in the UPML propagates with the same phase velocity as the incident wave, while simultaneously undergoing exponential decay along the x-axis normal to the interface between Regions 1 and 2. The attenuation factor is independent of frequency, although it is dependent on θ and the UPML conductivity σ_x.

7.5.2 Relationship to Berenger's Split-Field PML

Comparing the E- and H-fields transmitted into the UPML in (7.49) with the corresponding fields for Berenger's split-field PML in (7.22), we observe identical tangential fields and identical propagation characteristics. Further examination of (7.17) and (7.37) reveals that the two methods result in the same wave equation. Consequently, the plane waves satisfy the same dispersion relation.

However, upon comparing the normal electric fields E_x, we note that they differ by a factor of s_x. In fact, comparing the incident fields, it is clear that in the split-field formulation, E_x is continuous across the $x = 0$ boundary. However, for the UPML, E_x is discontinuous while $D_x = s_x^{-1} E_x$ is continuous. This implies that the two methods host different divergence theorems.

Within the UPML, Gauss' law for the E-field is explicitly written as

$$\nabla \cdot \boldsymbol{D} = \nabla \cdot \left(\varepsilon \bar{\bar{s}} \boldsymbol{E}\right) = \frac{\partial}{\partial x}\left(\varepsilon s_x^{-1} E_x\right) + \frac{\partial}{\partial y}\left(\varepsilon s_x E_y\right) + \frac{\partial}{\partial z}\left(\varepsilon s_x E_z\right) = 0 \tag{7.50}$$

This implies that $D_x = \varepsilon s_x^{-1} E_x$ must be continuous across the $x = 0$ interface, since no sources are assumed here. It was shown that ε must be continuous across the interface for a perfectly matched condition. Thus, D_x and $s_x^{-1} E_x$ must be continuous across $x = 0$. Comparing (7.49) with (7.41), this is indeed true for the TE_z-polarized wave.

Next, consider Gauss' law for Berenger's split-field PML formulation. The ∇ operator in the stretched-coordinate space of interest is defined as

$$\nabla = \hat{x}\frac{\partial}{s_x \partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z} \tag{7.51}$$

Therefore, we can express the divergence of the electric flux density as

$$\frac{1}{s_x}\frac{\partial}{\partial x}(\varepsilon E_x) + \frac{\partial}{\partial y}(\varepsilon E_y) + \frac{\partial}{\partial z}(\varepsilon E_z) = 0 \tag{7.52}$$

Since ε is continuous across the boundary and s_x^{-1} occurs outside the derivative, both E_x and D_x are continuous.

In summary, Berenger's split-field PML and the UPML have the same propagation characteristics, since they both result in the same wave equation. They also have identical reflection properties in a discrete FDTD implementation, as observed by Gedney [39] and proven analytically by Berenger [37]. However, the two formulations have different Gauss' laws. Hence, the E- and H-field components that are normal to the PML interface are different.

7.5.3 A Generalized Three-Dimensional Formulation

We now show that properly defining a general constitutive tensor $\bar{\bar{s}}$ allows the UPML medium to be used throughout the entire FDTD space lattice. This tensor provides for both a lossless, isotropic medium in the primary computation zone, *and* individual UPML absorbers adjacent to the outer lattice boundary planes for mitigation of spurious wave reflections.

For a matched condition, the time-harmonic Maxwell's curl equations in the UPML can be written in their most general form as

$$\nabla \times \breve{H} = j\omega\varepsilon\bar{\bar{s}}\breve{E} \quad ; \qquad \nabla \times \breve{E} = -j\omega\mu\bar{\bar{s}}\breve{H} \tag{7.53a, b}$$

where $\bar{\bar{s}}$ is the diagonal tensor defined by

$$\begin{aligned}\bar{\bar{s}} &= \begin{bmatrix} s_x^{-1} & 0 & 0 \\ 0 & s_x & 0 \\ 0 & 0 & s_x \end{bmatrix}\begin{bmatrix} s_y & 0 & 0 \\ 0 & s_y^{-1} & 0 \\ 0 & 0 & s_y \end{bmatrix}\begin{bmatrix} s_z & 0 & 0 \\ 0 & s_z & 0 \\ 0 & 0 & s_z^{-1} \end{bmatrix} \\ &= \begin{bmatrix} s_y s_z s_x^{-1} & 0 & 0 \\ 0 & s_x s_z s_y^{-1} & 0 \\ 0 & 0 & s_x s_y s_z^{-1} \end{bmatrix}\end{aligned} \tag{7.54}$$

Allowing for a nonunity real part κ, the multiplicative components of the diagonal elements of $\bar{\bar{s}}$ are given by

$$s_x = \kappa_x + \frac{\sigma_x}{j\omega\varepsilon} \quad ; \quad s_y = \kappa_y + \frac{\sigma_y}{j\omega\varepsilon} \quad ; \quad s_z = \kappa_z + \frac{\sigma_z}{j\omega\varepsilon} \tag{7.55a, b, c}$$

The general tensor in (7.54) is sufficient to describe the anisotropic PML in the entire FDTD volume. It turns out that it is sufficient to express s_x, s_y, and s_z, or more appropriately, κ_x, κ_y, κ_z, σ_x, σ_y, and σ_z as one-dimensional functions. That is:

$$\kappa_x(x) = \begin{cases} \kappa_x'(x) & x \le x_{\min},\ x \ge x_{\max} \\ 1 & x_{\min} < x < x_{\max} \end{cases}, \quad \sigma_x(x) = \begin{cases} \sigma_x'(x) & x \le x_{\min},\ x \ge x_{\max} \\ 0 & x_{\min} < x < x_{\max} \end{cases} \tag{7.56a}$$

$$\kappa_y(y) = \begin{cases} \kappa_y'(y) & y \le y_{\min},\ y \ge y_{\max} \\ 1 & y_{\min} < y < y_{\max} \end{cases}, \quad \sigma_y(y) = \begin{cases} \sigma_y'(y) & y \le y_{\min},\ y \ge y_{\max} \\ 0 & y_{\min} < y < y_{\max} \end{cases} \tag{7.56b}$$

$$\kappa_z(z) = \begin{cases} \kappa_z'(z) & z \le z_{\min},\ z \ge z_{\max} \\ 1 & z_{\min} < z < z_{\max} \end{cases}, \quad \sigma_z(z) = \begin{cases} \sigma_z'(z) & z \le z_{\min},\ z \ge z_{\max} \\ 0 & z_{\min} < z < z_{\max} \end{cases} \tag{7.56c}$$

Thus, in the interior region of the FDTD space, the tensor becomes the unit dyad, whereas in the PML regions, the tensor is properly expressed. Note that, as with the split-field PML, two tensor coefficients are nonunity in dihedral corner regions, and all three tensor coefficients are nonunity in trihedral corner regions.

The generalized constitutive tensor defined in (7.54) is no longer uniaxial by strict definition, but rather is anisotropic. However, the anisotropic PML is still referenced as uniaxial, since it is uniaxial in the nonoverlapping PML regions.

7.5.4 Inhomogeneous Media

Often, the PML is used to terminate an inhomogeneous material region that is assumed to extend infinitely. One example is a printed circuit constructed on an infinite substrate backed by a ground plane. A second example is an optical fiber that is assumed to be infinitely long in the axial direction. In such cases, the inhomogeneous material region extends through the PML to the outer boundary of the FDTD lattice. It can be shown that the PML is indeed perfectly matched to such a medium [43]. However, care must be taken to properly choose the PML parameters to maintain a stable and accurate formulation.

Consider an x-normal UPML boundary. Let an inhomogeneous dielectric $\varepsilon(y, z)$, which is assumed to be piecewise constant in the transverse y- and z-directions, extend into the UPML. From fundamental electromagnetic theory, $D_y = \varepsilon E_y$ must be continuous across any y-normal boundary, and $D_z = \varepsilon E_z$ must be continuous across any z-normal boundary. Then, from Gauss' law for the UPML in (7.50), we see that s_x must be independent of y and z to avoid surface charge at the boundaries of the discontinuity.

In the previous discussions, the dielectric in the UPML was assumed to be homogeneous. For this case in (7.55a), $s_x = \kappa_x + \sigma_x / j\omega\varepsilon$ was chosen. However, if $\varepsilon = \varepsilon(y, z)$ and is piecewise constant, then s_x is also piecewise constant in the transverse directions. Thus, surface charge densities result at the material boundaries as predicted by Gauss' law in (7.50), due to the derivative of a discontinuous function. This nonphysical charge leads to an ill-posed formulation. To avoid this, s_x must be independent of y and z. This holds only if σ_x/ε is maintained constant. This can be done in a brute-force manner by modifying σ_x in the transverse direction such that $\sigma_x(y, z)/\varepsilon(y, z)$ is a constant. A much simpler approach is to normalize σ_x by the relative permittivity, rewriting (7.55a) as

$$s_x = \kappa_x + \sigma_x' / j\omega\varepsilon_0 \tag{7.57}$$

where ε_0 is the free-space permittivity. In this case, σ_x' is simply a constant in the transverse y- and z-directions, although it is still scaled along the normal x-direction. Now, Gauss' law is satisfied within the UPML, leading to a well-posed formulation. This also leads to a material-independent formulation of the UPML [24, 39, 43, 44 – 49].

Next, consider Berenger's split-field PML. To understand the constraints of this technique in an inhomogeneous medium, it is simpler to work with the stretched-coordinate representation of the equations. In stretched coordinates, Gauss' law is represented in (7.52). Here, it appears that there are no further constraints on s_x. However, conservation laws require that the charge continuity equation be derived from Ampere's law and Gauss' law. To this end, the divergence of Ampere's law in (7.29) is performed in the stretched coordinates, using (7.52). We see that the divergence of the curl of $\boldsymbol{H}$ is zero only if s_x is independent of the transverse coordinates y and z. This holds only if σ_x/ε is independent of y and z. Again, this can be easily managed by representing s_x by (7.57), thus leading to a material-independent PML.

In summary, an infinite inhomogeneous medium can be terminated by a split-field or stretched-coordinate PML, or by a UPML medium. All are perfectly matched to arbitrary electromagnetic waves impinging upon the PML boundary. The method is accurate and stable, provided that the PML parameters s_w (s_x, s_y, or s_z) are posed to be independent of the transverse directions. This can be readily accomplished by normalizing σ_w by the relative permittivity, and hence posing $s_w = \kappa_w + \sigma_w'/j\omega\varepsilon_0$, where σ_w' is constant in the transverse direction.

7.6 THEORETICAL PERFORMANCE OF THE PML

7.6.1 The Continuous Space

When used to truncate an FDTD lattice, the PML has a thickness d and is terminated by the outer boundary of the lattice. If the outer boundary is assumed to be a PEC wall, as shown in Fig. 7.1, finite power reflects back into the primary computation zone. For a wave impinging upon the PML at angle θ relative to the w-directed surface normal, this reflection can be computed using transmission line analysis, yielding

$$R(\theta) = e^{-2\sigma_w \eta d \cos\theta} \tag{7.58}$$

Here, η and σ_w are, respectively, the PML's characteristic wave impedance and its conductivity, referred to propagation in the w-direction. In the context of an FDTD simulation, $R(\theta)$ is referred to as the "reflection error," since it is a nonphysical reflection due to the PEC wall that backs the PML. We note that the reflection error is the same for the split-field and stretched-coordinate PML and the UPML, since all support the same wave equation. This error decreases exponentially with σ_w and d. However, the reflection error increases as $\exp(\cos\theta)$, reaching the worst case for $\theta = 90°$. At this grazing angle of incidence, $R = 1$ and the PML is completely ineffective. To be useful within an FDTD simulation, we want $R(\theta)$ to be as small as possible. Clearly, for a thin PML, we must have σ_w as large as possible in order to reduce $R(\theta)$ to acceptably small levels, especially for θ approaching 90°.

7.6.2 The Discrete Space

Grading of the PML Loss Parameters

Theoretically, reflectionless wave transmission can take place across a PML interface regardless of the local step-discontinuity in σ and σ^* presented to the continuous impinging electromagnetic field. However, in FDTD or any discrete representation of Maxwell's equations, numerical artifacts arise due to the finite spatial sampling. Consequently, implementing PML as a single step-discontinuity of σ and σ^* in the FDTD lattice can lead to significant spurious wave reflection at the PML surface.

To reduce this reflection error, Berenger proposed that the PML losses along the direction normal to the interface gradually rise from zero [2]. Assuming such a grading, the PML remains matched, as seen from the stretched-coordinate theory in Section 7.4. Pursuing this idea, we consider as an example an x-directed plane wave impinging at angle θ upon a PEC-backed PML slab of thickness d, with the front planar interface located in the $x = 0$ plane (chosen here for convenience). Assuming the graded PML conductivity profile $\sigma_x(x)$, we have from (7.26) and (7.22) or (7.49) the reflection factor

$$R(\theta) = e^{-2\eta\cos\theta\int_0^d \sigma_x(x)\,dx} \tag{7.59}$$

Polynomial Grading. Several profiles have been suggested for grading $\sigma_x(x)$ [and $\kappa_x(x)$ in the context of the UPML]. The most successful use a polynomial or geometric variation of the PML loss with depth x [3]. Polynomial grading is simply

$$\sigma_x(x) = (x/d)^m \sigma_{x,\max} \quad ; \qquad \kappa_x(x) = 1 + (\kappa_{x,\max} - 1)\cdot(x/d)^m \tag{7.60a, b}$$

This increases the value of the PML σ_x from 0 at $x = 0$, the surface of the PML, to $\sigma_{x,\max}$ at $x = d$, the PEC outer boundary. Similarly, for the UPML, κ_x increases from 1 at $x = 0$ to $\kappa_{x,\max}$ at $x = d$. Substituting (7.60) into (7.59) yields

$$R(\theta) = e^{-2\eta\,\sigma_{x,\max}\,d\cos\theta/(m+1)} \tag{7.61}$$

For a fixed d, polynomial grading provides two parameters: $\sigma_{x,\max}$ and m. A large m yields a $\sigma_x(x)$ distribution that is relatively flat near the PML surface. However, deeper within the PML, σ_x increases more rapidly than for small m. In this region, the field amplitudes are substantially decayed, and reflections due to the discretization error contribute less. Typically, $3 \le m \le 4$ has been found to be nearly optimal for many FDTD simulations [3, 6, 14, 39].

For polynomial grading, the PML parameters can be readily determined for a given error estimate. For example, let m, d, and the desired reflection error $R(0)$ be known. Then, from (7.61), $\sigma_{x,\max}$ is computed as

$$\sigma_{x,\max} = -\frac{(m+1)\ln[R(0)]}{2\eta d} \tag{7.62}$$

Geometric Grading. The PML loss profile is defined for this case as [6]

$$\sigma_x(x) = \left(g^{1/\Delta}\right)^x \sigma_{x,0} \quad ; \qquad \kappa_x(x) = \left[(\kappa_{\max})^{1/d}\, g^{1/\Delta}\right]^x \tag{7.63a, b}$$

where $\sigma_{x,0}$ is the PML conductivity at its surface, g is the scaling factor, and Δ is the FDTD space increment. Here, the PML conductivity increases from $\sigma_{x,0}$ at its surface to $g^{d/\Delta}\sigma_{x,0}$ at the PEC outer boundary. Substituting (7.63) into (7.59) results in

$$R(\theta) = e^{-2\eta\,\sigma_{x,0}\,\Delta\,(g^{d/\Delta}-1)\cos\theta/\ln g} \tag{7.64}$$

For a fixed d, geometric grading provides two parameters: g and $\sigma_{x,0}$. $\sigma_{x,0}$ must be small to minimize the initial discretization error, and the metric $g > 1$ governs the rate of increase of the conductivity within the PML. Large values of g flatten the conductivity profile near $x = 0$, and steepen it deeper into the PML. Usually, g, d, and $R(0)$ are predetermined. This yields

$$\sigma_{x,0} = -\frac{\ln[R(0)]\ln(g)}{2\eta\Delta\left(g^{d/\Delta}-1\right)} \tag{7.65}$$

Typically, $2 \le g \le 3$ has been found to be nearly optimal for many FDTD simulations.

Discretization Error

The design of an effective PML requires balancing the theoretical reflection error, $R(\theta)$, and the numerical discretization error. For example, (7.62) provides $\sigma_{x,\max}$ for a polynomial-graded conductivity, given a predetermined $R(0)$ and m. If $\sigma_{x,\max}$ is small, the primary reflection from the PML is due to its PEC backing. Equation (7.59) provides a fairly accurate approximation of the reflection error for this case. However, we normally choose $\sigma_{x,\max}$ to be as large as possible to minimize $R(\theta)$. Unfortunately, if $\sigma_{x,\max}$ is too large, the discretization error due to the FDTD approximation dominates, and the actual reflection error is potentially orders of magnitude higher than what (7.59) predicts. Consequently, there is an optimal choice for $\sigma_{x,\max}$ that balances reflection from the PEC outer boundary and discretization error.

It was postulated by Berenger [3, 6] that the largest discretization error manifested as reflection occurs at $x = 0$, the PML surface. Any wave energy that penetrates further into the PML and then is reflected undergoes attenuation both before and after its point of reflection, and typically is not as large a contribution. Thus, it is desirable to minimize the discontinuity at $x = 0$. As discussed earlier, one way to achieve this is by flattening the PML loss profile near $x = 0$. However, if the subsequent rise of loss with depth is too rapid, reflections from deeper within the PML can dominate.

Through an extensive series of numerical experiments, it was demonstrated in [39, 43] that, for a broad range of applications, an optimal choice for a 10-cell-thick, polynomial-graded PML is $R(0) \equiv e^{-16}$. For a 5-cell-thick PML, $R(0) \equiv e^{-8}$ is optimal. From (7.62), this leads to an optimal choice for $\sigma_{x,\max}$ for polynomial grading:

$$\sigma_{x,\text{opt}} \approx -\frac{(m+1)\cdot(-16)}{(2\eta)\cdot(10\Delta)} = \frac{0.8(m+1)}{\eta\Delta} \tag{7.66}$$

This expression assumes a homogeneous medium. It can be further generalized to inhomogeneous, dispersive, or nonlinear media by using the stretched-coordinate coefficient in (7.57). Following the same principle, $\sigma_{x,\text{opt}}$ is generalized to [39, 43]

$$\sigma_{x,\text{opt}} \approx \frac{0.8(m+1)}{\eta_0 \Delta \sqrt{\varepsilon_{r,\text{eff}}\,\mu_{r,\text{eff}}}} \tag{7.67}$$

where η_0 is the free-space wave impedance, Δ is the lattice-cell dimension along the normal stencil, and $\varepsilon_{r,\text{eff}}$ and $\mu_{r,\text{eff}}$ are constants that represent the effective relative permittivity and permeability, respectively. For inhomogeneous or dispersive materials, $\varepsilon_{r,\text{eff}}$ and $\mu_{r,\text{eff}}$ can be chosen to be a mean value based on the material parameters, or based on a physical parameter such as the wavenumber of the fundamental mode expected to impinge on the PML boundary. This choice for $\sigma_{x,\text{opt}}$ has proven to be quite robust for many applications, as will be illustrated in Section 7.11.

7.7 COMPLEX FREQUENCY-SHIFTED TENSOR

7.7.1 Introduction

The performance of the PML in the previous sections considered only the absorption of propagating waves with near-normal angles of incidence. However, we note that, when implementing the PML within the discrete FDTD domain, large spurious reflections can occur for cases where the numerical waves impinging on the PML are near grazing incidence. This arises when the PML terminates a highly elongated domain. In addition, large PML reflections can occur when the impinging wave spectrum is highly evanescent. This arises when the PML boundary is placed very close to a field singularity such as that existing at a metallic edge or corner, or in the near-field of an antenna.

Consider the specular reflection of a normally incident TE_z-polarized wave from the front PML interface. From (7.21), this can be written directly as:

$$\Gamma_0 = \frac{1 - \sqrt{s_x^*(0)/s_x(0)}}{1 + \sqrt{s_x^*(0)/s_x(0)}} \tag{7.68}$$

where we have intentionally expressed this in terms of the electric and magnetic tensor coefficients s_x and s_x^*, respectively, and s_x and s_x^* are expressed via (7.57). The stretched coordinates are assumed to be spatially scaled using (7.60). Then, at the front PML interface, $s_x(0) = 1$. In the discrete space, the magnetic tensor and electric tensor parameters are staggered by one-half cell. Due to this discretization, the magnetic tensor coefficient is nonzero at locations one-half cell from the interface. Thus,

$$\Gamma_0 = \frac{1 - \sqrt{s_x(\Delta_x/2)}}{1 + \sqrt{s_x(\Delta_x/2)}} \tag{7.69}$$

where Δ_x is the FDTD lattice-cell dimension. Defining s_x from (7.55a), this is rewritten as

$$\Gamma_0 = \frac{\sqrt{1 + \sigma_x(\Delta_x/2)/j\omega\varepsilon_0} - 1}{\sqrt{1 + \sigma_x(\Delta_x/2)/j\omega\varepsilon_0} + 1} = \frac{\sqrt{\omega - j\sigma_x(\Delta_x/2)/\varepsilon_0} - \sqrt{\omega}}{\sqrt{\omega - j\sigma_x(\Delta_x/2)/\varepsilon_0} + \sqrt{\omega}} \tag{7.70}$$

where it is assumed for simplicity that $\kappa_x(\Delta_x/2) = 1$. In the frequency range $\omega >> \sigma_x(\Delta_x/2)/\varepsilon_0$, $\Gamma_0 \approx 0$. At low frequencies, where $\omega << \sigma_x(\Delta_x/2)/\varepsilon_0$, $\Gamma_0 \approx 1$. Consequently, slowly varying fields or waves with a long time-interaction with the PML interface can suffer from large reflections [4, 6].

A break frequency between the high-reflection, low-frequency regime and the low-reflection, high-frequency regime can be defined from (7.70) as [3]:

$$f_c = \frac{\sigma_x(\Delta_x/2)}{2\pi\varepsilon_0} = 1.8 \times 10^{10}\,\sigma_x(\Delta_x/2) \tag{7.71}$$

Thus, when $f > f_c$, most of the impinging field is transmitted into the PML, and when $f < f_c$, most of the impinging field is reflected. In the time domain, this allows small reflections from the PML interface up to a time of approximately

$$t_c = \frac{1}{f_c} = \frac{1}{1.8 \times 10^{10}\,\sigma_x(\Delta_x/2)} \tag{7.72}$$

Relatively large reflections can occur beyond this time, as verified in numerical experiments by Berenger [3]. An important consequence is that PML can perform poorly for highly oblique incident waves and slowly moving evanescent waves that inherently have long-term interactions with the PML boundary.

These undesirable late-time reflections from the PML interface can be mitigated by either decreasing $\sigma_x(\Delta_x/2)$ or increasing the size of the problem domain. However, such options either increase $R(\theta)$ or increase the computational burden. Thus, again, discretization error and theoretical reflection error must be balanced.

7.7.2 Strategy to Reduce Late-Time (Low-Frequency) Reflections

One reason for the large reflection coefficient at low frequencies is that the tensor coefficient s_x has a pole (or is singular) at zero frequency. That is, $s_x \to -j\infty$ as $\omega \to 0$. However, this problem can be circumvented by shifting the pole off the origin into the upper-half complex plane. To this end, a choice for the tensor coefficient s_x that has proven beneficial is [36, 50]:

$$s_x = \kappa_x + \frac{\sigma_x}{a_x + j\omega\varepsilon_0} \tag{7.73}$$

This has been referred to as the *complex frequency-shifted* (CFS) tensor coefficient [36, 37, 43]. Assuming $\kappa_x \geq 1$, $\sigma_x \geq 0$, and $a_x \geq 0$ and real, this choice for s_x is both causal and stable [51].

The reflection coefficient can again be estimated for the CFS-PML tensor using (7.68). In this case, it is assumed that s_x is polynomial scaled and $\sigma_x(0) = 0$, $\kappa_x(0) = 1$, $a_x(0) \neq 0$, and $s_x(0) = 1$. For the magnetic tensor coefficient, s_x^* is nonzero at $\Delta_x/2$. The reflection coefficient is then approximated as

$$\Gamma_{\text{CFS}} = \frac{\sqrt{1 + \dfrac{\sigma_x(\Delta_x/2)}{a_x(\Delta_x/2) + j\omega\varepsilon_0}} - 1}{\sqrt{1 + \dfrac{\sigma_x(\Delta_x/2)}{a_x(\Delta_x/2) + j\omega\varepsilon_0}} + 1} \tag{7.74}$$

Again, for simplicity it is assumed that $\kappa_x(\Delta_x/2) = 1$. In the high-frequency limit that $\omega >> a_x(\Delta_x/2)/\varepsilon_0$ and $\omega >> \sigma_x(\Delta_x/2)/\varepsilon_0$, we see that $\Gamma_{\text{CFS}} \to 0$. On the other hand, in the low-frequency limit that $\omega << a_x(\Delta_x/2)/\varepsilon_0$, we have

$$\Gamma_{\text{CFS}}\big|_{\text{low-}f\text{ limit}} \approx \frac{\sqrt{1 + \dfrac{\sigma_x(\Delta_x/2)}{a_x(\Delta_x/2)}} - 1}{\sqrt{1 + \dfrac{\sigma_x(\Delta_x/2)}{a_x(\Delta_x/2)}} + 1} \tag{7.75}$$

Assuming that $\sigma_x(\Delta_x/2)/a_x(\Delta_x/2) << 1$, then

$$\Gamma_{\text{CFS}}\big|_{\text{low-}f\text{ limit}} \approx \frac{\sigma_x(\Delta_x/2)}{4a_x(\Delta_x/2)} \tag{7.76}$$

Comparing this with (7.70), we see immediately that the CFS tensor coefficient can significantly reduce the specular reflection error from the front PML interface at low frequencies. (See [52] for a more detailed discussion of the reflection phenomena associated with the CFS PML.)

To fully understand the behavior of the CFS PML, we also need to explore the attenuation of a wave propagating within this medium. We assume that this wave has a complex propagation constant $\gamma_x^{inc} = \alpha_x^{inc} + j\beta_x^{inc}$, where α_x^{inc} represents attenuation. Following the derivation in Section 7.5.1, the wave exhibits an exponential dependence along the normal axis of the form $\exp(-\gamma_x^{inc} s_x x)$. Assuming the CFS tensor coefficient in (7.73), this is expanded as

$$e^{-\left(\alpha_x^{inc} + j\beta_x^{inc}\right)\left(\kappa_x + \frac{\sigma_x}{a_x + j\omega\varepsilon_0}\right)x} = e^{-\left(\alpha_x^{inc}\kappa_x + j\beta_x^{inc}\kappa_x + \alpha_x^{inc}\frac{\sigma_x}{a_x + j\omega\varepsilon_0} + j\beta_x^{inc}\frac{\sigma_x}{a_x + j\omega\varepsilon_0}\right)x} \tag{7.77}$$

A number of observations can be made from (7.77). The first is that κ_x amplifies the physical attenuation of the incident wave α_x^{inc} in the CFS PML. Next, in the limit that $\omega >> a_x/\varepsilon_0$, the CFS PML absorbs to the same degree as the standard PML. However, if $\omega << a_x/\varepsilon_0$, the CFS PML conductivity σ_x offers no additional attenuation. In fact, if the wave is purely propagating (i.e., $\alpha_x^{inc} = 0$), the CFS PML provides little if any attenuation at low frequencies. The break frequency for the CFS PML is thus quantified as

$$f_a = a_x / 2\pi\varepsilon_0 \tag{7.78}$$

such that if $f << f_a$, little absorption of a propagating wave is administered by the PML. Conversely, high absorption is realized if $f >> f_a$.

As a consequence, while the CFS PML significantly reduces the reflection error at low frequencies, propagating waves within the PML can be appreciably less attenuated. Fortunately, this dilemma can be resolved with a spatial scaling of a_x. As predicted by (7.74) and (7.76), a_x should be relatively large near the front PML interface, since the reflection coefficient is minimized when $\sigma_x(\Delta_x/2)/a_x(\Delta_x/2) << 1$. Inside the PML, a_x should be reduced so that the low-frequency spectrum of the waves propagating through the PML can be appropriately attenuated. Therefore, $a_x(x)$ should have a *maximum value* at the front PML interface, and decay to zero within the PML. To this end, a recommended polynomial scaling function for a_x is given by

$$a_x(x) = \begin{cases} a_{x,\max}\left(\dfrac{d-x}{d}\right)^{m_a}, & 0 \le x \le d \\ 0, & \text{otherwise} \end{cases} \tag{7.79}$$

where the front PML boundary is assumed to be located at $x = 0$, m_a is the scaling order, and d is the thickness of the PML.

7.8 EFFICIENT IMPLEMENTATION OF UPML IN FDTD

This section discusses mapping the UPML presented in Section 7.5 into the discrete FDTD space. The FDTD approximation is derived from the time-harmonic Maxwell's curl equations and tensor coefficients within the UPML medium, as defined in (7.53) to (7.55). Since the effective medium presented by the UPML is both dispersive and anisotropic, new techniques must be introduced to derive the discrete field-update equations. In this section, an *auxiliary differential equation* (ADE) approach is used, leading to an efficient formulation.

7.8.1 Derivation of the Finite-Difference Expressions

Starting with (7.53a) and (7.54), Ampere's law in a matched UPML is expressed as

$$\begin{bmatrix} \frac{\partial \breve{H}_z}{\partial y} - \frac{\partial \breve{H}_y}{\partial z} \\ \frac{\partial \breve{H}_x}{\partial z} - \frac{\partial \breve{H}_z}{\partial x} \\ \frac{\partial \breve{H}_y}{\partial x} - \frac{\partial \breve{H}_x}{\partial y} \end{bmatrix} = j\omega\varepsilon \begin{bmatrix} \frac{s_y s_z}{s_x} & 0 & 0 \\ 0 & \frac{s_x s_z}{s_y} & 0 \\ 0 & 0 & \frac{s_x s_y}{s_z} \end{bmatrix} \begin{bmatrix} \breve{E}_x \\ \breve{E}_y \\ \breve{E}_z \end{bmatrix} \tag{7.80}$$

where we repeat for convenience (7.55), the assumed tensor coefficients for general media:

$$s_x = \kappa_x + \frac{\sigma_x}{j\omega\varepsilon_0} \; ; \quad s_y = \kappa_y + \frac{\sigma_y}{j\omega\varepsilon_0} \; ; \quad s_z = \kappa_z + \frac{\sigma_z}{j\omega\varepsilon_0} \tag{7.81a, b, c}$$

Directly inserting (7.81) into (7.80) and then transforming into the time domain would lead to a convolution between the tensor coefficients and the E-field. This is not advisable, since implementing this convolution would be computationally intensive in this form. A more efficient approach is to define the proper constitutive relationship to decouple the frequency-dependent terms [24, 39]. Specifically, let

$$\breve{D}_x = \varepsilon \frac{s_z}{s_x} \breve{E}_x \; ; \qquad \breve{D}_y = \varepsilon \frac{s_x}{s_y} \breve{E}_y \; ; \qquad \breve{D}_z = \varepsilon \frac{s_y}{s_z} \breve{E}_z \tag{7.82a, b, c}$$

Then, (7.80) is rewritten as

$$\begin{bmatrix} \frac{\partial \breve{H}_z}{\partial y} - \frac{\partial \breve{H}_y}{\partial z} \\ \frac{\partial \breve{H}_x}{\partial z} - \frac{\partial \breve{H}_z}{\partial x} \\ \frac{\partial \breve{H}_y}{\partial x} - \frac{\partial \breve{H}_x}{\partial y} \end{bmatrix} = j\omega \begin{bmatrix} s_y & 0 & 0 \\ 0 & s_z & 0 \\ 0 & 0 & s_x \end{bmatrix} \begin{bmatrix} \breve{D}_x \\ \breve{D}_y \\ \breve{D}_z \end{bmatrix} \tag{7.83}$$

Now, we substitute s_x, s_y, and s_z from (7.81) into (7.83), and then apply the inverse Fourier transform using the identity $j\omega f(\omega) \rightarrow (\partial/\partial t) f(t)$. This yields an equivalent system of time-domain differential equations for (7.83):

$$\begin{bmatrix} \dfrac{\partial H_z}{\partial y} - \dfrac{\partial H_y}{\partial z} \\ \dfrac{\partial H_x}{\partial z} - \dfrac{\partial H_z}{\partial x} \\ \dfrac{\partial H_y}{\partial x} - \dfrac{\partial H_x}{\partial y} \end{bmatrix} = \frac{\partial}{\partial t} \begin{bmatrix} \kappa_y & 0 & 0 \\ 0 & \kappa_z & 0 \\ 0 & 0 & \kappa_x \end{bmatrix} \begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} + \frac{1}{\varepsilon_0} \begin{bmatrix} \sigma_y & 0 & 0 \\ 0 & \sigma_z & 0 \\ 0 & 0 & \sigma_x \end{bmatrix} \begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} \tag{7.84}$$

The system of equations in (7.84) can be discretized on the standard Yee lattice. Normal leapfrogging in time as described in Chapter 3 is suitable, wherein the loss terms are averaged in time according to the semi-implicit scheme. This leads to explicit time-stepping expressions for D_x, D_y, and D_z. For example, the D_x update is given by

$$D_x\Big|_{i+1/2,j,k}^{n+1} = \left(\frac{2\varepsilon_0\kappa_y - \sigma_y\Delta t}{2\varepsilon_0\kappa_y + \sigma_y\Delta t}\right) D_x\Big|_{i+1/2,j,k}^{n} + \left(\frac{2\varepsilon_0\Delta t}{2\varepsilon_0\kappa_y + \sigma_y\Delta t}\right) \cdot \left(\frac{H_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} - H_z\Big|_{i+1/2,j-1/2,k}^{n+1/2}}{\Delta y} - \frac{H_y\Big|_{i+1/2,j,k+1/2}^{n+1/2} - H_y\Big|_{i+1/2,j,k-1/2}^{n+1/2}}{\Delta z}\right) \tag{7.85}$$

Next, we focus on (7.82a–c). For example, we consider (7.82a). After multiplying both sides by s_x, and substituting for s_x and s_z from (7.81a, c), we have

$$\left(\kappa_x + \frac{\sigma_x}{j\omega\varepsilon_0}\right)\breve{D}_x = \varepsilon\left(\kappa_z + \frac{\sigma_z}{j\omega\varepsilon_0}\right)\breve{E}_x \tag{7.86}$$

Multiplying both sides by $j\omega$ and transforming into the time domain leads to

$$\frac{\partial}{\partial t}\left(\kappa_x D_x\right) + \frac{\sigma_x}{\varepsilon_0} D_x = \varepsilon\left[\frac{\partial}{\partial t}\left(\kappa_z E_x\right) + \frac{\sigma_z}{\varepsilon_0} E_x\right] \tag{7.87a}$$

Similarly, from (7.82b, c), we obtain

$$\frac{\partial}{\partial t}\left(\kappa_y D_y\right) + \frac{\sigma_y}{\varepsilon_0} D_y = \varepsilon\left[\frac{\partial}{\partial t}\left(\kappa_x E_y\right) + \frac{\sigma_x}{\varepsilon_0} E_y\right] \tag{7.87b}$$

$$\frac{\partial}{\partial t}\left(\kappa_z D_z\right) + \frac{\sigma_z}{\varepsilon_0} D_z = \varepsilon\left[\frac{\partial}{\partial t}\left(\kappa_y E_z\right) + \frac{\sigma_y}{\varepsilon_0} E_z\right] \tag{7.87c}$$

The time derivatives in (7.87) are discretized using standard Yee leapfrogging and time-averaging the loss terms. This yields explicit time-stepping expressions for E_x, E_y, and E_z. For example, the E_x update is given by

$$E_x\Big|_{i+1/2,j,k}^{n+1} = \left(\frac{2\varepsilon_0\kappa_z - \sigma_z\Delta t}{2\varepsilon_0\kappa_z + \sigma_z\Delta t}\right) E_x\Big|_{i+1/2,j,k}^{n} + \left[\frac{1}{(2\varepsilon_0\kappa_z + \sigma_z\Delta t)\,\varepsilon}\right] \cdot \left[(2\varepsilon_0\kappa_x + \sigma_x\Delta t)\, D_x\Big|_{i+1/2,j,k}^{n+1} - (2\varepsilon_0\kappa_x - \sigma_x\Delta t)\, D_x\Big|_{i+1/2,j,k}^{n}\right] \tag{7.88}$$

Overall, updating the components of $\boldsymbol{E}$ in the UPML requires two steps in sequence: (1) obtaining the new values of the components of $\boldsymbol{D}$ according to (7.85), and (2) using these new $\boldsymbol{D}$ components to obtain new values of the $\boldsymbol{E}$ components according to (7.88).

A similar two-step procedure is required to update the components of $\boldsymbol{H}$ in the UPML. Starting with Faraday's law in (7.53b) and (7.54), the first step involves developing the updates for the components of $\boldsymbol{B}$. For example, the update for B_x is given by

$$B_x\Big|_{i,j+1/2,k+1/2}^{n+3/2} = \left(\frac{2\varepsilon_0\kappa_y - \sigma_y\Delta t}{2\varepsilon_0\kappa_y + \sigma_y\Delta t}\right) B_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} - \left(\frac{2\varepsilon_0\Delta t}{2\varepsilon_0\kappa_y + \sigma_y\Delta t}\right) \cdot \left(\frac{E_z\Big|_{i,j+1,k+1/2}^{n+1} - E_z\Big|_{i,j,k+1/2}^{n+1}}{\Delta y} - \frac{E_y\Big|_{i,j+1/2,k+1}^{n+1} - E_y\Big|_{i,j+1/2,k}^{n+1}}{\Delta z}\right) \tag{7.89}$$

The second step involves updating the $\boldsymbol{H}$ components in the UPML using the values of the $\boldsymbol{B}$ components just obtained with (7.89) and analogous B-field updates:

$$H_x\Big|_{i,j+1/2,k+1/2}^{n+3/2} = \left(\frac{2\varepsilon_0\kappa_z - \sigma_z\Delta t}{2\varepsilon_0\kappa_z + \sigma_z\Delta t}\right) H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} + \left[\frac{1}{(2\varepsilon_0\kappa_z + \sigma_z\Delta t)\,\mu}\right] \cdot \left[(2\varepsilon_0\kappa_x + \sigma_x\Delta t)\, B_x\Big|_{i,j+1/2,k+1/2}^{n+3/2} - (2\varepsilon_0\kappa_x - \sigma_x\Delta t)\, B_x\Big|_{i,j+1/2,k+1/2}^{n+1/2}\right] \tag{7.90}$$

Here, the dual constitutive relation $\breve{B}_x = \mu(s_z/s_x)\breve{H}_x$ is chosen. Similar expressions can be derived for the remaining two H-field components.

It can be shown that numerical stability within the Courant limit is obtained when all six field components in the UPML are updated based on the two-step procedure outlined in (7.85), (7.88), (7.89), and (7.90) [13]. Further, it can be shown that the discrete fields satisfy Gauss' law, and that the UPML is well-posed [52, 53].

7.8.2 Computer Implementation of the UPML

Each $\boldsymbol{E}$ and $\boldsymbol{H}$ component within the UPML is computed using an explicit two-step time-marching scheme as derived in (7.85) and (7.88) for E_x, and in (7.89) and (7.90) for H_x. Based on these updates, the UPML is easily and efficiently implemented within the framework of existing FDTD codes. We now illustrate this in FORTRAN using the time-stepping expressions for E_x given in (7.85) and (7.88). First, at the beginning of the computer program, we precompute six coefficient arrays to be used in the field updates:

$$\texttt{C1(j)} = \frac{2\varepsilon_0\kappa_y(j) - \sigma_y(j)\Delta t}{2\varepsilon_0\kappa_y(j) + \sigma_y(j)\Delta t} \tag{7.91a}$$

$$\texttt{C2(j)} = \frac{2\varepsilon_0\Delta t}{2\varepsilon_0\kappa_y(j) + \sigma_y(j)\Delta t} \tag{7.91b}$$

$$\texttt{C3(k)} = \frac{2\varepsilon_0\kappa_z(k) - \sigma_z(k)\Delta t}{2\varepsilon_0\kappa_z(k) + \sigma_z(k)\Delta t} \tag{7.91c}$$

$$\texttt{C4(k)} = \frac{1}{2\varepsilon_0\kappa_z(k) + \sigma_z(k)\Delta t} \tag{7.91d}$$

$$\texttt{C5(i)} = 2\varepsilon_0\kappa_x(i) + \sigma_x(i)\Delta t \tag{7.91e}$$

$$\texttt{C6(i)} = 2\varepsilon_0\kappa_x(i) - \sigma_x(i)\Delta t \tag{7.91f}$$

Defining the field-updating coefficients in this manner permits a unified treatment of both the lossless interior working volume and the UPML slabs. In effect, UPML is assumed to fill the entire FDTD space lattice. We set $\sigma_w = 0$ and $\kappa_w = 1$ in the working volume to model free space. However, in the UPML slabs, σ_w and κ_w are assumed to have the polynomial-graded profile given in (7.60), or the geometric-graded profile given in (7.63), along the normal axes of the UPML slabs. As a result, the coefficients in (7.91) vary in only one dimension.

When defining the coefficient arrays (7.91a–f), it is critical to assign the proper value to the UPML loss parameters. To this end, σ_w and κ_w are computed at a physical coordinate using (7.60a, b) or (7.63a, b). The appropriate choice of the physical coordinate is at the edge center of the discrete field $E_x(i, j, k)$, which is $[(i + 1/2)\Delta x, j\Delta y, k\Delta z]$. Thus, in (7.91e, f), $\sigma_x(i)$ and $\kappa_x(i)$ are computed at the physical coordinate $(i + 1/2)\Delta x$. Similarly, in (7.91a, b), $\sigma_y(j)$ and $\kappa_y(j)$ are computed at the physical coordinate $j\Delta y$; and in (7.91c, d), $\sigma_z(k)$ and $\kappa_z(k)$ are computed at the physical coordinate $k\Delta z$. This is similarly done for the updates of E_y and E_z. The UPML loss parameters for the H-fields are chosen at the lattice face centers. For example, the physical coordinate of the discrete field $H_x(i, j, k)$ is $[i\Delta x, (j + 1/2)\Delta y, (k + 1/2)\Delta z]$. Thus, for the update of $H_x(i, j, k)$, $\sigma_x(i)$ and $\kappa_x(i)$ are computed at the physical coordinate $i\Delta x$; $\sigma_y(j)$ and $\kappa_y(j)$ are computed at the physical coordinate $(j + 1/2)\Delta y$; and $\sigma_z(k)$ and $\kappa_z(k)$ are computed at the physical coordinate $(k + 1/2)\Delta z$. This is similarly done for the updates of H_y and H_z.

Given the above "all-UPML" strategy, and assuming that the physical medium is a lossless and potentially inhomogeneous dielectric, then the FORTRAN program segment that executes the time-stepping of E_x *everywhere* in the FDTD space lattice can be written as a simple triply-nested loop:

```
do 10 k=2,nz-1
  do 10 j=2,ny-1
    do 10 i=1,nx-1
      dstore = dx(i,j,k)
      dx(i,j,k) = C1(j)*dx(i,j,k)
                + C2(j)*( (hz(i,j,k) - hz(i,j-1,k)) / deltay -
                          (hy(i,j,k) - hy(i,j,k-1)) / deltaz )
      ex(i,j,k) = C3(k)*ex(i,j,k)
              + C4(k)*( C5(i)*dx(i,j,k) - C6(i)*dstore)/eps(i,j,k)
10 continue
```

(7.92)

Assuming UPML throughout the entire FDTD lattice has the limitation that the flux densities D_x and B_x must be stored everywhere in the lattice. This essentially doubles the memory. However, this approach offers the significant advantage of simplifying the modification of existing FDTD codes. An alternative is to write a triply-nested loop for the interior fields, and separate loops for the various UPML slabs (segregating the corner regions). In this case, the auxiliary variables need to be stored only in the UPML region, leading to memory savings, at the cost of complexity in programming.

7.9 EFFICIENT IMPLEMENTATION OF CPML IN FDTD

The UPML presented in the previous section has been widely used by practitioners, and has been incorporated into a number of commercial FDTD software packages. Yet, as seen in Section 7.7, the use of the classical tensor coefficients defined in (7.55) can lead to large reflection errors for highly evanescent waves or for late-time, low-frequency interactions. It was found that using the CFS tensor coefficient with the proper scaling can overcome these limitations. An implementation of the CFS PML for FDTD was originally proposed by Gedney using an ADE formulation [43, Section 5.6.3]. However, a more efficient implementation was published by Roden and Gedney in [36] based on a recursive-convolution technique. This has since been referred to as the CPML formulation. The CPML is based on the stretched-coordinate form of the PML presented in Section 7.4. As will be shown in the remainder of this chapter, the CPML formulation is more accurate than the classical UPML, more efficient, and better suited for the application of domains with generalized materials.

7.9.1 Derivation of the Finite-Difference Expressions

Denoting $\mathcal{F}^{-1}$ the inverse Fourier transform operator, we first define the relation

$$\bar{s}_w = \mathcal{F}^{-1}\left[1/s_w(\omega)\right], \qquad w = x, y, z \tag{7.93}$$

where $\{s_w\}$ are the CFS stretched-coordinate tensor coefficients defined by (7.73). Then, from the stretched-coordinate form of the PML presented in Section 7.4, we can express the time-dependent Ampere's law as

$$\frac{\partial \boldsymbol{D}}{\partial t} = \hat{\boldsymbol{x}}\left(\bar{s}_y * \frac{\partial}{\partial y}H_z - \bar{s}_z * \frac{\partial}{\partial z}H_y\right) + \hat{\boldsymbol{y}}\left(\bar{s}_z * \frac{\partial}{\partial z}H_x - \bar{s}_x * \frac{\partial}{\partial x}H_z\right) + \hat{\boldsymbol{z}}\left(\bar{s}_x * \frac{\partial}{\partial x}H_y - \bar{s}_y * \frac{\partial}{\partial y}H_x\right) \tag{7.94}$$

In addition, Fourier transform theory yields

$$\bar{s}_w(t) = \mathcal{F}^{-1}\left(\frac{1}{\kappa_w + \dfrac{\sigma_w}{a_w + j\omega\varepsilon_0}}\right) = \frac{\delta(t)}{\kappa_w} - \frac{\sigma_w}{\varepsilon_0 \kappa_w^2} e^{-\left(\frac{\sigma_w}{\varepsilon_0\kappa_w} + \frac{a_w}{\varepsilon_0}\right)t} u(t) \equiv \frac{\delta(t)}{\kappa_w} + \zeta_w(t) \tag{7.95}$$

where $u(t)$ is the unit step function, and $\delta(t)$ is the unit impulse function. Now, from (7.95), Ampere's law in (7.94) can be expressed more concisely as

$$\frac{\partial \boldsymbol{D}}{\partial t} = \hat{\boldsymbol{x}}\left(\frac{1}{\kappa_y}\frac{\partial}{\partial y}H_z - \frac{1}{\kappa_z}\frac{\partial}{\partial z}H_y + \zeta_y * \frac{\partial}{\partial y}H_z - \zeta_z * \frac{\partial}{\partial z}H_y\right) + \hat{\boldsymbol{y}}\left(\frac{1}{\kappa_z}\frac{\partial}{\partial z}H_x - \frac{1}{\kappa_x}\frac{\partial}{\partial x}H_z + \zeta_z * \frac{\partial}{\partial z}H_x - \zeta_x * \frac{\partial}{\partial x}H_z\right) + \hat{\boldsymbol{z}}\left(\frac{1}{\kappa_x}\frac{\partial}{\partial x}H_y - \frac{1}{\kappa_y}\frac{\partial}{\partial y}H_x + \zeta_x * \frac{\partial}{\partial x}H_y - \zeta_y * \frac{\partial}{\partial y}H_x\right) \tag{7.96}$$

A dual expression can be derived for Faraday's law:

$$-\frac{\partial \boldsymbol{B}}{\partial t} = \hat{\boldsymbol{x}}\left(\frac{1}{\kappa_y}\frac{\partial}{\partial y}E_z - \frac{1}{\kappa_z}\frac{\partial}{\partial z}E_y + \zeta_y * \frac{\partial}{\partial y}E_z - \zeta_z * \frac{\partial}{\partial z}E_y\right) + \hat{\boldsymbol{y}}\left(\frac{1}{\kappa_z}\frac{\partial}{\partial z}E_x - \frac{1}{\kappa_x}\frac{\partial}{\partial x}E_z + \zeta_z * \frac{\partial}{\partial z}E_x - \zeta_x * \frac{\partial}{\partial x}E_z\right) + \hat{\boldsymbol{z}}\left(\frac{1}{\kappa_x}\frac{\partial}{\partial x}E_y - \frac{1}{\kappa_y}\frac{\partial}{\partial y}E_x + \zeta_x * \frac{\partial}{\partial x}E_y - \zeta_y * \frac{\partial}{\partial y}E_x\right) \tag{7.97}$$

The objective now is to derive explicit updates for the electric fields from (7.96) or the magnetic fields from (7.97) in the context of an FDTD discretization. The challenge is to efficiently implement discrete convolutions arising from $\zeta_w * (\partial H_v/\partial w)$. If implemented naively, this could be prohibitively expensive in computer resources. However, given the form of $\zeta_w(t)$, excellent efficiency can be realized by using the *recursive-convolution* (RC) technique reported by Luebbers and Hunsberger [54].

The first step is to derive $Z_w(m)$, the discrete impulse response of $\zeta_w(t)$:

$$\begin{aligned} Z_w(m) &= \int_{m\Delta t}^{(m+1)\Delta t} \zeta_w(\tau)\, d\tau \\ &= -\frac{\sigma_w}{\varepsilon_0 \kappa_w^2} \int_{m\Delta t}^{(m+1)\Delta t} e^{-\left(\frac{\sigma_w}{\varepsilon_0 \kappa_w} + \frac{a_w}{\varepsilon_0}\right)\tau} d\tau = c_w e^{-\left(\frac{\sigma_w}{\varepsilon_0 \kappa_w} + \frac{a_w}{\varepsilon_0}\right) m\Delta t} \end{aligned} \tag{7.98}$$

where

$$c_w = \frac{\sigma_w}{\sigma_w \kappa_w + \kappa_w^2 a_w} \left[e^{-\left(\frac{\sigma_w}{\varepsilon_0 \kappa_w} + \frac{a_w}{\varepsilon_0}\right)\Delta t} - 1 \right] \tag{7.99}$$

Now, the convolution can be implemented in a discrete manner by using a piecewise-constant approximation for $\zeta_w(t)$ and $\partial H_v/\partial w$:

$$\psi_{w,v}(n) = \zeta_w(t) * \frac{\partial}{\partial w} H_v(t)\bigg|_{t=n\Delta t} \approx \sum_{m=0}^{n-1} Z_w(m) \frac{\partial}{\partial w} H_v(n-m) \tag{7.100}$$

In this form, calculating the discrete convolution ψ at time $n\Delta t$ requires order(n) floating-point operations. If implemented directly in an FDTD simulation, the computational burden would be intolerable. Instead, a recursive relationship that efficiently computes the time advancement of ψ can be derived [53]. This is expressed as

$$\psi_{w,v}(n) = b_w \psi_{w,v}(n-1) + c_w \frac{\partial}{\partial w} H_v(n) \tag{7.101}$$

where

$$b_w = e^{-\left(\frac{\sigma_w}{\varepsilon_0 \kappa_w} + \frac{a_w}{\varepsilon_0}\right)\Delta t} \tag{7.102}$$

and c_w is given in (7.99). In this form, $\psi_{w,v}(n)$ can be advanced in time with great efficiency using a simple time-marching procedure.

By employing this recursive discrete convolution, an efficient FDTD formulation can be derived for (7.96) and (7.97). As an example, we shall consider the x-projection of Ampere's law in (7.96), assuming for generality that the host medium has a finite conductivity σ. Applying the proper constitutive relationship between the electric flux density $\boldsymbol{D}$ and the electric field intensity $\boldsymbol{E}$, we obtain:

$$\frac{\partial}{\partial t}(\varepsilon E_x) + \sigma E_x = \left(\frac{1}{\kappa_y}\frac{\partial}{\partial y}H_z - \frac{1}{\kappa_z}\frac{\partial}{\partial z}H_y\right) + \left(\zeta_y * \frac{\partial}{\partial y}H_z - \zeta_z * \frac{\partial}{\partial z}H_y\right) \tag{7.103}$$

Using the discrete time and space approximations of Chapter 3, and the RC approximation of the convolution integrals, (7.103) is expressed in discrete form as

$$\begin{aligned}
&\varepsilon_{i+1/2,j,k}\left(\frac{E_x\big|_{i+1/2,j,k}^{n+1/2} - E_x\big|_{i+1/2,j,k}^{n-1/2}}{\Delta t}\right) + \sigma_{i+1/2,j,k}\left(\frac{E_x\big|_{i+1/2,j,k}^{n+1/2} + E_x\big|_{i+1/2,j,k}^{n-1/2}}{2}\right) \\
&= \left(\frac{H_z\big|_{i+1/2,j+1/2,k}^{n} - H_z\big|_{i+1/2,j-1/2,k}^{n}}{\kappa_y \Delta y} - \frac{H_y\big|_{i+1/2,j,k+1/2}^{n} - H_y\big|_{i+1/2,j,k-1/2}^{n}}{\kappa_z \Delta z}\right) \\
&+ \psi_{E_{x,y}}\big|_{i+1/2,j,k}^{n} - \psi_{E_{x,z}}\big|_{i+1/2,j,k}^{n}
\end{aligned} \tag{7.104}$$

where $\psi_{E_{x,y}}$ and $\psi_{E_{x,z}}$ are discrete unknowns stored only in PML regions with y-normal and z-normal interface boundaries, respectively. These unknowns are updated as follows:

$$\psi_{E_{x,y}}\big|_{i+1/2,j,k}^{n} = b_{y_j}\psi_{E_{x,y}}\big|_{i+1/2,j,k}^{n-1} + c_{y_j}\left(\frac{H_z\big|_{i+1/2,j+1/2,k}^{n} - H_z\big|_{i+1/2,j-1/2,k}^{n}}{\Delta y}\right) \tag{7.105a}$$

$$\psi_{E_{x,z}}\big|_{i+1/2,j,k}^{n} = b_{z_k}\psi_{E_{x,z}}\big|_{i+1/2,j,k}^{n-1} + c_{z_k}\left(\frac{H_y\big|_{i+1/2,j,k+1/2}^{n} - H_y\big|_{i+1/2,j,k-1/2}^{n}}{\Delta z}\right) \tag{7.105b}$$

Note that the discrete coefficients b_{y_j} and c_{y_j} are nonzero only in PML regions with y-normal interface boundaries. These coefficients are computed using (7.102) and (7.99), respectively, with the scaled tensor parameters κ_y, σ_y, and a_y calculated at the discrete location $j\Delta y$, which is the y-coordinate of the discrete grid edge center $(i + 1/2, j, k)$. Coefficients b_{z_k} and c_{z_k} are nonzero only in PML regions with z-normal interface boundaries, and are calculated with the scaled tensor parameters κ_z, σ_z, and a_z computed via the appropriate spatial scaling at the discrete location $k\Delta z$, which is the z-coordinate of the discrete grid edge center $(i + 1/2, j, k)$.

The explicit update for E_x is derived directly from (7.104), and is given by

$$\begin{aligned}
E_x\big|_{i+1/2,j,k}^{n+1/2} &= C_a\big|_{i+1/2,j,k} E_x\big|_{i+1/2,j,k}^{n-1/2} \\
&+ C_b\big|_{i+1/2,j,k}\left(\begin{aligned}&\frac{H_z\big|_{i+1/2,j+1/2,k}^{n} - H_z\big|_{i+1/2,j-1/2,k}^{n}}{\kappa_{y_j}\Delta y} - \frac{H_y\big|_{i+1/2,j,k+1/2}^{n} - H_y\big|_{i+1/2,j,k-1/2}^{n}}{\kappa_{z_k}\Delta z} \\ &+ \psi_{E_{x,y}}\big|_{i+1/2,j,k}^{n} - \psi_{E_{x,z}}\big|_{i+1/2,j,k}^{n}\end{aligned}\right)
\end{aligned} \tag{7.106}$$

where

$$C_a\big|_{i+1/2,j,k} = \left(1 - \frac{\sigma_{i+1/2,j,k}\Delta t}{2\varepsilon_{i+1/2,j,k}}\right)\Bigg/\left(1 + \frac{\sigma_{i+1/2,j,k}\Delta t}{2\varepsilon_{i+1/2,j,k}}\right) \tag{7.107a}$$

$$C_b\big|_{i+1/2,j,k} = \left(\frac{\Delta t}{\varepsilon_{i+1/2,j,k}}\right)\Bigg/\left(1 + \frac{\sigma_{i+1/2,j,k}\Delta t}{2\varepsilon_{i+1/2,j,k}}\right) \tag{7.107b}$$

Similar expressions are derived for E_y and E_z by simply permuting the (i, j, k) indices and (x, y, z).

We can also derive dual equations for the $\boldsymbol{H}$ components. For example, from the x-projection of Faraday's law in (7.97), the explicit update for H_x is derived as

$$\begin{aligned} H_x\big|_{i,j+1/2,k+1/2}^{n+1} = {} & D_a\big|_{i,j+1/2,k+1/2} H_x\big|_{i,j+1/2,k+1/2}^{n} \\ & - D_b\big|_{i,j+1/2,k+1/2}\left(\begin{array}{l} \dfrac{E_z\big|_{i,j+1,k+1/2}^{n+1/2} - E_z\big|_{i,j,k+1/2}^{n+1/2}}{\kappa_{y_{j+1/2}}\Delta y} - \dfrac{E_y\big|_{i,j+1/2,k+1}^{n+1/2} - E_y\big|_{i,j+1/2,k}^{n+1/2}}{\kappa_{z_{k+1/2}}\Delta z} \\ + \psi_{H_{x,y}}\big|_{i,j+1/2,k+1/2}^{n+1/2} - \psi_{H_{x,z}}\big|_{i,j+1/2,k+1/2}^{n+1/2} \end{array}\right) \end{aligned} \tag{7.108}$$

where

$$D_a\big|_{i,j+1/2,k+1/2} = \left(1 - \frac{\dot{\sigma}_{i,j+1/2,k+1/2}\Delta t}{2\mu_{i,j+1/2,k+1/2}}\right)\Bigg/\left(1 + \frac{\dot{\sigma}_{i,j+1/2,k+1/2}\Delta t}{2\mu_{i,j+1/2,k+1/2}}\right) \tag{7.109a}$$

$$D_b\big|_{i,j+1/2,k+1/2} = \left(\frac{\Delta t}{\mu_{i,j+1/2,k+1/2}}\right)\Bigg/\left(1 + \frac{\dot{\sigma}_{i,j+1/2,k+1/2}\Delta t}{2\mu_{i,j+1/2,k+1/2}}\right) \tag{7.109b}$$

and where $\psi_{H_{x,y}}$ and $\psi_{H_{x,z}}$ are discrete unknowns stored, respectively, only in PML regions with y-normal and z-normal interface boundaries. These unknowns are updated as follows:

$$\psi_{H_{x,y}}\big|_{i,j+1/2,k+1/2}^{n+1/2} = b_{y_{j+1/2}}\psi_{H_{x,y}}\big|_{i,j+1/2,k+1/2}^{n-1/2} + c_{y_{j+1/2}}\left(\frac{E_z\big|_{i,j+1,k+1/2}^{n+1/2} - E_z\big|_{i,j,k+1/2}^{n+1/2}}{\Delta y}\right) \tag{7.110a}$$

$$\psi_{H_{x,z}}\big|_{i,j+1/2,k+1/2}^{n+1/2} = b_{z_{k+1/2}}\psi_{H_{x,z}}\big|_{i,j+1/2,k+1/2}^{n-1/2} + c_{z_{k+1/2}}\left(\frac{E_y\big|_{i,j+1/2,k+1}^{n+1/2} - E_y\big|_{i,j+1/2,k}^{n+1/2}}{\Delta z}\right) \tag{7.110b}$$

Note that the discrete coefficients $b_{y_{j+1/2}}$ and $c_{y_{j+1/2}}$ are nonzero only in PML regions with y-normal interface boundaries. These coefficients are computed using (7.102) and (7.99), respectively, with the scaled tensor parameters κ_y, σ_y, and a_y calculated at the discrete location $(j+1/2)\Delta y$, which is the y-coordinate of the discrete grid face center $(i, j+1/2, k+1/2)$. Coefficients $b_{z_{k+1/2}}$ and $c_{z_{k+1/2}}$ are nonzero only in PML regions with z-normal interface boundaries, and are calculated with the scaled tensor parameters κ_z, σ_z, and a_z computed at the discrete location $(k+1/2)\Delta z$, which is the z-coordinate of the discrete grid face center $(i, j+1/2, k+1/2)$. We see that similar updates are derived for H_y and H_z by simply permuting the (i, j, k) indices and (x, y, z).

Upon comparing the E_x time-stepping relation of (7.106) to (3.29a) of Chapter 3, and comparing the H_x time-stepping relation of (7.108) to (3.30a), we observe that the CPML field updates are identical to those of the standard FDTD algorithm, with the exception of the convolutional terms and the scaling parameters. The scaling parameters κ_y and κ_z are identified as real stretching coordinates since they physically lead to an effective scaling of the mesh in the PML region, much like the nonuniform grid presented in Chapter 11, Section 11.2. We also observe that the convolutional terms act as induced time-dependent polarization currents whose radiation cancels the wave propagating through the PML, rather than radiating externally.

Finally, we note that a more accurate approximation of the time convolution based on the piecewise-linear recursive-convolution (PLRC) technique (reviewed in Chapter 9, Section 9.3) could be employed. Then, the update of $\psi_{w,v}(n)$ would take the form

$$\psi_{w,v}^{\text{PLRC}}(n) = \chi_w^{\text{PLRC}} \psi_{w,v}^{\text{PLRC}}(n-1) + \xi_w^{\text{PLRC}} \frac{\partial H_v(n)}{\partial w} + \zeta_w^{\text{PLRC}} \frac{\partial H_v(n-1)}{\partial w} \tag{7.111}$$

where χ_w^{PLRC}, ξ_w^{PLRC}, and ζ_w^{PLRC} are coefficients to be determined. PLRC update (7.111) requires one additional time history of $\partial H_v/\partial w$ relative to RC update (7.101). Implementing this would require either updating $\psi_{w,v}(n)$ in the H-field updating (thereby complicating the algorithm), or storing the value of $\partial H_v/\partial w$ at the previous time-step (thereby increasing the required memory). In general, experience to date indicates that the RC approximation provides sufficient accuracy, and leads to a very simple implementation of the CPML algorithm, as shown in the next section.

7.9.2 Computer Implementation of the CPML

The computer implementation of the CPML for FDTD is very straightforward. Electric fields can be updated throughout the entire FDTD domain using (7.106) and (7.107), much in the same way as the classical FDTD update scheme. Then, in a separate loop, the convolutional terms are updated and then injected into the fields. As an example, we shall consider the E_x update. Following (7.106) and (7.107), the field-update loop is written in FORTRAN as

```
do 10 k=2,nz-1
  do 10 j=2,ny-1
    do 10 i=1,nx-1
      ex(i,j,k) = ca_x(i,j,k)*ex(i,j,k)
                + cb_x(i,j,k)*((hz(i,j,k) - hz(i,j-1,k)) / kedy(j) -
                               (hy(i,j,k) - hy(i,j,k-1)) / kedz(k) )
10 continue
```

(7.112)

where kedy(j) = $\kappa_y(j\Delta y)\Delta y$ and kedz(k) = $\kappa_z(k\Delta z)\Delta z$. Note that the tangential electric field on the PEC outer boundary of the FDTD space lattice is not updated, since this field remains zero for all time-steps.

Analogous loops are written for the updates of ey and ez. After updating all the electric fields, the discrete convolutional terms psi are updated and superimposed into the electric fields only in the PML regions. A representative loop through the base y-PML region of thickness nypml1 - 1 cells is given by

```
  do 10 k=1,nz-1
    do 10 j=2,nypml1
      do 10 i=1,nx-1
        psi_Exy1(i,j,k) = be_y(j)* psi_Exy1(i,j,k)
                        + ce_y(j)*((hz(i,j,k) - hz(i,j-1,k)) / dy
        psi_Ezy1(i,j,k) = be_y(j)* psi_Ezy1(i,j,k)
                        + ce_y(j)*((hx(i,j,k) - hx(i,j-1,k)) / dy
        ex(i,j,k) = ex(i,j,k) + cb_x(i,j,k) * psi_Exy1(i,j,k)
        ez(i,j,k) = ez(i,j,k) - cb_z(i,j,k) * psi_Ezy1(i,j,k)
10 continue
```

(7.113)

where

$$\texttt{be\_y(j)} = \exp(\sigma_y^e(\texttt{j}) / \kappa_y^e(\texttt{j}) + a_y^e(\texttt{j})) * \Delta t / \varepsilon_0 \tag{7.114a}$$

$$\texttt{ce\_y(j)} = \sigma_y^e(\texttt{j}) * (\texttt{be\_y(j)} - 1) / (\sigma_y^e(\texttt{j}) + \kappa_y^e(\texttt{j}) * a_y^e(\texttt{j})) / \kappa_y^e(\texttt{j}) \tag{7.114b}$$

and where, assuming the polynomial scaling of (7.60a, b), we have

$$\sigma_y^e(\texttt{j}) = \sigma_{y1}^{\max} * (\ (\texttt{nypml1} - \texttt{j}) / (\texttt{nypml1} - 1)\ )^{**}\texttt{m} \tag{7.115a}$$

$$\kappa_y^e(\texttt{j}) = 1 + (\kappa_{y1}^{\max} - 1) * ((\texttt{nypml1} - \texttt{j}) / (\texttt{nypml1} - 1)\ )^{**}\texttt{m} \tag{7.115b}$$

and finally, from the scaling of (7.79), we have

$$a_y^e(\texttt{j}) = a_{y1}^{\max} * (\ (\texttt{j} - 1) / (\texttt{nypml1} - 1)\ )^{**}\texttt{ma} \tag{7.115c}$$

Analogous loops can be written for the remaining PML regions to update the discrete convolutional terms psi and then superimpose these terms into their corresponding E-field components. For example, in the z-PML region, the discrete convolutional term psi_Exz1 is updated and superimposed into ex. In the same loop, the discrete convolutional term psi_Eyz1 is also updated and superimposed into ey.

For completeness, we now write the field-update loop for H_x, following (7.108):

```
do 10 k=1,nz-1
  do 10 j=1,ny-1
    do 10 i=2,nx-1
      hx(i,j,k) = da_x(i,j,k)*hx(i,j,k)
                - db_x(i,j,k)*((ez(i,j+1,k) - ez(i,j,k)) / khdy(j) -
                               (ey(i,j,k+1) - ey(i,j,k)) / khdz(k) )
10 continue
```
(7.116)

where `khdy(j)` $= \kappa_y[(j-1/2)\Delta y]\Delta y$ and `khdz(k)` $= \kappa_z[(k-1/2)\Delta z]\Delta z$. We note that the normal H-field at the PEC outer boundary of the FDTD space lattice is not updated, since this field remains zero at all time-steps.

Analogous loops are written for the updates of `hy` and `hz`. After updating all the H-fields in this manner, the discrete convolutional terms `psi` are updated and superimposed into the H-fields only in the PML regions. A representative loop through the base y-PML region is given by

```
do 10 k=1,nz-1
  do 10 j=1,nypml1-1
    do 10 i=1,nx-1
      psi_Hxy1(i,j,k) = bh_y(j)* psi_Hxy1(i,j,k)
                      + ch_y(j)*((ez(i,j+1,k) - ez(i,j,k)) / dy
      psi_Hzy1(i,j,k) = bh_y(j)* psi_Hzy1(i,j,k)
                      + ch_y(j)*((ex(i,j+1,k) - ex(i,j,k)) / dy
      hx(i,j,k) = hx(i,j,k) - db_x(i,j,k) * psi_Hxy1(i,j,k)
      hz(i,j,k) = hz(i,j,k) + db_z(i,j,k) * psi_Hzy1(i,j,k)
10 continue
```
(7.117)

where

$$\texttt{bh\_y(j)} = \exp(\sigma_y^h(j)/\kappa_y^h(j) + a_y^h(j)) * \Delta t/\varepsilon_0 \tag{7.118a}$$

$$\texttt{ch\_y(j)} = \sigma_y^h(j)*(\texttt{bh\_y(j)} - 1)/(\sigma_y^h(j) + \kappa_y^h(j)*a_y^h(j))/\kappa_y^h(j) \tag{7.118b}$$

and where, assuming the polynomial scaling of (7.60a, b), we have

$$\sigma_y^h(j) = \sigma_{y1}^{max} * ((\texttt{nypml1} - j - 1/2) / (\texttt{nypml1} - 1)) ** m \tag{7.119a}$$

$$\kappa_y^h(j) = 1 + (\kappa_{y1}^{max} - 1) * ((\texttt{nypml1} - j - 1/2) / (\texttt{nypml1} - 1)) ** m \tag{7.119b}$$

and finally, from the scaling of (7.79), we have

$$a_y^h(j) = a_{y1}^{max} * ((j - 1/2) / (\texttt{nypml1} - 1)) ** ma \tag{7.119c}$$

An important distinction is observed between the PML coefficients used for the E- and H-field updates. The discrete coefficients in (7.115) computed for the E-field updates are evaluated at integer multiples of the cell dimension Δy. On the other hand, the coefficients in (7.119) used for the H-field updates are computed at half-integer multiples of Δy. The distinction is that the coefficients are computed at the edge centers of the primary grid for the E-field updates, and at the face centers for the H-field updates. By correctly following this strategy, the CPML implementation presented in (7.112) to (7.119) is stable.

The efficiency of this implementation can be slightly improved by following some other simple strategies, such as weighting the field intensities by the scaled lengths $\kappa_w dw$, as suggested in Chapter 11, Section 11.2, or by indirectly addressing the physical material parameter arrays `ca_w`, `cb_w`, `da_w`, and `db_w` for memory efficiency, as suggested in Chapter 3, Section 3.6.5. Note also that the method can be readily extended to nonuniform grids following the procedure in Chapter 11, Section 11.2, and to the nonorthogonal grids of Chapter 11, Sections 11.4 to 11.6.

Finally, it was discussed at the end of Section 7.8 that UPML (and similarly split-field PML) is most efficiently and simply implemented by extending the required auxiliary fields throughout the entire FDTD lattice. This is not the case for CPML, which requires storage of its auxiliary convolutional variables `psi` *only* in the PML region. As a consequence, CPML is inherently more storage-efficient than UPML. Furthermore, UPML requires two auxiliary variables per field component to terminate lossy and dispersive media [24, 43], and three auxiliary variables per field component to apply the CFS technique [43]. On the other hand, the CPML formulation remains *unchanged* for all of these cases.

7.10 APPLICATION OF CPML IN FDTD TO GENERAL MEDIA

7.10.1 Introduction

An important feature of the CPML formulation is that it can be readily extended to anisotropic, dispersive, and nonlinear media. Since the field updates in (7.112) and (7.116) are written separately from the CPML updates in (7.113) and (7.117), there is no additional complexity in adding CPML to an FDTD code already designed to treat such general media. This is because, at its most basic level, CPML has nothing to do with the $\boldsymbol{D} - \boldsymbol{E}$ and $\boldsymbol{B} - \boldsymbol{H}$ constitutive relationships (i.e., how fluxes are related to field intensities) that normally characterize a material. This distinguishes CPML from UPML and split-field PML. Rather, CPML is applied as stretched coordinates to the curl operator, which is applied only to the field intensities. While the UPML and split-field PML formulations have been applied to general media, CPML provides a simpler alternative to such applications, since its formulation remains unchanged.

7.10.2 Example: Application of CPML to the Debye Medium

This section focuses on how to implement a CPML termination of the FDTD lattice when it hosts a dispersive medium. We shall find it instructive to provide a specific implementation for the single-pole Debye medium from which more general implementations can be inferred. (See Chapter 9 for a detailed treatment of FDTD modeling of electromagnetic wave propagation in dispersive and nonlinear media located within the interior of the space lattice.)

In the phasor domain, the relative permittivity of a single-pole Debye medium is given by

$$\varepsilon_r(\omega) = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + j\omega\tau} - j\frac{\sigma_s}{\omega\varepsilon_0} \tag{7.120}$$

where ε_s is the static relative permittivity, ε_∞ is the infinite-frequency limit for the relative permittivity, τ is the pole relaxation time, and σ_s is the conductivity. In the Debye medium, Ampere's law can be expressed as

$$\nabla \times \breve{\boldsymbol{H}} = j\omega\breve{\boldsymbol{D}} = j\omega\varepsilon_0\varepsilon_r(\omega)\breve{\boldsymbol{E}} \tag{7.121}$$

From (7.120), this can be written as

$$\nabla \times \breve{\boldsymbol{H}} = j\omega\varepsilon_0\varepsilon_\infty\breve{\boldsymbol{E}} + \sigma_s\breve{\boldsymbol{E}} + \breve{\boldsymbol{J}}_d \tag{7.122}$$

where $\breve{\boldsymbol{J}}_d$ is the polarization current due to the Debye pole given by

$$\breve{\boldsymbol{J}}_d = j\omega\varepsilon_0\left(\frac{\varepsilon_s - \varepsilon_\infty}{1 + j\omega\tau}\right)\breve{\boldsymbol{E}} \tag{7.123}$$

The expressions in (7.122) and (7.123) are then transformed into the time domain and written as a pair of coupled differential equations:

$$\nabla \times \boldsymbol{H} = \varepsilon_0\varepsilon_\infty\frac{\partial}{\partial t}\boldsymbol{E} + \sigma_s\boldsymbol{E} + \boldsymbol{J}_d \tag{7.124}$$

$$\boldsymbol{J}_d + \tau\frac{\partial}{\partial t}\boldsymbol{J}_d = \varepsilon_0\Delta\varepsilon\frac{\partial}{\partial t}\boldsymbol{E} \tag{7.125}$$

where $\Delta\varepsilon = \varepsilon_s - \varepsilon_\infty$.

Now, assume that we want to model a Debye medium extending to infinity through the use of CPML. This can be done by padding the FDTD space lattice with CPML, and extending the Debye medium through the entire CPML. Following the procedure used to derive (7.103), and applying (7.122), the x-projection of Ampere's law in the CPML region is then expressed as

$$\varepsilon_0\varepsilon_\infty\frac{\partial E_x}{\partial t} + \sigma_s E_x + J_{d_x} = \left(\frac{1}{\kappa_y}\frac{\partial H_z}{\partial y} - \frac{1}{\kappa_z}\frac{\partial H_y}{\partial z}\right) + \left(\zeta_y * \frac{\partial H_z}{\partial y} - \zeta_z * \frac{\partial H_y}{\partial z}\right) \tag{7.126}$$

where the right-hand side is independent of the media type, as discussed previously.

A standard finite-difference procedure can then be applied to formulate explicit update expressions for (7.126) and (7.125). Following a standard Yee scheme for (7.126), second-order accurate difference approximations would be of the form

$$\varepsilon_0\varepsilon_\infty\left(\frac{E_x\big|_{i+1/2,j,k}^{n+1/2}-E_x\big|_{i+1/2,j,k}^{n-1/2}}{\Delta t}\right)+\sigma_s\left(\frac{E_x\big|_{i+1/2,j,k}^{n+1/2}+E_x\big|_{i+1/2,j,k}^{n-1/2}}{2}\right)+J_{d_x}\big|_{i+1/2,j,k}^{n}=$$

$$\left[\begin{array}{l}\dfrac{H_z\big|_{i+1/2,j+1/2,k}^{n}-H_z\big|_{i+1/2,j-1/2,k}^{n}}{\kappa_{y_j}\Delta y}-\dfrac{H_y\big|_{i+1/2,j,k+1/2}^{n}-H_y\big|_{i+1/2,j,k-1/2}^{n}}{\kappa_{z_k}\Delta z}\\ +\ \psi_{E_{x,y}}\big|_{i+1/2,j,k}^{n}-\psi_{E_{x,z}}\big|_{i+1/2,j,k}^{n}\end{array}\right] \tag{7.127}$$

Then, from (7.125), we have

$$\left(\frac{J_{d_x}\big|_{i+1/2,j,k}^{n+1/2}+J_{d_x}\big|_{i+1/2,j,k}^{n-1/2}}{2}\right)+\tau\left(\frac{J_{d_x}\big|_{i+1/2,j,k}^{n+1/2}-J_{d_x}\big|_{i+1/2,j,k}^{n-1/2}}{\Delta t}\right)=\varepsilon_0\Delta\varepsilon\left(\frac{E_x\big|_{i+1/2,j,k}^{n+1/2}-E_x\big|_{i+1/2,j,k}^{n-1/2}}{\Delta t}\right) \tag{7.128}$$

Equation (7.128) can be used to derive an explicit update equation for J_{d_x}:

$$J_{d_x}\big|_{i+1/2,j,k}^{n+1/2}=k_d\,J_{d_x}\big|_{i+1/2,j,k}^{n-1/2}+\beta_d\left(\frac{E_x\big|_{i+1/2,j,k}^{n+1/2}-E_x\big|_{i+1/2,j,k}^{n-1/2}}{\Delta t}\right) \tag{7.129}$$

where

$$k_d=(2\tau-\Delta t)/(2\tau+\Delta t)\,;\qquad \beta_d=(2\varepsilon_0\Delta\varepsilon\Delta t)/(2\tau+\Delta t) \tag{7.130a, b}$$

In (7.127), $J_{d_x}\big|^n$ is evaluated as:

$$J_{d_x}\big|_{i+1/2,j,k}^{n}=\left(\frac{J_{d_x}\big|_{i+1/2,j,k}^{n-1/2}+J_{d_x}\big|_{i+1/2,j,k}^{n+1/2}}{2}\right)$$

$$=\frac{1}{2}(1+k_d)\,J_{d_x}\big|_{i+1/2,j,k}^{n-1/2}+\frac{\beta_d}{2\Delta t}\left(E_x\big|_{i+1/2,j,k}^{n+1/2}-E_x\big|_{i+1/2,j,k}^{n-1/2}\right) \tag{7.131}$$

After inserting (7.131) into (7.128), we obtain an explicit update equation for the electric field as

$$E_x\big|_{i+1/2,j,k}^{n+1/2}=C_a\big|_{i+1/2,j,k}\,E_x\big|_{i+1/2,j,k}^{n-1/2}$$

$$+\ C_b\big|_{i+1/2,j,k}\left[\begin{array}{l}\dfrac{H_z\big|_{i+1/2,j+1/2,k}^{n}-H_z\big|_{i+1/2,j-1/2,k}^{n}}{\kappa_{y_j}\Delta y}-\dfrac{H_y\big|_{i+1/2,j,k+1/2}^{n}-H_y\big|_{i+1/2,j,k-1/2}^{n}}{\kappa_{z_k}\Delta z}\\ +\ \psi_{E_{x,y}}\big|_{i+1/2,j,k}^{n}-\psi_{E_{x,z}}\big|_{i+1/2,j,k}^{n}-\dfrac{1}{2}(1+k_d)\,J_{d_x}\big|_{i+1/2,j,k}^{n-1/2}\end{array}\right] \tag{7.132}$$

where

$$C_a\big|_{i+1/2,j,k} = \frac{2\varepsilon_0\varepsilon_{\infty i+1/2,j,k} - \sigma_{s\,i+1/2,j,k}\Delta t + \beta_{d\,i+1/2,j,k}}{2\varepsilon_0\varepsilon_{\infty i+1/2,j,k} + \sigma_{s\,i+1/2,j,k}\Delta t + \beta_{d\,i+1/2,j,k}} \tag{7.133a}$$

$$C_b\big|_{i+1/2,j,k} = \frac{2\Delta t}{2\varepsilon_0\varepsilon_{\infty i+1/2,j,k} + \sigma_{s\,i+1/2,j,k}\Delta t + \beta_{d\,i+1/2,j,k}} \tag{7.133b}$$

Finally, the CPML convolutional terms $\psi_{E_{w,v}}$ are updated using the expressions in (7.105a, b).

If we compare (7.132) and (7.133) with (7.106) and (7.107), the only difference is the expression for the update coefficients in (7.133a, b) and the inclusion of the Debye polarization current and the polarization current update. The implementation details are identical to those in (7.112) to (7.115) plus some minor additions specific to the Debye media. Specifically, the Debye polarization current density update in (7.129) must be included in the field-update loop in (7.112). Another subtle difference is that since $J_{d_x}\big|^{n+1/2}$ is a function of $E_x\big|^{n+1/2}$, a partial update of $J_{d_x}\big|^{n+1/2}$ to include the contribution of the updated CPML variables $\psi_{E_{w,v}}$ to the electric field should also be included in the CPML update loop in (7.113). Finally, the H-field updates are identical to those presented in (7.108) to (7.110).

7.11 NUMERICAL EXPERIMENTS WITH PML

This section presents examples of using PML to terminate FDTD grids for a number of typical applications. The goal is to provide an understanding of how the PML tensor coefficients and grading functions impact the effectiveness of the PML as an ABC. In this manner, the reader can better understand how to properly choose these parameters. The CPML formulation discussed in Section 7.9 is used for all simulations unless otherwise noted.

7.11.1 Current Source Radiating in an Unbounded Two-Dimensional Region

Fig. 7.2 illustrates the first example, a y-directed electric current source centered in a 40×40-cell TE_z FDTD grid. The source is constant along its axis, and has the time variation of a differentiated Gaussian pulse:

$$J_y(x_0, y_0, t) = -2\left[(t - t_0)/t_w\right]\exp\left\{-\left[(t - t_0)/t_w\right]^2\right\} \tag{7.134}$$

where $t_w = 26.53$ ps and $t_0 = 4t_w$.

The FDTD grid has 1-mm-square cells and a time-step of 0.99 times the Courant limit. The E-field is probed at two points, A and B, as shown in Fig. 7.2. Point A is in the same y-plane as the source and two cells from the PML, and point B is two cells from the bottom and side PMLs. Time-stepping runs over 1,000 iterations, well past the steady-state response. Both 6-cell and 10-cell PML ABCs are used with polynomial grading as defined in (7.60), with $m = 3$.

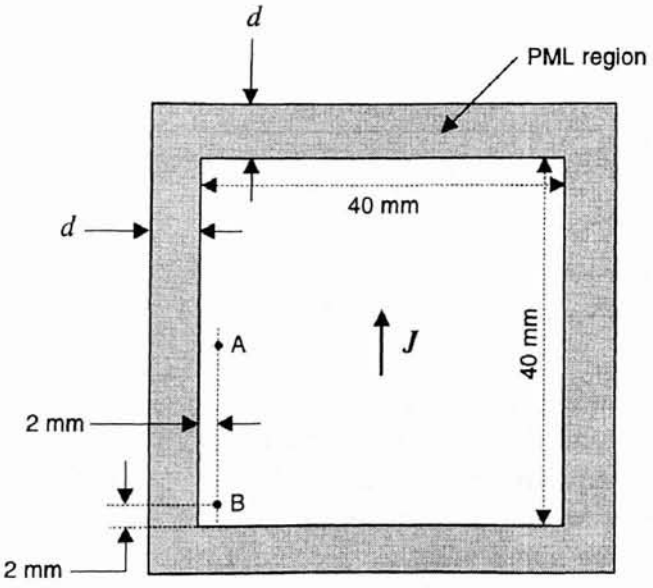


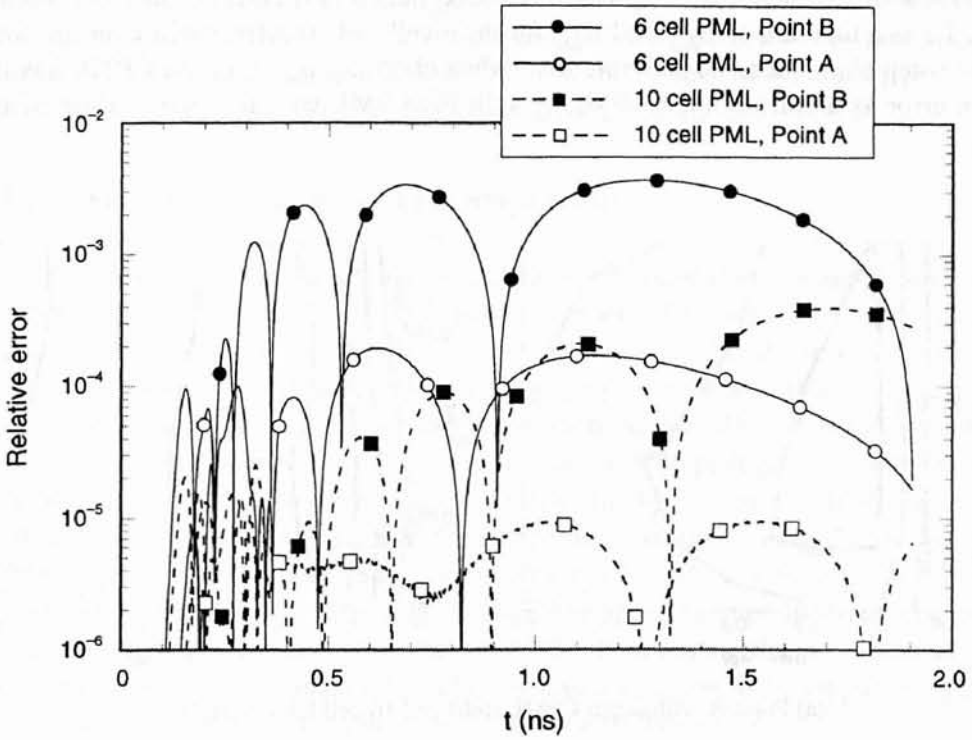
Fig. 7.2 y-directed electric current source centered in a two-dimensional 40×40-mm TE_z FDTD grid. The working volume is surrounded by PML of thickness d. E-fields are probed at points A and B.

For this case, the reference solution $E_{\text{ref}}|_{i,j}^{n}$ at grid location (i, j) and time-step n is obtained using a large $1{,}040 \times 1{,}040$-cell grid. An identical current source is centered within this grid, and the field-observation point (i, j) is at the same position relative to the source as in the test grid. The reference grid is sufficiently large such that there are no reflections from its outer boundaries during the time-stepping span of interest. This allows a relative error to be defined as

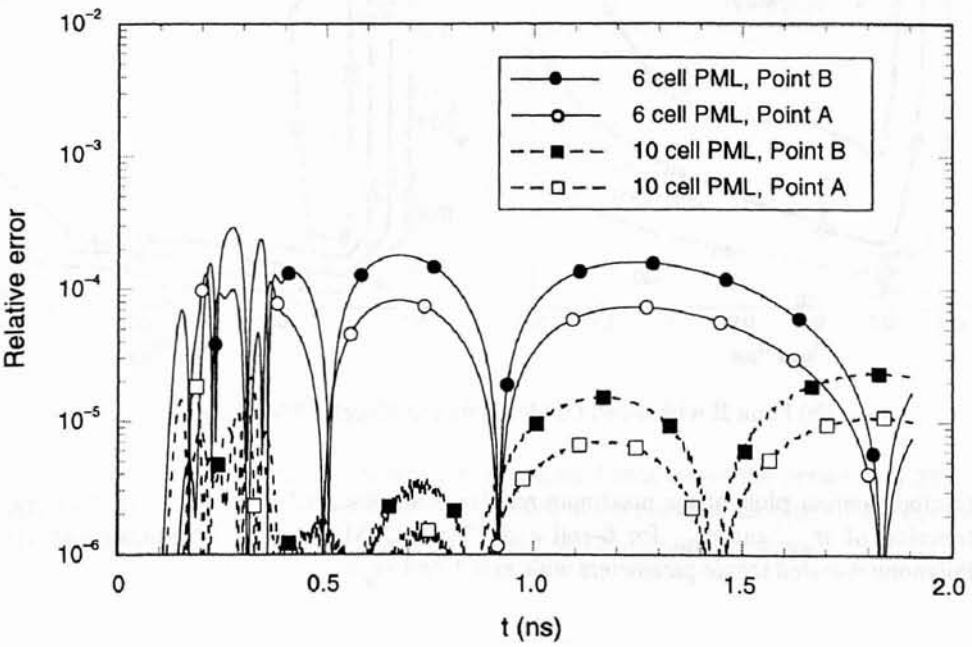
$$\text{Rel. error}|_{i,j}^{n} = \left| E|_{i,j}^{n} - E_{\text{ref}}|_{i,j}^{n} \right| \Big/ \left| E_{\text{ref max}}|_{i,j} \right| \tag{7.135}$$

where $E_{\text{ref max}}|_{i,j}$ is the maximum amplitude of the reference field at (i, j), as observed during the time-stepping span of interest.

Fig. 7.3 graphs the relative error calculated using (7.135) at points A and B of Fig. 7.2 over the first 1,000 time-steps of the FDTD run for 6-cell and 10-cell PMLs. In Fig. 7.3(a), each PML has polynomial-scaled σ with $m = 3$, $\sigma_{\max} = \sigma_{\text{opt}}$ [where σ_{opt} is given by (7.67)], $\kappa_{\max} = 1$, and $a_{\max} = 0$, yielding the properties of UPML. In Fig. 7.3(b), all PML parameters are the same, with the exception that $a_{\max} = 0.2$ and $m_a = 1$, yielding the properties of CPML. Comparing Figs. 7.3(a) and 7.3(b), we see that the CPML provides a clearly reduced error, especially at B, where the wave impinges on the PML obliquely at 45°. The CPML appears to mitigate the two sources of error for a wave impinging obliquely upon PML: (1) degraded bulk absorption by the PML, as per (7.59), relative to the absorption at normal incidence; and (2) degraded transparency of the front PML interface, as per Section 7.7, due to the prolonging of wave interactions with this interface relative to what occurs at normal incidence. The effect of the latter phenomenon can be observed in Fig. 7.3(a), where the reflection error at B increases in the late time, and is very slow to decay.



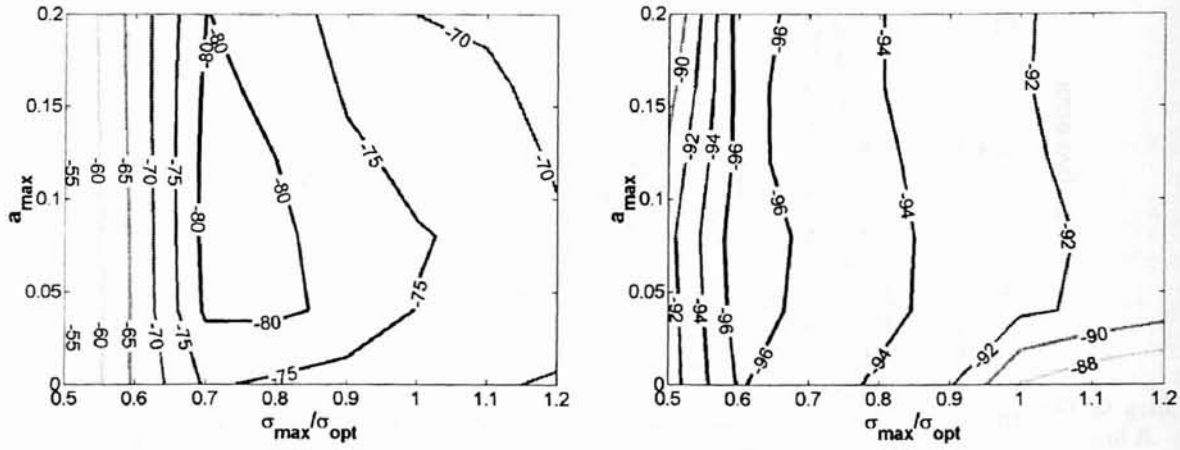
(a) $a_{max} = 0$ (equivalent to UPML).



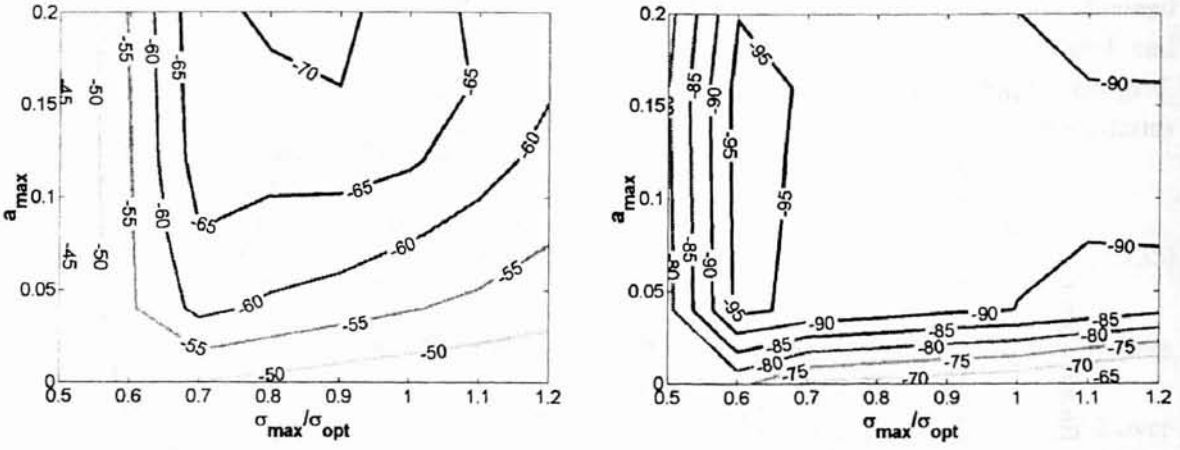
(b) CPML with $a_{max} = 0.2$ and $m_a = 1$.

Fig. 7.3 Relative error at points A and B of Fig. 7.2 over 1,000 time-steps for 6-cell and 10-cell PMLs with $\sigma_{max} = \sigma_{opt}$, $\kappa_{max} = 1$, and $m = 3$.

Fig. 7.4 shows decibel contour plots of the maximum relative error observed at points A and B in Fig. 7.2 as a function of σ_{max} and a_{max} for the 6-cell and 10-cell CPMLs over the course of a 1,000-time-step simulation. Again note that, when choosing $a_{max} = 0$, the CPML has the same reflection error as a conventional UPML or split-field PML with the same values of σ_{max} and κ_{max} [37].



(a) Point A with 6-cell CPML (left) and 10-cell CPML (right).



(b) Point B with 6-cell CPML (left) and 10-cell CPML (right).

Fig. 7.4 Decibel contour plots of the maximum relative error observed at points A and B in Fig. 7.2 as a function of σ_{max} and a_{max} for 6-cell and 10-cell CPMLs over a 1,000-time-step simulation. Polynomial-scaled tensor parameters with $m = 3$ and $m_a = 1$.

From Fig. 7.4(a), we see that the reflection error at A is only weakly dependent on a_{max}. However, from Fig. 7.4(b), we see that the reflection error at B can be quite strongly dependent on a_{max}. In fact, up to a 20 dB reduction in error relative to classical split-field PML or UPML can be achieved with the proper choice of a_{max}. This improved accuracy comes at no additional cost when using the CPML algorithm.

While there is an optimum choice of σ_{max} and a_{max} in Fig. 7.4, it is clear that there is a broad range of these parameters that lead to an excellent ABC with exceptionally low error levels. In fact, for this discretization, one would expect roughly three digits of accuracy from the FDTD simulation. Consequently, a 6-cell CPML is sufficiently accurate for this problem.

7.11.2 Highly Elongated Domains and Edge Singularities

The previous example involved a simple homogeneous two-dimensional space with a square domain. While it is necessary for an ABC to perform well for such an example, it does not demonstrate the full potential of the CPML formulation. Fig. 7.5 illustrates a more taxing example: three-dimensional electromagnetic scattering by an elongated thin PEC plate (having dimensions 25×100 mm). The FDTD model is discretized spatially with uniform cubic cells spanning $\Delta = 1$ mm, and temporally with a time-step $\Delta t = 1.9066$ ps (equal to $0.99 \times$ the Courant limit). Excitation is provided by a z-directed electric dipole located 1 mm above one of the corners of the plate. The time signature of the excitation is the differentiated Gaussian pulse of (7.134), with a half-width $t_w = 53$ ps and a time delay $t_0 = 4t_w$. In this example, we shall study the computed electric field in the plane of the plate at its opposite corner, in a direction normal to the plate edge at a distance of 1 mm from the edge. This field has significant evanescent content due to the presence of the edge singularity. There is also a creeping wave supported by the plate that requires a long-time interaction of the fields with the PML boundary.

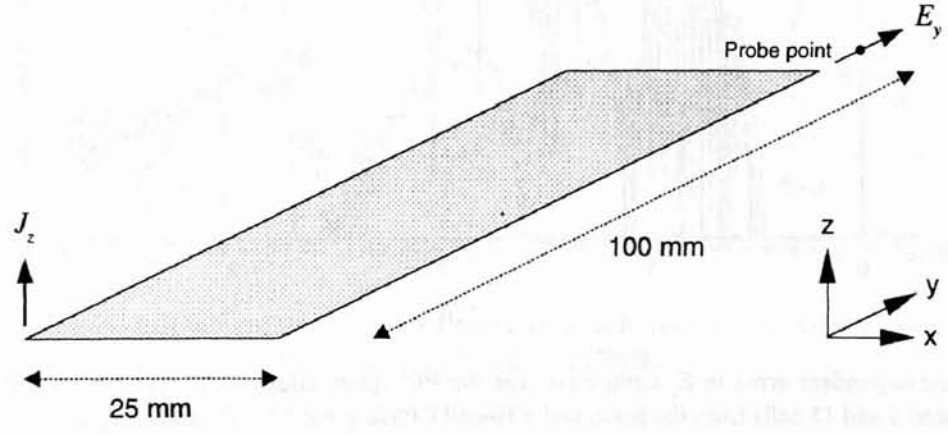


Fig. 7.5 Geometry of a vertical electric current dipole placed 1 mm above the corner of a 25×100 mm thin PEC plate. The horizontal E-field is computed in the plane of the plate at its opposite corner, in a direction normal to the plate edge at a distance of 1 mm from the edge.

To assess the accuracy of the simulation, we compare the FDTD results with reference data obtained using a significantly larger FDTD problem domain ($201 \times 276 \times 176$ cells). The reference domain includes an optimized 10-cell CPML ABC at each of its six outermost boundaries, and 78 free-space cells between the plate and the front CPML boundary. The error in the trial FDTD simulations using much smaller space lattices is then computed using (7.135).

Fig. 7.6 illustrates the error in the calculated E_y at the probe point for three trial simulations in which the FDTD lattice is terminated with either a 10-cell UPML or a 10-cell CPML. In the first trial, the UPML is placed only 3 cells from the plate in all directions, yielding an overall lattice size of $51 \times 126 \times 26$ cells. Standard UPML tensor coefficients are assumed as per (7.55) with $\sigma_{max} = 0.75\sigma_{opt}$, $\kappa_{max} = 15$, and $m = 3$ polynomial scaling. In the second trial, the same UPML coefficients and scaling are assumed, but the UPML is placed 15 cells from the PEC plate in all directions, yielding an overall lattice size of $75 \times 150 \times 50$ cells. In the third trial, the CPML is placed only 3 cells from the plate in all directions. Here, a_w is linearly scaled ($m_a = 1$) with $a_{max} = 0.24$, and $m = 3$ polynomial scaling is applied to σ_w and κ_w for $\sigma_{max} = \sigma_{opt}$ and $\kappa_{max} = 15$.

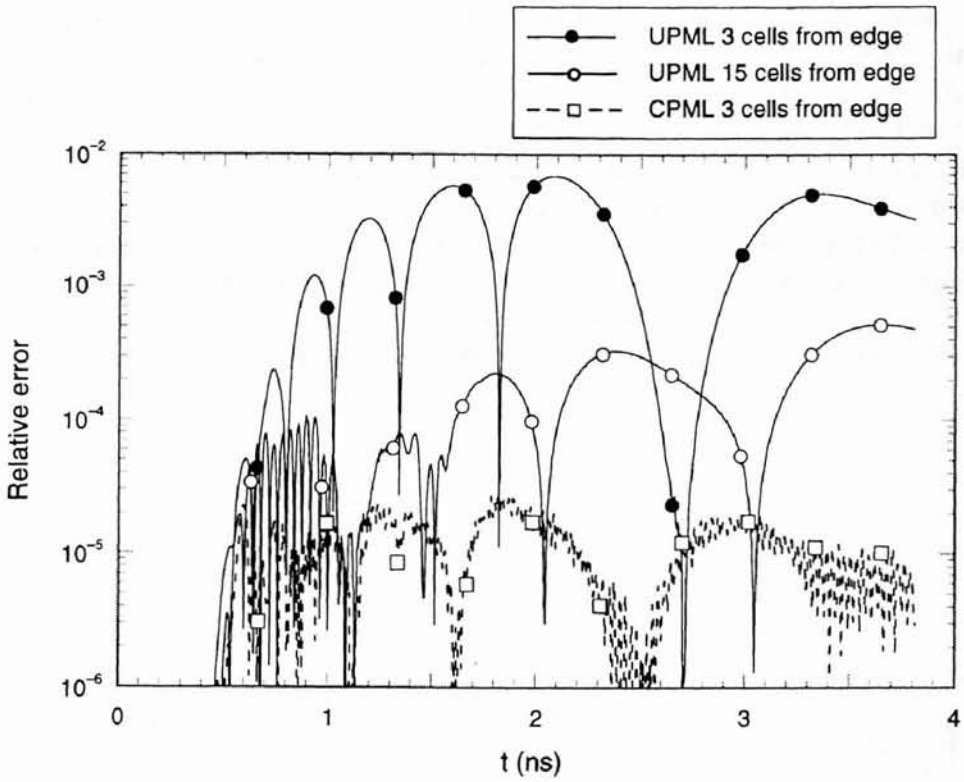
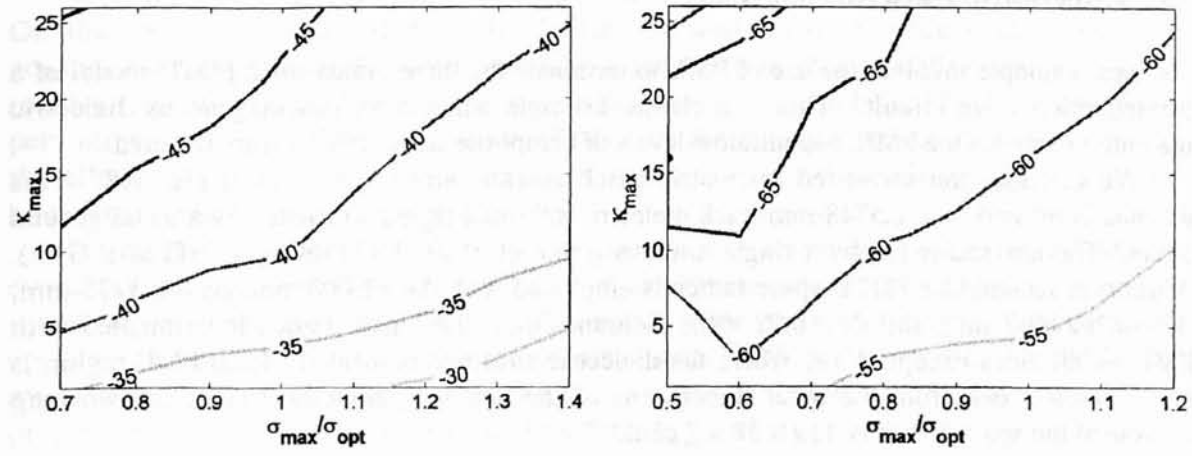
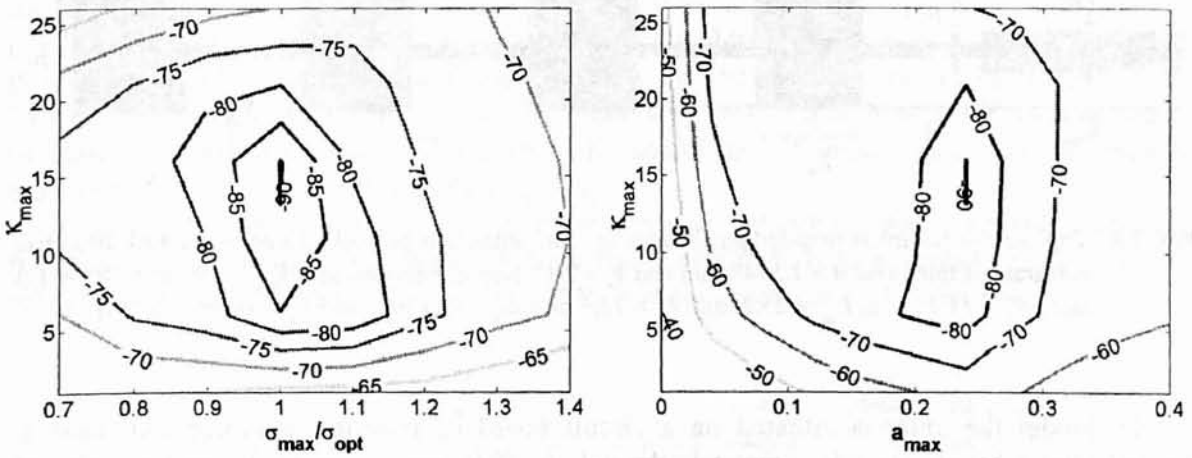


Fig. 7.6 Time-dependent error in E_y computed near the PEC plate edge of Fig. 7.5 for 10-cell UPMLs placed 3 and 15 cells from the plate, and a 10-cell CPML placed 3 cells from the plate.

From Fig. 7.6, we see that the closely located UPML exhibits significant late-time error. Specifically, for the UPML placed only 3 cells from the PEC plate, the error is on the order of 1% relative to the amplitude of the leading pulse. Such error may be intolerable if an accurate simulation of the creeping wave is required. As expected, if the UPML is moved 15 cells from the plate, the error drops by about 10:1. However, using the CPML placed only 3 cells from the plate is even better, with the error reduced by yet another one to two orders-of-magnitude. This means that the CPML formulation permits a much smaller space lattice to be employed while retaining accuracy. In this example, using the closely located CPML simultaneously reduces the required computer memory and running time by about 70% (relative to the UPML placed 15 cells from the edge of the plate), *and* improves accuracy by better than one full digit!



(a) UPML placed 3 cells from the plate (left) and 15 cells from the plate (right).



(b) CPML placed 3 cells from the plate with $m_a = 1$ and $a_{max} = 0.24$ (left) and $\sigma_{max} = \sigma_{opt}$ (right).

Fig. 7.7 Decibel contour plots of the maximum relative error in E_y observed at the probe point in Fig. 7.5 for $m = 3$ polynomial-scaled 10-cell-thick UPML and CPML.

Fig. 7.7 shows decibel contour plots of the maximum relative error in E_y observed at the probe point in Fig. 7.5 as a function of κ_{max} and σ_{max} for the UPML, and κ_{max}, σ_{max}, and a_{max} for the CPML. In both cases, the PML is 10 cells thick and polynomial scaled with $m = 3$ (determined in other trials to provide the most accurate results.) From Fig. 7.7(a), we see that significant error results if the UPML is located too close to the plate. This is predominately due to the presence of evanescent waves in the singular fields near the plate edge, and the long time interaction of the creeping waves. If the UPML is placed further away from the plate edge, then this error is reduced. From Fig. 7.7(b), we see that the CPML is significantly more accurate than the UPML even when placed very close to the plate. Here, optimum CPML parameters are in the ranges $0.8\sigma_{opt} < \sigma_{max} < 1.4\sigma_{opt}$, $7 < \kappa_{max} < 20$, and $0.15 < a_{max} < 0.3$. These yield reflection errors below –70 dB (or greater than 3.5 digits of accuracy), which is more than sufficient for this discretization.

7.11.3 Microstrip Patch Antenna Array

The next example involves the use of PML to terminate the three-dimensional FDTD model of a printed microwave circuit. This is a classic example wherein an inhomogeneous dielectric medium penetrates the PML, and ultralow levels of composite wave reflection are required.

We consider the series-fed microstrip patch antenna array illustrated in Fig. 7.8. This antenna is printed on a 1.5748-mm-thick dielectric substrate ($\varepsilon_r = 2.1$) backed by a metal ground plane. The antenna is fed by a single microstrip line of width 3.93 mm ($Z_0 \cong 56\Omega$ at 9 GHz). A uniform rectangular FDTD space lattice is employed with $\Delta x = 1.008$ mm, $\Delta y = 0.3275$ mm, $\Delta z = 0.2624667$ mm, and $\Delta t = 0.99 \times$ the Courant limit. The space lattice is terminated with PML on all sides except at $-z$, where the dielectric substrate is located. Each PML region is placed *only 1 cell* from the outer dimensions of the antenna geometry. Thus, the working volume of the space lattice is $119 \times 38 \times 7$ cells.

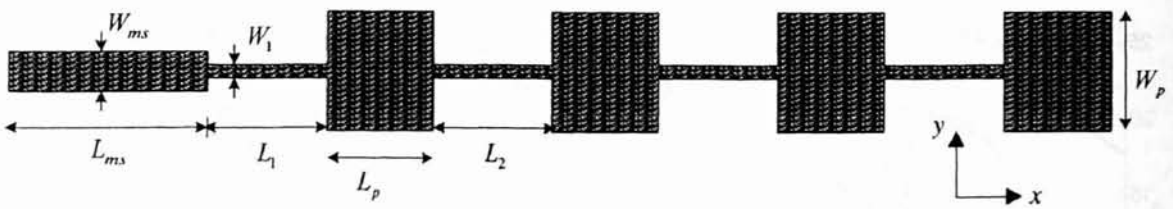


Fig. 7.8 Top view of the microstrip-fed patch antenna. The antenna is printed on a metal-backed dielectric substrate of thickness $h = 1.5748$ mm and $\varepsilon_r = 2.1$. Key dimensions are $W_{ms} = 3.93$ mm, $W_1 = 1.3$ mm, $W_p = 11.79$ mm, $L_{ms} = 25.2$ mm, $L_1 = 13.1$ mm, $L_2 = 12.1$ mm, and $L_p = 10.08$ mm [55].

To model the antenna situated on a circuit board of infinite transverse dimensions, the dielectric substrate extends completely through the PML regions on the $\pm x$- and $\pm y$-sides of the space lattice, ending at the outer PEC planes backing the PML. To model a matched transmission line exciting the antenna, the microstrip feed extends completely through the PML at the $-x$-side of the lattice to the outer PEC backing plane. Table 7.1 lists the parameters used for each CPML region.

TABLE 7.1

CPML Parameters Used for the Microstrip Patch Antenna Array Model of Fig. 7.8

CPML Region	$\varepsilon_{r_{eff}}$	m	$\sigma_{max}/\sigma_{opt}$	σ_{max}	κ_{max}	a_{max}	m_a
$-x$	1.55	3	1.0	6.8	1	0.0	n/a
$+x$	1.55	3	1.0	6.8	15	0.2	1
$-y$	1.55	3	1.0	20.8	15	0.2	1
$+y$	1.55	3	1.0	20.8	15	0.2	1
$-z$	–	–	–	–	–	–	
$+z$	1.0	3	1.0	32.4	15	0.2	1

With regard to Table 7.1, σ_{opt} can be chosen for each CPML region according to (7.67). On the $\pm x$- and $\pm y$-sides of the space lattice, the medium is inhomogeneous. Since the constitutive parameters scale only along the normal direction of each CPML and are constant relative to the transverse directions, only a single value of σ_{opt} can be used. Thus, an *effective* permittivity must be used, as suggested by (7.67). For this example, we choose a mean value for the effective permittivity on the sidewalls. Note that, on the +z boundary, $\varepsilon_{r_{eff}} = 1$, since this boundary is entirely in the air region.

As discussed in Section 7.7, while choosing $\kappa_{max} > 1$ and $a_{max} > 0$ helps to attenuate evanescent modes, these values can degrade the absorption of propagating modes. Therefore, we set $\kappa_{max} = 1$ and $a_{max} = 0$ in the $-x$ CPML region, where the waves impinging on the CPML are predominately due to the purely propagating, quasi-TEM mode supported by the feeding microstrip. In contrast, waves impinging on the other PML regions are a superposition of propagating and evanescent modes, and therefore κ_{max} and a_{max} are set to reasonable values, as determined in the previous section.

The driving-point voltage waveform is calculated for three different trial PML terminations which use the parameters of Table 7.1 and are placed only 1 cell from the outer dimensions of the antenna geometry: (1) 10-cell UPML ($a_{max} = 0$ for all regions), (2) 6-cell CPML, and (3) 10-cell CPML. In addition, a reference solution is obtained by increasing the separation of the 10-cell CPML from the antenna geometry by 50 space cells, yielding an overall $219 \times 138 \times 57$-cell lattice. In all cases, the microstrip feed is driven by a soft voltage source exciting the fundamental quasi-TEM mode. This source has a Gaussian-pulse time signature of half-width $t_w = 31.831$ ps and delay time $t_0 = 4t_w$.

Fig. 7.9 illustrates the relative error in the driving-point voltage waveform (normalized to the maximum source voltage). In the late time, the 6-cell CPML reduces error by more than 10:1 relative to the 10-cell UPML. The 10-cell CPML is about a factor of 10:1 better yet.

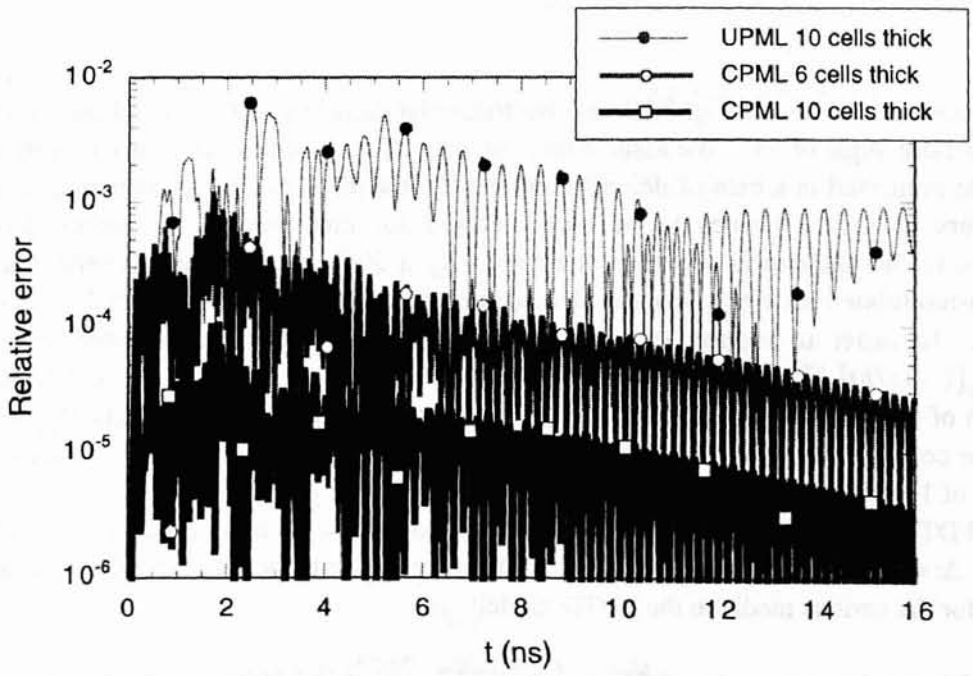


Fig. 7.9 Relative error in the driving-point voltage waveform for a 10-cell UPML, a 6-cell CPML, and a 10-cell CPML placed only 1 cell from the outer dimensions of the antenna geometry of Fig. 7.8.

7.11.4 Dispersive Media

Accurate and efficient modeling of electromagnetic waves in dispersive and nonlinear media is an important need for research in bioelectromagnetics, nonlinear optics, and geophysics. The PML was the first absorbing boundary condition for FDTD to allow such modeling for broadband excitations [24]. As discussed previously in this chapter, the CPML formulation is ideal for such applications since it has the advantage of material-independence. While some fine tuning may be required to optimize the choice of CPML parameters, this technique has proven to be very robust and broadly applicable to all types of materials.

In this section, we study the termination of an FDTD space lattice with CPML to model the radiation of a broadband bow-tie antenna that is potentially useful for detection and hyperthermia treatment of breast cancer [56, 57]. Fig. 7.10 illustrates the geometry of the bow-tie antenna.

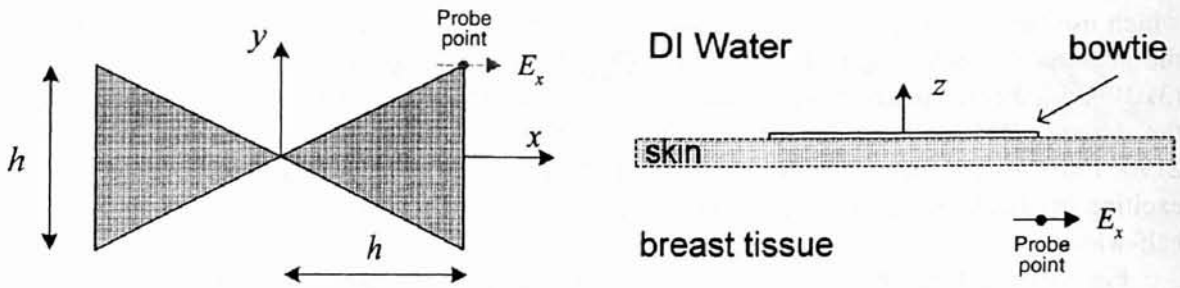


Fig. 7.10 Bow-tie antenna potentially useful for detection and hyperthermia treatment of breast cancer [56, 57]: (a) top view; (b) side view.

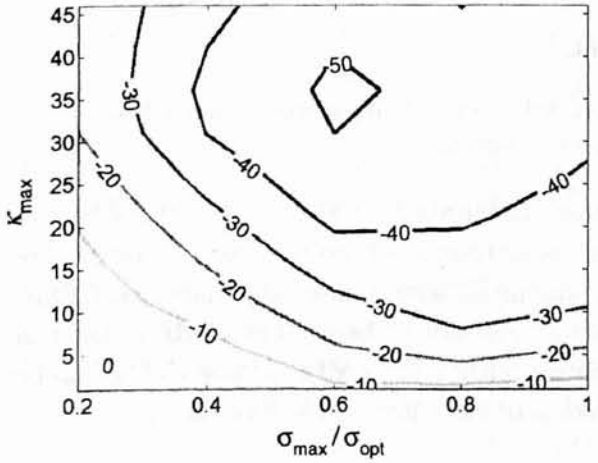
The bow-tie antenna of Fig. 7.10 has two triangular elements of length and base width $h = 4$ cm with a flare angle of 53°. We assume that the antenna is placed directly on 1-mm-thick breast skin while immersed in a bath of deionized water. Below the skin (along $-z$) is modeled several centimeters of breast tissues to be characterized for electromagnetic power deposition. Excitation to the antenna is provided at its apex by a 50Ω source which generates a 0.22-ns Gaussian-modulated carrier pulse (6-GHz center frequency, 4-GHz spectral width, no dc content). In order to reduce end reflections, the antenna has the variable conductivity $\sigma(x) = \sigma_0[1 - (x/h)] / [1 + (\sigma_0/\sigma_{1/2} - 2)\cdot(x/h)]$ S/m, where x/h is the normalized distance along the length of the antenna, $\sigma_0 = 3.27 \times 10^7$ S/m is the conductivity at the apex, and $\sigma_{1/2} = 1.0$ S/m. The finite conductivity is modeled as a sheet resistance, with the antenna assumed to have a thickness of 1 mm.

The FDTD model uses the spatial discretization $\Delta x = \Delta y = 0.5$ mm and $\Delta z = 0.25$ mm, and a time-step $\Delta t = 0.99 \times$ the Courant limit. The following Debye dispersive material parameters are assumed for the various media in the FDTD model:

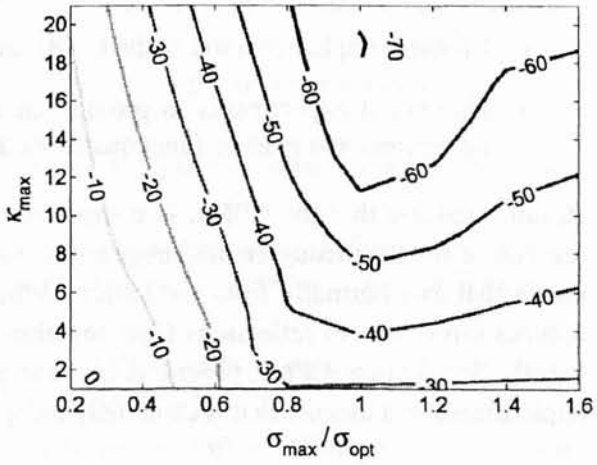
deionized water	$\varepsilon_\infty = 4.55$	$\varepsilon_s = 77.1$	$\sigma_s = 0.0002$	$\tau = 7.4$ ps
skin	$\varepsilon_\infty = 4$	$\varepsilon_s = 37$	$\sigma_s = 1.1$	$\tau = 7.23$ ps
breast tissue	$\varepsilon_\infty = 7$	$\varepsilon_s = 10$	$\sigma_s = 0.15$	$\tau = 7$ ps

The FDTD lattice is terminated by $m = 3$ polynomial-scaled PML: UPML in the first trial, and CPML in the second trial. Each PML is 10 cells thick and is placed 3 cells above the antenna and 3 cells from the edges of the antenna. Each uses a value of σ_{opt} predicted from (7.67) for $\varepsilon_{r_{eff}} = 4$ on the side boundaries; $\varepsilon_{r_{eff}} = 7$ on the lower boundary (terminating the breast-tissue region); and $\varepsilon_{r_{eff}} = 4.55$ on the upper boundary (terminating the DI-water region). For the CPML, a_w is linearly scaled with $m_a = 1$. To quantify the error, we probe the electric field E_x at the position $(x, y) = (h, h/2)$ and 2 mm below the skin in the breast tissue, and compare with a reference FDTD model conducted on a much larger space lattice.

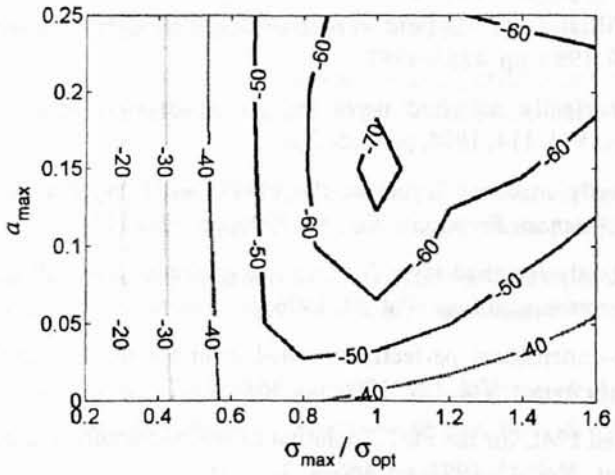
Figs. 7.11(a, b) are decibel contour plots of the maximum error at the probe point as a function of κ_{max} and σ_{max} for the UPML and the CPML, respectively, with a fixed value $a_{max} = 0.20$ used for the CPML. The CPML gives approximately 20 dB improved accuracy for this discretization. Fig. 7.11(c) completes the study for the CPML by plotting its error as a function of a_{max} and σ_{max} for a fixed $\kappa_{max} = 21$.



(a) UPML.



(b) CPML with fixed $a_{max} = 0.20$.



(c) CPML with fixed $\kappa_{max} = 21$.

Fig. 7.11 Decibel contour plots of the maximum relative error in E_x observed at the probe point in Fig. 7.10 for $m = 3$ polynomial-scaled 10-cell-thick UPML and CPML.

7.12 SUMMARY AND CONCLUSIONS

This chapter reviewed the theoretical foundation and numerical implementation of perfectly matched layer absorbing boundary conditions in Cartesian FDTD space lattices. Parametric studies of PML performance under a variety of conditions were reported. The review included:

- Berenger's original split-field PML concept;
- The stretched-coordinate PML formulation;
- Theoretical development of the UPML;
- Theoretical performance of the PML;
- The CFS-tensor coefficients;
- Efficient implementation of the UPML in FDTD;
- Efficient implementation of the CPML in FDTD;
- Numerical experiments to provide an understanding of the effect of the PML parameters and grading functions on PML effectiveness.

Results indicate that the CPML is a superior means to terminate FDTD space lattices having regions of homogeneous or inhomogeneous, lossless or lossy, and nondispersive or dispersive media that exit normally from the lattice. When employed with optimized parameters, CPML reduces errors due to reflections from the absorbing boundary by better than 20 dB relative to UPML. In addition, CPML programming is simplified relative to UPML, because CPML can be implemented in a manner that is essentially independent of the material to be terminated.

REFERENCES

[1] Holland, R., and J. Williams, "Total-field versus scattered-field finite-difference," *IEEE Trans. Nuclear Science,* Vol. 30, 1983, pp. 4583–4587.

[2] Berenger, J. P., "A perfectly matched layer for the absorption of electromagnetic waves," *J. Computational Physics,* Vol. 114, 1994, pp. 185–200.

[3] Berenger, J. P., "Perfectly matched layer for the FDTD solution of wave-structure interaction problems," *IEEE Trans. Antennas Propagat.,* Vol. 51, 1996, pp. 110–117.

[4] Berenger, J. P., "A perfectly matched layer for free-space simulations in finite-difference computer codes," *Annales des Telecommunications,* Vol. 51, 1996, pp. 39–46.

[5] Berenger, J. P., "Three-dimensional perfectly matched layer for the absorption of electromagnetic waves," *J. Computational Physics,* Vol. 127, 1996, pp. 363–379.

[6] Berenger, J. P., "Improved PML for the FDTD solution of wave-structure interaction problems," *IEEE Trans. Antennas Propagat.,* Vol. 45, 1997, pp. 466–473.

[7] Berenger, J. P., "Numerical reflection of evanescent waves from perfectly matched layers," *Proc. IEEE Antennas Propagat. Soc. Intl. Symp.,* Montreal, Canada, Vol. 3, 1997, pp. 1888–1891.

[8] 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 Guided Wave Lett.*, Vol. 4, 1994, pp. 268–270.

[9] Chew, W. C., and W. H. Weedon, "A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates," *IEEE Microwave Guided Wave Lett.*, Vol. 7, 1994, pp. 599–604.

[10] Mittra, R., and U. Pekel, "A new look at the perfectly matched layer (PML) concept for the reflectionless absorption of electromagnetic waves," *IEEE Microwave Guided Wave Lett.*, Vol. 5, 1995, pp. 84–87.

[11] Andrew, W. V., C. A. Balanis, and P. A. Tirkas, "A comparison of the Berenger perfectly matched layer and the Lindman higher-order ABCs for the FDTD method," *IEEE Microwave Guided Wave Lett.*, Vol. 5, 1995, pp. 192–194.

[12] Cai, M., and H. Zhou, "Proof of perfectly matched layer conditions in three dimensions," *Electronics Lett.*, Vol. 31, 1995, pp. 1675–1676.

[13] Nehrbass, J. W., J. F. Lee, and R. Lee, "Stability analysis for perfectly matched layered absorbers," *Electromagnetics,* Vol. 16, 1996, pp. 385–389.

[14] Wu, Z., and J. Fang, "High-performance PML algorithms," *IEEE Microwave Guided Wave Lett.*, Vol. 6, 1996, pp. 335–337.

[15] Wittwer, D. C., and R. W. Ziolkowski, "How to design the imperfect Berenger PML," *Electromagnetics,* Vol. 16, 1996, pp. 465–468.

[16] Veihl, J. C., and R. Mittra, "An efficient implementation of Berenger's perfectly matched layer (PML) for finite-difference time-domain mesh truncation," *IEEE Microwave Guided Wave Lett.*, Vol. 6, 1996, pp. 94–96.

[17] De Moerloose, J., and M. A. Stuchley, "Reflection analysis of PML ABC's for low-frequency applications," *IEEE Microwave Guided Wave Lett.*, Vol. 6, 1996, pp. 177–179.

[18] 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 Guided Wave Lett.*, Vol. 4, 1994, pp. 344–345.

[19] Chen, J. C., and K. Li, "Quartic perfectly matched layers for dielectric waveguides and gratings," *Microwave Optical Tech. Lett.*, Vol. 10, 1995, pp. 319–323.

[20] Wu, Z., and J. Fang, "Numerical implementation and performance of perfectly matched layer boundary condition for waveguide structures," *IEEE Trans. Microwave Theory Tech.*, Vol. 43, 1995, pp. 2676–2683.

[21] Fang, J., and Z. Wu, "Generalized perfectly matched layer—an extension of Berenger's perfectly matched layer boundary condition," *IEEE Microwave Guided Wave Lett.*, Vol. 5, 1995, pp. 451–453.

[22] Bernabeu Verdu, J., R. Gillard, K. Moustadir, and J. Citerne, "An extension of the PML technique to the FDTD analysis of multilayer planar circuits and antennas," *Microwave Optical Tech. Lett.*, Vol. 10, 1995, pp. 323–327.

[23] Pasik, M. F., G. Aguirre, and A. C. Cangellaris, "Application of the PML absorbing boundary condition to the analysis of patch antennas," *Electromagnetics,* Vol. 16, 1996, pp. 435–449.

[24] Gedney, S. D., "An anisotropic PML absorbing media for FDTD simulation of fields in lossy dispersive media," *Electromagnetics,* Vol. 16, 1996, pp. 399–415.

[25] Lau, Y. C., M. S. Leong, and P. S. Kooi, "Extension of Berenger's PML boundary condition in matching lossy medium and evanescent waves," *Electronics Lett.*, Vol. 32, 1996, pp. 974–976.

[26] Fang, J., and Z. Wu, "Generalized perfectly matched layer for the absorption of propagating and evanescent waves in lossless and lossy media," *IEEE Trans. Microwave Theory Tech.*, Vol. 14, 1996, pp. 2216–2222.

[27] Liu, Q. H., "An FDTD algorithm with perfectly matched layers for conductive media," *Microwave Optical Tech. Lett.*, Vol. 14, 1997, pp. 134–137.

[28] Roden, J. A., S. D. Gedney, M. P. Kesler, J. G. Maloney, and P. H. Harms, "Time-domain analysis of periodic structures at oblique incidence: Orthogonal and nonorthogonal FDTD implementations," *IEEE Trans. Microwave Theory Tech.*, Vol. 46, 1998, pp. 420–427.

[29] Fang, J., and Z. Wu, "Closed-form expression of numerical reflection coefficient at PML interfaces and optimization of PML performance," *IEEE Microwave Guided Wave Lett.*, Vol. 6, 1996, pp. 332–334.

[30] Sullivan, D. M., "A simplified PML for use with the FDTD method," *IEEE Microwave Guided Wave Lett.*, Vol. 6, 1996, pp. 97–99.

[31] Sullivan, D. M., "An unsplit step three-dimensional PML for use with the FDTD," *IEEE Microwave Guided Wave Lett.*, Vol. 7, 1997, pp. 184–186.

[32] Rappaport, C. M., "Perfectly matched absorbing boundary conditions based on anisotropic lossy mapping of space," *IEEE Microwave Guided Wave Lett.*, Vol. 5, 1995, pp. 90–92.

[33] Teixeira, F. L., and W. C. Chew, "PML-FDTD in cylindrical and spherical coordinates," *IEEE Microwave Guided Wave Lett.*, Vol. 7, 1997, pp. 285–287.

[34] Teixeira, F. L., and W. C. Chew, "Analytical derivation of a conformal perfectly matched absorber for electromagnetic waves," *Microwave Optical Tech. Lett.*, Vol. 17, 1998, pp. 231–236.

[35] Teixeira, F. L. , K. P. Hwang, W. C. Chew, and J. M. Jin, "Conformal PML-FDTD schemes for electromagnetic field simulations: A dynamic stability study," *IEEE Trans. Antennas Propagat.*, Vol. 49, 2001, pp. 902–907.

[36] Roden, J. A., and S. D. Gedney, "Convolutional PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media," *Microwave Optical Tech. Lett.*, Vol. 27, 2000, pp. 334–339.

[37] Berenger, J.-P., "Numerical reflection from FDTD PMLs: A comparison of the split PML with the unsplit and CFS PMLs," *IEEE Trans. Antennas Propagat.*, Vol. 50, 2002, pp. 258–265.

[38] 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 Propagat.*, Vol. 43, 1995, pp. 1460–1463.

[39] Gedney, S. D., "An anisotropic perfectly matched layer absorbing media for the truncation of FDTD lattices," *IEEE Trans. Antennas Propagat.*, Vol. 44, 1996, pp. 1630–1639.

[40] Ziolkowski, R. W., "Time-derivative Lorentz materials and their utilization as electromagnetic absorbers," *Phys. Rev. E,* Vol. 55, 1997, pp. 7696–7703.

[41] Teixeira, F. L., "On aspects of the physical realizability of perfectly matched absorbers for electromagnetic waves," *Radio Science*, Vol. 38, 2003.

[42] Berenger, J. P., "Calcul de la diffraction a l'aide d'une méthode aux différence finies," Actes du colloque CEM, CNFRS-URSI, Trégastel France, June 1983.

[43] Gedney, S., "The perfectly matched layer absorbing medium," Ch. 5 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Taflove, A., (ed.), Norwood, MA, Artech House, 1998.

[44] Zhao, A. P., "Generalized-material-independent PML absorbers used for the FDTD simulation of electromagnetic waves in 3D arbitrary anisotropic dielectric and magnetic media," *IEEE Trans. Microwave Theory Tech.*, Vol. 46, 1998, pp. 1511–1513.

[45] Teixeira, F. L., and W. C. Chew, "A general approach to extend Berenger's absorbing boundary condition to anisotropic and dispersive media," *IEEE Trans. Antennas Propagat.*, Vol. 46, 1998, pp. 1386–1387.

[46] Teixeira, F. L., and W. C. Chew, "General closed-form PML constitutive tensors to match arbitrary bianisotropic and dispersive linear media," *IEEE Microwave Guided Wave Lett.*, Vol. 8, 1998, pp. 223–225.

[47] Teixeira, F. L., W. C. Chew, M. L. Oristaglio, and T. Wang, "Perfectly matched layer and piecewise-linear recursive convolution for the FDTD solution of the 3D dispersive half- space problem," *IEEE Trans. Magnetics*, Vol. 34, 1998, pp. 2747–2750.

[48] Teixeira, F. L., W. C. Chew, M. Straka, M. L. Oristaglio, and T. Wang, "Finite-difference time-domain simulation of ground penetrating radar on dispersive, inhomogeneous, and conductive soils," *IEEE Trans. Geoscience Remote Sensing*, Vol. 36, 1998, pp. 1928–1937.

[49] Zhao, A. P., "Application of the material-independent PML absorbers to the FDTD analysis of electromagnetic waves in nonlinear media," *Microwave Optical Tech. Lett.*, Vol. 17, 1998, pp. 164–168.

[50] Kuzuoglu, M. and R. Mittra, "Frequency dependence of the constitutive parameters of causal perfectly matched anisotropic absorbers," *IEEE Microwave Guided Wave Lett.*, Vol. 6, 1996, pp. 447–449.

[51] E. Becache, P. G. Petropoulos, and S. D. Gedney, "On the long-time behavior of unsplit perfectly matched layers," *IEEE Trans. Antennas Propagat.*, Vol. 52, 2004, pp. 1335–1342.

[52] Abarbanel, S., and D. Gottlieb, "A mathematical analysis of the PML method," *J. Computational Physics,* Vol. 134, 1997, pp. 357–363.

[53] Teixeira, F. L., and W. C. Chew, "On causality and dynamic stability of perfectly matched layers for FDTD simulations," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, 1999, pp. 775–785.

[54] Luebbers, R. J., and F. Hunsberger, "FDTD for Nth-order dispersive media," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 1297–1301.

[55] Wang, C. F., J. M. Jin, L. W. Li, P. S. Kooi, and M. S. Leong, "A multilevel BCG-FFT method for the analysis of a microstrip antenna and array," *Microwave Optical Tech. Lett.,* vol. 27, pp. 20–23.

[56] Hagness, S. C., A. Taflove, and J. E. Bridges, "Three-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: Design of an antenna-array element," *IEEE Trans. Antennas Propagat.,* Vol. 37, 1999, pp. 783–791.

[57] Converse, M., E. J. Bond, S. C. Hagness, and B. D. Van Veen, "Ultrawide-band microwave space-time beamforming for hyperthermia treatment of breast cancer: A computational feasibility study," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, 2004, pp. 1876–1889.

PROJECTS

P7.1 Implement Berenger's split-field PML ABC in a two-dimensional TE_z FDTD grid. Use a cubic increase of loss with depth within the PML, and have the PML thickness and $R(0)$ as user-specified parameters. Compare the local and global error to the best that was obtained with the analytical ABCs implemented in Chapter 6.

P7.2 Implement the CPML ABC in the two-dimensional TE_z FDTD grid of Fig. 7.2. Replicate the results graphed in Fig. 7.3. Verify the optimum choices of CPML parameters shown in the contour plots of Fig. 7.4.

P7.3 Implement the CPML ABC in a three-dimensional FDTD space lattice for the PEC plate problem of Fig. 7.5, and replicate the results graphed in Fig. 7.6. Verify the optimum choices of CPML parameters shown in the contour plots of Fig. 7.7.

P7.4 Implement the CPML ABC in a three-dimensional FDTD space lattice for the microstrip patch antenna array problem of Fig. 7.8, and replicate the results graphed in Fig. 7.9.

P7.5 Implement the CPML ABC in a three-dimensional FDTD space lattice for the bow-tie antenna breast-hyperthermia problem of Fig. 7.10. Verify the optimum choices of CPML parameters shown in the contour plots of Fig. 7.11.

Chapter 8

Near-to-Far-Field Transformation

Allen Taflove, Xu Li, and Susan Hagness

8.1 INTRODUCTION

The provision of a well-defined scattered-field region in the zoned FDTD space lattice, as described in Chapter 5, permits a systematic *near-to-far-field* (NTFF) transformation, which is derived in this chapter. Using the near-field data obtained in a single FDTD modeling run, this transformation efficiently and accurately calculates the complete far-field bistatic scattering response of an illuminated structure for a single illumination angle, or the complete radiation pattern of an antenna. There is *no* need to extend the FDTD space lattice to the far field.

The discussion begins in the phasor domain with the two-dimensional TM_z mode. Using Green's theorem, we show that scattered or radiated $\boldsymbol{E}$ and $\boldsymbol{H}$ fields that are tangential to a closed virtual surface containing the scatterer or antenna of interest can be integrated to provide the far-field response. The virtual surface is independent of the nature of the structure being modeled, and can have a fixed rectangular shape to conform with a Cartesian FDTD grid. This yields the powerful surface equivalence theorem in two dimensions, which is subsequently extended to the general three-dimensional case. The resulting analytical NTFF expressions have been implemented in many FDTD codes. Next, we derive the time-domain NTFF transformation, which permits a direct computation of scattered or radiated field-versus-time waveforms. Finally, we conclude with a recent simple modification of the NTFF procedure that greatly improves the accuracy in calculating the backscatter from strongly forward-scattering objects such as biological cells illuminated by light, and certain types of low-observable vehicles.

8.2 TWO-DIMENSIONAL TRANSFORMATION, PHASOR DOMAIN

Consider the two-dimensional TM_z scattering or radiation geometry of Fig. 8.1, wherein the scattered or radiated field phasors $\breve{E}_z$, $\breve{H}_x$, and $\breve{H}_y$ are involved. We assume that an arbitrary structure of interest is enclosed by the arbitrary contour C_a having the unit outward normal vector $\hat{\boldsymbol{n}}_a$. Further, we assume the existence of the infinite-radius circular contour C_∞ centered at the coordinate system origin, and having the unit outward normal vector $\hat{\boldsymbol{n}}_\infty = \hat{\boldsymbol{r}}$. Finally, we assume that C_a and C_∞ are connected by an infinitely thin "umbilical" path to form a single continuous, closed contour that encloses the surface S. A positive (counterclockwise or right-hand) direction is assumed about this combined contour, so that the surface unit normal $d\hat{\boldsymbol{S}}$ points in the "thumb" direction. Note that $\hat{\boldsymbol{n}}_\infty$ and $\hat{\boldsymbol{n}}_a$ have opposite orientations relative to S.

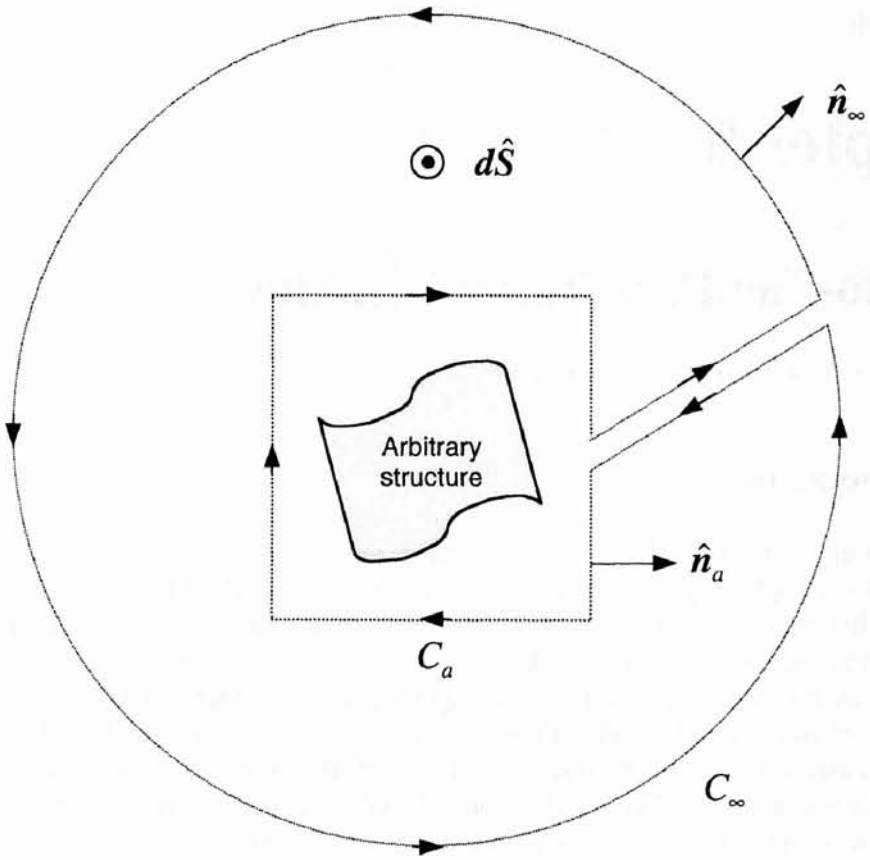


Fig. 8.1 Two-dimensional scattering or radiation geometry showing the integration contours used in the derivation of the near-to-far-field transformation.

8.2.1 Application of Green's Theorem

By Green's theorem [1] applied to the scalar functions $E_z(\boldsymbol{r})$ and $G(\boldsymbol{r}|\boldsymbol{r}')$, we have

$$\int_S \left[\breve{E}_z(\boldsymbol{r}') \left(\nabla^2\right)' G(\boldsymbol{r}|\boldsymbol{r}') - G(\boldsymbol{r}|\boldsymbol{r}') \left(\nabla^2\right)' \breve{E}_z(\boldsymbol{r}') \right] ds'$$

$$= \oint_{C_\infty} \left[\breve{E}_z(\boldsymbol{r}') \frac{\partial G(\boldsymbol{r}|\boldsymbol{r}')}{\partial r'} - G(\boldsymbol{r}|\boldsymbol{r}') \frac{\partial \breve{E}_z(\boldsymbol{r}')}{\partial r'} \right] dC'$$

$$- \oint_{C_a} \left[\breve{E}_z(\boldsymbol{r}') \frac{\partial G(\boldsymbol{r}|\boldsymbol{r}')}{\partial n_a'} - G(\boldsymbol{r}|\boldsymbol{r}') \frac{\partial \breve{E}_z(\boldsymbol{r}')}{\partial n_a'} \right] dC' \tag{8.1}$$

where $\boldsymbol{r}$ is an observation point in the two-dimensional space, $\boldsymbol{r}'$ is a source point, and dC' is a differential path element along the combined C_a and C_∞ source contour. The negative sign before the second integral term on the right-hand side of (8.1) results from the opposite orientations of $\hat{\boldsymbol{n}}_\infty = \hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{n}}_a$ relative to S.

Let us consider the various integral terms of (8.1). For the C_∞ term, we can show that both $\breve{E}_z(\boldsymbol{r}')$ and $G(\boldsymbol{r}|\boldsymbol{r}')$ decay as $1/\sqrt{r'}$ in two dimensions as $r' \to \infty$. Using this information, we find that the contribution of the C_∞ integral term is zero in the limit as $r' \to \infty$:

$$\begin{aligned} \int_{C_\infty} &\approx \lim_{r'\to\infty}\left[2\pi r' \cdot \frac{1}{\sqrt{r'}} \cdot \frac{\partial(1/\sqrt{r'})}{\partial r'}\right] \\ &\approx r' \cdot \frac{1}{(r')^{1/2}} \cdot \frac{1}{(r')^{3/2}} \approx \frac{r'}{(r')^2} \to 0 \end{aligned} \tag{8.2}$$

Now consider the S integral term evaluated at an observation point $\boldsymbol{r}$ in S. From the definition of the Green function for time-harmonic systems, we have

$$\left(\nabla^2\right)' G(\boldsymbol{r}|\boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}') - k^2 G(\boldsymbol{r}|\boldsymbol{r}') \tag{8.3}$$

where δ is the Dirac delta function and k is the wavenumber. Further, from the Helmholtz equation, we have

$$\left(\nabla^2\right)' \breve{E}_z(\boldsymbol{r}') = -k^2 \breve{E}_z(\boldsymbol{r}') \tag{8.4}$$

Substituting (8.3) and (8.4) into the S integral term, we obtain

$$\begin{aligned} &\int_S \left\{ \breve{E}_z(\boldsymbol{r}') \cdot \left[\delta(\boldsymbol{r} - \boldsymbol{r}') - k^2 G(\boldsymbol{r}|\boldsymbol{r}')\right] - G(\boldsymbol{r}|\boldsymbol{r}') \cdot \left[-k^2 \breve{E}_z(\boldsymbol{r}')\right] \right\} ds' \\ &\qquad = \int_S \breve{E}_z(\boldsymbol{r}')\delta(\boldsymbol{r} - \boldsymbol{r}')\, ds' = \breve{E}_z(\boldsymbol{r}) \end{aligned} \tag{8.5}$$

Thus, (8.1) simplifies to

$$\begin{aligned} \breve{E}_z(\boldsymbol{r}) &= \oint_{C_a} \left[G(\boldsymbol{r}|\boldsymbol{r}') \frac{\partial \breve{E}_z(\boldsymbol{r}')}{\partial n_a'} - \breve{E}_z(\boldsymbol{r}') \frac{\partial G(\boldsymbol{r}|\boldsymbol{r}')}{\partial n_a'} \right] dC' \\ &= \oint_{C_a} \left[G(\boldsymbol{r}|\boldsymbol{r}')\, \hat{\boldsymbol{n}}_a' \cdot \nabla' \breve{E}_z(\boldsymbol{r}') - \breve{E}_z(\boldsymbol{r}')\, \hat{\boldsymbol{n}}_a' \cdot \nabla' G(\boldsymbol{r}|\boldsymbol{r}') \right] dC' \end{aligned} \tag{8.6}$$

8.2.2 Far-Field Limit

Now, consider the analytical form of the Green function. In two dimensions, $G(\boldsymbol{r}|\boldsymbol{r}')$ is given by the Hankel function expression

$$G(\boldsymbol{r}|\boldsymbol{r}') = \frac{j}{4} H_0^{(2)}(k|\boldsymbol{r}-\boldsymbol{r}'|) \tag{8.7}$$

Consider Fig. 8.2, which depicts an observation point $\boldsymbol{P} = \boldsymbol{r}$ in the far field many wavelengths away from any source points on C_a. Clearly, $\boldsymbol{P}$ is located such that $k|\boldsymbol{r}-\boldsymbol{r}'|$ is very large. For this case, it can be shown that the limiting expression for $G(\boldsymbol{r}|\boldsymbol{r}')$ is

$$\lim_{k|\boldsymbol{r}-\boldsymbol{r}'|\to\infty} G(\boldsymbol{r}|\boldsymbol{r}') = \frac{j^{3/2}}{\sqrt{8\pi k}} \frac{e^{-jk|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|^{1/2}} \tag{8.8}$$

Applying the law of cosines to the geometry of Fig. 8.2, we obtain for $|\boldsymbol{r}-\boldsymbol{r}'|^2$:

$$\begin{aligned} |\boldsymbol{r}-\boldsymbol{r}'|^2 &= |\boldsymbol{r}|^2 + |\boldsymbol{r}'|^2 - 2|\boldsymbol{r}||\boldsymbol{r}'|\cos(\phi-\phi') \\ &= |\boldsymbol{r}|^2 \cdot \left[1 - 2\frac{|\boldsymbol{r}'|}{|\boldsymbol{r}|}\cos(\phi-\phi') + \frac{|\boldsymbol{r}'|^2}{|\boldsymbol{r}|^2}\right] \end{aligned} \tag{8.9}$$

Taking the square root of both sides of (8.9), and then expanding the resulting right-hand side in a one-term binomial expansion (assuming $|\boldsymbol{r}'|/|\boldsymbol{r}| \ll 1$), yields

$$|\boldsymbol{r}-\boldsymbol{r}'| \cong |\boldsymbol{r}| \cdot \left[1 - \frac{|\boldsymbol{r}'|}{|\boldsymbol{r}|}\cos(\phi-\phi') + \frac{1}{2}\frac{|\boldsymbol{r}'|^2}{|\boldsymbol{r}|^2}\right] \cong r - r'\cos(\phi-\phi') \tag{8.10}$$

Here, for convenience, we denote $|\boldsymbol{r}|$ and $|\boldsymbol{r}'|$ as simply r and r', respectively. Repeating this square-root procedure on the results of (8.10), we obtain

$$|\boldsymbol{r}-\boldsymbol{r}'|^{1/2} \cong \left\{r\left[1 - \frac{r'}{r}\cos(\phi-\phi')\right]\right\}^{1/2} \cong r^{1/2}\left[1 - \frac{1}{2}\frac{r'}{r}\cos(\phi-\phi')\right] \cong r^{1/2} \tag{8.11}$$

Substituting the results of (8.10) and (8.11) into the Green function expression of (8.8) provides a more convenient limiting form of $G(\boldsymbol{r}|\boldsymbol{r}')$:

$$\lim_{k|\boldsymbol{r}-\boldsymbol{r}'|\to\infty} G(\boldsymbol{r}|\boldsymbol{r}') = \frac{j^{3/2}}{\sqrt{8\pi k r}} e^{-jk[r-r'\cos(\phi-\phi')]} = \frac{j^{3/2}}{\sqrt{8\pi k r}} e^{-jkr} e^{+jk\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} \tag{8.12}$$

With $G(\boldsymbol{r}|\boldsymbol{r}')$ written in this manner, it is now clear that

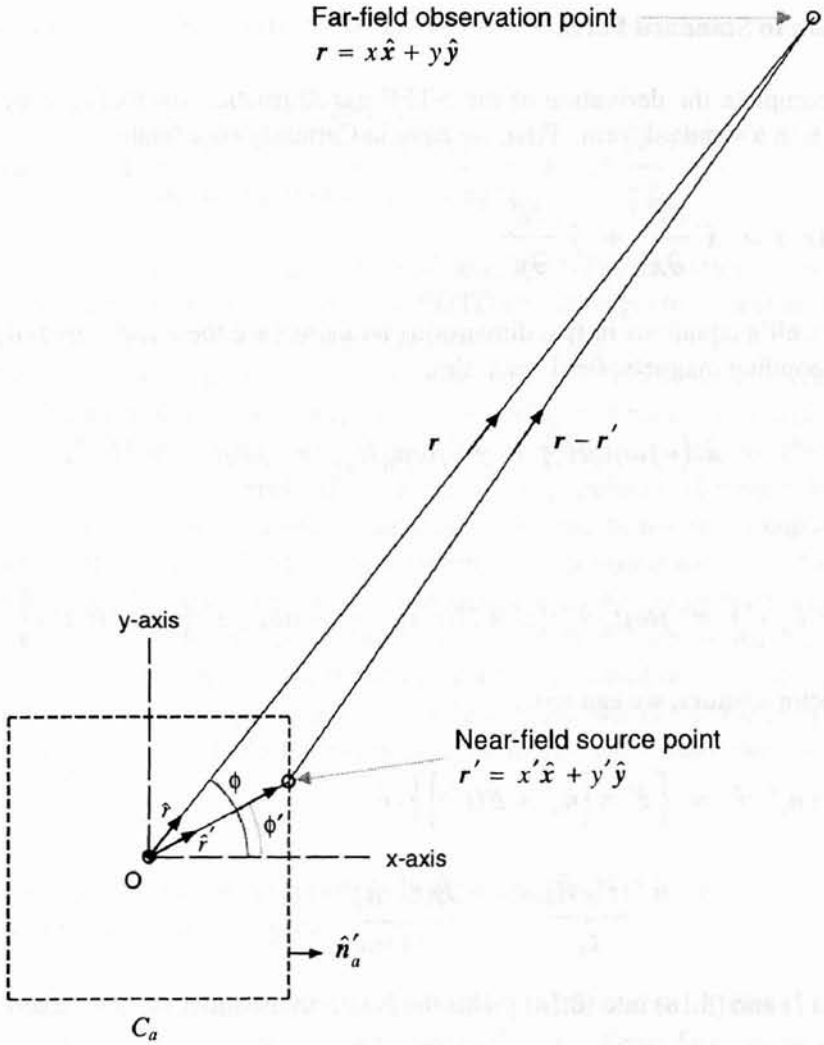


Fig. 8.2 Geometry of the far-field observation point relative to the near-field integration contour.

$$\lim_{k|\boldsymbol{r}-\boldsymbol{r}'|\to\infty} \nabla' G(\boldsymbol{r}|\boldsymbol{r}') = (jk\hat{\boldsymbol{r}}) \frac{j^{3/2}}{\sqrt{8\pi k r}} e^{-jkr} e^{+jk\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} \tag{8.13}$$

Substituting (8.12) and (8.13) into (8.6) results in the following expression for the far-field E_z:

$$\lim_{k|\boldsymbol{r}-\boldsymbol{r}'|\to\infty} \breve{E}_z(\boldsymbol{r}) = \frac{j^{3/2}}{\sqrt{8\pi k r}} e^{-jkr} \oint_{C_a} \left[\begin{array}{l} e^{+jk\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} \hat{\boldsymbol{n}}_a' \cdot \nabla' \breve{E}_z(\boldsymbol{r}') \\ - \breve{E}_z(\boldsymbol{r}') \hat{\boldsymbol{n}}_a' \cdot (jk\hat{\boldsymbol{r}}) e^{+jk\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} \end{array} \right] dC'$$

$$= \frac{e^{j3\pi/4}}{\sqrt{8\pi k r}} e^{-jkr} \oint_{C_a} \left[\hat{\boldsymbol{n}}_a' \cdot \nabla' \breve{E}_z(\boldsymbol{r}') - jk \breve{E}_z(\boldsymbol{r}') \hat{\boldsymbol{n}}_a' \cdot \hat{\boldsymbol{r}} \right] e^{+jk\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} dC' \tag{8.14}$$

8.2.3 Reduction to Standard Form

We will now complete the derivation of the NTFF transformation for this case by manipulating (8.14) to place it in a standard form. First, we have in Cartesian coordinates:

$$\nabla' \breve{E}_z(\boldsymbol{r}') = \hat{x}' \frac{\partial \breve{E}_z}{\partial x'} + \hat{y}' \frac{\partial \breve{E}_z}{\partial y'} \tag{8.15}$$

However, Maxwell's equations in two dimensions let us replace the x and y partial derivatives of E_z with corresponding magnetic field quantities:

$$\nabla' \breve{E}_z(\boldsymbol{r}') = \hat{x}' \left(-j\omega\mu_0 \breve{H}_y\right) + \hat{y}' \left(j\omega\mu_0 \breve{H}_x\right) = j\omega\mu_0 \, \hat{z}' \times \breve{\boldsymbol{H}}(\boldsymbol{r}') \tag{8.16}$$

Then, it follows that

$$\hat{\boldsymbol{n}}_a' \cdot \nabla' \breve{E}_z(\boldsymbol{r}') = j\omega\mu_0 \, \hat{\boldsymbol{n}}_a' \cdot \left[\hat{z}' \times \breve{\boldsymbol{H}}(\boldsymbol{r}')\right] = -j\omega\mu_0 \, \hat{z}' \cdot \left[\hat{\boldsymbol{n}}_a' \times \breve{\boldsymbol{H}}(\boldsymbol{r}')\right] \tag{8.17}$$

Exploiting a vector identity, we can write

$$\begin{aligned} \breve{E}_z(\boldsymbol{r}') \, \hat{\boldsymbol{n}}_a' \cdot \hat{\boldsymbol{r}} &= \left\{ \hat{z}' \times \left[\hat{\boldsymbol{n}}_a' \times \breve{\boldsymbol{E}}(\boldsymbol{r}')\right] \right\} \cdot \hat{\boldsymbol{r}} \\ &= \hat{\boldsymbol{n}}_a' \underbrace{(\hat{z}' \cdot \breve{\boldsymbol{E}})}_{\breve{E}_z} \cdot \hat{\boldsymbol{r}} - \underbrace{\breve{\boldsymbol{E}}(\hat{z}' \cdot \hat{\boldsymbol{n}}_a')}_{=0 \text{ here}} \cdot \hat{\boldsymbol{r}} \end{aligned} \tag{8.18}$$

Substituting (8.17) and (8.18) into (8.14) yields the NTFF transformation in standard form:

$$\begin{aligned} \lim_{k|\boldsymbol{r}-\boldsymbol{r}'|\to\infty} \breve{E}_z(\boldsymbol{r}) &= \frac{e^{-jkr}}{\sqrt{r}} \frac{e^{j(\pi/4)}}{\sqrt{8\pi k}} \oint_{C_a} \left\{ \begin{array}{l} \omega\mu_0 \, \hat{z}' \cdot \left[\hat{\boldsymbol{n}}_a' \times \breve{\boldsymbol{H}}(\boldsymbol{r}')\right] \\ + \, k \, \hat{z}' \times \left[\hat{\boldsymbol{n}}_a' \times \breve{\boldsymbol{E}}(\boldsymbol{r}')\right] \cdot \hat{\boldsymbol{r}} \end{array} \right\} e^{+jk\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} dC' \\ &= \frac{e^{-jkr}}{\sqrt{r}} \frac{e^{j(\pi/4)}}{\sqrt{8\pi k}} \oint_{C_a} \left[\begin{array}{l} \omega\mu_0 \, \hat{z}' \cdot \breve{\boldsymbol{J}}_{\text{eq}}(\boldsymbol{r}') \\ - \, k \, \hat{z}' \times \breve{\boldsymbol{M}}_{\text{eq}}(\boldsymbol{r}') \cdot \hat{\boldsymbol{r}} \end{array} \right] e^{+jk\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} dC' \end{aligned} \tag{8.19}$$

where $\breve{\boldsymbol{J}}_{\text{eq}} \equiv \hat{\boldsymbol{n}}_a \times \breve{\boldsymbol{H}}$ and $\breve{\boldsymbol{M}}_{\text{eq}} \equiv -\hat{\boldsymbol{n}}_a \times \breve{\boldsymbol{E}}$ are defined as the phasor tangential *equivalent electric and magnetic currents* observed at C_a. If we now identify a complex-valued pattern function $F(\phi)$:

$$F(\phi) \equiv \frac{e^{j(\pi/4)}}{\sqrt{8\pi k}} \oint_{C_a} \left[\omega\mu_0 \, \hat{z}' \cdot \breve{\boldsymbol{J}}_{\text{eq}}(\boldsymbol{r}') - k \, \hat{z}' \times \breve{\boldsymbol{M}}_{\text{eq}}(\boldsymbol{r}') \cdot \hat{\boldsymbol{r}}\right] e^{+jk\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} \, dC' \tag{8.20}$$

then the bistatic *radar cross section* (RCS) in two dimensions is defined as

$$\mathrm{RCS}(\phi) \equiv 2\pi \cdot \frac{\text{power scattered per unit angle in direction } \hat{r}}{\text{incident power per unit length}} \equiv 2\pi \frac{|F(\phi)|^2}{|\breve{E}_{\mathrm{inc}}|^2} \tag{8.21}$$

having the dimension of meters. As an historical note, the NTFF transformation of (8.19) was first published in the context of two-dimensional FDTD electromagnetic scattering simulations in [2], and was extended to three-dimensional FDTD simulations in [3].

It should be emphasized that C_a is *not* a physical surface. It is a *virtual surface* that is the locus of points in space where E- and H-field data are being compiled and integrated. Because C_a can have an arbitrary shape, we can conveniently assign it to lie along a rectangle in the scattered-field zone of the FDTD grid. If this rectangle is populated with time-domain E_z field components, the complex phasor values $\breve{E}_z$ of these components are first obtained via a DFT conducted concurrently with the FDTD time-stepping, as discussed below. Then, the phasor E-field data can be used directly in the numerical calculation of the integrals of (8.19) and (8.20), which can be implemented by any convenient integration rule (such as the trapezoidal rule). For this case, the time-domain tangential H-field components located one-half space cell to either side of C_a (due to the normal staggering of the field components in the FDTD grid) would be averaged at each time-step to obtain the necessary phasor H-field values directly on C_a, before undergoing the DFT.

8.3 OBTAINING PHASOR QUANTITIES VIA DISCRETE FOURIER TRANSFORMATION

The E- and H-field data used in the near-to-far-field transformation of (8.19) and (8.20) are phasor quantities. At each field point on the virtual surface C_a, these data can be efficiently and concurrently obtained for multiple frequencies with only one FDTD run. We need only provide an impulsive wideband electromagnetic excitation of the structure of interest, and perform a recursive DFT "on the fly" (i.e., concurrently with the FDTD time-stepping) for each frequency of interest. Computer storage is quite reasonable, with only two numbers (i.e., the field magnitude and phase) required to store the DFT results for each frequency at each field point on the virtual surface. Therefore, a *single* FDTD run can generate the complete far-field distribution of a structure (i.e., its bistatic RCS pattern or its radiation pattern) at *many* frequencies.

The formulation of the wideband excitation and the required DFT is best illustrated by considering the following annotated Fortran code segments:

```
dimension EZEFR(ib,nfreq,2), OMEGA(nfreq), RD(nfreq,2), SN(nfreq,2)
```

EZEFR stores the quadrature components of the phasor values $\breve{E}_z$ at ib grid locations along the front line of the rectangular near-to-far field contour C_a for each of nfreq frequencies. At each gridpoint i and frequency index number nf, EZEFR(i,nf,1) contains the real-valued amplitude of the sine component of the phasor and EZEFR(i,nf,2) contains the real-valued amplitude of the cosine component.

OMEGA stores the nfreq angular frequencies in the DFT at which the phasor values $\breve{E}_z$ are evaluated.

RD(nf,1) and RD(nf,2) store, respectively, the sine and cosine multipliers used in the DFT at frequency index number nf.

SN(nf,1) and SN(nf,2) store, respectively, the sine and cosine quadrature components of the incident field source phasor at frequency index number nf.

⋮

```
data ib, nfreq, OMEGA/ ... /, nmax, omegacenter, tau
```

ib defines the number of E_z components located along the front line of the rectangular near-to-far-field contour.

nfreq defines the number of frequency points in the DFT.

OMEGA specifies the nfreq angular frequencies in the DFT at which the phasor values $\breve{E}_z$ are evaluated.

nmax defines the maximum number of time-steps in the FDTD solution.

omegacenter defines the center frequency of the source spectrum.

tau defines the $1/e$ decay time of the source pulse.

⋮

```
    do 1 i=1,ib ;            nf=1,nfreq ;        k=1,2
1   EZEFR(i,nf,k) = 0.0

    do 2 nf=1,nfreq ;        k=1,2
2   SN(nf,k) = 0.0
```

The electric field and incident field source phasors are initialized to zero.

⋮

```
do 500 n=1,nmax
esource = exp(-(nΔt-3*tau)**2/tau**2) * sin(omegacenter*(nΔt-3*tau))
```

The main FDTD time-stepping loop is initiated, and the source waveform is evaluated for the current time-step. Note that the Gaussian pulse envelope of the source waveform is essentially zero at the initiation of time-stepping, and that the sinusoidal carrier has a zero-crossing at the peak of the Gaussian envelope nΔt = 3*tau. This causes esource to have odd symmetry in time about the peak of the envelope function, ensuring a spectrum concentrated about omegacenter and zero dc component. It is clear that adjusting tau and omegacenter permits specifying the position and width of the source spectrum.

⋮

Here the program implements the main time-stepping loops to update the fields everywhere in the FDTD grid.

⋮

```
do 110 nf=1,nfreq
RD(nf,1) = sin(OMEGA(nf)*nΔt)
RD(nf,2) = cos(OMEGA(nf)*nΔt)
```

The sine and cosine multipliers used in the DFT are calculated at each frequency of interest.

```
do 110 k=1,2
SN(nf,k) = SN(nf,k) + esource*RD(nf,k)
```

The sine and cosine quadrature components of the incident field source phasor SN are calculated using a simple recursive sum.

```
do 110 i=1,ib
EZEFR(i,nf,k) = EZEFR(i,nf,k) + E_z|_{i,j_front} *RD(nf,k)
```

The sine and cosine quadrature components of the electric field phasor EZEFR corresponding to each E_z component along the front line of the near-to-far-field contour, are calculated using a simple recursive sum. Note that this sum is updated concurrently with the FDTD time-stepping of the fields.

```
110 continue
```

⋮

The remainder of the main FDTD time-stepping loop is completed, including all other recursive sums implementing DFTs.

⋮

```
500 continue
```

Upon completion of the do 500 main FDTD time-stepping loop, there is also a termination of the recursive sums for the sine and cosine quadrature components of the incident field source phasor SN and the recursive sums for the quadrature components of the electric field phasor EZEFR.

⋮

```
do 600 nf=1,nfreq
store = sqrt(SN(nf,1)**2 + SN(nf,2)**2)
SN(nf,2) = -atan2(SN(nf,1), SN(nf,2))
SN(nf,1) = store
```

The final values of the recursive sums for SN are processed at each frequency in the DFT to yield the corresponding magnitude and phase of the source phasor stored in SN(nf,1) and SN(nf,2), respectively.

```
do 600 i=1,ib
store = sqrt(EZEFR(i,nf,1)**2 + EZEFR(i,nf,2)**2) / SN(nf,1)
EZEFR(i,nf,2) = -atan2(EZEFR(i,nf,1), EZEFR(i,nf,2)) - SN(nf,2)
EZEFR(i,nf,1) = store
```

The final values of the recursive sums for EZEFR are processed at each frequency in the DFT to yield the corresponding magnitude and phase of the source phasor stored in EZEFR(i,nf,1) and EZEFR(i,nf,2), respectively. In addition, the EZEFR values are normalized relative to the magnitude and phase of the incident field phasor at the corresponding frequency, just calculated above. The EZEFR data are now ready for insertion into (8.19) or (8.20) to permit computation of the far-field pattern at each of the nfreq frequencies that were tracked during FDTD time-stepping.

```
600 continue
```

The principal caution in this entire procedure is that nmax must be carefully chosen so that the impulse response of the structure of interest "rings down" to zero before the end of time-stepping at each observation point along the near-to-far-field virtual surface. Failure to decay the impulse response to zero causes inaccuracy of the computed DFT due to windowing of the true impulse response. To mitigate windowing effects for certain classes of structures that require many thousands or tens of thousands of time-steps to ring down, several techniques have been proposed to extrapolate the late-time impulse response from a windowed response before the DFT is applied [4, 5]. These are discussed in Chapter 15, Section 15.9, which reviews digital signal processing and spectrum-estimation techniques as applied to FDTD time-waveform data.

8.4 SURFACE EQUIVALENCE THEOREM

The implication of (8.19) is that knowledge of the equivalent electric and magnetic currents tangential to any closed contour surrounding a two-dimensional TM_z electromagnetic wave interaction structure is sufficient to obtain the far field via a simple integration of these currents around the contour. No other sources appear. In fact, we can think of the region within the observation contour as being *source-free* and *field-free*. This idea, illustrated in Fig. 8.3, forms the basis of the surface equivalence theorem [6, 7], a powerful concept in electromagnetic theory wherein actual sources such as an antenna are replaced by equivalent sources. Quoting from [7],

> "By the surface equivalence theorem, the fields outside an imaginary closed surface are obtained by placing over the closed surface suitable electric and magnetic current densities that satisfy the boundary conditions. The current densities are selected so that the fields inside the closed surface are zero and outside are equal to the radiation produced by the actual sources. Thus, the technique can be used to obtain the fields radiated outside a closed surface by sources enclosed within it. The formulation is exact but requires integration over the closed surface. The degree of accuracy depends on the knowledge of the tangential components of the fields over the closed surface."

Consider now the meaning of Fig. 8.3. Fig. 8.3(a) depicts the most general case of an electromagnetic wave interaction with an arbitrary three-dimensional structure. Following [7], we assume that $(\boldsymbol{E}_1, \boldsymbol{H}_1)$ filling all of space is generated by the action of physical electric and magnetic current sources $\boldsymbol{J}_1$ and $\boldsymbol{M}_1$ flowing on the surface of the structure. In Fig. 8.3(b), we assume that $\boldsymbol{J}_1$ and $\boldsymbol{M}_1$ are removed, and that a new field $(\boldsymbol{E}, \boldsymbol{H})$ now exists inside an arbitrary closed observation surface S that completely encloses the structure. However, we wish to observe the original field $(\boldsymbol{E}_1, \boldsymbol{H}_1)$ outside S. For this situation to satisfy the electromagnetic field boundary conditions on the tangential E- and H-field components at S, the following nonphysical electric and magnetic currents flowing tangentially along S must exist:

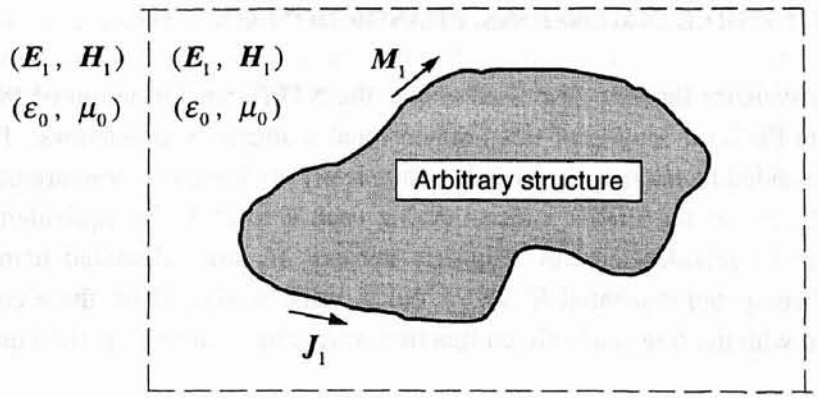
$$\boldsymbol{J}_S = \hat{\boldsymbol{n}} \times (\boldsymbol{H}_1 - \boldsymbol{H}) \; ; \qquad \boldsymbol{M}_S = -\hat{\boldsymbol{n}} \times (\boldsymbol{E}_1 - \boldsymbol{E}) \tag{8.22a,b}$$

where $\hat{\boldsymbol{n}}$ is the unit outward normal vector to S. The virtual electric and magnetic currents of (8.22) radiate into free space everywhere (inside and outside of S), and generate the original fields $(\boldsymbol{E}_1, \boldsymbol{H}_1)$ in the unbounded free-space region outside S. Since the fields within S can be anything (and we are not concerned with this region with regard to developing far-field information), it is useful to assume that $\boldsymbol{E}$ and $\boldsymbol{H}$ inside S are identically zero. Then, the equivalent problem of Fig. 8.3(b) reduces to Fig. 8.3(c), with the equivalent current densities

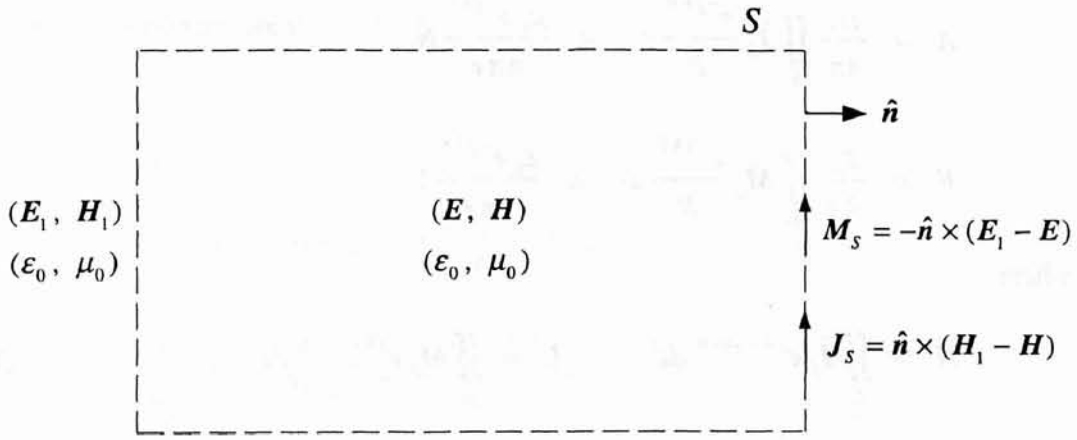
$$\boldsymbol{J}_S = \hat{\boldsymbol{n}} \times (\boldsymbol{H}_1 - \boldsymbol{H})\big|_{H=0} = \hat{\boldsymbol{n}} \times \boldsymbol{H}_1 \tag{8.23a}$$

$$\boldsymbol{M}_S = -\hat{\boldsymbol{n}} \times (\boldsymbol{E}_1 - \boldsymbol{E})\big|_{E=0} = -\hat{\boldsymbol{n}} \times \boldsymbol{E}_1 \tag{8.23b}$$

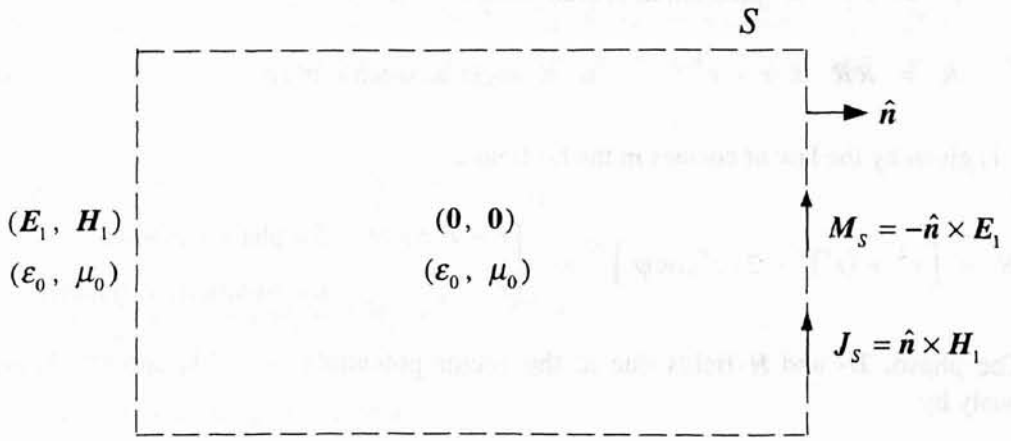
If applied in two dimensions, this concept of surface equivalent fields conforms with the rigorously derived integral expression of (8.19), which shows the far field being sourced by equivalent electric and magnetic currents tangential to a virtual contour completely enclosing the structure of interest.



(a) Original interaction geometry



(b) Intermediate equivalent problem



(c) Final equivalent problem

Fig. 8.3 Definition of electromagnetic fields and equivalent electric and magnetic virtual currents for the surface equivalence theorem. *After:* Balanis, *Advanced Engineering Electromagnetics*, Wiley, 1989.

8.5 EXTENSION TO THREE DIMENSIONS, PHASOR DOMAIN

Using the surface equivalence theorem discussed above, the NTFF transformation of (8.19) can be readily extended to FDTD modeling of three-dimensional scatterers and antennas. Here, the virtual surface is a six-sided rectangular locus S that completely encloses the structure of interest in the scattered-field zone of the FDTD lattice. Along each side of S, the equivalent phasor electric current $\breve{\boldsymbol{J}}_S$ and equivalent phasor magnetic current $\breve{\boldsymbol{M}}_S$ are calculated using DFTs applied to the FDTD-computed tangential H- and E-fields, respectively. Then, these equivalent currents are integrated with the free-space Green function weighting to obtain far-field quantities, as explained next.

Following the notation of [7], we can define a pair of vector potentials suitable for far-field computation:

$$\breve{\boldsymbol{A}} = \frac{\mu_0}{4\pi}\iint_S \breve{\boldsymbol{J}}_S \frac{e^{-jkR}}{R}\,ds' \cong \frac{\mu_0 e^{-jkr}}{4\pi r}\breve{\boldsymbol{N}} \tag{8.24}$$

$$\breve{\boldsymbol{F}} = \frac{\varepsilon_0}{4\pi}\iint_S \breve{\boldsymbol{M}}_S \frac{e^{-jkR}}{R}\,ds' \cong \frac{\varepsilon_0 e^{-jkr}}{4\pi r}\breve{\boldsymbol{L}} \tag{8.25}$$

where

$$\breve{\boldsymbol{N}} = \iint_S \breve{\boldsymbol{J}}_S e^{jkr'\cos\psi}\,ds' \quad ; \quad \breve{\boldsymbol{L}} = \iint_S \breve{\boldsymbol{M}}_S e^{jkr'\cos\psi}\,ds' \tag{8.26a, b}$$

$$\boldsymbol{r} = r\hat{\boldsymbol{r}} = \text{position of observation point } (x, y, z) \tag{8.26c}$$

$$\boldsymbol{r}' = r'\hat{\boldsymbol{r}}' \equiv \text{position of source point on } S\ (x', y', z') \tag{8.26d}$$

$$\boldsymbol{R} = R\hat{\boldsymbol{R}} \equiv \boldsymbol{r} - \boldsymbol{r}' \quad ; \quad \psi \equiv \text{angle between } \boldsymbol{r} \text{ and } \boldsymbol{r}' \tag{8.26e, f}$$

and R is given by the law of cosines in the far field as

$$R = \left[r^2 + (r')^2 - 2rr'\cos\psi\right]^{1/2} \cong \begin{cases} r - r'\cos\psi & \text{for phase variations} \\ r & \text{for amplitude variations} \end{cases} \tag{8.26g}$$

The phasor $\boldsymbol{E}$- and $\boldsymbol{H}$-fields due to the vector potentials of (8.24) and (8.25) are given rigorously by

$$\breve{\boldsymbol{E}} = -j\omega\left[\breve{\boldsymbol{A}} + \frac{1}{k^2}\nabla\left(\nabla\cdot\breve{\boldsymbol{A}}\right)\right] - \frac{1}{\varepsilon_0}\nabla\times\breve{\boldsymbol{F}} \tag{8.27}$$

$$\breve{\boldsymbol{H}} = -j\omega\left[\breve{\boldsymbol{F}} + \frac{1}{k^2}\nabla\left(\nabla\cdot\breve{\boldsymbol{F}}\right)\right] + \frac{1}{\mu_0}\nabla\times\breve{\boldsymbol{A}} \tag{8.28}$$

Neglecting terms in (8.27) and (8.28) that diminish as order$(1/r^2)$ or faster, and neglecting radial field components of negligible amplitude compared to the θ and ϕ components, we obtain the following simplification in the far field. Here, the $\boldsymbol{E}$ components are given by

$$\breve{E}_r \cong 0 \tag{8.29a}$$

$$\breve{E}_\theta \cong -j\omega\left(\breve{A}_\theta + \eta_0 \breve{F}_\phi\right) = -\frac{jke^{-jkr}}{4\pi r}\left(\breve{L}_\phi + \eta_0 \breve{N}_\theta\right) \tag{8.29b}$$

$$\breve{E}_\phi \cong -j\omega\left(\breve{A}_\phi - \eta_0 \breve{F}_\theta\right) = +\frac{jke^{-jkr}}{4\pi r}\left(\breve{L}_\theta - \eta_0 \breve{N}_\phi\right) \tag{8.29c}$$

The far-field $\boldsymbol{H}$ components are given by

$$\breve{H}_r \cong 0 \tag{8.30a}$$

$$\breve{H}_\theta \cong +\frac{j\omega}{\eta_0}\left(\breve{A}_\phi - \eta_0 \breve{F}_\theta\right) = +\frac{jke^{-jkr}}{4\pi r}\left(\breve{N}_\phi - \breve{L}_\theta/\eta_0\right) \tag{8.30b}$$

$$\breve{H}_\phi \cong -\frac{j\omega}{\eta_0}\left(\breve{A}_\theta + \eta_0 \breve{F}_\phi\right) = -\frac{jke^{-jkr}}{4\pi r}\left(\breve{N}_\theta + \breve{L}_\phi/\eta_0\right) \tag{8.30c}$$

where $\eta_0 = \sqrt{\mu_0/\varepsilon_0}$ is the intrinsic impedance of free space.

Given the Cartesian geometry of the FDTD space lattice and its NTFF transformation surface S, it is convenient to first calculate $\breve{\boldsymbol{N}}$ and $\breve{\boldsymbol{L}}$ in rectangular coordinates, and then transform to spherical coordinates for substitution into (8.29) and (8.30). Again following [7], we have

$$\breve{\boldsymbol{N}} = \iint_S \left(\hat{\boldsymbol{x}}\breve{J}_x + \hat{\boldsymbol{y}}\breve{J}_y + \hat{\boldsymbol{z}}\breve{J}_z\right) e^{+jkr'\cos\psi}\, ds' \tag{8.31}$$

$$\breve{\boldsymbol{L}} = \iint_S \left(\hat{\boldsymbol{x}}\breve{M}_x + \hat{\boldsymbol{y}}\breve{M}_y + \hat{\boldsymbol{z}}\breve{M}_z\right) e^{+jkr'\cos\psi}\, ds' \tag{8.32}$$

and the corresponding θ and ϕ components:

$$\breve{N}_\theta = \iint_S \left(\breve{J}_x\cos\theta\cos\phi + \breve{J}_y\cos\theta\sin\phi - \breve{J}_z\sin\theta\right) e^{+jkr'\cos\psi}\, ds' \tag{8.33a}$$

$$\breve{N}_\phi = \iint_S \left(-\breve{J}_x\sin\phi + \breve{J}_y\cos\phi\right) e^{+jkr'\cos\psi}\, ds' \tag{8.33b}$$

$$\breve{L}_\theta = \iint_S \left(\breve{M}_x \cos\theta\cos\phi + \breve{M}_y \cos\theta\sin\phi - \breve{M}_z \sin\theta\right) e^{+jkr'\cos\psi}\, ds' \tag{8.34a}$$

$$\breve{L}_\phi = \iint_S \left(-\breve{M}_x \sin\phi + \breve{M}_y \cos\phi\right) e^{+jkr'\cos\psi}\, ds' \tag{8.34b}$$

The integral expressions of (8.33) and (8.34) can now be numerically implemented, taking into account the location and orientation of each face of the integration surface S. Assuming that S is a rectangular box of side dimensions $2x_0$, $2y_0$, $2z_0$ located symmetrically about the coordinate origin, there are six faces to be accounted for in the integration process. These may be grouped into three pairs of faces whose kernel functions have common properties, as follows.

1. *The two faces of S located at* $x' = \pm x_0$

   Nonzero components of $\breve{J}_S$ and $\breve{M}_S$: J_y, $\breve{J}_z$, M_y, $\breve{M}_z$

   Exponential phase term:

$$\begin{aligned} r'\cos\psi &= \boldsymbol{r}'\cdot\hat{\boldsymbol{r}} = \left(\pm x_0\,\hat{\boldsymbol{x}} + y'\,\hat{\boldsymbol{y}} + z'\,\hat{\boldsymbol{z}}\right)\cdot\left(\hat{\boldsymbol{x}}\sin\theta\cos\phi + \hat{\boldsymbol{y}}\sin\theta\sin\phi + \hat{\boldsymbol{z}}\cos\theta\right) \\ &= \pm x_0\sin\theta\cos\phi + y'\sin\theta\sin\phi + z'\cos\theta \end{aligned}$$

   Integration limits: $-y_0 \le y' \le y_0$, $-z_0 \le z' \le z_0$; $ds' = dy'\,dz'$

2. *The two faces of S located at* $y' = \pm y_0$

   Nonzero components of $\breve{J}_S$ and $\breve{M}_S$: $\breve{J}_x$, $\breve{J}_z$, $\breve{M}_x$, $\breve{M}_z$

   Exponential phase term:

$$\begin{aligned} r'\cos\psi &= \boldsymbol{r}'\cdot\hat{\boldsymbol{r}} = \left(x'\,\hat{\boldsymbol{x}} \pm y_0\,\hat{\boldsymbol{y}} + z'\,\hat{\boldsymbol{z}}\right)\cdot\left(\hat{\boldsymbol{x}}\sin\theta\cos\phi + \hat{\boldsymbol{y}}\sin\theta\sin\phi + \hat{\boldsymbol{z}}\cos\theta\right) \\ &= x'\sin\theta\cos\phi \pm y_0\sin\theta\sin\phi + z'\cos\theta \end{aligned}$$

   Integration limits: $-x_0 \le x' \le x_0$, $-z_0 \le z' \le z_0$; $ds' = dx'\,dz'$

3. *The two faces of S located at* $z' = \pm z_0$

   Nonzero components of $\breve{J}_S$ and $\breve{M}_S$: $\breve{J}_x$, J_y, $\breve{M}_x$, M_y

   Exponential phase term:

$$\begin{aligned} r'\cos\psi &= \boldsymbol{r}'\cdot\hat{\boldsymbol{r}} = \left(x'\,\hat{\boldsymbol{x}} + y'\,\hat{\boldsymbol{y}} \pm z_0\,\hat{\boldsymbol{z}}\right)\cdot\left(\hat{\boldsymbol{x}}\sin\theta\cos\phi + \hat{\boldsymbol{y}}\sin\theta\sin\phi + \hat{\boldsymbol{z}}\cos\theta\right) \\ &= x'\sin\theta\cos\phi + y'\sin\theta\sin\phi \pm z_0\cos\theta \end{aligned}$$

   Integration limits: $-x_0 \le x' \le x_0$, $-y_0 \le y' \le y_0$; $ds' = dx'\,dy'$

Finally, assuming knowledge of $\breve{N}_\theta$, $\breve{N}_\phi$, $\breve{L}_\theta$, and $\breve{L}_\phi$, we can write the following expression for the time-averaged Poynting vector of the scattered field:

$$\begin{aligned} P_{\text{scat}} &= \frac{1}{2}\text{Re}\left(\breve{E}_\theta \breve{H}_\phi^*\right) + \frac{1}{2}\text{Re}\left(-\breve{E}_\phi \breve{H}_\theta^*\right) \\ &= \frac{k^2}{32\pi^2 \eta_0 r^2}\left(\left|\breve{L}_\phi + \eta_0 \breve{N}_\theta\right|^2 + \left|\breve{L}_\theta - \eta_0 \breve{N}_\phi\right|^2\right) \end{aligned} \tag{8.35}$$

Upon substitution of (8.35) into the definition of the bistatic RCS in three dimensions:

$$\text{RCS}(\theta, \phi) \equiv \lim_{r\to\infty}\left(4\pi r^2 \frac{P_{\text{scat}}}{P_{\text{inc}}}\right) \tag{8.36}$$

we obtain the final expression for RCS (having the dimension of square meters):

$$\text{RCS}(\theta, \phi) = \frac{k^2}{8\pi \eta_0 P_{\text{inc}}}\left(\left|\breve{L}_\phi + \eta_0 \breve{N}_\theta\right|^2 + \left|\breve{L}_\theta - \eta_0 \breve{N}_\phi\right|^2\right) \tag{8.37}$$

where P_{inc} is the power density in the illuminating wave. Note that (8.37) is a function of only θ and ϕ, *not* r.

It is clear that the integral expressions of (8.33) and (8.34) can be numerically evaluated as discussed above to obtain the required phasor quantities. Thus, a single FDTD run modeling a plane-wave excitation having a specific incidence angle and polarization provides the scattered near-field data along S to permit calculation of the complete far-field bistatic RCS pattern. For an antenna problem, there is no incident plane wave, but (8.37) is still usable to obtain the far-field radiation pattern from the radiated near field along S. Here, the normalizing factor P_{inc} can be taken as the total power exciting the antenna.

8.6 TIME-DOMAIN NEAR-TO-FAR-FIELD TRANSFORMATION

Luebbers et al. [8] reported an efficient "on-the-fly" time-domain NTFF transformation that calculates the time waveforms of the scattered or radiated E- and H-fields at selected angular locations in the far field. These calculations are performed concurrently with the normal FDTD time-stepping.

Luebbers' method involves setting up time-dimensioned arrays for the far-field vector potentials. Each array element is determined by conducting a recursive (running) sum of contributions from the time-domain electric and magnetic current sources just computed via FDTD on S. These contributions are delayed in time according to the propagation delay between a source element on S and the far-field observation point. If the far-field bistatic RCS pattern or antenna radiation pattern is required at specific frequencies, the field-versus-time waveform obtained in this manner can be postprocessed via an FFT.

We now summarize this approach. To correlate as much as possible with the notation used in [8], we start with (8.29b) and (8.29c) and make the substitution $k = 2\pi/\lambda_0$:

$$\breve{E}_\theta \equiv -\frac{jke^{-jkr}}{4\pi r}\left(\breve{L}_\phi + \eta_0 \breve{N}_\theta\right) = -\frac{je^{-jkr}}{2\lambda_0 r}\left(\breve{L}_\phi + \eta_0 \breve{N}_\theta\right) \tag{8.38}$$

$$\breve{E}_\phi \equiv +\frac{jke^{-jkr}}{4\pi r}\left(\breve{L}_\theta - \eta_0 \breve{N}_\phi\right) = +\frac{je^{-jkr}}{2\lambda_0 r}\left(\breve{L}_\theta - \eta_0 \breve{N}_\phi\right) \tag{8.39}$$

These are, respectively, (3) and (4) of [8], but the distance from the origin to the far-field observation point is denoted here by r, rather than the R symbol used in [8]. Next, defining

$$\breve{\boldsymbol{W}} \equiv \frac{je^{-jkr}}{2\lambda_0 r}\breve{\boldsymbol{N}}\,; \qquad \breve{\boldsymbol{U}} \equiv \frac{je^{-jkr}}{2\lambda_0 r}\breve{\boldsymbol{L}} \tag{8.40, 8.41}$$

as in (5) and (6) of [8], we can rewrite (8.38) and (8.39) more simply as

$$\breve{E}_\theta \equiv -\eta_0 \breve{W}_\theta - \breve{U}_\phi\,; \qquad \breve{E}_\phi \equiv -\eta_0 \breve{W}_\phi + \breve{U}_\theta \tag{8.42, 8.43}$$

We can now apply the inverse Fourier transformation to each term of (8.40) and (8.41), using the basic definitions of $\breve{\boldsymbol{N}}$ and $\breve{\boldsymbol{L}}$ in (8.31) and (8.32), respectively. This yields the following time-domain relations corresponding to (8.40) and (8.41):

$$\boldsymbol{W}(\boldsymbol{r}, t) = \frac{1}{4\pi rc}\frac{\partial}{\partial t}\left[\iint_S \boldsymbol{J}_S\left(t - \frac{r - \boldsymbol{r}'\cdot\hat{\boldsymbol{r}}}{c}\right)dS'\right] \tag{8.44}$$

$$\boldsymbol{U}(\boldsymbol{r}, t) = \frac{1}{4\pi rc}\frac{\partial}{\partial t}\left[\iint_S \boldsymbol{M}_S\left(t - \frac{r - \boldsymbol{r}'\cdot\hat{\boldsymbol{r}}}{c}\right)dS'\right] \tag{8.45}$$

In addition, inverse Fourier transformation of (8.42) and (8.43) yields

$$E_\theta(\boldsymbol{r}, t) \equiv -\eta_0 W_\theta(\boldsymbol{r}, t) - U_\phi(\boldsymbol{r}, t) \tag{8.46}$$

$$E_\phi(\boldsymbol{r}, t) \equiv -\eta_0 W_\phi(\boldsymbol{r}, t) + U_\theta(\boldsymbol{r}, t) \tag{8.47}$$

Equations (8.44) to (8.47) form the theoretical basis of Luebbers' time-domain NTFF transformation. These correspond to Equations (7) to (10) of [8], respectively, with minor differences in notation.

We note that the argument of $\boldsymbol{J}_S$ and $\boldsymbol{M}_S$ in (8.44) and (8.45) implies a time delay τ_d between the appearance of H- and E-fields on virtual surface S, and their impact upon the far-field vector potentials $\boldsymbol{W}$ and $\boldsymbol{U}$. This time delay is given by

$$\tau_d = \frac{r - \boldsymbol{r}'\cdot\hat{\boldsymbol{r}}}{c} = \frac{r - r'\cos\psi}{c} \tag{8.48}$$

In fact, this time delay is precisely the propagation delay of an electromagnetic wave in free space over the distance R given by (8.26g), which is between the source point $\boldsymbol{r}'$ on S and the far-field observation point $\boldsymbol{r}$.

We now consider the implementation of (8.44) to (8.47) in the context of the FDTD method. The overall strategy is to evaluate the integrals of (8.44) and (8.45) in Cartesian coordinates along the six planar faces of S at each time-step, thereby obtaining the Cartesian components of the incremental values of $\boldsymbol{W}$ and $\boldsymbol{U}$ at each cell of each face at that time-step. These increments are added to the respective running sums for the time samples of the $\boldsymbol{W}$ and $\boldsymbol{U}$ waveforms, assuming appropriate source-to-observation-point time delays τ_d. Upon conclusion of time-stepping, the Cartesian components of the final accumulated $\boldsymbol{W}$ and $\boldsymbol{U}$ waveforms are converted to θ and ϕ components in spherical coordinates for direct substitution into (8.46) and (8.47).

To illustrate the core of this process, we follow the example of (11) in [8], and use as a sample excitation the magnetic current

$$M_z \hat{z} = -\hat{y} \times E_x \hat{x} = E_x \hat{z} \tag{8.49}$$

flowing within a rectangular patch $\Delta x \Delta z$ located at $\boldsymbol{r}' = \left(x' \hat{x} + y_0 \hat{y} + z' \hat{z}\right)$ on the $y = +y_0$ face of S. From (8.45), this magnetic current yields the following additive increment to $\boldsymbol{U}(\boldsymbol{r}, t)$:

$$\Delta \boldsymbol{U} = \Delta U_z \hat{z} = \frac{1}{4\pi r c} \frac{\partial}{\partial t}\left(M_z \hat{z}\, \Delta x \Delta z\right) = \frac{\Delta x \Delta z}{4\pi r c} \frac{\partial}{\partial t}\left(E_x \hat{z}\right) \tag{8.50}$$

This increment contributes to $\boldsymbol{U}(\boldsymbol{r}, t)$ after the time delay τ_d of (8.48), expressed as the following fractional number of time-steps:

$$f = \frac{\tau_d}{\Delta t} = \frac{r - r' \cos\psi}{c\, \Delta t} \tag{8.51}$$

Assuming a second-order-accurate central-difference realization of the time derivative of (8.50) evaluated at the half time-step point $n + 1/2$, and assuming standard FDTD notation, (8.50) and (8.51) can be concisely combined to yield the following recursive (running) sum:

$$U_z\big|_{\boldsymbol{r}}^{n+1/2+f} = U_z\big|_{\boldsymbol{r}}^{n+1/2+f} + \frac{\Delta x \Delta z}{4\pi r c \Delta t}\left(E_x\big|_{\boldsymbol{r}'}^{n+1} - E_x\big|_{\boldsymbol{r}'}^{n}\right) \tag{8.52}$$

We see that (8.52) conveys the full meaning of (8.45) in that the time derivative of the magnetic current on the elemental surface patch at time-step $n + 1/2$ appears in the far-field vector potential after a delay of f time-steps, corresponding to the propagation delay between the surface patch and the observation point. The recursive sum notation appearing in (8.52) serves to remind us that magnetic currents flowing on other patches on S, not necessarily at the same time-step, can also provide a contribution to $U_z\big|_{\boldsymbol{r}}^{n+1/2+f}$ which must be summed immediately after they are obtained in the normal FDTD time-stepping.

The last step in the process is to conform (8.52) with the fact that $U_z|_r^n$ is, in fact, a finite array of time samples at M discrete time-steps. With the time-step delay f of (8.51) being a decimal fraction, it will be rare, indeed, that a delayed value of the additive increment in (8.50) falls precisely on one of the time samples of U_z. A decision must be made regarding how to apportion the delayed value of the increment to the two discrete values of U_z that lie to either side in time. The approach followed by [8] is to use linear interpolation. Adopting a slightly different notation than [8], let

$$nn = \mathrm{INT}(n + 1/2 + f) \tag{8.53}$$

where INT is a function that extracts the greatest integer in the argument. Then,

$$a = (n + 1/2 + f) - nn \tag{8.54}$$

is the fractional time-step ($0 \le a \le 1$) between the true time location of the delayed increment and the U_z sample $U_z|_r^{nn}$ that occurs just before. Conversely, $(1 - a)$ is the fractional time-step between the true time location of the delayed increment and the U_z sample $U_z|_r^{nn+1}$ that immediately follows. Then, we have the following apportionment for the running sum:

$$U_z|_r^{nn} = U_z|_r^{nn} + (1 - a)\cdot\frac{\Delta x\,\Delta z}{4\pi r c\,\Delta t}\left(E_x|_{r'}^{n+1} - E_x|_{r'}^{n}\right) \tag{8.55a}$$

$$U_z|_r^{nn+1} = U_z|_r^{nn+1} + a\cdot\frac{\Delta x\,\Delta z}{4\pi r c\,\Delta t}\left(E_x|_{r'}^{n+1} - E_x|_{r'}^{n}\right) \tag{8.55b}$$

Implementing (8.55) for each surface patch on S for each time-step of the FDTD modeling run completes the process of evaluating the U_z integral of (8.45), thereby completely evolving the far-field vector potential waveform $U_z(\boldsymbol{r}, t)$. We need only ensure that M, the total number of time-value storage locations assigned to the $U_z|_r^n$ array, sufficiently exceeds $NMAX$, the number of time-steps in the FDTD run, to allow for delayed contributions arriving from all parts of S at the conclusion of time-stepping. Assuming that the maximum distance between any two points on S is $s\Delta$, and that the time-step relation $2c\Delta t = \Delta$ is used, it is sufficient to set

$$M = NMAX + 2s \tag{8.56}$$

to ensure that there are enough storage locations in the $U_z|_r^n$ array to deal with arbitrary angular orientations of the far-field observation point $\boldsymbol{r}$.

It is clear that the process discussed above can be implemented in an analogous manner for $U_x(\boldsymbol{r}, t)$ and $U_y(\boldsymbol{r}, t)$ in (8.45), as well as for the three Cartesian components of $\boldsymbol{W}(\boldsymbol{r}, t)$ in (8.44). For the latter, two additional considerations are that:

1. The FDTD $\boldsymbol{H}$ components used to obtain $\boldsymbol{J}_S$ are shifted by $0.5\Delta t$ relative to the $\boldsymbol{E}$ components used to implement (8.45).
2. There is a sign reversal in the definition of the equivalent current.

To illustrate these changes, consider the following expressions that are analogous to (8.52) to (8.55). These expressions provide the contribution to the running sum for W_z along the $y = +y_0$ face of S. First, we have the true location of the time-delayed J_S contribution:

$$W_z\Big|_r^{n+f} = W_z\Big|_r^{n+f} - \frac{\Delta x \Delta z}{4\pi r c \Delta t}\left(H_x\Big|_{r'}^{n+1/2} - H_x\Big|_{r'}^{n-1/2}\right) \tag{8.57}$$

Next, there are the time interpolation factors:

$$nn = \text{INT}(n + f) \qquad a = (n + f) - nn \tag{8.58a, b}$$

Finally, the apportionment for the running sum:

$$W_z\Big|_r^{nn} = W_z\Big|_r^{nn} - (1-a)\cdot\frac{\Delta x \Delta z}{4\pi r c \Delta t}\left(H_x\Big|_{r'}^{n+1/2} - H_x\Big|_{r'}^{n-1/2}\right) \tag{8.59a}$$

$$W_z\Big|_r^{nn+1} = W_z\Big|_r^{nn+1} - a\cdot\frac{\Delta x \Delta z}{4\pi r c \Delta t}\left(H_x\Big|_{r'}^{n+1/2} - H_x\Big|_{r'}^{n-1/2}\right) \tag{8.59b}$$

In total, six one-dimensional arrays, each containing $M = NMAX + 2s$ storage locations, are required to store the time variation of the six Cartesian vector components of $\boldsymbol{W}(\boldsymbol{r}, t)$ and $\boldsymbol{U}(\boldsymbol{r}, t)$ for each (θ, ϕ) in the far field at which a field-versus-time waveform is required. Upon completion of the filling of these arrays, the following rectangular-to-spherical vector component conversion is performed:

$$W_\theta(\boldsymbol{r}, t) = W_x(\boldsymbol{r}, t)\cos\theta\cos\phi + W_y(\boldsymbol{r}, t)\cos\theta\sin\phi - W_z(\boldsymbol{r}, t)\sin\theta \tag{8.60a}$$

$$W_\phi(\boldsymbol{r}, t) = -W_x(\boldsymbol{r}, t)\sin\phi + W_y(\boldsymbol{r}, t)\cos\phi \tag{8.60b}$$

$$U_\theta(\boldsymbol{r}, t) = U_x(\boldsymbol{r}, t)\cos\theta\cos\phi + U_y(\boldsymbol{r}, t)\cos\theta\sin\phi - U_z(\boldsymbol{r}, t)\sin\theta \tag{8.61a}$$

$$U_\phi(\boldsymbol{r}, t) = -U_x(\boldsymbol{r}, t)\sin\phi + U_y(\boldsymbol{r}, t)\cos\phi \tag{8.61b}$$

The time waveforms of the vector potentials of (8.60) and (8.61) can now be inserted into (8.46) and (8.47) to directly obtain the desired time waveforms of the E-field at observation point $\boldsymbol{r}$ in the far zone, completing the calculation. From (8.44) and (8.45), we note that the amplitudes of these waveforms diminish as $1/r$, so that it is possible to obtain a normalized far-field response that is independent of the distance from the origin simply by multiplying (8.46) and (8.47) by r. In fact, this is how [8] presented its time-waveform results. Finally, we note that [8] reported a postprocessing FFT of its normalized far-field waveforms to yield the wideband RCS response of generic PEC flat plates. Very good agreement with frequency-domain MoM data was reported.

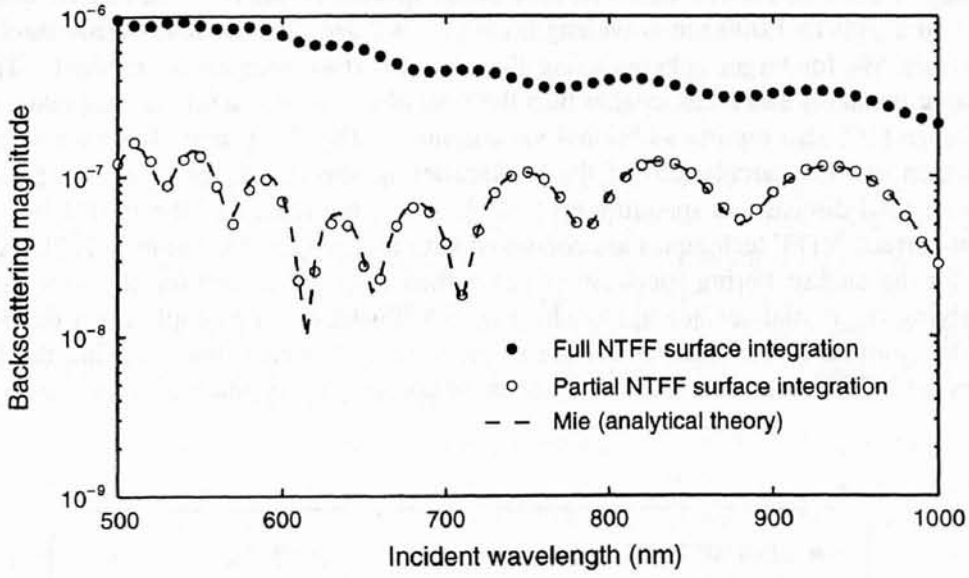
8.7 MODIFIED NTFF PROCEDURE TO MORE ACCURATELY CALCULATE BACKSCATTERING FROM STRONGLY FORWARD-SCATTERING OBJECTS

Backscattered electromagnetic waves in the microwave frequency range are routinely used in a wide variety of detection and remote-sensing applications. At much higher frequencies, it has recently been demonstrated that backscattered light spectra from biological tissues can provide valuable medical diagnostic information, including potential noninvasive means for early-stage cancer detection [9–11]. In fact, promising emerging applications in optical tissue diagnosis have motivated a series of detailed FDTD studies of light scattering by complex material shapes in the 1λ to 10λ size range [12–14]. Such objects are generally characterized by strong forward-scattering lobes [15]. (See also Chapter 16, Section 16.21.)

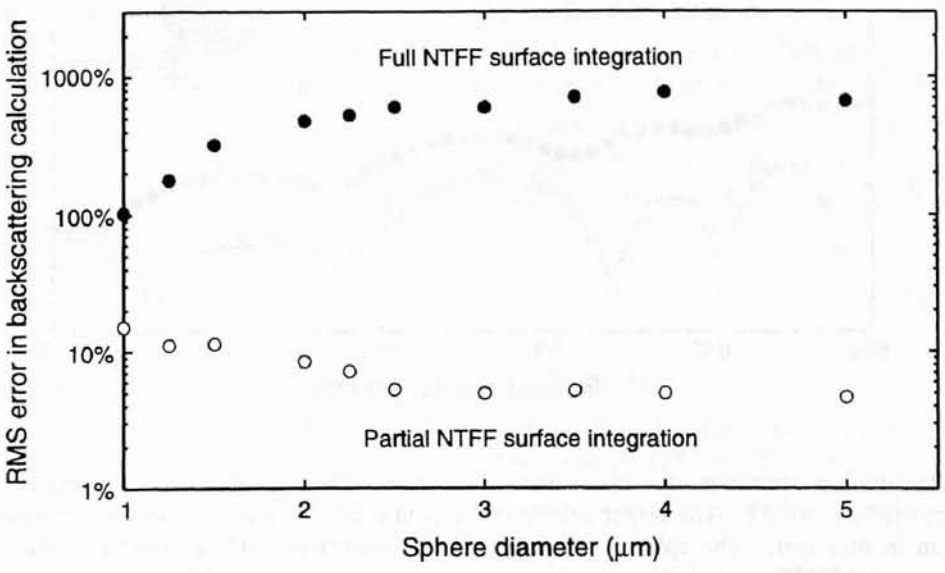
Unfortunately, using the basic NTFF procedure discussed earlier in this chapter, it has been found to be difficult to apply FDTD to accurately calculate the backscattering spectra of strongly forward-scattering objects. Numerical experiments reveal that, for such objects, small-percentage errors in the FDTD-calculated fields along the portion of the NTFF integration surface located in the forward-scattering region result in large-percentage errors in the far-field backscattering [16]. In essence, this is a "subtraction-noise" problem. That is, the hundreds or even thousands of relatively large field values collected over the NTFF surface in the forward-scattering region must nearly cancel upon integration (addition) to correctly yield the relatively small values in the backscattered far field. Consequently, converged backscattering results require unusually high-precision FDTD calculations of the near field, with spatial resolutions much finer than $\lambda/20$ (which usually suffices for accurate forward-scattering results). This requirement for very fine grid resolution poses obvious and significant computational burdens.

Instead of trying to improve the FDTD near-field calculation, [16] proposes a simple alternative: just *omit* the forward plane of the NTFF surface when performing the integration to calculate the far-field backscattering. This approach is clearly an approximation to the rigorous surface-equivalence theorem, in which integration over a closed virtual surface completely surrounding the scattering object is specified. However, this approximation can be justified by the fact that the Poynting vectors of the scattered waves propagating through the forward plane of the NTFF integration surface approximately point in the forward direction, and therefore should have only a small contribution to the backscattered far field. In essence, the working hypothesis of [16] is that, for strongly forward-scattering objects, the gain in accuracy realized by omitting the forward-plane fields (with their relatively high level of numerical "noise") from the NTFF integration substantially outweighs the loss of accuracy caused by the incomplete surface integration.

We now summarize numerical evidence reported in [16] that supports this hypothesis. Fig. 8.4(a) illustrates the backscattering spectrum of a single, plane-wave-illuminated, 3-μm-diameter dielectric sphere of relative permittivity $\varepsilon_r = 1.21$. The sphere is mapped with a staircased surface into a uniform, three-dimensional FDTD space lattice of cubic cell size $\Delta = 25$ nm. Two FDTD backscattering spectra are shown for this sphere: one calculated with the full NTFF surface integration (as specified earlier in this chapter), and the other calculated with a partial surface integration that omits the NTFF plane in the forward-scattering region. These spectra are graphed over a 500 to 1,000 nm incident wavelength range in comparison with the Mie far-field analytical solution. From Fig. 8.4(a), we see that omitting the forward plane of the NTFF surface results in a much more accurate calculation of the backscattering at all wavelengths examined, with no additional requirement for computational resources.



(a)



(b)

Fig. 8.4 (a) Backscattering spectrum of a single, plane-wave-illuminated, 3-μm-diameter dielectric sphere of relative permittivity $\varepsilon_r = 1.21$. The sphere is mapped with a staircased surface into a uniform, three-dimensional FDTD space lattice of cubic cell size $\Delta = 25$ nm. Two FDTD backscattering spectra are shown for this sphere: one calculated with the full NTFF surface integration, and the other calculated with a partial surface integration that omits the NTFF plane in the forward-scattering region. (b) RMS error of the backscattering calculations over the 500 to 1,000 nm incident wavelength range as functions of sphere diameter. *Source:* Li et al., *IEEE Antennas and Wireless Propagation Lett.*, 2005, pp. 35–38, © 2005 IEEE.

Fig. 8.4(b) repeats the calculation for a series of $\varepsilon_r = 1.21$ dielectric spheres with diameters ranging from 1 to 5 μm. Here, a root-mean-square (RMS) error functional is used to characterize the accuracy of each FDTD-calculated backscattering spectrum relative to the analytical solution over the entire 500 to 1,000 nm wavelength range. We see that the RMS error stabilizes at approximately 5% for larger spheres using the partial-surface integration method. This is a factor of approximately 100 times smaller than the error observed with a full-surface integration.

Reference [16] also reports additional validations of the NTFF partial-surface integration method which involve calculations of the backscattering spectra of aggregations of dielectric spheres with total dimensions spanning up to 20λ. Here, the results of the FDTD full-surface and partial-surface NTFF techniques are compared with the analytical solution of [17]. An RMS accuracy for the backscattering spectrum of better than 10% is reported for all cases examined when applying the partial-surface approach. Fig. 8.5 illustrates an example for a three-sphere aggregation. Similar to the findings for the single sphere, it is clear that omitting the forward plane from the NTFF integration greatly improves the accuracy in calculating backscattering.

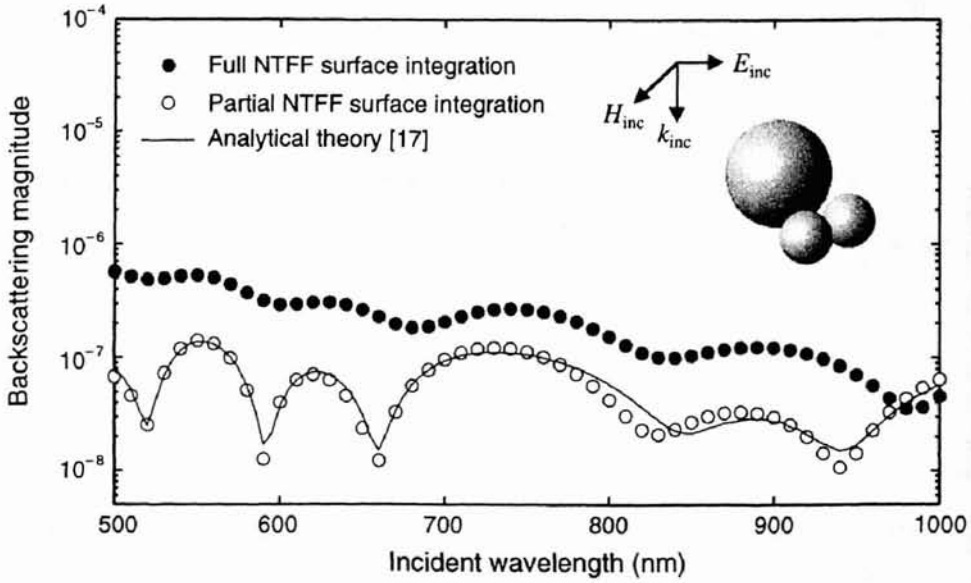


Fig. 8.5 Backscattering spectrum of a plane-wave-illuminated triplet of dielectric spheres of relative permittivity $\varepsilon_r = 1.21$. The larger sphere is 2 μm in diameter, and the two smaller spheres are 1 μm in diameter. The spheres are mapped with staircased surfaces into a uniform, three-dimensional FDTD space lattice of cubic cell size $\Delta = 25$ nm. Two FDTD backscattering spectra are shown: one calculated with the full NTFF surface integration, and the other calculated with a partial surface integration that omits the NTFF plane in the forward-scattering region. *Source:* Li et al., *IEEE Antennas and Wireless Propagation Lett.*, 2005, pp. 35–38, © 2005 IEEE.

In addition to having applications in numerical studies related to the biophotonics of individual living cells and clusters of cells, the method proposed in [16] could provide improved FDTD modeling of the backscattering cross section of certain types of low-observable aerospace vehicles. These vehicles would be in the class of structures that are designed to shift the major scattering lobes away from the backscatter direction.

8.8 SUMMARY

This chapter discussed the basis of frequency-domain and time-domain near-to-far-field transformations suitable for use in FDTD simulations. The discussion began by using the phasor-domain Green's theorem in two dimensions to prove that scattered or radiated $\boldsymbol{E}$ and $\boldsymbol{H}$ fields tangential to a virtual surface enclosing the structure being modeled can be integrated to provide the complete far-field pattern. This virtual surface is independent of the shape or composition of the structure being modeled, and can have a fixed rectangular shape to conform with the Cartesian FDTD lattice. It was next shown that the phasor field data needed for this calculation can be efficiently obtained during the FDTD run by running a concurrent DFT on the time-stepped $\boldsymbol{E}$ and $\boldsymbol{H}$ field components tangential to the designated virtual surface in the lattice.

These discussions motivated a review of a powerful phasor-domain surface equivalence theorem. This theorem allows calculation of the sinusoidal scattered or radiated far fields in general three-dimensional FDTD simulations.

The chapter continued with a review of the theory and numerical implementation of a three-dimensional time-domain NTFF transformation which permits direct computation of scattered or radiated pulses at a set of angles in the far field. This procedure can be used instead of the phasor-domain transformation discussed earlier, if the required data are time waveforms, and there are relatively few far-field observation angles of interest.

Finally, the chapter concluded with a recent simple modification of the NTFF procedure that greatly improves the accuracy in calculating the backscatter from strongly forward-scattering objects, such as biological cells illuminated by light and certain types of low-observable vehicles. The modification involves simply omitting the forward plane of the NTFF surface when performing the integration to calculate the far-field backscattering. This avoids the "subtraction-noise" problem posed by the requirement for near cancellation of the relatively large field values collected on this plane when integrated into the far backscattered field.

REFERENCES

[1] Jackson, J. D., *Classical Electrodynamics*, New York: Wiley, 1975.

[2] Umashankar, K. R., and A. Taflove, "A novel method to analyze electromagnetic scattering of complex objects," *IEEE Trans. Electromagn. Compat.*, Vol. 24, 1982, pp. 397–405.

[3] Taflove, A., and K. R. Umashankar, "Radar cross section of general three-dimensional structures," *IEEE Trans. Electromagn. Compat.*, Vol. 25, 1983, pp. 433–440.

[4] Ko, W., and R. Mittra, "A combination of FDTD and Prony methods for analyzing microwave integrated circuits," *IEEE Trans. Microwave Theory Tech.*, Vol. 39, 1991, pp. 2176–2181.

[5] Houshmand, B., T. W. Huang, and T. Itoh, "Microwave structure characterization by a combination of FDTD and system identification methods," *IEEE Microwave Guided Wave Lett.*, Vol. 3, 1993, pp. 262–264.

[6] Schelkunoff, S. A., "Some equivalence theorems of electromagnetics and their application to radiation problems," *Bell System Technical J.*, Vol. 15, 1936, pp. 92–112.

[7] Balanis, C. A., *Advanced Engineering Electromagnetics*, New York: Wiley, 1989.

[8] Luebbers, R. J., K. S. Kunz, M. Schneider, and F. Hunsberger, "A finite-difference time-domain near zone to far zone transformation," *IEEE Trans. Antennas Propagat.*, Vol. 39, 1991, pp. 429–433.

[9] Backman, V., M. B. Wallace, L. T. Perelman, J. T. Arendt, R. Gurjar, M. G. Muller, Q. Zhang, G. Zonios, E. Kline, T. McGillican, S. Shapshay, T. Valdez, J. V. Dam, K. Badizadegan, J. M. Crawford, M. Fitzmaurice, S. Kabani, H. S. Levin, M. Seiler, R. R. Dasari, I. Itzkan, and M. S. Feld, "Detection of preinvasive cancer cells in situ," *Nature,* Vol. 406, 2000, pp. 35–36.

[10] Kim, Y. L., Y. Liu, R. K. Wali, H. K. Roy, M. J. Goldberg, A. K. Kromine, K. Chen, and V. Backman, "Simultaneous measurement of angular and spectral properties of light scattering for characterization of tissue microarchitecture and its alteration in early precancer," *IEEE J. Selected Topics Quant. Elect.,* Vol. 9, 2003, pp. 243–256.

[11] Roy, H. K., Y. Liu, R. Wali, Y. L. Kim, A. K. Kromine, M. J. Goldberg, and V. Backman, "Four-dimensional elastic light-scattering fingerprints as preneoplastic markers in the rat model of colon carcinogenesis," *Gastroenterology,* Vol. 126, 2004, pp. 1071–1081.

[12] Drezek, R., A. Dunn, and R. Richards-Kortum, "Light scattering from cells: Finite-difference time-domain simulations and goniometric measurements," *Appl. Opt.,* Vol. 38, 1999, pp. 3651–3661.

[13] Li, X., Z. Chen, J. Gong, A. Taflove, and V. Backman, "Analytical techniques for addressing forward and inverse problems of light scattering by irregularly shaped particles," *Optics Lett.,* Vol. 29, 2004, pp. 1239–1241.

[14] Li, X., Z. Chen, A. Taflove, and V. Backman, "Equiphase-sphere approximation for analysis of light scattering by arbitrarily-shaped nonspherical particles," *Appl. Opt.,* Vol. 43, 2004, pp. 4497–4505.

[15] Ishimaru, A., *Wave Propagation and Scattering in Random Media*, New York: Academic, 1978.

[16] Li, X., A. Taflove, and V. Backman, "Modified FDTD near-to-far field transformation for improved backscattering calculation of strongly forward-scattering objects," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 4, 2005, pp. 35–38.

[17] Xu, Y. L., and R. T. Wang, "Electromagnetic scattering by an aggregate of spheres: Theoretical and experimental study of the amplitude scattering matrix," *Phys. Rev. E,* Vol. 58, 1998, pp. 3931–3948.

PROJECT

P8.1 Implement the phasor-domain near-to-far-field transformation in a two-dimensional TM_z FDTD code. Write a DFT subroutine to obtain the necessary phasor data at a rectangular virtual surface surrounding the structure of interest. Conduct numerical experiments modeling the illumination of a flat PEC plate or square PEC cylinder to obtain the complete bistatic RCS pattern. Compare the results with data in the literature, such as [2].

Chapter 9

Dispersive, Nonlinear, and Gain Materials

Allen Taflove, Susan Hagness, Wojciech Gwarek, Masafumi Fujii, and Shih-Hui Chang

9.1 INTRODUCTION

Current and emerging technological applications in aerial and ground-penetrating radar, high-power microwaves, bioelectromagnetics, biophotonics, nanophotonics, and microlasers can involve electromagnetic wave interactions at the wavelength or subwavelength scale with materials having frequency-dispersive dielectric properties that potentially exhibit nonlinearity and/or gain. Computational electrodynamics modeling requires four key electromagnetic characteristics of such materials to be accounted:

- *Linear dispersion.* A material's permittivity and/or permeability varies with frequency at low intensities of the wave's E- and H-fields.
- *Nonlinearity.* A material's permittivity and/or permeability varies with the intensity of the wave's local E-field, H-field, or Poynting power flux (usually at high intensities).
- *Nonlinear dispersion.* The strength of a material's nonlinearity varies with the sinusoidal frequency of the electromagnetic wave.
- *Gain.* An active dielectric material (such as in a laser) transfers power from an external source (the pump) to progressively increase the amplitude of an electromagnetic wave propagating within the material. Gain can also be frequency-dispersive and nonlinear.

This chapter discusses several approaches that allow the FDTD method to model dielectric materials exhibiting linear dispersion, nonlinearity, nonlinear dispersion, and gain. Topics covered include: (1) a summary of generic material dispersions, including Debye, Lorentz, and Drude media; (2) the *piecewise-linear recursive convolution* (PLRC) method for linear dispersive media; (3) the *auxiliary differential equation* (ADE) method for linear dispersive and nonlinear dispersive passive media, and for saturable dispersive active gain media; (4) analysis of linear magnetized ferrites; and (5) coupling of quantum effects in an active atomic system to Maxwell's equations, with application to the detailed analysis of a laser's pumping dynamics. As with all FDTD techniques, these approaches enforce the vector-field boundary conditions at interfaces of dissimilar media at the time scale of a small fraction of the impinging pulse width or carrier period. As a result, these approaches are almost completely general.

9.2 GENERIC ISOTROPIC MATERIAL DISPERSIONS

This chapter first considers three important generic classes of linear, isotropic material dispersions amenable to FDTD modeling at the macroscopic scale: (1) the Debye relaxation, (2) the Lorentzian resonance, and (3) the Drude model of metals. Note that these cases are defined using the $\exp(j\omega t)$ convention for phasor (Fourier-domain) quantities.

9.2.1 Debye Media

Debye media are characterized by a complex-valued, frequency-domain susceptibility function $\chi(\omega)$ that has one or more real poles at separate frequencies. For a single-pole Debye medium, we have

$$\chi_p(\omega) = \frac{\varepsilon_{s,p} - \varepsilon_{\infty,p}}{1 + j\omega\tau_p} \equiv \frac{\Delta\varepsilon_p}{1 + j\omega\tau_p} \tag{9.1}$$

where $\varepsilon_{s,p}$ is the static or zero-frequency relative permittivity, $\varepsilon_{\infty,p}$ is the relative permittivity at infinite frequency, $\Delta\varepsilon_p$ is the change in relative permittivity due to the Debye pole, and τ_p is the pole relaxation time. The real-valued time-domain susceptibility function $\chi(t)$ is obtained by inverse Fourier transformation of (9.1), yielding the decaying exponential function

$$\chi_p(t) = \frac{\Delta\varepsilon_p}{\tau_p} e^{-t/\tau_p} U(t) \tag{9.2}$$

where $U(t)$ is the unit step. For a Debye medium having P poles, we extend (9.1) to express the relative permittivity as

$$\varepsilon(\omega) = \varepsilon_\infty + \sum_{p=1}^{P} \frac{\Delta\varepsilon_p}{1 + j\omega\tau_p} \tag{9.3}$$

9.2.2 Lorentz Media

Lorentz media are characterized by a complex-valued, frequency-domain susceptibility function $\chi(\omega)$ that has one or more pairs of complex-conjugate poles. For a Lorentz medium characterized by a single pole pair, we have

$$\chi_p(\omega) = \frac{\Delta\varepsilon_p \omega_p^2}{\omega_p^2 + 2j\omega\delta_p - \omega^2} \tag{9.4}$$

where $\Delta\varepsilon_p = \varepsilon_{s,p} - \varepsilon_{\infty,p}$ is the change in relative permittivity due to the Lorentz pole pair, ω_p is the frequency of the pole pair (the undamped resonant frequency of the medium), and δ_p is the damping coefficient. The real-valued time-domain susceptibility function $\chi(t)$ is obtained by inverse Fourier transformation of (9.4), yielding the exponentially decaying sinusoidal function

$$\chi_p(t) = \frac{\Delta\varepsilon_p \omega_p^2}{\sqrt{\omega_p^2 - \delta_p^2}} e^{-\delta_p t} \sin\left(\sqrt{\omega_p^2 - \delta_p^2}\, t\right) U(t) \tag{9.5}$$

For a Lorentz medium having P pole pairs, we extend (9.4) to express the relative permittivity as

$$\varepsilon(\omega) = \varepsilon_\infty + \sum_{p=1}^{P} \frac{\Delta\varepsilon_p \omega_p^2}{\omega_p^2 + 2j\omega\delta_p - \omega^2} \tag{9.6}$$

9.2.3 Drude Media

At optical wavelengths, it may be important to treat electromagnetic wave interactions with metals using a dispersive formulation to properly account for the physics of internal electron motion. At the macroscopic scale, the Drude model has become widely used for such modeling. For a single-pole Drude medium, we have

$$\chi_p(\omega) = -\frac{\omega_p^2}{\omega^2 - j\omega\gamma_p} \tag{9.7}$$

where ω_p is the Drude pole frequency and γ_p is the inverse of the pole relaxation time. The real-valued time-domain susceptibility function $\chi(t)$ is obtained by inverse Fourier transformation of (9.7), yielding the function

$$\chi_p(t) = \frac{\omega_p^2}{\gamma_p}\left(1 - e^{-\gamma_p t}\right) U(t) \tag{9.8}$$

For a Drude medium having P poles, we extend (9.7) to express the relative permittivity as

$$\varepsilon(\omega) = \varepsilon_\infty - \sum_{p=1}^{P} \frac{\omega_p^2}{\omega^2 - j\omega\gamma_p} \tag{9.9}$$

9.3 PIECEWISE-LINEAR RECURSIVE-CONVOLUTION METHOD, LINEAR MATERIAL CASE

Luebbers et al. [1–3] reported the development of an efficient "on-the-fly" PLRC approach to modeling electromagnetic wave interactions with linear materials that have combinations of multiple Debye and Lorentz dispersions. The discussion that follows, adapted primarily from [3], considers only dispersive electric permittivity. Results for dispersive magnetic permeability can be readily obtained by using duality.

9.3.1 General Formulation

At any point in a linear dispersive medium, the time-domain electric flux density $D(t)$ is related to the time-domain electric field intensity $E(t)$ by the convolution

$$\boldsymbol{D}(t) = \varepsilon_0 \varepsilon_\infty \boldsymbol{E}(t) + \varepsilon_0 \int_{\tau=0}^{t} \boldsymbol{E}(t-\tau)\, \chi(\tau)\, d\tau \tag{9.10}$$

Using the notation $D^n = D(n\Delta t)$ and $E^n = E(n\Delta t)$, (9.10) can be written in time-discrete form as

$$\boldsymbol{D}^n = \varepsilon_0 \varepsilon_\infty \boldsymbol{E}^n + \varepsilon_0 \int_{\tau=0}^{n\Delta t} \boldsymbol{E}(n\Delta t-\tau)\, \chi(\tau)\, d\tau \tag{9.11}$$

where ε_0 is the permittivity of free space and $\chi(t)$ is the susceptibility time-function of the medium. In the present discussion, $\chi(t)$ is given by (9.2) and (9.5), respectively, for Debye and Lorentz dispersive media having a single pole or single pole pair.

Previous recursive-convolution formulations [1, 2] implementing (9.11) approximated the continuous time function $E(t)$ by a constant value over each time-step Δt. However, greater accuracy results if a linear approximation is used. For example, over any given time interval $[i\Delta t, (i+1)\Delta t]$, $E(t)$ can be approximated by linearly interpolating the values of E at the beginning and end of the time interval:

$$\boldsymbol{E}(t) = \boldsymbol{E}^i + \left(\frac{\boldsymbol{E}^{i+1} - \boldsymbol{E}^i}{\Delta t} \right) \cdot (t - i\Delta t) \tag{9.12}$$

We see that implementing (9.12) separately over each Δt in the process results in an overall piecewise-linear temporal approximation of the E-field.

Preparing to use the piecewise-linear temporal approximation for $E(t)$ given in (9.12) in the convolution integral of (9.11), we first rewrite (9.12) in a time-reversed and offset form:

$$\boldsymbol{E}(n\Delta t - \tau) = \boldsymbol{E}^{n-m} + \left(\frac{\boldsymbol{E}^{n-m-1} - \boldsymbol{E}^{n-m}}{\Delta t} \right) \cdot (\tau - m\Delta t) \tag{9.13}$$

After substituting (9.13) into (9.11), we obtain after some manipulation

$$\boldsymbol{D}^n = \varepsilon_0 \varepsilon_\infty \boldsymbol{E}^n + \varepsilon_0 \sum_{m=0}^{n-1} \left[\boldsymbol{E}^{n-m} \chi^m + \left(\boldsymbol{E}^{n-m-1} - \boldsymbol{E}^{n-m} \right) \xi^m \right] \tag{9.14}$$

where

$$\chi^m = \int_{m\Delta t}^{(m+1)\Delta t} \chi(\tau)\, d\tau \; ; \qquad \xi^m = \frac{1}{\Delta t} \int_{m\Delta t}^{(m+1)\Delta t} (\tau - m\Delta t)\, \chi(\tau)\, d\tau \tag{9.15a, b}$$

After substituting (9.14) for $\boldsymbol{D}^n$ (and a similar expression for $\boldsymbol{D}^{n+1}$) into the discrete-time form of Ampere's law

$$\nabla \times \boldsymbol{H}^{n+1/2} = \frac{\boldsymbol{D}^{n+1} - \boldsymbol{D}^n}{\Delta t} \tag{9.16}$$

we obtain the general form of the FDTD update equation for the E-field:

$$\boldsymbol{E}^{n+1} = \left(\frac{\varepsilon_\infty - \xi^0}{\varepsilon_\infty - \xi^0 + \chi^0}\right)\boldsymbol{E}^n + \left(\frac{\Delta t / \varepsilon_0}{\varepsilon_\infty - \xi^0 + \chi^0}\right)\nabla \times \boldsymbol{H}^{n+1/2}$$
$$+ \left(\frac{1}{\varepsilon_\infty - \xi^0 + \chi^0}\right)\sum_{m=0}^{n-1}\left[\boldsymbol{E}^{n-m}\Delta\chi^m + \left(\boldsymbol{E}^{n-m-1} - \boldsymbol{E}^{n-m}\right)\Delta\xi^m\right] \tag{9.17}$$

where

$$\Delta\chi^m = \chi^m - \chi^{m+1} \; ; \quad \Delta\xi^m = \xi^m - \xi^{m+1} \tag{9.18a, b}$$

We now designate the summation in (9.17) as the new time-dependent vector variable $\boldsymbol{\psi}$:

$$\boldsymbol{\psi}^n = \sum_{m=0}^{n-1}\left[\boldsymbol{E}^{n-m}\Delta\chi^m + \left(\boldsymbol{E}^{n-m-1} - \boldsymbol{E}^{n-m}\right)\Delta\xi^m\right] \tag{9.19}$$

For Debye and Lorentz media, (9.19) can be efficiently evaluated by using a recursion relation called in [3] the *recursive accumulator*. This is of the form

$$\boldsymbol{\psi}^n = \left(\Delta\chi^0 - \Delta\xi^0\right)\boldsymbol{E}^n + \Delta\xi^0\,\boldsymbol{E}^{n-1} + C_{\text{rec}}\,\boldsymbol{\psi}^{n-1} \tag{9.20}$$

where (as will be shown) $\Delta\chi^0$, $\Delta\xi^0$, and C_{rec} are real-valued constants for Debye media and complex-valued for Lorentz media. Now, E-field update equation (9.17) becomes

$$\boldsymbol{E}^{n+1} = \left(\frac{\varepsilon_\infty - \xi^0}{\varepsilon_\infty - \xi^0 + \chi^0}\right)\boldsymbol{E}^n + \left(\frac{\Delta t / \varepsilon_0}{\varepsilon_\infty - \xi^0 + \chi^0}\right)\nabla \times \boldsymbol{H}^{n+1/2}$$
$$+ \left(\frac{1}{\varepsilon_\infty - \xi^0 + \chi^0}\right)\boldsymbol{\psi}^n \tag{9.21}$$

Equations (9.20) and (9.21) permit an efficient updating of the E-field without the need to explicitly evaluate the convolutional sum embedded in (9.17). The primary algorithmic requirement is the storage and updating of an additional real or complex vector variable $\boldsymbol{\psi}$, representing the recursive accumulator, for each $\boldsymbol{E}$ in the space lattice.

In proceeding from time-step n to time-step $n+1$, each $\boldsymbol{E}^n$ is first stored in the temporary vector variable $\boldsymbol{E}_{\text{store}}$. Then, $\boldsymbol{E}^{n+1}$ is obtained by applying (9.21) to $\boldsymbol{E}^n$, $\nabla\times\boldsymbol{H}^{n+1/2}$, and $\boldsymbol{\psi}^n$. Then, $\boldsymbol{\psi}^{n+1}$ is computed by applying (9.20) to $\boldsymbol{E}^{n+1}$, $\boldsymbol{E}_{\text{store}}$, and $\boldsymbol{\psi}^n$. The entire cycle then repeats. This technique works because the value of $\boldsymbol{\psi}$ is zero at $n = 0$.

9.3.2 Application to Debye Media

To use the PLRC algorithm, the quantities χ^m, $\Delta\chi^m$, ξ^m, $\Delta\xi^m$, and C_{rec} must be calculated. All of these depend on the susceptibility function of the dispersive medium. Substituting the time-domain susceptibility function (9.2) for Debye media into (9.15a) and (9.15b), we obtain for the p'th Debye pole

$$\chi_p^m = \Delta\varepsilon_p\left(1 - e^{-\Delta t/\tau_p}\right)e^{-m\Delta t/\tau_p} \tag{9.22a}$$

$$\xi_p^m = \frac{\Delta\varepsilon_p\,\tau_p}{\Delta t}\left[1 - (\Delta t/\tau_p + 1)\,e^{-\Delta t/\tau_p}\right]e^{-m\Delta t/\tau_p} \tag{9.22b}$$

The quantities $\Delta\chi^m$ and $\Delta\xi^m$ are found by using (9.15a, b) in (9.18a, b). Noting that

$$\Delta\chi_p^{m+1} = \Delta\chi_p^m\,e^{-\Delta t/\tau_p}\;; \qquad \Delta\xi_p^{m+1} = \Delta\xi_p^m\,e^{-\Delta t/\tau_p} \tag{9.23a, b}$$

and following a procedure similar to that outlined in the appendix of [1], the recursion relation for $\boldsymbol{\psi}^n$ is given by

$$\boldsymbol{\psi}_p^n = \left(\Delta\chi_p^0 - \Delta\xi_p^0\right)\boldsymbol{E}^n + \Delta\xi_p^0\,\boldsymbol{E}^{n-1} + \boldsymbol{\psi}_p^{n-1}e^{-\Delta t/\tau_p} \tag{9.24}$$

All quantities in (9.24) are real valued.

9.3.3 Application to Lorentz Media

For Lorentz media, the real-valued time-domain susceptibility function $\chi_p(t)$ given in (9.5) does not lead to simple recursion relations for χ^m and ξ^m. This problem is circumvented as in [2] by introducing the complex-valued quasi-time-domain function

$$\hat{\chi}_p(t) = -j\,\gamma_p\,e^{(-\alpha_p + j\beta_p)t}\,U(t) \tag{9.25}$$

where the carat " ^ " denotes a complex number and

$$\alpha_p = \delta_p\;; \qquad \beta_p = \sqrt{\omega_p^2 - \delta_p^2}\;; \qquad \gamma_p = \Delta\varepsilon_p\,\omega_p^2/\beta_p \tag{9.26a, b, c}$$

We see that $\chi_p(t)$ is simply the real part of $\hat{\chi}_p(t)$. Substituting $\hat{\chi}_p(t)$ for $\chi(t)$ in (9.15a, b), we obtain for the p'th Lorentz pole pair

$$\hat{\chi}_p^m = \frac{-j\gamma_p}{\alpha_p - j\beta_p}\left[1 - e^{(-\alpha_p + j\beta_p)\Delta t}\right] e^{(-\alpha_p + j\beta_p) m\Delta t} \tag{9.27a}$$

$$\hat{\xi}_p^m = \frac{j\gamma_p / \Delta t}{(\alpha_p - j\beta_p)^2}\left\{\left[(\alpha_p - j\beta_p)\Delta t + 1\right] e^{(-\alpha_p + j\beta_p)\Delta t} - 1\right\} e^{(-\alpha_p + j\beta_p) m\Delta t} \tag{9.27b}$$

These have the properties

$$\hat{\chi}_p^{m+1} = \hat{\chi}_p^m e^{(-\alpha_p + j\beta_p)\Delta t} \quad ; \quad \hat{\xi}_p^{m+1} = \hat{\xi}_p^m e^{(-\alpha_p + j\beta_p)\Delta t} \tag{9.28a, b}$$

and are related to the real-valued quantities χ_p^m and ξ_p^m by

$$\chi_p^m = \mathrm{Re}\left(\hat{\chi}_p^m\right) \quad ; \quad \xi_p^m = \mathrm{Re}\left(\hat{\xi}_p^m\right) \tag{9.29a, b}$$

The values of χ_p^0 and ξ_p^0 found using (9.27a, b) and (9.29a, b) with $m = 0$ can be substituted directly into the E-field update equation (9.21).

A complex-valued recursive accumulator $\hat{\boldsymbol{\psi}}^n$ can be defined for the p'th Lorentz pole pair as

$$\hat{\boldsymbol{\psi}}_p^n = \sum_{m=0}^{n-1}\left[\boldsymbol{E}^{n-m}\Delta\hat{\chi}_p^m + \left(\boldsymbol{E}^{n-m-1} - \boldsymbol{E}^{n-m}\right)\Delta\hat{\xi}_p^m\right] \tag{9.30}$$

where

$$\Delta\hat{\chi}_p^m = \hat{\chi}_p^m - \hat{\chi}_p^{m+1} \quad ; \quad \Delta\hat{\xi}_p^m = \hat{\xi}_p^m - \hat{\xi}_p^{m+1} \tag{9.31a, b}$$

Again applying the procedure outlined in the appendix of [1], the recursion relation for $\hat{\boldsymbol{\psi}}^n$ is given by

$$\hat{\boldsymbol{\psi}}_p^n = \left(\Delta\hat{\chi}_p^0 - \Delta\hat{\xi}_p^0\right)\boldsymbol{E}^n + \Delta\hat{\xi}_p^0\,\boldsymbol{E}^{n-1} + \hat{\boldsymbol{\psi}}_p^{n-1} e^{(-\alpha_p + j\beta_p)\Delta t} \tag{9.32}$$

Noting (9.29), and comparing (9.30) with (9.19), we see that

$$\boldsymbol{\psi}_p^n = \mathrm{Re}\left(\hat{\boldsymbol{\psi}}_p^n\right) \tag{9.33}$$

Thus, in the PLRC algorithm for Lorentz media, the complex-valued recursive accumulator is updated using (9.32). However, the E-field update equation (9.21) uses the real-valued recursive accumulator found via (9.33).

9.3.4 Numerical Results

Reference [3] reports applying the PLRC method to calculate the reflection coefficient and transmitted field for a plane-wave pulse normally incident from free space onto a Lorentz dispersive half-space. Referring to (9.4), the Lorentz half-space in this example is characterized by a single pole pair with the parameters $\varepsilon_{s,p} = 3.0$, $\varepsilon_{\infty,p} = 1.5$, $\omega_p = 2\pi \times 20$ GHz, and $\delta_p = 0.1\omega_p$. The one-dimensional FDTD model has $\Delta = 0.25$ mm and $\Delta t = 0.833$ ps (the Courant stability limit). This corresponds to an FDTD grid density in the half-space medium of about 22 cells per wavelength at the pole-pair frequency of 20 GHz, and about 10 cells per wavelength at 100 GHz, the approximate upper bound of the spectrum of the impinging Gaussian pulse.

Fig. 9.1 shows the E-field distribution after 900 time-steps within the Lorentz dispersive half-space. Curves are shown for the exact solution, the PLRC-FDTD method discussed here, and the original RC-FDTD method introduced in [2]. We see that the PLRC results have greatly improved accuracy relative to the original RC method. In particular, the PLRC method picks up the sinusoidal behavior of the field in the precursor region located ahead of the main body of the transmitted pulse, whereas the RC method erroneously indicates a damped response with little oscillation. Further, the PLRC method provides a much more accurate calculation of the positive and negative peaks of the field distribution in the main body of the pulse.

The PLRC method can be extended to model dispersive media characterized by a finite sum of Debye poles and Lorentz pole pairs, by the existence of a fixed conductivity σ in addition to the dispersion, and by the Drude dispersion. These extensions are left as homework exercises.

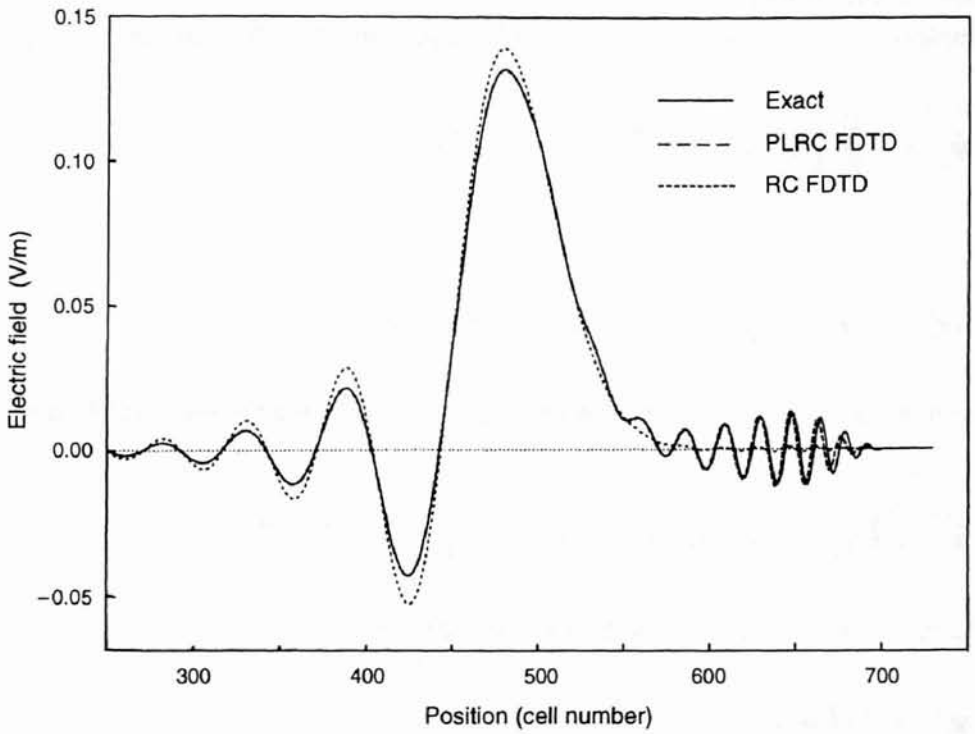


Fig. 9.1 Comparison of RC-FDTD, PLRC-FDTD, and exact results for the E-field distribution after 900 time-steps within a single pole-pair Lorentz dispersive half-space subjected to an impulsive illumination. *Source:* Kelley and Luebbers, *IEEE Trans. Antennas and Propagation*, 1996, pp. 792–797, © 1996 IEEE.

9.4 AUXILIARY DIFFERENTIAL EQUATION METHOD, LINEAR MATERIAL CASE

The second category of methods for implementing FDTD models of dispersive materials utilizes time-domain auxiliary differential equations linking the polarization and the electric flux density [4, 5]. These equations are time-stepped in synchronism with Maxwell's curl equations, yielding a composite self-consistent system. ADE methods have the same second-order accuracy as the PLRC techniques discussed in Section 9.3. In addition, their time-domain basis makes modeling of arbitrary nonlinear dispersive media particularly attractive.

Reference [6] reported an efficient reformulation of the ADE method of [4, 5] for a dispersive material characterized by a frequency-domain susceptibility function having multiple (P) real poles or pole pairs. This reformulation eliminates the previous need to solve a system of P linear equations at each $\boldsymbol{E}$ component at each time-step. Using the approach of [6], the ADE method now requires an equal or smaller number of unknowns to be stored than the corresponding PLRC scheme, and fewer floating-point operations, since no complex-number arithmetic is needed. This section discusses the formulation of the ADE technique of [6] for dispersive dielectric media characterized by multiple Debye, Lorentz, or Drude poles.

9.4.1 Formulation for Multiple Debye Poles

Consider a multiterm Debye dispersive medium having a total of P poles in its susceptibility response. At any particular $\boldsymbol{E}$ observation point, Ampere's law in the time domain can be expressed for this medium as

$$\nabla \times \boldsymbol{H} = \varepsilon_0 \varepsilon_\infty \frac{\partial \boldsymbol{E}}{\partial t} + \sigma \boldsymbol{E} + \sum_{p=1}^{P} \boldsymbol{J}_p \tag{9.34}$$

where $\boldsymbol{J}_p$ is the polarization current associated with the p'th Debye pole. The goal of the ADE technique is to develop a simple time-stepping scheme for $\boldsymbol{J}_p$ that can be updated synchronously with (9.34).

To this end, we express the relative permittivity of the Debye medium in the frequency domain as

$$\varepsilon(\omega) = \varepsilon_\infty + \sum_{p=1}^{P} \frac{\Delta\varepsilon_p}{1 + j\omega\tau_p} \tag{9.35}$$

where p denotes the number of the Debye pole, $\Delta\varepsilon_p = \varepsilon_{s,p} - \varepsilon_{\infty,p}$ is the change in relative permittivity due to the p'th pole, and τ_p is the pole relaxation time. A phasor polarization current is associated with each pole by

$$\breve{\boldsymbol{J}}_p = \varepsilon_0 \Delta\varepsilon_p \left(\frac{j\omega}{1 + j\omega\tau_p} \right) \breve{\boldsymbol{E}} \tag{9.36}$$

An efficient means to obtain $\boldsymbol{J}_p$ from (9.36) is to first multiply both sides of this equation by $(1 + j\omega\tau_p)$. This gives

$$\breve{\boldsymbol{J}}_p + j\omega\tau_p \breve{\boldsymbol{J}}_p = \varepsilon_0 \Delta\varepsilon_p j\omega \breve{\boldsymbol{E}} \tag{9.37a}$$

Conveniently exploiting the differentiation theorem for the Fourier transform, we perform an inverse Fourier transformation of each term of (9.37a):

$$\boldsymbol{J}_p + \tau_p \frac{\partial \boldsymbol{J}_p}{\partial t} = \varepsilon_0 \Delta\varepsilon_p \frac{\partial \boldsymbol{E}}{\partial t} \tag{9.37b}$$

Equation (9.37b) is the required ADE for $\boldsymbol{J}_p$. This can be easily and accurately implemented in an FDTD code using the semi-implicit scheme, wherein yet-to-be-computed fields at time-step $n+1$ are used to create an update formula for a field known at time-step n. Using this strategy, we write the following finite-difference expression for (9.37b), centered at $n+1/2$:

$$\left(\frac{\boldsymbol{J}_p^{n+1} + \boldsymbol{J}_p^n}{2}\right) + \tau_p\left(\frac{\boldsymbol{J}_p^{n+1} - \boldsymbol{J}_p^n}{\Delta t}\right) = \varepsilon_0 \Delta\varepsilon_p\left(\frac{\boldsymbol{E}^{n+1} - \boldsymbol{E}^n}{\Delta t}\right) \tag{9.38}$$

Solving (9.38) for $\boldsymbol{J}_p^{n+1}$, we obtain

$$\boldsymbol{J}_p^{n+1} = k_p \boldsymbol{J}_p^n + \beta_p\left(\frac{\boldsymbol{E}^{n+1} - \boldsymbol{E}^n}{\Delta t}\right) \tag{9.39}$$

where

$$k_p = \frac{1 - \Delta t/2\tau_p}{1 + \Delta t/2\tau_p}; \qquad \beta_p = \frac{\varepsilon_0 \Delta\varepsilon_p \Delta t/\tau_p}{1 + \Delta t/2\tau_p} \tag{9.40a, b}$$

The second component of the ADE algorithm involves solving (9.34) for $\boldsymbol{E}^{n+1}$. We again use a semi-implicit scheme centered at $n+1/2$. This requires knowledge of $\boldsymbol{J}_p^{n+1/2}$, which can be obtained from (9.39) by

$$\boldsymbol{J}_p^{n+1/2} = \frac{1}{2}\left(\boldsymbol{J}_p^n + \boldsymbol{J}_p^{n+1}\right) = \frac{1}{2}\left[\left(1 + k_p\right)\boldsymbol{J}_p^n + \frac{\beta_p}{\Delta t}\left(\boldsymbol{E}^{n+1} - \boldsymbol{E}^n\right)\right] \tag{9.41}$$

Now, we can evaluate (9.34) at time-step $n+1/2$:

$$\begin{aligned}\nabla \times \boldsymbol{H}^{n+1/2} = {} & \varepsilon_0 \varepsilon_\infty\left(\frac{\boldsymbol{E}^{n+1} - \boldsymbol{E}^n}{\Delta t}\right) + \sigma\left(\frac{\boldsymbol{E}^{n+1} + \boldsymbol{E}^n}{2}\right) \\ & + \frac{1}{2}\sum_{p=1}^{P}\left[\left(1 + k_p\right)\boldsymbol{J}_p^n + \frac{\beta_p}{\Delta t}\left(\boldsymbol{E}^{n+1} - \boldsymbol{E}^n\right)\right]\end{aligned} \tag{9.42}$$

Upon collecting like terms, we obtain the following explicit time-stepping relation for $\boldsymbol{E}$:

$$\boldsymbol{E}^{n+1} = \left(\frac{2\varepsilon_0 \varepsilon_\infty + \sum_{p=1}^{P} \beta_p - \sigma \Delta t}{2\varepsilon_0 \varepsilon_\infty + \sum_{p=1}^{P} \beta_p + \sigma \Delta t} \right) \boldsymbol{E}^n + \left(\frac{2\Delta t}{2\varepsilon_0 \varepsilon_\infty + \sum_{p=1}^{P} \beta_p + \sigma \Delta t} \right) \cdot \left[\nabla \times \boldsymbol{H}^{n+1/2} - \frac{1}{2} \sum_{p=1}^{P} \left(1 + k_p\right) \boldsymbol{J}_p^n \right] \tag{9.43}$$

Thus, the ADE-FDTD algorithm for modeling a dispersive medium with P Debye poles is a three-step fully explicit procedure. Starting with the assumed known (stored) component values of $\boldsymbol{E}^n$, $\boldsymbol{J}_p^n$, and $\boldsymbol{H}^{n+1/2}$, we first calculate the new $\boldsymbol{E}^{n+1}$ components using (9.43). Second, we calculate the new $\boldsymbol{J}_p^{n+1}$ components using (9.39) applied to the just-computed set of $\boldsymbol{E}^{n+1}$ components. Finally, $\boldsymbol{H}^{n+3/2}$ is obtained from $\boldsymbol{H}^{n+1/2}$ and $\boldsymbol{E}^{n+1}$ in the usual manner from the Yee realization of Faraday's law, and the complete cycle starts again. Only P additional real variables are required for each $\boldsymbol{E}$ component.

9.4.2 Formulation for Multiple Lorentz Pole Pairs

We next consider a multiterm Lorentz dispersive medium having a total of P pole pairs in its susceptibility response. Ampere's law in the time domain for this medium is again given by (9.34); and again the goal of the ADE technique is to develop a simple time-stepping scheme for $\boldsymbol{J}_p$, which can be updated synchronously with (9.34).

To this end, we express the relative permittivity of the Lorentz medium in the frequency domain as

$$\varepsilon(\omega) = \varepsilon_\infty + \sum_{p=1}^{P} \frac{\Delta\varepsilon_p \omega_p^2}{\omega_p^2 + 2j\omega\delta_p - \omega^2} \tag{9.44}$$

where p denotes the number of the Lorentz pole pair, $\Delta\varepsilon_p = \varepsilon_{s,p} - \varepsilon_{\infty,p}$ is the change in relative permittivity due to the p'th pole pair, ω_p is the undamped frequency of the p'th pole pair, and δ_p is its damping factor. A phasor polarization current is associated with each pole pair by

$$\breve{\boldsymbol{J}}_p = \varepsilon_0 \Delta\varepsilon_p \omega_p^2 \left(\frac{j\omega}{\omega_p^2 + 2j\omega\delta_p - \omega^2} \right) \breve{\boldsymbol{E}} \tag{9.45}$$

An efficient means to obtain $\boldsymbol{J}_p$ from (9.45) is to first multiply both sides of this equation by $(\omega_p^2 + 2j\omega\delta_p - \omega^2)$. This gives

$$\omega_p^2 \breve{\boldsymbol{J}}_p + 2j\omega\delta_p \breve{\boldsymbol{J}}_p - \omega^2 \breve{\boldsymbol{J}}_p = \varepsilon_0 \Delta\varepsilon_p \omega_p^2 j\omega \breve{\boldsymbol{E}} \tag{9.46a}$$

Conveniently exploiting the differentiation theorem for the Fourier transform, we perform an inverse Fourier transformation of each term of (9.46a):

$$\omega_p^2 \boldsymbol{J}_p + 2\delta_p \frac{\partial \boldsymbol{J}_p}{\partial t} + \frac{\partial^2 \boldsymbol{J}_p}{\partial t^2} = \varepsilon_0 \Delta\varepsilon_p \omega_p^2 \frac{\partial \boldsymbol{E}}{\partial t} \tag{9.46b}$$

Equation (9.46b) is the required ADE for $\boldsymbol{J}_p$. This can be easily and accurately implemented in an FDTD code using the semi-implicit scheme, wherein yet-to-be-computed fields at time-step $n+1$ are used to create an update formula for a field known at time-step n. Using this strategy, we write the following finite-difference expression for (9.46b), centered at time-step n:

$$\omega_p^2 \boldsymbol{J}_p^n + 2\delta_p \left(\frac{\boldsymbol{J}_p^{n+1} - \boldsymbol{J}_p^{n-1}}{2\Delta t} \right) + \left[\frac{\boldsymbol{J}_p^{n+1} - 2\boldsymbol{J}_p^n + \boldsymbol{J}_p^{n-1}}{(\Delta t)^2} \right] = \varepsilon_0 \Delta\varepsilon_p \omega_p^2 \left(\frac{\boldsymbol{E}^{n+1} - \boldsymbol{E}^{n-1}}{2\Delta t} \right) \tag{9.47}$$

Solving (9.47) for $\boldsymbol{J}_p^{n+1}$, we obtain

$$\boldsymbol{J}_p^{n+1} = \alpha_p \boldsymbol{J}_p^n + \xi_p \boldsymbol{J}_p^{n-1} + \gamma_p \left(\frac{\boldsymbol{E}^{n+1} - \boldsymbol{E}^{n-1}}{2\Delta t} \right) \tag{9.48}$$

where

$$\alpha_p = \frac{2 - \omega_p^2 (\Delta t)^2}{1 + \delta_p \Delta t}; \quad \xi_p = \frac{\delta_p \Delta t - 1}{\delta_p \Delta t + 1}; \quad \gamma_p = \frac{\varepsilon_0 \Delta\varepsilon_p \omega_p^2 (\Delta t)^2}{1 + \delta_p \Delta t} \tag{9.49a, b, c}$$

The second component of the ADE algorithm involves solving (9.34) for $\boldsymbol{E}^{n+1}$. To this end, we use a semi-implicit scheme centered at time-step $n+1/2$. This requires knowledge of $\boldsymbol{J}_p^{n+1/2}$, which can be obtained from (9.48) by

$$\boldsymbol{J}_p^{n+1/2} = \frac{1}{2}\left(\boldsymbol{J}_p^n + \boldsymbol{J}_p^{n+1} \right) = \frac{1}{2}\left[\left(1 + \alpha_p\right) \boldsymbol{J}_p^n + \xi_p \boldsymbol{J}_p^{n-1} + \frac{\gamma_p}{2\Delta t} \left(\boldsymbol{E}^{n+1} - \boldsymbol{E}^{n-1} \right) \right] \tag{9.50}$$

Now, we can evaluate (9.34) at time-step $n+1/2$:

$$\begin{aligned} \nabla \times \boldsymbol{H}^{n+1/2} = {} & \varepsilon_0 \varepsilon_\infty \left(\frac{\boldsymbol{E}^{n+1} - \boldsymbol{E}^n}{\Delta t} \right) + \sigma \left(\frac{\boldsymbol{E}^{n+1} + \boldsymbol{E}^n}{2} \right) \\ & + \frac{1}{2} \sum_{p=1}^{P} \left[\left(1 + \alpha_p\right) \boldsymbol{J}_p^n + \xi_p \boldsymbol{J}_p^{n-1} + \frac{\gamma_p}{2\Delta t} \left(\boldsymbol{E}^{n+1} - \boldsymbol{E}^{n-1} \right) \right] \end{aligned} \tag{9.51}$$

Upon collecting like terms, we obtain the following explicit time-stepping relation for $\boldsymbol{E}$:

$$\boldsymbol{E}^{n+1} = C_1 \boldsymbol{E}^{n-1} + C_2 \boldsymbol{E}^{n} + C_3 \cdot \left\{ \nabla \times \boldsymbol{H}^{n+1/2} - \frac{1}{2} \sum_{p=1}^{P} \left[\left(1 + \alpha_p\right) \boldsymbol{J}_p^{n} + \xi_p \boldsymbol{J}_p^{n-1} \right] \right\} \tag{9.52}$$

where

$$C_1 = \frac{\frac{1}{2} \sum_{p=1}^{P} \gamma_p}{2\varepsilon_0 \varepsilon_\infty + \frac{1}{2} \sum_{p=1}^{P} \gamma_p + \sigma \Delta t} \tag{9.53a}$$

$$C_2 = \frac{2\varepsilon_0 \varepsilon_\infty - \sigma \Delta t}{2\varepsilon_0 \varepsilon_\infty + \frac{1}{2} \sum_{p=1}^{P} \gamma_p + \sigma \Delta t} \tag{9.53b}$$

$$C_3 = \frac{2 \Delta t}{2\varepsilon_0 \varepsilon_\infty + \frac{1}{2} \sum_{p=1}^{P} \gamma_p + \sigma \Delta t} \tag{9.53c}$$

Thus, the ADE-FDTD algorithm for modeling a dispersive medium with P Lorentz pole pairs is a three-step fully explicit procedure. Starting with the assumed known (stored) component values of $\boldsymbol{E}^{n-1}$, $\boldsymbol{E}^{n}$, $\boldsymbol{J}_p^{n-1}$, $\boldsymbol{J}_p^{n}$, and $\boldsymbol{H}^{n+1/2}$, we first calculate the new $\boldsymbol{E}^{n+1}$ components using (9.52). Second, we calculate the new $\boldsymbol{J}_p^{n+1}$ components using (9.48) applied to the just-computed set of $\boldsymbol{E}^{n+1}$ components. Finally, $\boldsymbol{H}^{n+3/2}$ is obtained from $\boldsymbol{H}^{n+1/2}$ and $\boldsymbol{E}^{n+1}$ in the usual manner from the Yee realization of Faraday's law, and the complete cycle starts again. Only $3P$ additional real variables are required for each $\boldsymbol{E}$ component.

9.4.3 Formulation for Multiple Drude Poles

We last consider a multiterm Drude dispersive medium having a total of P poles in its susceptibility response. Ampere's law in the time domain for this medium is again given by (9.34); and again the goal of the ADE technique is to develop a simple time-stepping scheme for $\boldsymbol{J}_p$ which can be updated synchronously with (9.34).

To this end, we express the relative permittivity of the Drude medium in the frequency domain as

$$\varepsilon(\omega) = \varepsilon_\infty - \sum_{p=1}^{P} \frac{\omega_p^2}{\omega^2 - j\omega\gamma_p} \tag{9.54}$$

where p denotes the number of the Drude pole, ω_p is the frequency of the p'th pole, and γ_p is the inverse of the pole relaxation time. A phasor polarization current is associated with each pole by

$$\breve{\boldsymbol{J}}_p = -j\omega\varepsilon_0\left(\frac{\omega_p^2}{\omega^2 - j\omega\gamma_p}\right)\breve{\boldsymbol{E}} \tag{9.55}$$

An efficient means to obtain $\boldsymbol{J}_p$ from (9.55) is to first multiply both sides of this equation by $(\omega^2 - j\omega\gamma_p)$. This gives

$$\omega^2\breve{\boldsymbol{J}}_p - j\omega\gamma_p\breve{\boldsymbol{J}}_p = -j\omega\varepsilon_0\omega_p^2\breve{\boldsymbol{E}} \tag{9.56a}$$

Conveniently exploiting the differentiation theorem for the Fourier transform, we perform an inverse Fourier transformation of each term of (9.56a):

$$\frac{\partial^2\boldsymbol{J}_p}{\partial t^2} + \gamma_p\frac{\partial\boldsymbol{J}_p}{\partial t} = \varepsilon_0\omega_p^2\frac{\partial\boldsymbol{E}}{\partial t} \tag{9.56b}$$

Integrating once with respect to time, and omitting the trivial constant solution, yields

$$\frac{\partial\boldsymbol{J}_p}{\partial t} + \gamma_p\boldsymbol{J}_p = \varepsilon_0\omega_p^2\boldsymbol{E} \tag{9.56c}$$

Equation (9.56c) is the required ADE for $\boldsymbol{J}_p$. This can be easily and accurately implemented in an FDTD code using the semi-implicit scheme, wherein yet-to-be-computed fields at time-step $n+1$ are used to create an update formula for a field known at time-step n. Using this strategy, we write the following finite-difference expression for (9.56c), centered at time-step $n+1/2$:

$$\left(\frac{\boldsymbol{J}_p^{n+1} - \boldsymbol{J}_p^n}{\Delta t}\right) + \gamma_p\left(\frac{\boldsymbol{J}_p^{n+1} + \boldsymbol{J}_p^n}{2}\right) = \varepsilon_0\omega_p^2\left(\frac{\boldsymbol{E}^{n+1} + \boldsymbol{E}^n}{2}\right) \tag{9.57}$$

Solving (9.57) for $\boldsymbol{J}_p^{n+1}$, we obtain

$$\boldsymbol{J}_p^{n+1} = k_p\boldsymbol{J}_p^n + \beta_p\left(\boldsymbol{E}^{n+1} + \boldsymbol{E}^n\right) \tag{9.58}$$

where

$$k_p = \frac{1 - \gamma_p\Delta t/2}{1 + \gamma_p\Delta t/2}\ ; \qquad \beta_p = \frac{\omega_p^2\varepsilon_0\Delta t/2}{1 + \gamma_p\Delta t/2} \tag{9.59a, b}$$

The second component of the ADE algorithm involves solving (9.34) for $\boldsymbol{E}^{n+1}$. We again use a semi-implicit scheme centered at $n+1/2$. This requires knowledge of $\boldsymbol{J}_p^{n+1/2}$, which can be obtained from (9.58) by

$$\boldsymbol{J}_p^{n+1/2} = \frac{1}{2}\left(\boldsymbol{J}_p^n + \boldsymbol{J}_p^{n+1}\right) = \frac{1}{2}\left[\left(1+k_p\right)\boldsymbol{J}_p^n + \beta_p\left(\boldsymbol{E}^{n+1}+\boldsymbol{E}^n\right)\right] \tag{9.60}$$

Now, we can evaluate (9.34) at time-step $n+1/2$:

$$\nabla\times\boldsymbol{H}^{n+1/2} = \varepsilon_0\,\varepsilon_\infty\left(\frac{\boldsymbol{E}^{n+1}-\boldsymbol{E}^n}{\Delta t}\right) + \sigma\left(\frac{\boldsymbol{E}^{n+1}+\boldsymbol{E}^n}{2}\right) + \frac{1}{2}\sum_{p=1}^{P}\left[\left(1+k_p\right)\boldsymbol{J}_p^n + \beta_p\left(\boldsymbol{E}^{n+1}+\boldsymbol{E}^n\right)\right] \tag{9.61}$$

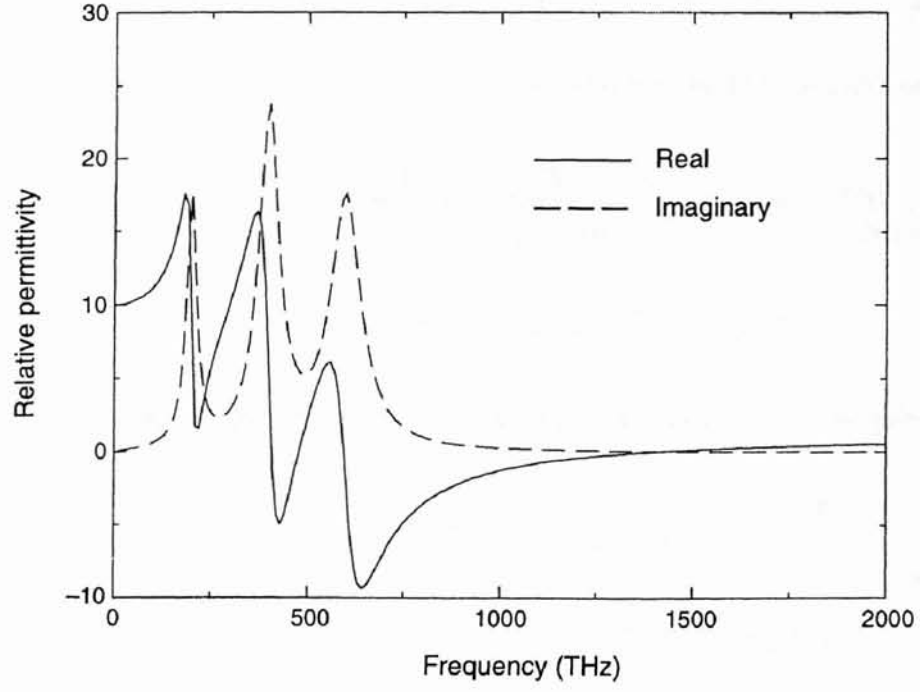
Upon collecting like terms, we obtain the following explicit time-stepping relation for $\boldsymbol{E}$:

$$\boldsymbol{E}^{n+1} = \left(\frac{2\varepsilon_0\,\varepsilon_\infty - \Delta t\sum_{p=1}^{P}\beta_p - \sigma\Delta t}{2\varepsilon_0\,\varepsilon_\infty + \Delta t\sum_{p=1}^{P}\beta_p + \sigma\Delta t}\right)\boldsymbol{E}^n + \left(\frac{2\Delta t}{2\varepsilon_0\,\varepsilon_\infty + \Delta t\sum_{p=1}^{P}\beta_p + \sigma\Delta t}\right)\cdot\left[\nabla\times\boldsymbol{H}^{n+1/2} - \frac{1}{2}\sum_{p=1}^{P}\left(1+k_p\right)\boldsymbol{J}_p^n\right] \tag{9.62}$$

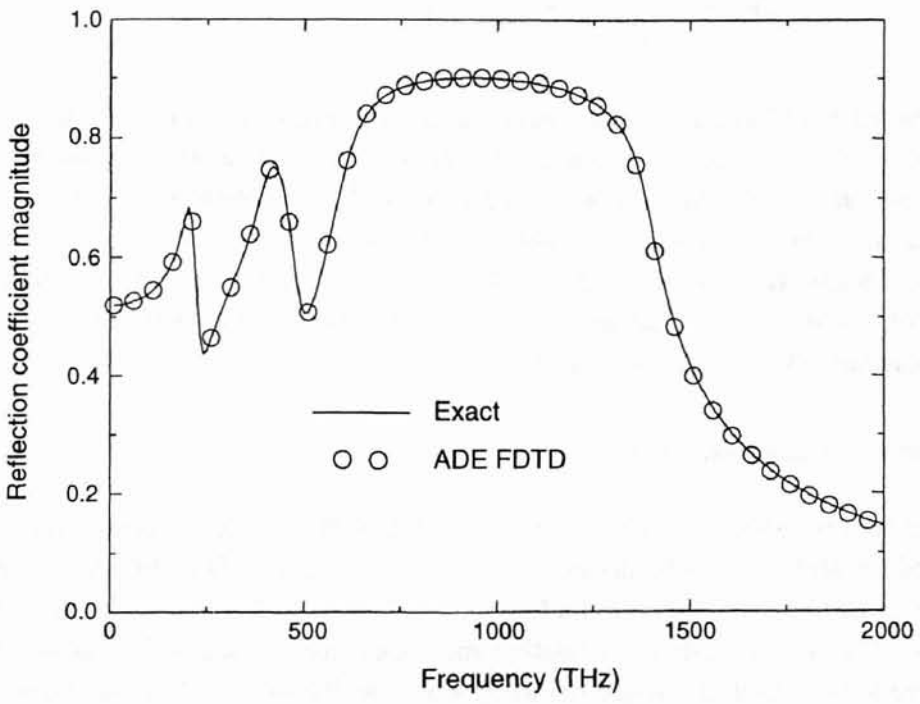
Thus, the ADE-FDTD algorithm for modeling a dispersive medium with P Drude poles is a three-step fully explicit procedure. Starting with the assumed known (stored) component values of $\boldsymbol{E}^n$, $\boldsymbol{J}_p^n$, and $\boldsymbol{H}^{n+1/2}$, we first calculate the new $\boldsymbol{E}^{n+1}$ components using (9.62). Second, we calculate the new $\boldsymbol{J}_p^{n+1}$ components using (9.58) applied to the just-computed set of $\boldsymbol{E}^{n+1}$ components. Finally, $\boldsymbol{H}^{n+3/2}$ is obtained from $\boldsymbol{H}^{n+1/2}$ and $\boldsymbol{E}^{n+1}$ in the usual manner from the Yee realization of Faraday's law, and the complete cycle starts again. Only P additional real variables are required for each $\boldsymbol{E}$ component.

9.4.4 Illustrative Numerical Results

Fig. 9.2 illustrates the accuracy achievable using ADE-FDTD to model reflection from a medium characterized by multiple Lorentzian pole-pairs. Here, ADE-FDTD results for the reflection coefficient of a half-space comprised of a hypothetical dispersive material are compared with the exact solution obtained using monochromatic impedance theory. The FDTD data over the entire spectral range are obtained in a single run using a narrow illuminating Gaussian pulse, and then taking the ratio of the DFTs of the reflected and incident waveforms. With $\varepsilon_s = 10$ and $\varepsilon_\infty = 1$, the half-space is assumed to have three moderately undamped Lorentzian resonances in the optical range: ($f_1 = 2\times10^{14}$ Hz, $\delta_1 = 0.5f_1$), ($f_2 = 4\times10^{14}$ Hz, $\delta_2 = 0.5f_2$), and ($f_3 = 6\times10^{14}$ Hz, $\delta_3 = 0.5f_3$). The ADE-FDTD and exact results agree to within 0.1% at all points shown.



(a) Real and imaginary parts of the permittivity.



(b) Comparison of FDTD and exact results.

Fig. 9.2 Demonstration of the accuracy of ADE-FDTD modeling of reflection from a dispersive half-space characterized by three Lorentzian pole-pairs in the optical range.

9.5 MODELING OF LINEAR MAGNETIZED FERRITES

Magnetized ferrites are used extensively in microwave circuit technology, with specific applications in the design of circulators and isolators. This section presents a new approach to the FDTD modeling of magnetized ferrites. An equivalent lumped circuit of an FDTD cell filled with such a medium is developed. Then, the lumped circuit is used to propose a new FDTD algorithm. This algorithm is verified on canonical examples, and is shown to be simple, accurate, robust, and computationally more effective than the previously published approaches.

FDTD modeling of microwave devices using magnetized ferrites has followed two distinct paths. In the first [7–9], the ferrite is characterized in the time domain by the equation of motion of the magnetization vector. The resulting equations are discretized, and supplement the FDTD algorithm in a manner that imposes modifications of the Yee space lattice. Here, at each spatial point where a standard mesh includes only a magnetic field component perpendicular to the dc magnetization of the ferrite, we must also define the magnetic induction component and another magnetic field component [7]. In this approach, the anisotropic and dispersive properties of the ferrite are inseparably bound together in a new form of the FDTD mesh.

In the second approach [10], the ferrite is characterized in the frequency domain by the Polder vector, which in a lossless case can be expressed as

$$\mu = \mu_0 \begin{bmatrix} \mu_1 & -j\kappa & 0 \\ j\kappa & \mu_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{9.63}$$

with

$$\mu_1 = 1 + \frac{\omega_0 \omega_m}{\omega_0^2 - \omega^2} \quad ; \qquad \kappa = \frac{\omega \omega_m}{\omega_0^2 - \omega^2} \tag{9.64a, b}$$

$$\omega_0 = H_i \gamma \quad ; \qquad \omega_m = M_s \gamma \tag{9.64c, d}$$

where γ is the gyromagnetic ratio, M_s is the saturation magnetization, and H_i is the static biasing magnetic field assumed to be z-directed. Here, the basic FDTD mesh is not modified, and the ferrite's anisotropic properties are modeled by direct mutual influence of the magnetic field components perpendicular to the dc magnetization. The dispersive properties of the ferrite are treated separately. The frequency dependence of these properties is transformed into the time domain through a convolution wherein four complex-valued convolution terms must be updated recursively for each FDTD cell [10]. Thus, for each such cell, four complex (or eight real) variables need to be additionally stored.

This section presents an alternative approach [11] that is simpler and computationally more effective than [7–10]. As in [10], we use a basically unmodified Yee cell, and separate the anisotropic and dispersive properties of the ferrite. However, we account for the material dispersion by considering a simple lumped equivalent resistor-inductor-capacitor (RLC) circuit.

9.5.1 Equivalent RLC Model

We first consider the two-dimensional lossless TM_z case involving field components (E_z, H_x, H_y) propagating between two electric planes, $z = 0$ and $z = a$. Introducing a finite-difference discretization of space with cell size $\Delta = a$, we can write two scalar components of Maxwell's curl equations in the phasor domain as

$$\breve{j}\omega C\breve{V}\big|_{i,j} = \breve{I}_x\big|_{i-1/2,j} - \breve{I}_x\big|_{i+1/2,j} + \breve{I}_y\big|_{i,j-1/2} - \breve{I}_y\big|_{i,j+1/2} \tag{9.65}$$

$$\breve{j}\omega L\breve{I}_x\big|_{i+1/2,j} = \breve{V}\big|_{i,j} - \breve{V}\big|_{i+1,j} \tag{9.66}$$

where $\breve{j} = \sqrt{-1}$, $C = \varepsilon a$, $L = \mu a$, and

$$\breve{V}\big|_{i,j} = a\breve{E}_z\big|_{i,j} \; ; \qquad \breve{I}_x\big|_{i,j} = a\breve{H}_y\big|_{i,j} \; ; \qquad \breve{I}_y\big|_{i,j} = -a\breve{H}_x\big|_{i,j} \tag{9.67a, b, c}$$

The same equations describe an equivalent lumped circuit [12] as presented in Fig. 9.3(a). Let us now consider a ferrite material with dc magnetization along the z-axis, described by the tensor defined in (9.63). Equation (9.65) does not change, while (9.66) is modified to

$$\breve{j}\omega a\mu_0\left[(1+\chi)\breve{I}_x\big|_{i+1/2,j} - \breve{j}\frac{\omega}{\omega_0}\chi\breve{I}_y\big|_{i+1/2,j}\right] = \breve{V}\big|_{i,j} - \breve{V}\big|_{i+1,j} \tag{9.68}$$

with

$$\chi = \frac{\omega_0\omega_m}{\omega_0^2 - \omega^2} \tag{9.69}$$

We note that χ of (9.69) corresponds to a parallel LC circuit resonating at $\omega_0^2 = 1/L_1C_1$ and with $L_1/L = \omega_m/\omega_0$. This is reflected in the equivalent circuit shown in Fig. 9.3(b). We also note that the term $\breve{V}_{Cx}\big|_{i+1/2,j} = -\breve{j}\omega a\mu_0\chi\breve{I}_x\big|_{i+1/2,j}$ describes voltage across the parallel LC circuit. The current $\breve{I}_{Cx}\big|_{i+1/2,j}$ flowing through the capacitor is proportional to that voltage multiplied by $\breve{j}\omega$. Thus, in the anisotropic term in (9.68), we can eliminate the dispersive factor χ after replacing $\breve{I}_y\big|_{i+1/2,j}$ with $\breve{I}_{Cy}\big|_{i+1/2,j}$. Equation (9.68) then transforms to

$$\breve{j}\omega a\mu_0(1+\chi)\breve{I}_x\big|_{i+1/2,j} + \frac{1}{\omega_0 C_1}\breve{I}_{Cy}\big|_{i+1/2,j} = \breve{V}\big|_{i,j} - \breve{V}\big|_{i+1,j} \tag{9.70}$$

Thus, the voltage source describing the anisotropic term is driven by the capacitor current in the perpendicular arms. Since such a current is not available at the position of the source, it must be obtained by averaging the current flowing in the four adjacent perpendicular arms:

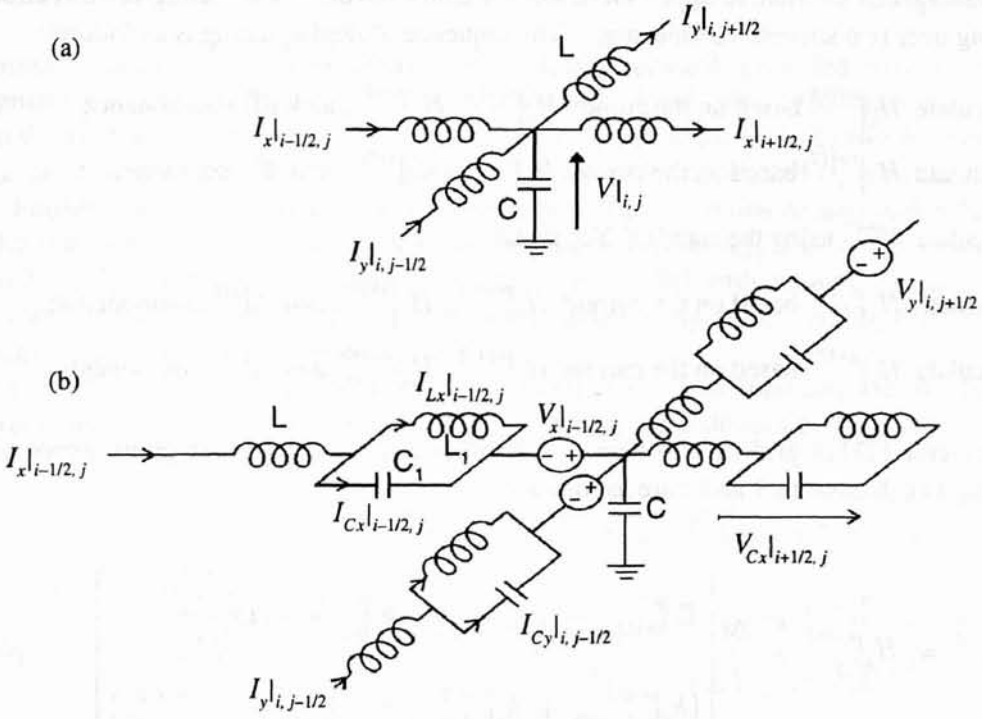


Fig. 9.3 Lumped equivalent circuit of a discretized two-dimensional TM_z structure: (a) simple dielectric case; (b) magnetized ferrite case. *Adapted from*: Gwarek and Moryc, *IEEE Microwave and Wireless Components Lett.*, 2004, pp. 331–333, © 2004 IEEE.

$$\breve{V}_x\big|_{i-1/2,\,j} = \left(\breve{I}_{Cy}\big|_{i,\,j-1/2} + \breve{I}_{Cy}\big|_{i,\,j+1/2} + \breve{I}_{Cy}\big|_{i-1,\,j-1/2} + \breve{I}_{Cy}\big|_{i-1,\,j+1/2}\right) \Big/ 4\omega_0 C_1 \qquad (9.71)$$

9.5.2 Time-Stepping Algorithm

The development of the FDTD time-stepping algorithm is based upon inverse Fourier transformation of equations of the type of (9.65) and (9.70). In the circuit arm describing component H_y, we need to define (and update at each time-step) I_x, which corresponds directly to H_y. We also define and update two other variables, one describing V_{Cx} and the other describing I_{Lx}, or alternatively, $I_{Cx} = I_x - I_{Lx}$. The latter choice is computationally more effective. I_x is calculated from the time-domain discretization of (9.70). Then, we update V_{Cx}, knowing the current injected into the capacitor over the previous iteration; and we update I_{Lx}, knowing the voltage across L_1. The analogous operations are performed in the arms corresponding to H_x. The remaining equation uses standard FDTD time-stepping.

The sequence of time-stepping the H_x and H_y fields must be carefully considered since, in equations like (9.68) or (9.70) containing the effects of off-diagonal anisotropy, it is assumed that we should know (at the same instant of time) H_x to calculate H_y, and H_y to calculate H_x. The implicit relation between H_x and H_y can be easily resolved in the one-dimensional case [10]. The three-dimensional situation additionally involves the need for space-averaging of these components, which makes direct solving of the implicit relation impractical. In previous work reported in the literature, it was solved by either extrapolation or interpolation in time [8].

An advantageous alternative approach discussed below involves alternating the direction of field updating over two successive time-steps. The sequence of field updating is as follows:

1. Calculate $H_y\big|^{n+1/2}$ based on the current $H_y\big|^{n-1/2}$, $H_x\big|^{n-1/2}$, and $E\big|^n$ components;
2. Calculate $H_x\big|^{n+1/2}$ based on the current $H_x\big|^{n-1/2}$, $H_y\big|^{n+1/2}$, and $E\big|^n$ components;
3. Calculate $E\big|^{n+1}$ using the standard Yee update;
4. Calculate $H_x\big|^{n+3/2}$ based on the current $H_x\big|^{n+1/2}$, $H_y\big|^{n+1/2}$, and $E\big|^{n+1}$ components;
5. Calculate $H_y\big|^{n+3/2}$ based on the current $H_y\big|^{n+1/2}$, $H_x\big|^{n+3/2}$, and $E\big|^{n+1}$ components.

For a square-cell FDTD grid of cell size Δ and time-step Δt, the time-stepping expressions corresponding to calculations 1 and 2 are as follows:

$$H_y\big|_{i,j}^{n+1/2} = H_y\big|_{i,j}^{n-1/2} + \frac{\Delta t}{L}\left[\begin{array}{l} E_z\big|_{i+1/2,j}^{n} - E_z\big|_{i-1/2,j}^{n} - p_y\big|_{i,j}^{n} - 0.25\sqrt{\dfrac{L_1}{C_1}}\cdot \\ \left(h_x\big|_{i-1,j-1/2}^{n-1/2} + h_x\big|_{i-1,j+1/2}^{n-1/2} + h_x\big|_{i,j-1/2}^{n-1/2} + h_x\big|_{i,j+1/2}^{n-1/2}\right) \end{array}\right] \tag{9.72}$$

$$h_y\big|_{i,j}^{n+1/2} = h_y\big|_{i,j}^{n-1/2} + H_y\big|_{i,j}^{n+1/2} - H_y\big|_{i,j}^{n-1/2} + \frac{\Delta t}{L_1} p_y\big|_{i,j}^{n} \tag{9.73}$$

$$p_y\big|_{i,j}^{n+1} = p_y\big|_{i,j}^{n} + \frac{\Delta t}{C_1} h_y\big|_{i,j}^{n+1/2} \tag{9.74}$$

$$H_x\big|_{i,j}^{n+1/2} = H_x\big|_{i,j}^{n-1/2} + \frac{\Delta t}{L}\left[\begin{array}{l} E_z\big|_{i,j-1/2}^{n} - E_z\big|_{i,j+1/2}^{n} - p_x\big|_{i,j}^{n} + 0.25\sqrt{\dfrac{L_1}{C_1}}\cdot \\ \left(h_y\big|_{i-1/2,j-1}^{n+1/2} + h_y\big|_{i+1/2,j-1}^{n+1/2} + h_y\big|_{i-1/2,j}^{n+1/2} + h_y\big|_{i+1/2,j}^{n+1/2}\right) \end{array}\right] \tag{9.75}$$

$$h_x\big|_{i,j}^{n+1/2} = h_x\big|_{i,j}^{n-1/2} + H_x\big|_{i,j}^{n+1/2} - H_x\big|_{i,j}^{n-1/2} + \frac{\Delta t}{L_1} p_x\big|_{i,j}^{n} \tag{9.76}$$

$$p_x\big|_{i,j}^{n+1} = p_x\big|_{i,j}^{n} + \frac{\Delta t}{C_1} h_x\big|_{i,j}^{n+1/2} \tag{9.77}$$

where $L = \mu\Delta$, $L_1 = L\omega_m/\omega_0$, $C_1 = 1/(\omega_0^2 L_1)$, and p_x, p_y, h_x, and h_y are four auxiliary algorithm variables required to be stored and updated. These variables correspond, respectively, to V_{Cx}, V_{Cy}, I_{Cx}, and I_{Cy} in Fig. 9.3(b). Analogous updates are implemented for calculations 4 and 5.

9.5.3 Extension to the Three-Dimensional Case, Including Loss

Extension to the three-dimensional case is straightforward, provided that we maintain dc magnetization along the z-axis. In this case, the right-hand side of (9.70), which describes H_y (and the analogous equation for H_x), need only be supplemented with the two remaining terms of the E-field integration. The remaining four equations for E_x, E_y, E_z, and H_z are classical FDTD.

Furthermore, extension to the lossy case is not difficult. When an attenuation factor appears in the Polder vector [9], the equivalent-circuit scheme can be supplemented with a conductance $G_1 = 2\alpha\omega_0 C_1$ (where α is a damping constant) in parallel with L_1 and C_1, and a resistance $R = \alpha\omega_0 L_1$ in series with L. This modifies the updating equations for V_{Cx} and I_x in a manner typical to FDTD relations for lossy media.

Overall, for a cubic-cell FDTD lattice of cell size Δ and time-step Δt, the time-stepping expressions corresponding to calculations 1 and 2 for the three-dimensional lossy case are:

$$H_y\Big|_{i,j,k}^{n+1/2} = D_a H_y\Big|_{i,j,k}^{n-1/2} + D_b \left[\begin{array}{l} E_z\Big|_{i+1/2,j,k}^{n} - E_z\Big|_{i-1/2,j,k}^{n} + E_x\Big|_{i,j,k-1/2}^{n} - E_x\Big|_{i,j,k+1/2}^{n} \\ - p_y\Big|_{i,j,k}^{n} - 0.25\sqrt{\dfrac{L_1}{C_1}} \cdot \left(\begin{array}{l} h_x\Big|_{i-1,j-1/2,k}^{n-1/2} + h_x\Big|_{i-1,j+1/2,k}^{n-1/2} \\ + h_x\Big|_{i,j-1/2,k}^{n-1/2} + h_x\Big|_{i,j+1/2,k}^{n-1/2} \end{array} \right) \end{array} \right] \tag{9.78}$$

$$h_y\Big|_{i,j,k}^{n+1/2} = h_y\Big|_{i,j,k}^{n-1/2} + H_y\Big|_{i,j,k}^{n+1/2} - H_y\Big|_{i,j,k}^{n-1/2} + \frac{\Delta t}{L_1} p_y\Big|_{i,j,k}^{n} \tag{9.79}$$

$$p_y\Big|_{i,j,k}^{n+1} = \left(\frac{1 - G_1\Delta t/2C_1}{1 + G_1\Delta t/2C_1} \right) p_y\Big|_{i,j,k}^{n} + \left(\frac{\Delta t/C_1}{1 + G_1\Delta t/2C_1} \right) h_y\Big|_{i,j,k}^{n+1/2} \tag{9.80}$$

$$H_x\Big|_{i,j,k}^{n+1/2} = D_a H_x\Big|_{i,j,k}^{n-1/2} + D_b \left[\begin{array}{l} E_z\Big|_{i,j-1/2,k}^{n} - E_z\Big|_{i,j+1/2,k}^{n} + E_y\Big|_{i,j,k+1/2}^{n} - E_y\Big|_{i,j,k-1/2}^{n} \\ - p_x\Big|_{i,j,k}^{n} + 0.25\sqrt{\dfrac{L_1}{C_1}} \cdot \left(\begin{array}{l} h_y\Big|_{i-1/2,j-1,k}^{n+1/2} + h_y\Big|_{i+1/2,j-1,k}^{n+1/2} \\ + h_y\Big|_{i-1/2,j,k}^{n+1/2} + h_y\Big|_{i+1/2,j,k}^{n+1/2} \end{array} \right) \end{array} \right] \tag{9.81}$$

$$h_x\Big|_{i,j,k}^{n+1/2} = h_x\Big|_{i,j,k}^{n-1/2} + H_x\Big|_{i,j,k}^{n+1/2} - H_x\Big|_{i,j,k}^{n-1/2} + \frac{\Delta t}{L_1} p_x\Big|_{i,j,k}^{n} \tag{9.82}$$

$$p_x\Big|_{i,j,k}^{n+1} = \left(\frac{1 - G_1\Delta t/2C_1}{1 + G_1\Delta t/2C_1} \right) p_x\Big|_{i,j,k}^{n} + \left(\frac{\Delta t/C_1}{1 + G_1\Delta t/2C_1} \right) h_x\Big|_{i,j,k}^{n+1/2} \tag{9.83}$$

where L, L_1, C_1, p_x, p_y, h_x, and h_y are defined as before; $D_a = (2L - R\Delta t)/(2L + R\Delta t)$; and $D_b = 2\Delta t/(2L + R\Delta t)$. Analogous updates are implemented for calculations 4 and 5.

It is interesting to note that G_1 influences the modeling results only in the frequency range very close to ω_0. In most practical ferrite applications, the interesting frequencies are much higher than ω_0, and thus $G_1 = 0$ can be assumed without any significant error.

9.5.4 Illustrative Numerical Results

Fig. 9.4 illustrates the accuracy achievable using the magnetized ferrite model discussed above. In this example, a rectangular waveguide of width $w_1 = 22.86$ mm is assumed to be nonsymmetrically loaded with a ferrite slab of width $w_2 = w_1/3$. The ferrite slab is transversly magnetized in z-direction with $H_i = 200$ Oe (1.59×10^4 A/m), and is assumed to have the parameters $\varepsilon_r = 9$, $4\pi M_s = 2{,}000$G (0.2T), and $\alpha = 0.02$. In this example, the nonsymmetric location of the ferrite slab within the waveguide cross section is chosen to yield different phase constants β for the forward and backward waves.

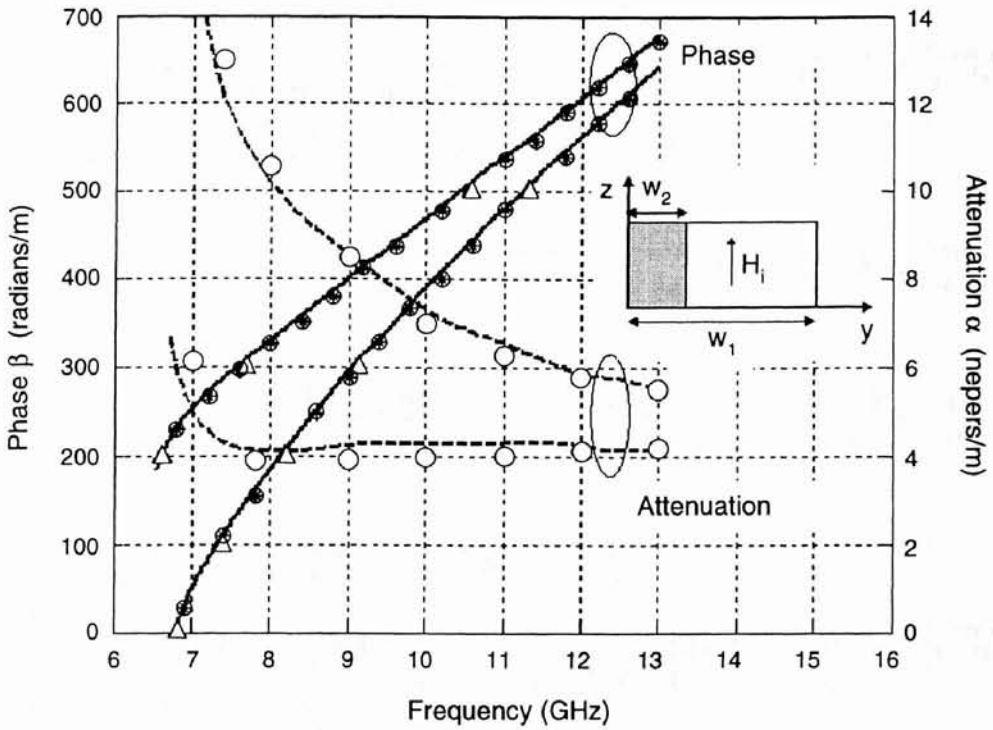


Fig. 9.4 Phase and attenuation constants of the TE_{10} mode of a rectangular waveguide nonsymmetrically loaded with a ferrite slab: continuous line – calculated using the method of this section; circled stars – simulations after [9]; triangles – analytical results after [8]; circles – analytical results after [9]. *Source*: Gwarek and Moryc, *IEEE Microwave and Wireless Components Lett.*, 2004, pp. 331–333, © 2004 IEEE.

From Fig. 9.4, we see that the FDTD modeling results obtained using the method of this section (obtained with a cell size $\Delta = 0.5$ mm) agree very well with the result of simulations presented in [9], and are within 1% of the analytical results quoted in [8]. Taking into account the fairly coarse $\lambda/15$ grid resolution at 13 GHz, we conclude that the level of modeling error is close to that introduced by the intrinsic FDTD grid dispersion.

9.5.5 Comparison of Computer Resources

Table 9.1 compares the computer resources needed for the method discussed in this section with three previously reported methods for FDTD modeling of magnetized ferrites [8–10]. This table lists the minimum number of variables needed to be stored for each FDTD cell, and the operation count (the minimum number of floating-point operations per cell per iteration). Assuming the use of uniform square or cubic meshing, this comparison has been conducted for both lossless and lossy ferrite cases in two- and three-dimensional problems. We note that, when taking into account the complex-number operations needed in the method of [10], the multiplication of two complex numbers was assumed to be equivalent to six real floating-point operations. Assumed also in the method of [10] was the use of a simple space averaging of the contribution of the off-diagonal terms, similar to that employed in the method of this section in (9.75).

TABLE 9.1

Comparison of Various FDTD Methods for Modeling Magnetized Ferrites
Adapted from: Gwarek and Moryc, *IEEE Microwave and Wireless Components Lett.*, 2004, pp. 331–333, © 2004 IEEE.

Method	*Required Number of Stored Variables*		*Operation Count*			
			2-D Solver		*3-D Solver*	
	2-D Solver	*3-D Solver*	$\alpha > 0$	$\alpha = 0$	$\alpha > 0$	$\alpha = 0$
This section [11]	7	10	39	35	58	54
Reference [9]	7	10	41		61	
Reference [8]	7	10		37		56
Reference [10]	11	14	77		86	

From Table 9.1, we see that the method discussed in this section compares favorably with previous methods in terms of computer effort, being slightly more conserving than those of [8] and [9], and considerably more conserving than that of [10].

However, the method discussed in this section does have one significant computational advantage relative to those of [8] and [9]: excellent numerical stability. Specifically, it was stated in [8] that a very fine grid discretization of 30 to 100 cells per wavelength is needed to achieve a stable solution in practical cases. In contrast, no stability problems have been observed when using the method of this section, even for a grid resolution as coarse as $\lambda/15$ (realized in Fig. 9.4 at the upper end of the excitation spectrum).

Moreover, in what is an important advantage in deepening our theoretical understanding, formal numerical stability and dispersion analyses of the FDTD ferrite model of this section can be conducted by separating the resonant and anisotropic effects. In [8] and [9], the resonant and anisotropic effects were inherently bound, making their numerical stability analysis very difficult. Indeed, no formal stability analyses of these algorithms have been reported to date. Such analyses are currently in progress for the algorithm of this section.

9.6 AUXILIARY DIFFERENTIAL EQUATION METHOD, NONLINEAR DISPERSIVE MATERIAL CASE

The behavior of electromagnetic fields in nonlinear dielectrics is a central topic in nonlinear optics technology, currently of great interest in the areas of lasers, communications, and high-speed digital switching. In principle, the electrodynamics of nonlinear media can be determined by solving Maxwell's equations, subject to the assumption that the electric polarization has a nonlinear relation to the electric field magnitude. However, until recently, the resulting nonlinear equations have not been attacked directly. Rather, approximations have been made that result in the class of *generalized nonlinear Schrödinger equations* (GNLSE) [13, 14]. The least approximate methods for GNLSE solve nonlinear scalar equations for the envelope of the propagating optical pulses. For example, the split-step Fourier method [14] is often used to simulate the propagation of optical pulses in low-loss fibers over very long optical distances, and the propagating beam method [15] has been used to model directional couplers.

References [16–21] reported auxiliary differential equation FDTD techniques that permit direct time integration of the full-vector Maxwell's equations for the important case of materials characterized by both a linear dispersion and a third-order Kerr and Raman dispersive nonlinearity. The optical carrier is retained in these approaches, which are robust and allow modeling ultrashort optical pulse interactions with complex two- and three-dimensional micron-scale structures. Recently, a simplified and computationally more efficient ADE-FDTD technique with the same capabilities was reported by Fujii et al. [22]. This section reviews the basic strategy of ADE-FDTD techniques for this problem, and the improved algorithm presented by Fujii et al. in [22].

9.6.1 Strategy

The overall strategy employed here is to apply normal Yee leapfrogging to time-step the electric flux density $\boldsymbol{D}$ via

$$\frac{\partial \boldsymbol{D}}{\partial t} = \nabla \times \boldsymbol{H} \tag{9.84}$$

and then calculate the electric field $\boldsymbol{E}$ (for subsequent application in the normal Yee time-stepping of $\boldsymbol{H}$) by applying the constitutive relation

$$\boldsymbol{D} = \varepsilon_0 \varepsilon_\infty \boldsymbol{E} + \boldsymbol{P}^{\mathrm{L}} + \boldsymbol{P}^{\mathrm{NL}} \tag{9.85}$$

Here, we separate the dielectric polarization into two parts: a linear part P^{L} and a nonlinear part P^{NL} [14]. Independent (uncoupled) time-domain auxiliary differential equations are developed to time-step P^{L} and P^{NL}, thereby avoiding the computationally inefficient simultaneous equations required in the formulation of previous ADE-FDTD approaches for this problem.

In the discussion to follow, we shall assume that the linear polarization P^{L} has two contributing additive terms: P_{Debye}, due to a single-pole Debye relaxation; and P_{Lorentz}, due to a single-pole-pair Lorentzian dispersion. Furthermore, we shall assume that the nonlinear polarization P^{NL} has two contributing additive third-order terms: P_{Kerr}, due to an instantaneous Kerr nonlinearity; and P_{Raman}, due to a dispersive Raman nonlinearity.

9.6.2 Contribution of the Linear Debye Polarization

As stated in Section 9.2.1, Debye media are characterized by a complex-valued, frequency-domain susceptibility function $\chi(\omega)$ that has one or more real poles at separate frequencies. For convenience, we repeat (9.1), which represents a single-pole Debye medium:

$$\chi_p(\omega) = \frac{\varepsilon_{s,p} - \varepsilon_{\infty,p}}{1 + j\omega\tau_p} \equiv \frac{\Delta\varepsilon_p}{1 + j\omega\tau_p}$$

where $\varepsilon_{s,p}$ is the static or zero-frequency relative permittivity, $\varepsilon_{\infty,p}$ is the relative permittivity at infinite frequency, $\Delta\varepsilon_p$ is the change in relative permittivity due to the Debye pole, and τ_p is the pole relaxation time. The electric polarization is then given by

$$\breve{P}_{\text{Debye}} = \frac{\varepsilon_0 \Delta\varepsilon_p}{1 + j\omega\tau_p} \breve{E} \tag{9.86}$$

For simplicity, let us assume that the electric field has a single vector component, E. After multiplying (9.86) through by $(1 + j\omega\tau_p)$ and transforming to the time domain, we obtain the following auxiliary differential equation:

$$P_{\text{Debye}} + \tau_p \frac{\partial P_{\text{Debye}}}{\partial t} = \varepsilon_0 \Delta\varepsilon_p E \tag{9.87}$$

Finite-differencing (9.87) centered at time-step $n + 1/2$ yields the following update for P_{Debye}:

$$P_{\text{Debye}}^{n+1} = \underbrace{\left(\frac{2\tau_p - \Delta t}{2\tau_p + \Delta t}\right)}_{a_{\text{Debye}}} P_{\text{Debye}}^{n} + \underbrace{\left(\frac{\varepsilon_0 \Delta\varepsilon_p \Delta t}{2\tau_p + \Delta t}\right)}_{b_{\text{Debye}}} \left(E^{n+1} + E^{n}\right) \tag{9.88}$$

where Δt is the time-step.

9.6.3 Contribution of the Linear Lorentz Polarization

As stated in Section 9.2.2, Lorentz media are characterized by a complex-valued, frequency-domain susceptibility function $\chi(\omega)$ that has one or more pairs of complex-conjugate poles. For convenience, we repeat (9.4), which represents a single-pole-pair Lorentz medium:

$$\chi_p(\omega) = \frac{\Delta\varepsilon_p \omega_p^2}{\omega_p^2 + 2j\omega\delta_p - \omega^2}$$

where $\Delta\varepsilon_p = \varepsilon_{s,p} - \varepsilon_{\infty,p}$ is the change in relative permittivity due to the Lorentz pole pair, ω_p is the frequency of the pole pair (the undamped resonant frequency of the medium), and δ_p is the damping coefficient. The electric polarization is then given by

$$\breve{P}_{\text{Lorentz}} = \frac{\varepsilon_0 \Delta\varepsilon_p \omega_p^2}{\omega_p^2 + 2j\omega\delta_p - \omega^2} \breve{E} \tag{9.89}$$

After multiplying (9.89) through by $(\omega_p^2 + 2j\omega\delta_p - \omega^2)$ and transforming to the time domain, we obtain the following auxiliary differential equation:

$$\omega_p^2 P_{\text{Lorentz}} + 2\delta_p \frac{\partial P_{\text{Lorentz}}}{\partial t} + \frac{\partial^2 P_{\text{Lorentz}}}{\partial t^2} = \varepsilon_0 \Delta\varepsilon_p \omega_p^2 E \tag{9.90}$$

Finite-differencing (9.90) centered at time-step n yields the following update for P_{Lorentz}:

$$P_{\text{Lorentz}}^{n+1} = \left[\frac{2 - \omega_p^2(\Delta t)^2}{\delta_p \Delta t + 1}\right] P_{\text{Lorentz}}^{n} + \left(\frac{\delta_p \Delta t - 1}{\delta_p \Delta t + 1}\right) P_{\text{Lorentz}}^{n-1} + \left[\frac{\varepsilon_0 \Delta\varepsilon_p \omega_p^2 (\Delta t)^2}{\delta_p \Delta t + 1}\right] E^n \tag{9.91}$$

9.6.4 Contributions of the Third-Order Nonlinear Polarization

In general, the third-order nonlinear polarization is given by the time convolution between the third-order susceptibility function $\chi^{(3)}(t_1, t_2, t_3)$ and the electric field for the different time scales represented by the arguments. For a relatively simple model of the electron response, it reduces to the Born-Oppenheimer approximation [23]:

$$P^{\text{NL}}(t) = \varepsilon_0 \chi_0^{(3)} E(t) \int_{-\infty}^{\infty} g(t - t') E^2(t') dt' \tag{9.92a}$$

where $\chi_0^{(3)}$ is the strength of the third-order nonlinearity. The causal response function $g(t)$ is normalized so that

$$\int_{-\infty}^{\infty} g(t) dt = 1 \tag{9.92b}$$

Equation (9.92) accounts only for nonresonant third-order processes, including phonon interactions and nonresonant electronic effects. To model these responses, we let [24]

$$g(t) = \alpha \delta(t) + (1 - \alpha) g_{\text{Raman}}(t) \tag{9.93a}$$

where $\delta(t)$ is a Dirac delta function that models Kerr nonresonant virtual electronic transitions on the order of 1 fs or less, and $g_{\text{Raman}}(t)$ is given by

$$g_{\text{Raman}}(t) = \left(\frac{\tau_1^2 + \tau_2^2}{\tau_1 \tau_2^2}\right) e^{-t/\tau_2} \sin(t/\tau_1)\, U(t) \tag{9.93b}$$

that models transient Raman scattering. Effectively, $g_{\text{Raman}}(t)$ models a single Lorentzian line centered on the optical phonon frequency $1/\tau_1$ and having a bandwidth of $1/\tau_2$, the reciprocal phonon lifetime. Note that α is a real-valued constant in the range $0 \le \alpha \le 1$ which parameterizes the relative strengths of the Kerr and Raman interactions.

Instantaneous Kerr Nonlinear Polarization

The polarization contributed by the instantaneous Kerr nonlinearity is given by [23]

$$P_{\text{Kerr}}(t) = \varepsilon_0 \chi_0^{(3)} E(t) \int_{-\infty}^{\infty} \alpha \delta(t-t') E^2(t') dt' = \alpha \varepsilon_0 \chi_0^{(3)} E^3(t) \tag{9.94}$$

The corresponding update equation is simply

$$P_{\text{Kerr}}^{n+1} = \alpha \varepsilon_0 \chi_0^{(3)} \left(E^{n+1}\right)^3 \tag{9.95}$$

Raman Nonlinear Polarization

From (9.92a), the polarization contributed by the Raman effect can be expressed as the convolution

$$P_{\text{Raman}}(t) = \varepsilon_0 E(t) \left[\chi_{\text{Raman}}^{(3)}(t) * E^2(t)\right] \tag{9.96a}$$

where

$$\chi_{\text{Raman}}^{(3)}(t) = (1-\alpha) \chi_0^{(3)} g_{\text{Raman}}(t) \tag{9.96b}$$

Following [22], we now introduce an auxiliary variable for the convolution

$$S(t) = \chi_{\text{Raman}}^{(3)}(t) * E^2(t) \tag{9.97}$$

The Fourier transform of (9.97) leads to

$$S(\omega) = \chi_{\text{Raman}}^{(3)}(\omega) \cdot \mathcal{F}\left[E^2(t)\right] \tag{9.98}$$

where $\mathcal{F}$ denotes the Fourier transform, and the frequency-domain response function is given by

$$\chi_{\text{Raman}}^{(3)}(\omega) = \frac{(1-\alpha) \chi_0^{(3)} \omega_{\text{Raman}}^2}{\omega_{\text{Raman}}^2 + 2j\omega\delta_{\text{Raman}} - \omega^2} \tag{9.99}$$

where

$$\omega_{\text{Raman}} = \sqrt{\frac{\tau_1^2 + \tau_2^2}{\tau_1^2 \tau_2^2}} \quad ; \qquad \delta_{\text{Raman}} = 1/\tau_2 \tag{9.100a, b}$$

We next rewrite S as

$$S(\omega) = \frac{(1-\alpha)\chi_0^{(3)}\omega_{\text{Raman}}^2}{\omega_{\text{Raman}}^2 + 2j\omega\delta_{\text{Raman}} - \omega^2}\,\mathcal{F}\left[E^2(t)\right] \tag{9.101}$$

and multiply (9.101) through by $(\omega_{\text{Raman}}^2 + 2j\omega\delta_{\text{Raman}} - \omega^2)$. Inverse Fourier transformation to the time domain then yields the following auxiliary differential equation:

$$\omega_{\text{Raman}}^2 S + 2\delta_{\text{Raman}}\frac{\partial S}{\partial t} + \frac{\partial^2 S}{\partial t^2} = (1-\alpha)\chi_0^{(3)}\omega_{\text{Raman}}^2 E^2 \tag{9.102}$$

Finite-differencing (9.102) centered at time-step n yields the following update for S:

$$S^{n+1} = \left[\frac{2-\omega_{\text{Raman}}^2(\Delta t)^2}{\delta_{\text{Raman}}\Delta t + 1}\right]S^n + \left(\frac{\delta_{\text{Raman}}\Delta t - 1}{\delta_{\text{Raman}}\Delta t + 1}\right)S^{n-1} + \left[\frac{(1-\alpha)\chi_0^{(3)}\omega_{\text{Raman}}^2(\Delta t)^2}{\delta_{\text{Raman}}\Delta t + 1}\right]\left(E^n\right)^2 \tag{9.103}$$

Since from (9.96a) and (9.97) we have

$$P_{\text{Raman}}(t) = \varepsilon_0 E(t) S(t) \tag{9.104}$$

the update equation for the Raman polarization is simply

$$P_{\text{Raman}}^{n+1} = \varepsilon_0 E^{n+1} S^{n+1} \tag{9.105}$$

where S^{n+1} is given by (9.103).

9.6.5 Electric Field Update

The time-stepping expressions for the linear and nonlinear polarization contributions obtained in (9.88), (9.91), (9.95), and (9.105) can now be substituted into (9.85), given in expanded form by

$$D^{n+1} = \varepsilon_0\varepsilon_\infty E^{n+1} + P_{\text{Debye}}^{n+1} + P_{\text{Lorentz}}^{n+1} + P_{\text{Kerr}}^{n+1} + P_{\text{Raman}}^{n+1} \tag{9.106}$$

Isolating the E^{n+1} term, we obtain

$$E^{n+1} = \left(D^{n+1} - P_{\text{Debye}}^{n+1} - P_{\text{Lorentz}}^{n+1} - P_{\text{Kerr}}^{n+1} - P_{\text{Raman}}^{n+1}\right)\Big/\varepsilon_0\varepsilon_\infty \tag{9.107}$$

We note that the Debye, Kerr, and Raman polarization terms on the right-hand side of (9.107) include E^{n+1}, and furthermore, the Kerr polarization term incorporates $(E^{n+1})^3$. Therefore, (9.107) is nonlinear in E^{n+1}. Following the approach reported in [17], a Newton iteration procedure can be applied to (9.107) to determine E^{n+1}. This procedure is given by

$$E^{<m+1>} = \frac{D^{n+1} - a_{\text{Debye}} P^{n}_{\text{Debye}} - b_{\text{Debye}} E^{n} - P^{n+1}_{\text{Lorentz}}}{\varepsilon_0 \varepsilon_\infty + b_{\text{Debye}} + \alpha \varepsilon_0 \chi^{(3)}_0 \left(E^{<m>}\right)^2 + \varepsilon_0 S^{n+1}} \tag{9.108}$$

for $m = 0, 1, 2, \ldots$ Here, a_{Debye} and b_{Debye} are defined in (9.88), and $E^{<m>}$ denotes the approximation of E^{n+1} at the mth iteration of the Newton procedure. For convenience, the initial estimate of E^{n+1} to begin the iterative process is simply E^n; that is, $E^{<0>} = E^n$.

9.6.6 Illustrative Numerical Results for Temporal Solitons

This section presents selected published results illustrating the dynamics of temporal solitons, calculated using ADE-FDTD nonlinear dispersion algorithms. The materials investigated are assumed to exhibit a single linear Lorentz polarization and the third-order Kerr and Raman nonlinearities.

References [16, 17] reported modeling a 50-fs pulsed optical signal source switched on at $t = 0$ at the surface $x = 0$ of a material half-space. The material is assumed to be characterized by the following parameters:

Linear Lorentz dispersion: $\varepsilon_s = 5.25$; $\varepsilon_\infty = 2.25$; $\omega_p = 4 \times 10^{14}$ sec^{-1} ; $\delta_p = 2 \times 10^{9}$ sec^{-1}

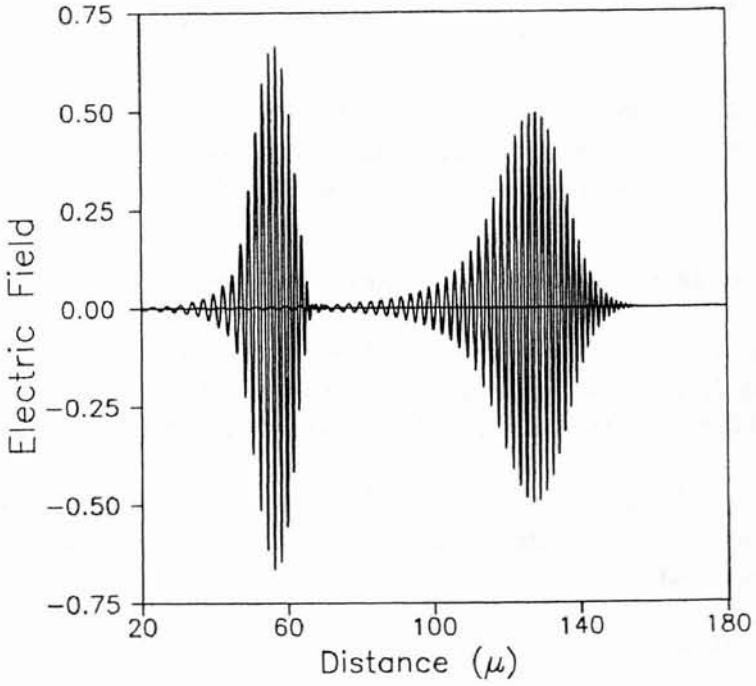
Nonlinear response: $\chi^{(3)}_0 = 7 \times 10^{-2}$ (V/m)$^{-2}$; $\alpha = 0.7$; $\tau_1 = 12.2$ fs ; $\tau_2 = 32$ fs

These parameters represent a nonphysical material which has a group velocity dispersion and nonlinearity selected to demonstrate temporal soliton formation over short propagation spans of less than 200 μm. The impinging pulse is assumed to have unity amplitude of its E-field, the carrier frequency $f_c = 1.37 \times 10^{14}$ Hz ($\lambda_0 = 2.19$ μm), and a hyperbolic secant envelope function with a characteristic time constant of 14.6 fs. Approximately seven cycles of the optical carrier are contained within the pulse envelope, and the center of the pulse coincides with a zero crossing of the sinusoid. The grid resolution is set at 5 nm ($\equiv \lambda_0/400$) to limit the numerical phase velocity error to approximately one part in 10^5, very small compared to the physical dispersions being modeled.

Fig. 9.5, taken from [16, 17], depicts the results of the dispersive and nonlinear dispersive ADE-FDTD computations. In Fig. 9.5(a), the computed rightward propagating pulse for the linear Lorentz dispersive case ($\chi_0^{(3)}$ temporarily set to zero) is graphed at $n = 20{,}000$ and 40,000 time-steps. This corresponds to pulse propagation to depths of $x = 55$ μm and 126 μm at times of 487 fs and 973 fs, respectively, after initiation. Here, the assumed linear dispersion causes the computed pulse to broaden and have diminished amplitude during its propagation. In addition, the pulse exhibits a carrier frequency modulation that is greater than f_c on the leading side of the pulse, and less than f_c on the trailing side of the pulse.

Fig. 9.5(b) graphs the optical pulse propagation when the assumed dispersive nonlinearity is actuated. Upon the precise choice of $\chi_0^{(3)}$ and the initial pulse amplitude, this yields a rightward propagating *temporal soliton* that retains its amplitude and width. Here the pulse-width spreading effect caused by the assumed linear dispersion is exactly balanced by the pulse-width sharpening effect caused by the assumed nonlinearity. Fig. 9.5(b) also shows the formation of a small precursor (daughter) pulse, which has spectral content at approximately three times the frequency of the soliton. We note that all of these phenomena have been similarly calculated using the efficient ADE-FDTD algorithm introduced in [22] and reviewed in this section.

(a)



(b)

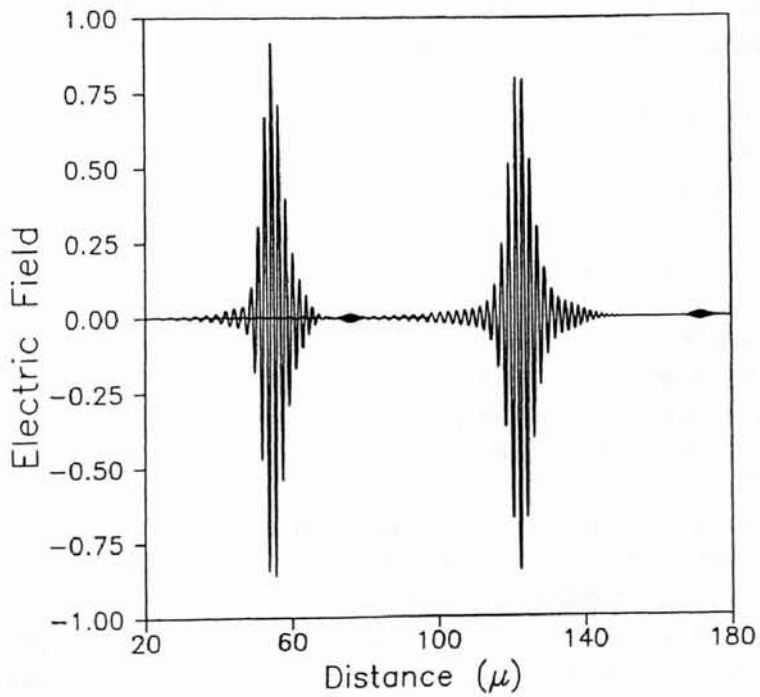


Fig. 9.5 ADE-FDTD study of temporal soliton formation in a nonphysical, nonlinear dispersive medium: (a) calculated optical carrier pulse after propagating 55 μm and 126 μm in the linear Lorentz dispersive medium; (b) corresponding results for the nonlinear dispersive case, showing the formation of a soliton and a small precursor pulse. *Source:* Goorjian and Taflove, *Optics Letters*, 1992, pp. 180–182.

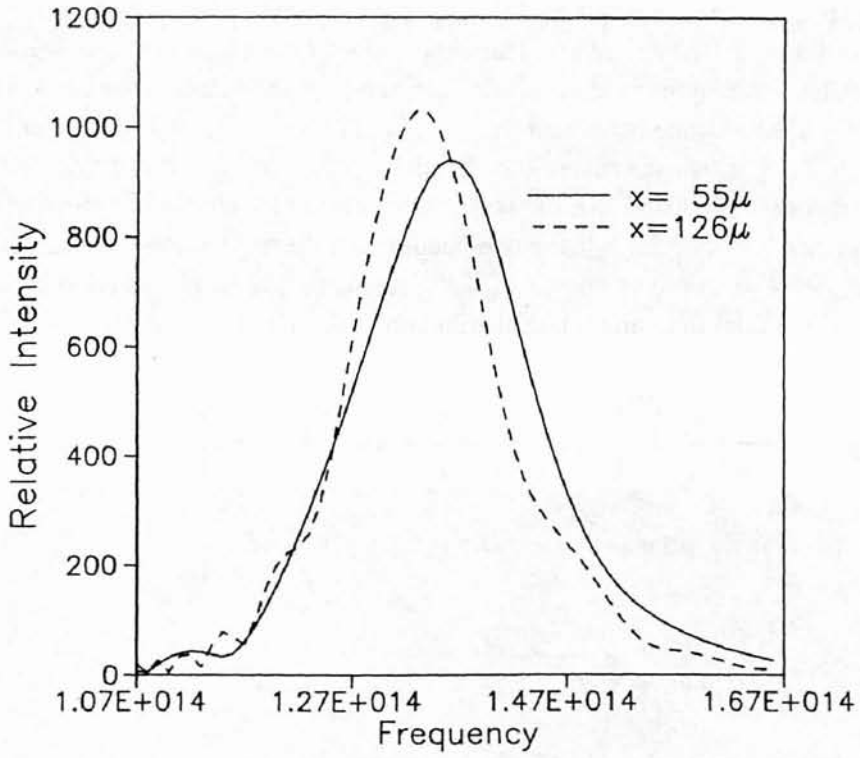


Fig. 9.6 Redshift of the Fourier spectrum of the propagating soliton of Fig. 9.5(b). *Source:* Goorjian and Taflove, *Optics Letters*, 1992, pp. 180–182.

Fig. 9.6 graphs the Fourier spectrum of the temporal soliton of Fig. 9.5(b) at $n = 20{,}000$ and 40,000 time-steps. This figure indicates a 4-THz redshift and a sharpening of the soliton spectrum as the computed pulse propagates. These phenomena have also been predicted from the GNLSE theory as being caused by the Lorentz nonlinear dispersion.

9.6.7 Illustrative Numerical Results for Spatial Solitons

Reference [19] reported FDTD modeling of *spatial optical soliton* propagation and mutual deflection in a homogeneous, physical glass medium that has an instantaneous Kerr nonlinearity. Here, the models calculate two-dimensional TM_z sinusoidal optical beam propagation with the field components E_z, H_x, and H_y. Spatial soliton formation is observed when the transverse spreading of the beam caused by linear diffraction is balanced by the transverse self-focusing of the beam caused by the nonlinearity. The propagation medium is assumed to be Type-RN Corning glass with the linear and nonlinear refractive indexes $n_0 = 2.46$ and $n_2 = 1.25 \times 10^{-18}$ m^2/W, respectively. The beam carrier frequency is 2.31×10^{14} Hz ($\lambda_0 = 1.3$ μm), the initial peak E_z field is 6.87×10^9 V/m, and the initial transverse distribution of E_z is a hyperbolic secant function with a spatial width (full-width at half-maximum – FWHM) of 0.65 μm.

Fig. 9.7(a) illustrates the results of the FDTD simulation of parallel, copropagating, equal-amplitude spatial solitons separated by 1.05 μm center-to-center. Here, the beams are assumed to have a carrier phase difference of π. This model shows the beam-to-beam *repulsion* expected from previous NLSE analyses, with increasing separation of the beams as they propagate.

Fig. 9.7(b) illustrates the results for the same conditions as Fig. 9.7(a), except that the beams are assumed to have in-phase carriers (i.e., a carrier phase difference of zero). Previous NLSE analyses for this case predicted that the two beams interact by alternately attracting, coalescing, repelling, and then recoalescing. If the two beams have the appropriate amplitudes and spacing, the attraction and repulsion is periodic. In fact, [25] states that two in-phase fundamental spatial solitons having an initial field-amplitude distribution in the transverse direction of

$$A(y) = \frac{1}{k\,w}\left(\frac{n_0}{n_2}\right)^{1/2}\left[\operatorname{sech}\left(\frac{y-y_0}{w}\right) + \operatorname{sech}\left(\frac{y+y_0}{w}\right)\right] \tag{9.109}$$

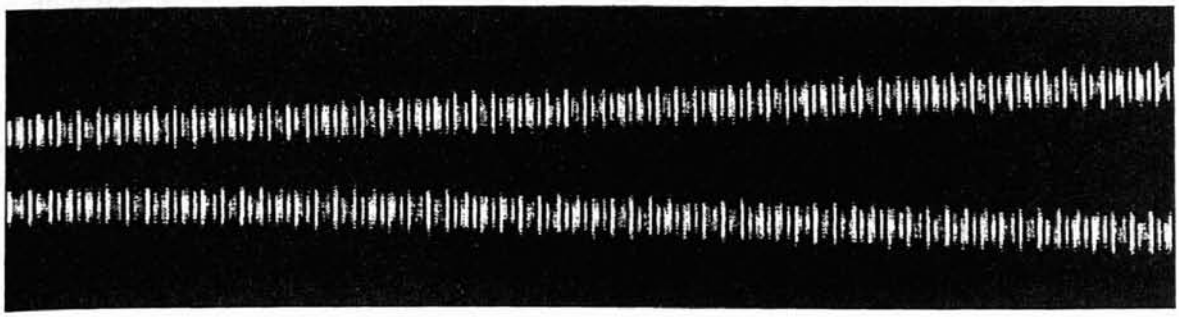
repeatedly coalesce in the propagation direction with a period of

$$x_p = \frac{2\,x_0 \sinh\left(2y_0/w\right)\cosh\left(y_0/w\right)}{2y_0/w + \sinh\left(2y_0/w\right)} \tag{9.110}$$

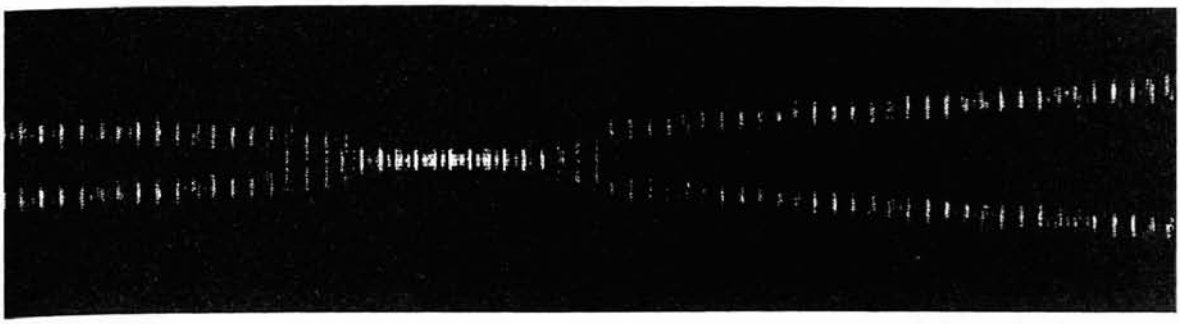
based on the NLSE theory of [26]. Here, w is the characteristic width of the hyperbolic secant function, $y_0 = 1.42w$, $2y_0$ is the center-to-center separation of the two beams, and $x_0 = \pi^2 n_0 w^2/\lambda$ is the soliton period. For the choice of parameters used in the FDTD simulations, the predicted repetition period is $x_p = 9$ μm. However, as shown in Fig. 9.7(b), the FDTD calculations show only a single beam coalescence and then subsequent beam divergence to arbitrarily large separations, for an effective $x_p = \infty$.

It is important to understand why the nonlinear FDTD Maxwell's equations model does not agree with the NLSE prediction in this case. The first possibility is a flawed FDTD simulation, due to either inadequate grid resolution or inadequate isolation of the beam-interaction region from the outer grid boundaries (where the fairly reflective second-order Mur ABC was used at the time). However, tests with progressively finer space-time resolution and larger grid size yielded no changes. Therefore, it was concluded that the original FDTD model was numerically converged and sufficiently free of the outer boundary artifact to yield plausible results.

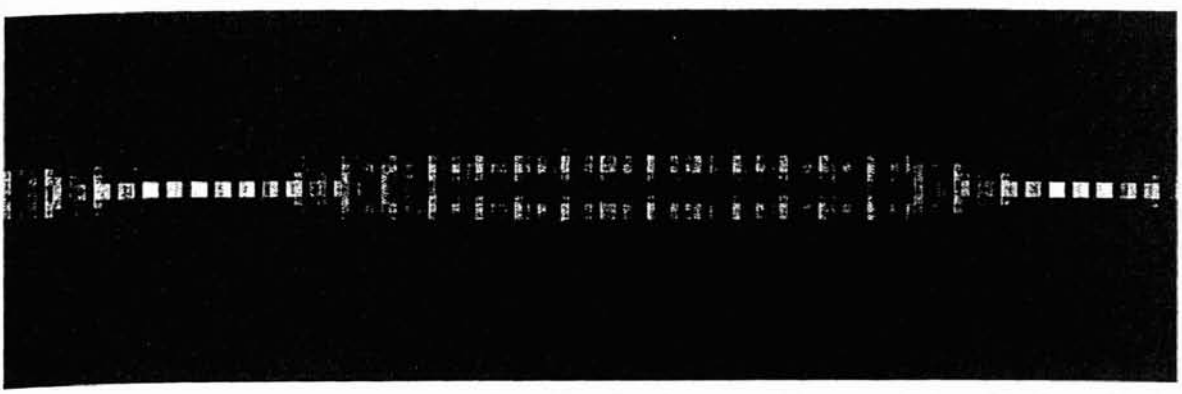
The second possibility is a flawed NLSE simulation, because the optical beamwidth is too narrow, and diffraction effects are not properly accounted. Since it is known that additional terms in the NLSE are required to model higher-order effects for temporal solitons, it is reasonable that NLSE modeling of copropagating spatial solitons becomes more physically meaningful if the two beams are widened relative to the optical wavelength, while maintaining the same ratio of beamwidth-to-beam separation. This would reduce linear beam diffraction effects, hopefully bringing the test case into the region of validity for the simple NLSE model. To test this possibility, [19] reported additional FDTD simulations with a progressively increased intensity beamwidth B_I and beam-to-beam separation, keeping the dielectric wavelength λ_d constant. The results of one such simulation are shown in Fig. 9.7(c).



(a)



(b)



(c)

Fig. 9.7 FDTD computations of the E_z field of optically narrow, mutually interacting spatial solitons in Type-RN Corning glass: (a) repulsion for relative carrier phase = π; (b) single coalescence and subsequent divergence for relative carrier phase = 0; (c) periodic beam recoalescence after doubling the intensity beamwidth and separation parameters of the simulated beams, but keeping the wavelength constant. *Source:* Joseph and Taflove, *IEEE Photonics Technology Letters*, 1994, pp. 1251–1254, © 1994 IEEE.

After an initial doubling of the beamwidth and beam separation, the FDTD-predicted spatial solitons begin to qualitatively show the recoalescence behavior predicted by NLSE, but with a 38% longer recoalescence period x_p than the NLSE value. After a second doubling, the FDTD and NLSE predictions for x_p show better agreement, differing by 13%. Finally, a third increase of the beamwidth and beam separation dropped the difference between the FDTD and NLSE predictions for x_p to only 5%. These results are shown in Table 9.2.

TABLE 9.2

Progressive Agreement of FDTD and NLSE Results for the Recoalescence Period x_p of Copropagating Spatial Solitons as the Ratio of the Beamwidth to Wavelength Increases

B_I, *FWHM* (μm)	B_I/λ_d	x_p (μm) *NLSE*	x_p (μm) *FDTD*	*Difference*
0.65	1.22	9	∞	∞%
1.3	2.46	34	47	38%
2.6	4.9	135	153	13%
3.9	7.36	305	320	5%

We conclude that there is a strong likelihood that copropagating, optically narrow spatial solitons have only a single coalescence, followed by indefinite separation. The FDTD model appears to properly predict the behavior of such solitons, both in the regime where the standard NLSE model breaks down ($B_I/\lambda_d < 1$), and the regime where the standard NLSE model is valid ($B_I/\lambda_d >> 1$). The paraxial approximation inherent to NLSE, according to [27], accounts only for zeroth-order linear diffraction effects. Since the FDTD model implements the fundamental Maxwell's curl equations, it makes no assumption about a preferred propagation direction. It naturally accounts for energy transport in arbitrary transverse directions, and should be exact for the computed optical electromagnetic fields up to the limit set by the grid resolution.

The single-time spatial soliton coalescence behavior indicated by the FDTD modeling studies discussed above provides the basis for an all-optical "AND" gate operating at time scales in the order of 100 fs. This gate would consist of a Kerr-type nonlinear interaction region with a pair of input and output waveguides on each side. Optical signal and control pulses would be fed in at the left edge, interact in the nonlinear medium, and then couple into receptor waveguides. In the absence of the control beam, the signal beam would propagate with zero deflection. In the presence of the control beam, and depending upon its carrier phase relative to the signal pulse, there would be either a single coalescence and then deflection to a collecting waveguide, or deflection without coalescence.

9.7 AUXILIARY DIFFERENTIAL EQUATION METHOD, MACROSCOPIC MODELING OF SATURABLE, DISPERSIVE OPTICAL GAIN MATERIALS

This section reviews an ADE-FDTD formulation [28] for modeling at the macroscopic scale materials that have saturable, frequency-dependent gain at optical wavelengths. By macroscopic, we mean that conventional dielectric parameters such as the permittivity ε and the conductivity σ are adjusted to yield the desired gain. No attempt is made here to model at the atomic level the rates of electron transitions and photon interactions that underlie this gain. However, Section 9.8 provides a detailed discussion of FDTD modeling of just such an atomic-level gain model.

9.7.1 Theory[1]

The ADE-FDTD method described in this section models a saturable, frequency-dependent optical gain through the constitutive relation for the current density. The basic idea underlying this approach, using a negative conductance for gain, is an extension of a decades-old concept in microwave engineering and electronics. For example, the tunnel diode has a long history of use as a microwave negative-resistance amplifier.

We assume that the optical gain medium is homogeneously broadened, wherein the atoms of the gain medium are indistinguishable and have the same atomic transition frequency ω_0. The relaxation processes are included in a phenomenological manner through the time constant T_2. This dipole relaxation time accounts for the fact that any phase coherence introduced into the system of atoms by the action of the E-field is lost in a finite time interval once the field is turned off. Therefore, the small-signal linear gain is governed in the frequency domain by a single Lorentzian profile having a width determined by T_2. The gain also exhibits a large-signal saturation, which is due to the decrease of the population inversion with field intensity.

For simplicity, the algorithm is described in the context of a one-dimensional problem with field components E_z and H_y propagating along the x-direction through a nonmagnetic, isotropic medium. In this case, Maxwell's curl equations are

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu_0}\frac{\partial E_z}{\partial x} \quad ; \qquad J_z + \varepsilon_r \varepsilon_0 \frac{\partial E_z}{\partial t} = \frac{\partial H_y}{\partial x} \tag{9.111a, b}$$

The frequency-dependent conductivity that relates the E-field and the current density is given by

$$\sigma(\omega) = \frac{J_z(\omega)}{E_z(\omega)} = \left(\frac{1}{1 + I/I_s}\right)\cdot\left[\frac{\sigma_0/2}{1 + j(\omega - \omega_0)T_2} + \frac{\sigma_0/2}{1 + j(\omega + \omega_0)T_2}\right] \tag{9.112}$$

where Hermitian symmetry is assumed for the Lorentzian gain profile, σ_0 is related to the peak value of the gain set by the pumping level and the resulting population inversion, and I_s is the saturation intensity. Since $\sigma(\omega)$ accounts for any gain (or loss) in the medium, the susceptibility can be reduced to a frequency-independent real constant, such that $D_z = \varepsilon E_z$ where $\varepsilon = \varepsilon_r \varepsilon_0$.

---

[1] For a more detailed physical description of the physics of a homogeneously broadened gain medium, the reader is referred to a quantum electronics textbook (for example, see [29, 30]).

To see how this gives gain, consider the linear case where $I << I_s$. The right-hand side of (9.112) simplifies and is separated into real and imaginary parts:

$$\sigma(\omega) = \sigma_R(\omega) + j\sigma_I(\omega)$$

$$= \frac{\sigma_0\left[1+\left(\omega_0^2+\omega^2\right)T_2^2\right]}{\left[1+\left(\omega_0^2-\omega^2\right)T_2^2\right]^2+4\omega^2T_2^2} + \frac{j\sigma_0\,\omega\,T_2\left[-1+\left(\omega_0^2-\omega^2\right)T_2^2\right]}{\left[1+\left(\omega_0^2-\omega^2\right)T_2^2\right]^2+4\omega^2T_2^2} \tag{9.113}$$

Upon assuming a propagating wave of the form

$$E_z(x,t) = E_0\, e^{-\alpha x}\, e^{j(\omega t-\beta x)} \tag{9.114}$$

and that the material is low-gain (or low-loss) such that

$$\sigma_R(\omega) << \omega\varepsilon_0\,\varepsilon_{r,\text{eff}}(\omega) \tag{9.115a}$$

where the effective relative permittivity is

$$\varepsilon_{r,\text{eff}}(\omega) = \varepsilon_r - \sigma_I(\omega)/\omega\varepsilon_0 \tag{9.115b}$$

it can be shown that [28]

$$\alpha(\omega) \cong \frac{\sigma_R(\omega)}{2c\varepsilon_0 n_{\text{eff}}} \quad ; \qquad \beta(\omega) \cong n_{\text{eff}}\frac{\omega}{c} \tag{9.116a, b}$$

where $n_{\text{eff}} = \sqrt{\varepsilon_{r,\text{eff}}}$. From (9.116a), we conclude that if $\sigma_R(\omega)$ is negative, then α is negative, providing gain. According to (9.113), this is the case when σ_0 is chosen to be negative. Physically, this corresponds to a population inversion.

Taking the inverse Fourier transformation of (9.112) provides the following auxiliary differential equation that is solved simultaneously with (9.111b):

$$(1+\omega_0^2T_2^2)J_z + 2T_2\frac{\partial J_z}{\partial t} + T_2^2\frac{\partial^2 J_z}{\partial t^2} = s\sigma_0 E_z + s\sigma_0 T_2\frac{\partial E_z}{\partial t} \tag{9.117}$$

Here, $s \equiv (1+I/I_s)^{-1}$ is the saturation coefficient that contains feedback information concerning the latest peak E-field. The saturation intensity I_s is treated as a frequency-independent constant, and the spatially dependent intensity I is treated as a feedback parameter in time. An auxiliary variable F_z is defined in order to write (9.117) as two first-order equations:

$$F_z \equiv \frac{\partial J_z}{\partial t} \tag{9.118}$$

$$(1 + \omega_0^2 T_2^2) J_z + 2 T_2 F_z + T_2^2 \frac{\partial F_z}{\partial t} = s \sigma_0 E_z + s \sigma_0 T_2 \frac{\partial E_z}{\partial t} \tag{9.119}$$

The coupled differential equations (9.111b), (9.118), and (9.119) are converted to finite-difference expressions using second-order central-differences at time-step $n + 1/2$. Solving the resulting system for $F_z|_i^{n+1}$, $J_z|_i^{n+1}$, and $E_z|_i^{n+1}$ yields the following explicit update equations:

$$F_z|_i^{n+1} = A_1 \left(H_y|_{i+1/2}^{n+1/2} - H_y|_{i-1/2}^{n+1/2} \right) + A_2 E_z|_i^n + A_3 J_z|_i^n + A_4 F_z|_i^n \tag{9.120}$$

$$J_z|_i^{n+1} = J_z|_i^n + \frac{\Delta t}{2} \left(F_z|_i^{n+1} + F_z|_i^n \right) \tag{9.121}$$

$$E_z|_i^{n+1} = E_z|_i^n - \frac{\Delta t}{2 \varepsilon} \left(J_z|_i^{n+1} + J_z|_i^n \right) + \frac{\Delta t}{\varepsilon \Delta x} \left(H_y|_{i+1/2}^{n+1/2} - H_y|_{i-1/2}^{n+1/2} \right) \tag{9.122}$$

where

$$A_1 = \frac{4 \Delta t \, s(i) \sigma_0 (\Delta t + 2 T_2)}{B \Delta x} \; ; \qquad A_2 = \frac{8 \varepsilon s(i) \sigma_0 \Delta t}{B} \tag{9.123a, b}$$

$$A_3 = -\frac{4 \Delta t \left[2 \varepsilon (1 + \omega_0^2 T_2^2) + s(i) \sigma_0 (\Delta t + 2 T_2) \right]}{B} \tag{9.123c}$$

$$A_4 = -\frac{8 \varepsilon T_2 (\Delta t - T_2)}{B} - \frac{(\Delta t)^2 \left[2 \varepsilon (1 + \omega_0^2 T_2^2) + s(i) \sigma_0 (\Delta t + 2 T_2) \right]}{B} \tag{9.123d}$$

$$B = 8 \varepsilon T_2 (\Delta t + T_2) + (\Delta t)^2 \left[2 \varepsilon (1 + \omega_0^2 T_2^2) + s(i) \sigma_0 (\Delta t + 2 T_2) \right] \tag{9.123e}$$

$$s(i) = \left[1 + I(i)/I_s \right]^{-1} ; \qquad I(i) = 0.5 c n \varepsilon_0 \left(E_z|_i^{\text{peak}} \right)^2 \tag{9.123f}$$

For a linear medium, $s(i) = 1$, since the intensity is negligible compared to the saturation intensity. For a nonlinear medium, $s(i)$ is updated as follows. If $|E|$ at time-step n is greater than $|E|$ at $n - 1$ at the same location in space, then $s(i)$ is updated using the latest value of E. On the other hand, if $|E|$ at n has decreased from its previous value, then $s(i)$ is not updated, and hence remains based on the most recent peak E-field value. In this manner, intensity feedback in the time domain retains, as much as possible, its frequency-domain meaning. Note that since the feedback is performed independently at each grid location, this model represents a spatially inhomogeneously broadened medium in which spatial hole burning can occur.

The following update expression for H_y is obtained at time-step n in the standard manner from (9.111a):

$$H_y\Big|_{i+1/2}^{n+1/2} = H_y\Big|_{i+1/2}^{n-1/2} + \frac{\Delta t}{\mu_0 \Delta x}\left(E_z\Big|_{i+1}^{n} - E_z\Big|_{i}^{n}\right) \tag{9.124}$$

This completes the FDTD algorithm for a saturable, single-Lorentzian optical gain medium. We see that this algorithm is a four-step process given by (9.120) to (9.122) and (9.124), which retains the explicit nature and field storage of the original. It is easily extended to two- and three-dimensional problems with heterogeneous media. More complicated gain spectra can be approximated using a linear combination of Lorentzians.

9.7.2 Validation Studies

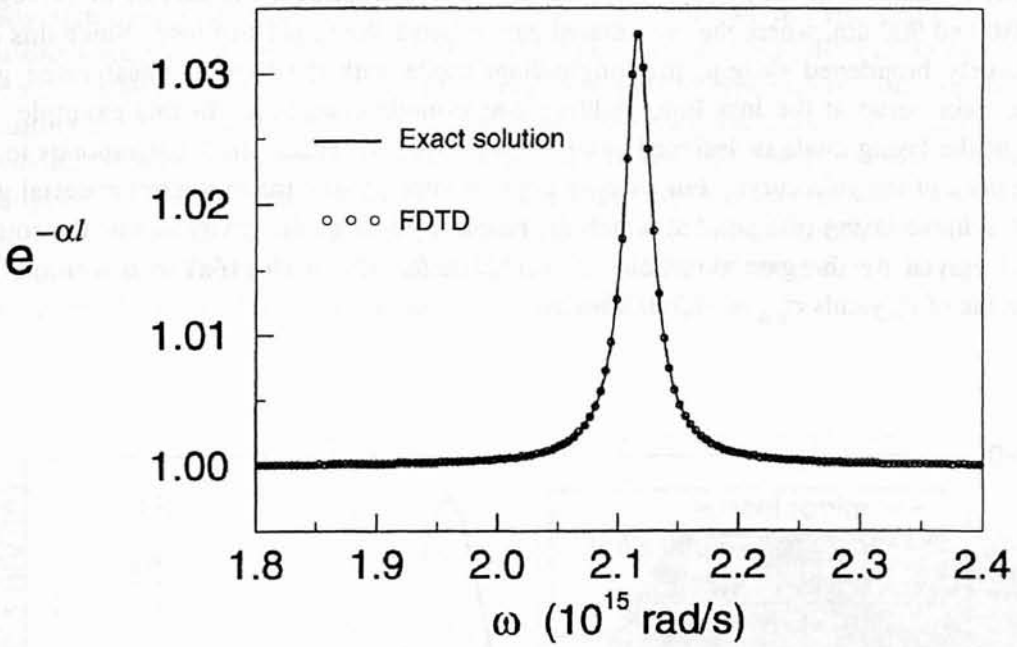
Linear Dispersive Optical Gain Medium

The accuracy of the dispersive-gain FDTD algorithm is first established for a linear gallium arsenide (GaAs) medium. In the example of [28], the complex-valued propagation factor $\gamma = \alpha + j\beta$ of a plane wave in a one-dimensional FDTD grid is computed for the grid filled with a homogeneous GaAs gain medium characterized by $\lambda_0 = 0.89$ μm, $T_2 = 0.07$ ps (gain spectrum bandwidth = 5 THz), and $n = 3.59$. A Gaussian carrier pulse of center frequency ω_0 and bandwidth greater than 5 THz excites the grid. The time waveforms of the FDTD-calculated E-fields are recorded at two observation points x_1 and x_2 separated by $l = x_2 - x_1 = \lambda_0/n$. By taking the ratio of the DFTs of these two waveforms, the propagation factor γ is calculated over the full bandwidth of the pulse.

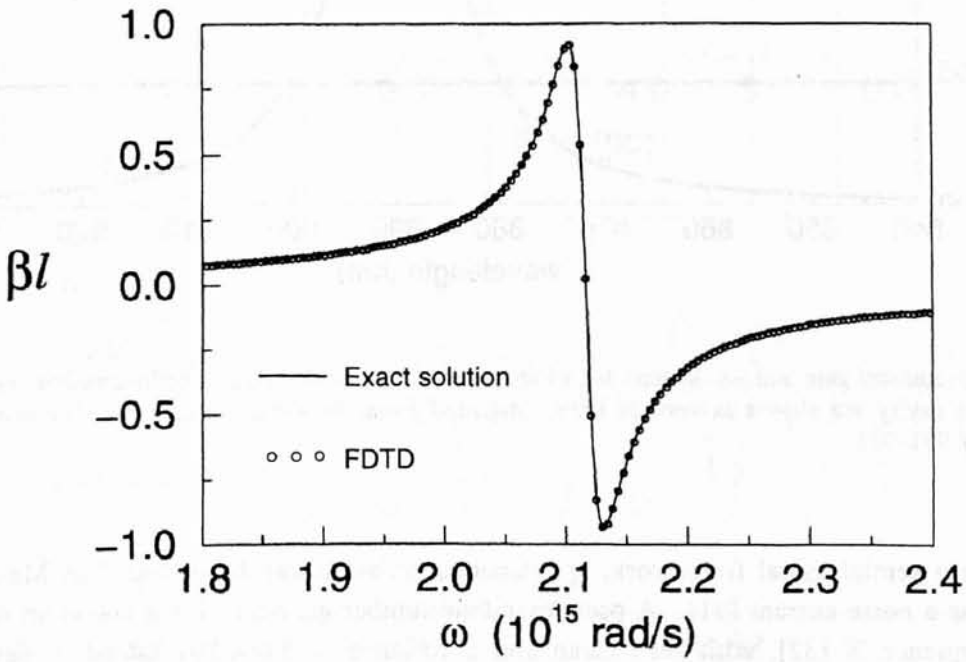
Fig. 9.8 compares the FDTD-computed amplification factor $e^{-\alpha l}$ and phase βl with the exact values for a very-fine grid resolution of $\Delta x = \lambda_0/400n = 0.62$ nm. (βl is corrected by $n\omega l/c$ to give the phase due solely to the presence of the gain medium.) Over the complete frequency range, the error is less than two parts per 10,000 in $e^{-\alpha l}$, and less than 0.006° in βl. This indicates the potential for high accuracy of the FDTD dispersive-gain model over an extremely large instantaneous bandwidth.

Saturable Dispersive Optical Gain Medium

Next, the accuracy of the dispersive-gain algorithm is established for a saturable dispersive gain medium. The model used for this study [28] consists of a one-dimensional Fabry-Perot laser cavity of length $l = 12.4$ μm filled entirely with a GaAs gain medium characterized by $\lambda_0 = 0.89$ μm, $T_2 = 0.07$ ps, $n = 3.59$, and $I_s = 65.2$ kW/cm^2. The mirrors at the ends of the laser cavity are formed by planar GaAs / air interfaces. Therefore, the reflectivity R of each end facet is independent of frequency. The mth resonance for this simple cavity is given by $\lambda_m = 2nl/m$.



(a) amplification factor.



(b) phase in degrees.

Fig. 9.8 Comparison of FDTD results and exact theory for the amplification factor and phase spectra of an optical pulse propagating a distance of one dielectric wavelength in a linear GaAs gain medium with $\sigma_0 = -5{,}000$ S/m. *Source:* Hagness et al., *Radio Science*, 1996, pp. 931–941.

Fig. 9.9 shows an above-threshold unsaturated gain curve for $\sigma_0 = -7{,}000$ S/m superimposed on the cavity resonances marked by the vertical lines. Three longitudinal modes lie in the region between 880 and 900 nm, where the unsaturated gain exceeds the round-trip loss. Since this is a homogeneously broadened system, the longitudinal mode with the highest unsaturated gain clamps the gain curve at the loss line, yielding single-mode operation. In this example, the frequency of the lasing mode is designed to be the $m = 100$ resonance, which corresponds to λ_0; that is, the peak of the gain curve. For a cavity with no internal loss, the minimum material gain required to achieve lasing (the point at which the round-trip loss of the cavity equals the round-trip gain) is given by the gain threshold $\alpha_{th} = (1/2l)\ln R$. Using (9.116a) to determine the threshold value of σ_0 yields $\sigma_{0,th} = -1{,}760$ S/m.

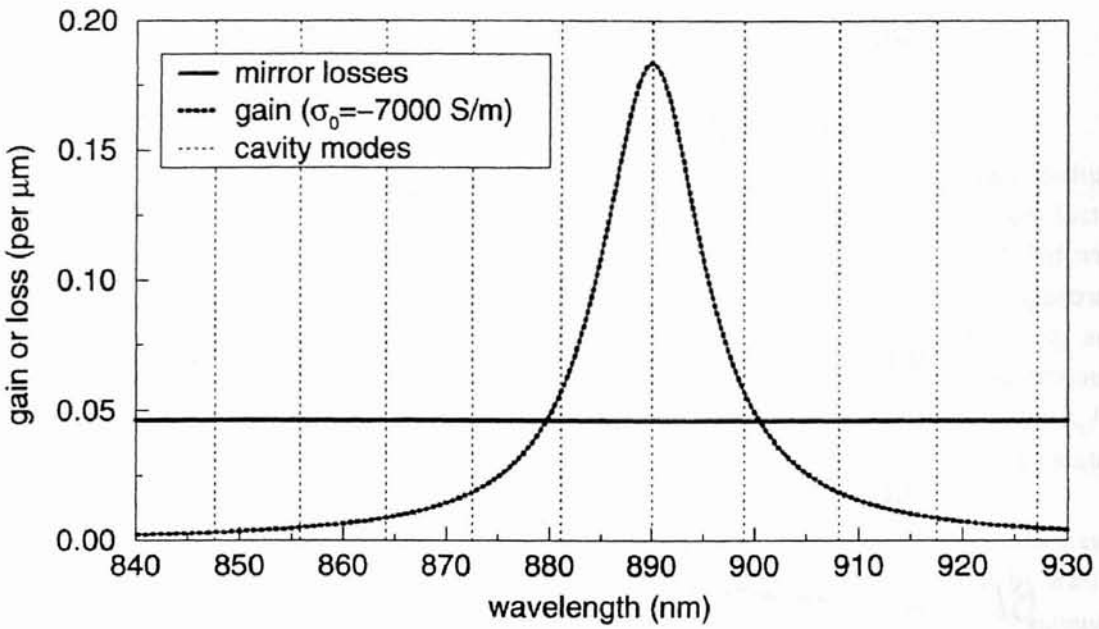


Fig. 9.9 Unsaturated gain and loss spectra for a Fabry-Perot microcavity laser. The longitudinal modes of the cavity are shown as vertical lines. *Adapted from:* Hagness et al., *Radio Science*, 1996, pp. 931–941.

Within a semiclassical framework, spontaneous emission can be included in Maxwell's equations as a noise current [31]. A pseudorandom number generator for a Gaussian random number sequence X [32], with zero mean and a variance of 1.0×10^{-6} (standard deviation $\sigma_x = 0.001$), is used to implement a noise current in the FDTD model of the laser cavity. The noise current is implemented in the FDTD algorithm at each time-step as an additive virtual current source at an E-field component located within the laser cavity:

$$E_z\big|_i^{n+1} = E_z\big|_i^{n+1} + \mathrm{X} \tag{9.125}$$

Fig. 9.10(a) shows the FDTD-computed time evolution of the *E*-field outside the laser cavity for $\sigma_0 = -7{,}000$ S/m. The oscillations build up from the noise background and then level off to a steady-state amplitude as the gain saturates. Fig. 9.10(b) shows an expanded time scale of the sinusoidal steady-state region of Fig. 9.10(a). This illustrates that the FDTD model correctly predicts a lasing frequency of ω_0. The simulation is then repeated for smaller values of σ_0, obtaining for each run the output intensity from the steady-state data. For gain levels above threshold, the output intensity is found to vary linearly with σ_0, a behavior that is expected for a homogeneously broadened system. Extrapolating the output intensity data towards zero yields $\sigma_{0,\text{th}} = -1{,}780$ S/m, a value that corresponds closely to $-1{,}760$ S/m calculated above. Thus, this study of a simple Fabry-Perot laser demonstrates the validity of using FDTD to provide accurate predictions of the gain threshold and lasing wavelength.

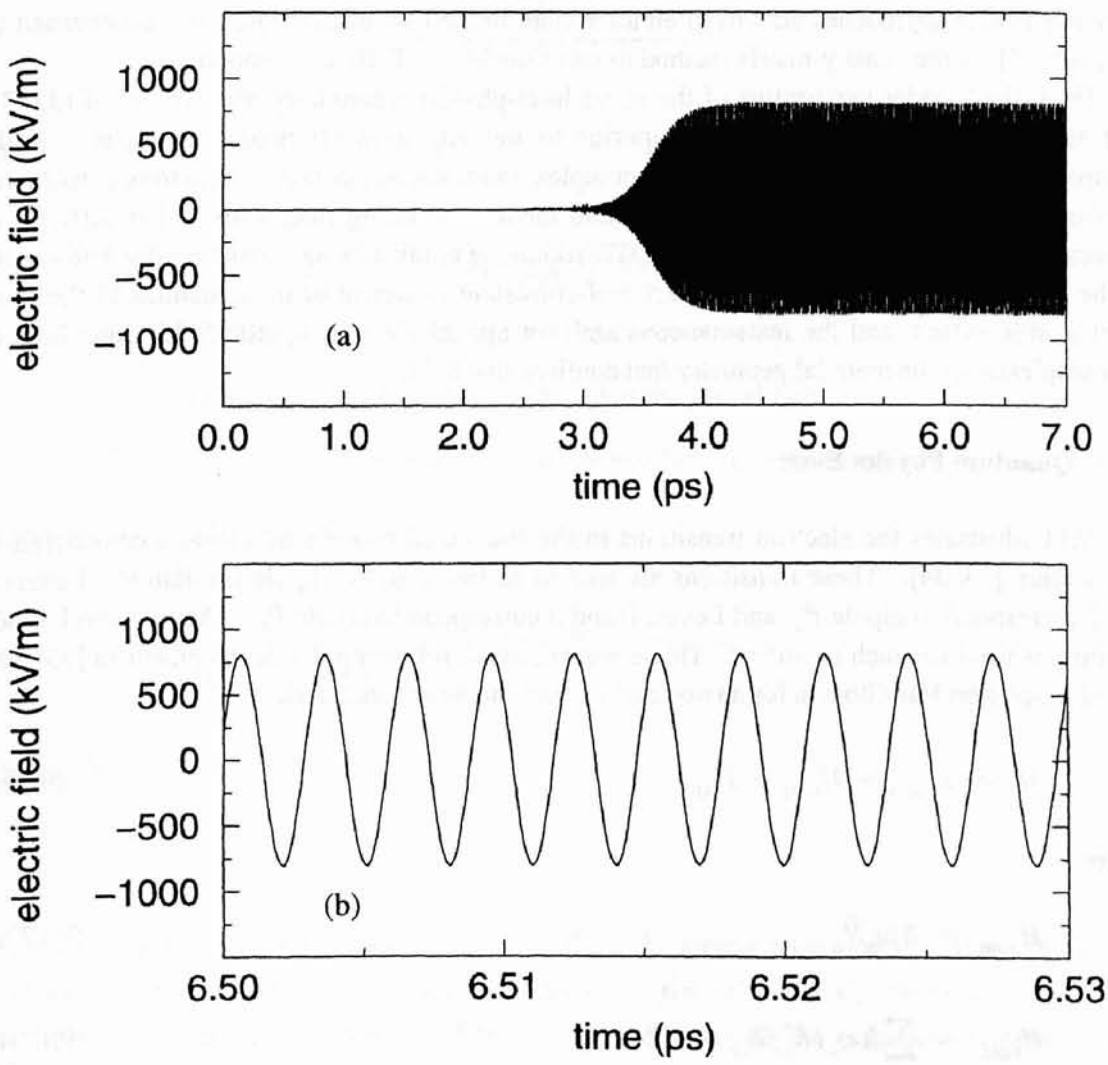


Fig. 9.10 (a) FDTD-computed time evolution of the *E*-field outside the Fabry-Perot laser cavity; (b) expanded time scale of the steady-state region showing a single-mode oscillation at ω_0. *Source:* Hagness et al., *Radio Science*, 1996, pp. 931–941.

9.8 AUXILIARY DIFFERENTIAL EQUATION METHOD, MODELING OF LASING ACTION IN A FOUR-LEVEL TWO-ELECTRON ATOMIC SYSTEM

This section reviews a new ADE-FDTD formulation [33, 34] for modeling lasing dynamics. Unlike the method of [28] reviewed in Section 9.7, the new technique is a semiclassical approach, wherein we treat the atom quantum mechanically and the electromagnetic wave classically. The new model incorporates quantized electron energies that specify four possible energy levels for each of two interacting electrons. Electron transitions between these energy levels are governed by four coupled differential (rate) equations and the *Pauli Exclusion Principle* (PEP) [30]. In addition, there are two oscillator-type differential equations for temporal evolution of the induced polarization densities.

From the FDTD perspective, the new ADE formulation appears to be an advance relative to previously reported techniques for semiclassical modeling of optical interactions with materials. Previous FDTD approaches employed either a more limited set of rate equations ungoverned by the PEP [35], or the density-matrix method to solve the Maxwell-Bloch equation [36].

From the broader perspective of the entire laser-physics community, the method of [33, 34] that is reviewed in this section is superior to the conventional modal expansion of the electromagnetic field. When dealing with complex, inhomogeneous laser geometries constructed in photonic crystals or light-localizing random media, the lasing modes are either difficult or impossible to calculate. The new ADE-FDTD technique eliminates the need for prior knowledge of the lasing modes. It provides a robust, self-consistent treatment of the dynamics of the four-level atomic system and the instantaneous ambient optical electromagnetic field, regardless of the complexity of the material geometry that confines this field.

9.8.1 Quantum Physics Basis

Fig. 9.11 illustrates the electron transitions in the four-level two-electron model considered in this section [33, 34]. These transitions are treated as two coupled dipole oscillators. Levels 1 and 2 correspond to dipole $\boldsymbol{P}_a$, and Levels 0 and 3 correspond to dipole $\boldsymbol{P}_b$. The two-level Bloch equation is used for each oscillator. This is now derived, following the development in [33, 34]. The atom-photon Hamiltonian for a two-level system can be expressed as:

$$H = H_{\text{Atom}} + H_{\text{Field}} + H_{\text{AF}} \tag{9.126}$$

where

$$H_{\text{Atom}} = \hbar \omega_a \hat{N}_u \tag{9.127a}$$

$$H_{\text{Field}} = \sum_{k,\sigma} \hbar \omega_k (\hat{a}^+_{k\sigma} \hat{a}_{k\sigma} + 1/2) \tag{9.127b}$$

$$H_{\text{AF}} = -\hat{\boldsymbol{\mu}} \cdot \hat{\boldsymbol{E}} \tag{9.127c}$$

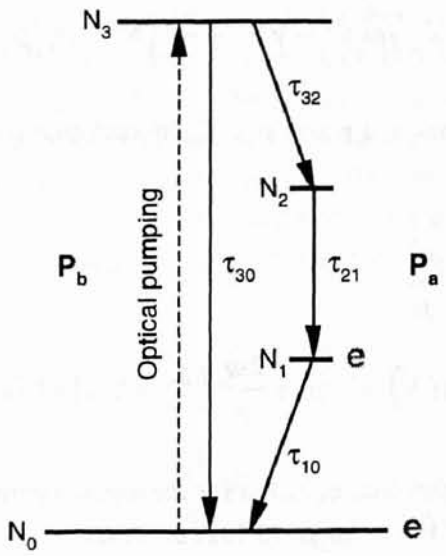


Fig. 9.11 Four-level two-electron model [33, 34].

$\hat{N}_u = |u><u| =$ number operator for the upper-level $|u>$

$\hat{N}_g = |g><g| =$ number operator for the upper-level $|g>$

$\hbar\omega_a =$ energy difference between $|u>$ and ground level $|g>$

$\hat{a}^+_{k\sigma} =$ photon-creation operator; $k =$ mode number; $\sigma =$ polarization direction

In (9.126) and (9.127), the electric field is given by

$$\hat{\boldsymbol{E}} = -\frac{\partial \hat{\boldsymbol{A}}}{\partial t} = j\sum_{k\sigma}\sqrt{\frac{\hbar\omega_k}{2\varepsilon_0 V}}\,\boldsymbol{e}_{k\sigma}\left(\hat{a}_{k\sigma}e^{jkx} - \hat{a}^+_{k\sigma}e^{-jkx}\right) \tag{9.128}$$

and the dipole operator is given by

$$\hat{\boldsymbol{\mu}} = -e\hat{\boldsymbol{r}} = \sum_{ll'}\boldsymbol{\mu}_{ll'}\hat{V}_{ll'} \tag{9.129}$$

where e is the electron charge and $\hat{V}_{ll'} = |l><l'|$ is the atomic transition operator. For a two-level system, the dipole operator can be written as $\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}\hat{V}^\dagger + \boldsymbol{\mu}^*\hat{V}$, where $\boldsymbol{\mu} = \langle u|e\hat{\boldsymbol{r}}|g\rangle$. The first-order differential equations of the dipole operator with the empirical dephasing term $\gamma\hat{V}$ are given by:

$$\frac{d\hat{V}}{dt} = \frac{j}{\hbar}\left[\hat{H}, \hat{V}\right] = -j\omega_a\hat{V} - \gamma\hat{V} + \frac{\omega_a}{\hbar}\left[\hat{N}_2 - \hat{N}_1\right]\boldsymbol{\mu}\cdot\hat{\boldsymbol{A}} \tag{9.130}$$

$$\frac{d\hat{V}^+}{dt} = \frac{j}{\hbar}\left[\hat{H}, \hat{V}^\dagger\right] = j\omega_a \hat{V}^+ - \gamma \hat{V}^+ + \frac{\omega_a}{\hbar}\left[\hat{N}_2 - \hat{N}_1\right]\boldsymbol{\mu}^* \cdot \hat{\boldsymbol{A}} \tag{9.131}$$

The direction of the dipole along the z-axis is defined such that $\boldsymbol{\mu} = \mu \boldsymbol{e}_z$. Therefore, the dipole operators are

$$\begin{aligned}\frac{d\hat{\mu}}{dt} &= \mu \frac{d\hat{V}^\dagger}{dt} + \mu^* \frac{d\hat{V}}{dt} \\ &= j\omega_a\left(\mu \hat{V}^\dagger - \mu^* \hat{V}\right) - \gamma\hat{\mu} + \frac{2\omega_a}{\hbar}\left[\hat{N}_2 - \hat{N}_1\right]|\mu|^2\left(\boldsymbol{e}_z \cdot \hat{\boldsymbol{A}}\right)\end{aligned} \tag{9.132}$$

Next, the second-order differential equation for the dipole operators in a two-level system is derived. Using $d/dt\,(\mu\hat{V}^\dagger - \mu^*\hat{V}) = j\omega_a\hat{\mu}$, (9.132) becomes

$$\frac{d^2\hat{\mu}}{dt^2} + 2\gamma\frac{d\hat{\mu}}{dt} + \left(\omega_a^2 + \frac{2\omega_a^2}{\hbar^2}|\mu|^2\hat{A}_z^2\right)\hat{\mu} = \frac{2\omega_a|\mu|^2}{\hbar}\hat{E}_z\left(\hat{N}_2 - \hat{N}_1\right) \tag{9.133}$$

The polarization density P_a(t) between |1> and |2> is then

$$\frac{d^2P_a}{dt^2} + \gamma_a\frac{dP_a}{dt} + \left[\omega_{12}^2 + \frac{2\omega_{12}^2}{\hbar^2}(\boldsymbol{\mu}_a \cdot \boldsymbol{A})^2\right]P_a = \frac{2\omega_{12}|\mu_a|^2}{\hbar}E_z(N_1 - N_2) \tag{9.134}$$

where ω_{12} is the resonance frequency, μ_a is the dipole matrix element between levels |1> and |2>, γ_a is the dephasing rate for $P_a(t)$, $\boldsymbol{A}$ is the vector potential, and N_i is the atomic population density in level i. A similar equation holds for the polarization density $P_b(t)$ between |0> and |3>:

$$\frac{d^2P_b}{dt^2} + \gamma_{30}\frac{dP_b}{dt} + \left[\omega_{30}^2 + \frac{2\omega_{30}^2}{\hbar^2}(\boldsymbol{\mu}_b \cdot \boldsymbol{A})^2\right]P_b = \frac{2\omega_{30}|\mu_b|^2}{\hbar}E_z(N_0 - N_3) \tag{9.135}$$

where ω_{30} is the pumping frequency. The Rabi oscillation term $2\omega_{30}^2(\boldsymbol{\mu}_b \cdot \boldsymbol{A})^2/\hbar^2$ in (9.135) is neglected in the subsequent development since it is important only when the external electric field is very high. This yields the governing equations [33, 34]:

$$\frac{d^2\boldsymbol{P}_a}{dt^2} + \gamma_a\frac{d\boldsymbol{P}_a}{dt} + \omega_a^2\boldsymbol{P}_a = \zeta_a(N_2 - N_1)\boldsymbol{E} \tag{9.136}$$

$$\frac{d^2\boldsymbol{P}_b}{dt^2} + \gamma_b\frac{d\boldsymbol{P}_b}{dt} + \omega_b^2\boldsymbol{P}_b = \zeta_b(N_3 - N_0)\boldsymbol{E} \tag{9.137}$$

where $\boldsymbol{P}_a$ and $\boldsymbol{P}_b$ have the resonant frequencies ω_a and ω_b, $\hbar\omega_a$ is the energy difference between Levels 1 and 2, $\hbar\omega_b$ is the energy difference between Levels 0 and 3, and $\zeta_a = 6\pi\varepsilon_0 c^3/\omega_{21}^2\tau_{21}$. In (9.136) and (9.137), the driving terms are proportional to the population differences, and the damping coefficients γ_a and γ_b simulate the nonradiation loss. For the example discussed later, $\gamma_a = \gamma_b = 10^{-13}$ s^{-1}. Furthermore, note that the electric field E is an instantaneous value that is composed of contributions from both the pumping and emission signals.

The population operator for the upper level is expressed as

$$\frac{d\hat{N}_u}{dt} = \frac{j}{\hbar}[\hat{H}, \hat{N}_u] = -\gamma(1 - N_g)\hat{N}_u - \frac{\omega_a}{\hbar}\hat{\boldsymbol{\mu}}\cdot\hat{\boldsymbol{A}}$$

$$= -\gamma(1 - N_g)\hat{N}_u + \frac{1}{\hbar\omega_a}\frac{\partial\hat{\boldsymbol{\mu}}}{\partial t}\cdot\hat{\boldsymbol{E}} \tag{9.138}$$

where N_u and N_g are defined in the context of (9.126) and (9.127). Here, we include the spontaneous decay to the lower level by the term $-\gamma(1 - N_g)\hat{N}_u$.

Due to the PEP, the presence of electrons within one energy level reduces the efficiency of the pumping or relaxation from other levels, since each quantum state can be occupied by only one electron. Similar to a semiconductor band structure, for interband interactions ($3 \rightarrow 0$) or ($2 \rightarrow 1$), this results in pumping blocking that takes the form $\mu = \mu_0(1 - N)$. Here, μ_0 is the quantum efficiency when there are no electrons in the active region, and N is the electron population density probability. As a consequence, the quantum efficiency drops by a factor of $(1 - N)$. Similar to interband relaxations, intraband transitions ($3 \rightarrow 2$) or ($1 \rightarrow 0$) are reduced by a factor $(1 - N)$ due to the Fermi distribution of the electron population within the band.

By coupling the dipole oscillator $\boldsymbol{P}_a$ (between levels 1 and 2) with $\boldsymbol{P}_b$ (between levels 0 and 3), the preceding considerations lead to the following rate equations for the electron populations within the four energy levels [33, 34]:

$$\frac{dN_3}{dt} = -\frac{N_3(1 - N_2)}{\tau_{32}} - \frac{N_3(1 - N_0)}{\tau_{30}} + \frac{1}{\hbar\omega_b}\boldsymbol{E}\cdot\frac{d\boldsymbol{P}_b}{dt} \tag{9.139}$$

$$\frac{dN_2}{dt} = \frac{N_3(1 - N_2)}{\tau_{32}} - \frac{N_2(1 - N_1)}{\tau_{21}} + \frac{1}{\hbar\omega_a}\boldsymbol{E}\cdot\frac{d\boldsymbol{P}_a}{dt} \tag{9.140}$$

$$\frac{dN_1}{dt} = \frac{N_2(1 - N_1)}{\tau_{21}} - \frac{N_1(1 - N_0)}{\tau_{10}} - \frac{1}{\hbar\omega_a}\boldsymbol{E}\cdot\frac{d\boldsymbol{P}_a}{dt} \tag{9.141}$$

$$\frac{dN_0}{dt} = \frac{N_3(1 - N_0)}{\tau_{30}} + \frac{N_1(1 - N_0)}{\tau_{10}} - \frac{1}{\hbar\omega_b}\boldsymbol{E}\cdot\frac{d\boldsymbol{P}_b}{dt} \tag{9.142}$$

Here, N_i is the electron population density probability in Level i, and τ_{ij} is the decay time constant between levels i and j. The electron populations vary with pumping $\boldsymbol{E}\cdot(d\boldsymbol{P}/dt)$ and spontaneous emission decay $(N_i - N_j)/\tau_{ij}$. Electrons in Level 3 spontaneously decay to Levels 2 and 0 with decay time constants τ_{32} and τ_{30}, respectively. Electrons in Level 2 spontaneously decay to Level 1 with decay time constant τ_{21}.

9.8.2 Coupling to Maxwell's Equations

The quantum physics model that resulted in (9.136), (9.137), (9.139), (9.140), (9.141) and (9.142) couples to the macroscopic Maxwell's equations via Ampere's law [33, 34]:

$$\frac{d\boldsymbol{E}}{dt} = \frac{1}{\varepsilon}\nabla\times\boldsymbol{H} - \frac{1}{\varepsilon}N_{\text{density}}\left(\frac{d\boldsymbol{P}_a}{dt} + \frac{d\boldsymbol{P}_b}{dt}\right) \tag{9.143}$$

This permits investigating both transient processes as well as the asymptotic behavior at longer time scales.

9.8.3 Time-Stepping Algorithm

We now summarize the time-stepping algorithm used in [33, 34], assuming that the following variables are known at the beginning of the update: $\boldsymbol{P}_a^n$, $\boldsymbol{P}_b^n$, $\boldsymbol{E}^n$, $\boldsymbol{H}^{n+1/2}$, $N_1^{n-1/2}$, $N_3^{n-1/2}$, N_0^n, and N_2^n. Note that, where variables are required at time levels midway between their formal evaluation points, simple averaging is used. This yields the familiar "semi-implicit" formalism that results in fully explicit updates.

Step. 1. Update $\boldsymbol{P}_a$ and $\boldsymbol{P}_b$ to time-level $n+1$ by implementing (9.136) and (9.137) with explicit second-order temporal finite-differences centered at n. For example, (9.136) yields

$$\boldsymbol{P}_a^{n+1} = C_{a,1}\boldsymbol{P}_a^n + C_{a,2}\boldsymbol{P}_a^{n-1} + C_{a,3}\zeta_a\left(N_2 - N_1\right)\boldsymbol{E}^n \tag{9.144}$$

where

$$C_{a,1} = \frac{2-\omega_a^2(\Delta t)^2}{1+\gamma_a\Delta t/2};\quad C_{a,2} = -\frac{1-\gamma_a\Delta t/2}{1+\gamma_a\Delta t/2};\quad C_{a,3} = \frac{(\Delta t)^2}{1+\gamma_a\Delta t/2} \tag{9.145a, b, c}$$

$\boldsymbol{P}_b$ is updated to time-level $n+1$ in an analogous manner.

Step 2. With $\boldsymbol{P}_a$ and $\boldsymbol{P}_b$ now available at time-levels n and $n+1$, update $\boldsymbol{E}$ to time level $n+1$ by implementing (9.143) in the normal Yee manner with an explicit temporal finite-difference centered about $n+1/2$.

Step 3. Update N_3 to time-level $n+1/2$ by implementing (9.139) with an explicit temporal finite-difference centered about n:

$$N_3^{n+1/2} = \frac{\tau_{30}\tau_{32}\Delta t}{2\tau_{30}\tau_{32} + \tau_{30}(1-N_2^n)\Delta t + \tau_{32}(1-N_0^n)\Delta t} \cdot \left[N_3^{n-1/2}\left(\frac{N_0^n - 1}{\tau_{30}} + \frac{N_2^n - 1}{\tau_{32}} + \frac{2}{\Delta t} \right) + \frac{\boldsymbol{E}^n \cdot \left(\boldsymbol{P}_b^{n+1} - \boldsymbol{P}_b^{n-1} \right)}{\hbar \omega_b \Delta t} \right] \tag{9.146}$$

Step 4. Update N_1 to time-level $n+1/2$ by implementing (9.141) with an explicit temporal finite-difference centered about n:

$$N_1^{n+1/2} = \frac{\tau_{10}\tau_{21}\Delta t}{2\tau_{10}\tau_{21} + \tau_{10}N_2^n\Delta t + \tau_{21}(1-N_0^n)\Delta t} \cdot \left[N_1^{n-1/2}\left(\frac{N_0^n - 1}{\tau_{10}} - \frac{N_2^n}{\tau_{21}} + \frac{2}{\Delta t} \right) + \frac{2N_2^n}{\tau_{21}} - \frac{\boldsymbol{E}^n \cdot \left(\boldsymbol{P}_a^{n+1} - \boldsymbol{P}_a^{n-1} \right)}{\hbar \omega_a \Delta t} \right] \tag{9.147}$$

Step 5. Update N_2 to time-level $n+1$ by implementing (9.140) with an explicit temporal finite-difference centered about $n+1/2$:

$$N_2^{n+1} = \frac{\tau_{21}\tau_{32}\Delta t}{2\tau_{21}\tau_{32} + \tau_{21}N_3^{n+1/2}\Delta t + \tau_{32}(1-N_1^{n+1/2})\Delta t} \cdot \left[N_2^n\left(\frac{N_1^{n+1/2} - 1}{\tau_{21}} - \frac{N_3^{n+1/2}}{\tau_{32}} + \frac{2}{\Delta t} \right) + \frac{2N_3^{n+1/2}}{\tau_{32}} + \frac{(\boldsymbol{E}^{n+1} + \boldsymbol{E}^n) \cdot \left(\boldsymbol{P}_a^{n+1} - \boldsymbol{P}_a^n \right)}{\hbar \omega_a \Delta t} \right] \tag{9.148}$$

Step 6. Update N_0 to time-level $n+1$ by implementing (9.142) with an explicit temporal finite-difference centered about $n+1/2$:

$$N_0^{n+1} = \frac{\tau_{10}\tau_{30}\Delta t}{2\tau_{10}\tau_{30} + \tau_{10}N_3^{n+1/2}\Delta t + \tau_{30}N_1^{n+1/2}\Delta t} \cdot \left[N_0^n\left(\frac{-N_1^{n+1/2}}{\tau_{10}} - \frac{N_3^{n+1/2}}{\tau_{30}} + \frac{2}{\Delta t} \right) + \frac{N_1^{n+1/2}}{\tau_{10}} + \frac{N_3^{n+1/2}}{\tau_{30}} - \frac{(\boldsymbol{E}^{n+1} + \boldsymbol{E}^n) \cdot \left(\boldsymbol{P}_b^{n+1} - \boldsymbol{P}_b^n \right)}{\hbar \omega_b \Delta t} \right] \tag{9.149}$$

Step 7. Update $\boldsymbol{H}$ to time-level $n+3/2$ via normal Yee time-stepping of the Maxwell-Faraday Law. This completes the algorithm. Note that this sequence of updates avoids the need to use the predictor-corrector algorithm of [36].

9.8.4 Illustrative Results

Fig. 9.12 illustrates laser-modeling results obtained using the ADE-FDTD model discussed in this section [34]. The laser is assumed to have a one-dimensional, optically pumped, single-defect, *distributed Bragg reflector* (DBR) cavity. Each DBR has three layers of refractive indices alternating between $n = 1.0$ and 2.0, with thickness 375 and 187.5 nm, respectively. The gap between the DBRs is 750 nm, corresponding to a cold-cavity defect mode of 1.5 μm. This mode has a cold-cavity *quality factor* (Q) of 100, which is sufficiently small to achieve lasing with a relatively short FDTD running time.

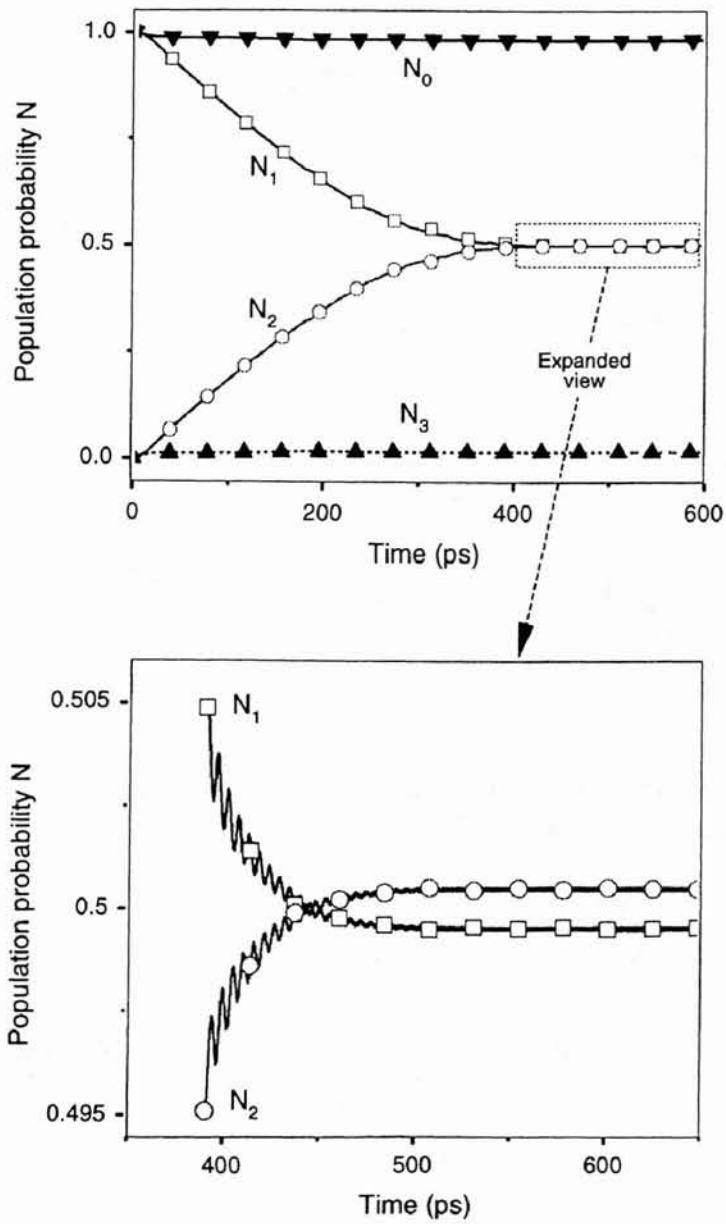


Fig. 9.12 Electron population density probability showing the inversion between Levels 1 and 2. *Source:* Chang and Taflove, *Optics Express*, 2004, pp. 3827–3833.

In each DBR, the passband is centered at 0.75 μm, which allows pump light at this wavelength to escape the cavity. Furthermore, the stopband is centered at 1.5 μm, which confines the lasing mode at this wavelength within the cavity. To computationally extract the lasing signal, the calculated output time-waveform is recorded on one side of the DBR cavity, transformed to the frequency domain with an FFT, windowed by a flat-topped Gaussian filter function, and finally transformed back to the time domain with another FFT.

In the present example [34], $\tau_{32} = \tau_{10} = 100$ fs and $\tau_{21} = \tau_{30} = 300$ ps, which takes into account the dephasing time. The initial population density is $N_1 = N_0 = 5 \times 10^{23}/\text{m}^3$. With these parameters, the required density probability of population inversion is estimated to be approximately 8.4×10^{-4}.

Fig. 9.12 shows the calculated time evolution of the electron populations in all four levels. After the onset of pumping, the population in Level 1 starts to decrease, and the population in Level 2 starts to increase. This is because pumping moves electrons from Level 1 to Level 0, then to Level 3, and finally to Level 2. Because of the relatively long decay time from Level 2 to Level 1, the population of Level 2 increases until inversion relative to Level 1 is reached. As shown in Fig. 9.13, the laser output signal jumps upward at this point, followed by fluctuations that have been predicted in the literature [37].

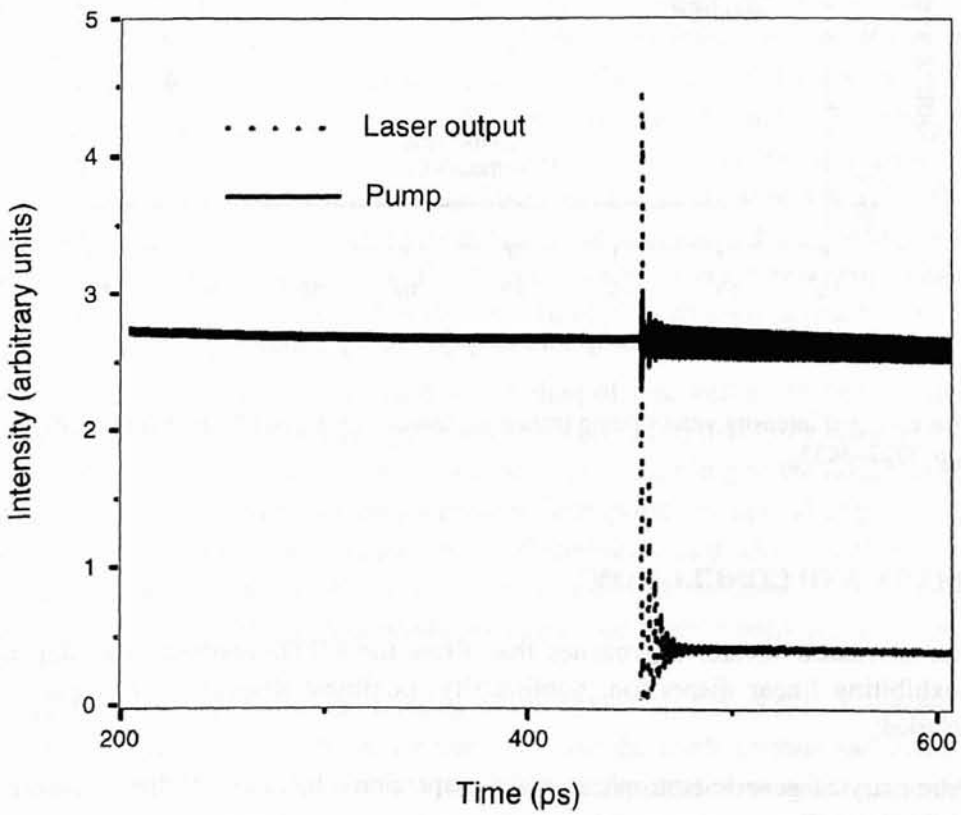


Fig. 9.13 Intensity output of the pump and laser output signals. *Source:* Chang and Taflove, *Optics Express*, 2004, pp. 3827–3833.

To obtain the lasing threshold, the output intensity is plotted versus the pump intensity as shown in Fig. 9.14 for four cases: (1) two-electron model with PEP, (2) two-electron model without PEP, (3) one-electron model with PEP, and (4) one-electron model without PEP. For both cases with the PEP, the lasing threshold is evidenced by a physically plausible sudden jump of the output intensity. On the other hand, omission of the PEP for the one-electron case results in a nonphysical zero-threshold laser operation. Further, omission of the PEP for the two-electron case results in a nonphysical situation where both electrons occupy the ground state, and there is no light output for the pump intensity considered. These calculations illustrate the importance of incorporating the Pauli Exclusion Principle.

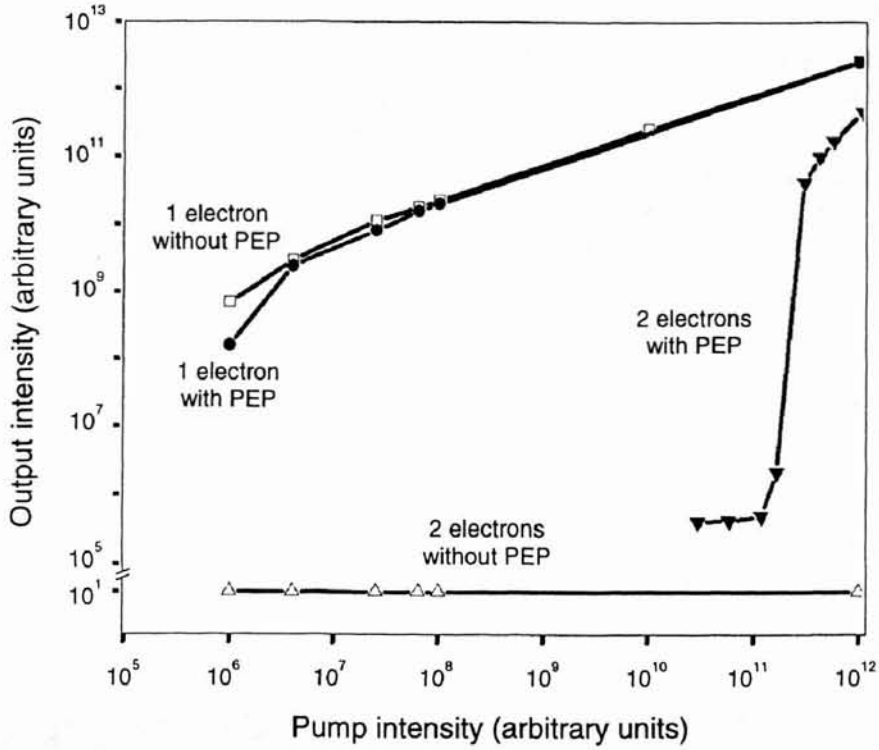


Fig. 9.14 Laser output intensity versus pump intensity. *Source:* Chang and Taflove, *Optics Express*, 2004, pp. 3827–3833.

9.9 SUMMARY AND CONCLUSIONS

This chapter discussed several approaches that allow the FDTD method to model dielectric materials exhibiting linear dispersion, nonlinearity, nonlinear dispersion, and gain. Topics covered included:

- A summary of generic isotropic material dispersions, including Debye, Lorentz, and Drude media;
- The piecewise-linear recursive convolution technique for modeling linear dispersions, with application to Debye and Lorentz media;

- The auxiliary differential equation technique for modeling linear dispersions characterized by multiple Debye, Lorentz, or Drude poles;
- Analysis of linear magnetized ferrites using a simple lumped equivalent resistor-inductor-capacitor circuit model;
- The auxiliary differential equation technique for modeling combined Debye and Lorentz linear dispersions and third-order Kerr and Raman nonlinearities, with application to analyzing the formation, propagation, and interactions of temporal and spatial optical solitons;
- The auxiliary differential equation technique for modeling saturable, dispersive optical gain materials at the macroscopic level, with application to estimating the output wavelength and power of a simple one-dimensional laser cavity;
- The auxiliary differential equation technique for coupling quantum effects in a four-level, two-electron atomic system to Maxwell's equations, with application to the detailed analysis of a laser's pumping dynamics and the variation of the intensity of its light output relative to the intensity of the pump.

In particular, the FDTD techniques reviewed in the last three sections of this chapter permit detailed modeling of the electromagnetic wave behavior of a wide variety of passive and active photonic materials. Unlike prior analytical or numerical approaches in this area, FDTD is almost completely general. It assumes nothing about: (1) the homogeneity or isotropy of the optical medium; (2) the magnitude of any nonlinearity; (3) the nature of the material's linear and nonlinear dispersions, except that these dispersions can be approximated by combinations of generic relaxations; (4) the shape, duration, polarization, and numbers of interacting optical pulses; and (5) the direction of wave propagation or scattering. Consequently, FDTD provides much richer physical detail and is more robust than prior techniques for modeling micron-scale photonic devices and integrated circuits. FDTD permits optical structures comparable in size to one wavelength to be modeled without neglecting the physics of wave diffraction or the effects of nanometer-scale features.

We note at this point that the algorithmic coupling of Maxwell's equations to the quantum mechanical description of electron energy levels in an atomic system (described in Section 9.8) opens possibilities for a subsequent deeper theoretical understanding of the fascinating transition regime between fully quantum and fully classical descriptions of optical physics. This could potentially enable the future engineering design of nanoscale quantum optical devices of great technological importance, especially in the area of quantum computing. It is expected that theoretical and algorithmic research pursuing such goals will significantly intensify over the next several years.

Finally, while proving numerical stability of the FDTD procedure for the general nonlinear case is difficult, experience to date shows that modeling the nonlinearities and gains of optical materials require no more restrictive a time-step than is needed to ensure stability for the corresponding linear dispersive case. Numerical stability has been observed for literally hundreds of thousands of time-steps.

REFERENCES

[1] Luebbers, R. J., and F. Hunsberger, "FDTD for Nth-order dispersive media," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 1297–1301.

[2] Luebbers, R. J., D. Steich, and K. Kunz, "FDTD calculation of scattering from frequency-dependent materials," *IEEE Trans. Antennas Propagat.*, Vol. 41, 1993, pp. 1249–1257.

[3] Kelley, D. F., and R. J. Luebbers, "Piecewise linear recursive convolution for dispersive media using FDTD," *IEEE Trans. Antennas Propagat.*, Vol. 44, 1996, pp. 792–797.

[4] Kashiwa, T., and I. Fukai, "A treatment by FDTD method of dispersive characteristics associated with electronic polarization," *Microwave Optics Tech. Lett.*, Vol. 3, 1990, pp. 203–205.

[5] 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 Lett.*, Vol. 16, 1991, pp. 1412–1414.

[6] Okoniewski, M., M. Mrozowski, and M. A. Stuchly, "Simple treatment of multi-term dispersion in FDTD," *IEEE Microwave Guided Wave Lett.*, Vol. 7, 1997, pp. 121–123.

[7] Pereda, J. A., L. A. Vielva, A. Vegas, and A. Prieto, "A treatment of magnetized ferrites using the FDTD method," *IEEE Microwave Guided Wave Lett.*, Vol. 3, 1993, pp. 136–138.

[8] Okoniewski, M., and E. Okoniewska, "FDTD analysis of magnetized ferrites: A more efficient approach," *IEEE Microwave Guided Wave Lett.*, Vol. 4, 1994, pp. 169–171.

[9] Pereda, J. A., L. A. Vielva, M. A. Solano, A. Vegas, and A. Prieto, "FDTD analysis of magnetized ferrites: Application to the calculation of dispersion characteristics of ferrite-loaded waveguides," *IEEE Trans. Microwave Theory Tech.*, Vol. 43, 1995, pp. 350–356.

[10] Kunz, K., and R. Luebbers, *The Finite-Difference Time-Domain Method for Electromagnetics*, Ch. 15, Orlando, FL: CRC Press, 1993.

[11] Gwarek, W. K., and A. Moryc, "An alternative approach to FDTD analysis of magnetized ferrites," *IEEE Microwave Wireless Components Lett.*, Vol. 14, 2004, pp. 331–333.

[12] Gwarek, W., "Analysis of an arbitrarily shaped planar circuit — A time-domain approach," *IEEE Trans. Microwave Theory Tech.*, Vol. 33, 1985, pp. 1067–1072.

[13] Hasagawa, A., *Optical Solitons in Fibers*, New York: Springer-Verlag, 1990.

[14] Agrawal, G. P., *Nonlinear Fiber Optics*, New York: Academic Press, 1989.

[15] Feit, M., and J. W. Fleck, Jr., "Three-dimensional analysis of a directional coupler exhibiting a Kerr nonlinearity," *IEEE J. Quantum Electronics*, Vol. 24, 1988.

[16] 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 Lett.*, Vol. 17, 1992, pp. 180–182.

[17] Goorjian, P. M., A. Taflove, R. M. Joseph, and S. C. Hagness, "Computational modeling of femtosecond optical solitons from Maxwell's equations," *IEEE J. Quantum Electronics*, Vol. 28, 1992, pp. 2416–2422.

[18] Joseph, R. M., P. M. Goorjian, and A. Taflove, "Direct time integration of Maxwell's equations in 2D dielectric waveguides for propagation and scattering of femtosecond electromagnetic solitons," *Optics Lett.*, Vol. 18, 1993, pp. 491–493.

[19] Joseph, R. M., and A. Taflove, "Spatial soliton deflection mechanism indicated by FDTD Maxwell's equations modeling," *IEEE Photonics Tech. Lett.*, Vol. 6, 1994, pp. 1251–1254.

[20] Ziolkowski, R. W., and J. B. Judkins, "Applications of the nonlinear finite difference time domain (NL-FDTD) method to pulse propagation in nonlinear media: Self-focusing and linear / nonlinear interfaces," *Radio Science*, Vol. 28, 1993, pp. 901–911.

[21] Ziolkowski, R. W., and J. B. Judkins, "Full-wave vector Maxwell equations modeling of self-focusing of ultrashort optical pulses in a nonlinear Kerr medium exhibiting a finite response time," *J. Opt. Soc. America B*, Vol. 10, 1993, pp. 186–198.

[22] Fujii, M., M. Tahara, I. Sakagami, W. Freude, and P. Russer, "High-order FDTD and auxiliary differential equation formulation of optical pulse propagation in 2D Kerr and Raman nonlinear dispersive media," *IEEE J. Quantum Electronics*, Vol. 40, 2004, pp. 175–182.

[23] Hellworth, R. W., "Third-order optical susceptibility of liquids and solids," *J. Prog. Quantum Electronics*, Vol. 5, 1977, pp. 1–68.

[24] Blow, K., and D. Wood, "Theoretical description of transient stimulated Raman scattering in optical fibers," *IEEE J. Quantum Electronics*, Vol. 25, 1989, pp. 2665–2673.

[25] Aitchison, J. S., A. M. Weiner, Y. Silberberg, D. E. Leaird, M. K. Oliver, J. L. Jackel, and P. W. E. Smith, "Experimental observation of spatial soliton interactions," *Optics Lett.*, Vol. 16, 1991, pp. 15–17.

[26] Desem, C., and P. L. Chu, "Reducing soliton interaction in single-mode optical fibres," *IEE Proc.*, Vol. 134, 1987, pp. 145–151.

[27] Lax, M., W. H. Louisell, and W. B. McKnight, "From Maxwell to paraxial wave optics," *Phys. Rev. A*, Vol. 11, 1975, p. 1365.

[28] Hagness, S. C., R. M. Joseph, and A. Taflove, "Subpicosecond electrodynamics of distributed Bragg reflector microlasers: Results from finite difference time domain simulations," *Radio Science*, Vol. 31, 1996, pp. 931–941.

[29] Yariv, A., *Quantum Electronics*, 3rd ed., New York: Wiley, 1989.

[30] Siegman, A., *Lasers*, Mill Valley, CA: University Science Books, 1986.

[31] Agrawal, G. P., and N. K. Dutta, *Semiconductor Lasers*, 2nd ed., New York: Van Nostrand Reinhold, 1993.

[32] Stearns, D. S., and R. David, *Signal Processing Algorithms Using Fortran and C*, Englewood Cliffs, NJ: PTR Prentice Hall, 1993.

[33] Huang, Y., *Simulation of Semiconductor Materials Using FDTD Method*, M.S. thesis, Northwestern University, Evanston, IL, May 2002.

[34] Chang, S.-H., and A. Taflove, "Finite-difference time-domain model of lasing action in a four-level two-electron atomic system," *Optics Express*, Vol. 12, 2004, pp. 3827–3833.

[35] Nagra, A. S., and R. A. York, "FDTD analysis of wave propagation in nonlinear absorbing and gain media," *IEEE Trans. Antennas Propagat.*, Vol. 46, 1998, pp. 334–340.

[36] Ziolkowski, R. W., J. M. Arnold, and D. M. Gogny, "Ultrafast pulse interactions with two-level atoms," *Phys. Rev. A*, Vol. 52, 1995, pp. 3082–3094.

[37] Singh, J., *Semiconductor Optoelectronics Physics and Technology*, New York: McGraw-Hill, 1995.

PROBLEMS

9.1 Extend the PLRC method of Section 9.3 to a material having a dielectric susceptibility characterized by a finite sum of Debye poles.

9.2 Extend the PLRC method of Section 9.3 to a material having a dielectric susceptibility characterized by a finite sum of Lorentz pole-pairs.

9.3 Extend the PLRC method of Section 9.3 to a material having a dielectric susceptibility characterized by a finite sum of Debye poles and a finite sum of Lorentz pole-pairs.

9.4 Extend the PLRC method of Section 9.3 to a material having a magnetic susceptibility characterized by a finite sum of Debye poles.

9.5 Extend the PLRC method of Section 9.3 to a material having a magnetic susceptibility characterized by a finite sum of Lorentz pole-pairs.

9.6 Extend the ADE method of Section 9.4 to a material having a dielectric susceptibility characterized by a finite sum of Debye poles and a finite sum of Lorentz pole pairs.

9.7 Extend the ADE method of Section 9.4 to a material having a magnetic susceptibility characterized by a finite sum of Debye poles.

9.8 Extend the ADE method of Section 9.4 to a material having a magnetic susceptibility characterized by a finite sum of Lorentz pole-pairs.

PROJECTS

P9.1 Construct a one-dimensional FDTD code using the PLRC technique of Section 9.3 and replicate the results shown in Fig. 9.1 for the pulse transmitted into a single pole-pair Lorentz half-space.

P9.2 Construct a one-dimensional FDTD code using the ADE technique of Section 9.4 and replicate the results shown in Fig. 9.1 for the pulse transmitted into a single pole-pair Lorentz half-space.

P9.3 Construct a one-dimensional FDTD code using the ADE technique of Section 9.6 and replicate the results shown in Fig. 9.5(b) for the temporal soliton transmitted into a nonlinear half-space.

P9.4 Construct a two-dimensional FDTD code using the ADE technique of Section 9.6 and replicate qualitatively the results shown in Fig. 9.7 for spatial solitons propagating in a nonlinear half-space.

P9.5 Construct a one-dimensional FDTD code using the ADE technique of Section 9.7 and replicate the results shown in Fig. 9.8 for the amplification factor and phase spectra.

Chapter 10

Local Subcell Models of Fine Geometrical Features

Allen Taflove, Malgorzata Celuch-Marcysiak, and Susan Hagness

10.1 INTRODUCTION

A fundamental problem arises in any grid-based numerical modeling tool for basic science and engineering simulations. Namely, the distance scales over which key physical processes or material / structural properties must be resolved can range over several orders of magnitude. The modeler is left with the following options:

1. Use a variable lattice of space cells to precisely model the shape and features of a structure. During a separate preprocessing step, use sophisticated mesh-generation software to automatically grade the cells' sizes and shapes from one region to another. The resulting mesh completely contains the structure and accommodates the full range of its characteristic distance scales and features.

2. Use a simpler and much more uniform mesh than option 1. Approximate the physical properties of the fine spatial details by somehow building them into the local mesh cells that are immediately adjacent to the details.

From the twin standpoints of maintaining geometrical fidelity and second-order numerical accuracy everywhere in the space grid, option 1 is the choice if: (1) computer software is available to generate the mesh, (2) the mesh generation does not dominate the total simulation time, and (3) the mesh generation avoids introducing space cells that are so strangely shaped that the numerical algorithm becomes unstable. Recall that numerical instability is an ever-present concern for an explicit time-stepping algorithm such as FDTD, due to the requirement that the time-step have an upper bound related to the space increment.

The accuracy of the local subcell approach may be less than that of a properly implemented global mesh generation. However, for a number of important modeling problems, the subcell approach can have computer-resource and mesh-generation requirements that are relaxed enough relative to those of global mesh generation to prompt its use as the more desirable alternative. This chapter discusses FDTD technology that implements local subcell models for a background Cartesian mesh. For the reader interested in pursuing option 1, see Chapter 11 for a detailed discussion of a generalized FDTD algorithm on a globally variable array of space cells, and Chapter 19 for details regarding a robust new hybrid FDTD / finite-element algorithm that combines the computational efficiency of Yee's Cartesian space lattice with the geometrical modeling flexibility of finite elements.

This chapter first reviews the premise that the integral forms of Faraday's and Ampere's laws yield the free-space Yee algorithm as originally derived from the differential forms of these laws (Maxwell's curl equations). The initial application of this insight yields the two simplest local subcell approaches: the diagonal split-cell model for nonconforming PEC surfaces, and the average permittivity-permeability model for material surfaces. Subsequent discussions apply contour-path models to accurately simulate subcell slots in conducting plates, subcell-diameter wires, and curved PEC and dielectric surfaces. For the thin-wire case, it is shown that the contour-path approach allows incorporation of singularities of the E- and H-fields immediately adjacent to the wire. The subcell models are then extended to treat thin material sheets, surface impedance and the skin effect, and thin coatings on a PEC surface. The chapter concludes with the application of the subcell approach to provide accurate rest-frame models of PEC surfaces subject to relativistic translation and vibration.

10.2 BASIS OF CONTOUR-PATH FDTD MODELING

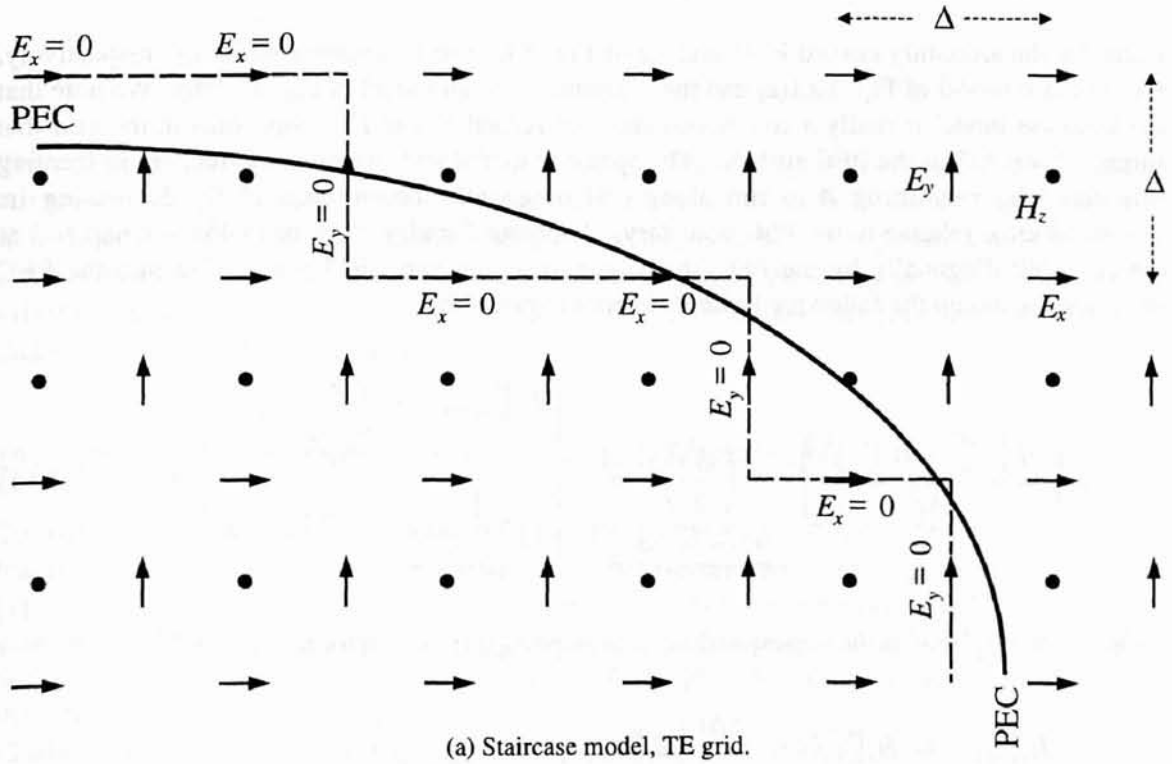
The Yee algorithm for FDTD was originally interpreted as a direct approximation of the pointwise derivatives of Maxwell's time-dependent curl equations by numerical central differences. Although this interpretation is useful for understanding how FDTD models electromagnetic wave propagation away from 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, curved surfaces, and thin material layers requiring subcell spatial resolution.

References [1 – 4] reported that FDTD modeling can be extended to wires, slots, and curved surfaces by departing from Yee's original pointwise derivative interpretation. As discussed in Chapter 3, Section 3.6.8, 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.

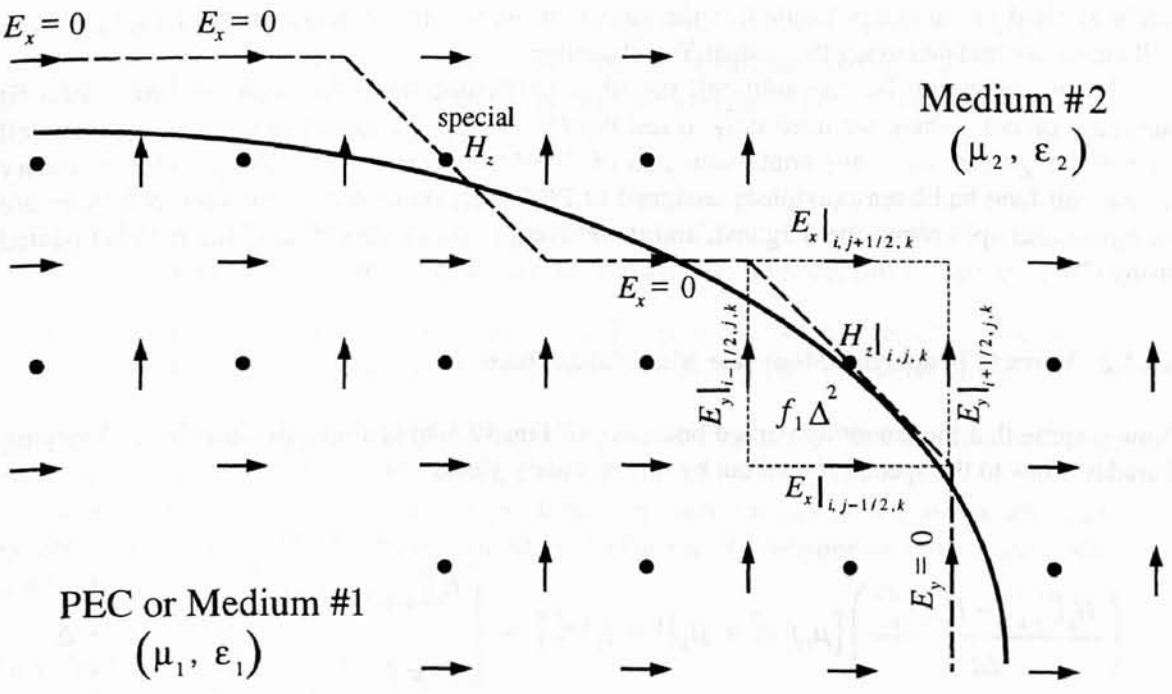
This meshing results in the filling of the FDTD modeled space by a three-dimensional "chain-link" array of intersecting orthogonal contours. The presence of wires, slots, and curved surfaces can be accounted for by incorporating appropriate field behavior into the contour and surface integrals implementing 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 model to deal with physical quantities such as electromotive forces and magnetomotive forces developed when completing one circuit about a Faraday's or Ampere's law contour path, and magnetic flux and electric displacement current when performing the surface integrations for the patches bounded by the respective contours. See Section 3.6.8 for a demonstration of the equivalence of the Yee and contour-path interpretations for the free-space case.

10.3 THE SIMPLEST CONTOUR-PATH SUBCELL MODELS

The two simple subcell models discussed here represent a minimal, yet surprisingly effective, approach to dealing with PEC and material structures whose bounding surfaces do not conform to grid planes in a uniform Cartesian mesh. Both are easily derived from the contour-path concepts of Sections 3.6.8. We shall refer to Fig. 10.1 in our discussion.



(a) Staircase model, TE grid.



(b) Diagonal split-cell model for PEC surface, average-properties model for material interface.

Fig. 10.1 The simplest subcell models for PEC surfaces and material interfaces.

10.3.1 Diagonal Split-Cell Model for PEC Surfaces

Consider the smoothly curved PEC surface of Fig. 10.1 and its approximation by, respectively, the staircase model of Fig. 10.1(a) and the diagonal split-cell model of Fig. 10.1(b). We note that the staircase model is really a continuous chain of zeroed E_x and E_y components in the grid that forms a "best fit" to the PEC surface. The split-cell model adds a degree of freedom in locating this chain by permitting it to run along cell diagonals, thereby potentially decreasing its positional error relative to the PEC boundary. Applying Faraday's law of (3.45) of Chapter 3 at the cell split diagonally by the PEC chain, and assuming zero field penetration into the PEC structure, we obtain the following finite-difference expression:

$$\mu_0 \left(\frac{H_z\big|_{i,j,k}^{n+1/2} - H_z\big|_{i,j,k}^{n-1/2}}{\Delta t} \right) \cdot \underbrace{\left(\frac{\Delta^2}{2} \right)}_{\substack{\text{area of split cell} \\ \text{carrying magnetic flux}}} = \left(E_x\big|_{i,j+1/2,k}^{n} - \underbrace{E_x\big|_{i,j-1/2,k}^{n}}_{\text{0 field in PEC}} + \underbrace{E_y\big|_{i-1/2,j,k}^{n}}_{\text{0 field in PEC}} - E_y\big|_{i+1/2,j,k}^{n} \right) \Delta \tag{10.1}$$

Isolating $H_z\big|_{i,j,k}^{n+1/2}$ yields the corresponding time-stepping expression for the split cell:

$$H_z\big|_{i,j,k}^{n+1/2} = H_z\big|_{i,j,k}^{n-1/2} + \frac{2\Delta t}{\mu_0 \Delta} \left(E_x\big|_{i,j+1/2,k}^{n} - E_y\big|_{i+1/2,j,k}^{n} \right) \tag{10.2}$$

Only H-field components located at the center of such split cells are updated using (10.2). All others are updated using the normal Yee algorithm.

Mesh generation for the split-cell model is predicated upon first constructing a best-fit staircase model. Then, we need only to test the H_x, H_y, or H_z component centered in each cell face of the grid to see if any contiguous pair of E-field components circulating on the boundary of the cell face had been previously assigned as PEC (i.e., set to zero). All such cell faces are designated as split along the diagonal, and the H-component at the center of the face is updated using (10.2).

10.3.2 Average Properties Model for Material Surfaces

Now assume that the smoothly curved boundary of Fig. 10.1(b) is a material interface. Applying Faraday's law to the special H_z cell cut by this boundary yields

$$\left(\frac{H_z\big|_{i,j,k}^{n+1/2} - H_z\big|_{i,j,k}^{n-1/2}}{\Delta t} \right) \left[\mu_1 f_1 \Delta^2 + \mu_2 (1 - f_1) \Delta^2 \right] = \left(E_x\big|_{i,j+1/2,k}^{n} - E_x\big|_{i,j-1/2,k}^{n} + E_y\big|_{i-1/2,j,k}^{n} - E_y\big|_{i+1/2,j,k}^{n} \right) \Delta \tag{10.3}$$

where $0 \le f_1 \le 1$ is the fraction of the cell area embedded in Medium #1. Isolating $H_z\big|_{i,j,k}^{n+1/2}$ yields the corresponding time-stepping expression:

$$H_z\Big|_{i,j,k}^{n+1/2} = H_z\Big|_{i,j,k}^{n-1/2} + \frac{\Delta t}{[\mu_1 f_1 + \mu_2(1-f_1)]\,\Delta}\begin{pmatrix} E_x\Big|_{i,j+1/2,k}^{n} - E_x\Big|_{i,j-1/2,k}^{n} \\ + E_y\Big|_{i-1/2,j,k}^{n} - E_y\Big|_{i+1/2,j,k}^{n} \end{pmatrix} \tag{10.4}$$

The cell permeability μ is seen to be the weighted average of the μ on each side of the interface. An analogous development leads to a weighted average for σ^* in this cell, and for ε and σ in E-cells cut by a dielectric interface. While this simple average-properties model does not rigorously enforce continuity of the tangential field components, it nonetheless provides an effective modeling tool. See Section 10.6.3 for a discussion of how this approach is used to treat dielectric structures having curved surfaces.

10.4 THE CONTOUR-PATH MODEL OF THE NARROW SLOT

To illustrate how the contour-path concept provides the basis for FDTD modeling of geometrical features requiring subcell resolution, we consider the two-dimensional slot geometry of Fig. 10.2 [1]. The slot is assumed to provide a subcell air gap that penetrates a planar PEC screen of finite size and thickness subjected to TE_z illumination. Faraday's law contour paths C_1, C_2, and C_3 are used to derive special FDTD updates for the H_z components located immediately adjacent to the screen.

The following summarizes the a priori assumptions concerning the near-field physics built into the Faraday's law models of Fig. 10.2.

- *Contour* C_1 (away from the slot). Field components H_z and E_y have no variation in the y-direction (perpendicular to the screen). Evaluated at the x-midpoint of contour C_1, components H_z and E_x represent the average values of their respective fields over the full x-interval.
- *Contour* C_2 (at the opening of the slot). Component H_z represents its average value within the free-space region enclosed by C_2. E_y has no variation in the y-direction. Component E_x represents its average value over the full x-interval.
- *Contour* C_3 (within the slot). Component H_z represents its average value over the full y-interval. H_z and E_x have no variation in the x-direction (across the slot gap).

Finally, for all three contours, the portions of the contours located within the conducting screen are assumed to have zero E- and H-fields.

After applying Faraday's law for the three contours subject to the above assumptions, the following special FDTD relations are obtained for the H_z components immediately adjacent to the screen.

Away from the Slot (Contour C_1):

$$\frac{H_z\Big|_{x,y_0}^{n+1/2} - H_z\Big|_{x,y_0}^{n-1/2}}{\Delta t} \equiv \frac{\left(E_y\Big|_{x-\Delta/2,\,y_0}^{n} - E_y\Big|_{x+\Delta/2,\,y_0}^{n}\right)\cdot(\Delta/2+\alpha) - E_x\Big|_{x,\,y_0-\Delta/2}^{n}\,\Delta}{\mu_0\Delta(\Delta/2+\alpha)} \tag{10.5}$$

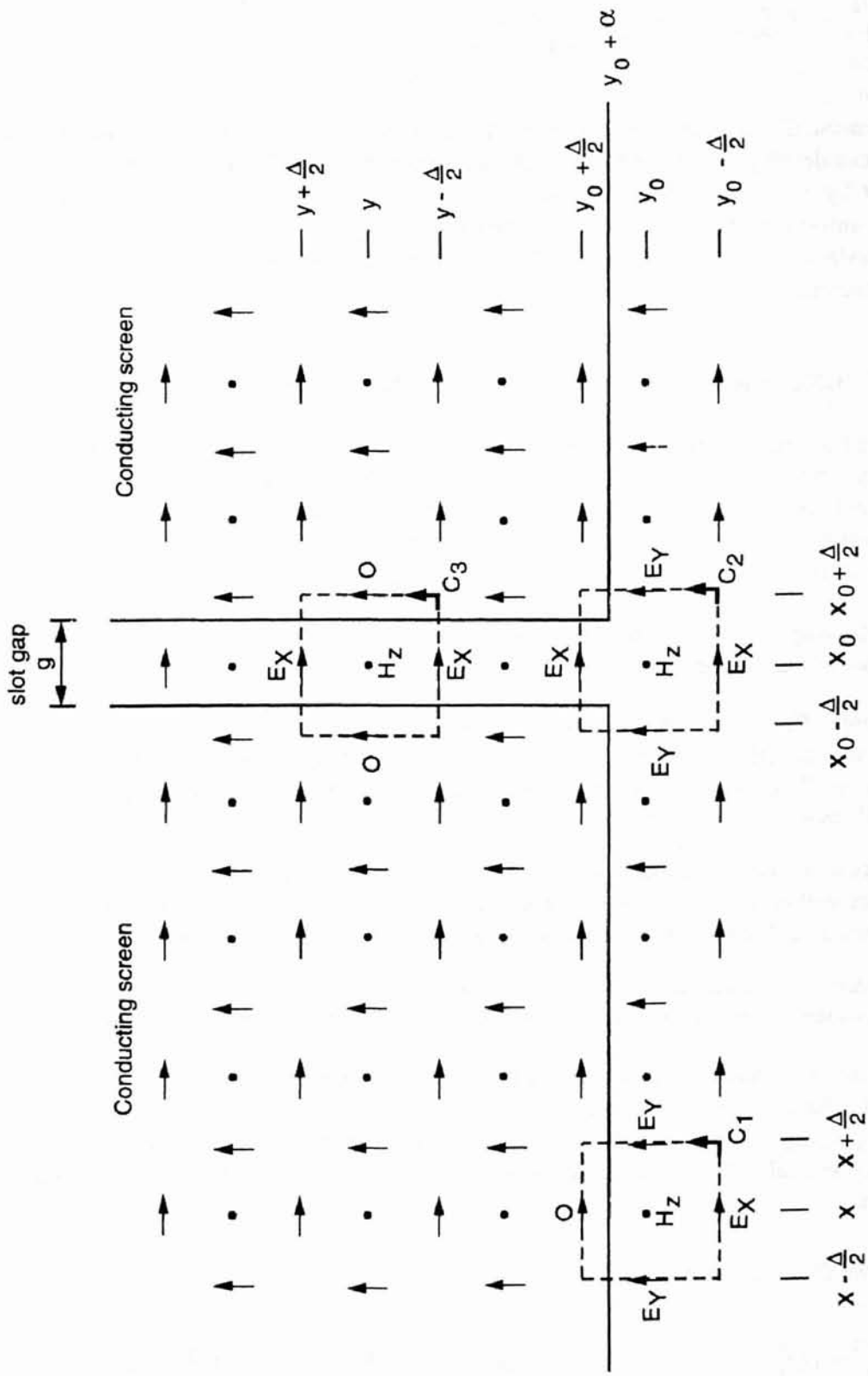


Fig. 10.2 Faraday's law contour paths for the narrow slot in a PEC screen. *Source:* Taflove et al., *IEEE Trans. Antennas and Propagation*, 1988, pp. 247–257, © 1988 IEEE.

At the Opening (Aperture) of the Slot (Contour C_2):

$$\frac{H_z\big|_{x_0,y_0}^{n+1/2} - H_z\big|_{x_0,y_0}^{n-1/2}}{\Delta t} \equiv \frac{\left[\begin{array}{l} E_x\big|_{x_0,\, y_0+\Delta/2}^{n}\, g \;-\; E_x\big|_{x_0,\, y_0-\Delta/2}^{n}\, \Delta \;+ \\ \left(E_y\big|_{x_0-\Delta/2,\, y_0}^{n} \;-\; E_y\big|_{x_0+\Delta/2,\, y_0}^{n}\right)\cdot(\Delta/2+\alpha) \end{array}\right]}{\mu_0\left[(\Delta/2+\alpha)\Delta \;+\; (\Delta/2-\alpha)\, g\right]} \tag{10.6}$$

Within the Slot (Contour C_3):

$$\frac{H_z\big|_{x_0,y}^{n+1/2} - H_z\big|_{x_0,y}^{n-1/2}}{\Delta t} \equiv \frac{E_x\big|_{x_0,\, y+\Delta/2}^{n}\, g \;-\; E_x\big|_{x_0,\, y-\Delta/2}^{n}\, g}{\mu_0\, g\, \Delta} \tag{10.7}$$

Here, the slot gap width g cancels on the right-hand side, reducing the FDTD relation for H_z in the slot to that of a $\pm y$-directed plane wave in free space.

In each of (10.5), (10.6), and (10.7), isolation of the $H_z\big|^{n+1/2}$ term on the left-hand side yields the corresponding FDTD time-stepping relation that conveys the electromagnetic field physics of the special cell. We also note that no $\boldsymbol{E}$ or $\boldsymbol{H}$ components in the FDTD grid other than the H_z components immediately adjacent to the screen require modified time-stepping relations. This is a very desirable general characteristic of the contour-path approach, recurring when modeling thin wires and curved surfaces, as will be seen.

We now consider a study of the accuracy of the Faraday's law contour-path model for a narrow slot having a subcell air gap [1]. Here, coarse-grid FDTD results, using the contour-path approach to treat the slot air gap as a fraction of one space cell, are compared to two different high-resolution numerical calculations. Fig. 10.3(a) shows the geometry of the slot that penetrates through a PEC shielding screen. The screen is $\lambda_0/10$ thick and extends $\lambda_0/2$ to each side of a straight slot having a gap dimension of $\lambda_0/40$. Excitation is provided by a TE_z-polarized plane wave at broadside incidence to the screen. Three types of predictive data are compared for the magnitude and phase of the gap-field (E_x) distribution within the slot along an observation locus extending from the lit side to the shadow side of the screen:

1. Contour-path FDTD model with $\Delta = \lambda_0/10$ (slot air gap equals 1/4 cell);
2. High-resolution FDTD model with $\Delta = \lambda_0/40$ (slot air gap equals 1 cell);
3. Very-high-resolution phasor-domain method-of-moments model which treats the slotted screen as a pure scattering geometry. $\lambda_0/400$ resolution in the slot is required for convergence.

We see from Fig. 10.3 that the low-resolution contour-path FDTD model agrees very well with both sets of high-resolution data in both the magnitude and the phase of the gap E_x field. Of particular interest in the area of electromagnetic pulse and high-power microwave effects is the ability of the contour-path model to accurately compute the peak E_x field in the slot. This field can cause arcing across the air gap if sufficiently intense. As discussed in [1], the coarse-grid contour-path FDTD model calculates the peak gap E_x field very accurately, even if the slot gap is closed to as little as $\lambda_0/1{,}000$ (1/100th of a space cell).

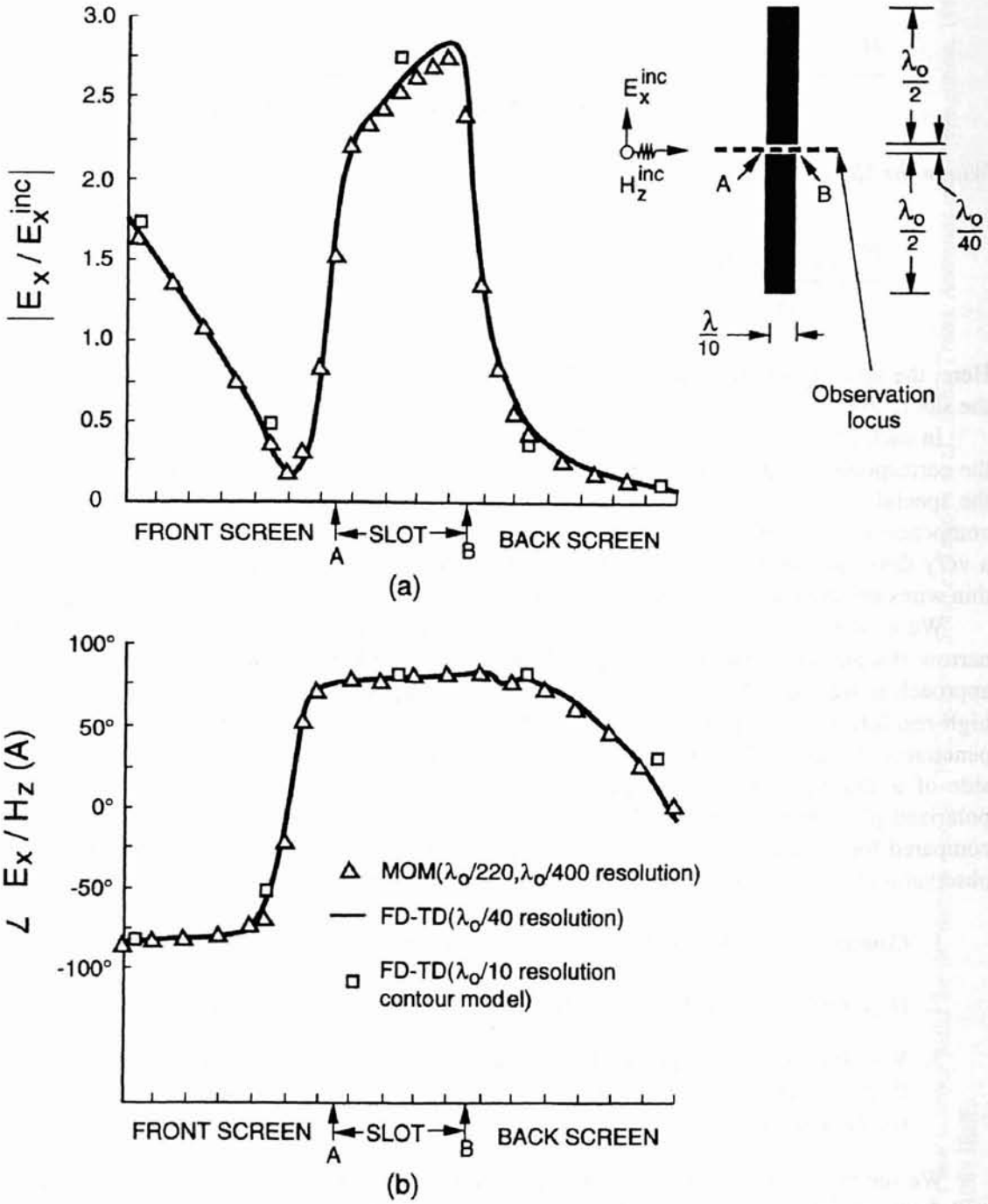


Fig. 10.3 Comparison of contour-path FDTD and frequency-domain moment-method results for the gap electric field distribution in a slotted PEC screen for broadside TE_z illumination: (a) magnitude; (b) phase. *Source:* Taflove et al., *IEEE Trans. Antennas and Propagation*, 1988, pp. 247–257, © 1988 IEEE.

10.5 THE CONTOUR-PATH MODEL OF THE THIN WIRE

The contour-path interpretation permits incorporation of near-field physics, yielding special-purpose time-stepping expressions that are *not* obvious from the pure finite-difference perspective. An excellent example of this involves the interaction of an electromagnetic wave with a PEC wire of circular cross section having a subcell diameter [2, 5]. Fig. 10.4 illustrates the geometry used in deriving special FDTD time-stepping expressions for the circumferential (looping) $\boldsymbol{H}$-components and radial $\boldsymbol{E}$-components immediately adjacent to a z-aligned thin wire. Although only H_y and E_x fields to the right of the wire are shown, the analysis is easily generalized to the other adjacent looping $\boldsymbol{H}$ and radial $\boldsymbol{E}$ components.

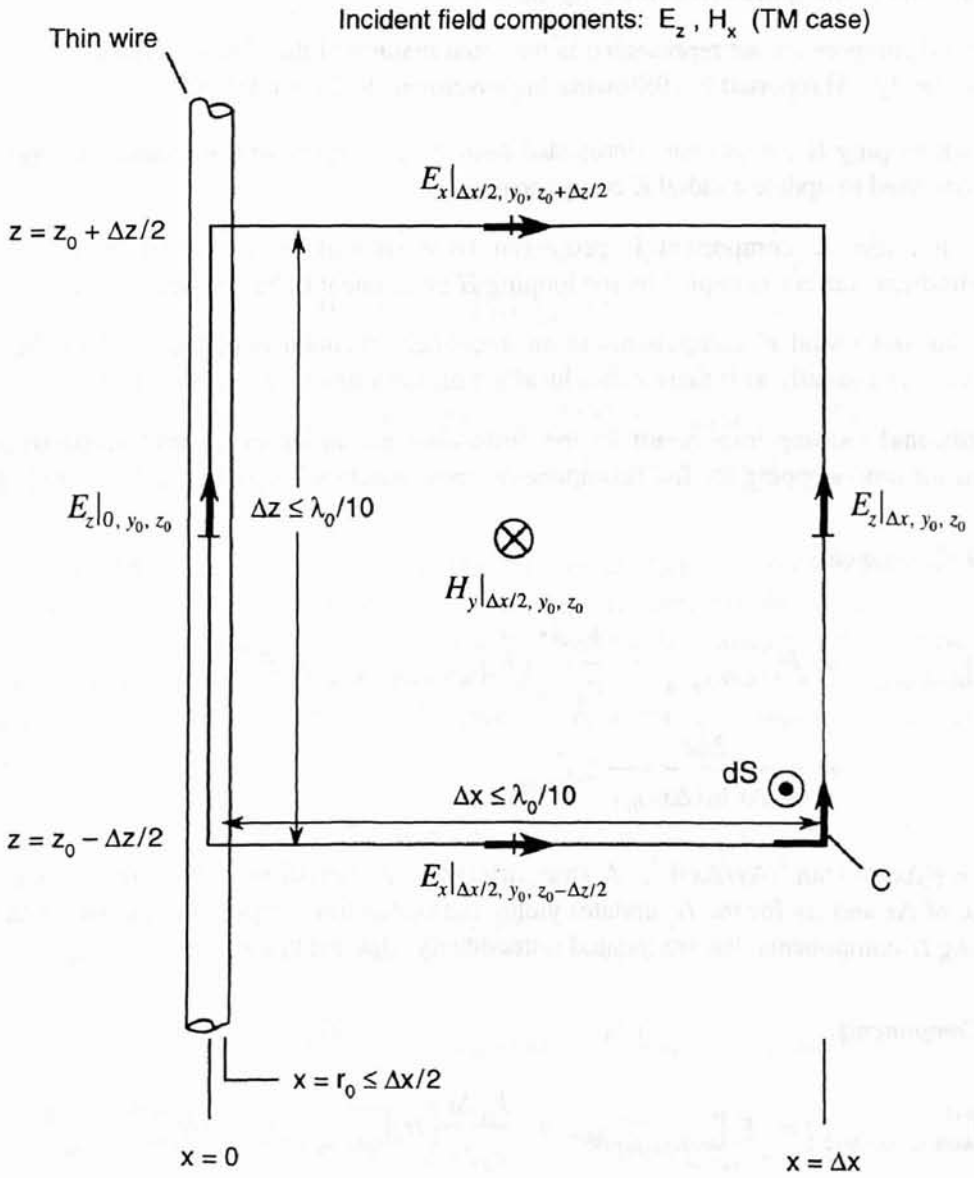


Fig. 10.4 Faraday's law contour path for the thin wire. *Source:* Umashankar et al., *IEEE Trans. Antennas and Propagation*, 1987, pp. 1248–1257, © 1987 IEEE.

The basic idea in the thin-wire model of [2] is that we can assume a static-field spatial distribution for the $\boldsymbol{E}$ and $\boldsymbol{H}$ components immediately adjacent to the wire. In effect, we assume that these components behave as if they are sourced by the local wire current and charge.

- Each looping $\boldsymbol{H}$ component nearest to the wire (for example, $H_y|_{\Delta x/2,\, y_0,\, z_0}$ in Fig. 10.4) varies as $1/r$, where r is the radial distance from the wire axis.
- Each radial $\boldsymbol{E}$ component nearest to the wire (for example, $E_x|_{\Delta x/2,\, y_0,\, z_0 \pm \Delta z/2}$ in Fig. 10.4) also varies as $1/r$.
- Each axial $\boldsymbol{E}$ component within the wire (for example, $E_z|_{0,\, y_0,\, z_0}$ in Fig. 10.4) is set to zero. This implies that no displacement current is allowed to flow in the axial direction in the space cell containing the wire.

All other field components are represented in the usual manner of the Yee algorithm.

Subsequently, [5] reported the following improvements to the model of [2]:

- Each looping $\boldsymbol{H}$ component is projected onto the its respective Cartesian cell edge when used to update a radial $\boldsymbol{E}$ component.
- Each radial $\boldsymbol{E}$ component is projected from its Cartesian cell facet to the cylindrical surface occupied by the looping $\boldsymbol{H}$ component to be updated.
- Axial and radial $\boldsymbol{E}$ components at an unconnected (open-circuited) end of the wire vary spatially as if there exists locally a uniform line-charge distribution.

These additional assumptions result in the following examples of special finite-difference expressions for time-stepping the field components immediately adjacent to the thin wire [5]:

Looping $\boldsymbol{H}$ Components:

$$H_y\Big|^{n+1/2}_{\Delta x/2,\, y_0,\, z_0} = H_y\Big|^{n-1/2}_{\Delta x/2,\, y_0,\, z_0} + \frac{k_E\,\Delta t}{\mu_0\,\Delta z}\left(E_x\Big|^{n}_{\Delta x/2,\, y_0,\, z_0-\Delta z/2} - E_x\Big|^{n}_{\Delta x/2,\, y_0,\, z_0+\Delta z/2}\right) + \frac{2\Delta t}{\mu_0\,\Delta x \ln(\Delta x/r_0)} E_z\Big|^{n}_{\Delta x,\, y_0,\, z_0} \tag{10.8}$$

where $k_E = [(\Delta x/\Delta y)\tan^{-1}(\Delta y/\Delta x)]^{-1}$. A straightforward permutation of the subscripts (with an interchange of Δx and Δy for the H_x update) yields analogous time-stepping relations for the three other looping $\boldsymbol{H}$ components that are located immediately adjacent to the wire at $z = z_0$.

Radial $\boldsymbol{E}$ Components:

$$E_x\Big|^{n+1}_{\Delta x/2,\, y_0,\, z_0-\Delta z/2} = E_x\Big|^{n}_{\Delta x/2,\, y_0,\, z_0-\Delta z/2} + \frac{k_H\,\Delta t}{\varepsilon_0\,\Delta z}\left(H_y\Big|^{n+1/2}_{\Delta x/2,\, y_0,\, z_0-\Delta z} - H_y\Big|^{n+1/2}_{\Delta x/2,\, y_0,\, z_0}\right) + \frac{\Delta t}{\varepsilon_0\,\Delta y}\left(H_z\Big|^{n+1/2}_{\Delta x/2,\, y_0+\Delta y/2,\, z_0-\Delta z/2} - H_z\Big|^{n+1/2}_{\Delta x/2,\, y_0-\Delta y/2,\, z_0-\Delta z/2}\right) \tag{10.9}$$

where $k_H = (\Delta x/\Delta y)\tan^{-1}(\Delta y/\Delta x)$. A straightforward permutation of the subscripts (with an interchange of Δx and Δy for the E_y update) yields analogous time-stepping relations for the three other radial $\boldsymbol{E}$ components that are located immediately adjacent to the wire at $z = z_0$.

End of Open-Circuited Wire, Radial $\boldsymbol{E}$ *Components Located at* $z = z_{\text{top}}$:

$$\begin{aligned} E_x\Big|^{n+1}_{\Delta x/2,\, y_0,\, z_{\text{top}}} &= E_x\Big|^{n}_{\Delta x/2,\, y_0,\, z_{\text{top}}} + (k_M/k_E)\, E^{\text{ave}}_{x,q}\Big|^{n+1/2}_{\Delta x/2,\, y_0,\, z_{\text{top}}} \\ &+ \frac{\Delta t}{\varepsilon_0 \Delta z}\left(k_H H_y\Big|^{n+1/2}_{\Delta x/2,\, y_0,\, z_{\text{top}}-\Delta z/2} - H_y\Big|^{n+1/2}_{\Delta x/2,\, y_0,\, z_{\text{top}}+\Delta z/2} \right) \\ &+ \frac{\Delta t}{\varepsilon_0 \Delta y}\left(H_z\Big|^{n+1/2}_{\Delta x/2,\, y_0+\Delta y/2,\, z_{\text{top}}} - H_z\Big|^{n+1/2}_{\Delta x/2,\, y_0-\Delta y/2,\, z_{\text{top}}} \right) \end{aligned} \tag{10.10}$$

where $k_E = [(\Delta x/\Delta y)\tan^{-1}(\Delta y/\Delta x)]^{-1}$, $k_M = (1 - r_0/\Delta x)$, and

$$E^{\text{ave}}_{x,q}\Big|^{n+1/2}_{\Delta x/2,\, y_0,\, z_{\text{top}}} \cong \left[\begin{array}{l} 1/\sqrt{2} + 0.166(r_0/\Delta x) - 3.65(r_0/\Delta x)^2 + \\ 8.91(r_0/\Delta x)^3 - 16.1(r_0/\Delta x)^4 \end{array} \right] E_{r,q}\Big|^{n+1/2} \tag{10.11}$$

$$E_{r,q}\Big|^{n+1/2} = \left(\frac{-I\big|^{n+1/2}_{z_{\text{top}}-\Delta z/2}\, \Delta t}{2\pi^2 \varepsilon_0} \right) \frac{(\Delta x/2 - r_0)\mathrm{K}(k_1) + (\Delta x/2 + r_0)\mathrm{E}(k_1)}{\Delta x\left[(\Delta x/2)^2 - r_0^2\right]} \tag{10.12}$$

In (10.12), K and E are complete elliptic integrals of the first and second kind, respectively, and $k_1 = 2(r_0 \Delta x/2)^{1/2}/(\Delta x/2 + r_0)$. Here, the current I is computed from the circulation of the looping H-field around the wire in a plane one-half space cell below its end. A straightforward permutation of the subscripts (with an interchange of Δx and Δy for the E_y update) yields analogous time-stepping relations for the three other radial $\boldsymbol{E}$ components that are located immediately adjacent to the wire at $z = z_{\text{top}}$.

End of Open-Circuited Wire, Axial $\boldsymbol{E}$ *Component Located at* $z = z_{\text{top}} + \Delta z/2$:

$$\begin{aligned} E_z\Big|^{n+1}_{0,\, y_0,\, z_{\text{top}}+\Delta z/2} &= E_x\Big|^{n}_{0,\, y_0,\, z_{\text{top}}+\Delta z/2} + E^{\text{ave}}_{z,q}\Big|^{n+1/2}_{0,\, y_0,\, z_{\text{top}}+\Delta z/2} \\ &+ \frac{\Delta t}{\varepsilon_0 \Delta y}\left(H_x\Big|^{n+1/2}_{0,\, y_0-\Delta y/2,\, z_{\text{top}}+\Delta z/2} - H_x\Big|^{n+1/2}_{0,\, y_0+\Delta y/2,\, z_{\text{top}}+\Delta z/2} \right) \\ &+ \frac{\Delta t}{\varepsilon_0 \Delta x}\left(H_y\Big|^{n+1/2}_{\Delta x/2,\, y_0,\, z_{\text{top}}+\Delta z/2} - H_y\Big|^{n+1/2}_{-\Delta x/2,\, y_0,\, z_{\text{top}}+\Delta z/2} \right) \end{aligned} \tag{10.13}$$

where

$$E_{z,q}^{\text{ave}}\Big|_{0,y_0,z_{\text{top}}+\Delta z/2}^{n+1/2} \equiv \left[\begin{array}{l}\pi/6 \;-\; 0.111(r_0/\Delta z) \;+ \\ 3.31(r_0/\Delta z)^2 \;-\; 2.51(r_0/\Delta z)^3\end{array}\right] E_{z,q}\Big|^{n+1/2} \tag{10.14}$$

$$E_{z,q}\Big|^{n+1/2} = \left(\frac{-I\Big|_{z_{\text{top}}-\Delta z/2}^{n+1/2}\Delta t}{8\pi\varepsilon_0}\right)\frac{\Delta z}{\left[(\Delta z/2)^2 + r_0^2\right]^{3/2}} \tag{10.15}$$

Reference [5] reported numerical experiments measuring the degree of accuracy of the thin-wire model reviewed above. In one test, an $L = 0.305$m dipole antenna was modeled within a uniform FDTD space lattice for two cell sizes, 21 and 41 cells per dipole length (14.5 and 7.44 mm, respectively). The lattice was terminated with an eight-cell PML, and a simple delta-gap hard source was used to excite the antenna. In the impedance calculations, voltage values were interpolated in time to compensate for their half time-step offset relative to the current values. Benchmark data were calculated using the phasor-domain moment-method software, NEC-2.

Fig. 10.5 graphs the calculated $|S_{11}|$ of the dipole over the frequency range 0.3 to 0.6 GHz. We see that the improved thin-wire model of [5] accurately predicts the resonant frequency. Even when implemented in the low-resolution mesh, the model of [5] yields better results than the original model of [2] run at high resolution. Reference [5] concludes that the improvement in accuracy is largely due to the treatment of the ends of the dipole, as per (10.10) to (10.15).

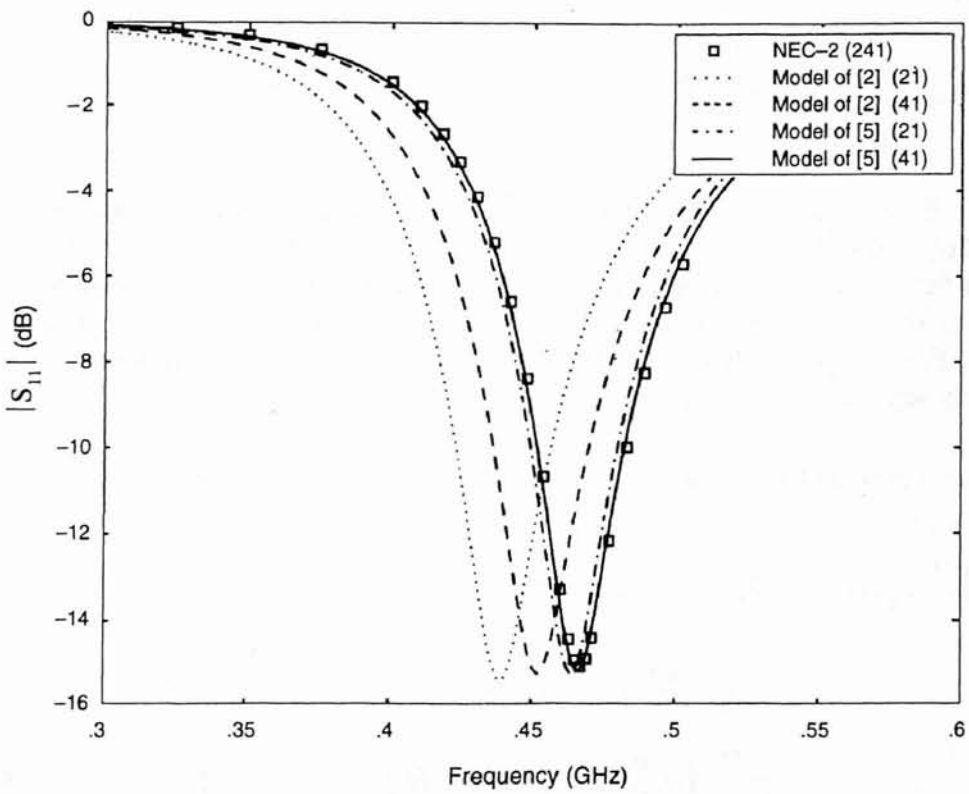


Fig. 10.5 Calculated $|S_{11}|$ of the dipole antenna. The number in the legend refers to the number of segments or cells used per antenna length L. Wire radius $r_0 = L/820$. *Source:* Mäkinen et al., *IEEE Trans. Microwave Theory and Techniques*, 2002, pp. 1245–1255, © 2002 IEEE.

Fig. 10.6 graphs the calculated input impedance of the dipole at frequency $f = c/2L$ over a wide range of wire radii. Even when implemented in the low-resolution mesh, the improved thin-wire model of [5] is more accurate than the original model of [2] run at high resolution.

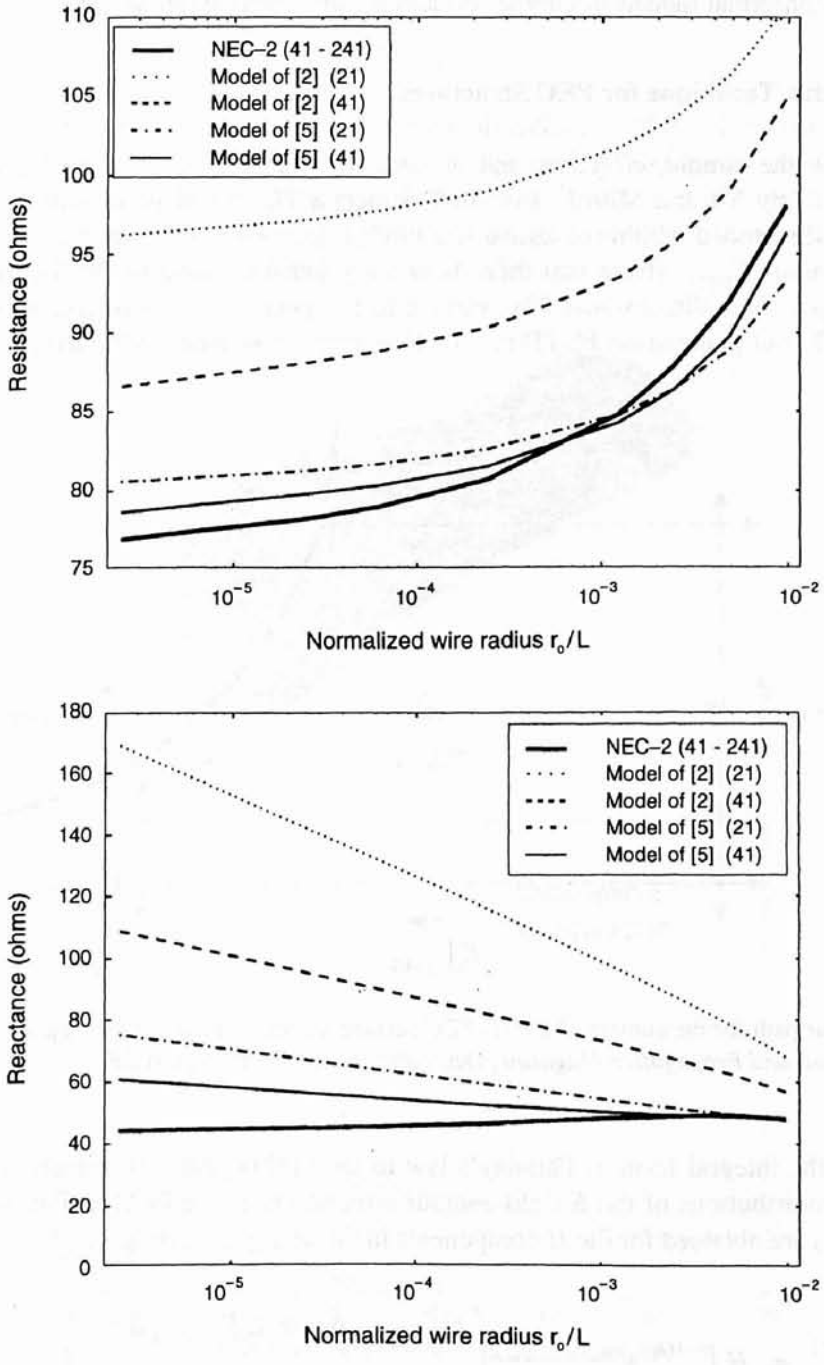


Fig. 10.6 Calculated input impedance of the dipole antenna at frequency $f = c/2L$ over a wide range of wire radii. The number in the legend refers to the number of segments or cells used per antenna length L. *Source:* Mäkinen et al., *IEEE Trans. Microwave Theory and Techniques*, 2002, pp. 1245–1255, © 2002 IEEE.

10.6 LOCALLY CONFORMAL MODELS OF CURVED SURFACES

The contour-path technique was proposed in [3, 4] to implement conformal models of structures having curved surfaces within a Cartesian FDTD lattice. This technique has evolved to have significantly improved numerical stability and accuracy [6 – 8]. The basis and illustrative results of such locally conformal models of curved surfaces are presented in this section.

10.6.1 Yu-Mittra Technique for PEC Structures

We now review the simple, effective, and numerically stable conformal modeling technique introduced in [8] by Yu and Mittra. Fig. 10.7 depicts a TE_z cut-plane in a three-dimensional FDTD lattice. Embedded within the lattice is a PEC structure whose surface intersects this cut-plane along contour C_{PEC}. (Note that there is an easy generalization to the TE_x and TE_y cut-planes, allowing a three-dimensional PEC surface to be specified by its intersections with each TE_x, TE_y, and TE_z cut-plane in the FDTD mesh.) Free space is assumed to be to the left of C_{PEC}.

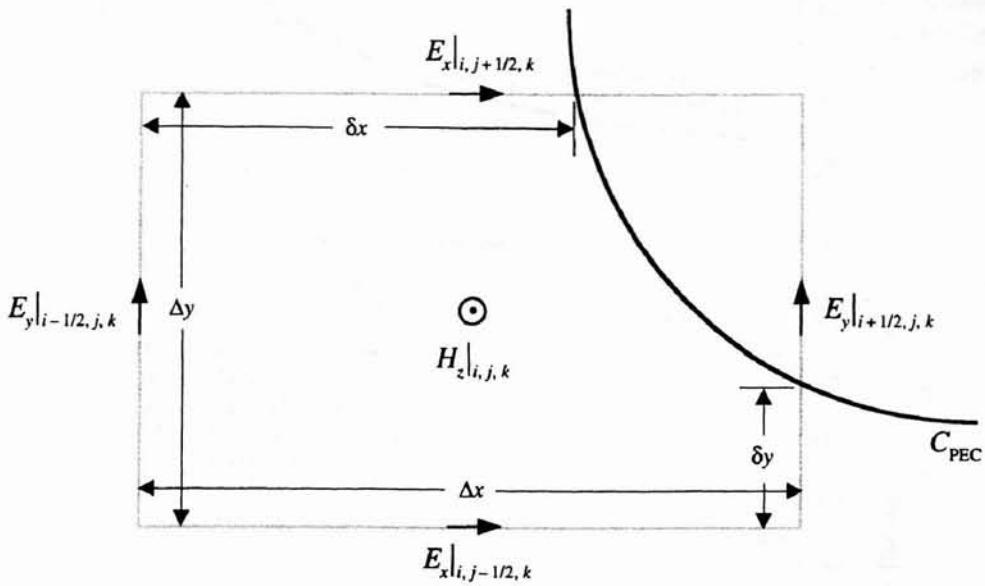


Fig. 10.7 Contour path for the conformal FDTD PEC surface model. *Adapted from:* Yu and Mittra, *IEEE Antennas and Propagation Magazine*, Oct. 2000, pp. 28–39, © 2000 IEEE.

Upon applying the integral form of Faraday's law to the FDTD grid cell intersecting C_{PEC}, and discarding the contributions of the E-field contour integral inside the PEC region, the following modified updates are obtained for the H-components in the cells intersecting C_{PEC}:

$$H_z\Big|_{i,j,k}^{n+1/2} = H_z\Big|_{i,j,k}^{n-1/2} + \frac{\Delta t}{\mu_0 \Delta x \Delta y}\begin{pmatrix} E_y\Big|_{i-1/2,j,k}^{n}\Delta y + E_x\Big|_{i,j+1/2,k}^{n}\delta x \\ - E_y\Big|_{i+1/2,j,k}^{n}\delta y - E_x\Big|_{i,j-1/2,k}^{n}\Delta x \end{pmatrix} \tag{10.16}$$

Note that the entire area $\Delta x\,\Delta y$ of the grid cell is used. Further, each E-component is updated exactly as in the usual Yee manner, whether located inside or outside of the PEC region.

10.6.2 Illustrative Results for PEC Structures

We now discuss three representative studies of locally conformal FDTD models of PEC structures having curved surfaces. The goal is to demonstrate the key characteristics of these models, especially their improved accuracy relative to the traditional Yee staircasing method.

Twisted Elliptical Cross Section Waveguide Cavity

Reference [9] reported application of the contour-path technique of [6, 7] to analyze the resonant frequency of a PEC cavity comprised of a twisted waveguide of elliptical cross section. Here, the benchmark was a highly resolved phasor-domain finite-element calculation using Ansoft HFSS. Fig. 10.8 shows the resonator geometry and compares the error in its resonant frequency calculated by the contour-path and staircase FDTD models.

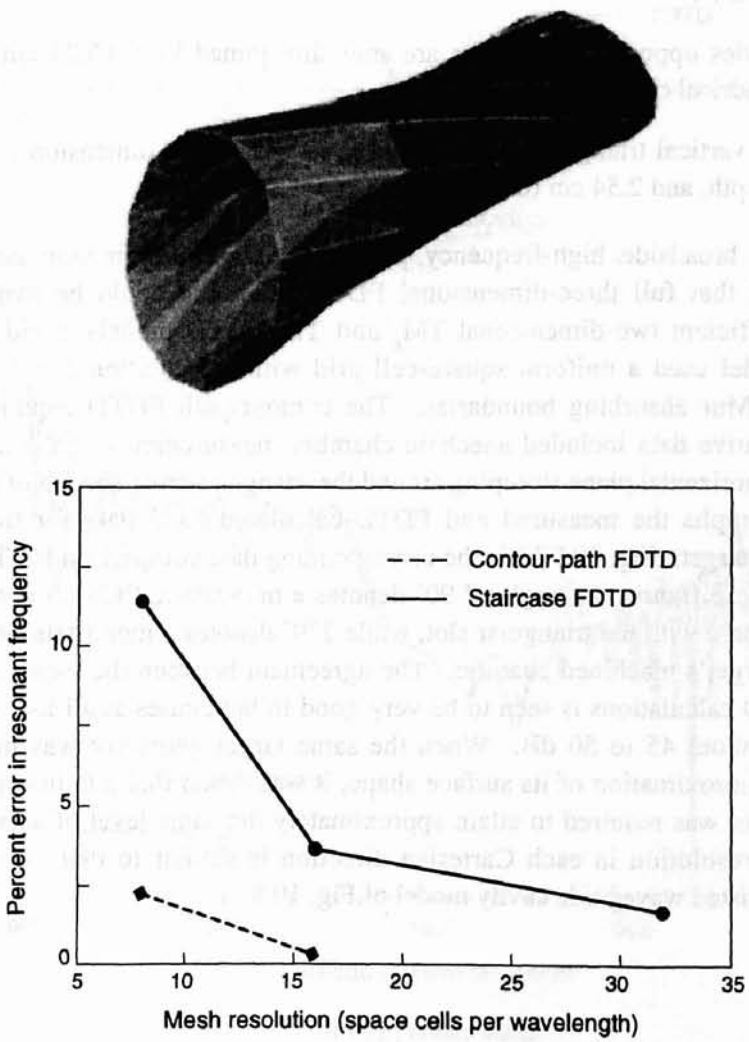


Fig. 10.8 Twisted elliptical PEC waveguide cavity and comparison of error in its fundamental-mode resonant frequency calculated using contour-path and staircase FDTD. *Source:* Waldschmidt and Taflove, *IEEE Trans. Antennas and Propagation*, 2004, pp. 1658–1664, © 2004 IEEE.

From Fig. 10.8, we see that contour-path FDTD yields an error at a mesh resolution of 8 cells per wavelength which is comparable to that obtained using staircased FDTD at 32 cells per wavelength. This represents a very significant computer-storage and running-time reduction of approximately $(32/8)^3:1 = 64:1$ and $(32/8)^4:1 = 256:1$, respectively.

Winglike Object

The second example is a detailed study conducted in 1990 of the monostatic RCS at 10 GHz of a winglike aluminum plate having acute 11° angles for its leading and trailing edges [10]. The target specifications follow:

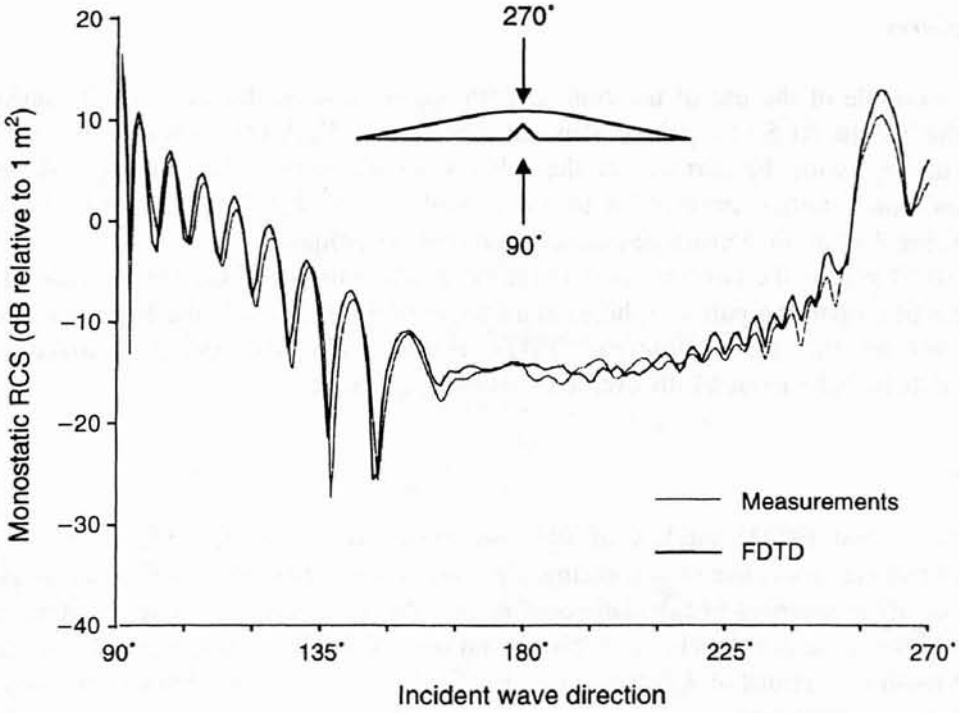
1. Length = 30.48 cm ($10\lambda_0$) in the vertical (z) direction, width = 25.4 cm ($8.47\lambda_0$);
2. Cross section shape (in the horizontal x-y plane) of an isosceles triangle with base angles = 11.31°;
3. Triangle sides opposite to the base are smoothly joined by a 15.24 cm ($5.1\lambda_0$) radius cylindrical chamfer;
4. Base has a vertical triangular slot centered in its span. Slot dimensions: 1.27 cm ($0.42\lambda_0$) depth, and 2.54 cm ($0.85\lambda_0$) width.

Because of the broadside, high-frequency, plane-wave target illumination used in this study, it was determined that full three-dimensional FDTD modeling could be avoided. Instead, computationally efficient two-dimensional TM_z and TE_z FDTD models could provide useful results. Each model used a uniform square-cell grid with a resolution $\Delta = 1.5$ mm $= \lambda_0/20$, and second-order Mur absorbing boundaries. The contour-path FDTD technique of [3] was applied.[1] Comparative data included anechoic chamber measurements of the monostatic RCS at 1° intervals in a horizontal plane sweeping around the triangular cross section of the object.

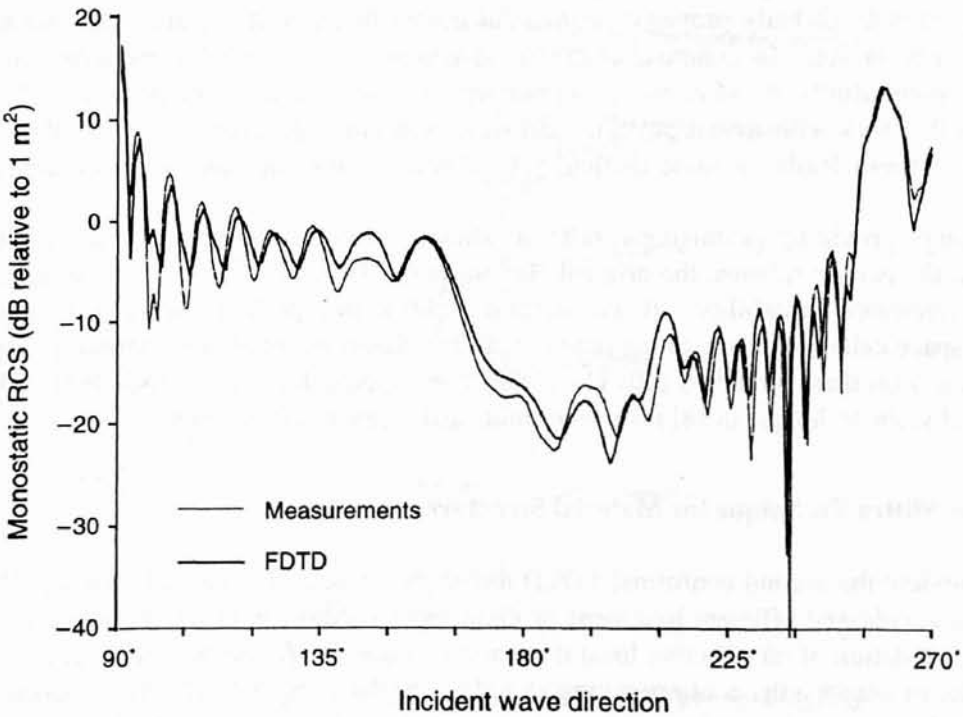
Fig. 10.9(a) graphs the measured and FDTD-calculated RCS data for the case of TM_z illumination of the target. Fig. 10.9(b) is the corresponding data comparison for TE_z illumination of the target. In these figures, an angle of 90° denotes a monostatic RCS observation broadside to the target's flat base with the triangular slot, while 270° denotes a monostatic RCS observation broadside to the target's machined chamfer. The agreement between the measurements and the contour-path FDTD calculations is seen to be very good in both cases at all look angles. Useful dynamic range is about 45 to 50 dB. When the same target geometry was modeled using a staircased FDTD approximation of its surface shape, it was found that a uniform grid resolution of $\Delta = \lambda_0/80$ or finer was required to attain approximately the same level of accuracy. This 4:1 advantage in grid resolution in each Cartesian direction is similar to that noted above in the discussion of the twisted waveguide cavity model of Fig. 10.8.

---

[1] Unlike the later, more robust methods of [6–8], stretched cells were a feature of the algorithms of [3, 4], which caused an unpredictable (geometry-dependent) late-time numerical instability. Nevertheless, in many cases, such as the models of the winglike object and the sphere-pair reviewed in this section, the algorithms of [3, 4] were able to attain accurate, converged results before the onset of instability.



(a) TM_z polarization.



(b) TE_z polarization.

Fig. 10.9 Validation of the two-dimensional contour-path FDTD model of [3] for the 10-GHz monostatic radar cross section of a winglike metal object having a V-shaped vertical slot. *Source:* Taflove, Report to General Dynamics, 1990.

Pair of Spheres

The final example of the use of the contour-path approach to model curved PEC surfaces is a study of the bistatic RCS of a pair of identical spheres [4]. Each sphere was assumed to have a diameter of $1\lambda_0$, with the surfaces of the spheres separated by a $1\lambda_0$ air gap. A cubic-cell background space lattice provided a uniform resolution of $\lambda_0/20$. Comparative data were obtained using the phasor-domain generalized multipole technique.[2]

Fig. 10.10 graphs the two sets of data for the *E*-plane bistatic RCS for the case of a plane wave impinging upon the pair of spheres at an angle of 45° relative to the line connecting their centers. We see that the contour-path FDTD results differ from the generalized multipole technique data by only about ±1 dB over a 35-dB dynamic range.

Discussion

Locally conformal FDTD models of PEC structures with curved surfaces provide clear advantages relative to the use of staircasing. In a number of disparate simulations ranging from studies of cavity resonances to calculations of monostatic and bistatic RCS, the conformal FDTD models achieve accuracy levels at mesh resolutions of $\lambda_0/20$ that staircased models might achieve at mesh resolutions of $\lambda_0/80$ or finer. Especially in three-dimensional simulations, this is an overwhelming advantage.

Locally conformal FDTD models may also provide advantages relative to the use of unstructured and partially structured space lattices. Away from the PEC surface, the locally conformal models globally propagate numerical modes in the uniform and well-characterized Cartesian Yee lattice. In contrast, unstructured-lattice models globally propagate numerical modes in nonuniform, non-Cartesian meshes which have numerical dispersion and stability properties that vary with spatial position, and from problem to problem. This variability of the background mesh leads to some difficulty in characterizing the baseline accuracy of such models.

Although providing promising results, as shown above in the examples of the winglike object and the pair of spheres, the original contour-path FDTD models of [3, 4] were subject to late-time numerical instability. It was determined that this problem arose from the use of stretched space cells and the resulting requirement for "borrows" of $\boldsymbol{E}$ components to implement Faraday's law on these stretched cells [11]. While the approach of [11] satisfactorily addressed this difficulty, the technique of [8] is more accurate and simpler to implement.

10.6.3 Yu-Mittra Technique for Material Structures

We now review the second conformal FDTD technique reported by Yu and Mittra [12], which provides a simple and efficient treatment of dielectric structures with curved surfaces. Based upon the calculation of an effective local dielectric constant for $\boldsymbol{E}$ components, this technique is much easier to use than the contour-path model of [3] or the effective-dielectric-constant method of [13]. It shows a significant improvement in accuracy relative to staircasing that is comparable to either of the two previous approaches.

---

[2]Provided by Art Ludwig of General Research Corporation.

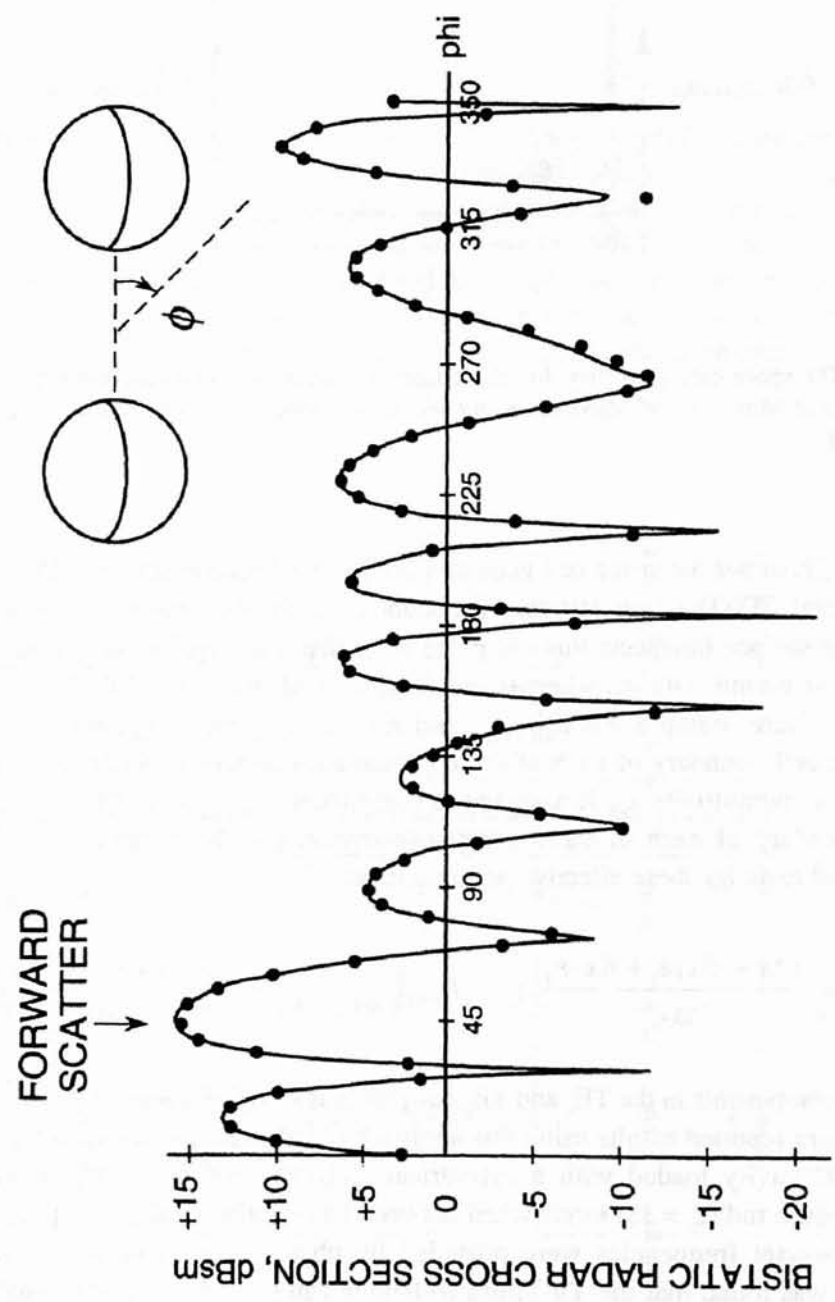


Fig. 10.10 Validation of the three-dimensional contour-path FDTD model of the bistatic radar cross section of two $1\lambda_0$-diameter PEC spheres separated by a $1\lambda_0$ air gap, oblique incidence case. Legend: solid line = FDTD model; dots = generalized multipole technique. *Source:* Jurgens and Taflove, *IEEE Trans. Antennas and Propagation*, 1993, pp. 1703–1708, © 1993 IEEE.

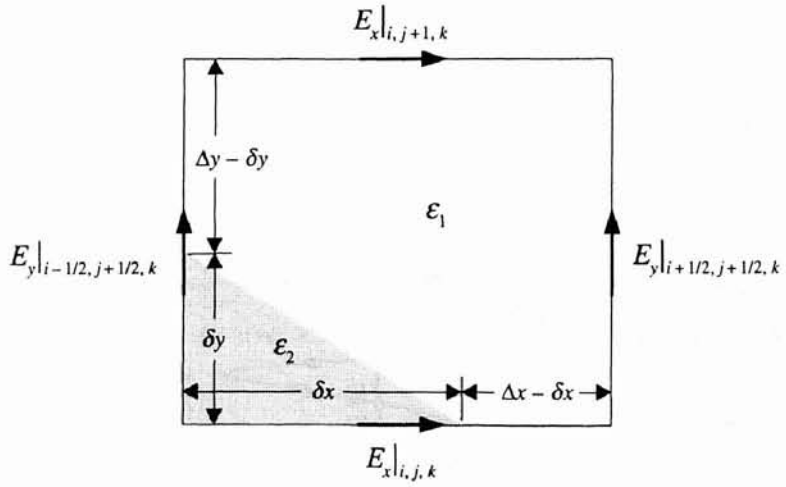


Fig. 10.11 FDTD space-cell geometry for the conformal dielectric structure model. *Adapted from:* Yu and Mittra, *IEEE Microwave and Wireless Components Lett.*, 2001, pp. 25–27, © 2001 IEEE.

Fig. 10.11 illustrates the space-cell geometry for the Yu-Mittra model in a TE_z cut-plane in a three-dimensional FDTD space lattice. Embedded within the space lattice is a dielectric structure whose surface intersects this cut plane such that a triangular part of the cell is filled with dielectric of permittivity ε_2, whereas the remainder of the cell is filled with dielectric of permittivity ε_1. Here, components $E_x|_{i,j+1,k}$ and $E_y|_{i+1/2,j+1/2,k}$ are assigned the permittivity ε_1, since the entire cell boundary of each of these components resides in medium 1. On the other hand, an effective permittivity ε_{eff} is assigned to components $E_x|_{i,j,k}$ and $E_y|_{i-1/2,j+1/2,k}$, since part of the cell boundary of each of these components resides in both media. A simple linear weighting is used to assign these effective permittivities:

$$\varepsilon_{x,\text{eff}}\big|_{i,j,k} = \frac{(\Delta x - \delta x)\varepsilon_1 + \delta x \cdot \varepsilon_2}{\Delta x} \; ; \qquad \varepsilon_{y,\text{eff}}\big|_{i-1/2,j+1/2,k} = \frac{(\Delta y - \delta y)\varepsilon_1 + \delta y \cdot \varepsilon_2}{\Delta y} \tag{10.17}$$

Extension to E-components in the TE_x and TE_y cut-planes is straightforward.

Yu and Mittra reported results using this approach to calculate the resonant frequencies of a rectangular PEC cavity loaded with a cylindrical dielectric rod ($\varepsilon_r = 38$), as well as for a cylindrical dielectric rod ($\varepsilon_r = 38$) sandwiched between two parallel PEC plates [12]. Benchmark data for the resonant frequencies were provided by phasor-domain numerical or analytical techniques. It was found that the Yu-Mittra technique can provide accuracy equal to or better than that previously reported in [13], which in turn was significantly better than that achievable using staircasing. Furthermore, the Yu-Mittra technique is simpler to implement because it does not require area or volume calculations pertaining to how space cells are cut in three dimensions by arbitrarily oriented surfaces. Hence, the mesh generation for this conformal technique is quite simple.

10.7 MALONEY-SMITH TECHNIQUE FOR THIN MATERIAL SHEETS

This section reviews the basis and illustrative results of a contour-path method to model planar material sheets of subcell thickness, where the sheet is perpendicular to one of the major axes of the FDTD space lattice. As reported by Maloney and Smith [14], this technique permits accurate treatment of both thin dielectric and conducting sheets.

10.7.1 Basis

Fig. 10.12 illustrates embedding a material sheet of thickness $d < \Delta/2$ and properties ε_s, σ_s, and $\mu_s = \mu_0$ in a three-dimensional free-space FDTD lattice. The cross section of the sheet in a lattice plane j = constant is shown. The sheet is perpendicular to the x-axis and lies between $x = i\Delta$ and $x = (i + 1/2)\Delta$. Yee meshing is used everywhere except for special cells that contain the sheet. In these cells, the E_x component (normal to the sheet) is split into two parts, $E_{x,\text{in}}$ and $E_{x,\text{out}}$, where "in" or "out" denotes that the component is evaluated either inside or outside of the sheet. This splitting permits modeling the jump discontinuity of E_x across the air-material boundary due to the change in ε and σ.

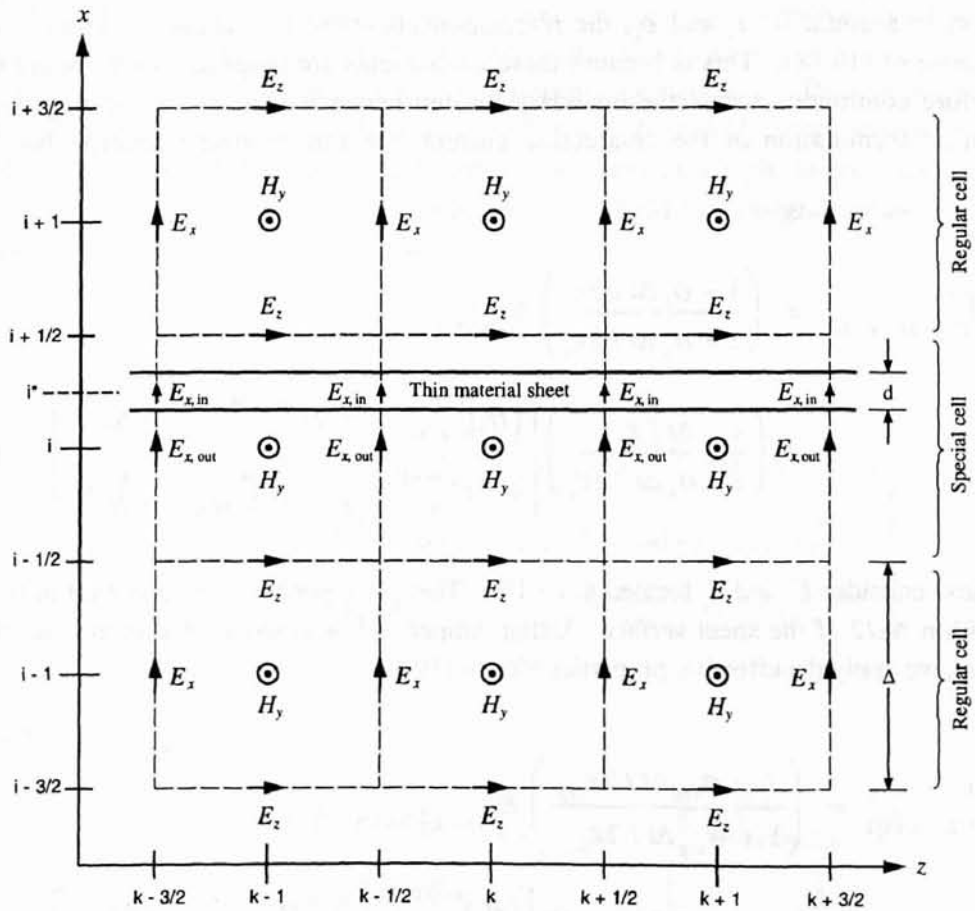


Fig. 10.12 A slice of the three-dimensional FDTD lattice at a j = constant plane, showing the locations of the field components used in the thin-sheet model. *Source:* Maloney and Smith, *IEEE Trans. Antennas and Propagation*, 1992, pp. 323–330, © 1992 IEEE.

In Fig. 10.12, we note that there is no discontinuity of the normal H_x component at the boundary of the thin material sheet, since its permeability is assumed to be that of free space. Therefore, H_x need not be split. Furthermore, all tangential components of the field (such as H_y and E_z) need not be split, since they are continuous across the material boundary.

We first consider $E_{x,\text{out}}$ and $E_{x,\text{in}}$ located, respectively, at i and i^* just outside and inside of the sheet. An Ampere's law contour can be constructed in the y-z plane parallel to the sheet to surround $E_{x,\text{out}}$. This contour lies entirely in free space. Thus, the update expression for $E_{x,\text{out}}$ is the same as for E_x in a regular free-space cell:

$$E_{x,\text{out}}\Big|^{n+1}_{i,\,j-1/2,\,k+1/2} = E_{x,\text{out}}\Big|^{n}_{i,\,j-1/2,\,k+1/2} + \frac{\Delta t}{\varepsilon_0}\left[\begin{array}{l}\left(H_y\Big|^{n+1/2}_{i,\,j-1/2,\,k} - H_y\Big|^{n+1/2}_{i,\,j-1/2,\,k+1}\right)\Big/\Delta z \\ + \left(H_z\Big|^{n+1/2}_{i,\,j,\,k+1/2} - H_z\Big|^{n+1/2}_{i,\,j-1,\,k+1/2}\right)\Big/\Delta y\end{array}\right] \tag{10.18}$$

where the coordinates are those of [14]. For $E_{x,\text{in}}$, the surrounding Ampere's law contour is located in a y-z plane entirely inside the sheet. While this location modifies the updating coefficients to account for ε_s and σ_s, the H-components circulating about the contour are the same as those of (10.18). This is because these components are tangential to the sheet surface and therefore continuous across the boundary located between i and i^*. Following the usual semi-implicit formulation of the conduction current, the time-stepping relation for $E_{x,\text{in}}$ is therefore

$$\begin{aligned} E_{x,\text{in}}\Big|^{n+1}_{i^*,\,j-1/2,\,k+1/2} &= \left(\frac{1-\sigma_s\,\Delta t/2\varepsilon_s}{1+\sigma_s\,\Delta t/2\varepsilon_s}\right) E_{x,\text{in}}\Big|^{n}_{i^*,\,j-1/2,\,k+1/2} \\ &+ \left(\frac{\Delta t/\varepsilon_s}{1+\sigma_s\,\Delta t/2\varepsilon_s}\right)\left[\begin{array}{l}\left(H_y\Big|^{n+1/2}_{i,\,j-1/2,\,k} - H_y\Big|^{n+1/2}_{i,\,j-1/2,\,k+1}\right)\Big/\Delta z \\ + \left(H_z\Big|^{n+1/2}_{i,\,j,\,k+1/2} - H_z\Big|^{n+1/2}_{i,\,j-1,\,k+1/2}\right)\Big/\Delta y\end{array}\right] \end{aligned} \tag{10.19}$$

We next consider E_y and E_z located at $i+1/2$. These components are tangential to the sheet and lie within $\Delta x/2$ of the sheet surface. Using Ampere's law contours that pierce the sheet at right angles, we apply the effective-properties idea of (10.17):

$$\begin{aligned} E_y\Big|^{n+1}_{i+1/2,\,j,\,k+1/2} &= \left(\frac{1-\sigma_{\text{avg}}\,\Delta t/2\varepsilon_{\text{avg}}}{1+\sigma_{\text{avg}}\,\Delta t/2\varepsilon_{\text{avg}}}\right) E_y\Big|^{n}_{i+1/2,\,j,\,k+1/2} \\ &+ \left(\frac{\Delta t/\varepsilon_{\text{avg}}}{1+\sigma_{\text{avg}}\,\Delta t/2\varepsilon_{\text{avg}}}\right)\left[\begin{array}{l}\left(H_x\Big|^{n+1/2}_{i+1/2,\,j,\,k+1} - H_x\Big|^{n+1/2}_{i+1/2,\,j,\,k}\right)\Big/\Delta z \\ + \left(H_z\Big|^{n+1/2}_{i,\,j,\,k+1/2} - H_z\Big|^{n+1/2}_{i+1,\,j,\,k+1/2}\right)\Big/\Delta x\end{array}\right] \end{aligned} \tag{10.20}$$

$$E_z\Big|_{i+1/2,\,j-1/2,\,k}^{n+1} = \left(\frac{1-\sigma_{\text{avg}}\Delta t/2\varepsilon_{\text{avg}}}{1+\sigma_{\text{avg}}\Delta t/2\varepsilon_{\text{avg}}}\right) E_z\Big|_{i+1/2,\,j-1/2,\,k}^{n} + \left(\frac{\Delta t/\varepsilon_{\text{avg}}}{1+\sigma_{\text{avg}}\Delta t/2\varepsilon_{\text{avg}}}\right)\left[\begin{array}{l}\left(H_y\Big|_{i+1,\,j-1/2,\,k}^{n+1/2} - H_y\Big|_{i,\,j-1/2,\,k}^{n+1/2}\right)\Big/\Delta x \\ + \left(H_x\Big|_{i+1/2,\,j-1,\,k}^{n+1/2} - H_x\Big|_{i+1/2,\,j,\,k}^{n+1/2}\right)\Big/\Delta y\end{array}\right] \tag{10.21}$$

where ε_{avg} and σ_{avg} are the average permittivity and conductivity, respectively, within the Ampere's law contour, and are given by

$$\varepsilon_{\text{avg}} \equiv \left(1-\frac{d}{\Delta x}\right)\varepsilon_0 + \frac{d}{\Delta x}\varepsilon_s\,; \qquad \sigma_{\text{avg}} \equiv \frac{d}{\Delta x}\sigma_s \tag{10.22a, b}$$

We next consider the H_x components located at $i+1/2$. These components are normal to the sheet and lie within $\Delta x/2$ of the sheet surface. A Faraday's law contour which lies entirely in free space can be constructed in the y-z plane parallel to the sheet to surround each H_x. Thus, the update expression is the same as for H_x in a regular free-space cell.

Finally, we consider the H_y and H_z components located at i. These components are tangential to the sheet and lie at the sheet surface. For each H_y and H_z component, we construct a surrounding Faraday's law contour that pierces the sheet at a right angle. Evaluating the magnetic flux integral in the enclosed patch, and the E-field line integral around the contour, yields the following time-stepping relations:

$$H_y\Big|_{i,\,j-1/2,\,k}^{n+1/2} = H_y\Big|_{i,\,j-1/2,\,k}^{n-1/2} + \frac{\Delta t}{\mu_0\,\Delta x\,\Delta z}\left[\begin{array}{l}(\Delta x-d)\left(E_{x,\text{out}}\Big|_{i,\,j-1/2,\,k+1/2}^{n} - E_{x,\text{out}}\Big|_{i,\,j-1/2,\,k-1/2}^{n}\right) \\ + d\left(E_{x,\text{in}}\Big|_{i^*,\,j-1/2,\,k+1/2}^{n} - E_{x,\text{in}}\Big|_{i^*,\,j-1/2,\,k-1/2}^{n}\right) \\ + \Delta z\left(E_z\Big|_{i-1/2,\,j-1/2,\,k}^{n} - E_z\Big|_{i+1/2,\,j-1/2,\,k}^{n}\right)\end{array}\right] \tag{10.23}$$

$$H_z\Big|_{i,\,j,\,k+1/2}^{n+1/2} = H_z\Big|_{i,\,j,\,k+1/2}^{n-1/2} + \frac{\Delta t}{\mu_0\,\Delta x\,\Delta y}\left[\begin{array}{l}(\Delta x-d)\left(E_{x,\text{out}}\Big|_{i,\,j+1/2,\,k+1/2}^{n} - E_{x,\text{out}}\Big|_{i,\,j-1/2,\,k+1/2}^{n}\right) \\ + d\left(E_{x,\text{in}}\Big|_{i^*,\,j+1/2,\,k+1/2}^{n} - E_{x,\text{in}}\Big|_{i^*,\,j-1/2,\,k+1/2}^{n}\right) \\ + \Delta y\left(E_y\Big|_{i-1/2,\,j,\,k+1/2}^{n} - E_y\Big|_{i+1/2,\,j,\,k+1/2}^{n}\right)\end{array}\right] \tag{10.24}$$

10.7.2 Illustrative Results

Maloney and Smith reported two tests of the accuracy of their subcell technique for FDTD modeling of the thin material sheet [14]:

1. Comparison with the exact solution for the attenuation and phase characteristics of a two-dimensional PEC parallel-plate waveguide having a thin lossy sheet centered within its walls;

2. Comparison with measurements of the reflected voltage from a resistively loaded monopole antenna modeled in a cylindrical, rotationally symmetric FDTD grid.

The goals and results of these tests are now summarized. We shall refer to Fig. 10.13 for the results of the waveguide modeling test.

Waveguide Modeling Test

The Maloney-Smith subcell model of the thin material sheet treats the normal and tangential components of the E-field at the sheet surface very differently. The normal component is split into two parts: $E_{x,\text{out}}$, time-stepped without direct reference to the sheet material properties ε_s and σ_s; and $E_{x,\text{in}}$, time-stepped with direct reference to ε_s and σ_s. Tangential components E_y and E_z are time-stepped using the cell-averaged material properties approach. Reference [14] reported a test of this model for separate cases, in which either the normal or tangential component of the E-field of the impinging wave would be more significant within the sheet. This would test the validity of the physics modeling for the disparate cases of ohmic losses within the sheet resulting from currents flowing either normally or tangentially within the sheet.

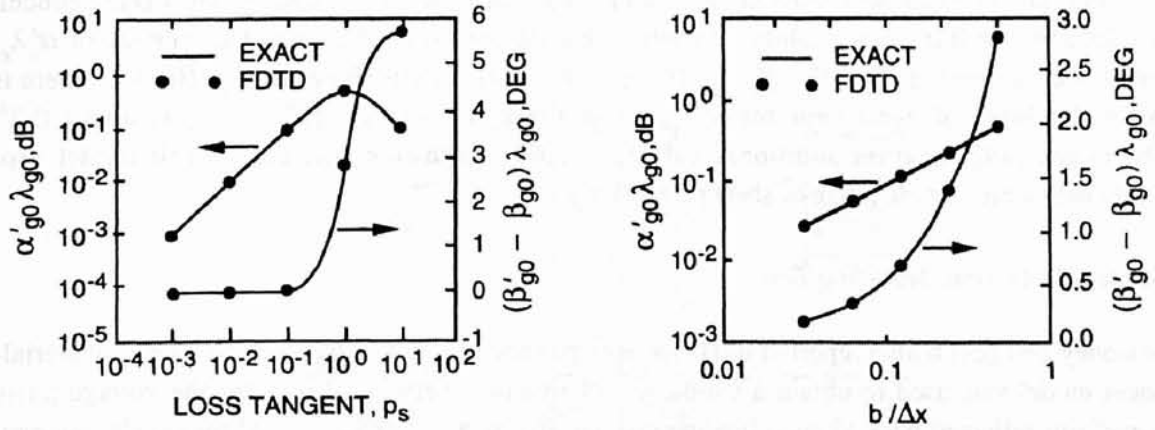
The two-dimensional PEC parallel-plate waveguide was selected as the basis for the tests. This is because it can support the propagation of both: (1) the TEM mode, having an E-field perpendicular to the guide walls, and (2) a spectrum of TM modes, each having E-field components both parallel and perpendicular to the guide walls. By modeling the loading of the waveguide with a thin material sheet parallel to and centered between the walls, and separately exciting the TEM or TM mode, the orientation of the E-field of the propagating wave relative to the thin material sheet could be precisely controlled. Further, there exists an exact solution for the propagation constant of such a loaded waveguide for both modes.

The following are the parameters of the sheet-loaded waveguide model considered in [14] and repeated here:

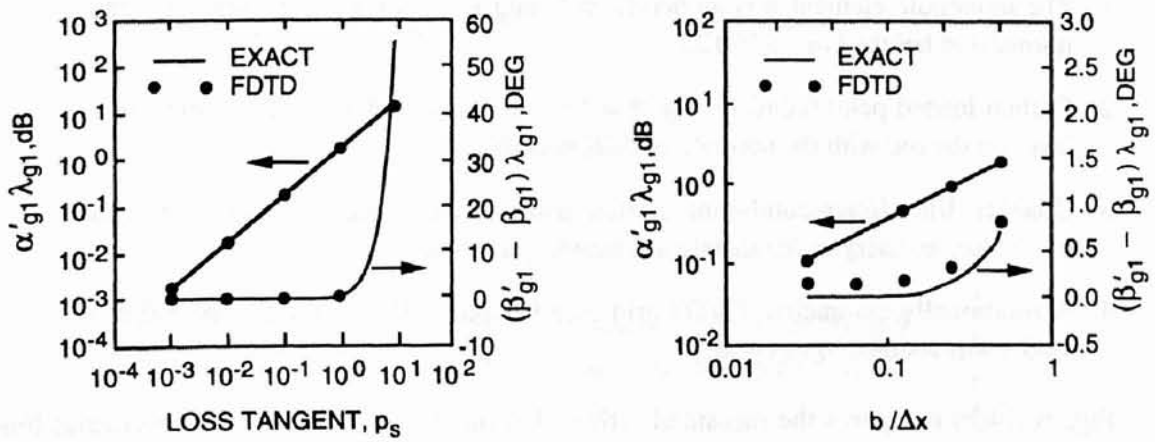
1. Waveguide plate separation = $2a = 32\Delta$ (square grid cells). TM_1 mode cutoff frequency = $\omega_c = \pi c/2a$, where c = free-space speed of light.

2. Material sheet thickness = $d = 2b$; permittivity = $\varepsilon_s = \varepsilon_0$.

3. Monochromatic excitation with the TEM mode frequency = $0.1\omega_c$; TM_1 mode frequency = $\sqrt{2}\,\omega_c$. No change in Δ for these two cases, so that the grid resolution was an ultrafine $\Delta = \lambda_0/640$ for the TEM mode, and a still very fine $\Delta = \lambda_0/45.255$ for the TM_1 mode.



(a) TEM mode, function of sheet loss tangent (left) and fractional-cell thickness (right).



(b) TM_1 mode, function of sheet loss tangent (left) and fractional-cell thickness (right).

Fig. 10.13 Validation of contour-path FDTD calculations of the normalized attenuation and phase constants for propagation in a PEC parallel-plate waveguide loaded by a subcell-thick lossy sheet. *Source:* Maloney and Smith, *IEEE Trans. Antennas and Propagation*, 1992, pp. 323–330, © 1992 IEEE.

Fig. 10.13 depicts the principal results of this study. Each of the four graphs in this figure plots $\alpha'_g\lambda_g$, the wave attenuation per guide wavelength within the sheet-loaded waveguide; and $(\beta'_g - \beta_g)\lambda_g$, the corresponding difference in the wave phase between the loaded and unloaded guide. The pair of graphs on the left show these propagation factors as a function of the sheet loss tangent p_s, with the sheet-thickness parameter fixed at $b/\Delta = 0.5$. This thickness was selected as the worst-case example, since the assumptions underlying the thin-sheet approximation have improved validity as $b/\Delta \to 0$. The pair of graphs on the right show the propagation factors as a function of the fractional-cell thickness of the sheet, with the sheet loss tangent fixed at $p_s = 1$.

For the TEM-mode results shown in Fig. 10.13(a), the predictions of the FDTD subcell model agree with the exact solution to better than 1% over several orders of magnitude of $\alpha'_g \lambda_g$, and to better than 0.1° for $(\beta'_g - \beta_g)\lambda_g$. For the TM_1-mode results shown in Fig. 10.13(b), there is a similar level of agreement for $\alpha'_g \lambda_g$, while the agreement for $(\beta'_g - \beta_g)\lambda_g$ is about 0.2°. Reference [14] provides additional validation studies showing that the subcell model also accurately treats a wide range of sheet permittivity.

Monopole Antenna Modeling Test

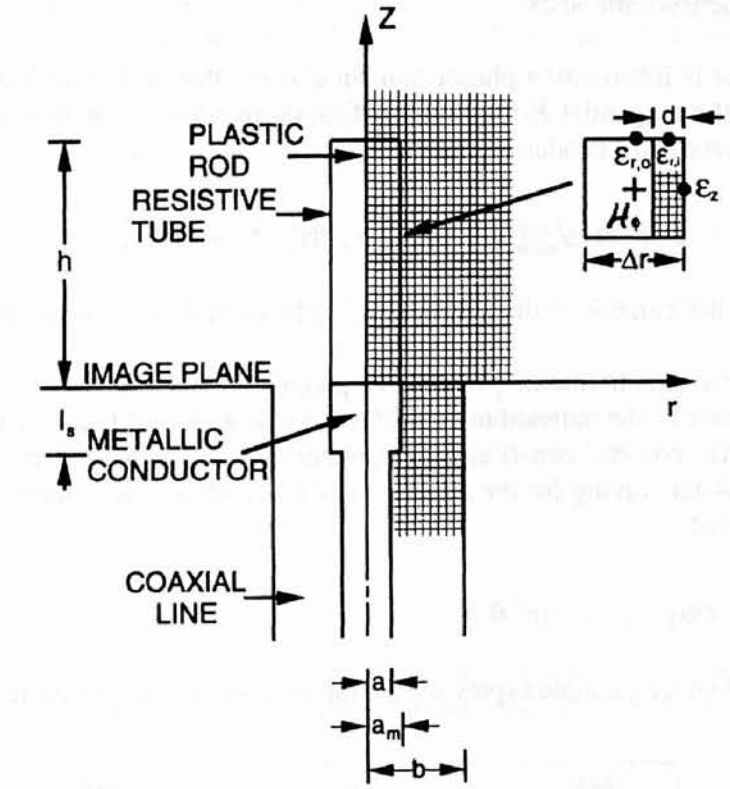
Maloney and Smith also reported in [14] a very promising result in which their subcell material-sheet model was used to obtain a validation relative to experimental data for the voltage pulse waveform reflected back along a feeding coaxial line by a resistively loaded monopole antenna. Fig. 10.14(a) depicts their antenna and feed geometry. Referring to this figure, the details of the monopole antenna system and FDTD model are as follows:

1. The monopole element was an acrylic rod with $\varepsilon_r = 2.45$, mean radius = a_m, and normalized height $h/a_m = 29.12$.
2. Carbon-loaded paint (conductivity $\sigma_s = 200$ S/m) was used to create a conducting layer on the rod with the normalized thickness $d/a_m = 0.147$.
3. Coaxial line inner-conductor radius $a/a_m = 0.893$, outer-conductor radius $b = 2.30a$, and acrylic rod metal-base length $l_s = 4.61a$.
4. A rotationally symmetric FDTD grid (see Chapter 12) was used. The subcell model was adapted to this grid.

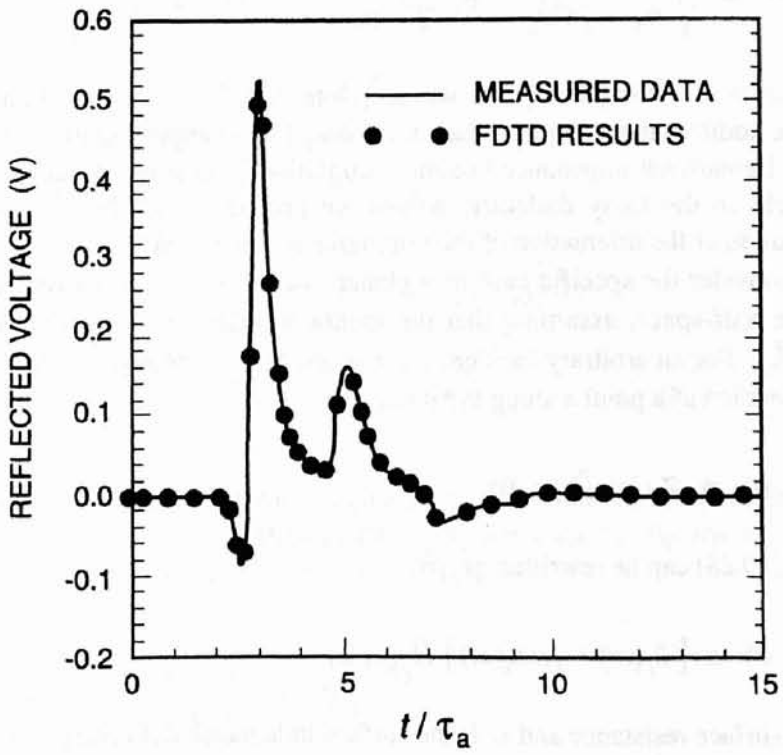
Fig. 10.14(b) compares the measured reflected voltage time waveform in the coaxial line with the FDTD modeling results. The incident Gaussian pulse for this study was assumed to have a peak amplitude of 1.0V and a normalized decay time constant equal to $0.162(h/c)$, where h/c was the characteristic time scale of the antenna. Excellent agreement between the measured and predicted results is apparent, confirming the accuracy of the thin-sheet model for this case. Comparison with earlier FDTD modeling and experimental results for a PEC monopole [15] showed a reduction in the reflections from the end of the antenna caused by the resistive loading.

10.8 SURFACE IMPEDANCE

Surface impedance boundary conditions (SIBCs) have been used in electromagnetic wave analyses since the 1940s [16–18]. In the context of FDTD modeling, SIBCs can be used to calculate the fields outside a lossy dielectric or conducting structure *without* having to model its interior. This avoids the need to resolve the decay of penetrating fields due to the skin effect. Instead, the free-space cells immediately adjacent to the surface are provided with a special time-stepping relation that conveys much of the physics of the exact surface fields and permits a large reduction in computer burdens. This section will discuss and compare several FDTD approaches to modeling SIBCs.



(a)



(b)

Fig. 10.14 Validation of contour-path FDTD calculation of reflected voltage time waveform in a coaxial line feeding a monopole antenna loaded by a thin sheet of resistive paint. *Source:* Maloney and Smith, *IEEE Trans. Antennas and Propagation*, 1992, pp. 323–330, © 1992 IEEE.

10.8.1 The Monochromatic SIBC

Surface impedance is inherently a phasor-domain concept that is founded upon the Ohm's law relation between the tangential E-field and surface-current vector phasors at a point $\boldsymbol{r}$ on the interface of free space and a conducting medium:

$$\breve{\boldsymbol{E}}_{\tan}(\boldsymbol{r}) = Z_s(\boldsymbol{r}, \omega)\, \breve{\boldsymbol{J}}_{\text{surface}}(\boldsymbol{r}) = Z_s(\boldsymbol{r}, \omega)\left[\hat{\boldsymbol{n}}(\boldsymbol{r}) \times \breve{\boldsymbol{H}}(\boldsymbol{r})\right] \tag{10.25}$$

where $Z_s(\boldsymbol{r}, \omega)$ is the complex-valued surface impedance and $\hat{\boldsymbol{n}}(\boldsymbol{r})$ is the unit surface-normal vector.

We consider a time-harmonic plane wave propagating in free space (medium number 1) at the angle θ relative to the outward normal of a lossy dielectric half-space (medium number 2). The lossy dielectric has the constitutive parameters ε_2, μ_2, and σ_2, with the loss tangent $p_2 = \sigma_2/\omega\varepsilon_2 > 1$. After solving for the reflected and transmitted components of the field subject to the assumption that

$$\left| \varepsilon_2 - j\sigma_2/\omega \right| >> \varepsilon_0 \sin^2\theta \tag{10.26}$$

is satisfied, (10.25) yields a simple expression for the monochromatic surface impedance:

$$Z_s(\omega) \cong \sqrt{\frac{j\omega\mu_2}{\sigma_2 + j\omega\varepsilon_2}} \cong \sqrt{\frac{j\omega\mu_2}{\sigma_2}} = (1+j)\sqrt{\frac{\omega\mu_2}{2\sigma_2}} \tag{10.27}$$

for any orientation of the incident plane wave. [Note that, in (10.27), the final approximation is based upon the additional assumption that $\sigma_2 >> \omega\varepsilon_2$.] The approximation of (10.27), which is designated the Leontovich impedance boundary condition [16], is equivalent to assuming that the transmitted field in the lossy dielectric half-space propagates in the direction normal to the interface regardless of the orientation of the impinging incident field.

We next consider the specific case of a planar interface at $z = 0$ between free space and the lossy dielectric half-space, assuming that the monochromatic field has the phasor components $\breve{H}_y$, $\breve{E}_x$, and $\breve{E}_z$. For an arbitrary incident wave in the free-space region, (10.25) to (10.27) yield the following relation at a point x along the interface:

$$\breve{E}_x(x, 0) \cong Z_s(\omega)\, \breve{H}_y(x, 0) \tag{10.28}$$

Using (10.27), (10.28) can be rewritten as

$$\breve{E}_x(x, 0) \cong \left[R_s(\omega) + j\omega L_s(\omega)\right] \breve{H}_y(x, 0) \tag{10.29}$$

where R_s is the surface resistance and L_s is the surface inductance defined by

$$R_s(\omega) = \sqrt{\frac{\omega\mu_2}{2\sigma_2}}\,; \qquad L_s(\omega) = \sqrt{\frac{\mu_2}{2\omega\sigma_2}} \tag{10.30}$$

In the monochromatic case, the frequency dependence of R_s and L_s can be removed by evaluating these quantities at a particular frequency and treating them as constants [19]. Equation (10.29) can then be written as

$$\breve{E}_x(x, 0) \equiv \left[R_s + j\omega L_s\right] \breve{H}_y(x, 0) \tag{10.31}$$

Inverse Fourier transformation of (10.31) yields the time-domain condition:

$$E_x(x, 0, t) \equiv R_s H_y(x, 0, t) + L_s \frac{\partial}{\partial t} H_y(x, 0, t) \tag{10.32}$$

This condition requires that the E- and H-fields be located at the same space-time point $(x, 0, t)$. In the context of the interleaved Yee mesh, this can be interpreted as [19]

$$E_x\Big|^n_{i,\,0} \equiv R_s H_y\Big|^n_{i,\,-1/2} + L_s \left.\frac{\partial H_y}{\partial t}\right|^n_{i,-1/2} \tag{10.33}$$

using the H_y component located in free space $\Delta z/2$ from the boundary as an approximation of the H_y at the boundary. Applying the contour-path method and Faraday's law to this H_y yields

$$\begin{aligned} \mu_0 \Delta x \Delta z \frac{\partial}{\partial t} H_y\Big|^n_{i,\,-1/2} &= \left(E_z\Big|^n_{i+1/2,\,-1/2} - E_z\Big|^n_{i-1/2,\,-1/2}\right)\Delta z \\ &+ \left(E_x\Big|^n_{i,\,-1} - E_x\Big|^n_{i,\,0}\right)\Delta x \end{aligned} \tag{10.34}$$

Using (10.33) to substitute for the final E-field term yields

$$\begin{aligned} \mu_0 \Delta x \Delta z \frac{\partial}{\partial t} H_y\Big|^n_{i,\,-1/2} &= \left(E_z\Big|^n_{i+1/2,\,-1/2} - E_z\Big|^n_{i-1/2,\,-1/2}\right)\Delta z \\ &+ E_x\Big|^n_{i,\,-1}\Delta x - \left(R_s H_y\Big|^n_{i,\,-1/2} + L_s \frac{\partial}{\partial t} H_y\Big|^n_{i,\,-1/2}\right)\Delta x \end{aligned} \tag{10.35}$$

We now group the time-derivative terms, approximate the time derivative using a central difference, use a semi-implicit representation of the $R_s H_y$ term, and finally solve for the updated H_y. This yields the final form of the monochromatic SIBC [19]:

$$\begin{aligned} H_y\Big|^{n+1/2}_{i,\,-1/2} &= \left(\frac{\mu_0 \Delta z + L_s - R_s \Delta t/2}{\mu_0 \Delta z + L_s + R_s \Delta t/2}\right) H_y\Big|^{n-1/2}_{i,\,-1/2} \\ &+ \frac{\Delta t}{\mu_0 \Delta z + L_s + R_s \Delta t/2}\left[\left(E_z\Big|^n_{i+1/2,\,-1/2} - E_z\Big|^n_{i-1/2,\,-1/2}\right)\cdot\left(\frac{\Delta z}{\Delta x}\right) + E_x\Big|^n_{i,\,-1}\right] \end{aligned} \tag{10.36}$$

10.8.2 Convolution-Based Models of the Frequency-Dependent SIBC

This section summarizes two FDTD formulations of the frequency-dependent SIBC based upon a time-domain convolution approach [19, 20]. These were reported in consecutive papers in the January 1992 issue of *IEEE Transactions on Antennas and Propagation.*

Method of Beggs et al.

This technique [19] is founded on the impedance boundary condition of (10.27) and (10.28). Here, however, (10.28) is cast in a slightly different form:

$$\breve{E}_x(x, 0) \equiv j\omega \left[\frac{Z_s(\omega)}{j\omega} \right] \breve{H}_y(x, 0) \tag{10.37}$$

to facilitate the subsequent transformation to the time domain. For convenience, we define the modified surface impedance Z_s' as

$$Z_s'(\omega) \equiv \frac{Z_s(\omega)}{j\omega} = \sqrt{\frac{\mu_2}{j\omega\sigma_2}}\,; \qquad Z_s'(s=j\omega) = \sqrt{\frac{\mu_2}{\sigma_2}}\,\frac{1}{\sqrt{s}} \tag{10.38}$$

This allows (10.37) to be rewritten as

$$\breve{E}_x(x, 0) \equiv Z_s'(\omega)\left[j\omega \breve{H}_y(x, 0)\right] \tag{10.39}$$

Inverse Fourier transformation of (10.39) yields the time-domain convolution

$$E_x(x, 0, t) \equiv Z_s'(t) * \left[\frac{\partial}{\partial t} H_y(x, 0, t) \right] \tag{10.40}$$

where $Z_s'(t)$ is the equivalent surface-impedance impulse response obtained by inverse Laplace transformation of (10.38):

$$Z_s'(t) = \begin{cases} 0 & t < 0 \\ \sqrt{\dfrac{\mu_2}{\pi\sigma_2 t}} & t > 0 \end{cases} \tag{10.41}$$

Similar to (10.32), (10.40) requires collocation of the E- and H-fields at the same space-time point $(x, 0, t)$. Following [19], (10.40) is interpreted as

$$E_x\big|_{i,\,0}^{n} \equiv \sqrt{\frac{\mu_2}{\pi\sigma_2 n\Delta t}} * \left.\frac{\partial H_y}{\partial t}\right|_{i,\,-1/2}^{n} \tag{10.42}$$

where the nearest-neighbor H_y component located in free space $\Delta z/2$ from the boundary is used as an approximation of the H_y at the boundary.

Now the contour-path method and Faraday's law are applied to this same H_y component, yielding again (10.34). Using (10.42) to substitute for the final E-field term in (10.34) provides

$$\mu_0 \Delta x \Delta z \frac{\partial}{\partial t} H_y\Big|_{i,\,-1/2}^{n} = \left(E_z\Big|_{i+1/2,\,-1/2}^{n} - E_z\Big|_{i-1/2,\,-1/2}^{n}\right)\Delta z + E_x\Big|_{i,\,-1}^{n}\Delta x - \sqrt{\frac{\mu_2}{\pi\sigma_2\, n\Delta t}} * \frac{\partial H_y}{\partial t}\Big|_{i,\,-1/2}^{n}\Delta x \tag{10.43}$$

The convolution in (10.43) can be expressed as a discrete summation to give

$$\mu_0 \Delta x \Delta z \frac{\partial}{\partial t} H_y\Big|_{i,\,-1/2}^{n} = \left(E_z\Big|_{i+1/2,\,-1/2}^{n} - E_z\Big|_{i-1/2,\,-1/2}^{n}\right)\Delta z + E_x\Big|_{i,\,-1}^{n}\Delta x - \sqrt{\frac{\mu_2\,\Delta t}{\pi\sigma_2}} \sum_{m=0}^{n-1} \left\{ \frac{\partial H_y\Big|_{i,\,-1/2}^{n-m}}{\partial[(n-m)\Delta t]} Z_0(m) \right\} \Delta x \tag{10.44}$$

where $Z_0(m)$ is the discrete impulse response given by [21]

$$Z_0(m) \equiv \int_{m-1/2}^{m+1/2} \frac{1}{\sqrt{\alpha}}\, d\alpha \tag{10.45}$$

In (10.45), $Z_0(m)$ is obtained by assuming that the fields are piecewise constant in time. Further, if $m = 0$, the lower limit of the integration interval is 0.

After implementing a central-difference approximation of the time derivatives in (10.44) and solving for the updated value of H_y, the following time-stepping expression is obtained:

$$H_y\Big|_{i,\,-1/2}^{n+1/2} = H_y\Big|_{i,\,-1/2}^{n-1/2} - Z_1 \sum_{m=0}^{n-1} \left[\left(H_y\Big|_{i,\,-1/2}^{n-m+1/2} - H_y\Big|_{i,\,-1/2}^{n-m-1/2} \right) Z_0(m) \right] + \frac{\Delta t}{\mu_0 \Delta z} \left[\left(E_z\Big|_{i+1/2,\,-1/2}^{n} - E_z\Big|_{i-1/2,\,-1/2}^{n} \right) \cdot \left(\frac{\Delta z}{\Delta x} \right) + E_x\Big|_{i,\,-1}^{n} \right] \tag{10.46}$$

where

$$Z_1 \equiv \frac{1}{\mu_0 \Delta z} \sqrt{\frac{\mu_2 \Delta t}{\pi \sigma_2}} \tag{10.47}$$

Equation (10.46) is the full-sum convolutional formulation of the dispersive SIBC of [19]. While suitable for computer implementation, (10.46) is so cumbersome as to be impractical for most problems, requiring storage and processing of *all* past values of the tangential H-field.

This computational difficulty can be overcome by using Prony's method (discussed in Section 15.8.1 of Chapter 15) to expand the kernel function of the convolution in a sum of exponentials. Then, we can apply the recursive-sum technique (discussed in Section 9.3 of Chapter 9) to update a sum for each exponential term. Following this strategy, the following convergent expansion can be constructed [19]:

$$Z_0(m) \cong \sum_{l=1}^{N} a_l e^{\alpha_l m} \tag{10.48}$$

This leads directly to the final form of the dispersive SIBC of [19]:

$$H_y\Big|_{i,\,-1/2}^{n+1/2} = H_y\Big|_{i,\,-1/2}^{n-1/2} - \frac{Z_1}{1 + Z_1 Z_0(0)} \sum_{l=1}^{N} \psi_l\Big|_{i,\,-1/2}^{n} + \frac{\Delta t}{\mu_0 \Delta z\left[1 + Z_1 Z_0(0)\right]}\left[\left(E_z\Big|_{i+1/2,\,-1/2}^{n} - E_z\Big|_{i-1/2,\,-1/2}^{n}\right)\cdot\left(\frac{\Delta z}{\Delta x}\right) + E_x\Big|_{i,\,-1}^{n}\right] \tag{10.49}$$

where the lth recursive sum is defined by

$$\psi_l\Big|_{i,\,-1/2}^{n} = \left(H_y\Big|_{i,\,-1/2}^{n-1/2} - H_y\Big|_{i,\,-1/2}^{n-3/2}\right) a_l e^{\alpha_l} + e^{\alpha_l} \psi_l\Big|_{i,\,-1/2}^{n-1} \tag{10.50a}$$

$$\psi_l\Big|_{i,\,-1/2}^{1} = \psi_l\Big|_{i,\,-1/2}^{0} = 0 \tag{10.50b}$$

Implementing recursive equations (10.49) and (10.50) is much more efficient than directly implementing (10.45) and (10.46). This is because the recursive formulation is local in time, requiring computer storage of only the N running sums ψ_l at each boundary point. (For the examples considered in [19], $N = 10$ provided an adequate approximation.) Further, only two multiplications and two additions per running sum are needed at each time-step. An additional feature is that the preprocessing time needed to calculate the time-domain modified impedance function of (10.41) and its exponential-series approximation of (10.48) is negligible compared with the normal FDTD requirement.

Method of Maloney and Smith

Somewhat similar to the method of [19], this convolutional technique [20] was developed independently and published concurrently. It departs directly from transformation of (10.25) into the time domain, rather than from (10.39). Thus, in place of (10.40), the following relation is used to reconstruct the surface E-field:

$$E_x(x, 0, t) \cong Z_s(t) * H_y(x, 0, t) \tag{10.51}$$

Inverse Laplace transformation of (10.27) for the case $\mu_2 = \mu_0$ provides the following equivalent surface-impedance impulse response suitable for use in (10.51):

$$Z_s(t) = \frac{\eta_0}{\sqrt{\varepsilon_{2r}}} \left\{ b e^{bt} \left[I_0(bt) + I_1(bt) \right] U(t) + \delta(t) \right\} \tag{10.52}$$

where $b \equiv -\sigma_2/2\varepsilon_2$, I_0 and I_1 are the modified Bessel functions of the first kind of order zero and one, $U(t)$ is the Heaviside unit step function, and $\delta(t)$ is the Dirac delta function. Assuming that the fields are piecewise linear in time, the convolution integral is discretized to produce the value of surface E-field that will subsequently be used in (10.34). The nearest-neighbor and time-extrapolation approximations of the H-field are applied, yielding

$$E_x\Big|_{i,\,0}^{n} \equiv \frac{\eta_0}{\sqrt{\varepsilon_{2r}}} \left(H_y\Big|_{i,\,-1/2}^{n-1/2} + \sum_{m=0}^{n-1} F(m) H_y\Big|_{i,\,-1/2}^{n-m} \right) \tag{10.53}$$

where

$$F(m) = \int_{\gamma=m-1}^{\gamma=m+1} \left(1 - |\gamma - m|\right) b\Delta t\, e^{\gamma b \Delta t} \left[I_0(\gamma b \Delta t) + I_1(\gamma b \Delta t) \right] d\gamma \tag{10.54}$$

and the lower limit for the integral of (10.54) is set to 0 if $m = 0$. Now, Prony's method is applied to approximate the function $F(m)$ with a sum of complex exponentials, by direct analogy to the approximation of $Z_0(m)$ in (10.48). This leads to the following recursive implementation of (10.53):

$$E_x\Big|_{i,\,0}^{n} \equiv \frac{\eta_0}{\sqrt{\varepsilon_{2r}}} \left(H_y\Big|_{i,\,-1/2}^{n-1/2} + \sum_{l=1}^{N} G_l\Big|_{i,\,-1/2}^{n} \right) \tag{10.55}$$

where the lth recursive sum is defined by

$$G_l\Big|_{i,\,-1/2}^{n} = a_l H_y\Big|_{i,\,-1/2}^{n-1/2} + e^{\alpha_l} G_l\Big|_{i,\,-1/2}^{n-1} \;; \qquad G_l\Big|_{i,\,-1/2}^{0} = 0 \tag{10.56a, b}$$

Validation Studies

Beggs et al. reported FDTD results in [19] testing their SIBC formulation for one- and two-dimensional problems. Their first study involved modeling the illumination of a conducting half-space by a normally incident plane wave having the time-dependence of a Gaussian pulse. The SIBC model replaced the half-space with the boundary condition of either (10.36) or (10.49) implemented on the interface between free space and the lossy medium. For this example, $\varepsilon_2 = \varepsilon_0$, $\mu_2 = \mu_0$, and $\sigma_2 = 2$ and 20 S/m. Both the magnitude and phase of the reflection coefficient were calculated and compared to the analytical solution.

The results of this study indicated reflection-coefficient errors of 1% to 4% when the temporal width of the Gaussian pulse was resolved with $256\Delta t$. Upon changing the grid space-time resolution, the errors varied as $O(\Delta z)$, consistent with the first-order-accurate rectangular-rule integration assumed for the convolution.

The second study reported in [19] involved modeling the illumination of an infinitely long, square conducting cylinder centered in a two-dimensional TM grid. The cylinder was 96 mm square with the material parameters $\varepsilon_2 = \varepsilon_0$, $\mu_2 = \mu_0$, and $\sigma_2 = 20$ S/m. It was assumed to be illuminated by a Gaussian pulse having a 10-GHz bandwidth and temporally resolved with $64\Delta t$. Two modeling cases were considered: (1) uniform gridding (including the cylinder interior) with $\Delta = 0.5$ mm $= 0.1\lambda_2$ at 10 GHz, and (2) the cylinder replaced by the SIBC of (10.49) applied at what would have been the cylinder surface, with $\Delta = 3$ mm $= 0.1\lambda_0$ at 10 GHz. Using both the normal FDTD model and the dispersive SIBC model, the scattering width was calculated versus frequency for scattering angles of 0° and 30°. The sets of data agreed to within approximately ±1 dB across the band up to 10 GHz.

Maloney and Smith [20] first studied the excitation of a lossy dielectric half-space by an infinitely long line current flowing in free space parallel to the interface at a height $z = d$. The current waveform $I(t)$ was a differentiated Gaussian pulse of unity amplitude. Using the recursive-convolution SIBC (10.55) applied at $z = 0$ to replace the half-space, FDTD simulations developed data for the tangential E-field along the interface. These data were compared with the exact solution obtained using the plane-wave spectrum approach [18]. The following parameters were selected for this example: $\varepsilon_2 = \varepsilon_0$, loss tangent $p_{2p} = 30$ at ω_p [ω_p = frequency of the peak amplitude of the spectrum $I(\omega)$], $\Delta t = \tau_p/32$ [τ_p = time from the zero-crossing to the peak of $I(t)$], and normalized height of line current = $k_p d = 0.25$ and 1.0 ($k_p = \omega_p/c$).

Fig. 10.15 compares the exact and FDTD-SIBC results of [20] for the time waveforms of the total and reflected tangential E-field at a point on the interface laterally displaced by d relative to the line current. The FDTD-SIBC data agree very well with the exact results for both heights. However, we note from [20] that this agreement diminishes if p_{2p} is reduced to 3, because of a basic limitation of the SIBC formulation *not* associated with its FDTD implementation.

The second study reported in [20] involved modeling the propagation of a monochromatic TEM mode in a parallel-plate waveguide composed of lossy walls separated by a $\lambda_0/10$ air gap. Here, the wall loss tangent p ranged from 1 to 10^5 for a relative permittivity of 1. This range was large enough to transform the wall properties from those of a lossy dielectric to those of a good conductor. Three data sets were considered: (1) the exact solution for α and β (the real and imaginary parts of the propagation constant) with no SIBC assumed, (2) the analytical solution for α and β for the case of the SIBC of (10.27) applied at the waveguide walls, and (3) the FDTD solution for α and β employing the recursive-convolution SIBC of (10.55). A square-cell grid with $\Delta = \lambda_0/80$ was employed.

Fig. 10.16 depicts the principal results of this study. This figure plots $\alpha\lambda_0$, the wave attenuation per free-space wavelength within the lossy-wall waveguide, and $(\beta - \beta_0)\lambda_0$, the corresponding difference in the wave phase between the lossy-wall waveguide and the PEC-wall waveguide. We see that the analytical SIBC and its FDTD numerical realization agree very well over the entire range of wall loss tangents. We also see that both the analytical and FDTD SIBCs correspond very well with the exact solution for $\alpha\lambda_0$ when the wall loss tangent $p > 3$. In fact, this high level of accuracy extends over orders-of-magnitude of p and $\alpha\lambda_0$. Similar accuracy is obtained for $(\beta - \beta_0)\lambda_0$ when $p > 30$. Overall, these results show that the FDTD-SIBC approach of [20] can properly model the impact of wall loss upon wave-propagation characteristics, an important engineering result.

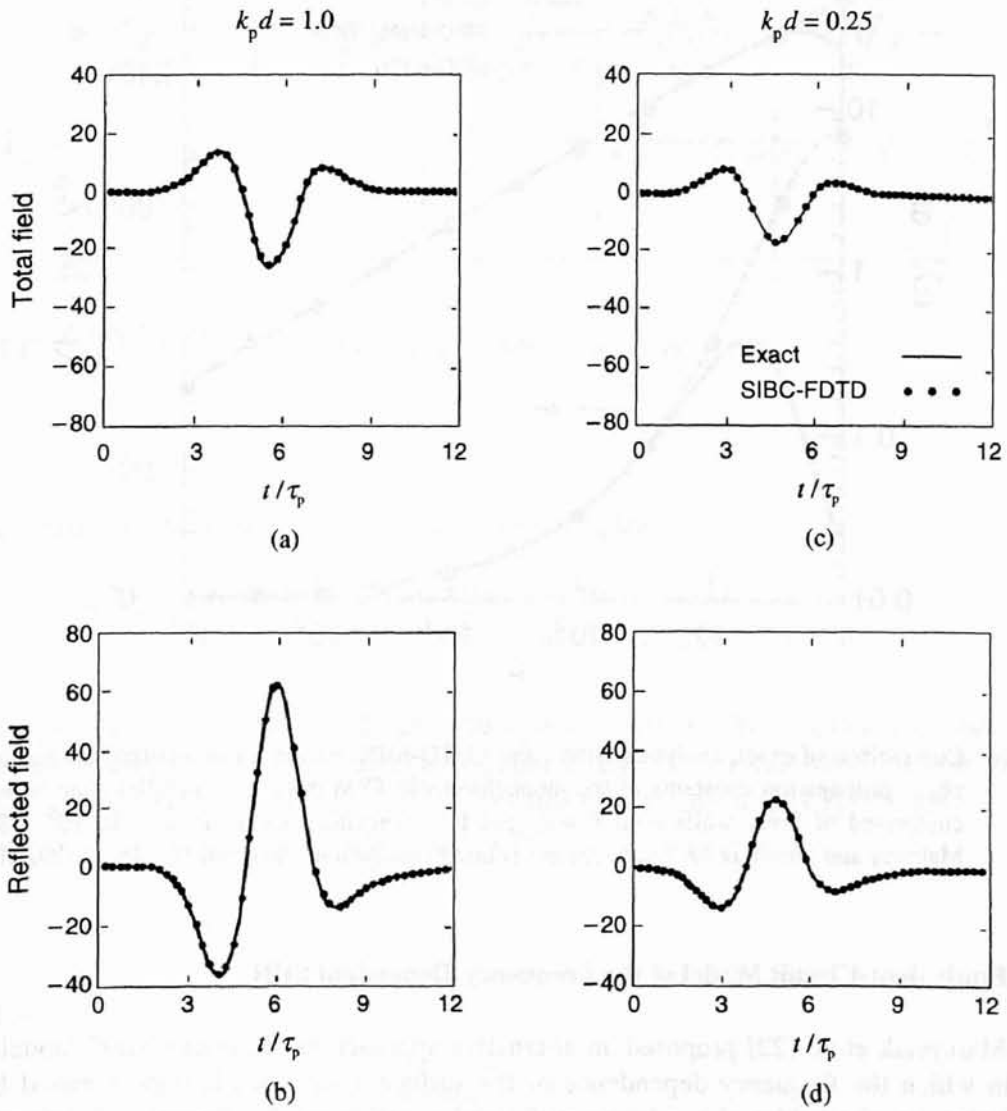


Fig. 10.15 Validation of the FDTD-SIBC model of a lossy half-space excited by an infinitely long impulsive line current flowing in free space parallel to the interface at a height $z = d$. Both the FDTD and exact time waveforms of the tangential E-field are graphed at a point on the interface at a lateral displacement of $x = d$ away from the line current. (a) Total-field time waveform, current at height $k_p d = 1.0$; (b) reflected-field time waveform, current at height $k_p d = 1.0$; (c) total-field time waveform, current at height $k_p d = 0.25$; (d) reflected-field time waveform, current at height $k_p d = 0.25$. *Source:* Maloney and Smith, *IEEE Trans. Antennas and Propagation*, 1992, pp. 38–48, © 1992 IEEE.

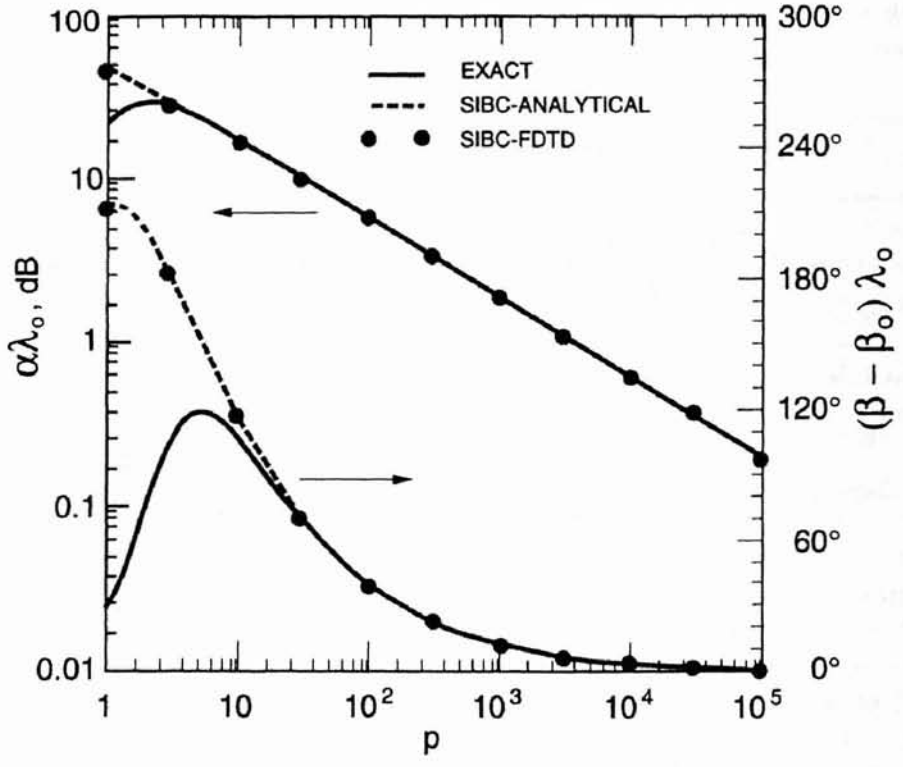


Fig. 10.16 Comparison of exact, analytical SIBC, and FDTD-SIBC results for normalized attenuation and phase propagation constants of the monochromatic TEM mode in a parallel-plate waveguide comprised of lossy walls with $\varepsilon_2 = \varepsilon_0$ and loss tangents varying from 1 to 10^5. *Source:* Maloney and Smith, *IEEE Trans. Antennas and Propagation*, 1992, pp. 38–48, © 1992 IEEE.

10.8.3 Equivalent-Circuit Model of the Frequency-Dependent SIBC

Celuch-Marcysiak et al. [22] proposed an alternative approach to dispersive SIBC modeling in FDTD in which the frequency-dependence of the surface admittance is approximated by the input admittance of a ladder of conductance (G) and capacitance (C) elements, as illustrated in Fig. 10.17. While an infinitely long ladder is theoretically needed to match the required admittance characteristic over all frequencies, in practice a ladder of only 8 to 20 GC units provides a modeling bandwidth of one to two decades in frequency.

This model was derived in the context of extending Gwarek's independently developed, two-dimensional equivalent-circuit model of Maxwell's equations [23] to implement conformal modeling in three dimensions. In [23], time-stepping of the electromagnetic field vector components was interpreted as repeated updates of voltages across capacitors and currents through inductors. In essence, the values of capacitances and inductances varied throughout the model, being proportional to the local material permittivity and permeability, but also being related to the shapes of (not necessarily rectangular) individual space cells. This same equivalent-circuit idea was subsequently shown to be useful in formulating other useful extensions to the basic FDTD method such as modeling linear magnetized ferrites, as discussed in Section 9.5 of Chapter 9. Here, it will be used to derive an efficient FDTD SIBC model.

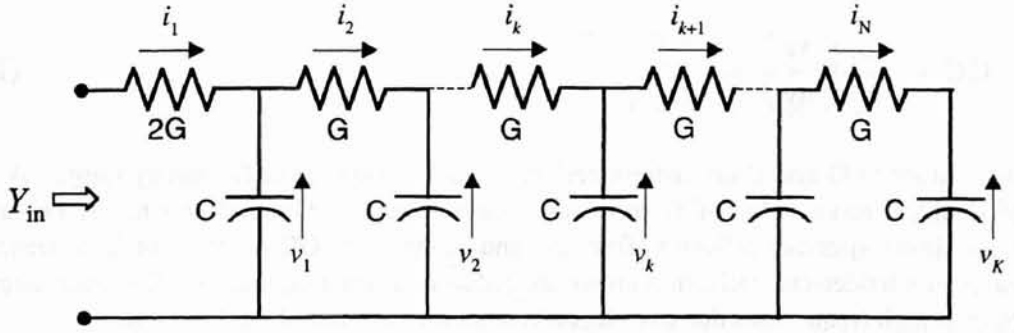


Fig. 10.17 Configuration of the equivalent-circuit SIBC model (*After:* [22].)

Basis

Consider (10.34) reformulated in terms of field-integral quantities:

$$c_y \frac{\partial}{\partial t} h_y\Big|^n_{i,\,-1/2} = e_z\Big|^n_{i+1/2,\,-1/2} - e_z\Big|^n_{i-1/2,\,-1/2} + e_x\Big|^n_{i,\,-1} - e_x\Big|^n_{i,\,0} \tag{10.57}$$

where $h_y = H_y \Delta y$, $e_z = E_z \Delta z$, $e_x = E_x \Delta x$, and $c_y = \mu_0 \Delta x \Delta z / \Delta y$. This equation is conveniently interpreted as updating voltage h_y across capacitor c_y due to four currents e flowing in or out through four branches. Current $e_x|_{i,\,0}$, corresponding to the boundary tangential E-field, is to be related to voltage $h_y|_{i,\,-1/2}$ via admittance Y:

$$e_x\Big|_{i,\,0} = Y\, h_y\Big|_{i,\,-1/2} \tag{10.58}$$

such that

$$Y(\omega) = \frac{\Delta x}{\Delta y} Z_s(\omega) = \frac{\Delta x}{\Delta y} \sqrt{\frac{j\omega\mu_2}{\sigma_2}} \tag{10.59}$$

We see that the SIBC of [22] is based on the same analytical expression for the surface impedance as (10.27) in the SIBC of [19], but scaled by the cell aspect ratio due to the integral notation being used. However, it substantially differs from this approach, as well as that of [20]. With the technique of [22], there is *no* inverse Fourier transformation of a phasor-domain impedance relation, and *no* numerical convolution. Instead, we implement a finite ladder of GC elements (as in Fig. 10.17) at each surface e_x, realizing that an infinite ladder of this type provides an input admittance having exactly the required frequency dependence:

$$Y_{in}(\omega) = \sqrt{j\omega GC} \tag{10.60}$$

Equating (10.59) and (10.60) yields the following relation for the product of G and C:

$$GC = \frac{\mu_2}{\sigma_2}\left(\frac{\Delta x}{\Delta y}\right)^2 \tag{10.61}$$

The actual values of G and C are determined by the desired operating frequency range. A small value of C and a large value of G provides accurate results at high frequencies, but at low frequencies allows spurious reflection from the end of the finite GC ladder. As C increases and G decreases, the ladder-end reflection errors are reduced at low frequencies. However, accuracy deteriorates at high frequencies due to reflections from the first unit of the GC ladder.

Numerical experiments [22] have shown that a good balance between low- and high-frequency operation of the GC ladder is obtained by: (1) providing each GC unit with $\pi/4$ phase shift at the center frequency f_0 of the desired operating band, and (2) using 2G for the first unit of the ladder (as shown in Fig. 10.17). The resulting recommended values of G and C are

$$G = \sqrt{\frac{2\pi f_0 \mu_2}{\sigma_2}}\,\frac{\Delta x}{\Delta y}\;; \qquad C = \sqrt{\frac{\mu_2}{2\pi f_0 \sigma_2}}\,\frac{\Delta x}{\Delta y} \tag{10.62a, b}$$

At each time step, the sequence of voltage and current updates along the GC ladder is given by

$$v_k^n = v_k^{n-1} + \frac{\Delta t}{C}\left(i_k^{n-1} - i_{k+1}^{n-1}\right) \qquad k = 1, 2, \ldots, K \tag{10.63a}$$

$$i_k^n = G\left(v_{k-1}^n - v_k^n\right) \qquad k = 2, 3, \ldots, K \tag{10.63b}$$

$$i_1^n = 2G\left(h_y\big|_{i,\,-1/2}^{n-1/2} - v_1^n\right) \equiv e_x\big|_{i,\,0}^n \tag{10.63c}$$

Finally, the value of $e_x\big|_{i,\,0}^n$ obtained in (10.63c) is used in (10.57) to update the boundary H-field $h_y\big|_{i,\,-1/2}$ from time-step $n-1/2$ to time-step $n+1/2$.

We note that the updates of (10.63) differ somewhat from the FDTD leapfrog scheme. This is because the GC ladder imposes a different time synchronism than the equivalent LC ladder of the FDTD space lattice. Specifically, the ladder currents in (10.63b) are correctly synchronized with the voltages across the respective capacitances. In (10.63a), the capacitor voltages are updated explicitly (with first-order accuracy in time) from the currents. There is also an explicit relation in time between the boundary H-field and the first ladder current, as in (10.63c).

Validation Study

Reference [22] reported using the equivalent-circuit SIBC technique to model TEM transmission loss in a 30-cm-long, 50Ω coaxial line (inner and outer conductor radii equal to 3.04 and 7 mm), wherein both conductors were assumed to have $\sigma = 10^5$ S/m. Fig. 10.18 compares the results for four and eight GC ladder units of Fig. 10.17 with the analytical solution from dc to 20 GHz.

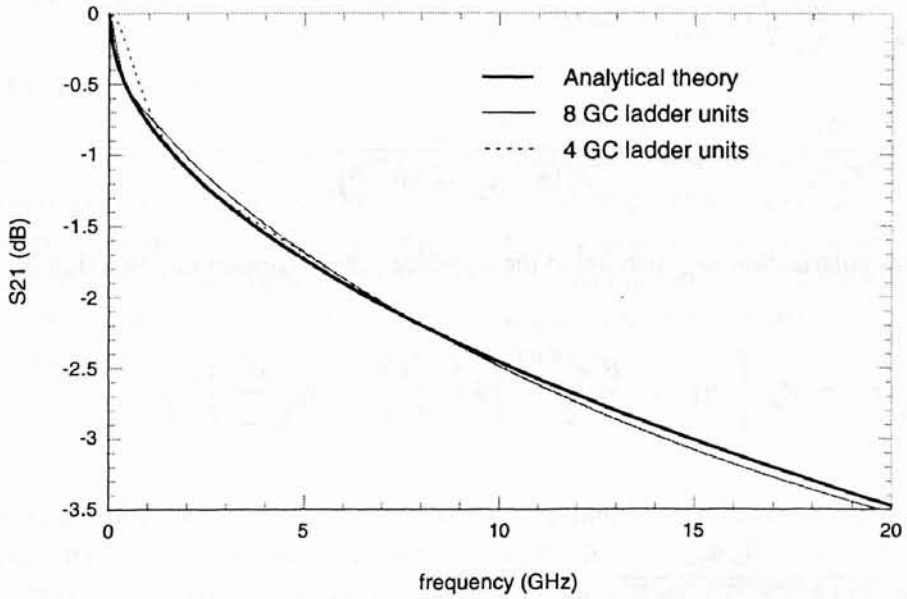


Fig. 10.18 Transmission through a 30-cm length of coaxial line (conductor $\sigma = 10^5$ S/m) calculated using four and eight GC ladder units of Fig. 10.17 to model the frequency-dependent SIBC from dc to 20 GHz (*After:* [22].)

From Fig. 10.18, we see that a ladder containing as few as four GC units gave an excellent approximation over the 10:1 frequency band 2 to 20 GHz. Extended coverage at lower frequencies well below 1 GHz required simply increasing to eight the number of GC units in the ladder at each surface tangential E-field component.

10.8.4 Sources of Error

There are two primary sources of error in the SIBC models reviewed above: (1) neglecting the incident-wave angle for media with small values of permittivity and loss tangent, and (2) using a nearest-neighbor approximation and time extrapolation of the surface $H_{\tan}$. The first error source was studied by Kellali et al. [24], who reported an extension of the convolutional SIBC method of [20] to the oblique incidence case. From [24], the modified $Z_s(t)$ function corresponding to (10.52) for a vertically polarized wave (H_{inc} parallel to the interface) impinging from free space at incidence angle θ is given by

$$Z_{s_v}(t) = Z_{v_0}\left\{\begin{array}{l} \delta(t) + \dfrac{B' e^{-B't/2}}{2}\left[I_1\left(\dfrac{B't}{2}\right) - I_0\left(\dfrac{B't}{2}\right)\right] + \\ (B' - B)\left\{e^{-Bt} + \dfrac{B'}{2}\displaystyle\int_0^t e^{-B(t-\tau)-B'\tau/2}\left[I_1\left(\dfrac{B'\tau}{2}\right) - I_0\left(\dfrac{B'\tau}{2}\right)\right] d\tau\right\} \end{array}\right\} \tag{10.64}$$

where

$$Z_{v_0} = \frac{\eta_0}{\varepsilon_{2r}} \sqrt{\varepsilon_{2r}\mu_{2r} - \sin^2\theta} \tag{10.65a}$$

$$B = \frac{\sigma_2}{\varepsilon_0 \varepsilon_{2r}} \; ; \qquad B' = \frac{\sigma_2 \mu_{2r}}{\varepsilon_0 \left(\varepsilon_{2r}\mu_{2r} - \sin^2\theta\right)} \tag{10.65b}$$

For horizontal polarization (E_{inc} parallel to the interface), the corresponding function is

$$Z_{s_h}(t) = Z_{h_0} \left\{ \delta(t) + \frac{B' e^{-B't/2}}{2} \left[I_1\left(\frac{B't}{2}\right) - I_0\left(\frac{B't}{2}\right) \right] \right\} \tag{10.66}$$

where

$$Z_{h_0} = \frac{\eta_0 \mu_{2r}}{\sqrt{\varepsilon_{2r}\mu_{2r} - \sin^2\theta}} \tag{10.67}$$

The remainder of the development proceeds in the same manner as (10.53) to (10.56). We note that these refinements complicate the implementation of the convolutional SIBC model of [20], but markedly extend the range of its applications to low-permittivity half-spaces.

The second error source arises from the fact that the SIBC models reviewed here require knowledge of H_{tan} at the same space-time location on the structure surface as that of E_{tan}. This conflicts with the spatial and temporal interleaving of $\boldsymbol{E}$ and $\boldsymbol{H}$ that forms the basis of the Yee algorithm. For the techniques reviewed in this section, the required H_{tan} at the surface is approximated by the parallel H-component located one-half space cell away from the surface in free space, and one-half time-step earlier in time. This causes the time-domain SIBC methods to have only a first-order convergence. Roden and Gedney [21] proposed two approaches to improve temporal convergence. An improved spatial extrapolation of the near-surface H-field is also conceptually feasible. In fact, for the equivalent-circuit technique, this can be readily accomplished by adjusting the parameters of the first GC unit.

10.8.5 Discussion

We can categorize FDTD models of the frequency-dependent SIBC into two basic types: convolution-based techniques and equivalent-circuit-based approaches. The first type involves a recursive convolution of the surface tangential H-field with an equivalent impedance temporal impulse response obtained by preprocessing the kernel function into a sum of exponentials. The second type solves for the currents and voltages along a lumped-circuit GC ladder, which yields a driving-point temporal response analogous to that exhibited by the tangential E- and H-fields at the surface of a conductor.

The equivalent-circuit approach has two potential advantages: (1) robust and predictable stability and convergence for the high conductivities σ of typical metals, and (2) no signal preprocessing is needed. While computer memory must be provided for the voltages and currents associated with the two to eight units of each GC ladder, the convolution-based approaches also require storage for each of the recursive sums, which may number 10 or more, depending upon σ.

10.9 THIN COATINGS ON A PEC SURFACE

10.9.1 Method of Lee at al.

In Chapter 9, the auxiliary-differential-equation technique was presented as an alternative to the recursive-convolution method for FDTD modeling of dispersive materials. This section summarizes the basis of such a technique reported by Lee et al. [25] for modeling a thin, lossy dielectric coating on a PEC surface. For consistency with the previous discussion, the same notation for fields and material properties is maintained. This notation differs somewhat from that presented in [25].

Basis

Consider a planar PEC surface located at $z = 0$ having a thin, lossy material coating of thickness h. The coating has the constitutive parameters $\varepsilon_2 = \varepsilon_{2r}\varepsilon_0$, $\mu_2 = \mu_{2r}\mu_0$, and $\sigma_2 \neq 0$. Subject to the assumptions involved in (10.27), the tangential E-field phasor at the free-space / coating interface can be related to its normal derivative by

$$\breve{E}_x(x, -h) = \frac{\eta_{2r}}{k_0} \tan\left(k_0 k_{2r} h\right) \frac{\partial \breve{E}_x}{\partial z}(x, -h) \tag{10.68}$$

where k_0 is the free-space wavenumber, and

$$\eta_{2r} = \sqrt{\frac{\mu_{2r}}{\varepsilon_{2r} - j\sigma_2/\omega\varepsilon_0}}\ ; \qquad k_{2r} = \sqrt{\mu_{2r}\left(\varepsilon_{2r} - j\sigma_2/\omega\varepsilon_0\right)} \tag{10.69a, b}$$

are the coating's relative wave impedance and relative wavenumber, respectively.

Equation (10.68) can be rewritten as:

$$\breve{E}_x(x, -h) = \frac{\eta_{2r}}{k_0} \frac{\sin\left(k_0 k_{2r} h\right)}{\cos\left(k_0 k_{2r} h\right)} \frac{\partial \breve{E}_x}{\partial z}(x, -h) \tag{10.70}$$

Using Taylor's series expansions of the sine and cosine functions, (10.70) can be expressed as the ratio of two power series in the argument $k_0 k_{2r} h$:

$$\breve{E}_x(x, -h) = \frac{\eta_{2r}}{k_0} \left[\frac{(k_0 k_{2r} h) - \dfrac{(k_0 k_{2r} h)^3}{6} + \dfrac{(k_0 k_{2r} h)^5}{120} - \cdots}{1 - \dfrac{(k_0 k_{2r} h)^2}{2} + \dfrac{(k_0 k_{2r} h)^4}{24} - \cdots} \right] \frac{\partial \breve{E}_x}{\partial z}(x, -h) \tag{10.71}$$

Now, cross-multiplying (10.71) yields

$$\left(1 - \frac{k_0^2 k_{2r}^2 h^2}{2} + \frac{k_0^4 k_{2r}^4 h^4}{24} - \ldots\right) \breve{E}_x(x, -h)$$

$$= \eta_{2r} k_{2r} h \left(1 - \frac{k_0^2 k_{2r}^2 h^2}{6} + \frac{k_0^4 k_{2r}^4 h^4}{120} - \ldots\right) \frac{\partial \breve{E}_x}{\partial z}(x, -h) \qquad (10.72)$$

Retaining the first three terms of each expansion and substituting $k_0 = \omega/c$ and k_{2r} of (10.69b) into (10.72), we obtain the final phasor-domain relation between the tangential E-field and its normal derivative:

$$\left[1 + \frac{h^2 \mu_{2r}}{2c^2}\left(\frac{j\omega\sigma_2}{\varepsilon_0} - \omega^2 \varepsilon_{2r}\right) + \frac{h^4 \mu_{2r}^2}{24 c^4}\left(-\frac{\omega^2 \sigma_2^2}{\varepsilon_0^2} - \frac{2j\omega^3 \varepsilon_{2r}\sigma_2}{\varepsilon_0} + \omega^4 \varepsilon_{2r}^2\right)\right] \breve{E}_x(x, -h)$$

$$= \mu_{2r} h \left[1 + \frac{h^2 \mu_{2r}}{6c^2}\left(\frac{j\omega\sigma_2}{\varepsilon_0} - \omega^2 \varepsilon_{2r}\right) + \frac{h^4 \mu_{2r}^2}{120 c^4}\left(-\frac{\omega^2 \sigma_2^2}{\varepsilon_0^2} - \frac{2j\omega^3 \varepsilon_{2r}\sigma_2}{\varepsilon_0} + \omega^4 \varepsilon_{2r}^2\right)\right] \frac{\partial \breve{E}_x}{\partial z}(x, -h) \qquad (10.73)$$

At this point, the inverse Fourier transform is applied term by term to (10.73) to provide the corresponding time-domain differential equation. The following transform pairs make this possible by inspection:

$$\frac{\partial}{\partial t} \Leftrightarrow j\omega\,; \qquad \frac{\partial^2}{\partial t^2} \Leftrightarrow -\omega^2 \qquad (10.74a, b)$$

$$\frac{\partial^3}{\partial t^3} \Leftrightarrow -j\omega^3; \qquad \frac{\partial^4}{\partial t^4} \Leftrightarrow \omega^4 \qquad (10.74c, d)$$

Applying the Fourier transform pairs of (10.74) to (10.73) and retaining terms only through $O(h^2)$ yields the following first-order time-domain SIBC:

$$\left[1+\left(\frac{h^2\mu_{2r}\sigma_2}{2\varepsilon_0 c^2}\right)\frac{\partial}{\partial t}+\left(\frac{h^2\mu_{2r}\varepsilon_{2r}}{2c^2}\right)\frac{\partial^2}{\partial t^2}\right]E_x(x,\,-h,\,t)$$

$$=\mu_{2r}h\left[1+\left(\frac{h^2\mu_{2r}\sigma_2}{6\varepsilon_0 c^2}\right)\frac{\partial}{\partial t}+\left(\frac{h^2\mu_{2r}\varepsilon_{2r}}{6c^2}\right)\frac{\partial^2}{\partial t^2}\right]\frac{\partial E_x}{\partial z}(x,\,-h,\,t) \qquad (10.75)$$

Retaining terms through $O(h^4)$ yields the following second-order time-domain SIBC:

$$\left[\begin{array}{l}1+\left(\dfrac{h^2\mu_{2r}\sigma_2}{2\varepsilon_0 c^2}\right)\dfrac{\partial}{\partial t}+\left(\dfrac{h^2\mu_{2r}\varepsilon_{2r}}{2c^2}+\dfrac{h^4\mu_{2r}^2\sigma_2^2}{24\varepsilon_0^2 c^4}\right)\dfrac{\partial^2}{\partial t^2}\\ +\left(\dfrac{h^4\mu_{2r}^2\varepsilon_{2r}\sigma_2}{12\varepsilon_0 c^4}\right)\dfrac{\partial^3}{\partial t^3}+\left(\dfrac{h^4\mu_{2r}^2\varepsilon_{2r}^2}{24c^4}\right)\dfrac{\partial^4}{\partial t^4}\end{array}\right]E_x(x,\,-h,\,t)$$

$$=\mu_{2r}h\left[\begin{array}{l}1+\left(\dfrac{h^2\mu_{2r}\sigma_2}{6\varepsilon_0 c^2}\right)\dfrac{\partial}{\partial t}+\left(\dfrac{h^2\mu_{2r}\varepsilon_{2r}}{6c^2}+\dfrac{h^4\mu_{2r}^2\sigma_2^2}{120\varepsilon_0^2 c^4}\right)\dfrac{\partial^2}{\partial t^2}\\ +\left(\dfrac{h^4\mu_{2r}^2\varepsilon_{2r}\sigma_2}{60\varepsilon_0 c^4}\right)\dfrac{\partial^3}{\partial t^3}+\left(\dfrac{h^4\mu_{2r}^2\varepsilon_{2r}^2}{120c^4}\right)\dfrac{\partial^4}{\partial t^4}\end{array}\right]\frac{\partial E_x}{\partial z}(x,\,-h,\,t) \qquad (10.76)$$

The SIBC of (10.75) or (10.76) can now be applied at the interface of free-space and the thin coating, thereby eliminating the need to employ a fine grid to model the coating or the underlying PEC structure. To obtain an expression for the required normal derivative of the E-field at the point of application of the SIBC, a quadratic interpolating polynomial was developed in [25] using the following three unevenly spaced datapoints:

1. $E_x = 0$ at the surface of the PEC ($z = 0$);
2. Finite E_x at the free-space / thin-film interface ($z = -h$), computed by applying (10.75) or (10.76);
3. Finite E_x in free space one cell away from the free-space / thin-film interface ($z = -h - \Delta z$), computed via normal Yee time-stepping.

Differentiating the interpolating polynomial with respect to z and evaluating at $z = -h$, we obtain the following relation for the normal derivative:

$$\left.\frac{\partial E_x}{\partial z}\right|_{x,\,-h} = -\left[\frac{h_{\text{eff}}}{(h_{\text{eff}} + \Delta z)\,\Delta z}\right] E_x|_{x,\,-h-\Delta z} - \left(\frac{\Delta z - h_{\text{eff}}}{h_{\text{eff}}\,\Delta z}\right) E_x|_{x,\,-h} \tag{10.77}$$

where the effective coating thickness $h_{\text{eff}} = h\sqrt{\varepsilon_{2r}}$ accounts for the slowing of the wave velocity in the coating due to its dielectric constant. Central-differencing can be used to implement the time derivatives in (10.75) and (10.76).

Discussion

For typical grid resolutions, the second-order SIBC of (10.76) was shown in a sample computation in [25] to provide less than 1% error from 0 to 10 GHz for the magnitude of the reflection coefficient of a coated PEC. For such problems, this approach appears to be competitive in accuracy with recursive-convolution methods, and has the advantage of not requiring the Prony's method preprocessing step.

10.9.2 Method of Kärkkäinen

Kärkkäinen [26] reported a subcell FDTD technique for modeling electrically thin dispersive layers, including coatings on a PEC surface. The layers and coatings can have a permittivity and permeability characterized in the frequency domain by multiple pole-pairs, which are treated in the time domain using an auxiliary differential equation technique. This model reduces to that reported by Maloney and Smith [14] for the case of dielectric and conductive layers. Since the Maloney-Smith model was discussed in detail earlier in Section 10.7, the reader is referred to [26] for the required modification of the field updates to accommodate the material dispersion.

10.10 RELATIVISTIC MOTION OF PEC BOUNDARIES

The engineering motivation for studying electromagnetic wave scattering by relativistically moving bodies has been provided by work in the generation of millimeter and submillimeter waves via the interaction of microwaves with rapidly moving plasma fronts or electron beams. Classical theory in this area models the physics of a moving reflecting surface by employing system transformations where the surface is at rest. This approach has been used to solve canonical moving-body problems, such as planar conducting and dielectric interfaces in uniform translation or vibration, uniformly moving or vibrating cylinders and spheres, and simple rotating shapes. However, difficulties arise for general scatterers, since closed-form solutions cannot be obtained when the scatterer shape, composition, translation, and surface vibration are arbitrary. Solution of such general problems is important in the context of understanding microwave beam interactions with moving or oscillating charged particle beams of finite cross section.

Although appearing at first glance to have a very different basis, the theory of scattering by relativistically moving PEC surfaces can be cast into the SIBC framework discussed in Section 10.8. As reported by Harfoush et al. [27], the key is to stay in the rest frame and apply the relativistic surface boundary condition of [28]. Since this boundary condition is instantaneous in time and therefore requires no convolution, it can be implemented more easily than those considered in Section 10.8.2. This section will discuss the basis and illustrative results of Harfoush's approach of [27].

10.10.1 Basis

Consider an electromagnetic wave in free space (medium 1) impinging upon a uniformly moving material half-space (medium 2). Let $\boldsymbol{v}$ and $\hat{\boldsymbol{n}}$ denote the velocity and the unit normal vector of the moving interface, respectively. Then, the following boundary conditions on the total electromagnetic field [28] must be satisfied at the interface for arbitrary velocities, $0 \le v < c$:

$$\hat{\boldsymbol{n}} \times \left(\boldsymbol{E}_2 - \boldsymbol{E}_1\right) = \left(\hat{\boldsymbol{n}} \cdot \boldsymbol{v}\right)\left(\boldsymbol{B}_2 - \boldsymbol{B}_1\right) \tag{10.78}$$

$$\hat{\boldsymbol{n}} \cdot \left(\boldsymbol{D}_2 - \boldsymbol{D}_1\right) = \rho_s \tag{10.79}$$

$$\hat{\boldsymbol{n}} \times \left(\boldsymbol{H}_2 - \boldsymbol{H}_1\right) = \boldsymbol{J}_s - \left(\hat{\boldsymbol{n}} \cdot \boldsymbol{v}\right)\left(\boldsymbol{D}_2 - \boldsymbol{D}_1\right) \tag{10.80}$$

$$\hat{\boldsymbol{n}} \cdot \left(\boldsymbol{B}_2 - \boldsymbol{B}_1\right) = 0 \tag{10.81}$$

where $\boldsymbol{E}_m$, $\boldsymbol{D}_m$, $\boldsymbol{H}_m$, and $\boldsymbol{B}_m$ are, respectively, the electric field, electric flux density, magnetic field, and magnetic flux density in media $m = 1, 2$ at the interface; ρ_s is the surface electric charge density; and $\boldsymbol{J}_s$ is the surface electric current density.

We note that (10.79) and (10.81) are unchanged from the ordinary zero-velocity boundary conditions that enforce continuity of the normally directed $\boldsymbol{D}$ and $\boldsymbol{B}$ across an interface. Further, (10.78) and (10.80) reduce to the ordinary enforcement of continuity of the tangential $\boldsymbol{E}$ and $\boldsymbol{H}$ across an interface if either $v = 0$ or $\hat{\boldsymbol{n}} \cdot \boldsymbol{v} = 0$. The latter implies that any motion of the half-space perpendicular to the surface normal (i.e., in the plane of the interface) has no impact upon the boundary condition. Only motion perpendicular to the plane of the interface has an effect. Without any loss of generality, we will assume in the subsequent development that the half-space has a zero-velocity component in the plane of the interface so that $\boldsymbol{v} = v_n \hat{\boldsymbol{n}}$.

Now, consider the application of (10.78) to the case of the PEC half-space. Here $\boldsymbol{E}_2 = 0$ and $\boldsymbol{B}_2 = 0$, and (10.78) can be rewritten as

$$\hat{\boldsymbol{n}} \times \boldsymbol{E}_1 = \left(\hat{\boldsymbol{n}} \cdot \boldsymbol{v}\right)\boldsymbol{B}_1 \tag{10.82}$$

Using the vector identity $\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = (\boldsymbol{A} \cdot \boldsymbol{C})\boldsymbol{B} - (\boldsymbol{A} \cdot \boldsymbol{B})\boldsymbol{C}$, we can expand the right-hand side of (10.82) to obtain

$$\hat{\boldsymbol{n}} \times \boldsymbol{E}_1 = \hat{\boldsymbol{n}} \times \left(\boldsymbol{B}_1 \times \boldsymbol{v}\right) + \left(\hat{\boldsymbol{n}} \cdot \boldsymbol{B}_1\right)\boldsymbol{v} \tag{10.83}$$

Now applying the cross product $-\hat{\boldsymbol{n}} \times (\ldots)$ to both sides of (10.83) yields

$$\boldsymbol{E}_1 = \boldsymbol{B}_1 \times \boldsymbol{v} - \hat{\boldsymbol{n}} \times \left(\hat{\boldsymbol{n}} \cdot \boldsymbol{B}_1\right)\boldsymbol{v} \tag{10.84}$$

However, $\hat{\boldsymbol{n}} \times \boldsymbol{v} = 0$ by the assumption that $\boldsymbol{v} = v_n \hat{\boldsymbol{n}}$, so that we have the following E-field boundary condition at the surface of the moving PEC half-space:

$$\boldsymbol{E}_1 = -\boldsymbol{v} \times \boldsymbol{B}_1 \tag{10.85}$$

Inspection of (10.85) reveals that unfamiliar physics results for $v > 0$. Namely, the tangential $\boldsymbol{E}$ at the surface of a moving PEC boundary can be *finite* rather than the required zero when the boundary is at rest. However, this does not result in an infinite surface current density, because the usual Ohm's law expression $\boldsymbol{J} = \sigma \boldsymbol{E}$ is no longer valid. Instead, for a uniformly moving conducting surface, the total induced current is the result of the conduction current plus an extra Lorentz force term. Defining $\beta \equiv v/c$, the total current is given by

$$\boldsymbol{J} = \frac{\sigma(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B})}{\sqrt{1 - \beta^2}} \tag{10.86}$$

where, for a PEC, $\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B} = 0$ by (10.85), and therefore $\boldsymbol{J}_s$ remains finite. In many cases, only small velocities are considered, and the term β^2 is negligible compared to 1.

For a moving PEC surface, relativistic boundary condition (10.85) linearly relates the local values of the instantaneous total tangential electric and magnetic fields at the surface. Because this relation is similar in form to the SIBCs discussed earlier in this chapter, it presents a similar problem: how to implement a space-time-coincident boundary condition in the context of the FDTD algorithm that does *not* calculate its E- and H-fields at the same points in space and time. The nearest-neighbor approximation of the surface H-field, used successfully in [19] and [20], is useless here because of the potentially significant positional change of the PEC surface during each $0.5\Delta t$ due to its relativistic velocity. Here, it is mandatory to derive an equivalent form of (10.85) that is consistent with the temporal interleaving of the field values in FDTD.

The derivation of the required relativistic boundary condition was reported in [27], and is summarized below. For consistency with earlier discussions in the chapter, the derivation assumes a $+x$-polarized incident plane wave in free space propagating in the $+z$-direction. This wave impinges upon a PEC half-space located at $z \geq 0$ that is in uniform motion with velocity $\pm v \hat{z}$. Therefore, the incident E- and B-fields in the free-space region $z < 0$ are given by

$$E_{x,\text{inc}}(z, t) = f(z - ct) \; ; \qquad B_{y,\text{inc}}(z, t) = \frac{1}{c} E_{x,\text{inc}}(z, t) \tag{10.87a, b}$$

where f is an arbitrary function. The reflected fields have the form

$$E_{x,\text{ref}}(z, t) = R f(z + ct) \; ; \qquad B_{y,\text{ref}}(z, t) = -\frac{1}{c} E_{x,\text{ref}}(z, t) \tag{10.88a, b}$$

Summing the incident and reflected components, the total B-field in the free-space region is

$$B_{y,\text{tot}}(z, t) = B_{y,\text{inc}} + B_{y,\text{ref}} = \frac{1}{c}\left(E_{x,\text{inc}} - E_{x,\text{ref}}\right) \tag{10.89}$$

but since $E_{x,\text{ref}} = E_{x,\text{tot}} - E_{x,\text{inc}}$, we have

$$B_{y,\text{tot}}(z, t) = \frac{1}{c}\left(2 E_{x,\text{inc}} - E_{x,\text{tot}}\right) \tag{10.90}$$

For the assumed velocity of the PEC half-space, the relativistic boundary condition obtained from (10.85) is given by

$$E_{x,\text{tot}} = \pm v B_{y,\text{tot}} = v B_{y,\text{tot}} \tag{10.91}$$

where it is clear that $E_{x,\text{tot}}$ takes the same sign as the velocity of the half-space. After substituting $B_{y,\text{tot}}$ from (10.90), (10.91) yields

$$E_{x,\text{tot}} = \frac{v}{c}\left(2E_{x,\text{inc}} - E_{x,\text{tot}}\right) \tag{10.92}$$

Upon collecting terms, we obtain the final expression for $E_{x,\text{tot}}$ at the moving boundary in terms of the incident field:

$$E_{x,\text{tot}} = \left(\frac{2v/c}{1+v/c}\right)E_{x,\text{inc}} = \left(\frac{2\beta}{1+\beta}\right)E_{x,\text{inc}} \tag{10.93}$$

Following an analogous procedure, we can also obtain the expression for $B_{y,\text{tot}}$ at the moving boundary in terms of the incident field:

$$B_{y,\text{tot}} = \left(\frac{2c/v}{1+c/v}\right)B_{y,\text{inc}} = \left(\frac{2}{1+\beta}\right)B_{y,\text{inc}} \tag{10.94}$$

Further, both of these relations can be generalized for the case where the half-space velocity $\boldsymbol{v}$ is arbitrarily oriented in the x-z plane by defining

$$\beta \equiv -\frac{1}{c}\hat{\boldsymbol{n}} \cdot \boldsymbol{v} \tag{10.95}$$

The negative sign in this expression results from the unit outward normal at the surface of the half-space pointing in the $-z$-direction.

Using (10.93) to (10.95), the values of the total tangential fields at any point on the moving PEC boundary, at any point in time, are given solely in terms of the known incident fields at that *same* point in space and time. This is much simpler than the convolutional relation used in Section 10.8.2 for general dispersive media. Assuming that the scatterer moves along a grid axis as in [27], the FDTD implementation of the relativistic boundary condition now becomes straightforward. For example, assume that the FDTD algorithm is ready to update the E-fields in the grid at a particular time-step. In sequence, the algorithm must:

1. Determine the instantaneous position of the PEC boundary in the FDTD grid from its assumed uniform motion;
2. Perform normal Yee updates for all $\boldsymbol{E}$ components in free space, except those at locations less than $\Delta/2$ from the boundary;

3. Apply (10.93) to calculate the tangential E-fields at the intersection of the PEC boundary and each grid line along which E-components are located;

4. Use linear interpolation for each remaining E-component by averaging the value at the PEC boundary computed in step 3 with the adjacent value in free space computed in step 2.

Now the simulation time jumps $\Delta t/2$, and the algorithm proceeds with the H-field updates in the grid. Each of the four steps listed above is followed by analogy.

For a wave obliquely incident upon the moving PEC boundary at an angle θ_{inc} relative to the normal, the relativistic boundary conditions of (10.93) and (10.94) must be modified to account for the velocity-dependent reflected-wave angle θ_{ref}. From the special theory of relativity, θ_{ref} satisfies the following relation [29]:

$$\cos\theta_{ref} = \frac{(1+\beta^2)\cos\theta_{inc} - 2\beta}{1 - 2\beta\cos\theta_{inc} + \beta^2} \tag{10.96}$$

A derivation similar to the normal-incidence case leads to the following relativistic boundary conditions suitable for FDTD implementation:

$$E_{tan,tot} = \pm\frac{\beta(\cos\theta_{ref} + \cos\theta_{inc})}{1 \pm \beta\cos\theta_{ref}} E_{tan,inc} \tag{10.97}$$

$$B_{tan,tot} = \frac{\cos\theta_{ref} + \cos\theta_{inc}}{(1 \pm \beta\cos\theta_{ref})\cos\theta_{inc}} B_{tan,inc} \tag{10.98}$$

where the fields refer to the total tangential values at the moving PEC boundary. The numerical implementation is now only slightly more complicated than for the normal-incidence case because of the angular dependence of the incident field values at the surface of the moving boundary.

10.10.2 Illustrative Results

We now review three numerical experiments reported by Harfoush et al. [27] to validate the FDTD relativistic boundary conditions discussed in Section 10.10.1. These experiments were designed to test this technique for both uniformly translating and vibrating PEC surfaces.

Uniformly Moving PEC Surface, Normal Incidence

The first example in [27] involved the illumination at normal incidence of a uniformly moving PEC surface by a plane wave of frequency ω_{inc}. This geometry led to the modeling of the *double-Doppler effect* [30], in which both the frequency and amplitude of the reflected wave are scaled relative to the incident wave by the same multiplying factor, $\alpha = (1-\beta)/(1+\beta)$, where β is defined in (10.95).

Using a uniform FDTD grid of resolution $\Delta = \lambda_0/20$ at ω_{inc}, eight velocities of the PEC surface were modeled. Three simulated the surface advancing toward the incident wave (β = –1/7, –1/5, and –1/3), and five simulated the surface receding from the incident wave (β = 1/7, 1/5, 1/4, 1/3, and 1/2). A DFT was applied to the field time waveforms to obtain the reflected spectrum. In all of the cases studied, it was found that the frequency of the main upshifted or downshifted reflected spectral component agreed with the exact double-Doppler theory up to the resolution limit of the DFT. Further, the amplitude of this component agreed with the theory to better than 1.5%, and spurious frequency components in the reflected spectrum were limited to less than 5% of the amplitude of the incident wave.

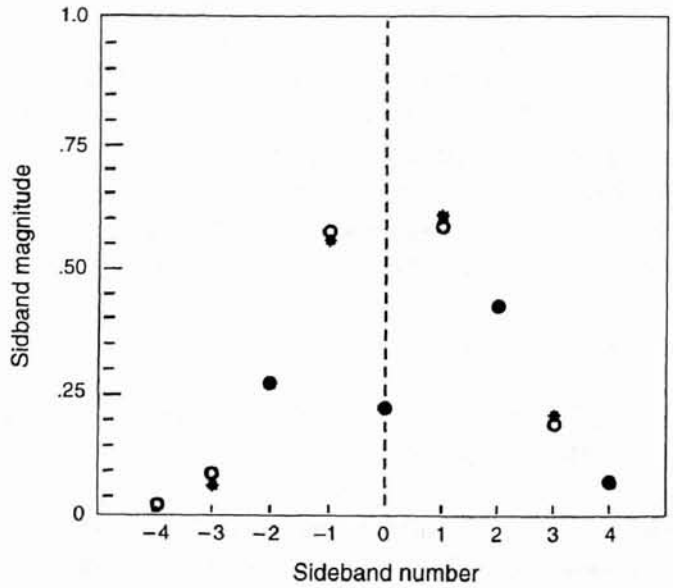
Vibrating PEC Surface, Normal Incidence

The second example in [27] involved the illumination at normal incidence of a PEC surface vibrating at frequency ω_v. Here, the surface was assumed to be located at $z = d\sin(\omega_v t)$, and the maximum normalized surface velocity was $\beta = \omega_v d/c = v_{max}/c$. The exact solution [31, 32] shows that the reflected spectrum is similar to that of an FM tone-modulated signal. Namely, the vibrating surface generates an infinite number of reflected spectral components having the frequencies $\omega_{inc} \pm m\omega_v$, where the amplitude of the m'th component is proportional to J_m, a Bessel function of the first kind of order m. The argument of J_m depends upon the amplitude and frequency of the vibration. A key element in the FDTD model reported in [27] was the use of the "Doppler approximation" [30, 31], where the relativistic boundary conditions of (10.93) and (10.94) were applied at each time-step based upon the assumption that the PEC surface moves with a uniform velocity equal to its instantaneous vibrational value.

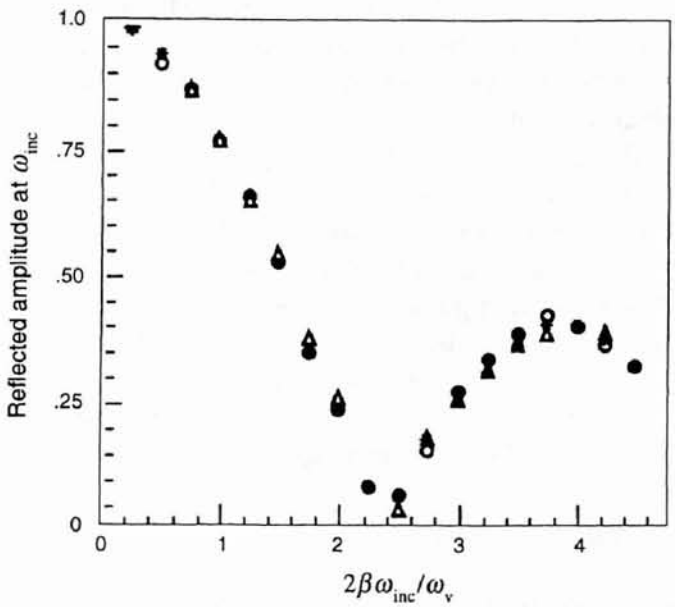
Fig. 10.19(a) compares the FDTD and exact results for the reflected spectrum of the specific case $\omega_v = 0.1\omega_{inc}$, $\beta = 0.1$, and grid resolution $\Delta = \lambda_0/20$ at ω_{inc}. There is less than 5% error in the magnitude of each spectral sideband, and only 0.27% error in the magnitude of the component at ω_{inc}. Fig. 10.19(b) shows that, upon varying the vibration parameters over a wide range, the FDTD-calculated magnitude of the reflected component at ω_{inc} obeys the exact $J_0(2\beta\omega_{inc}/\omega_v)$ dependence over an argument range of 0 to 5 for J_0, including the interesting null of the reflection at the zero of the Bessel function. The level of agreement with the exact solution is the same regardless of whether β is fixed at 0.1 and ω_{inc}/ω_v is varied from 0 to 25, or β is varied from 0 to 0.5 and ω_{inc}/ω_v is fixed at 5. These parametric studies are strongly supportive of the validity of the FDTD modeling approach for this case.

Vibrating PEC Surface, Oblique Incidence

The third example in [27] involved oblique plane-wave incidence on an infinite vibrating PEC surface using the relativistic boundary conditions of (10.97) and (10.98). From the FDTD standpoint, the primary challenge was to model an infinitely long structure in a finite grid. To avoid edge-diffraction effects, it was decided to model a long PEC slab having edges so far away from the observation point that they were causally isolated during a well-defined early time period. The scattered-field waveform observed during this time window was exactly that if the FDTD grid and PEC slab were infinite. From the analytical standpoint, this oblique-incidence problem was much more complicated than normal incidence, since it has no closed-form solution. As described in [33], the analytical solution can be written as an infinite series using a plane-wave expansion, but then the coefficients in the series must be obtained numerically.

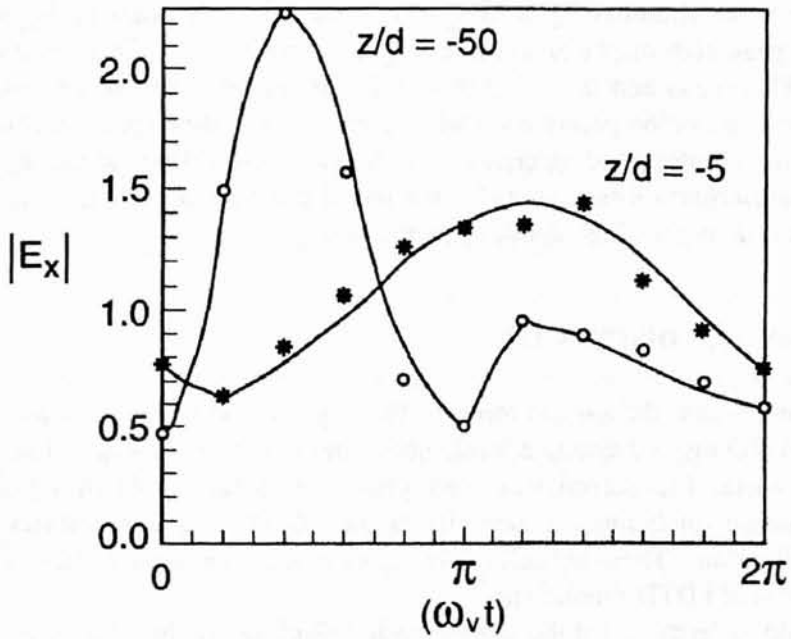


(a) Legend: ooo = FDTD data; ∗ ∗ = exact solution.

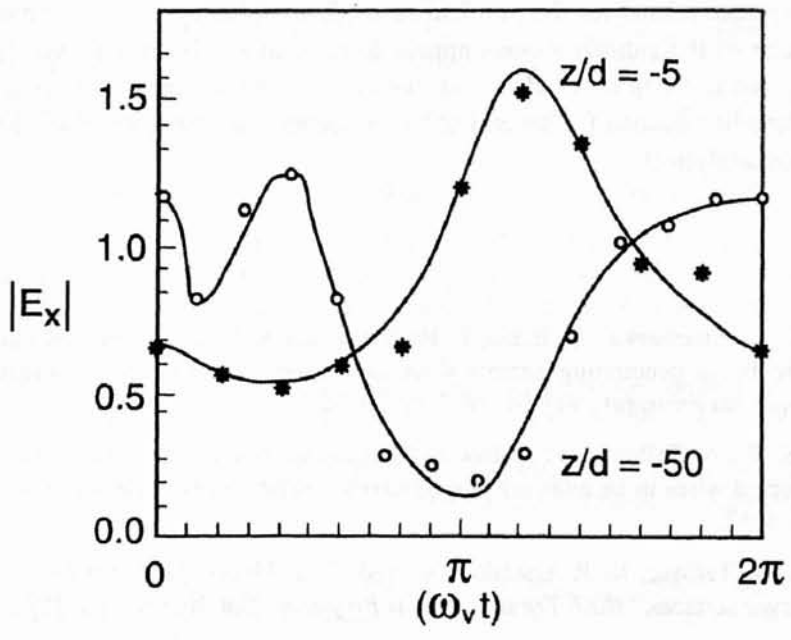


(b) Legend: ooo = FDTD data for $\omega_{inc}/\omega_v = 5$ and β varied from 0–0.5;
$\Delta\,\Delta$ = FDTD data for $\beta = 0.1$ and ω_{inc}/ω_v varied from 0–25; ∗ ∗ = exact solution.

Fig. 10.19 Validation of FDTD results for the reflected spectrum at normal incidence on an infinite PEC surface vibrating at a relativistic speed: (a) spectral sideband amplitudes for $\omega_v = 0.1\omega_{inc}$, $\beta = 0.1$, and grid resolution $\Delta = \lambda_0/20$ at ω_{inc}; (b) magnitude of reflected amplitude at ω_{inc} for a wide-ranging parametric study of surface velocities and vibration frequencies, showing the $J_0(2\beta\omega_{inc}/\omega_v)$ Bessel function dependence. *Source:* Harfoush et al., *IEEE Trans. Antennas and Propagation*, 1989, pp. 55–63, © 1989 IEEE.



(a) Incident angle $\theta_{inc} = 30°$. Legend: —— = analytical results; ∗ ∗ and ∘∘∘ = FDTD data.



(b) Incident angle $\theta_{inc} = 60°$. Legend: —— = analytical results; ∗ ∗ and ∘∘∘ = FDTD data.

Fig. 10.20 Validation of FDTD results for the envelope of the scattered E-field versus time at observation points $z = -5d$ and $z = -50d$ in free space in front of a relativistically vibrating PEC surface illuminated at oblique incidence. Vibration parameters: $\omega_v = 0.2\omega_{inc}$, $d = \lambda_0/2\pi$, $\beta = 0.2$. *Source:* Harfoush et al., *IEEE Trans. Antennas and Propagation*, 1989, pp. 55–63, © 1989 IEEE.

Fig. 10.20(a) compares the FDTD and infinite-series results for a test case involving a TM sinusoidal plane wave illuminating a $20\lambda_0 \times 3\lambda_0$ vibrating PEC slab at $\theta_{inc} = 30°$. Here, $\omega_v = 0.2\omega_{inc}$, the peak slab displacement is $d = \lambda_0/2\pi$, and $\beta = 0.2$. There is good agreement between the FDTD results and the solution of [33] for the envelope of the scattered E-field versus time at the observation points $z = -5d$ and $z = -50d$ in free space in front of the slab. Fig. 10.20(b) shows similar good agreement for an even more oblique angle, $\theta_{inc} = 60°$. These results indicate that Harfoush's method of [27] has promise for engineering models of microwave beam interactions with moving high-density charge fronts.

10.11 SUMMARY AND DISCUSSION

This chapter discussed how the integral forms of Faraday's law and Ampere's law can be used to develop for the FDTD method accurate local subcell models of: (1) narrow slots in conducting plates, (2) thin wires, (3) curved PEC and dielectric surfaces, (4) thin material sheets, (5) surface impedance (including the skin effect), and (6) PEC surfaces subject to relativistic translation and vibration. These subcell models greatly increase the accuracy, efficiency, and range of applications of FDTD simulations.

Readers should understand that the subcell models discussed in this chapter are perturbations of the basic Yee algorithm. Thus, the numerical stability criteria developed in Chapter 4 no longer apply. Because of the complexity of the stability derivations, in many cases an "engineering approach" is utilized wherein numerical experiments are conducted to find a time-step that produces stable results for the problem at hand. Despite the broad range of applications discussed here, none of the subcell models appear to have unusually restrictive stability bounds relative to those given in Chapter 4. However, we must *exercise care* in the choice of the time-step because the stability bounds for several of the procedures discussed in this chapter have not yet been established analytically.

REFERENCES

[1] Taflove, A., K. R. Umashankar, B. Beker, F. Harfoush, and K. S. Yee, "Detailed FDTD analysis of electromagnetic fields penetrating narrow slots and lapped joints in thick conducting screens," *IEEE Trans. Antennas Propagat.*, Vol. 36, 1988, pp. 247–257.

[2] 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 Propagat.*, Vol. 35, 1987, pp. 1248–1257.

[3] Jurgens, T. G., A. Taflove, K. R. Umashankar, and T. G. Moore, "Finite-difference time-domain modeling of curved surfaces," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 357–366.

[4] Jurgens, T. G., and A. Taflove, "Three-dimensional contour FDTD modeling of scattering from single and multiple bodies," *IEEE Trans. Antennas Propagat.*, Vol. 41, 1993, pp. 1703–1708.

[5] Mäkinen, R. M., J. S. Juntunen, and M. A. Kivikoski, "An improved thin-wire model for FDTD," *IEEE Trans. Microwave Theory Tech.*, Vol. 50, 2002, pp. 1245–1255.

[6] Dey, S., and R. Mittra, "A locally conformal finite-difference time-domain algorithm for modeling three-dimensional perfectly conducting objects," *IEEE Microwave Guided Wave Lett.*, Vol. 7, 1997, pp. 273–275.

[7] Dey, S., and R. Mittra, "A modified locally conformal finite-difference time-domain algorithm for modeling three-dimensional perfectly conducting objects," *Microwave Optical Tech. Lett.*, Vol. 17, 1998, pp. 349–352.

[8] Yu, W., and R. Mittra, "A conformal FDTD software package modeling antennas and microstrip circuit components," *IEEE Antennas Propagat. Magazine*, Vol. 42, Oct. 2000, pp. 28–39.

[9] Waldschmidt, G., and A. Taflove, "Three-dimensional CAD-based mesh generator for the Dey-Mittra conformal FDTD algorithm," *IEEE Trans. Antennas Propagat.*, Vol. 52, 2004, pp. 1658–1664.

[10] Taflove, A., Report on P.O. 4059045 to General Dynamics Fort Worth Division, March 1, 1990.

[11] Railton, C. J., and I. J. Craddock, "Stabilised CP-FDTD algorithm for the analysis of arbitrary 3D PEC structures," *Proc. Institute Electrical Engineers*, Vol. 143, part H, 1996, pp. 367–372.

[12] Yu, W., and R. Mittra, "A conformal finite difference time domain technique for modeling curved dielectric surfaces," *IEEE Microwave Wireless Components Lett.*, Vol. 11, 2001, pp. 25–27.

[13] Kaneda, N., B. Houshmand, and T. Itoh, "FDTD analysis of dielectric resonators with curved surfaces," *IEEE Trans. Microwave Theory Tech.*, Vol. 45, 1997, pp. 1645–1649.

[14] Maloney, J. G. and G. S. Smith, "The efficient modeling of thin material sheets in the FDTD method," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 323–330.

[15] 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 Propagat.*, Vol. 38, 1990, pp. 1059–1068.

[16] Leontovich, M. A., "Approximate boundary conditions for the electromagnetic field on the surface of a good conductor," *Investigations on Radiowave Propagation*, Pt. II., 5-12, Printing House of the Academy of Sciences, Moscow, USSR. Translation: V. A. Fock, "Diffraction, refraction, and reflection of radio waves," U. S. Air Force Cambridge Research Center Report TN-57-102, 1957.

[17] Senior, T. B. A., "Impedance boundary conditions for imperfectly conducting surfaces," *Appl. Sci. Res. B*, Vol. 8, 1960, pp. 418–436.

[18] Smith, G. S., "On the skin-effect approximation," *Am. J. Physics*, Vol. 58, 1990, pp. 996–1002.

[19] Beggs, J. H., R. J. Luebbers, K. S. Yee, and K. S. Kunz, "Finite-difference time-domain implementation of surface impedance boundary conditions," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 49–56. See also "Corrections," *IEEE Trans. Antennas Propagat.*, Vol. 41, 1993, p. 118.

[20] Maloney, J. G., and G. S. Smith, "The use of surface impedance concepts in the finite-difference time-domain method," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 38–48.

[21] Roden, J. A., and S. D. Gedney, "The efficient implementation of the surface impedance boundary condition in general curvilinear coordinates," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, 1999, pp. 1954–1963.

[22] Celuch-Marcysiak, M., W. K. Gwarek, and M. Sypniewski, "A simple and effective approach to FDTD modeling of structures including lossy metals," *Proc. Asia-Pacific Microwave Conf.*, Yokohama, Japan, Dec. 1998, pp. 991–993.

[23] Gwarek, W. K., "Analysis of an arbitrarily-shaped planar circuit – A time-domain approach," *IEEE Trans. Microwave Theory Tech.*, Vol. 33, 1985, pp. 1067–1072.

[24] Kellali, S., B. Jecko, and A. Reineix, "Implementation of a surface impedance formalism at oblique incidence in FDTD method," *IEEE Trans. Electromagn. Compat.*, Vol. 35, 1993, pp. 347–356.

[25] Lee, C. F., R. T. Shin, and J. A. Kong, "Time domain modeling of impedance boundary conditions," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, 1992, pp. 1847–1850.

[26] Kärkkäinen, M. K., "Subcell FDTD modeling of electrically thin dispersive layers," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, 2003, pp. 1774–1780.

[27] Harfoush, F., A. Taflove, and G. A. Kriegsmann, "A numerical technique for analyzing electromagnetic wave scattering from moving surfaces in one and two dimensions," *IEEE Trans. Antennas Propagat.*, Vol. 37, 1989, pp. 55–63.

[28] Costen, R. C., and Adamson, "Three dimensional derivation of the electrodynamic jump conditions and momentum-energy laws at a moving boundary," *Proc. IEEE*, Vol. 53, 1965.

[29] Pauli, W., *Theory of Relativity*, Elmsford, NY: Pergamon, 1958.

[30] Cooper, J., "Scattering of electromagnetic fields by a moving boundary: The one-dimensional case," *IEEE Trans. Antennas Propagat.*, Vol. 28, 1980, pp. 791–795.

[31] Van Bladel, J., and D. De Zutter, "Reflections from linearly vibrating objects: Plane mirror at normal incidence," *IEEE Trans. Antennas Propagat.*, Vol. 29, 1981, pp. 629–636.

[32] Kleinman, R. E., "Scattering by linearly vibrating objects," *IEEE Trans. Antennas Propagat.*, Vol. 27, 1979, pp. 344–352.

[33] De Zutter, D., "Reflections from linearly vibrating objects: Plane mirror at oblique incidence," *IEEE Trans. Antennas Propagat.*, Vol. 30, 1982, pp. 898–903.

SELECTED BIBLIOGRAPHY

Cangellaris, A. C., and D. B. Wright, "Analysis of the numerical error caused by the stair-stepped approximation of a conducting boundary in FDTD simulations of electromagnetic phenomena," *IEEE Trans. Antennas Propagat.*, Vol. 39, 1991, pp. 1518–1525.

Chavannes, N., and N. Kuster, "A novel 3D CPFDTD scheme for modeling grid nonnonformally aligned transmitter structures," *IEEE Trans. Antennas Propagat.*, Vol. 52, 2004, pp. 1324–1334.

Denecker, B., F. Olyslager, L. Knockaert, and D. De Zutter, "Generation of FDTD subcell equations by means of reduced order modeling," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 1806–1817.

Dridi, K. H., J. S. Hesthaven, and A. Ditkowski, "Staircase-free finite-difference time-domain formulation for general materials in complex geometries," *IEEE Trans. Antennas Propagat.*, Vol. 49, 2001, pp. 749–756.

Fang, J., and J. Ren, "A locally conformed FDTD algorithm of modeling arbitrary shape planar metal strips," *IEEE Trans. Microwave Theory Tech.*, Vol. 41, 1993, pp. 830–838.

Foroughipour, S. M., and K. P. Esselle, "The theory of a singularity-enhanced FDTD method for diagonal metal edges," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 312–321.

Holland, R., "Pitfalls of staircase meshing," *IEEE Trans. Electromagn. Compat.*, Vol. 35, 1993, pp. 434–439.

Holland R., and L. Simpson, "Finite difference analysis of EMP coupling to thin struts and wires," *IEEE Trans. Electromagn. Compat.*, Vol. 23, 1981, pp. 88–97.

Kärkkäinen, M. K., and S. A. Tretyakov, "Finite-difference time-domain model of interfaces with metals and semiconductors based on a higher order surface impedance boundary condition," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 2448–2455.

Kim, I. S., and W. J. R. Hoefer, "A local mesh refinement algorithm for the time-domain finite-difference method using Maxwell's curl equations," *IEEE Trans. Microwave Theory Tech.*, Vol. 38, 1990, pp. 812–815.

Kosmanis, T. I., and T. D. Tsiboukis, "A systematic and topologically stable conformal finite-difference time-domain algorithm for modeling curved dielectric interfaces in three dimensions," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, 2003, pp. 839–847.

Kunz, K. S., and L. Simpson, "A technique for increasing the resolution of finite-difference solutions to the Maxwell's equations," *IEEE Trans. Electromagn. Compat.*, Vol. 23, 1981, pp. 419–422.

Nadobny, J., D. Sullivan, W. Wlodarczyk, P. Deuflhard, and P. Wust, "A 3D tensor FDTD formulation for treatment of sloped interfaces in electrically inhomogeneous media," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 1760–1770.

Prescott, D. T., and N. V. Shuley, "A method for incorporating different sized cells into the FDTD analysis technique," *IEEE Microwave Guided Wave Lett.*, Vol. 2, 1992, pp. 434–436.

Railton, C. J., "An algorithm for the treatment of curved metallic laminas in the finite difference time domain method," *IEEE Trans. Microwave Theory Tech.*, Vol. 41, 1993, pp. 1429–1438.

Riley, D. J., and C. D. Turner, "Hybrid thin-slot algorithm for the analysis of narrow apertures in FDTD calculations," *IEEE Trans. Antennas Propagat.*, Vol. 38, 1990, pp. 1943–1950.

Riley, D. J., and C. D. Turner, "The inclusion of wall loss in finite difference time domain thin slot algorithms," *IEEE Trans. Electromagn. Compat.*, Vol. 33, 1991, pp. 304–311.

Shorthouse, D. B., and C. J. Railton, "The incorporation of static field solutions in the finite difference time domain method," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, 1992, pp. 986–994.

Tirkas, P. A., and K. R. Demarest, "Modeling of thin dielectric structures using the finite-difference time-domain technique," *IEEE Trans. Antennas Propagat.*, Vol. 39, 1991, pp. 1338–1345.

Wu, C.-T., Y.-H. Pang, and R. B. Wu, "An improved formalism for FDTD analysis of thin-slot problems by conformal mapping technique," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 2530–2533.

Wu, L.-K., and L.-T. Han, "Implementation and application of resistive sheet boundary condition in the FDTD method," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 628–633.

Xiao, T., and Q.-H. Liu, "A staggered upwind embedded boundary (SUEB) method to eliminate the FDTD staircasing error," *IEEE Trans. Antennas Propagat.*, Vol. 52, 2004, pp. 730–741.

Yee, K. S., J. S. Chen, and A. H. Chang, "Conformal finite-difference time-domain with overlapping grids," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 1068–1075.

Yee, K. S., K. Shlager, and A. H. Chang, "An algorithm to implement a surface impedance boundary condition for FDTD," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 833–837.

Zivanovic, S. S., K. S. Yee, and K. K. Mei, "A subgridding method for the time-domain finite-difference method to solve Maxwell's equations," *IEEE Trans. Microwave Theory Tech.*, Vol. 39, 1991, pp. 471–479.

PROJECTS

P10.1 Implement the diagonal split-cell model (Section 10.3.1) for an elliptical PEC cylinder in a TE_z code. Compare the monostatic RCS versus look angle calculated using this approach to that computed using a staircase approximation of the cylinder's surface. What is the grid resolution required for the staircase results to closely approximate the split-cell RCS data, if the latter were obtained using $\Delta = \lambda_0/20$? Compare relative computer burdens for the two methods.

P10.2 Implement the contour-path model of the narrow slot (Section 10.4) in a TE_z code. Replicate the results of Fig. 10.3(a) for the gap E-field in the slotted PEC screen.

P10.3 Implement the contour-path model of the thin wire (Section 10.5) in a three-dimensional code. Replicate the results of Fig. 10.6 for the input impedance of the dipole antenna for the case of 21 cells spanning the antenna length.

P10.4 Repeat P10.1, but use the Yu-Mittra technique (Section 10.6.1) in a TE_z code.

P10.5 Implement the Maloney-Smith thin-material-sheet model (Section 10.7) in a TE_z code. Then replicate the results of Fig. 10.13(a).

P10.6 Implement the Celuch-Marcysiak skin-effect model (Section 10.8.3) in both TE_z and TM_z codes for a parallel-plate waveguide. Then, verify each model relative to the analytical solution for the attenuation of a fundamental mode in the waveguide. Examine the bandwidths over which the solutions are accurate to within 5% of the exact values as a function of the number of GC units in the ladder networks.

Chapter 11

Nonuniform Grids, Nonorthogonal Grids, Unstructured Grids, and Subgrids

Stephen Gedney, Faiza Lansing, and Nicolas Chavannes

11.1 INTRODUCTION

The previous chapters have presented the FDTD algorithm as based on an orthogonal, regular Cartesian lattice. Due to the orthogonality of the grid and the uniform spacing of the gridpoints, the first-order derivatives of Maxwell's equations can be approximated using central-difference operators. This leads to a second-order accurate solution in both space and time. This error, however, is associated only with the differential operators. Additional error can be encountered through the boundary conditions imposed on the discrete fields.

It was shown that PEC and PMC boundaries are naturally treated by the basic FDTD algorithm. Further, discontinuities in material parameters are also naturally treated in a second-order-accurate manner. However, these boundary conditions assume that the physical boundary conforms to the orthogonal lattice. In the event that the grid does not conform to the shape of the physical boundary due to its curvature, additional error results. This is principally because the desired boundary conditions cannot be enforced directly on the physical boundary, but rather on an auxiliary boundary which is a staircased approximation of the physical boundary. Unfortunately, this may lead to a formulation that does not converge to the correct answer, no matter how fine the mesh is made to better resolve the boundary contour [1]. In studies where the field interaction is highly dependent on the shape of the boundary, this can lead to substantial error in the calculation. As a result, FDTD algorithms using boundary-fitted grids have been actively studied. Chapter 10 reviewed a popular class of techniques whereby globally orthogonal space lattices are provided with locally conformal subcell models to permit modeling of fine geometrical features.

This chapter discusses more general FDTD techniques that allow the usage of nonuniform grids, nonorthogonal grids, unstructured grids, and subgrids. These techniques implement either conformal meshes for structure surfaces, or provide enhanced spatial resolution for conventional staircasing. Specific topics to be covered include: (1) nonuniform orthogonal grids, (2) global curvilinear coordinates, (3) irregular nonorthogonal structured grids, (4) irregular nonorthogonal unstructured grids, (5) the planar generalized Yee algorithm, and (6) Cartesian subgrids. Note that all of these approaches fall within the general context of classical FDTD techniques. For emerging hybrid *finite-element time-domain* (FETD) / FDTD techniques similarly aimed at conformal or enhanced-resolution modeling, see Chapter 19.

11.2 NONUNIFORM ORTHOGONAL GRIDS

The FDTD algorithm is second-order-accurate by nature of the central-difference approximations used to realize the first-order spatial and temporal derivatives. This leads to a discrete approximation for the fields based on a uniform orthogonal lattice. Unfortunately, structures with fine geometrical features cannot always conform to the edges of the uniform lattice. Further, it is often desirable to have a refined lattice in localized regions, such as near sharp edges or corners, to accurately model the local field phenomena. With a uniform lattice, this results in global refinement of the mesh density. Since such a high level of refinement is not needed in all regions, this leads to an unnecessary increase in the computational burden. Chapter 10 addressed this issue by using a local subcell approach to model fine features. However, a reduction in the actual cell size is sometimes necessary to more accurately model the local fields, rather than a subcell model. This can be implemented with the use of nonuniform grids.

A quasi-nonuniform grid FDTD algorithm was introduced by Sheen [2]. This method is based on reducing the grid size by exactly one-third. By choosing the subgrid to be exactly one-third, the spatial derivatives of the fields at the interface between the two regions can be expressed using central-difference approximations, resulting in a second-order-accurate formulation. This technique was successfully applied to a number of microwave circuit and antenna problems [2, 3]. However, this method is limited to specific geometries that conform to this specialized grid.

It is clear that more general geometries could be handled by a grid with arbitrary spacing. Unfortunately, central differences can no longer be used to evaluate the spatial derivatives of the fields for such a grid, leading to first-order error. However, it was demonstrated by Monk that, although this formulation does lead to first-order error locally, it results in second-order error globally [4, 5]. This is known as *supraconvergence* [4–7]. In this section, a supraconvergent FDTD algorithm based on nonuniform meshing is presented.

A three-dimensional nonuniform lattice is introduced. The vertices of the lattice are defined by the general one-dimensional coordinates:

$$\{x_i;\ i=1,N_x\};\qquad \{y_j;\ j=1,N_y\};\qquad \{z_k;\ k=1,N_z\} \tag{11.1}$$

The edge lengths between vertices are also defined as

$$\{\Delta x_i = x_{i+1} - x_i;\ \ i=1,N_x-1\};\qquad \{\Delta y_j = y_{j+1} - y_j;\ \ j=1,N_y-1\};$$

$$\{\Delta z_k = z_{k+1} - z_k;\ \ k=1,N_z-1\} \tag{11.2}$$

We introduce the following notation defining the cell and edge centers in the nonuniform space:

$$x_{i+1/2} = x_i + \Delta x_i/2;\qquad y_{j+1/2} = y_j + \Delta y_j/2;\qquad z_{k+1/2} = z_k + \Delta z_k/2 \tag{11.3}$$

A set of dual edge lengths representing the distances between the edge centers is then introduced:

$$\{h_i^x = (\Delta x_i + \Delta x_{i-1})/2;\ \ i=2,N_x\};\qquad \{h_j^y = (\Delta y_j + \Delta y_{j-1})/2;\ \ j=2,N_y\};$$

$$\{h_k^z = (\Delta z_k + \Delta z_{k-1})/2;\ \ k=2,N_z\} \tag{11.4}$$

Finally, the E- and H-fields in the nonuniform grid are denoted as in the following examples:

$$E_x\Big|^{n}_{i+1/2,j,k} \equiv E_x(x_{i+1/2}, y_j, z_k, n\Delta t) \tag{11.5a}$$

$$H_x\Big|^{n+1/2}_{i,j+1/2,k+1/2} \equiv H_x[x_i, y_{j+1/2}, z_{k+1/2}, (n+1/2)\Delta t\,] \tag{11.5b}$$

The nonuniform FDTD algorithm is based on a discretization of Maxwell's equations in their integral form; specifically, Faraday's law and Ampere's law:

$$\oint_L \boldsymbol{E}\cdot d\boldsymbol{L} = -\frac{\partial}{\partial t}\iint_A \boldsymbol{B}\cdot d\boldsymbol{A} - \iint_A \boldsymbol{M}\cdot d\boldsymbol{A} \tag{11.6}$$

$$\oint_L \boldsymbol{H}\cdot d\boldsymbol{L} = \frac{\partial}{\partial t}\iint_A \boldsymbol{D}\cdot d\boldsymbol{A} + \iint_A \sigma\boldsymbol{E}\cdot d\boldsymbol{A} + \iint_A \boldsymbol{J}\cdot d\boldsymbol{A} \tag{11.7}$$

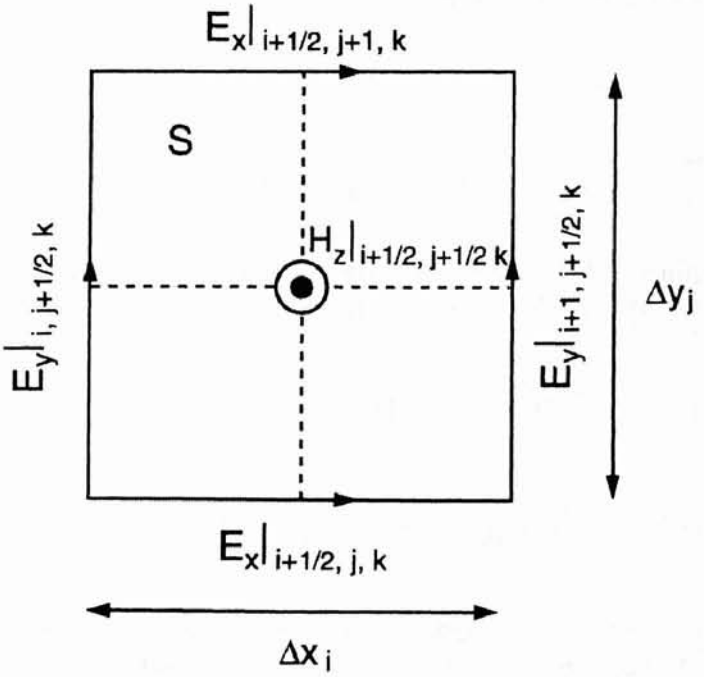
The surface integral in (11.6) is performed over a lattice cell face, and the contour integral is performed over the edges bounding the face, as illustrated in Fig. 11.1(a). Similarly, the surface integral in (11.7) is performed over a dual-lattice cell face.

Evaluating (11.6) and (11.7) over the cell faces using the discrete field approximations in (11.5), and evaluating the time derivatives using central-difference approximations leads to

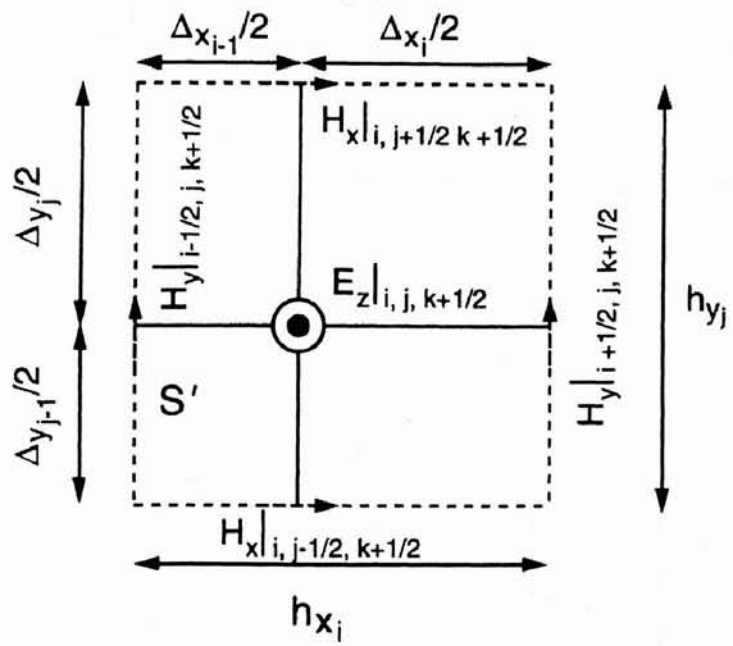
$$\begin{aligned} &E_x\Big|^{n}_{i+1/2,j+1,k}\Delta x_i - E_x\Big|^{n}_{i+1/2,j,k}\Delta x_i - E_y\Big|^{n}_{i+1,j+1/2,k}\Delta y_j + E_y\Big|^{n}_{i,j+1/2,k}\Delta y_j \\ &= -\left[\mu_{i+1/2,j+1/2,k}\left(\frac{H_z\Big|^{n+1/2}_{i+1/2,j+1/2,k} - H_z\Big|^{n-1/2}_{i+1/2,j+1/2,k}}{\Delta t}\right) + M_z\Big|^{n+1/2}_{i+1/2,j+1/2,k}\right]\Delta x_i\,\Delta y_j \end{aligned} \tag{11.8}$$

$$\begin{aligned} &H_x\Big|^{n+1/2}_{i,j+1/2,k+1/2}h_{x_i} - H_x\Big|^{n+1/2}_{i,j-1/2,k+1/2}h_{x_i} - H_y\Big|^{n+1/2}_{i+1/2,j,k+1/2}h_{y_j} + H_y\Big|^{n+1/2}_{i-1/2,j,k+1/2}h_{y_j} \\ &= \left[\begin{array}{l} \varepsilon_{i,j,k+1/2}\left(\dfrac{E_z\Big|^{n+1}_{i,j,k+1/2} - E_z\Big|^{n}_{i,j,k+1/2}}{\Delta t}\right) + \\ \dfrac{\sigma_{i,j,k+1/2}}{2}\left(\dfrac{E_z\Big|^{n+1}_{i,j,k+1/2} + E_z\Big|^{n}_{i,j,k+1/2}}{\Delta t}\right) + J_z\Big|^{n+1/2}_{i,j,k+1/2} \end{array}\right] h_{x_i}\,h_{y_j} \end{aligned} \tag{11.9}$$

where $\varepsilon_{i,j,k+1/2}$, $\sigma_{i,j,k+1/2}$, and $\mu_{i+1/2,j+1/2,k}$ are the averaged permittivity, conductivity, and permeability, respectively, about the grid edges [see (11.84a, b)]. Subsequently, this leads to an explicit update scheme:



(a) Lattice cell face bounded by grid edges, showing a dual-lattice edge passing through its center.



(b) The dual-lattice face bounded by dual edges.

Fig. 11.1 Lattice faces bounded by lattice edges defining surfaces of integration bounded by closed contours.

$$H_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} = H_z\Big|_{i+1/2,j+1/2,k}^{n-1/2} - \frac{\Delta t}{\mu_{i+1/2,j+1/2,k}} \cdot \left[\begin{array}{l} \dfrac{1}{\Delta y_j}\left(E_x\Big|_{i+1/2,j+1,k}^{n} - E_x\Big|_{i+1/2,j,k}^{n} \right) - \\ \dfrac{1}{\Delta x_i}\left(E_y\Big|_{i+1,j+1/2,k}^{n} - E_y\Big|_{i,j+1/2,k}^{n} \right) + M_z\Big|_{i+1/2,j+1/2,k}^{n+1/2} \end{array} \right] \tag{11.10}$$

$$E_z\Big|_{i,j,k+1/2}^{n+1} = \left(\frac{2\varepsilon_{i,j,k+1/2} - \sigma_{i,j,k+1/2}\,\Delta t}{2\varepsilon_{i,j,k+1/2} + \sigma_{i,j,k+1/2}\,\Delta t} \right) E_z\Big|_{i,j,k+1/2}^{n} + \left(\frac{2\Delta t}{2\varepsilon_{i,j,k+1/2} + \sigma_{i,j,k+1/2}\,\Delta t} \right) \cdot \left[\begin{array}{l} \dfrac{1}{h_{y_j}}\left(H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} - H_x\Big|_{i,j-1/2,k+1/2}^{n+1/2} \right) - \\ \dfrac{1}{h_{x_i}}\left(H_y\Big|_{i+1/2,j,k+1/2}^{n+1/2} - H_y\Big|_{i-1/2,j,k+1/2}^{n+1/2} \right) - J_z\Big|_{i,j,k+1/2}^{n+1/2} \end{array} \right] \tag{11.11}$$

Similar updates for the remaining field components are easily derived by permuting the indices in (11.10) and (11.11) in a right-handed manner. This leads to an explicit solution for the E- and H-fields, which is stable providing that

$$\Delta t < \frac{1}{c\sqrt{\dfrac{1}{\left(\Delta x_{i_{\min}}\right)^2} + \dfrac{1}{\left(\Delta y_{j_{\min}}\right)^2} + \dfrac{1}{\left(\Delta z_{k_{\min}}\right)^2}}} \tag{11.12}$$

where $\Delta x_{i_{\min}}$, $\Delta y_{j_{\min}}$, and $\Delta z_{k_{\min}}$ are the minimum edge lengths along the x-, y-, and z-directions, respectively, in the nonuniform grid.

The explicit updates for the H-fields in (11.10) are second-order-accurate in both space and time, since the vertices of the dual lattice are assumed to be located at the cell centers of the primary lattice. On the other hand, the explicit updates for the E-fields in (11.11) are only first-order-accurate in space. This results in local first-order error in regions where the grid is nonuniform. However, Monk has shown that despite the local first-order error, the nonuniform model is globally second-order-accurate [4, 5]. Specifically, the method is supraconvergent since it converges with a higher order accuracy than the local error mandates.

The principle of supraconvergence is now demonstrated through a numerical example. Referring to Figs. 11.2 and 11.3, we consider calculation of the resonant frequencies of a rectangular $0.5 \times 1.0 \times 0.25$m PEC cavity using the nonuniform FDTD method. A random grid spacing for x_i, y_j, and z_k is assumed such that

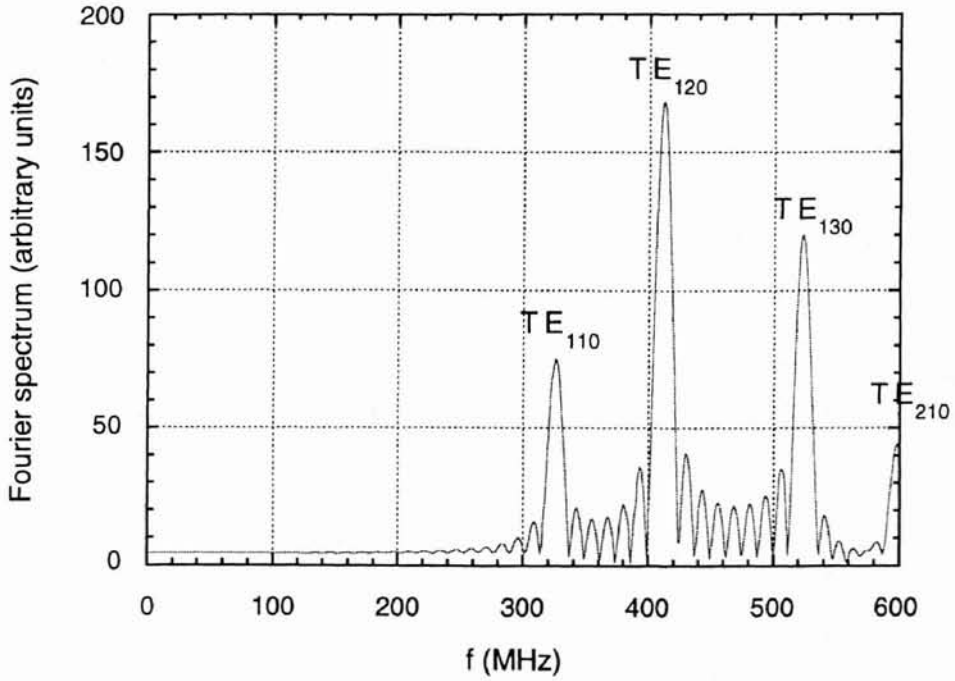


Fig. 11.2 Amplitude of the Fourier transform of the time-varying E-field probed in the resonant cavity for a $21 \times 41 \times 11$ nonuniform lattice with random grid spacing.

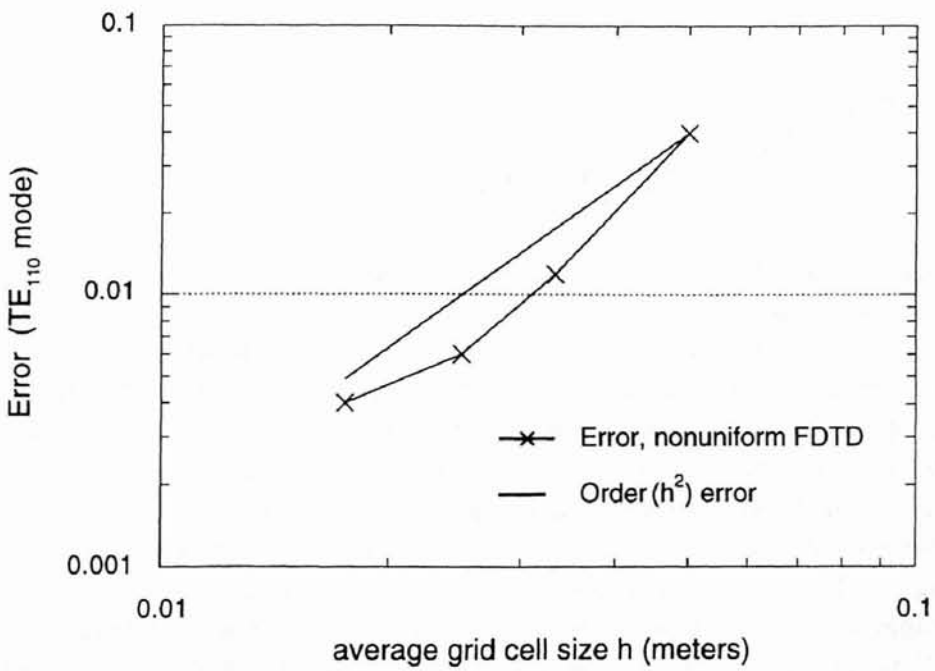


Fig. 11.3 Error convergence of the resonant frequency of the TE_{110} mode computed using the nonuniform FDTD algorithm.

$$\left\{ x_i = (i-1)\Delta x + 0.5\Re\Delta x \;;\; i = 1, N_x \right\}$$

$$\left\{ y_j = (j-1)\Delta y + 0.5\Re\Delta y \;;\; j = 1, N_y \right\} \qquad (11.13a, b, c)$$

$$\left\{ z_k = (k-1)\Delta z + 0.5\Re\Delta z \;;\; k = 1, N_z \right\}$$

where $-1/2 \leq \Re \leq 1/2$ is a random number. Note that the cavity boundaries remain fixed.

The interior of the cavity is excited with a time-varying magnetic dipole placed off a center axis. The magnetic current is z-directed and has the Gaussian time variation

$$\boldsymbol{M}(t) \;=\; \hat{z}\, e^{-(t-t_0)^2/T^2} \qquad (11.14)$$

where $T = 0.15$ ns and $t_0 = 0.45$ ns. Expressions analogous to (11.10) and (11.11) are used to time-step the H- and E-fields within the cavity region, subject to the standard PEC boundary condition on the cavity walls, $\hat{\boldsymbol{n}} \times \boldsymbol{E} = 0$. The time-varying E-field is then probed off a center axis to avoid the nulls of odd resonant modes, and the cavity resonant frequencies are extracted using an FFT, as per Chapter 15, Section 15.8.

Fig. 11.2 shows an example of the Fourier spectrum of the time-varying field within the cavity. Here, the spectral peaks correspond to the resonant frequencies of the cavity modes. Subsequently, the average grid cell size h is reduced, and the entire simulation is run again. Fig. 11.3 graphs the results of four such runs for the error of the nonuniform grid FDTD model in calculating the resonant frequency of the TE_{110} mode relative to the exact solution [8], as well as a generic order(h^2) accuracy slope. We see that the convergence of the resonant frequency is indeed second-order.

The nonuniform FDTD method is extremely well-suited for the analysis of planar microwave circuits. The geometrical details of such circuits are typically electrically small, leading to very small cell sizes. Further, microwave circuits are often located in an unbounded medium, requiring absorbing boundaries to be placed a sufficient distance from the circuit to avoid nonphysical reflections. For uniform meshing, these two characteristics can potentially combine to produce very large lattices. With a nonuniform grid, the local cell size can be altered such that field singularities near edges and corners are accurately modeled, while coarser cells are used in regions where the fields are better behaved. It is noted that the maximum cell size in coarser regions should still be less than $\lambda_{\min}/15$, where $\lambda_{\min}$ is the smallest wavelength corresponding to the upper bandwidth of the excitation.

As an example, consider the microstrip circuit in Fig. 11.4, a lowpass filter printed on a dielectric substrate backed by a PEC ground plane. The substrate and the air above are assumed to be infinite in extent. The traditional FDTD algorithm is first used to model this circuit with $\Delta x = 0.4064$ mm, $\Delta y = 0.4233$ mm, $\Delta z = 0.265$ mm, and $\Delta t = 0.441$ ps [9]. The lattice has a dimension of $80 \times 110 \times 16$ cells, and the total simulation requires 4,000 time-steps. The time-dependent line voltages V_1 and V_2 are calculated at the input and output ports, and the frequency-dependent S parameters are computed for the two-port device after a DFT (see Chapter 15, Section 15.2). We note that the circuit boundaries cannot be modeled exactly using the above discretization, since the widths of the lines are noninteger multiples of the grid spacing. For accuracy, the grid spacing is based on the dimensions of the resonant strip (the horizontal microstrip), since the resonant frequency is usually sensitive to small dimensional errors.

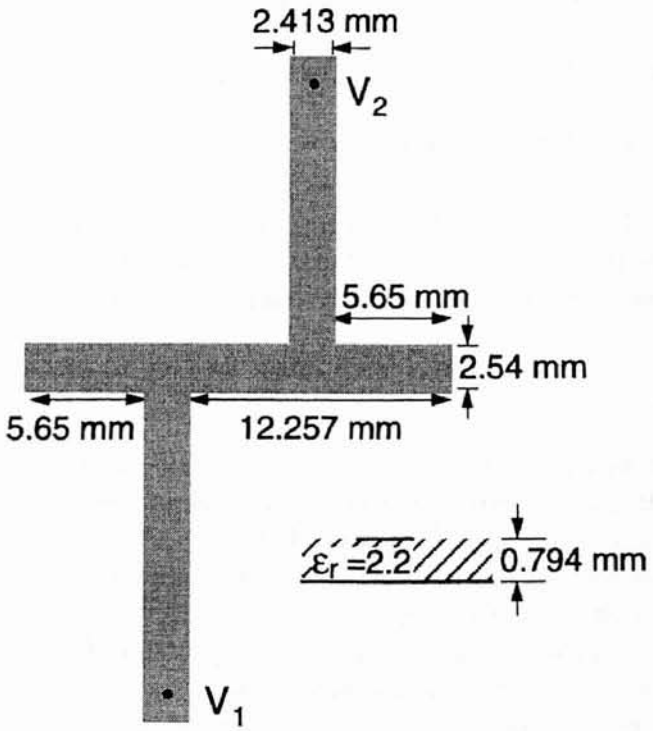


Fig. 11.4 Microstrip lowpass filter printed on a 0.794-mm substrate of permittivity $\varepsilon_r = 2.2$.

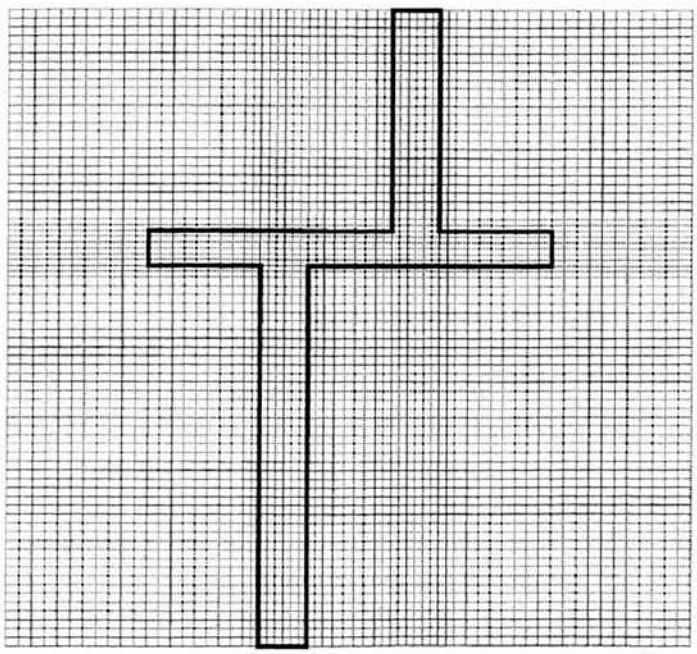


Fig. 11.5 Top view of the mesh used for the nonuniform grid FDTD simulation.

Next, the nonuniform FDTD grid illustrated in Fig. 11.5 is used to simulate the same circuit. This grid allows the microstrip boundary positions to be exactly modeled. Here, a global cell size of $\Delta x = 0.64$ mm, $\Delta y = 0.635$ mm, and $\Delta z = 0.265$ mm is used, along with mesh refinement, to more accurately calculate the fields near the microstrip edges. The overall lattice has a dimension of $64 \times 76 \times 16$ cells, and a time-step of 0.441 ps is used, resulting in 4,000 iterations. Because the physical locations of the voltage observation points are the same as used earlier for the uniform FDTD grid, the two calculations can be directly compared. Fig. 11.6 shows that there is excellent agreement of the S parameters calculated using the uniform and nonuniform grids, although the nonuniform grid provides a computer time savings of about 50%.

This example illustrates the usefulness of the nonuniform-grid FDTD algorithm. For complex and highly detailed circuits having rectangular geometries, the nonuniform FDTD algorithm is clearly a powerful and versatile technique. The nonuniform grid can easily conform to such geometries without substantially increasing the lattice dimensions and without losing accuracy. However, the reader is cautioned that the grid should not change size too rapidly. As a rule of thumb, one should try to maintain $0.5\Delta x_{i\pm1} \le \Delta x_i \le 2\Delta x_{i\pm1}$ to avoid large local error. If a large change in grid density is needed, this can be realized more accurately by scaling over a few cells. While this is not a strict rule, it provides confidence in the numerical solution.

11.3 LOCALLY CONFORMAL GRIDS, GLOBALLY ORTHOGONAL

This class of algorithms is based on globally orthogonal grids having only those space cells located immediately adjacent to curved boundary surfaces distorted to conform to these surfaces. Known as the conformal or contour-path FDTD approach, this technique has been used successfully for a number of applications [10, 11]. In particular, work by Yu and Mittra [12] has resulted in a simple approach for modeling curved and arbitrarily oriented PEC surfaces in the FDTD lattice while retaining good numerical stability. See Chapter 10, Section 10.6 for details and illustrative results of this class of conformal gridding techniques.

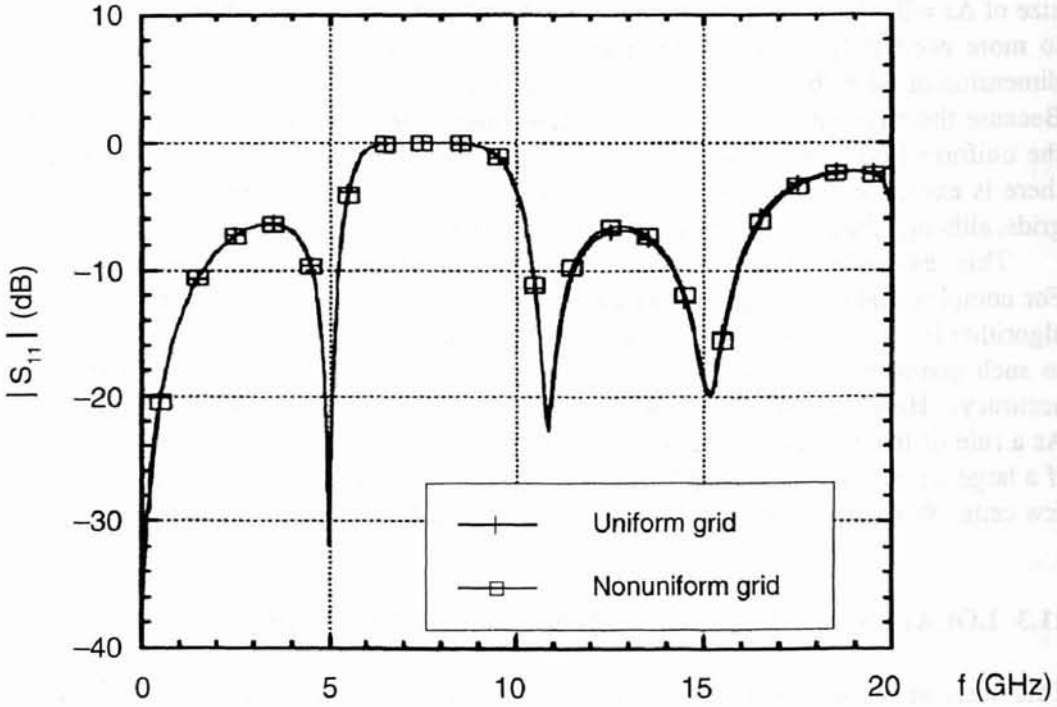
Locally conformal FDTD gridding methods are based on a regular orthogonal grid, with the exception that the grid is deformed in the near vicinity of the boundary interface. As a result, Maxwell's curl equations cannot be discretized in the deformed cells using simple central differences. Rather, the discretization in these cells is based on Maxwell's equations in integral form; specifically, Faraday's and Ampere's laws of (11.6) and (11.7). Here, special contour-path and surface integrals are performed for each deformed cell. As discussed in Chapter 10, such techniques can result in highly accurate solutions with little computational overhead.

11.4 GLOBAL CURVILINEAR COORDINATES

11.4.1 Nonorthogonal Curvilinear FDTD Algorithm

A number of investigations have demonstrated that an efficient FDTD algorithm can be formulated for a regular grid described by a global curvilinear coordinate system [13, 14]. This is shown by mapping Maxwell's curl equations from curvilinear to Cartesian coordinates. The advantage of such a scheme is that a general nonorthogonal curvilinear system can be introduced that conforms to the boundaries of the problem, and a uniform update scheme can be described for the three-dimensional space.

(a)



(b)

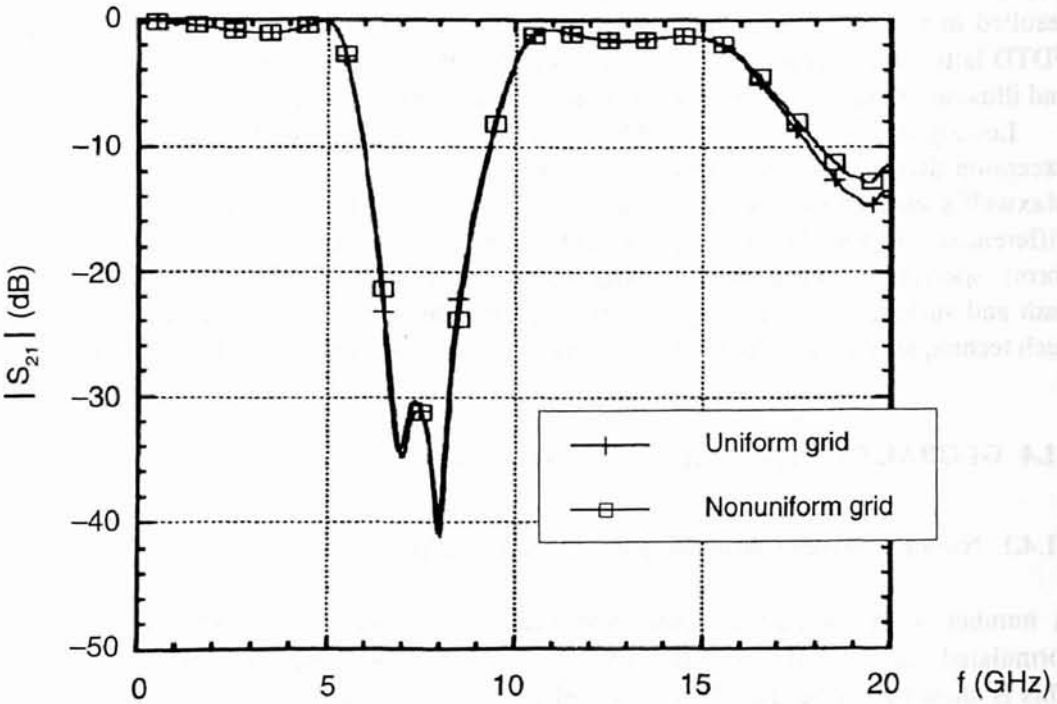


Fig. 11.6 Comparison of the S parameters for the lowpass filter of Fig. 11.5, computed using the uniform-grid and nonuniform-grid FDTD algorithms: (a) S_{11}; (b) S_{21}.

Referring to Fig. 11.7, consider a general nonorthogonal curvilinear coordinate system $(\boldsymbol{u}^1, \boldsymbol{u}^2, \boldsymbol{u}^3)$. Within this coordinate space, the differential length vector is described as

$$\boldsymbol{dr} = \sum_{i=1}^{3} \frac{\partial \boldsymbol{r}}{\partial u^i} du^i = \sum_{i=1}^{3} \boldsymbol{a}_i du^i \tag{11.15}$$

where the $\boldsymbol{a}_i$ are known as the unitary vectors. A reciprocal basis is also introduced, and is related to the unitary basis as

$$\boldsymbol{a}^1 = \frac{\boldsymbol{a}_2 \times \boldsymbol{a}_3}{\sqrt{g}}; \quad \boldsymbol{a}^2 = \frac{\boldsymbol{a}_3 \times \boldsymbol{a}_1}{\sqrt{g}}; \quad \boldsymbol{a}^3 = \frac{\boldsymbol{a}_1 \times \boldsymbol{a}_2}{\sqrt{g}} \tag{11.16}$$

where g is the determinant of the metric tensor with the elements

$$g_{i,j} = \sum_{k=1}^{3} \frac{\partial x^k}{\partial u^i} \frac{\partial x^k}{\partial u^j} = \boldsymbol{a}_i \cdot \boldsymbol{a}_j \tag{11.17}$$

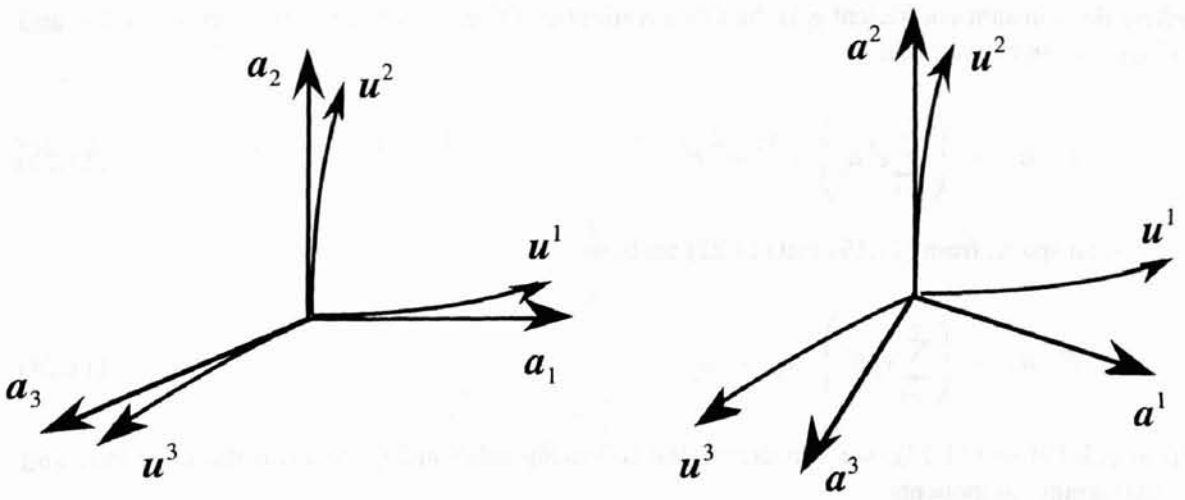


Fig. 11.7 Nonorthogonal curvilinear coordinate system $(\boldsymbol{u}^1, \boldsymbol{u}^2, \boldsymbol{u}^3)$: (a) unitary vectors $(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$; (b) reciprocal unitary vectors $(\boldsymbol{a}^1, \boldsymbol{a}^2, \boldsymbol{a}^3)$.

It can be further shown that

$$\sqrt{g} = \boldsymbol{a}_1 \cdot \boldsymbol{a}_2 \times \boldsymbol{a}_3 \tag{11.18}$$

From (11.16), it can be seen that the reciprocal vectors $\boldsymbol{a}^i$ are orthogonal to the unitary vectors $\boldsymbol{a}_j$, $\boldsymbol{a}_k$ $(i \neq j, k)$, and from (11.18) we derive the relationship that

$$\boldsymbol{a}^i \cdot \boldsymbol{a}_j = \delta_{i,j} \tag{11.19}$$

where $\delta_{i,j}$ is the Kronecker delta function. Further, the inverse metric is defined where

$$\boldsymbol{a}^i \cdot \boldsymbol{a}^j = g^{i,j} \tag{11.20}$$

and $g^{i,j}$ defines a metric tensor that is the inverse of the metric tensor $g_{i,j}$.

Using the notation of Stratton [15], a vector field in the curvilinear space can be represented by its *contravariant* or by its *covariant* components. To this end, the E-field vector can be expanded by the unitary vectors as

$$\boldsymbol{E} = \sum_{i=1}^{3} e^i \boldsymbol{a}_i \tag{11.21}$$

where the constant coefficient e^i is the i'th contravariant component of the E-field; or, it can be expanded using the reciprocal vectors as

$$\boldsymbol{E} = \sum_{i=1}^{3} e_i \boldsymbol{a}^i \tag{11.22}$$

where the constant coefficient e_i is the i'th covariant component of the E-field. From (11.19) and (11.21), it can be seen that

$$\boldsymbol{E} \cdot \boldsymbol{a}^j = \left(\sum_{i=1}^{3} e^i \boldsymbol{a}_i \right) \cdot \boldsymbol{a}^j = e^j \tag{11.23}$$

In the dual space, from (11.19) and (11.22) we have

$$\boldsymbol{E} \cdot \boldsymbol{a}_j = \left(\sum_{i=1}^{3} e_i \boldsymbol{a}^i \right) \cdot \boldsymbol{a}_j = e_j \tag{11.24}$$

From (11.19) to (11.24), we can derive the following relationship between the covariant and contravariant components:

$$e_i = \sum_{j=1}^{3} g_{i,j} e^j \quad ; \qquad e^i = \sum_{j=1}^{3} g^{i,j} e_j \tag{11.25a, b}$$

Since neither the unitary basis nor the reciprocal basis are unit vectors (i.e., with a unit amplitude), the e^i and e_i do not have units of volts per meter, but rather have dimensions that are dependent on the curvilinear space. However, from (11.17) to (11.24), contravariant and covariant components of E having units of volts per meter can be expressed by the scaled values

$$E^i = \sqrt{g_{i,i}}\, e^i \quad ; \qquad E_j = \sqrt{g^{j,j}}\, e_j \tag{11.26a, b}$$

Based on the above discussion, Maxwell's curl equations can be represented using the covariant and contravariant projections. To this end, Maxwell's curl equations in a source-free medium are expressed as

$$-\mu \frac{\partial \boldsymbol{H}}{\partial t} = \nabla \times \boldsymbol{E} \tag{11.27}$$

$$\varepsilon \frac{\partial \boldsymbol{E}}{\partial t} + \sigma \boldsymbol{E} = \nabla \times \boldsymbol{H} \tag{11.28}$$

Taking the dot product of (11.27) with $\boldsymbol{a}^i$ $(i = 1, 2, 3)$ leads to

$$-\mu \frac{\partial h^1}{\partial t} = \frac{1}{\sqrt{g}} \left(\frac{\partial e_3}{\partial u^2} - \frac{\partial e_2}{\partial u^3} \right) \tag{11.29a}$$

$$-\mu \frac{\partial h^2}{\partial t} = \frac{1}{\sqrt{g}} \left(\frac{\partial e_1}{\partial u^3} - \frac{\partial e_3}{\partial u^1} \right) \tag{11.29b}$$

$$-\mu \frac{\partial h^3}{\partial t} = \frac{1}{\sqrt{g}} \left(\frac{\partial e_2}{\partial u^1} - \frac{\partial e_1}{\partial u^2} \right) \tag{11.29c}$$

Similarly, taking the dot product of (11.28) with $\boldsymbol{a}^i$ $(i = 1, 2, 3)$ leads to

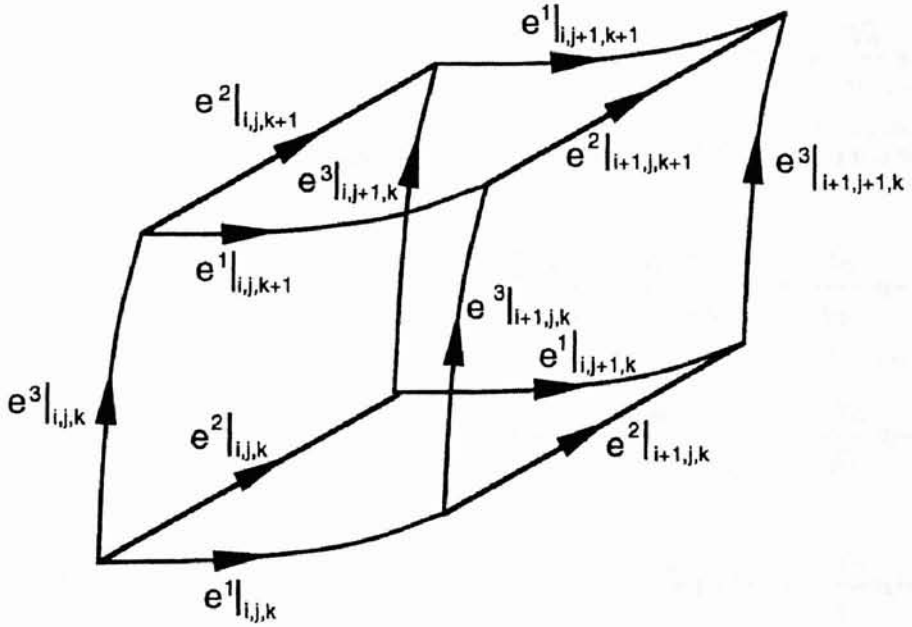
$$\varepsilon \frac{\partial e^1}{\partial t} + \sigma e^1 = \frac{1}{\sqrt{g}} \left(\frac{\partial h_3}{\partial u^2} - \frac{\partial h_2}{\partial u^3} \right) \tag{11.30a}$$

$$\varepsilon \frac{\partial e^2}{\partial t} + \sigma e^2 = \frac{1}{\sqrt{g}} \left(\frac{\partial h_1}{\partial u^3} - \frac{\partial h_3}{\partial u^1} \right) \tag{11.30b}$$

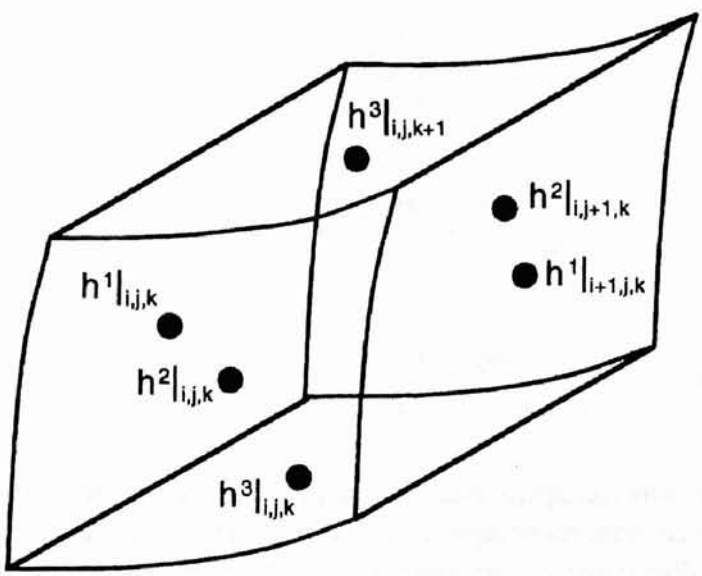
$$\varepsilon \frac{\partial e^3}{\partial t} + \sigma e^3 = \frac{1}{\sqrt{g}} \left(\frac{\partial h_2}{\partial u^1} - \frac{\partial h_1}{\partial u^2} \right) \tag{11.30c}$$

Using central-difference approximations in time and space, an explicit time-marching scheme can be derived from these equations. To this end, a dual-lattice is introduced in the nonorthogonal curvilinear space, for which a typical lattice cell is illustrated in Fig. 11.8. Subsequently, (11.29a) and (11.30a) lead to the updates

$$h^1\Big|_{i,j,k}^{n+1} = h^1\Big|_{i,j,k}^{n} - \frac{\Delta t}{\mu \sqrt{g}} \left[\left(\frac{e_3\Big|_{i,j+1,k}^{n+1/2} - e_3\Big|_{i,j,k}^{n+1/2}}{\Delta u^2} \right) - \left(\frac{e_2\Big|_{i,j,k+1}^{n+1/2} - e_2\Big|_{i,j,k}^{n+1/2}}{\Delta u^3} \right) \right] \tag{11.31}$$



(a) ~ electric field



(b) ~ magnetic field

Fig. 11.8 Electric and magnetic field discretization within a unit cell of the lattice, based on a nonorthogonal curvilinear coordinate system.

$$e^1\Big|_{i,j,k}^{n+1/2} = \left(\frac{2\varepsilon|_{i,j,k} - \sigma|_{i,j,k}\,\Delta t}{2\varepsilon|_{i,j,k} + \sigma|_{i,j,k}\,\Delta t}\right) e^1\Big|_{i,j,k}^{n-1/2} \tag{11.32}$$

$$+ \frac{2\Delta t}{\left(2\varepsilon|_{i,j,k} + \sigma|_{i,j,k}\,\Delta t\right)\sqrt{g}}\left[\left(\frac{h_3\big|_{i,j,k}^{n} - h_3\big|_{i,j-1,k}^{n}}{\Delta u^2}\right) - \left(\frac{h_2\big|_{i,j,k}^{n} - h_2\big|_{i,j,k-1}^{n}}{\Delta u^3}\right)\right]$$

where the (i, j, k) indices are referenced to Fig. 11.8. The updates of the remaining components of e and h can be derived by permuting the indices.

We note that the covariant components of the E- and H-fields appear on the right-hand sides of the above updating expressions for the contravariant components. Once the contravariant components of the fields are updated, the covariant components must first be computed using (11.25a) before the dual-field can be updated. This projection requires all three of the contravariant components. However, these components are not known in space at the same locations, and they must be averaged in space to maintain second-order accuracy. In the discrete space, the projections are thus expressed by

$$\begin{aligned} h_1\big|_{i,j,k}^{n} &= g_{1,1}\,h^1\big|_{i,j,k}^{n} + 0.25\,g_{1,2}\left(h^2\big|_{i-1,j,k}^{n} + h^2\big|_{i-1,j+1,k}^{n} + h^2\big|_{i,j,k}^{n} + h^2\big|_{i,j+1,k}^{n}\right) \\ &+ 0.25\,g_{1,3}\left(h^3\big|_{i-1,j,k}^{n} + h^3\big|_{i-1,j,k+1}^{n} + h^3\big|_{i,j,k}^{n} + h^3\big|_{i,j,k+1}^{n}\right) \end{aligned} \tag{11.33a}$$

$$\begin{aligned} e_1\big|_{i,j,k}^{n} &= g_{1,1}\,e^1\big|_{i,j,k}^{n} + 0.25\,g_{1,2}\left(e^2\big|_{i,j,k}^{n} + e^2\big|_{i,j-1,k}^{n} + e^2\big|_{i+1,j,k}^{n} + e^2\big|_{i+1,j-1,k}^{n}\right) \\ &+ 0.25\,g_{1,3}\left(e^3\big|_{i,j,k}^{n} + e^3\big|_{i,j,k-1}^{n} + e^3\big|_{i+1,j,k}^{n} + e^3\big|_{i+1,j,k-1}^{n}\right) \end{aligned} \tag{11.33b}$$

where the indices (i, j, k) are referenced to Fig. 11.8.

The field updates in (11.31) to (11.33) are in the contravariant space. The contravariant field components can be solved explicitly, providing that the time-step Δt satisfies the stability criterion defined in the next section. Finally, for both near-field and far-field calculations, the physical E- and H-fields can be derived from the contravariant fields using (11.26).

11.4.2 Stability Criterion

The stability criterion for the nonorthogonal curvilinear FDTD algorithm [16] can be derived in a manner similar to that for FDTD on a Cartesian grid, as discussed in Chapter 4. Assuming the three-dimensional space to have a homogeneous material profile and to be source-free, the E- (and H-) fields must satisfy the wave equation

$$\nabla \times \nabla \times \boldsymbol{E} + \frac{1}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = 0 \tag{11.34}$$

where c is the speed of light within the homogeneous material medium. Applying the vector identity $\nabla \times \nabla \times \boldsymbol{E} = -(\nabla \cdot \nabla)\boldsymbol{E} + \nabla(\nabla \cdot \boldsymbol{E})$ and Gauss' law $\nabla \cdot \boldsymbol{E} = 0$, then (11.34) becomes

$$\nabla^2 \boldsymbol{E} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = 0 \tag{11.35}$$

The E-field is now expanded into a superposition of the eigenmodes of (11.35), which can be expressed as the spectrum of plane waves. The FDTD algorithm must be stable for all components of the plane-wave expansion. Thus, it can be shown that if it is stable for an arbitrary plane wave, it is stable for all plane waves. We express

$$\boldsymbol{E}(u^1, u^2, u^3, t) = \boldsymbol{e}(t)\, e^{-j\boldsymbol{k}\cdot\boldsymbol{r}} \tag{11.36}$$

where $\boldsymbol{e}(t)$ is an arbitrary function of time. This can be further expressed as

$$\boldsymbol{k} \cdot \boldsymbol{r} = k_1 u^1 + k_2 u^2 + k_3 u^3 \tag{11.37}$$

and

$$k_i = \boldsymbol{k} \cdot \boldsymbol{a}_i \tag{11.38}$$

The task at hand is to determine the stability of the explicit time-stepping scheme in (11.31) to (11.33). To this end, the growth-factor α is defined, where

$$\alpha = e^{n+1} / e^n \tag{11.39}$$

Stability requires that $|\alpha| \leq 1$. To derive the stability criterion, central-difference approximations are used to discretize (11.35), leading to an explicit update scheme. It is sufficient to show stability for this explicit scheme, in order to prove stability for the explicit update scheme in (11.31) to (11.33).

The stability analysis proceeds as follows. In the general nonorthogonal curvilinear coordinate system, the ∇ operator is expressed as

$$\nabla = \boldsymbol{a}^1 \frac{\partial}{\partial u^1} + \boldsymbol{a}^2 \frac{\partial}{\partial u^2} + \boldsymbol{a}^3 \frac{\partial}{\partial u^3} \tag{11.40}$$

Using central-difference approximations, (11.40) and (11.36) lead to

$$\nabla^2 \boldsymbol{E} = (\nabla \cdot \nabla)\boldsymbol{E} = 2j \left[\sum_{i=1}^{3} \frac{\boldsymbol{a}^i}{\Delta u^i} \sin\left(\frac{k_i \Delta u^i}{2} \right) \right] \cdot 2j \left[\sum_{m=1}^{3} \frac{\boldsymbol{a}^m}{\Delta u^m} \sin\left(\frac{k_m \Delta u^m}{2} \right) \right] \boldsymbol{E} \tag{11.41}$$

From (11.41), and using a central-difference approximation of the second-order time derivative, (11.35) is expressed as

$$-4\left[\sum_{i=1}^{3}\frac{\boldsymbol{a}^i}{\Delta u^i}\sin\left(\frac{k_i\,\Delta u^i}{2}\right)\right]\cdot\left[\sum_{m=1}^{3}\frac{\boldsymbol{a}^m}{\Delta u^m}\sin\left(\frac{k_m\Delta u^m}{2}\right)\right]\boldsymbol{E}^n=\frac{1}{c^2}\frac{\boldsymbol{E}^{n+1}-2\boldsymbol{E}^n+\boldsymbol{E}^{n-1}}{(\Delta t)^2} \tag{11.42}$$

From (11.39), we assume $\boldsymbol{E}^{n+1}=\alpha\boldsymbol{E}^n$ and $\boldsymbol{E}^{n-1}=\boldsymbol{E}^n/\alpha$, leading to

$$-4\left[\sum_{i=1}^{3}\frac{\boldsymbol{a}^i}{\Delta u^i}\sin\left(\frac{k_i\,\Delta u^i}{2}\right)\right]\cdot\left[\sum_{m=1}^{3}\frac{\boldsymbol{a}^m}{\Delta u^m}\sin\left(\frac{k_m\,\Delta u^m}{2}\right)\right]\boldsymbol{E}^n=\frac{1}{c^2}\left(\frac{\alpha^2-2\alpha+1}{\alpha(\Delta t)^2}\right)\boldsymbol{E}^n \tag{11.43}$$

Finally, solving for the growth-factor α from (11.43), we obtain

$$\alpha=\left[1-2s^2(\Delta t)^2\right]\pm 2s\,\Delta t\sqrt{s^2(\Delta t)^2-1} \tag{11.44}$$

where

$$s^2=c^2\left[\sum_{i=1}^{3}\frac{\boldsymbol{a}^i}{\Delta u^i}\sin\left(\frac{k_i\,\Delta u^i}{2}\right)\right]\cdot\left[\sum_{m=1}^{3}\frac{\boldsymbol{a}^m}{\Delta u^m}\sin\left(\frac{k_m\,\Delta u^m}{2}\right)\right] \tag{11.45}$$

We see from (11.44) that $|\alpha|\le 1$ only if

$$s^2(\Delta t)^2\le 1 \tag{11.46}$$

This must be true for all plane waves, and therefore it can be said that

$$s^2\le c^2\left[\sum_{i=1}^{3}\frac{\boldsymbol{a}^i}{\Delta u^i}\right]\cdot\left[\sum_{m=1}^{3}\frac{\boldsymbol{a}^m}{\Delta u^m}\right]=c^2\sum_{i=1}^{3}\sum_{m=1}^{3}\frac{g^{i,m}}{\Delta u^i\,\Delta u^m} \tag{11.47}$$

where $g^{i,m}$ is defined in (11.20). Then, combining (11.47) and (11.46), we find that $|\alpha|\le 1$ if

$$\Delta t\le\frac{1}{c\sqrt{\sum_{i=1}^{3}\sum_{m=1}^{3}\frac{g^{i,m}}{\Delta u^i\,\Delta u^m}}} \tag{11.48}$$

Equation (11.48) is the stability criterion for the nonorthogonal curvilinear FDTD algorithm.

One check of this result is that the stability criterion of the nonorthogonal curvilinear FDTD algorithm should reduce to that derived for Cartesian space. In Cartesian space, we have

$$\boldsymbol{a}_1=\hat{\boldsymbol{x}}\;;\qquad \boldsymbol{a}_2=\hat{\boldsymbol{y}}\;;\qquad \boldsymbol{a}_3=\hat{\boldsymbol{z}} \tag{11.49a}$$

and

$$\Delta u^1 = \Delta x \; ; \quad \Delta u^2 = \Delta y \; ; \quad \Delta u^3 = \Delta z \tag{11.49b}$$

and from (11.16)

$$\boldsymbol{a}^1 = \hat{\boldsymbol{x}} \; ; \quad \boldsymbol{a}^2 = \hat{\boldsymbol{y}} \; ; \quad \boldsymbol{a}^3 = \hat{\boldsymbol{z}} \tag{11.49c}$$

From (11.20) and (11.48), it can be shown for this case that

$$\Delta t \leq \frac{1}{c\sqrt{\dfrac{1}{(\Delta x)^2} + \dfrac{1}{(\Delta y)^2} + \dfrac{1}{(\Delta z)^2}}} \tag{11.50}$$

which is the same stability bound as given in (4.54b) of Chapter 4.

11.5 IRREGULAR NONORTHOGONAL STRUCTURED GRIDS

The previous section presented an FDTD algorithm based on a grid defined in a general nonorthogonal curvilinear space. This algorithm assumes the grid is regular and uniform (i.e., that the unitary vectors $\boldsymbol{a}_i$ are constant throughout the space). Practitioners of mesh generation find this assumption to be quite restrictive, in that such a coordinate system can be derived only for specific geometries. In fact, more general techniques have been developed that are based on irregular nonorthogonal structured grids [16–31]. Note that these grids are still structured, in the sense that the grid vertices are described by three-dimensional arrays $x(i, j, k)$, $y(i, j, k)$, and $z(i, j, k)$. However, the unitary vectors can vary from cell to cell. This allows for a more general grid that can conform to highly irregular geometries.

A number of formulations have been introduced based on such grids, especially the class of *finite-volume time-domain* (FVTD) methods developed by Shankar et al. [18–20]. One principal difference between FVTD methods of this class and FDTD is that the former collocate rather than stagger the E- and H-fields in space and time. The coupled curl equations are then solved simultaneously by casting them into conservation form. The general nonorthogonal grid is mapped into a Cartesian grid using a local mapping, based on the assumption that the grid cell can be described by a local curvilinear coordinate system. This permits the physical coordinates to be mapped into a Cartesian coordinate system using the Jacobian of the transformation. Note that because the three components of $\boldsymbol{E}$ or $\boldsymbol{H}$ are collocated in space, the averaging used in (11.33a, b) is no longer necessary. The conservation form of Maxwell's equations is then solved using an explicit solution method such as the Lax-Wendroff upwind scheme.

An alternate technique of assuming a locally curvilinear coordinate system for nonorthogonal structured grids was introduced by Lee et al. [16, 22, 23], and by Fusco [30, 31]. Similar to FVTD, the nonorthogonal grid is mapped into a Cartesian grid through local Jacobian transformations. This is done numerically for a general grid by assuming the cells to be arbitrary parallelepipeds [16], and leads to an explicit time-stepping scheme, which is implemented in a manner similar to the FDTD algorithm presented in the previous section. However, there are subtle yet significant differences between this algorithm and the FVTD technique proposed in [18–20]. As discussed in Chapter 4, it is well-known that FDTD is inherently dispersive, and the same can be shown for Lee's algorithm. But, the FVTD method reported in [18–20] is

inherently *dissipative*. Further, boundary conditions are more naturally enforced using Lee's algorithm, since the fields are associated with the grid edges. Because the FVTD approach associates the fields with the vertices of the grid cells (*E*-fields) and the centroids of the grid cells (*H*-fields), one must enforce boundary conditions using *all three* components of the fields.

The general nonorthogonal FDTD algorithm can be derived in a manner similar to that of the nonorthogonal curvilinear FDTD method. To this end, each cell in the space lattice is approximated as a parallelepiped, as illustrated in Fig 11.9. Unitary vectors $\boldsymbol{A}_i$ are defined as the length vectors bounding the lattice cell originating from a cell vertex. Note that in this formulation, the $\boldsymbol{A}_i$ have the units of length. A reciprocal basis is also introduced. This is related to the unitary basis as

$$\boldsymbol{A}^1 = \frac{\boldsymbol{A}_2 \times \boldsymbol{A}_3}{\sqrt{g}} \; ; \qquad \boldsymbol{A}^2 = \frac{\boldsymbol{A}_3 \times \boldsymbol{A}_1}{\sqrt{g}} \; ; \qquad \boldsymbol{A}^3 = \frac{\boldsymbol{A}_1 \times \boldsymbol{A}_2}{\sqrt{g}} \tag{11.51}$$

where g is the determinant of the metric tensor with the elements

$$g_{i,j} = \boldsymbol{A}_i \cdot \boldsymbol{A}_j \tag{11.52}$$

It can be further shown that

$$\sqrt{g} = \boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_3) \tag{11.53}$$

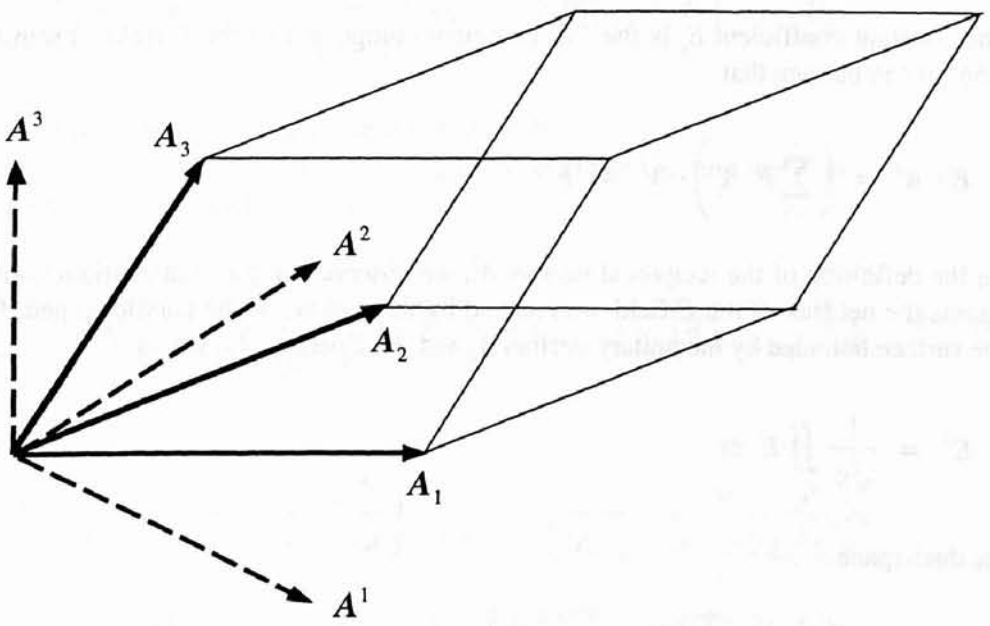


Fig. 11.9 Structured nonorthogonal space-lattice unit cell and the unitary and reciprocal vectors associated with the cell. *After:* Lee et al., *IEEE Trans. Microwave Theory and Techniques*, 1992, pp. 346–352, © 1992 IEEE.

which is equal to the volume of the parallelepiped. The reciprocal vectors $\boldsymbol{A}^i$ are normal to the faces of the parallelepiped bounded by the unitary vectors $\boldsymbol{A}_j$, $\boldsymbol{A}_k$ $(i \neq j, k)$. Further, their amplitudes are equal to the area of the face divided by the volume of the cell. Finally, from (11.51) and (11.53), we derive the relationship

$$\boldsymbol{A}^i \cdot \boldsymbol{A}_j = \delta_{i,j} \tag{11.54}$$

where $\delta_{i,j}$ is the Kronecker delta. The inverse metric is defined where

$$\boldsymbol{A}^i \cdot \boldsymbol{A}^j = g^{i,j} \tag{11.55}$$

and $g^{i,j}$ defines a metric that is the inverse of $g_{i,j}$.

As discussed in Section 11.4 for nonorthogonal curvilinear spaces, $\boldsymbol{E}$ and $\boldsymbol{H}$ are again represented by their contravariant and covariant components. To this end, the E-field is expanded in the unitary vectors as

$$\boldsymbol{E} = \sum_{i=1}^{3} E^i \boldsymbol{A}_i \tag{11.56}$$

where the constant coefficient E^i is the i'th contravariant component of the E-field. The E-field can also be expanded using the reciprocal vectors

$$\boldsymbol{E} = \sum_{i=1}^{3} E_i \boldsymbol{A}^i \tag{11.57}$$

where the constant coefficient E_i is the i'th covariant component of the E-field. From (11.54) and (11.56), it can be seen that

$$\boldsymbol{E} \cdot \boldsymbol{A}^j = \left(\sum_{i=1}^{3} E^i \boldsymbol{A}_i \right) \cdot \boldsymbol{A}^j = E^j \tag{11.58}$$

Based on the definition of the reciprocal vectors $\boldsymbol{A}^j$, we observe that the contravariant component E^j represents the net flux of the E-field, normalized by the volume of the parallelepiped, flowing across the surface bounded by the unitary vectors $\boldsymbol{A}_k$ and $\boldsymbol{A}_i$. Specifically, we have

$$E^j = \frac{1}{\sqrt{g}} \iint_{s_{k,i}} \boldsymbol{E} \cdot d\boldsymbol{s} \tag{11.59}$$

and in the dual-space

$$\boldsymbol{E} \cdot \boldsymbol{A}_j = \left(\sum_{i=1}^{3} E_i \boldsymbol{A}^i \right) \cdot \boldsymbol{A}_j = E_j \tag{11.60}$$

Based on the definition of the unitary vector $\boldsymbol{A}_j$, we observe that the covariant component E_j represents the net flow of the E-field along the lattice-cell edge. Thus,

$$E_j = \int_{C_j} \boldsymbol{E} \cdot d\boldsymbol{L} \tag{11.61}$$

From (11.54) to (11.60), we can derive the relationship between the covariant and contravariant components:

$$E_i = \sum_{j=1}^{3} g_{i,j} E^j \quad ; \qquad E^i = \sum_{j=1}^{3} g^{i,j} E_j \tag{11.62a, b}$$

For this formulation, it is convenient to express the relationship of the E- and H-fields through Maxwell's equations in integral form, specifically using Faraday's law and Ampere's law of (11.6) and (11.7), respectively. Faraday's law is evaluated in the discrete nonorthogonal space by choosing S to be the surfaces of the lattice cells, and C to be the edges bounding the faces. Similarly, Ampere's law is evaluated by choosing S' to be the faces of the dual-lattice cells, and C' to be the dual edges bounding these faces. These integrals are naturally performed using the covariant and contravariant components of the fields. To this end, we use (11.59) and (11.61) to express (11.6) and (11.7) as

$$-\mu \frac{\partial H^i\big|_{i,j,k}}{\partial t} = \frac{1}{\sqrt{g_{i,j,k}}} \left(E_j\big|_{i,j,k} + E_k\big|_{i,j+1,k} - E_j\big|_{i,j,k+1} - E_k\big|_{i,j,k} \right) \tag{11.63}$$

$$\varepsilon \frac{\partial E^i\big|_{i,j,k}}{\partial t} + \sigma E^i\big|_{i,j,k} = \frac{1}{\sqrt{g_{i,j,k}}} \left(H_j\big|_{i,j,k-1} + H_k\big|_{i,j,k} - H_j\big|_{i,j,k} - H_k\big|_{i,j-1,k} \right) \tag{11.64}$$

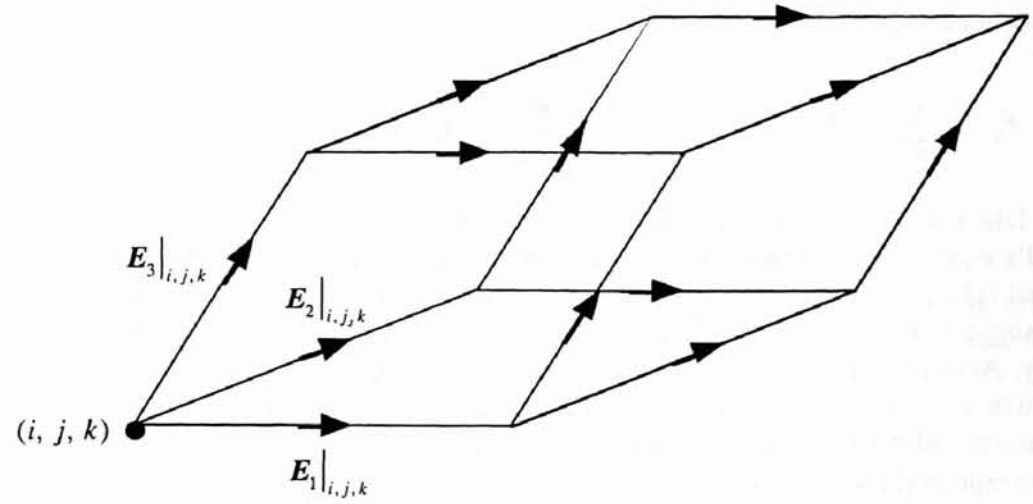
where the i, j, and k indices are referenced to Fig. 11.10.

Central-difference approximations are used to evaluate the time derivatives in (11.63) and (11.64), leading to the explicit updates

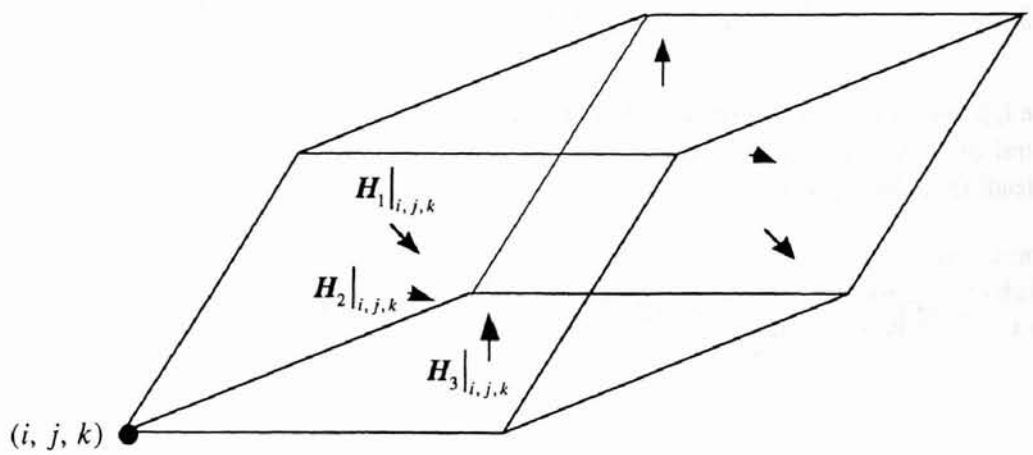
$$H^i\big|_{i,j,k}^{n+1/2} = H^i\big|_{i,j,k}^{n-1/2} - \frac{\Delta t}{\mu \sqrt{g_{i,j,k}}} \left[\left(E_k\big|_{i,j+1,k}^{n} - E_k\big|_{i,j,k}^{n} \right) - \left(E_j\big|_{i,j,k+1}^{n} - E_j\big|_{i,j,k}^{n} \right) \right] \tag{11.65a}$$

$$E^i\big|_{i,j,k}^{n+1} = \left(\frac{2\varepsilon_{i,j,k} - \sigma_{i,j,k} \Delta t}{2\varepsilon_{i,j,k} + \sigma_{i,j,k} \Delta t} \right) E^i\big|_{i,j,k}^{n} + \frac{2\Delta t}{\left(2\varepsilon_{i,j,k} + \sigma_{i,j,k} \Delta t \right) \sqrt{g_{i,j,k}}} \cdot \left[\left(H_k\big|_{i,j,k}^{n+1/2} - H_k\big|_{i,j-1,k}^{n+1/2} \right) - \left(H_j\big|_{i,j,k}^{n+1/2} - H_j\big|_{i,j,k-1}^{n+1/2} \right) \right] \tag{11.65b}$$

The updates of the remaining components of $\boldsymbol{E}$ and $\boldsymbol{H}$ can be derived by permuting the indices i, j, and k.



(a) Covariant electric fields.



(b) Covariant magnetic fields.

Fig. 11.10 Covariant electric and magnetic fields defined within a structured nonorthogonal-lattice unit cell at (i, j, k).

We observe that there is a direct resemblance between (11.65a, b) and the pair of updating equations (11.31) and (11.32). Specifically, if the general nonorthogonal space reduces to a global nonorthogonal curvilinear space, then these equations reduce exactly to the FDTD algorithm. One exception exists in (11.65a, b), where the Δu^i are effectively 1 because the physical edge lengths are incorporated into the unitary vectors, and thus the Δu^i are included in the $\sqrt{g_{i,j,k}}$. However, this formulation is much more general, since the unitary vectors are assumed to be cell-dependent, and the coordinates can be deformed in a general manner.

The right-hand sides of (11.65a, b) involve the covariant components of $\boldsymbol{E}$ and $\boldsymbol{H}$, whereas the components being time-stepped via the explicit update expressions are the contravariant components. Once the contravariant components of the fields are updated, the covariant components must be projected into the contravariant space using (11.62) before the dual field can be updated. This projection requires all three of the contravariant components. However, these components are not known in space at the same locations, and must be averaged in space to maintain second-order accuracy. In the discrete space, the projections are given by

$$H_i\Big|_{i,j,k}^{n+1/2} = g_{i,i}\, H^i\Big|_{i,j,k}^{n+1/2} + \frac{g_{i,j}}{4}\left(H^j\Big|_{i-1,j,k}^{n+1/2} + H^j\Big|_{i-1,j+1,k}^{n+1/2} + H^j\Big|_{i,j,k}^{n+1/2} + H^j\Big|_{i,j+1,k}^{n+1/2}\right)$$
$$+ \frac{g_{i,k}}{4}\left(H^k\Big|_{i-1,j,k}^{n+1/2} + H^k\Big|_{i-1,j,k+1}^{n+1/2} + H^k\Big|_{i,j,k}^{n+1/2} + H^k\Big|_{i,j,k+1}^{n+1/2}\right) \tag{11.66a}$$

$$E_i\Big|_{i,j,k}^{n+1} = g_{i,i}\, E^i\Big|_{i,j,k}^{n+1} + \frac{g_{i,j}}{4}\left(E^j\Big|_{i,j,k}^{n+1} + E^j\Big|_{i,j-1,k}^{n+1} + E^j\Big|_{i+1,j,k}^{n+1} + E^j\Big|_{i+1,j-1,k}^{n+1}\right)$$
$$+ \frac{g_{i,k}}{4}\left(E^k\Big|_{i,j,k}^{n+1} + E^k\Big|_{i,j,k-1}^{n+1} + E^k\Big|_{i+1,j,k}^{n+1} + E^k\Big|_{i+1,j,k-1}^{n+1}\right) \tag{11.66b}$$

where the metrics $g_{i\,i}$, $g_{i,j}$, and $g_{i,k}$ are calculated from cell (i, j, k). The remaining field components can be computed by permuting the i, j, k indices.

Finally, (11.65a, b) and (11.66a, b) define the explicit time-stepping scheme for the contravariant E- and H-fields using a general nonorthogonal lattice. Once the contravariant fields are computed, they can be mapped back to the physical space as

$$E^i = \sqrt{g_{i,i}}\, E^i \quad ; \qquad H^i = \sqrt{g_{i,i}}\, H^i \tag{11.67a, b}$$

where the direction of the field is along the unit vector of the reciprocal vectors (i.e., $\boldsymbol{A}^i / |\boldsymbol{A}^i|$).

The explicit update scheme in (11.65a, b) and (11.66a, b) is conditionally stable. However, due to the irregularity of the space lattice, a stability criterion cannot be derived in the global sense as was presented in Section 11.4, but rather is derived in a more local sense. Again, a growth-factor $\alpha = |E^{n+1}| / |E^n|$ is defined. Due to the explicit nature of the algorithm, α is not defined by the vector norm $\|E^n\|$, but rather is defined for each edge in the lattice. As before, numerical stability requires that all $\alpha \leq 1$. Following the analogy presented in Section 11.4, we can thus derive the stability criterion [16]

$$\Delta t \leq \frac{1}{c \sup\left(\sqrt{\sum_{l=1}^{3}\sum_{m=1}^{3} g^{l,m}}\right)} \tag{11.68}$$

where sup(*) denotes the maximum value throughout the i, j, k space. This implies the maximum time-step for stability is determined either by the smallest lattice cell, or by the possible presence in the lattice of highly oblique parallelepiped cells (i.e., parallelepipeds with very large interior angles become wedge-shaped).

A number of applications have been studied using the general nonorthogonal FDTD method, and the reader is referred to [16, 22–27] for further study. A principal advantage of this technique is that in the discrete space, all quantities can be described using a three-dimensional array. This eases the complexity of programming and the amount of memory overhead. Further, there are commercially available mesh generators that can automatically generate nonorthogonal structured space lattices.

However, many geometries are difficult to model using structured meshes due to high rates of surface curvature, fine features, and highly complex shapes. In such cases, a grid can develop undesirable characteristics that degrade accuracy and numerical stability, such as very large cell aspect ratios or extremely small cell edge lengths. An alternative for such problems is to use the more general unstructured gridding method, which is the topic of the next section.

11.6 IRREGULAR NONORTHOGONAL UNSTRUCTURED GRIDS

Unstructured grids are composed of arrays of general fitted polyhedral cells. While every cell shares a common face and edge, the vertices of the grid can no longer be listed in a regular sense. These meshes can conform to highly complex shapes while maintaining good cell aspect ratios and global uniformity. Unfortunately, the methods discussed in the previous section are not applicable to unstructured grids, and more general approaches must be used.

A general method to formulate an explicit time-marching solution of Maxwell's equations using unstructured grids was introduced by Madsen [32]. This was founded upon deriving an explicit operator based on Maxwell's equations in integral form, and evaluating the discrete surface and line integrals. A key need was to introduce a projection scheme similar to (11.66a, b), which projects the face-normal components of the fields onto the dual-grid edges. Due to the unstructured nature of the lattice, the precise approach of (11.66a, b) could not be implemented, and more general methods were investigated. It was shown that insufficient accuracy of the projection method could cause the numerical solution to become unstable after a very large number of time-steps. This problem was overcome, and a numerically stable, second-order-accurate projection of the normal fields onto the lattice edges was reported. This was later incorporated into other algorithms for microstrip line applications by Gedney and Lansing [33–36].

The class of algorithms presented in [32–36] provides an efficient and accurate means for treating three-dimensional electromagnetic wave problems based on unstructured grids, assuming that a second-order-accurate solver is used. This class is referred to here as the generalized Yee algorithm, since analysis reveals that it is truly a straightforward extension of Yee's classic space-time meshing approach [37].

11.6.1 Generalized Yee Algorithm

We now consider the extension of the Yee concept of interleaved E- and H-fields to a dual-space lattice formed by a primary and a secondary mesh. The primary mesh is composed of space-filling polyhedra, while the secondary (or dual) mesh is composed of the closed polyhedra whose edges connect the centroids of adjacent primary cells, thereby penetrating shared faces. Fig. 11.11 illustrates such an arrangement wherein adjoining primary and secondary cells are hexahedrons.

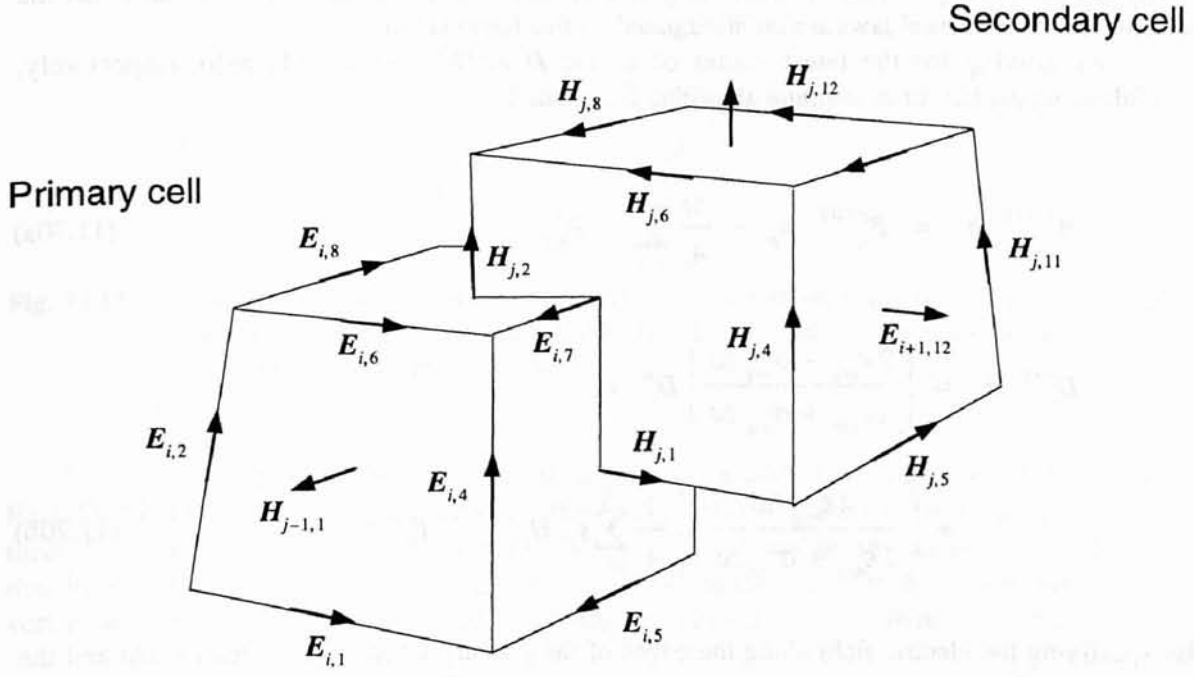


Fig. 11.11 A generalized unstructured Yee lattice using hexahedral cells.

In Fig. 11.11, an E-component is defined along each edge of a primary lattice cell, and an H-component is located along each secondary lattice cell edge. Faraday's law (11.6) and Ampere's law (11.7) are implemented by discretizing the surface and line integrals over each primary and secondary lattice cell face, respectively. The fields are assumed to be constant over each face and along each edge. Approximating the time derivative using a central difference, we obtain

$$\sum_{j=1}^{N_{e_i}} \boldsymbol{p}_j \cdot \boldsymbol{E}_{i,j}^{n} = -A_i \left(\frac{\boldsymbol{B}_i^{n+1/2} - \boldsymbol{B}_i^{n-1/2}}{\Delta t} \right) \cdot \hat{\boldsymbol{n}}_p \tag{11.69a}$$

$$\sum_{j=1}^{N_{e_i}} \boldsymbol{s}_j \cdot \boldsymbol{H}_{i,j}^{n+1/2} = A_i \left(\frac{\boldsymbol{D}_i^{n+1} - \boldsymbol{D}_i^n}{\Delta t} \right) \cdot \hat{\boldsymbol{n}}_s + \sigma_{\text{avg}} A_i \left(\frac{\boldsymbol{D}_i^{n+1} + \boldsymbol{D}_i^n}{2\varepsilon_{\text{avg}}} \right) \cdot \hat{n}_s + A_i \boldsymbol{J}_i^{n+1/2} \cdot \hat{\boldsymbol{n}}_s \tag{11.69b}$$

In (11.69), the superscripts denote the time index; N_{e_i} is the number of edges bounding the i'th face of the primary lattice in (11.69a) or the secondary lattice in (11.69b); A_i is the area of the i'th face; $\boldsymbol{p}_j$ and $\boldsymbol{s}_j$ are the length vectors of the j'th edge bounding the primary and secondary lattice faces, respectively; and ε_{avg} and σ_{avg} are the average material properties, to be discussed later. By performing a closed surface integral over each cell, it will also be shown later that the discrete forms of Gauss' laws are satisfied based on this formulation.

Upon solving for the latest values of $\boldsymbol{B}$ and $\boldsymbol{D}$ in (11.69a) and (11.69b), respectively, the following explicit time-stepping algorithm is obtained:

$$\boldsymbol{B}_i^{n+1/2} \cdot \hat{\boldsymbol{n}}_p = \boldsymbol{B}_i^{n-1/2} \cdot \hat{\boldsymbol{n}}_p - \frac{\Delta t}{A_i} \sum_{j=1}^{N_{e_i}} \boldsymbol{p}_j \cdot \boldsymbol{E}_{i,j}^n \tag{11.70a}$$

$$\boldsymbol{D}_i^{n+1} \cdot \hat{\boldsymbol{n}}_s = \left(\frac{2\varepsilon_{\text{avg}} - \sigma_{\text{avg}} \Delta t}{2\varepsilon_{\text{avg}} + \sigma_{\text{avg}} \Delta t} \right) \boldsymbol{D}_i^n \cdot \hat{\boldsymbol{n}}_s + \left(\frac{2\varepsilon_{\text{avg}} \Delta t}{2\varepsilon_{\text{avg}} + \sigma_{\text{avg}} \Delta t} \right) \left(\frac{1}{A_i} \sum_{j=1}^{N_{e_i}} \boldsymbol{s}_j \cdot \boldsymbol{H}_{i,j}^{n+1/2} - \boldsymbol{J}_i^{n+1/2} \cdot \hat{\boldsymbol{n}}_s \right) \tag{11.70b}$$

By specifying the electric field along the edges of the primary lattice cell at time $t = n\Delta t$ and the magnetic field normal to each face at time $t = (n-1/2)\Delta t$, the normal magnetic flux density at time $t = (n+1/2)\Delta t$ can be updated using (11.70a). Then, $\boldsymbol{H} = \boldsymbol{B}/\mu$ can be used to update the electric flux density in (11.70b).

If the fields are discretized over a regular orthogonal lattice, (11.70a, b) are equivalent to the Yee algorithm (see Problem 11.1). This holds because, in this case, $\hat{\boldsymbol{n}}_p = \hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{n}}_s = \hat{\boldsymbol{p}}$. However, as shown in Fig. 11.12, if the lattice is unstructured, then the edge vector is *not* necessarily parallel to the normal vector, and the normal field must be projected onto the complementary lattice edge. Since the normal field alone cannot uniquely describe the edge component, a secondary expression must be introduced.

The projection of the magnetic field onto the secondary lattice edge requires an interpolation of the local magnetic flux densities normal to adjacent faces. This is similar to what was required for the nonorthogonal structured lattice projection performed by (11.66a). However, due to the unstructured nature of the lattice, the adjacent faces are randomly oriented and numbered. Therefore, a more general projection scheme is required. It must be chosen such that: (1) the flux projected onto the edges has zero divergence in a charge-free medium, and (2) the time-stepping algorithm maintains numerical stability. The projection scheme of Madsen in [32] and slightly modified in [33, 34, 38] meets these criteria.

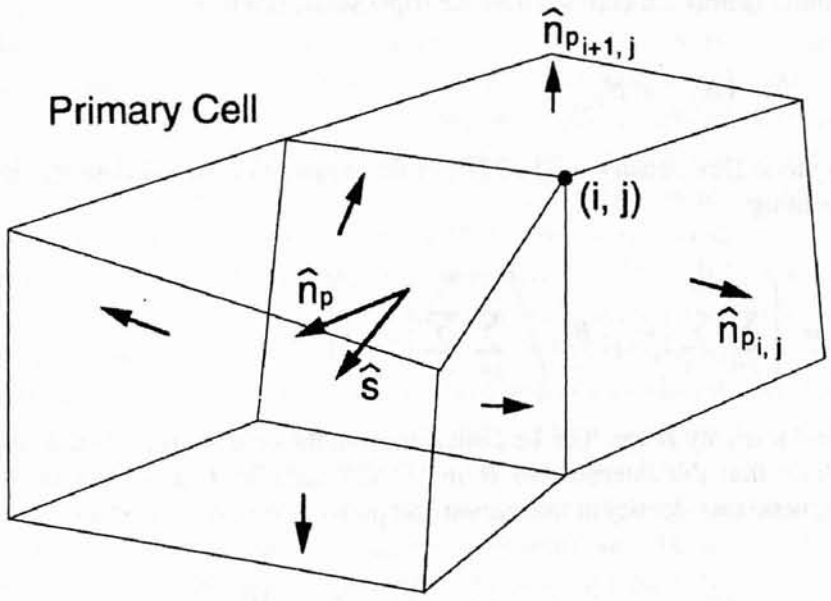


Fig. 11.12 Adjacent primary lattice cells sharing a common face with the magnetic fields normal to each face known. The face is penetrated by the edge of a dual cell directed along the unit vector $\hat{s}$, which is not necessarily parallel to $\hat{n}_p$.

Referring to Fig. 11.12, we assume that each face is shared by N_c cells, where $N_c = 2$ or 1. Each face is also bounded by N_e edges that connect N_e vertices, where the i'th vertex is shared by three faces of the j'th cell. Equation (11.70a) is used to update the normal magnetic flux densities passing through each face. Then, the magnetic flux density associated with the i'th vertex and the j'th cell can be computed by solving the following 3×3 system of equations:

$$\boldsymbol{B}_{i,j} \cdot \boldsymbol{N}_P = \boldsymbol{B} \cdot \boldsymbol{N}_P \tag{11.71a}$$

$$\boldsymbol{B}_{i,j} \cdot \boldsymbol{N}_{P_{i,j}} = \boldsymbol{B} \cdot \boldsymbol{N}_{P_{i,j}} \tag{11.71b}$$

$$\boldsymbol{B}_{i,j} \cdot \boldsymbol{N}_{P_{i+1,j}} = \boldsymbol{B} \cdot \boldsymbol{N}_{P_{i+1,j}} \tag{11.71c}$$

where $\boldsymbol{N}_P = A_i \hat{\boldsymbol{n}}_p$ is the area vector normal to the i'th primary lattice face. Since the right-hand side is known from (11.70a), then (11.71) is used to solve for the three orthogonal components of $\boldsymbol{B}_{i,j}$. Subsequently, this is performed for each of the vertices of the face ($i = 1, N_e$), and for each cell ($j = 1, N_c$) shared by the face. At this point, we note that $\boldsymbol{B}_{i,j}$ is not an interpolation for the total field, but rather a local value associated with the (i, j)'th corner shared by the face.

The magnetic flux density vector over the face is expressed by the interpolation of the local field values as

$$\boldsymbol{B} = \sum_{j=1}^{N_c} \sum_{i=1}^{N_e} |w_{i,j}| \, \boldsymbol{B}_{i,j} \Bigg/ \sum_{j=1}^{N_c} \sum_{i=1}^{N_e} |w_{i,j}| \tag{11.72}$$

where the weighting factors are computed by the triple scalar product

$$w_{i,j} = \boldsymbol{N}_P \cdot \left(\boldsymbol{N}_{P_{i,j}} \times \boldsymbol{N}_{P_{i+1,j}} \right) \tag{11.73}$$

Finally, the magnetic flux density in (11.72) can be uniquely projected onto the secondary cell edge by implementing

$$\boldsymbol{B} \cdot \hat{\boldsymbol{s}} = \left(\sum_{j=1}^{N_c} \sum_{i=1}^{N_{e_j}} \left| w_{i,j} \right| \boldsymbol{B}_{i,j} \Big/ \sum_{j=1}^{N_c} \sum_{i=1}^{N_{e_j}} \left| w_{i,j} \right| \right) \cdot \hat{\boldsymbol{s}} \tag{11.74}$$

The magnetic field intensity $\boldsymbol{H}$ can then be evaluated from the constitutive relationship.

We now show that the interpolated $\boldsymbol{B}$ in (11.72) satisfies Gauss' law in discrete form. Consider the magnetic flux density at the current and previous time-steps. Then:

$$\frac{1}{\Delta t} \left(\int_V \nabla \cdot \boldsymbol{B}^{n+1/2} \, dV - \int_V \nabla \cdot \boldsymbol{B}^{n-1/2} \, dV \right) = \int_V \nabla \cdot \frac{\partial \boldsymbol{B}^n}{\partial t} dV \tag{11.75}$$

where V is the volume of the primary cell. From the Divergence Theorem and Faraday's law, the right-hand side of (11.75) is expressed as

$$\begin{aligned} \int_V \nabla \cdot \frac{\partial \boldsymbol{B}^n}{\partial t} dV &= \oiint_S \frac{\partial \boldsymbol{B}^n}{\partial t} \cdot \boldsymbol{ds} \\ &= \sum_{i=1}^{N_f} \left(\iint_{S_i} \frac{\partial \boldsymbol{B}^n}{\partial t} \cdot \boldsymbol{ds}_i \right) = - \sum_{i=1}^{N_f} \left(\oint_{C_i} \boldsymbol{E}^n \cdot \boldsymbol{dL}_i \right) = 0 \end{aligned} \tag{11.76}$$

where S is the surface bounding the cell volume V, N_f is the number of faces of the cell, S_i is the i'th face of the cell, and C_i is the contour bounding S_i. This expression is zero since the edges of the cell are traversed twice in opposing directions when evaluating the line integral. Since this is true for each $\boldsymbol{B}$, the argument can be extended for the $\boldsymbol{B}$ appearing in (11.72). Finally, assuming that the initial distribution of $\boldsymbol{B}$ has zero divergence throughout the volume, then from (11.75) and (11.76) we have

$$\oiint_S \boldsymbol{B}^{n+1/2} \cdot \boldsymbol{ds} = 0 \tag{11.77}$$

and the numerical fields have zero divergence, as well.

By duality, the electric flux density is projected onto the primary lattice edges using

$$\boldsymbol{D} \cdot \hat{\boldsymbol{p}} = \left(\sum_{j=1}^{N_c} \sum_{i=1}^{N_{e_j}} \left| w_{i,j} \right| \boldsymbol{D}_{i,j} \Big/ \sum_{j=1}^{N_c} \sum_{i=1}^{N_{e_j}} \left| w_{i,j} \right| \right) \cdot \hat{\boldsymbol{p}} \tag{11.78}$$

where the weighting coefficients are $w_{i,j} = \boldsymbol{N}_s \cdot (\boldsymbol{N}_{S_{i,j}} \times \boldsymbol{N}_{S_{i+1,j}})$ and the $\boldsymbol{N}_S$ are the normal area vectors of the secondary lattice faces. It can be shown that the electric flux density in (11.78) also satisfies the discrete form of Gauss' law. Finally, the field updates can be performed by computing the vector fields normal to the primary or secondary lattice faces using (11.70a) or (11.70b), and then projecting them onto the edges of the secondary or primary lattices using (11.74) or (11.78), respectively. By introducing the correction terms in (11.74) and (11.78), it can be shown that second-order accuracy of the algorithm is maintained.

We recall that explicit time-stepping is conditionally stable, and the time-step Δt must be bounded to guarantee numerical stability. Due to the unstructured and irregular nature of this lattice, a von Neumann stability analysis similar to that presented in Section 11.4 would be extremely difficult to employ. A rigorous analysis of stability would calculate the eigenspectrum of the linear operations involved in the combined update expressions (11.70a, b), (11.74), and (11.78). However, these expressions are problem-dependent. The eigenvalue analysis would have to be conducted for each modeling case, a procedure that at the minimum would be computationally intensive, and at the maximum, prohibitive. However, through numerical experimentation, a very simple relationship has been established that provides an excellent estimate of the stability bound for tetrahedral, pentahedral, and hexahedral elements [34]:

$$\Delta t < \frac{1}{c \sup\left(\sqrt{\sum_{i=1}^{3} \frac{1}{L_i^2}}\right)} \tag{11.79}$$

Here, c is the speed of light and L_i ($i = 1, 2, 3$) are lengths of three edges in each cell sharing a common vertex. This expression provides an estimate for the maximum stable time-step that is likely within 10% of the actual upper bound. We see from (11.79) that the stability bound is determined primarily by the smallest edge length in the lattice, as was found earlier with the nonorthogonal structured-grid FDTD algorithm.

The advantage of the above formulation is that it is based on general unstructured lattices. Space cells can be any polyhedron (assuming that each face is uniquely shared by, at most, two polyhedra), and the number of polyhedra shared by each edge is arbitrary. This simplifies modeling complex three-dimensional geometries, and is well-suited for commercially available mesh-generation software.

There are some disadvantages, however. Since the lattice is unstructured and irregular, regular indexing such as was used for the structured nonorthogonal grid case in (11.64) and (11.65) cannot be used to store the lattice vertices and the fields. Rather, a more general storage scheme is needed. This is discussed later.

11.6.2 Inhomogeneous Media

The generalized Yee algorithm presented in the previous section is applicable for modeling electromagnetic fields in arbitrary, isotropic, dielectric and conducting media. However, it is assumed that the medium within each primary lattice cell is homogeneous. We now consider the inhomogeneous medium case.

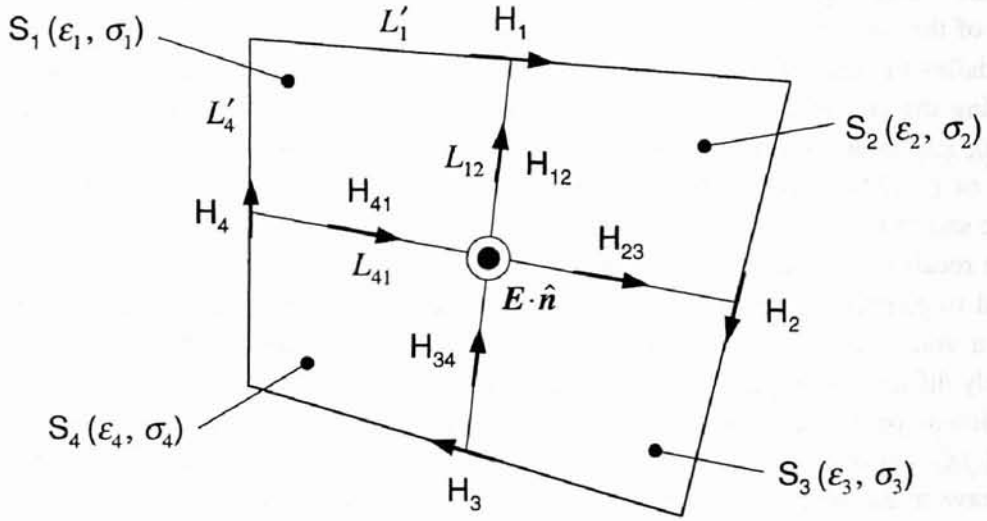


Fig. 11.13 Secondary cell face within an inhomogeneous medium.

Fig. 11.13 illustrates a quadrilateral face of a secondary cell of an irregular and unstructured lattice. The face is defined as the surface S bounded by the contour C. For the inhomogeneous medium case, we assume that the primary edge passing through the centroid of the face is tangential to the interface of a boundary shared by four materials characterized by $(\varepsilon_i, \sigma_i)$, $i = 1, 2, 3, 4$. This is inherently the case, since the edge defines the boundary of four primary lattice cells. Ampere's law must be valid over surface S. However, the permittivity of the face is ambiguous as referenced to the electric field passing through its centroid. Consequently, S is decomposed into four distinct surfaces S_i, each with the constant properties $(\varepsilon_i, \sigma_i)$.

Ampere's law can now be expressed for the geometry of Fig. 11.13 as

$$\sum_{i=1}^{4}\left(\frac{\partial}{\partial t}\iint_{S_i}\varepsilon_i\,\boldsymbol{E}\cdot\boldsymbol{ds}_i + \iint_{S_i}\sigma_i\,\boldsymbol{E}\cdot\boldsymbol{ds}_i\right) = \sum_{i=1}^{4}\left(\oint_{C_i}\boldsymbol{H}\cdot\boldsymbol{dL}_i\right) \tag{11.80}$$

where the contour C_i bounds S_i. The integrals are then performed over each S_i and the corresponding C_i. For example, in S_1 we have

$$\varepsilon_1 A_1\left(\frac{\boldsymbol{E}_1^{n+1} - \boldsymbol{E}_1^{n}}{\Delta t}\right)\cdot\hat{\boldsymbol{n}} + \sigma_1 A_1\left(\frac{\boldsymbol{E}_1^{n+1} + \boldsymbol{E}_1^{n}}{2}\right)\cdot\hat{\boldsymbol{n}}$$

$$= \left(L_1' H_1^{n+1/2} - L_{12} H_{12}^{n+1/2} - L_{41} H_{41}^{n+1/2} + L_4' H_4^{n+1/2}\right) \tag{11.81}$$

where $\boldsymbol{E}_1$ is the electric field in S_1, and $\hat{\boldsymbol{n}}$ is the unit normal. Similar expressions can be derived for the remaining three regions.

With the assumption that $\hat{\boldsymbol{n}}$ is tangential to the boundary interface, the requirement of continuity of the tangential $\boldsymbol{E}$ across an interface yields

$$\boldsymbol{E}_1 \cdot \hat{\boldsymbol{n}} \cong \boldsymbol{E}_2 \cdot \hat{\boldsymbol{n}} \cong \boldsymbol{E}_3 \cdot \hat{\boldsymbol{n}} \cong \boldsymbol{E}_4 \cdot \hat{\boldsymbol{n}} \cong \boldsymbol{E} \cdot \hat{\boldsymbol{n}} \tag{11.82}$$

Assuming that the edges extending above or below the boundary interface are close to normal, we add the four equations implied by (11.80). This results in

$$\begin{aligned} &\left(\varepsilon_1 A_1 + \varepsilon_2 A_2 + \varepsilon_3 A_3 + \varepsilon_4 A_4\right)\left(\frac{\boldsymbol{E}^{n+1} - \boldsymbol{E}^n}{\Delta t}\right) \cdot \hat{\boldsymbol{n}} \\ &\quad + \left(\sigma_1 A_1 + \sigma_2 A_2 + \sigma_3 A_3 + \sigma_4 A_4\right)\left(\frac{\boldsymbol{E}^{n+1} + \boldsymbol{E}^n}{2}\right) \cdot \hat{\boldsymbol{n}} \\ &\quad = \left(\boldsymbol{L}_1 \cdot \boldsymbol{H}_1^{n+1/2} + \boldsymbol{L}_2 \cdot \boldsymbol{H}_2^{n+1/2} + \boldsymbol{L}_3 \cdot \boldsymbol{H}_3^{n+1/2} + \boldsymbol{L}_4 \cdot \boldsymbol{H}_4^{n+1/2}\right) \end{aligned} \tag{11.83}$$

This permits definition of the weighted average permittivity and conductivity:

$$\varepsilon_{\text{avg}} = \left(\varepsilon_1 A_1 + \varepsilon_2 A_2 + \varepsilon_3 A_3 + \varepsilon_4 A_4\right) / A \tag{11.84a}$$

$$\sigma_{\text{avg}} = \left(\sigma_1 A_1 + \sigma_2 A_2 + \sigma_3 A_3 + \sigma_4 A_4\right) / A \tag{11.84b}$$

where A is the total face area. For a face bounded by an arbitrary number of edges, we find

$$\varepsilon_{\text{avg}} = \sum_{i=1}^{N_e} A_i \varepsilon_i \Bigg/ \sum_{i=1}^{N_e} A_i \quad ; \qquad \sigma_{\text{avg}} = \sum_{i=1}^{N_e} A_i \sigma_i \Bigg/ \sum_{i=1}^{N_e} A_i \tag{11.85a, b}$$

where A_i is the area of subregion S_i. This leads directly to (11.70b).

We note that this development can be applied to the structured-grid FDTD algorithms presented in the previous sections. Equation (11.84a, b) can be used to compute the effective dielectric constants at inhomogeneous dielectric interfaces in (11.32b) and (11.65b).

11.6.3 Practical Implementation of the Generalized Yee Algorithm

The numerical implementation of (11.70a, b), (11.74), and (11.78) could have a large computer burden due to its calculation and/or storage of the area of each cell face, unit normal, edge vector, and weighting factor. However, if treated properly, the algorithm can be implemented in an efficient manner. To this end, the generalized Yee algorithm can be thought of as a series of linear operations acting on the fields that are normal to the primary and secondary lattice faces. These operations can be expressed in their discrete form as matrix multiplications. For example, the explicit time-stepping relation of (11.70a) can be written as the linear equation

$$\{b_n\} = \{b_n\} - \left[A_h\right]\{e^p\} \tag{11.86}$$

where $\{b_n\}$ is a one-dimensional array representing the amplitudes of the magnetic flux density vectors normal to the lattice faces, and $\{e^p\}$ is a one-dimensional array representing the amplitudes of the electric field vectors projected onto the primary lattice edges. As can be seen from (11.70a), matrix $[A_h]$ is sparse, so that we need to store only a few nonzero entries in each row. Further, the update in (11.86) requires far fewer floating-point operations than (11.70a), leading to a much more computationally efficient algorithm.

The array $\{e^p\}$ must be computed from the array of normal electric flux-density vectors $\{d_n\}$ using (11.78). Again, this is a linear operation and can be expressed as

$$\{e^p\} = \left[A_{pe}\right]\{d_n\} \tag{11.87}$$

Matrix $[A_{pe}]$ is also very sparse, since it involves only local field interpolation. Further, in regions of orthogonality, the corresponding blocks of $[A_{pe}]$ are simply the identity matrix and need not be stored. Now, combining (11.86) and (11.87), the update of $\{b_n\}$ can be written as

$$\{b_n\} = \{b_n\} - \left[A_h\right]\left[A_{pe}\right]\{d_n\} \tag{11.88}$$

Of course, the matrices never need to be multiplied. Rather, each update simply requires two matrix-vector products. In a similar manner, starting with (11.70b) and (11.74), the time-stepping relation for the array of normal electric flux densities can be written as

$$\{d_n\} = \left[D_e\right]\{d_n\} - \left[A_e\right]\left[A_{ph}\right]\{b_n\} \tag{11.89}$$

where $[D_e]$ is a diagonal matrix.

The sparse matrices in (11.88) and (11.89) need to be constructed only once and then stored. This can be accomplished efficiently on a cell-by-cell basis [39]. To this end, the dual-lattice is never actually constructed. In fact, there is sufficient information to construct the matrices associated with the dual-lattice from the primary lattice alone. Upon completion of the matrix construction, we can begin a computationally efficient time-stepping process based upon the above matrix-vector multiplications.

11.7 A PLANAR GENERALIZED YEE ALGORITHM

The memory requirements of the generalized Yee algorithm can be greatly reduced by exploiting symmetries in the model. The focus of this section is on geometries such as microwave printed circuits having planar symmetry, namely three-dimensional geometries that can be uniquely specified by a projection onto a two-dimensional transverse plane [36]. Here, the entire circuit geometry can be described by an unstructured two-dimensional grid in the transverse plane, and by a regular grid in the third dimension. However, only the cell geometry of the two-dimensional grid needs to be stored. Thus, the sparse time-stepping matrices in (11.86) and (11.87) need be stored only for a single layer of grid cells. This relaxes the memory requirements of the algorithm to the point where it is as memory-efficient as the basic FDTD method.

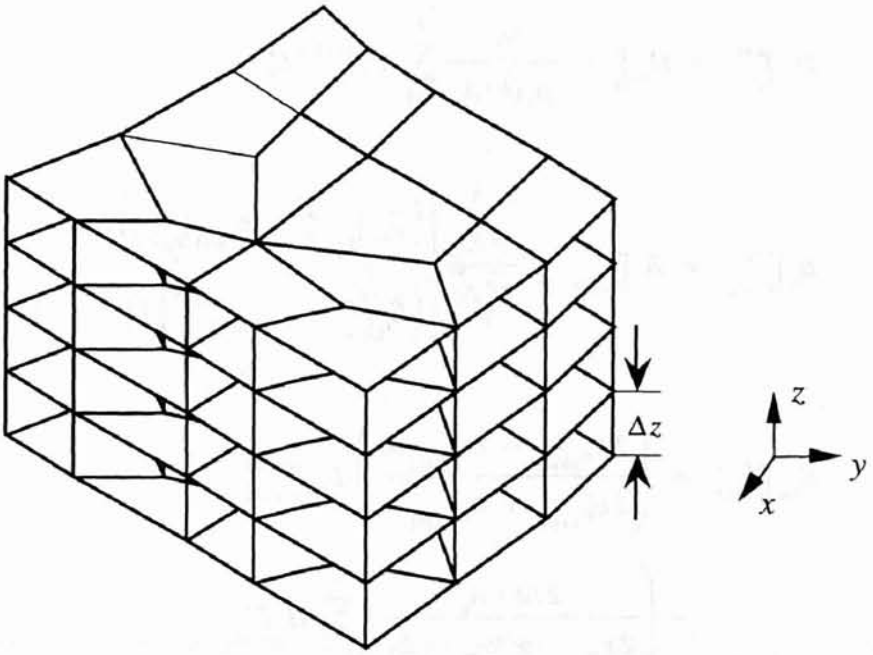


Fig. 11.14 An example of a primary PGY lattice composed of similar two-dimensional unstructured grids stacked in the vertical z-direction in a regular sense. *Source:* Gedney and Lansing, *International Journal for Numerical Modeling (Electronics Networks, Devices, and Fields)*, 1995, pp. 249–264.

Fig. 11.14 illustrates the gridding used for this *planar generalized Yee* (PGY) algorithm. The primary PGY grid is generated simply by extruding a two-dimensional unstructured grid in the vertical (z) direction, and then segmenting it at discrete heights. A secondary grid is staggered within the primary grid such that its vertices lie at the centroids of the primary grid cells, and the edges of the secondary grid connect the centroids by passing through the faces of the primary grid. The E- and H-fields are decomposed into transverse and vertical components

$$\boldsymbol{E} = \boldsymbol{E}_t + \hat{z} E_z \; ; \qquad \boldsymbol{H} = \boldsymbol{H}_t + \hat{z} H_z \tag{11.90}$$

Subsequently, the transverse E- and H-fields are mapped onto the horizontal edges of the primary and secondary grids, respectively; and the vertical E- and H-fields are mapped onto the vertical edges of the primary and secondary grids, respectively. These fields are assumed to be constant along their respective edge lengths, as well as over the dual-faces through which they pass.

11.7.1 Time-Stepping Expressions

Based on the above discretization, Faraday's and Ampere's laws are approximated by choosing the surfaces of integration to be the faces of the secondary and primary grids, respectively. The time derivative is then approximated using a central-difference expression that is second-order-accurate if the fields are staggered in time. This leads to the following explicit time-stepping expressions:

$$H_{z_j}\Big|_k^{n+1} = H_{z_j}\Big|_k^n - \frac{\Delta t}{\mu_j(k)\,A_p}\sum_{i=1}^{N_{p_j}} E_{t_i}\Big|_k^{n+1/2} L_i^p \tag{11.91}$$

$$B_{t_i}\Big|_{k+1/2}^{n+1} = B_{t_i}\Big|_{k+1/2}^{n} - \frac{c\,\Delta t}{L_i^p\,\Delta z}\left[\begin{array}{l}\left(E_{z_m}\Big|_{k+1/2}^{n+1/2} - E_{z_{m+1}}\Big|_{k+1/2}^{n+1/2}\right)\Delta z + \\ \left(E_{t_i}\Big|_{k+1}^{n+1/2} - E_{t_i}\Big|_{k}^{n+1/2}\right) L_i^p\end{array}\right] \tag{11.92}$$

$$\begin{aligned} E_{z_m}\Big|_{k+1/2}^{n+3/2} &= \left(\frac{2\varepsilon_{m_{k+1/2}} - \sigma_{m_{k+1/2}}\Delta t}{2\varepsilon_{m_{k+1/2}} + \sigma_{m_{k+1/2}}\Delta t}\right) E_{z_m}\Big|_{k+1/2}^{n+1/2} \\ &\quad - \left(\frac{2\Delta t / A_s}{2\varepsilon_{m_{k+1/2}} + \sigma_{m_{k+1/2}}\Delta t}\right)\sum_{i=1}^{N_{s_j}} H_{t_i}\Big|_{k+1/2}^{n+1} L_i^s \end{aligned} \tag{11.93}$$

$$\begin{aligned} D_{t_i}\Big|_k^{n+3/2} &= \left(\frac{2\varepsilon_{i_{k+1/2}} - \sigma_{i_{k+1/2}}\Delta t}{2\varepsilon_{i_{k+1/2}} + \sigma_{i_{k+1/2}}\Delta t}\right) D_{t_i}\Big|_k^{n+1/2} \\ &\quad - \left(\frac{2\varepsilon_{i_{k+1/2}}\,\Delta t \,/\, L_i^s\,\Delta z}{2\varepsilon_{i_{k+1/2}} + \sigma_{i_{k+1/2}}\Delta t}\right)\left[\begin{array}{l}\left(H_{z_j}\Big|_k^{n+1} - H_{z_{j+1}}\Big|_k^{n+1}\right)\Delta z + \\ \left(H_{t_i}\Big|_{k-1/2}^{n+1} - H_{t_i}\Big|_{k+1/2}^{n+1}\right) L_i^s\end{array}\right] \end{aligned} \tag{11.94}$$

where D_t and B_t are the electric and magnetic flux densities, respectively, in the transverse plane; k is the index along the z-direction; A_p and A_s are the areas of the primary and secondary grid faces, respectively; N_{p_i} and N_{s_j} are the number of edges bounding the i'th primary and the j'th secondary grid faces, respectively; and L_i^p and L_i^s are the lengths of the primary and secondary grid edges, respectively. The material parameters ε, μ, and σ are assumed to be piecewise homogeneous in both the z-direction as well as the transverse direction, and can be computed using the methods discussed earlier.

11.7.2 Projection Operators

Note that the flux densities updated in (11.92) and (11.94) are normal to the faces. However, the corresponding field intensities on the dual-edges passing through these faces are not necessarily normal to the faces. As a result, the flux densities must be projected onto the edges before the dual-fields can be updated. This requires the introduction of an auxiliary projection operator. The projection operators used in Section 11.6 are also used here. The principal difference is that, due to the orthogonality of the vertical and transverse fields, only fields within the transverse plane are needed for the interpolation. This results in a simplified form of (11.74) and (11.78).

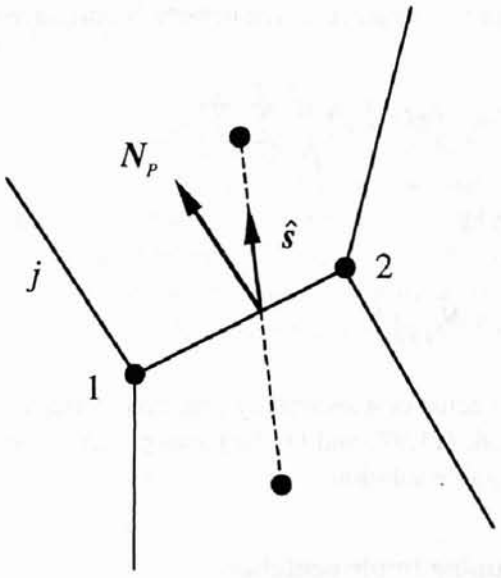


Fig. 11.15 Normal vector to a transverse primary face in the PGY grid, and a dual-edge passing through the face.

Consider the geometry of Fig. 11.15. Here, $\boldsymbol{N}_P$ is the normal area vector of a primary grid transverse face, and $\hat{s}$ is the unit vector along the dual-grid edge passing through the face. Note that the face is associated with an edge in the transverse plane. From (11.92), $\boldsymbol{B} \cdot \boldsymbol{N}_p$ is known at each face and provides the basis of the interpolation. We introduce a general magnetic flux-density vector $\boldsymbol{B}$, which is to be determined from the interpolation. The edge corresponding to the face is bounded by vertices 1 and 2, identified by the index $i = 1, 2$ in Fig. 11.15. Each vertex is also shared by two additional edges that share a common cell. Let j represent one of these edges, where $j = 1, 2$. The normal area vector to the j'th edge associated with the i'th vertex is $\boldsymbol{N}_{P_{i,j}}$.

Now, we define $\boldsymbol{B}_{i,j}$ to be the local estimate of the magnetic flux vector associated with the i'th vertex and the j'th edge. $\boldsymbol{B}_{i,j}$ is computed by solving the system

$$\boldsymbol{B}_{i,j} \cdot \boldsymbol{N}_P = \boldsymbol{B} \cdot \boldsymbol{N}_P \; ; \qquad \boldsymbol{B}_{i,j} \cdot \boldsymbol{N}_{P_{i,j}} = \boldsymbol{B} \cdot \boldsymbol{N}_{P_{i,j}} \tag{11.95}$$

where the right-hand side is known from (11.92). Introducing the weighting coefficient

$$w_{i,j} = \left| \hat{z} \cdot \left(\boldsymbol{N}_P \times \boldsymbol{N}_{P_{i,j}} \right) \right| \tag{11.96}$$

the magnetic flux density projected onto the dual-edge is expressed as

$$\boldsymbol{B} \cdot \hat{s} = \sum_{i=1}^{2} \sum_{j=1}^{2} w_{i,j} \left(\boldsymbol{B}_{i,j} \cdot \hat{s} \right) \Bigg/ \sum_{i=1}^{2} \sum_{j=1}^{2} w_{i,j} \tag{11.97}$$

A similar projection operator for the electric flux density is introduced as

$$\boldsymbol{D} \cdot \hat{\boldsymbol{p}} = \sum_{i=1}^{2} \sum_{j=1}^{2} w_{i,j} \left(\boldsymbol{D}_{i,j} \cdot \hat{\boldsymbol{p}} \right) \Big/ \sum_{i=1}^{2} \sum_{j=1}^{2} w_{i,j} \tag{11.98}$$

where the weights are given by

$$w_{i,j} = \left| \hat{z} \cdot \left(\boldsymbol{N}_S \times \boldsymbol{N}_{S_{i,j}} \right) \right| \tag{11.99}$$

and $\boldsymbol{N}_S$ is the normal area vector of a secondary grid cell through which a primary edge passes. As discussed in Section 11.6, (11.97) and (11.98) also preserve the divergenceless nature of the flux densities and lead to a stable solution.

11.7.3 Efficient Time-Stepping Implementation

The explicit solution method used to time-step the discrete E- and H-fields is based upon solving (11.91) to (11.94), (11.97), and (11.98). As discussed in Section 11.6, computing the parameters for these equations requires a significant number of floating-point operations. However, by employing standard finite-element-type techniques, the computational efficiency can be greatly enhanced by treating these expressions as linear operators.

To this end, the explicit time-stepping formulation for the PGY technique is expressed by the following set of linear equations, involving matrix operations on one-dimensional arrays of field components located within the grid (designated using a bracket { } notation):

$$\left\{ h_{z_k}^{n+1} \right\} = \left\{ h_{z_k}^{n} \right\} + \left[A_{h_z} \right] \left\{ e_{t_k}^{n+1/2} \right\} \tag{11.100}$$

$$\left\{ b_{t_{k+1/2}}^{n+1} \right\} = \left\{ b_{t_{k+1/2}}^{n} \right\} + \left[A_{h_t} \right] \begin{bmatrix} \left\{ e_{z_{k+1/2}}^{n+1/2} \right\} \\ \left\{ e_{t_{k+1}}^{n+1/2} \right\} \end{bmatrix} \tag{11.101}$$

$$\left\{ h_{t_{k+1/2}}^{n+1} \right\} = \left[A_{h_P} \right] \left\{ b_{t_{k+1/2}}^{n+1} \right\} \tag{11.102}$$

$$\left\{ e_{z_{k+1/2}}^{n+3/2} \right\} = \left[D_{e_z} \right] \left\{ e_{z_{k+1/2}}^{n+1/2} \right\} + \left[A_{e_z} \right] \left\{ h_{t_{k+1/2}}^{n+1} \right\} \tag{11.103}$$

$$\left\{ d_{t_k}^{n+3/2} \right\} = \left[D_{e_t} \right] \left\{ d_{t_k}^{n+1/2} \right\} + \left[A_{e_t} \right] \begin{bmatrix} \left\{ h_{z_{k+1}}^{n+1} \right\} \\ \left\{ h_{t_{k+1/2}}^{n+1} \right\} \end{bmatrix} \tag{11.104}$$

$$\left\{e_{t_k}^{n+3/2}\right\} = \left[A_{e_P}\right]\left\{d_{t_k}^{n+3/2}\right\} \tag{11.105}$$

In (11.100) to (11.105), the subscript k refers to the discrete height along the z-direction, the $[D]$'s are diagonal matrices, and the $[A]$'s are sparse matrices. Note that these matrices are associated only with the two-dimensional unstructured grid, since they are the same for all values of k. As a result, the additional computer memory required to store these matrices is nominal.

Overall, the field updates performed using (11.100) to (11.105) provide an efficient, second-order accurate computational approach to implement PGY models. Further, denoting L_i^p as horizontal edges sharing a common vertex, the solution is numerically stable [34, 38] if

$$\Delta t < \frac{1}{c \sup\left[\sqrt{\left(\frac{1}{\Delta z}\right)^2 + \sum_{i=1}^{2}\left(\frac{1}{L_i^p}\right)^2}\right]} \tag{11.106}$$

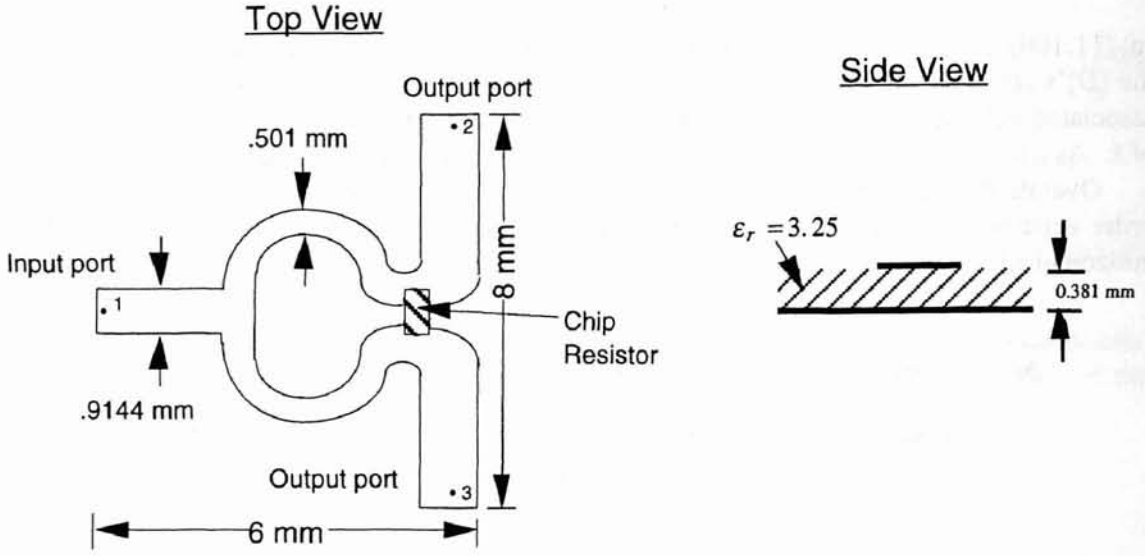
11.7.4 Modeling Example: 32-GHz Wilkinson Power Divider

Many digital and microwave circuit devices and antennas have a planar symmetry that is ideal for PGY modeling of their associated electromagnetic wave phenomena. The ability of the PGY technique to effectively implement an FDTD solution within an unstructured grid allows for the use of commercially available software for computer-aided design and mesh generation. These features greatly simplify the modeling process. Further, circuits with complex geometries are more accurately modeled with the use of unstructured nonorthogonal grids.

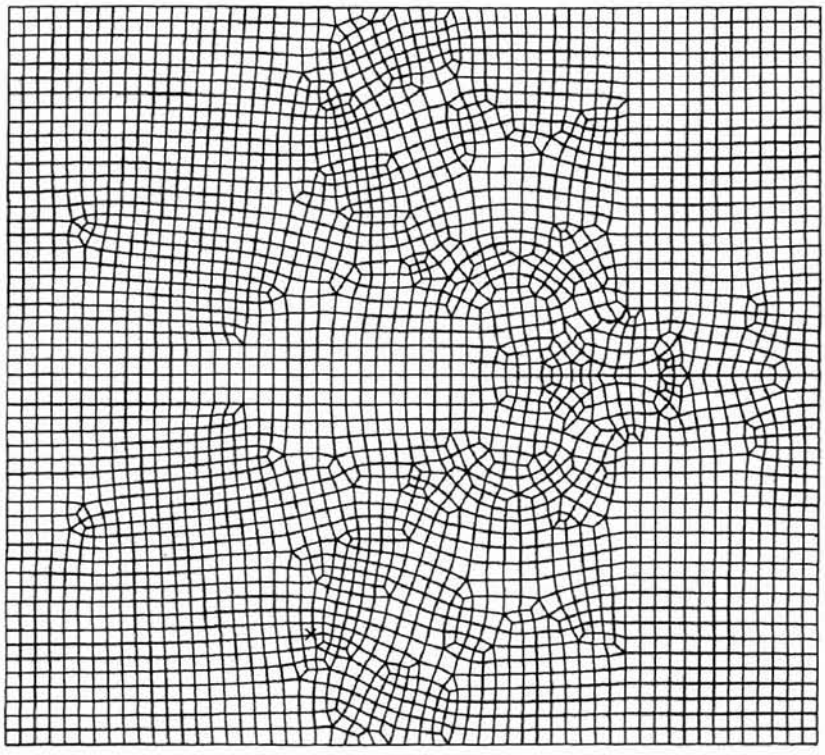
This is now demonstrated through the application of the PGY algorithm to model a printed-circuit Wilkinson power divider [34], as illustrated in Fig. 11.16(a). This device is designed for equal-phase 3-dB power division at 32 GHz, assuming matching to 50Ω microstrip lines. A 15-mil substrate of permittivity $\varepsilon_r = 3.25$ backed by a copper ground plane is assumed, along with a 100Ω chip resistor placed between ports 2 and 3 for isolation.

Fig. 11.16(b) depicts the horizontal cross section of the unstructured mesh used to model the device of Fig. 11.16(a). Each such mesh plane consists of 3,200 quadrilaterals. Along the vertical direction, there are 25 uniformly spaced lattice cells with $\Delta z = 0.0635$ mm. The fields on the outer boundary of the lattice are updated using the second-order Higdon ABC discussed in Chapter 6, Section 6.4. To maintain the second-order accuracy of this ABC, the mesh in the horizontal plane is padded with two layers of rectangular cells, slightly extending the problem domain. The device is excited by an ideal voltage source generating a Gaussian pulse, and the chip resistor is modeled as a lumped load, as discussed in Chapter 15, Section 15.9.2.

Fig. 11.17 shows the S parameter S_{11} computed for this device using the PGY technique. (See Chapter 15, Section 15.2 for the extraction of S parameters from FDTD data.) The S_{11} data are compared with results from a uniform Cartesian FDTD model ($\Delta x = \Delta y = 0.0573$ mm, $\Delta z = 0.0635$ mm) having five times as many cells. The Cartesian model requires fine resolution in the horizontal plane to limit errors due to: (1) staircasing the curved metal traces, and (2) imprecise modeling of the length of the curved quarter-wave transformers in the circuit. We see good agreement between the results of the relatively coarse PGY mesh and the fine FDTD grid. Similar agreement is found for the other S parameters of this device. This indicates the possibility of an 80% reduction in computer resources by using the PGY method.



(a) Geometry.



(b) Horizontal cross section of the unstructured grid.

Fig. 11.16 32-GHz Wilkinson power divider and its planar generalized Yee model. *Source:* Gedney and Lansing, *IEEE Trans. Microwave Theory Tech.*, 1996, pp. 1393–1400, © 1996 IEEE.

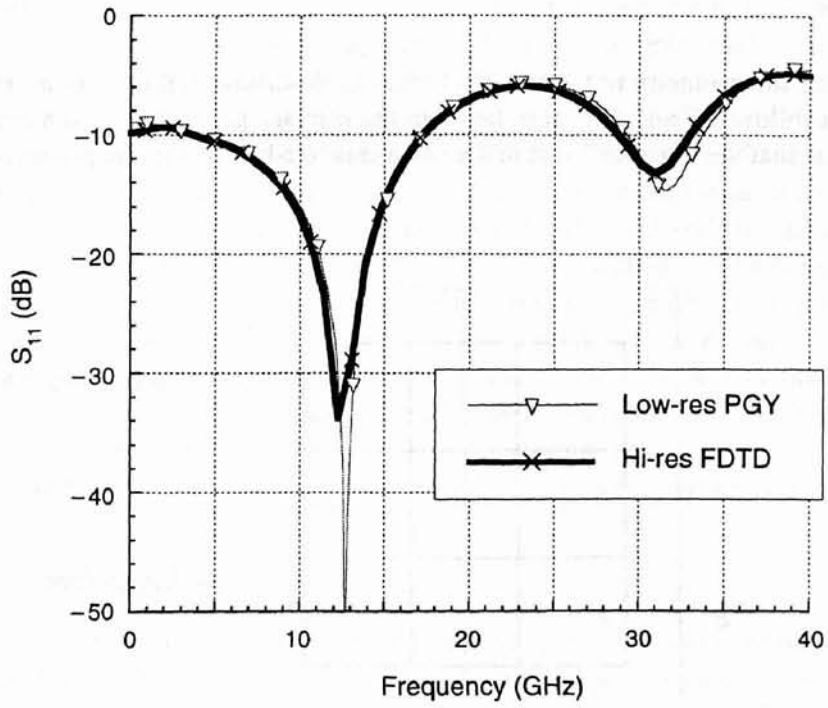


Fig. 11.17 Magnitude of S_{11} of the Wilkinson power divider of Fig. 11.16 computed using both low-resolution PGY and high-resolution Cartesian FDTD methods. *Source:* Gedney and Lansing, *IEEE Trans. Microwave Theory Tech.*, 1996, pp. 1393–1400, © 1996 IEEE.

11.8 CARTESIAN SUBGRIDS

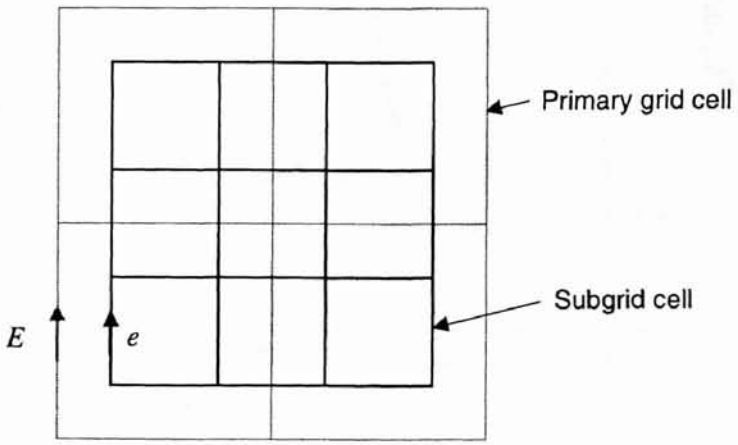
A variety of methods have been proposed to reduce staircasing errors in Cartesian FDTD grids and to refine the mesh resolution. Nonuniform (graded) meshes, discussed in Section 11.2, comprise one approach. However, since grading extends throughout an entire cross section of the space lattice (see Fig. 11.5), there may be an unnecessary expenditure of computer resources to store and process small grid cells located far from the geometrical detail of interest.

A more general approach is to use a subgrid. Subgrids are blocks of small mesh cells placed at specific locations within the computational domain without modifying the global mesh. A variety of subgridding schemes have been proposed, ranging from methods based on the Helmholtz equation [40, 41]; on curl and source matrices in the finite-integration technique [42]; and on Ampere's law in integral form [43]. Improvements resulting in broadened applicability include spatial and temporal interpolation for data transfer between the main grid and the subgrid [44 – 46], and using filtering techniques to reduce instabilities [47]. Whereas these schemes can show excellent performance, general applicability might not be guaranteed. Problems can arise due to numerical instability, inability to model PEC and dielectric materials entering the subgrid from the main grid, or significant reflections from the main-grid / subgrid interface.

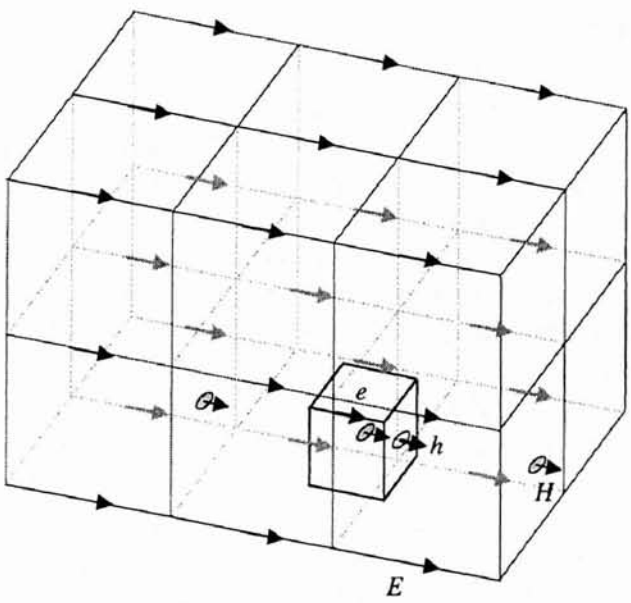
This section presents a subgridding technique that is focused on providing high performance with respect to numerical stability, low boundary reflections, material transitions, and computational requirements. Furthermore, the method is tailored to provide general applicability within automated CAD-supported modeling environments.

11.8.1 Geometry

Fig. 11.18 illustrates the geometry of the subgrid technique described in this section. Here and in the discussion that follows, E and H denote fields in the primary grid, and e and h denote fields in the subgrid. Note that the space-cell size in the subgrid is one-half that in the primary grid.



(a) Simplified plan view of a primary grid containing a 2:1 subgrid.



(b) Three-dimensional perspective view, showing
(E, e) and (H, h) components used for spatial interpolation.

Fig. 11.18 Geometry of a primary grid and a 2:1 subgrid.

The subgridding technique illustrated in Fig. 11.18 is designed to robustly achieve a spatial refinement factor of 2:1 at each stage from the exterior grid, and furthermore allow for systematic nesting of subgrids so that composite spatial refinements of 2^n:1 are achievable. This technique has been successful for up to five nested subgrids (overall refinements up to 32:1).

The spatial positioning of the subgrid shifted by one-quarter of the primary grid's cell dimensions in each direction, as shown in Fig. 11.18(a), has been found to lead to the smoothest transition, since all H-fields of the primary grid as well as the subgrid are collocated in a way that simplifies spatial interpolation and optimizes temporal updating. Within this implementation, the cell size does not have to be modeled equidistant for the entire three-dimensional subgrid, only for each Cartesian direction. Figure 11.18(b) depicts a detailed view of this setup with a single subgrid cell placed into a $3 \times 2 \times 2$-cell primary grid. As will be discussed, the electric and magnetic field components located on the outermost surfaces of the subgrid are used for information transfer from the primary grid to the subgrid, and the magnetic field components located within the subgrid are used for information transfer in the opposite direction.

11.8.2 Time-Stepping Scheme

For numerical stability, it is clear that time-stepping in the 2:1 subgrid must be conducted using one-half the time-step (i.e., at twice the rate of time-stepping) employed for the primary grid. This results in missing temporal information when using a common FDTD updating scheme for both the primary grid and subgrid. Most reported subgridding schemes use an extrapolation in time to determine the missing information. In contrast, the technique discussed here uses an *interpolation* in time, a procedure that is generally considered to be more accurate than extrapolation. Extensive numerical experimentation has shown that this approach achieves stability beyond 100,000 time-steps when using a time-step only 10% below the Courant limit, even for cases where a PEC traverses the transition region between the primary grid and subgrid.

We assume that the following sets of fields are known at the beginning of time-stepping: (1) all E-components through time-step $n-1$, (2) all H-components through time-step $n-0.5$, (3) all e-components through time-step $n-1$, (4) all h-components through time-step $n-0.75$, (5) all interior e-components away from the subgrid border region (designated e_{int}) through time-step $n-0.5$, and (6) all interior h-components away from the subgrid border region (designated h_{int}) through time-step $n-0.25$. The complete time-stepping algorithm follows:

1. Update all E components to time-step n.
2. Update all H components to time-step $n+0.5$.
3. Using a spatial-interpolation technique (to be described in the next section), calculate all e-components in the subgrid border region at time-step n from the surrounding E-components in the primary grid. For purposes of notation, we shall designate these subgrid border-region electric-field components as e_b^n. Furthermore, those e_b^n components which are located at the outermost surface of the subgrid are identified as e_s^n.
4. Using a spatial-interpolation technique (to be described in the next section), calculate all h-components in the subgrid border region at time-step $n+0.5$ from the surrounding H-components in the primary grid. For purposes of notation, we shall designate these subgrid border-region magnetic-field components as $h_b^{n+0.5}$.

5. Using the available data for e_s^n, e_s^{n-1}, and e_s^{n-2}, use a quadratic-function interpolation in time to calculate $e_s^{n-0.5}$.

6. Using the available data for $h_b^{n+0.5}$, $h_b^{n-0.75}$, and $h_b^{n-1.5}$, use a quadratic-function interpolation in time to calculate $h_b^{n-0.25}$.

7. Using the available data for $h^{n-0.75}$ and $e^{n-0.5}$ (also consisting of $e_b^{n-0.5}$ which contain the $e_s^{n-0.5}$ components updated in Step 5), apply the normal Yee algorithm to obtain an alternative evaluation of $h_b^{n-0.25}$.

8. Calculated the weighted average $0.35 \times h_b^{n-0.25}\big|_{\text{Step 6}} + 0.65 \times h_b^{n-0.25}\big|_{\text{Step 7}}$ for each $h_b^{n-0.25}$.

9. Using the available data for $e^{n-0.5}$ and $h^{n-0.25}$, apply the normal Yee algorithm to calculate e_{int}^n.

10. Using the available data for $h^{n-0.25}$ and e^n, apply the normal Yee algorithm to calculate $h_{\text{int}}^{n+0.25}$.

11. Calculate all $H^{n+0.25}$ fields within the subgrid via spatial interpolation of the adjacent $h^{n+0.25}$ fields.

12. Using the available data for e^n and $h^{n+0.25}$, apply the normal Yee algorithm to calculate $e_{\text{int}}^{n+0.5}$.

13. Using the available data for $h^{n+0.25}$ and $e_{\text{int}}^{n+0.5}$, apply the normal Yee algorithm to calculate $h_{\text{int}}^{n+0.75}$.

14. Calculate all $H^{n+0.75}$ fields within the subgrid via spatial interpolation of the adjacent $h_{\text{int}}^{n+0.75}$ fields.

15. Calculate all $H^{n+0.5}$ fields within the subgrid via simple averaging of the $H^{n+0.25}$ fields obtained in Step 11 and the $H^{n+0.75}$ fields obtained in Step 14.

This returns the sets of primary and subgrid fields to the original assumed state, except that all fields have been advanced by one complete time-step of the primary grid. The process can now repeat in a self-consistent manner.

11.8.3 Spatial Interpolation

As shown in Fig. 11.18(b), a total of $3 \times 9 = 27$ E-components distributed in three dimensions are used for the spatial interpolation of a single e-component in Step 3 of the time-stepping scheme. This interpolation is implemented using a three-dimensional cubic spline [48] to take advantage of its low-reflection properties [45]. For certain cases, particularly in the initial propagation stage, a basic cubic spline can lead to over- and underestimated values. To decrease error and therefore reduce instabilities, one can use a cubic smoothing spline [49], which allows a trade-off between minimizing the residual error and minimizing local variation.

In Step 4 of the time-stepping scheme, h-components are spatially interpolated by application of either one-dimensional or two-dimensional cubic splines on the correlated H-components, depending on their position. As stated earlier, this simplified interpolation relative to that required for the e-components results from the spatial positioning of the subgrid shifted by one-quarter of the primary grid's cell dimensions in each direction.

For a PEC cutting through the primary-grid / subgrid transition, a special interpolation must be performed if at least one of the field components is located inside the conductor. The modified scheme implements a two-dimensional or a one-dimensional spline interpolation, depending on the localization of the cell for a virtually introduced e-component, while including only those E-components that are not located inside the conducting material.

11.8.4 Numerical Stability Considerations

A major issue with subgridding schemes is reduced numerical stability. This problem arises due to the abrupt spatial and temporal transition from the primary grid to the subgrid, which requires non-Yee calculations to approximate the missing field components in the interface region. Due to the complex updating scheme of the algorithm discussed here (and of three-dimensional subgridding schemes in general), derivations of stability criteria for subgridding are difficult. Some work on stability has been reported for a two-dimensional subgrid based on a simple interpolation approach [50], and for a three-dimensional case [47].

In order to derive the stability criterion for the scheme discussed in this section, each of the steps within an updating cycle can be formulated as a relation between Yee field values at two subsequent sequential time-steps valid either inside or outside the subgrid region, which leads to a product of different matrix terms. Via Fourier mode analysis, the corresponding amplification matrix subsequently would be derived. To ensure stability of the system, all roots of the characteristic polynomial derived from the amplification matrix would be required to be less than or equal to one. An extensive discussion of this general procedure is found, for example, in [51] for the case of a two-dimensional problem using different time-stepping schemes. Detailed descriptions of this type of procedure can be found in [52, 53].

The derivation of the characteristic polynomial for the method presented in this section leads to a very complex system whose solution is beyond the scope of this presentation. However, by applying temporal interpolation instead of extrapolation, and by incorporating smoothing cubic splines, robust, stable solutions have been *demonstrated* using a time-step that is reduced by *only 10%* relative to the nominal CFL limit. The temporal interpolation process performs an averaging procedure in Step 8, weighting the influence of the missing electric and magnetic field values determined in the previous Steps 5, 6, and 7. This indirectly specifies the order of the discontinuity within the intersection region between the two grids. The averaging factors of 0.35 and 0.65 specified in Step 8 were determined by numerical experiments for a variety of different cases. By applying these factors, no tendencies for instability have been observed, even for testing up to *more than 100,000 time-steps*.

11.8.5 Reflection from the Interface of the Primary Grid and Subgrid

This section summarizes numerical experiments aimed at measuring spurious reflections from the interface of the primary grid and subgrid for a plane wave traversing the subgridding region. Fig. 11.19 illustrates the gridding geometry used for the tests.

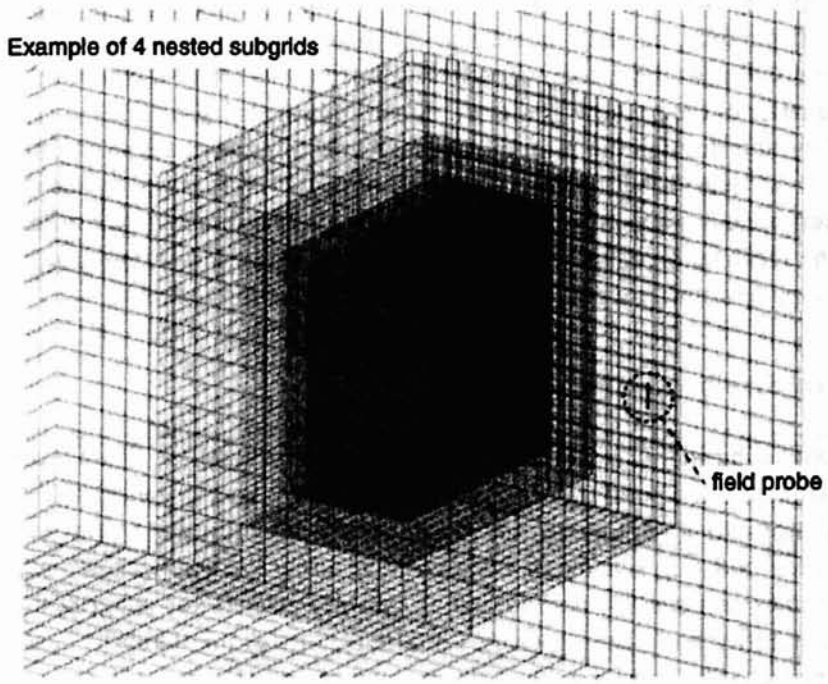


Fig. 11.19 Gridding geometry used for numerical experiments assessing spurious reflections from single or nested free-space subgrids embedded within a free-space primary grid.

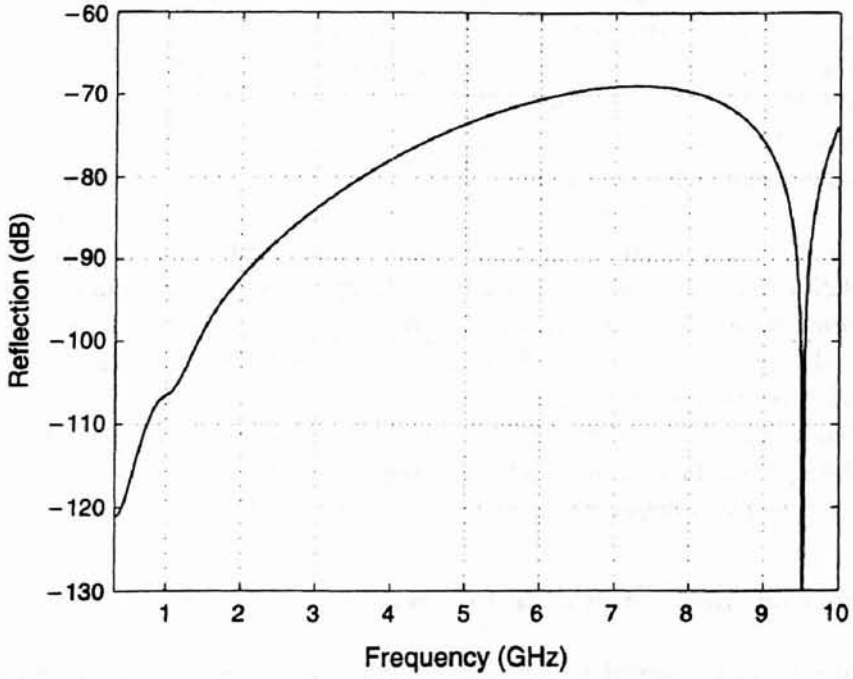


Fig. 11.20 Calculated spurious reflection from a single 2:1 free-space subgrid for $\lambda_0/30$ resolution of the free-space primary grid at 10 GHz.

In the first numerical experiment, a single 2:1 free-space subgrid spanning $31 \times 31 \times 23$ cells was embedded within a free-space primary grid of spatial resolution $\lambda_0/30$ at 10 GHz. Reflections from the subgrid were probed at a point in the primary grid three cells away from the transition region used for the spline interpolation. An impulsive plane-wave excitation of 10-GHz bandwidth was applied to the test primary-grid / subgrid geometry, as well as to a reference primary grid wherein the subgrid was deactivated. Fig. 11.20 graphs the calculated subgrid reflection as a function of frequency. We see that this reflection is below –70 dB across the entire spectrum of interest.

The second numerical experiment involved the embedding of a system of either triply-nested or quad-nested 2:1 free-space grids (as illustrated in Fig. 11.19) within the free-space primary grid. This arrangement provided either a maximum 8:1 or 16:1 reduction of the grid-cell size in going from the primary grid to the innermost subgrid. Except for the nesting, all simulation parameters were the same as for the first experiment.

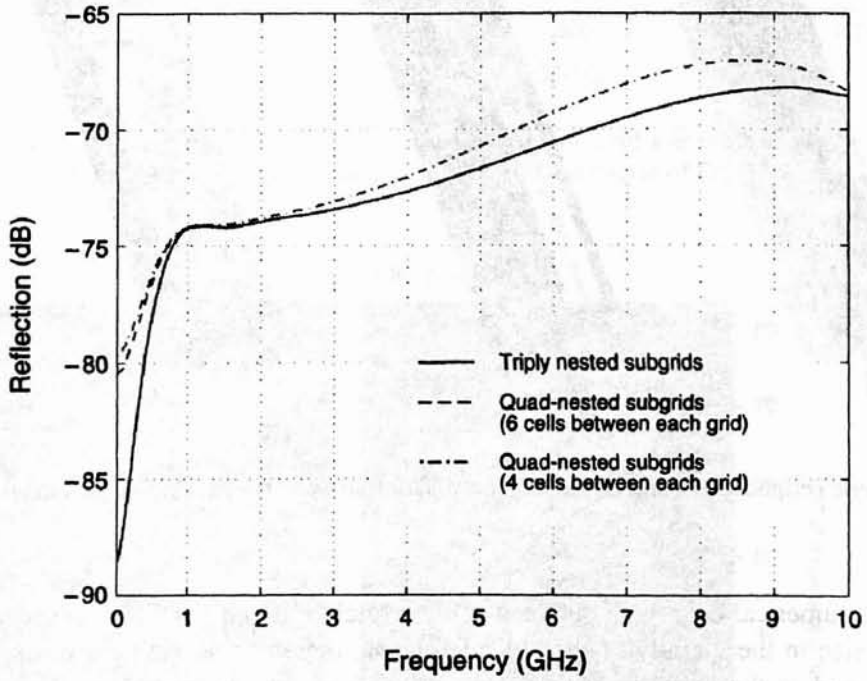


Fig. 11.21 Calculated spurious reflections from systems of triply-nested and quad-nested 2:1 free-space subgrids for $\lambda_0/30$ resolution of the free-space primary grid at 10 GHz.

Fig. 11.21 graphs the calculated reflection from this subgridding system as a function of frequency. Comparing Figs. 11.20 and 11.21, we see that, while using nested subgrids causes significantly increased reflections at lower frequencies, the worst-case reflection is approximately the same—that is, approximately –70 dB. Thus, for most FDTD simulations, embedding a nested subgridding system of the type discussed in this section has little impact upon the rest of the grid, regardless of the maximum degree of mesh refinement that is achieved. That is, the subgridding system is quite transparent.

11.8.6 Illustrative Results: Helical Antenna on Generic Cellphone at 900 MHz

We next discuss an experimental and code-to-code validation wherein the subgridding technique reviewed in this section was used to model a generic cellphone. As shown in Fig. 11.22, the phone was modeled as a $40 \times 16 \times 140$-mm metal box equipped with an axial-mode, 6.4-mm-diameter helical antenna [54]. The antenna was comprised of seven turns of 1.3-mm diameter wire, with a pitch angle of 7.2°. In the laboratory experiments, surface waves on the feeding coaxial cable were suppressed using a $\lambda/4$ stub and ferrite chokes.

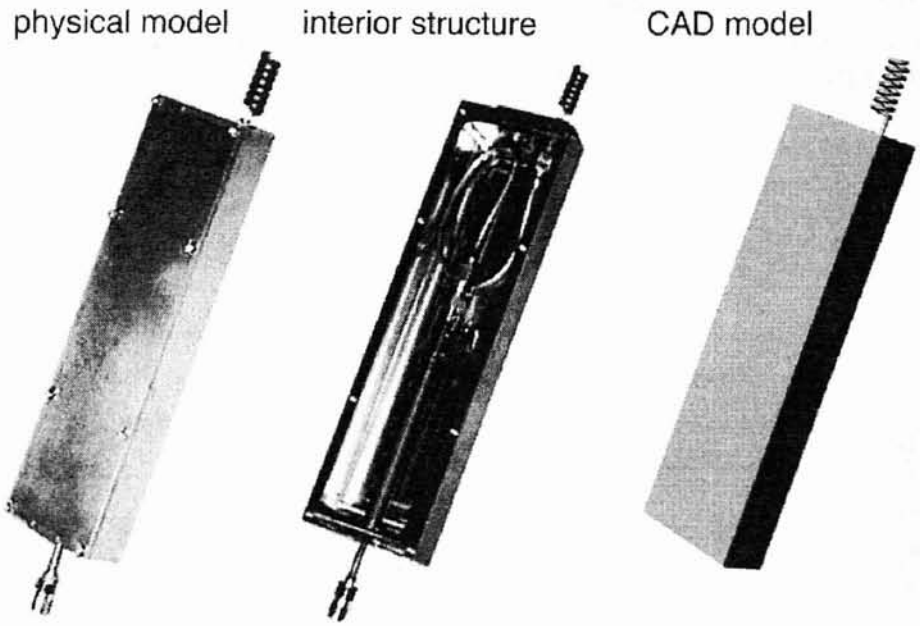


Fig. 11.22 Generic cellphone and antenna used in a validation study of the FDTD subgrid model.

Benchmark numerical data were obtained using a finely meshed FDTD model that provided 0.325-mm cell size in the vicinity of the helical antenna, thereby resolving the cross section of the thin wire used in the helix with 4×4 grid cells. The subgrid provided the same spatial resolution of the antenna, but connected to a much coarser primary mesh with a maximum cell size of 8 mm. For both FDTD models, a discrete source feed was placed in a 0.325-mm gap between the PEC box and the helix wire.

Fig. 11.23 compares the measured feedpoint impedance with the results of the finely meshed benchmark FDTD model and the subgrid FDTD model. We see that the two FDTD datasets are virtually identical, and correspond well with the measured values, even for the highly sensitive imaginary part of the impedance. The excellent agreement between the subgrid and benchmark FDTD models is verified in Fig. 11.24, which visualizes the differences in these models' results for the E- and H-fields along a vertical cut-plane through the antenna helix and feedpoint. Very small differences of less than 0.35 dB are seen here for both field types. Additional results (not shown here) indicate that the radiated (far-field) E-components calculated by these two models have similarly small differences, in the order of 0.5% or less.

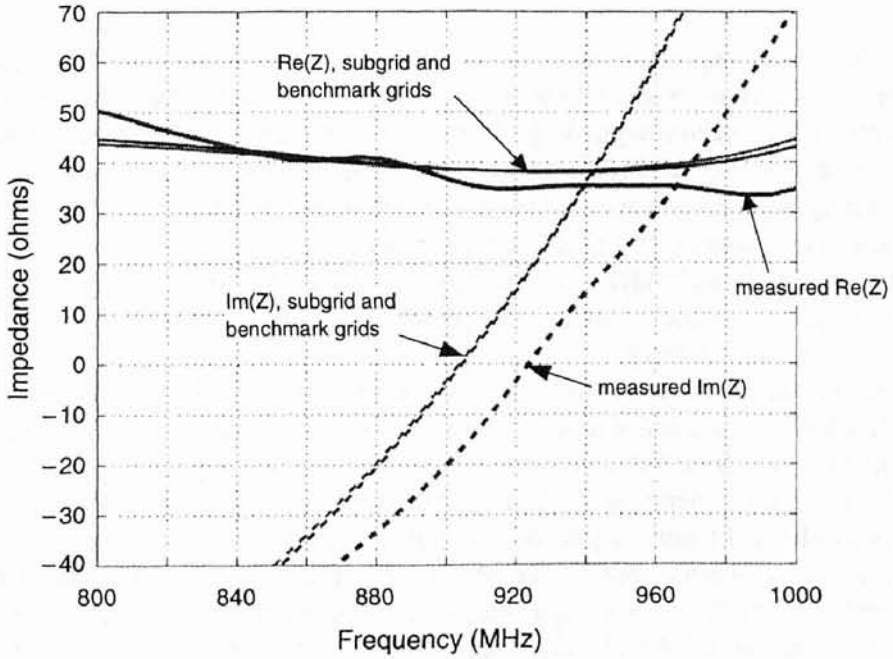


Fig. 11.23 Comparison of measured feedpoint impedance with results of the benchmark fine-grid FDTD model and the subgrid FDTD model of the generic cellphone of Fig. 11.22.

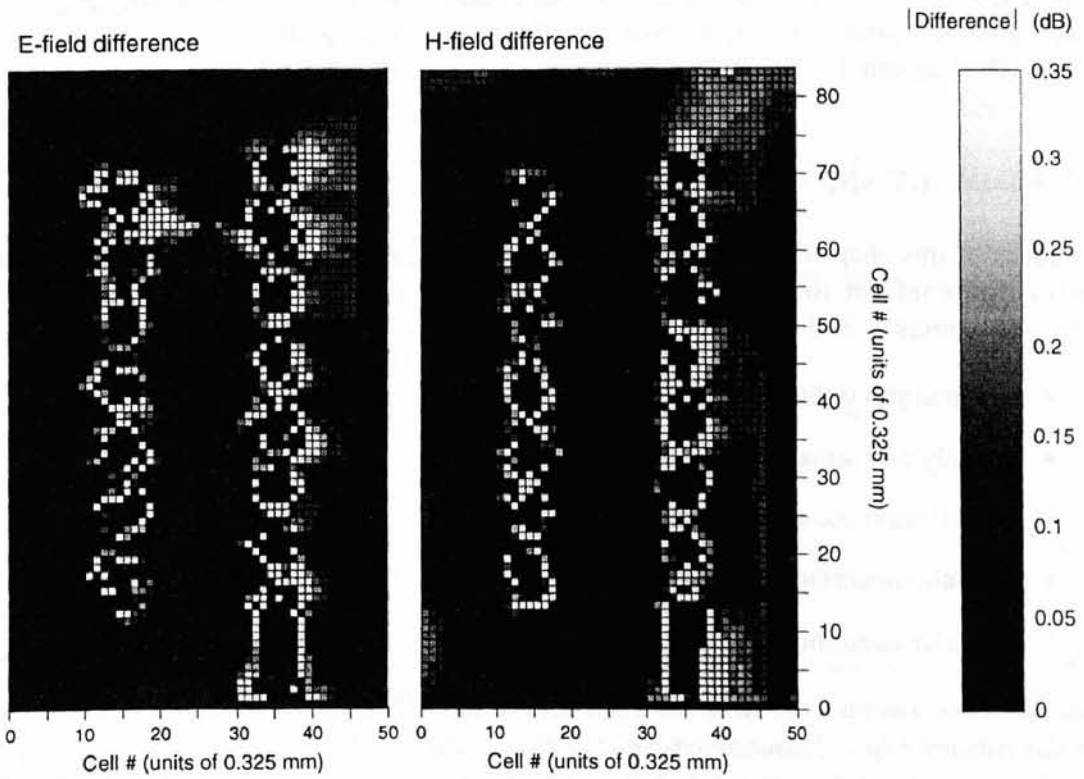


Fig. 11.24 Grayscale visualization of the differences in the subgrid and benchmark fine-grid FDTD models' results for the *E*- and *H*-fields along a vertical cut-plane through the antenna helix and feedpoint for the generic cellphone of Fig. 11.22.

11.8.7 Computational Efficiency

Commercial electromagnetic devices usually include structures that are considerably smaller and more complicated than those of benchmark problems. Although such fine details often do not significantly affect electromagnetic performance, it is risky to exclude this possibility a priori. For this reason, highly detailed simulations are often desired. An example of the geometric detail that may be necessary is provided in the case study of Chapter 14, Section 14.9, where the performance of a commercial cellphone is examined when the phone is located in free space and adjacent to human head models. The subgridding technique of this section is used in this case study, making it possible to model the phone in a computationally efficient manner while retaining engineering accuracy.

Modeling functional details of compact antennas may simultaneously require resolution of fine geometrical details smaller than $\lambda/1{,}000$, and much larger structures spanning one or more λ. An example is given by a bifilar antenna [55, 56] mounted on a generic phone and placed next to the ear of an inhomogeneous head model [57]. Here, key antenna details require a local spatial resolution of about 0.1 mm. By using a doubly nested subgrid for the antenna (0.1-mm inner subgrid resolution, 0.2-mm outer subgrid resolution) that is embedded within a primary mesh that is graded from 0.4 to 8 mm, a total of approximately 5×10^6 cells is required. For a sinusoidal simulation at 2.4 GHz, the steady state is reached after 30 periods, which corresponds to about 60,000 time-steps. This requires a running time of 1.5 hr using a 3.6-GHz Intel Pentium-4 computer, with a memory utilization of 370 MB. In comparison, using only an appropriate graded mesh without the antenna subgrid would elevate the running time to 17 hr and the required memory to 1.3 GB. In the worst case, using a hypothetical homogeneous grid would require about 2 months of running time *if* sufficient central memory could be found to provide the 700 GB of required storage. The latter is clearly infeasible at the present time.

11.9 SUMMARY AND CONCLUSIONS

One focus of this chapter was on the development of FDTD algorithms implemented on space lattices that conform to the surfaces of all structure boundaries in the problem domain. Five types of grids were discussed:

- Nonuniform orthogonal grids;
- Locally conformal grids, globally orthogonal;
- Global curvilinear coordinates;
- Irregular nonorthogonal structured grids;
- Irregular nonorthogonal unstructured grids.

A second focus was on a robust, stable, and computationally efficient subgridding technique that provides enhanced spatial resolution for conventional staircasing.

We conclude that the extension of the FDTD method to general space lattices, along with the application of robust subgridding techniques, are promising means to exploit commercially available software for computer-aided design and mesh generation. This greatly simplifies the FDTD modeling process, and enhances its accuracy for structures with complex geometries.

REFERENCES

[1] Cangellaris, A. C., and D. B. Wright, "Analysis of the numerical error caused by the stair-stepped approximation of a conducting boundary in FDTD simulations of electromagnetic phenomena," *IEEE Trans. Antennas Propagat.*, Vol. 39, 1993, pp. 1518–1525.

[2] Sheen, D., *Numerical Modeling of Microstrip Circuits and Antennas*, Ph.D. diss., MIT, Cambridge, MA, June 1991.

[3] Tulintseff, A., "The finite-difference time-domain method and computer program description applied to multilayered microstrip antenna and circuit configurations," AMT: 336.5-92-041, Tech. Rpt., Jet Propulsion Laboratory, May 1992.

[4] Monk, P., and E. Suli, "A convergence analysis of Yee's scheme on non-uniform grids," *SIAM J. Numerical Analysis*, Vol. 31, 1994, pp. 393–412.

[5] Monk, P., "Error estimates for Yee's method on non-uniform grids," *IEEE Trans. Magnetics*, Vol. 30, 1994, pp. 3200–3203.

[6] Manteuffel, T. A., and J. A. White, "The numerical solution of second-order boundary value problems on nonuniform meshes," *Math. Comp.*, Vol. 47, 1986, pp. 511–535.

[7] Kreiss, H., T. Manteuffel, B. Schwartz, B. Wendroff, and J. A. B. White, "Supraconvergent schemes on irregular meshes," *Math. Comp.*, Vol. 47, 1986, pp. 537–554.

[8] Balanis, C., *Advanced Engineering Electromagnetics*, New York: Wiley, 1989.

[9] Sheen, D., S. Ali, M. Abouzahra, and J. A. Kong, "Application of the three-dimensional finite-difference time-domain method to the analysis of planar microstrip circuits," *IEEE Trans. Microwave Theory Tech.*, Vol. 38, 1990, pp. 849–856.

[10] Jurgens, T. G., A. Taflove, K. Umashankar, and T. G. Moore, "Finite-difference time-domain modeling of curved surfaces," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 357–366.

[11] Jurgens, T. G., and A. Taflove, "Three-dimensional contour FDTD modeling of scattering from single and multiple bodies," *IEEE Trans. Antennas Propagat.*, Vol. 41, 1993, pp. 1703–1708.

[12] Yu, W., and R. Mittra, "A conformal FDTD software package modeling antennas and microstrip circuit components," *IEEE Antennas Propagat Magazine*, Vol. 42, Oct. 2000, pp. 28–39.

[13] Holland, R., "Finite-difference solution of Maxwell's equations in generalized nonorthogonal coordinates," *IEEE Trans. Nuclear Science*, Vol. 30, 1983, pp. 4589–4591.

[14] Mei, K. K., A. Cangellaris, and D. J. Angelakos, "Conformal time domain finite-difference method," *Radio Science*, Vol. 19, 1984, pp. 1145–1147.

[15] J. A. Stratton, *Electromagnetic Theory*, New York: McGraw-Hill, 1941.

[16] Lee, J.-F., R. Palendech, and R. Mittra, "Modeling three-dimensional discontinuities in waveguides using nonorthogonal FDTD algorithm," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, 1992, pp. 346–352.

[17] 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.

[18] Shankar, V., A. Mohammadian, and W. F. Hall, "A time-domain finite-volume treatment for the Maxwell equations," *Electromagnetics*, Vol. 10, 1990, pp. 127–145.

[19] Mohammadian, A. H., V. Shankar, and W. F. Hall, "Computation of electromagnetic scattering and radiation using a time-domain finite-volume discretization procedure," *Computer Physics Comm.*, Vol. 68, 1991, pp. 175–196.

[20] Mohammadian, A. H., V. Shankar, and W. F. Hall, "Application of time-domain finite-volume method to some radiation problems in two and three dimensions," *IEEE Trans. Magnetics*, Vol. 27, 1991, pp. 3841–3844.

[21] Holland, R., V. P. Cable, and L. C. Wilson, "Finite volume time-domain (FVTD) techniques for EM scattering," *IEEE Trans. Electromagn. Compat.*, Vol. 33, 1991, pp. 281–294.

[22] Lee, J.-F., "Numerical solutions of TM scattering using an obliquely Cartesian finite difference time domain algorithm," *IEE Proc. H (Microwaves, Antennas, Propagat.)*, Vol. 140, 1992, pp. 23–28.

[23] Lee, J.-F., "Finite difference time domain algorithm for non-orthogonal grids and its application to the solution of electromagnetic scattering problems," *Archiv Elektronik Uebertragungstechnik*, Vol. 46, 1992, pp. 328–335.

[24] Harms, P., J.-F. Lee, and R. Mittra, "A study of the nonorthogonal FDTD method versus the conventional FDTD technique for computing resonant frequencies of cylindrical cavities," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, 1992, pp. 741–746.

[25] Harms, P., J.-F. Lee, and R. Mittra, "Characterizing the cylindrical via discontinuity," *IEEE Trans. Microwave Theory Tech.*, Vol. 41, 1993, pp. 153–156.

[26] Hein, S., "Finite difference time-domain approximation of Maxwell's equations with non-orthogonal condensed TLM mesh," *Intl. J. Numerical Modeling: Electronic Networks, Devices, Fields*, Vol. 7, 1994, pp. 179–188.

[27] Kashiwa, T., T. Onishi, and I. Fukai, "Analysis of microstrip antennas on a curved surface using the conformal grids FDTD method," *IEEE Trans. Antennas Propagat.*, Vol. 42, 1994, pp. 423–426.

[28] Ray, S., "Numerical dispersion and stability characteristics of time-domain methods on nonorthogonal meshes," *IEEE Trans. Antennas Propagat.*, Vol. 41, 1993, pp. 233–235.

[29] Zhao, L., L. S. Tong, and R. Carter, "The influence of boundary conditions on resonant frequencies of cavities in 3-D FDTD algorithm using nonorthogonal coordinates," *IEEE Trans. Magnetics*, Vol. 30, 1994, pp. 3570–3573.

[30] Fusco, M., "FDTD algorithm in curvilinear coordinates," *IEEE Trans. Antennas Propagat.*, Vol. 38, 1990, pp. 78–89.

[31] Fusco, M., M. Smith, and L. Gordon, "A three-dimensional FDTD algorithm in curvilinear coordinates," *IEEE Trans. Antennas Propagat.*, Vol. 39, 1991, pp. 1463–1471.

[32] Madsen, N., "Divergence preserving discrete surface integral methods for Maxwell's equations using nonorthogonal unstructured grids," Tech. Rpt. UCRL-JC-109787, LLNL, Feb. 1992.

[33] Gedney, S., and F. Lansing, "Full wave analysis of printed microstrip devices using a generalized Yee algorithm," *IEEE Antennas Propagat. Symp. Dig.*, Ann Arbor, MI, 1993.

[34] Gedney, S., F. Lansing and D. Rascoe, "A generalized Yee-algorithm for the analysis of MMIC devices," *IEEE Trans. Microwave Theory Tech.*, Vol. 44, 1996, pp. 1393–1400.

[35] Gedney, S., and F. Lansing, "A parallel planar generalized Yee algorithm for the analysis of microwave circuit devices," *Intl. J. Numerical Modeling (Electronic Networks, Devices, Fields)*, Vol. 8, 1995, pp. 249–264.

[36] Gedney, S., F. Lansing, R. T. Kihm, N. Owona and K. L. Virga, "Simulating 'large' microwave circuits with the parallel planar generalized Yee algorithm," *Proc. IEEE Intl. Symp. Microwave Theory Tech.*, San Francisco, CA, 1996.

[37] Yee, K. S., "Numerical solution of initial boundary value problems in isotropic media," *IEEE Trans. Antennas Propagat.*, Vol. 14, 1966, pp. 302–307.

[38] Gedney, S., F. Lansing, and D. Rascoe, "Full wave analysis of microwave monolithic circuit devices using a generalized Yee algorithm based on unstructured grids," Tech. Rpt. 19186/8768, Jet Propulsion Laboratory, Pasadena, CA, Sept. 1994.

[39] Jin, J., *The Finite Element Method in Electromagnetics*, New York: Wiley, 1993.

[40] Zivanovic, S. S., K. S. Yee, and K. K. Mei, "A subgridding method for the time-domain finite-difference method to solve Maxwell's equations," *IEEE Trans. Microwave Theory Tech.*, Vol. 39, 1991, pp. 471–479.

[41] Nehrbass, J. W., and R. Lee, "Optimal finite-difference sub-gridding techniques applied to the Helmholtz equation," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 976–984.

[42] Thoma, P., and T. Weiland, "A consistent subgridding scheme for the finite difference time domain method," *Intl. J. Numerical Modeling*, Vol. 9, 1996, pp. 359–374.

[43] White, M. J., Z. Yun, and M. F. Iskander, "A new 3D FDTD multigrid technique with dielectric traverse capabilities," *IEEE Trans. Microwave Theory Tech.*, Vol. 49, 2001, pp. 422–430.

[44] Chevalier, M. W., R. J. Luebbers, and V. P. Cable, "FDTD local grid with material traverse," *IEEE Trans. Antennas Propagat.*, Vol. 45, 1997, pp. 411–421.

[45] Okoniewski, M., E. Okoniewska, and M. Stuchly, "Three-dimensional subgridding algorithm for FDTD," *IEEE Trans. Antennas Propagat.*, Vol. 45, 1997, pp. 422–429.

[46] Yu, W., and R. Mittra, "A new subgridding method for the finite-difference time-domain (FDTD) algorithm," *Microwave Optical Tech. Lett.*, Vol. 21, 1999, pp. 330–333.

[47] Vaccari, A., R. Pontalti, C. Malacarne, and L. Cristoforetti, "A robust and efficient subgridding algorithm for finite-difference time-domain simulations of Maxwell's equations," *J. Comp. Phys.*, Vol. 194, 2004, pp. 117–139.

[48] Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, 2nd ed., Cambridge, U.K.: Cambridge University Press, 1993.

[49] de Boor, C., *A Practical Guide to Splines*, New York: Springer-Verlag, 1978.

[50] Monk, P., "Sub-gridding FDTD schemes," *ACES J.*, Vol. 11, 1996, pp. 37–46.

[51] Schuster, C., *Simulation, Analysis, and Parameter Extraction of Electronic Components and Circuits Using the Finite Difference Time Domain Method*, Ph.D. thesis, *Diss. ETH Nr.13522*, Zurich, 2000.

[52] Morton, K. W., and D. F. Mayers, *Numerical Solution of Partial Differential Equations*, Cambridge, U.K.: Cambridge University Press, 1994.

[53] Thomas, J. W., *Numerical Partial Differential Equations: Finite Difference Methods*, Berlin, Germany: Springer, 1995.

[54] Burkhardt, M., N. Chavannes, K. Pokovic, T. Schmid, and N. Kuster, "Study on the FDTD performance of transmitters in complex environments," *Proc. ICECOM*, Dubrovnik, 1997, pp. 83–86.

[55] Leisten, O., Y. Vardaxoglou, T. Schmid, B. Rosenberger, E. Agboraw, N. Kuster, and G. Nicolaidis, "Miniature dielectric-loaded personal telephone antennas with low user exposure," *Electronics Lett.*, Vol. 34, 1998, pp. 1628–1629.

[56] Leisten, O., J. C. Vardaxoglou, and P. McEvoy, "Miniature dielectric-loaded antennas for personal telephony and mobility," *Microwave Engineering Europe*, Aug. 2000, pp. 31–34.

[57] Burkhardt, M., *Contributions Toward Uncertainty Assessments and Error Minimization of FDTD Simulations Involving Complex Dielectric Bodies*, Ph.D. thesis, *Diss. ETH Nr.13176*, Zurich, 1999.

PROBLEMS

11.1 Derive the explicit updates (11.10) and (11.11) for the nonuniform FDTD method from (11.6) and (11.7). Similarly derive the explicit updates for the remaining field components.

11.2 Looking ahead to Section 11.6.2, find general expressions for the averaged $\varepsilon_{i,j,k+1/2}$, $\sigma_{i,j,k+1/2}$, and $\mu_{i+1/2,j+1/2,k}$ found in (11.10) and (11.11), assuming that the material medium is three-dimensionally inhomogeneous and isotropic.

11.3 (a) Compute the determinant of the metric tensor g defined in (11.17).

(b) Show that $\sqrt{g} = \boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)$ (Equation (11.18)).

11.4 Take the dot product of (11.28) with the reciprocal vector $\boldsymbol{a}^1$ to derive (11.30a).

11.5 Using central-difference approximations for the spatial derivatives in (11.40), derive (11.41). Show that in the plane-wave space, the ∇ operator can be replaced by

$$\nabla = \sum_{i=1}^{3} \boldsymbol{a}^i \sin\left(\frac{k_i \Delta u^i}{2}\right)$$

Hence, derive (11.42).

11.6 Derive the FDTD algorithm for a Cartesian space based on the nonorthogonal curvilinear FDTD algorithm. Show that they are equivalent.

11.7 The explicit nonorthogonal curvilinear FDTD method can be applied to any curvilinear system, including orthogonal curvilinear coordinates. Assume that you wish to derive the FDTD algorithm for a spherical coordinate system where $(u^1, u^2, u^3) = (r, \theta, \phi)$. Then:

(a) Derive the unitary vectors $\boldsymbol{a}_i$.

(b) Derive the reciprocal vectors $\boldsymbol{a}^i$.

(c) Derive the metrics g_{ij} and g^{ij}.

(d) Derive the explicit update scheme for the electric and magnetic fields e_r, e_θ, e_ϕ, h_r, h_θ, and h_ϕ similar to (11.31) and (11.32).

(e) Are the projections from the contravariant components to the covariant components, as in (11.33a) and (11.33b), necessary? Why or why not?

(f) Derive the stability criterion for the spherical FDTD algorithm (assume that all grid-points lie in the region $r > 0$).

11.8 (a) Compute the determinant of the metric tensor g defined in (11.52).

(b) Define a parallelepiped bounded by the three unitary vectors $(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3)$. Show that the volume of the parallelepiped equals $\boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_3)$.

(c) Show that $\sqrt{g} = \boldsymbol{A}_1 \cdot (\boldsymbol{A}_2 \times \boldsymbol{A}_3)$.

11.9 Given a general parallelepiped within a nonorthogonal lattice defined by the eight vertices having the indexes $(x, y, z)_{i,j,k}$ for $i, j, k = 1, 2$.

(a) Compute the three unitary vectors $(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3)$ and the three reciprocal vectors $(\boldsymbol{A}^1, \boldsymbol{A}^2, \boldsymbol{A}^3)$. Assume that the common origin of the unitary vectors is $(x, y, z)_{1,1,1}$.

(b) Compute the metrics g_{ij}.

(c) Derive the expressions for the field updates in (11.65a, b) and (11.66a, b).

(d) Compute the stability criterion for the explicit update scheme using the metrics pertaining to this parallelepiped.

11.10 Assuming an unstructured mesh composed of general fitted hexahedral elements, derive (11.70a) and (11.70b). Show that if the elements become orthogonal (brick-shaped) with edge lengths Δx, Δy, and Δz, that (11.70a) and (11.70b) reduce to the standard FDTD algorithm.

11.11 Show that the discrete fields expressed in (11.70a) and (11.70b) satisfy Gauss' laws.

11.12 Show that the electric flux density in (11.78) explicitly satisfies Gauss' law.

11.13 Show that if the unstructured mesh is locally orthogonal (i.e., the local hexahedral cells are rectilinear), then the flux densities projected onto the dual edges, as expressed in (11.74) and (11.78), are identical to the flux densities normal to the face as updated in (11.70a, b).

11.14 From (11.91) to (11.99), derive expressions for the elements of the matrices in (11.100) to (11.105), assuming the use of two-dimensional mesh elements that are arbitrary quadrilaterals, and assuming uniform discretization along the z-direction.

PROJECTS

P11.1 You are given a general nonorthogonal structured lattice that is defined by the vertices $[x(i, j, k), y(i, j, k), z(i, j, k)]$ located within a volume Ω bounded by the surface $\partial\Omega$. Assume that the material parameters are homogeneous within the volume, and that on the surface $\partial\Omega$, the fields satisfy the boundary condition $\hat{\boldsymbol{n}} \times \boldsymbol{E} = 0$. Write a program to compute the time-varying E- and H-fields within Ω that are excited by a discrete electric current source aligned on the grid edge (i_S, j_S, k_S). The source is defined as

$$\boldsymbol{J}_s = \left(\boldsymbol{A}^s / |\boldsymbol{A}^s| \right) e^{-(t-t_0)^2/t_w^2} \quad : \quad t_0 = 45 \times 10^{-11} \text{ sec}; \quad t_w = 15 \times 10^{-11} \text{ sec}$$

(a) To test your program, model a $1 \times 2 \times 3$-cm rectangular cavity, using a lattice with uniform discretization $\Delta x = \Delta y = \Delta z = \Delta = 1$ mm. Compute the time-varying contravariant fields for the current source located at grid edge (3, 3, 3). Write to an external file the time-varying E-field at grid edge (4, 5, 6). Using an FFT algorithm, compute the frequency response of this time waveform, noting the resonant frequencies. Compare these with the exact cavity resonances. Double the spatial discretization and monitor the change in the percentage of error.

(b) Repeat (a) for a cylindrical cavity with a radius of 1 cm and a height of 2 cm. Construct a structured grid for this geometry (see [24]) using a discretization on the order of 1 mm, and excite the cavity with the same current source. Compute the resonant frequencies and compare with the exact values. Double the spatial discretization and monitor the change in the percentage of error.

P11.2 You are given an unstructured mesh composed of general fitted hexahedral elements. The geometrical description of the hexahedral elements is provided by a list of all the nodes of the mesh, which uniquely define the vertices of the hexahedral elements. The node coordinates are expressed by three one-dimensional arrays dimensioned as $x(N)$, $y(N)$, $z(N)$, where N is the total number of nodes. Another list is provided that gives the set of nodes for each element. Since there are eight nodes per hexahedron, the list is stored in a two-dimensional array that is dimensioned as $lhexn(8, M)$, where M is the total number of hexahedral elements.

(a) Write a subroutine to uniquely determine the edges and faces in the model, and construct the pointer array $lhexe(12, M)$, which lists the 12 edges of each hexahedron, and the pointer array $lhexf(6, M)$, which lists the 6 faces of each hexahedron. Use a regular ordering based on that of the nodes, so that the edges, faces, and nodes are easily referenced to one another.

(b) From (a), write a subroutine to construct the update matrices in (11.88) and (11.89). Use a cell-by-cell assembly paradigm for this purpose.

(c) Write another subroutine to implement the field updates of (11.88) and (11.89).

(d) Repeat Project P11.1 using the generalized Yee algorithm.

P11.3 You are given a two-dimensional unstructured mesh composed of general fitted quadrilateral elements. The geometrical description of the mesh is provided by a list of all the nodes that uniquely define the vertices of the quadrilateral elements. The node coordinates are expressed by two one-dimensional arrays dimensioned as $x(N)$ and $y(N)$, where N is the total number of nodes. Another list is provided that gives the set of nodes for each element. Since there are four nodes per quadrilateral, the list is stored in a two-dimensional array that is dimensioned as $lquadn(4, M)$, where M is the total number of quadrilaterals.

(a) Write a subroutine to uniquely determine the edges in the model, and construct the pointer array $lquade(4, M)$, which lists the four edges of each quadrilateral. Use a regular ordering based on that of the nodes so that the edges and nodes are easily referenced to one another.

(b) From (a), write a subroutine to construct the update matrices in (11.100) to (11.105). Use a cell-by-cell assembly paradigm for this purpose. Note that arrays $\{e_t\}$ and $\{h_t\}$ are associated with the two-dimensional grid edges. Further, array $\{e_z\}$ is associated with the nodes, and array $\{h_z\}$ is associated with the quadrilateral cells.

(c) Write another subroutine to implement the field updates of (11.100) to (11.105). Assume that the medium is inhomogeneous in the z-direction only.

(d) Repeat Project P11.1 using the PGY algorithm.

Chapter 12

Bodies of Revolution

Thomas Jurgens, Jeffrey Blaschak, and Gregory Saewert

12.1 INTRODUCTION

This chapter discusses the conformal FDTD modeling of *bodies of revolution* (BOR). The objects considered here are symmetric about an axis, leading to the natural use of cylindrical coordinates. The azimuthal (ϕ) dependence of the fields is expressed as a Fourier series. The algorithm is able to compute solutions for all the Fourier modes, provided it is rerun for each mode. This sort of algorithm can be alternatively labeled as 2.5-dimensional. Since the azimuthal field variation is analytically accounted for, there is no gridding in the ϕ-direction. This means the BOR-FDTD algorithm is two-dimensional in terms of computer resource usage.

Section 12.2 introduces the BOR expansion that will be used in this discussion. Section 12.3 derives the difference equations needed for the computation of field components away from the coordinate axis. Field components that lie on the coordinate axis must be dealt with differently, and are discussed in Section 12.4. Section 12.5 tackles the issue of algorithm stability. An ABC extended from the PML theory discussed in Section 7.4 is presented in Section 12.6. Section 12.7 deals with the application of this method to problems in accelerator physics.

The BOR-FDTD algorithm has been used to model electromagnetic pulse effects [1], electromagnetic wave scattering [2], subsurface interface radar [3], wake fields and impedances of particle accelerators [4–8], optical lenses [9], and guided waves [10].

12.2 FIELD EXPANSION

Our starting point for the derivation of the BOR-FDTD algorithm is the Ampere's and Faraday's laws in integral form:

$$\oint_C \boldsymbol{H}\cdot d\boldsymbol{l} = \iint_S \sigma \boldsymbol{E}\cdot d\boldsymbol{S} + \frac{\partial}{\partial t}\iint_S \boldsymbol{D}\cdot d\boldsymbol{S} \tag{12.1}$$

$$\oint_C \boldsymbol{E}\cdot d\boldsymbol{l} = -\frac{\partial}{\partial t}\iint_S \boldsymbol{B}\cdot d\boldsymbol{S} \tag{12.2}$$

where $\boldsymbol{D} = \varepsilon\boldsymbol{E}$ and $\boldsymbol{B} = \mu\boldsymbol{H}$. For the purposes of this discussion, we let $\sigma = 0$.

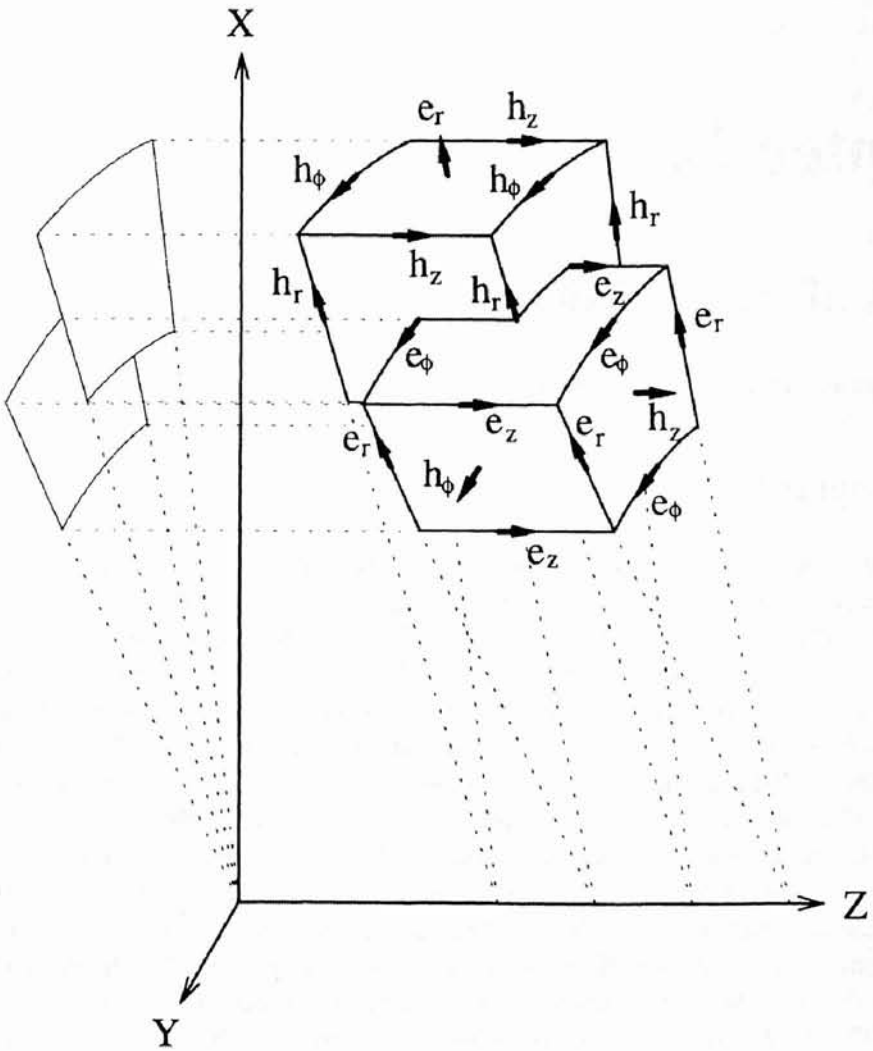


Fig. 12.1 Spatial relationship of the field components for the FDTD unit cell in cylindrical coordinates.

Fig. 12.1 illustrates the discretization of (12.1) and (12.2) with interlinked electric and magnetic path integrals. In contrast with Cartesian coordinates, the integral contours have a different shape depending upon the field component to be updated.

The next step is the analytical introduction of the azimuthal field variation. Assuming linearity of all media, we expand the fields in a Fourier series of sines and cosines:

$$\boldsymbol{E} = \sum_{m=0}^{\infty} \left(\boldsymbol{e}_u \cos m\phi + \boldsymbol{e}_v \sin m\phi \right) \tag{12.3}$$

$$\boldsymbol{H} = \sum_{m=0}^{\infty} \left(\boldsymbol{h}_u \cos m\phi + \boldsymbol{h}_v \sin m\phi \right) \tag{12.4}$$

where m is the mode number, u denotes the Fourier coefficients for the cosinusoidal dependence, and v denotes the coefficients for the sinusoidal dependence. In this expansion, $\boldsymbol{E}$ and $\boldsymbol{H}$ are dependent on r, ϕ, z, and t, while $\boldsymbol{e}_u$, $\boldsymbol{e}_v$, $\boldsymbol{h}_u$, and $\boldsymbol{h}_v$ are dependent on r, z, and t. Note that the dependence of the Fourier coefficients on m is not explicitly expressed, since the algorithm is designed to provide a solution one mode at a time.

12.3 DIFFERENCE EQUATIONS FOR OFF-AXIS CELLS

This section derives the difference equations for the off-axis E- and H-components, including the case where the path integrals are flexed to conform with local surface curvatures.

12.3.1 Ampere's Law Contour Path Integral to Calculate e_r

Fig. 12.2 illustrates the computation of an e_r component using a path integral in the $\phi - z$ coordinate surface.

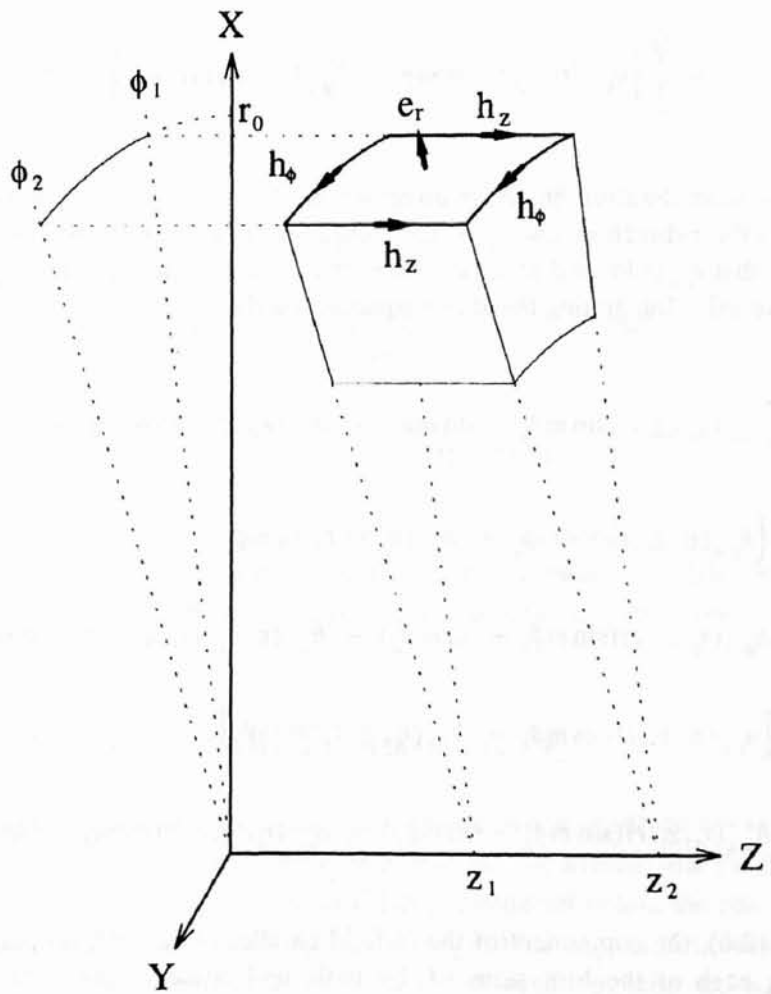


Fig. 12.2 Spatial relationship of the field components for the Ampere's law contour path integral used to calculate e_r.

From Fig. 12.2, Ampere's law yields

$$\varepsilon \frac{\partial}{\partial t} \int_{z_1}^{z_2} \int_{\phi_1}^{\phi_2} \left[e_{r,u}(r_0, zz, t) \cos m\phi + e_{r,v}(r_0, zz, t) \sin m\phi \right] r_0 \, d\phi \, dz$$

$$= \int_{z_1}^{z_2} \left[h_{z,u}(r_0, zz, t) \cos m\phi_2 + h_{z,v}(r_0, zz, t) \sin m\phi_2 \right] dz$$

$$+ \int_{\phi_2}^{\phi_1} \left[h_{\phi,u}(r_0, z_2, t) \cos m\phi + h_{\phi,v}(r_0, z_2, t) \sin m\phi \right] r_0 \, d\phi$$

$$+ \int_{z_2}^{z_1} \left[h_{z,u}(r_0, zz, t) \cos m\phi_1 + h_{z,v}(r_0, zz, t) \sin m\phi_1 \right] dz$$

$$+ \int_{\phi_1}^{\phi_2} \left[h_{\phi,u}(r_0, z_1, t) \cos m\phi + h_{\phi,v}(r_0, z_1, t) \sin m\phi \right] r_0 \, d\phi \qquad (12.5)$$

where the spatial discretization in the z-direction is $\Delta z = z_2 - z_1$, $zz = z_1 + \Delta z/2$, $e_{r,u}$ is the component of $\boldsymbol{e}_u$ in the r-direction, and $e_{r,v}$ is the component of $\boldsymbol{e}_v$ in the r-direction. The notation $e_{r,u}(r_0, zz, t)$ means that $e_{r,u}$ is located at the coordinates r_0 and zz at time t, where r_0 is the distance from the coordinate axis. Integrating the above equation yields

$$\frac{r_0 \varepsilon \Delta z}{m} \frac{\partial}{\partial t} \left[e_{r,u}(r_0, zz, t)(\sin m\phi_2 - \sin m\phi_1) - e_{r,v}(r_0, zz, t)(\cos m\phi_2 - \cos m\phi_1) \right]$$

$$= \Delta z \left[h_{z,u}(r_0, zz, t) \cos m\phi_2 + h_{z,v}(r_0, zz, t) \sin m\phi_2 \right]$$

$$+ \frac{r_0}{m} \left[h_{\phi,u}(r_0, z_2, t)(\sin m\phi_1 - \sin m\phi_2) - h_{\phi,v}(r_0, z_2, t)(\cos m\phi_1 - \cos m\phi_2) \right]$$

$$- \Delta z \left[h_{z,u}(r_0, zz, t) \cos m\phi_1 + h_{z,v}(r_0, zz, t) \sin m\phi_1 \right]$$

$$+ \frac{r_0}{m} \left[h_{\phi,u}(r_0, z_1, t)(\sin m\phi_2 - \sin m\phi_1) - h_{\phi,v}(r_0, z_1, t)(\cos m\phi_2 - \cos m\phi_1) \right] \qquad (12.6)$$

To arrive at (12.6), the component of the H-field parallel to the integral path is assumed to be constant along each of the four sides of the path, and equal to the value of the H-field component located along that side. In addition, the E-field normal to the surface of the patch enclosed by the path is assumed to be constant within the patch, and equal to the value of the E-field component located there.

Since ϕ_1 and ϕ_2 have arbitrary values, the cosine and sine terms can be collected in separate equations, which are listed below:

$$\left[-\Delta z\, h_{z,u}(r_0, zz, t) - \frac{r_0}{m} h_{\phi,v}(r_0, z_2, t) + \frac{r_0}{m} h_{\phi,v}(r_0, z_1, t)\right] \cos m\phi_1$$

$$= \left[\frac{r_0\, \varepsilon\, \Delta z}{m} \frac{\partial}{\partial t} e_{r,v}(r_0, zz, t)\right] \cos m\phi_1 \tag{12.7}$$

$$\left[\Delta z\, h_{z,u}(r_0, zz, t) + \frac{r_0}{m} h_{\phi,v}(r_0, z_2, t) - \frac{r_0}{m} h_{\phi,v}(r_0, z_1, t)\right] \cos m\phi_2$$

$$= \left[-\frac{r_0\, \varepsilon\, \Delta z}{m} \frac{\partial}{\partial t} e_{r,v}(r_0, zz, t)\right] \cos m\phi_2 \tag{12.8}$$

$$\left[-\Delta z\, h_{z,v}(r_0, zz, t) + \frac{r_0}{m} h_{\phi,u}(r_0, z_2, t) - \frac{r_0}{m} h_{\phi,u}(r_0, z_1, t)\right] \sin m\phi_1$$

$$= \left[-\frac{r_0\, \varepsilon\, \Delta z}{m} \frac{\partial}{\partial t} e_{r,u}(r_0, zz, t)\right] \sin m\phi_1 \tag{12.9}$$

$$\left[\Delta z\, h_{z,v}(r_0, zz, t) - \frac{r_0}{m} h_{\phi,u}(r_0, z_2, t) + \frac{r_0}{m} h_{\phi,u}(r_0, z_1, t)\right] \sin m\phi_2$$

$$= \left[\frac{r_0\, \varepsilon\, \Delta z}{m} \frac{\partial}{\partial t} e_{r,u}(r_0, zz, t)\right] \sin m\phi_2 \tag{12.10}$$

Note that (12.7) and (12.8) describe the same relationship between the field components and are therefore redundant, as are (12.9) and (12.10). The resulting two equations are listed as (12.19) and (12.25) in Section 12.3.4.

12.3.2 Ampere's Law Contour Path Integral to Calculate e_ϕ

Fig. 12.3 illustrates the computation of an e_ϕ component using an Ampere's law path integral in the $r-z$ coordinate surface. As with the e_r path integral, we assume that the H-field is constant along each side of the path, and the normal E-field is constant within the patch enclosed by the path. Designating $zz = z_1 + \Delta z/2$, $\Delta r = r_2 - r_1$, $rr = r_1 + \Delta r/2$, $e_{\phi,u}$ as the component of $\boldsymbol{e}_u$ in the ϕ-direction, and $e_{\phi,v}$ as the component of $\boldsymbol{e}_v$ in the ϕ-direction, then the equation for the e_ϕ integral is

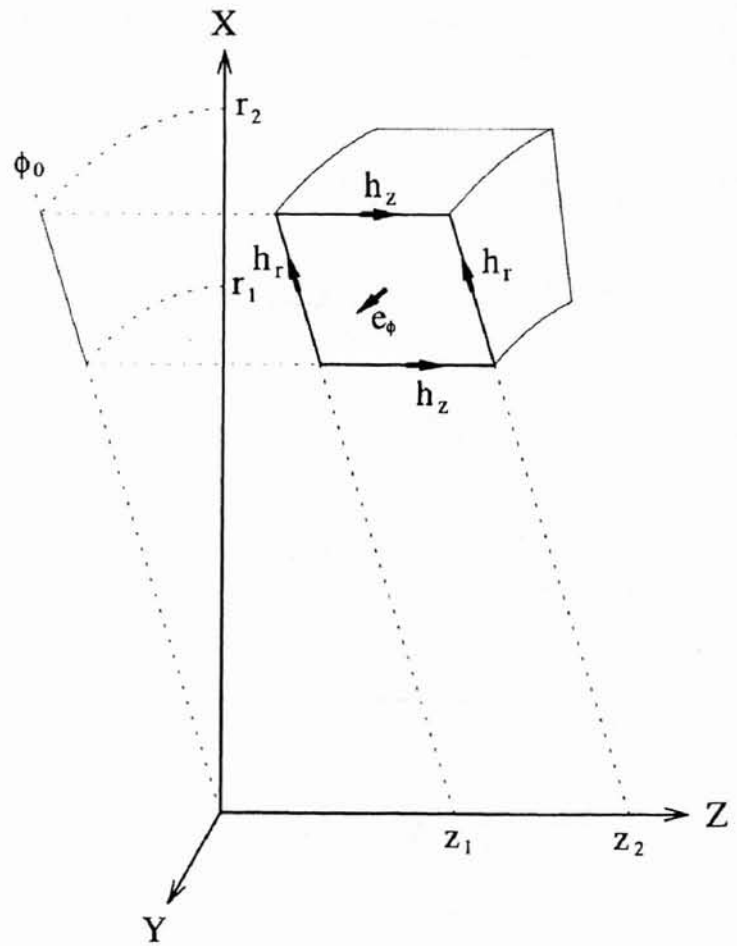


Fig. 12.3 Spatial relationship of the field components for the Ampere's law contour path integral used to calculate e_ϕ.

$$\varepsilon\frac{\partial}{\partial t}\int_{z_1}^{z_2}\int_{r_1}^{r_2}\left[e_{\phi,u}(rr,zz,t)\cos m\phi + e_{\phi,v}(rr,zz,t)\sin m\phi\right]dr\,dz$$

$$= \int_{z_1}^{z_2}\left[h_{z,u}(r_1,zz,t)\cos m\phi + h_{z,v}(r_1,zz,t)\sin m\phi\right]dz$$

$$+ \int_{r_1}^{r_2}\left[h_{r,u}(rr,z_2,t)\cos m\phi + h_{r,v}(rr,z_2,t)\sin m\phi\right]dr$$

$$+ \int_{z_2}^{z_1}\left[h_{z,u}(r_2,zz,t)\cos m\phi + h_{z,v}(r_2,zz,t)\sin m\phi\right]dz$$

$$+ \int_{r_2}^{r_1}\left[h_{r,u}(rr,z_1,t)\cos m\phi + h_{r,v}(rr,z_1,t)\sin m\phi\right]dr \qquad (12.11)$$

Analogous relations exist for $h_{z,u}$, $h_{z,v}$, $h_{r,u}$, and $h_{r,v}$. Now integrating (12.11), we obtain

$$\varepsilon \Delta z \Delta r \frac{\partial}{\partial t}\left[e_{\phi,u}(rr,zz,t)\cos m\phi + e_{\phi,v}(rr,zz,t)\sin m\phi\right]$$

$$= \Delta z\left[h_{z,u}(r_1,zz,t)\cos m\phi + h_{z,v}(r_1,zz,t)\sin m\phi\right]$$

$$+ \Delta r\left[h_{r,u}(rr,z_2,t)\cos m\phi + h_{r,v}(rr,z_2,t)\sin m\phi\right]$$

$$- \Delta z\left[h_{z,u}(r_2,zz,t)\cos m\phi + h_{z,v}(r_2,zz,t)\sin m\phi\right]$$

$$- \Delta r\left[h_{r,u}(rr,z_1,t)\cos m\phi + h_{r,v}(rr,z_1,t)\sin m\phi\right] \qquad (12.12)$$

Since ϕ is arbitrary, (12.20) and (12.26) of Section 12.3.4 follow from (12.12).

12.3.3 Ampere's Law Contour Path Integral to Calculate e_z

Fig. 12.4 illustrates the computation of an e_z component using an Ampere's law path integral in the $r-\phi$ coordinate surface. Again, we assume that the H-field is constant along each side of the path, and the normal E-field is constant within the patch enclosed by the path. The equation for the e_z integral is then

$$\varepsilon \frac{\partial}{\partial t}\int_{r_1}^{r_2}\int_{\phi_1}^{\phi_2}\left[e_{z,u}(rr,z,t)\cos m\phi + e_{z,v}(rr,z,t)\sin m\phi\right] r\, d\phi\, dr$$

$$= \int_{r_1}^{r_2}\left[h_{r,u}(rr,z,t)\cos m\phi_1 + h_{r,v}(rr,z,t)\sin m\phi_1\right] dr$$

$$+ \int_{\phi_1}^{\phi_2}\left[h_{\phi,u}(r_2,z,t)\cos m\phi + h_{\phi,v}(r_2,z,t)\sin m\phi\right] r_2\, d\phi$$

$$+ \int_{r_2}^{r_1}\left[h_{r,u}(rr,z,t)\cos m\phi_2 + h_{r,v}(rr,z,t)\sin m\phi_2\right] dr$$

$$+ \int_{\phi_2}^{\phi_1}\left[h_{\phi,u}(r_1,z,t)\cos m\phi + h_{\phi,v}(r_1,z,t)\sin m\phi\right] r_1\, d\phi \qquad (12.13)$$

Integration of the above equation gives

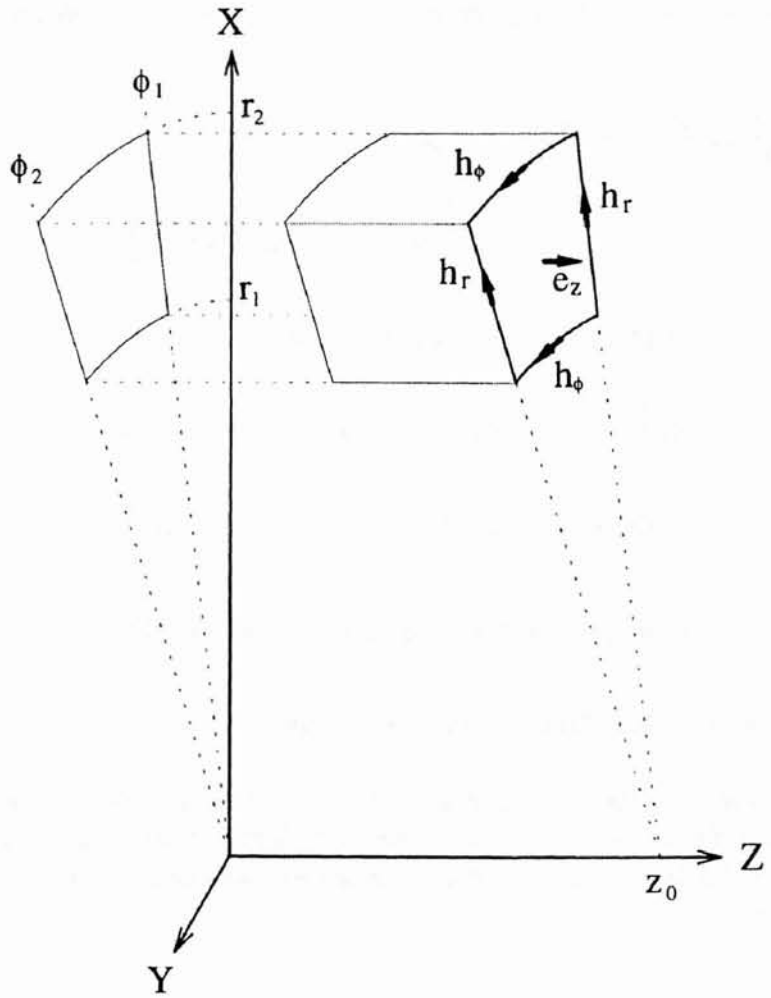


Fig. 12.4 Spatial relationship of the field components for the Ampere's law contour path integral used to calculate e_z.

$$\frac{\varepsilon(r_2^2 - r_1^2)}{2m}\frac{\partial}{\partial t}\left[e_{z,u}(rr,z,t)(\sin m\phi_2 - \sin m\phi_1) - e_{z,v}(rr,z,t)(\cos m\phi_2 - \cos m\phi_1)\right]$$

$$= \Delta r\left[h_{r,u}(rr,z,t)\cos m\phi_1 + h_{r,v}(rr,z,t)\sin m\phi_1\right]$$

$$+ \frac{r_2}{m}\left[h_{\phi,u}(r_2,z,t)(\sin m\phi_2 - \sin m\phi_1) - h_{\phi,v}(r_2,z,t)(\cos m\phi_2 - \cos m\phi_1)\right]$$

$$- \Delta r\left[h_{r,u}(rr,z,t)\cos m\phi_2 + h_{r,v}(rr,z,t)\sin m\phi_2\right]$$

$$- \frac{r_1}{m}\left[h_{\phi,u}(r_1,z,t)(\sin m\phi_2 - \sin m\phi_1) - h_{\phi,v}(r_1,z,t)(\cos m\phi_2 - \cos m\phi_1)\right] \qquad (12.14)$$

In (12.13) and (12.14), $\Delta r = r_2 - r_1$, $rr = r_1 + \Delta r/2$, $e_{z,u}$ is the component of e_u in the z-direction, and $e_{z,v}$ is the component of e_v in the z-direction. Analogous relations exist for $h_{\phi,u}$, $h_{\phi,v}$, $h_{r,u}$, and $h_{r,v}$.

Since ϕ_1 and ϕ_2 are arbitrary, the cosine and sine terms can be collected, giving the following set of equations:

$$\left[\Delta r\, h_{r,u}(rr,z,t) + \frac{r_2}{m} h_{\phi,v}(r_2,z,t) - \frac{r_1}{m} h_{\phi,v}(r_1,z,t)\right] \cos m\phi_1$$

$$= \left[\frac{\varepsilon(r_2^2 - r_1^2)}{2m} \frac{\partial}{\partial t} e_{z,v}(rr,z,t)\right] \cos m\phi_1 \tag{12.15}$$

$$\left[-\Delta r\, h_{r,u}(rr,z,t) - \frac{r_2}{m} h_{\phi,v}(r_2,z,t) + \frac{r_1}{m} h_{\phi,v}(r_1,z,t)\right] \cos m\phi_2$$

$$= \left[-\frac{\varepsilon(r_2^2 - r_1^2)}{2m} \frac{\partial}{\partial t} e_{z,v}(rr,z,t)\right] \cos m\phi_2 \tag{12.16}$$

$$\left[\Delta r\, h_{r,v}(rr,z,t) - \frac{r_2}{m} h_{\phi,u}(r_2,z,t) + \frac{r_1}{m} h_{\phi,u}(r_1,z,t)\right] \sin m\phi_1$$

$$= \left[-\frac{\varepsilon(r_2^2 - r_1^2)}{2m} \frac{\partial}{\partial t} e_{z,u}(rr,z,t)\right] \sin m\phi_1 \tag{12.17}$$

$$\left[-\Delta r\, h_{r,v}(rr,z,t) + \frac{r_2}{m} h_{\phi,u}(r_2,z,t) - \frac{r_1}{m} h_{\phi,u}(r_1,z,t)\right] \sin m\phi_2$$

$$= \left[\frac{\varepsilon(r_2^2 - r_1^2)}{2m} \frac{\partial}{\partial t} e_{z,u}(rr,z,t)\right] \sin m\phi_2 \tag{12.18}$$

Note that, again, (12.15) and (12.16) are redundant, as are (12.17) and (12.18). The two equations that result are (12.21) and (12.26) of Section 12.3.4.

12.3.4 Difference Equations

The difference equations for off-axis field components are given here. The equations derived in the previous three sections along with their H-field counterparts are listed first. This list is divided into two sets of six equations. The first set follows:

$$\frac{\partial}{\partial t}e_{r,v}(r_0,zz,t) = \frac{1}{\varepsilon\,\Delta z}\left[h_{\phi,v}(r_0,z_1,t) - h_{\phi,v}(r_0,z_2,t)\right] - \frac{m}{\varepsilon\,r_0}h_{z,u}(r_0,zz,t) \tag{12.19}$$

$$\frac{\partial}{\partial t}e_{\phi,u}(rr,zz,t) = \frac{1}{\varepsilon\,\Delta r}\left[h_{z,u}(r_1,zz,t) - h_{z,u}(r_2,zz,t)\right] + \frac{1}{\varepsilon\,\Delta z}\left[h_{r,u}(rr,z_2,t) - h_{r,u}(rr,z_1,t)\right] \tag{12.20}$$

$$\frac{\partial}{\partial t}e_{z,v}(rr,z,t) = \frac{2\,r_2}{\varepsilon(r_2^2 - r_1^2)}h_{\phi,v}(r_2,z,t) - \frac{2\,r_1}{\varepsilon(r_2^2 - r_1^2)}h_{\phi,v}(r_1,z,t) + \frac{2\,m\,\Delta r}{\varepsilon(r_2^2 - r_1^2)}h_{r,u}(rr,z,t) \tag{12.21}$$

$$\frac{\partial}{\partial t}h_{r,u}(r_0,zz,t) = \frac{1}{\mu\,\Delta z}\left[e_{\phi,u}(r_0,z_2,t) - e_{\phi,u}(r_0,z_1,t)\right] - \frac{m}{\mu\,r_0}e_{z,v}(r_0,zz,t) \tag{12.22}$$

$$\frac{\partial}{\partial t}h_{\phi,v}(rr,zz,t) = \frac{1}{\mu\,\Delta r}\left[e_{z,v}(r_2,zz,t) - e_{z,v}(r_1,zz,t)\right] + \frac{1}{\mu\,\Delta z}\left[e_{r,v}(rr,z_1,t) - e_{r,v}(rr,z_2,t)\right] \tag{12.23}$$

$$\frac{\partial}{\partial t}h_{z,u}(rr,z,t) = \frac{2\,r_1}{\mu(r_2^2 - r_1^2)}e_{\phi,u}(r_1,z,t) - \frac{2\,r_1}{\mu(r_2^2 - r_1^2)}e_{\phi,u}(r_2,z,t) + \frac{2\,m\,\Delta r}{\mu(r_2^2 - r_1^2)}e_{r,v}(rr,z,t) \tag{12.24}$$

The second set of six equations follows:

$$\frac{\partial}{\partial t}e_{r,u}(r_0,zz,t) = \frac{1}{\varepsilon\,\Delta z}\left[h_{\phi,u}(r_0,z_1,t) - h_{\phi,u}(r_0,z_2,t)\right] + \frac{m}{\varepsilon\,r_0}h_{z,v}(r_0,zz,t) \tag{12.25}$$

$$\frac{\partial}{\partial t}e_{\phi,v}(rr,zz,t) = \frac{1}{\varepsilon\,\Delta r}\left[h_{z,v}(r_1,zz,t) - h_{z,v}(r_2,zz,t)\right] + \frac{1}{\varepsilon\,\Delta z}\left[h_{r,v}(rr,z_2,t) - h_{r,v}(rr,z_1,t)\right] \tag{12.26}$$

$$\frac{\partial}{\partial t}e_{z,u}(rr,z,t) = \frac{2r_2}{\varepsilon(r_2^2 - r_1^2)}h_{\phi,u}(r_2,z,t) - \frac{2r_1}{\varepsilon(r_2^2 - r_1^2)}h_{\phi,u}(r_1,z,t) + \frac{2m\Delta r}{\varepsilon(r_2^2 - r_1^2)}h_{r,u}(rr,z,t) \tag{12.27}$$

$$\frac{\partial}{\partial t}h_{r,v}(r_0,zz,t) = \frac{1}{\mu\Delta z}\left[e_{\phi,v}(r_0,z_2,t) - e_{\phi,v}(r_0,z_1,t)\right] + \frac{m}{\mu r_0}e_{z,u}(r_0,zz,t) \tag{12.28}$$

$$\frac{\partial}{\partial t}h_{\phi,u}(rr,zz,t) = \frac{1}{\mu\Delta r}\left[e_{z,u}(r_2,zz,t) - e_{z,u}(r_1,zz,t)\right] + \frac{1}{\mu\Delta z}\left[e_{r,u}(rr,z_1,t) - e_{r,u}(rr,z_2,t)\right] \tag{12.29}$$

$$\frac{\partial}{\partial t}h_{z,v}(rr,z,t) = \frac{2r_1}{\mu(r_2^2 - r_1^2)}e_{\phi,v}(r_1,z,t) - \frac{2r_1}{\mu(r_2^2 - r_1^2)}e_{\phi,v}(r_2,z,t) + \frac{2m\Delta r}{\mu(r_2^2 - r_1^2)}e_{r,u}(rr,z,t) \tag{12.30}$$

The field components described in (12.19) to (12.24) are rotated 90° with respect to those described in (12.25) to (12.30). Note that the two sets of equations are independent. The independence of the two sets is lost if the material causes coupling between them, such as for anisotropic media.

Without loss of generality, the first set will be used in deriving the difference equations. Since only one set is being considered, the u and v subscripts will be omitted for the remainder of this section, without ambiguity. The time derivative in (12.19) to (12.24) is replaced by a central-difference approximation. Using the notation $i = r/\Delta r$, $k = z/\Delta z$, and $n = t/\Delta t$, the following equations result:

$$e_r\big|_{i+1/2,k+1/2}^{n+1/2} = e_r\big|_{i+1/2,k+1/2}^{n-1/2} + \frac{\Delta t}{\varepsilon\Delta z}\left(h_\phi\big|_{i+1/2,k}^{n} - h_\phi\big|_{i+1/2,k+1}^{n}\right) - \frac{m\Delta t/\varepsilon}{(i+1/2)\Delta r}h_z\big|_{i+1/2,k+1/2}^{n} \tag{12.31}$$

$$e_\phi\big|_{i,k+1/2}^{n+1/2} = e_\phi\big|_{i,k+1/2}^{n-1/2} + \frac{\Delta t}{\varepsilon\Delta r}\left(h_z\big|_{i-1/2,k+1/2}^{n} - h_z\big|_{i+1/2,k+1/2}^{n}\right) + \frac{\Delta t}{\varepsilon\Delta z}\left(h_r\big|_{i,k+1}^{n} - h_r\big|_{i,k}^{n}\right) \tag{12.32}$$

$$e_z\Big|_{i,k}^{n+1/2} = e_z\Big|_{i,k}^{n-1/2} + \frac{(i+1/2)\Delta t}{\varepsilon i \Delta r} h_\phi\Big|_{i+1/2,k}^{n} - \frac{(i-1/2)\Delta t}{\varepsilon i \Delta r} h_\phi\Big|_{i-1/2,k}^{n} + \frac{m\Delta t}{\varepsilon i \Delta r} h_r\Big|_{i,k}^{n} \tag{12.33}$$

$$h_r\Big|_{i,k}^{n+1} = h_r\Big|_{i,k}^{n} + \frac{\Delta t}{\mu \Delta z}\left(e_\phi\Big|_{i,k+1/2}^{n+1/2} - e_\phi\Big|_{i,k-1/2}^{n+1/2}\right) - \frac{m\Delta t}{\mu i \Delta r} e_z\Big|_{i,k}^{n+1/2} \tag{12.34}$$

$$h_\phi\Big|_{i+1/2,k}^{n+1} = h_\phi\Big|_{i+1/2,k}^{n} + \frac{\Delta t}{\mu \Delta r}\left(e_z\Big|_{i+1,k}^{n+1/2} - e_z\Big|_{i,k}^{n+1/2}\right) + \frac{\Delta t}{\mu \Delta z}\left(e_r\Big|_{i+1/2,k-1/2}^{n+1/2} - e_r\Big|_{i+1/2,k+1/2}^{n+1/2}\right) \tag{12.35}$$

$$h_z\Big|_{i+1/2,k+1/2}^{n+1} = h_z\Big|_{i+1/2,k+1/2}^{n} + \frac{i\Delta t}{(i+1/2)\mu \Delta r} e_\phi\Big|_{i,k+1/2}^{n+1/2} - \frac{(i+1)\Delta t}{(i+1/2)\mu \Delta r} e_\phi\Big|_{i+1,k+1/2}^{n+1/2} + \frac{m\Delta t}{(i+1/2)\mu \Delta r} e_r\Big|_{i+1/2,k+1/2}^{n+1/2} \tag{12.36}$$

12.3.5 Surface-Conforming Contour Path Integrals

When a field component lies near a material interface, the associated contour path used to calculate that field component can be deformed to conformally represent the geometry of the interface. The procedure used in implementing this idea is similar to that described for Cartesian coordinates [11]. As an example, Fig. 12.5 is a sketch of the Faraday's law contour path needed to calculate an h_ϕ component near a PEC surface. This results in the following equation:

$$-\mu A \frac{\partial}{\partial t}\left[h_{\phi,u}(rr,zz,t)\cos m\phi + h_{\phi,v}(rr,zz,t)\sin m\phi\right] = l_0\left[e_{z,u}(r_0,zz,t)\cos m\phi + e_{z,v}(r_0,zz,t)\sin m\phi\right] + l_2\left[e_{r,u}(rr,z_2,t)\cos m\phi + e_{r,v}(rr,z_2,t)\sin m\phi\right] - l_1\left[e_{r,u}(rr,z_1,t)\cos m\phi + e_{r,v}(rr,z_1,t)\sin m\phi\right] \tag{12.37}$$

where $l_0 = \Delta z$ is the length of the side of the contour containing e_z; l_1 and l_2 are the lengths of the sides containing e_r at $z = z_1$ and $z = z_2$, respectively; A is the area enclosed by the contour; $rr = r_0 + \Delta r/2$; and $zz = z_1 + \Delta z/2$. Since one side of the contour coincides with the PEC surface where the tangential E-field is zero, the contribution from that side of the path integral is zero.

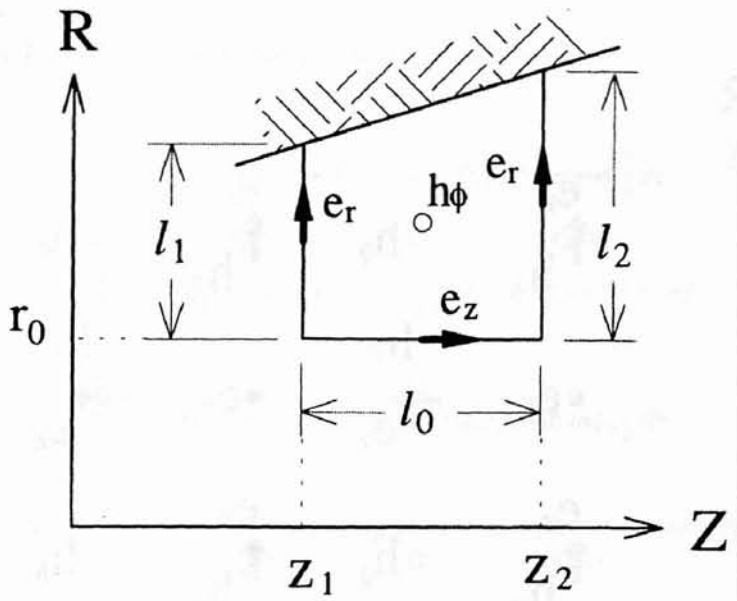


Fig. 12.5 Faraday's law contour path for calculating h_ϕ near a PEC surface.

The sine and cosine terms can be collected in two separate equations, either of which can be used. To be consistent with (12.31) to (12.36), (12.37) is further discretized, giving a difference equation for the $h_{\phi,v}$ component that has the form

$$h_\phi\Big|_{i+1/2,k}^{n+1} = h_\phi\Big|_{i+1/2,k}^{n} - \frac{\Delta t\,\Delta z}{\mu A} e_z\Big|_{i,k}^{n+1/2} + \frac{\Delta t}{\mu A}\left(l_1\, e_r\Big|_{i+1/2,k-1/2}^{n+1/2} - l_2\, e_r\Big|_{i+1/2,k+1/2}^{n+1/2}\right) \tag{12.38}$$

12.4 DIFFERENCE EQUATIONS FOR ON-AXIS CELLS

Fig. 12.6 shows that e_z, e_ϕ, and h_r field components lie on the z-axis. These components cannot be calculated using the integrals described in the previous section. The e_z components along the z-axis need to be computed only for $m = 0$, since these components equal zero for all $m > 0$. This follows from the fact that any constant-r Faraday's law path integral containing $r = 0$ integrates to zero for $m > 0$. The e_ϕ and h_r components need to be computed only for $m = 1$, since they are also identically zero for $m \neq 1$, due to the nature of the field distribution.

12.4.1 Ampere's Law Contour Path Integral to Calculate e_z on the z-Axis

Fig. 12.7 illustrates the geometry for computing an e_z component on the z-axis using an Ampere's law contour path integral in the $r - \phi$ coordinate surface. Upon applying (12.13), we obtain

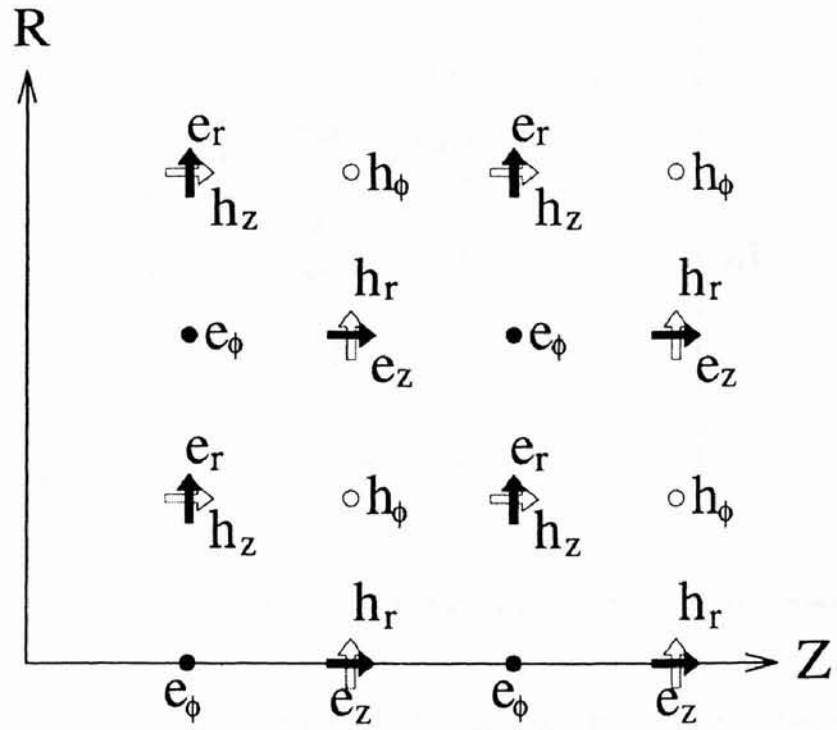


Fig. 12.6 Spatial relationship of the field components in the vicinity of the z-axis.

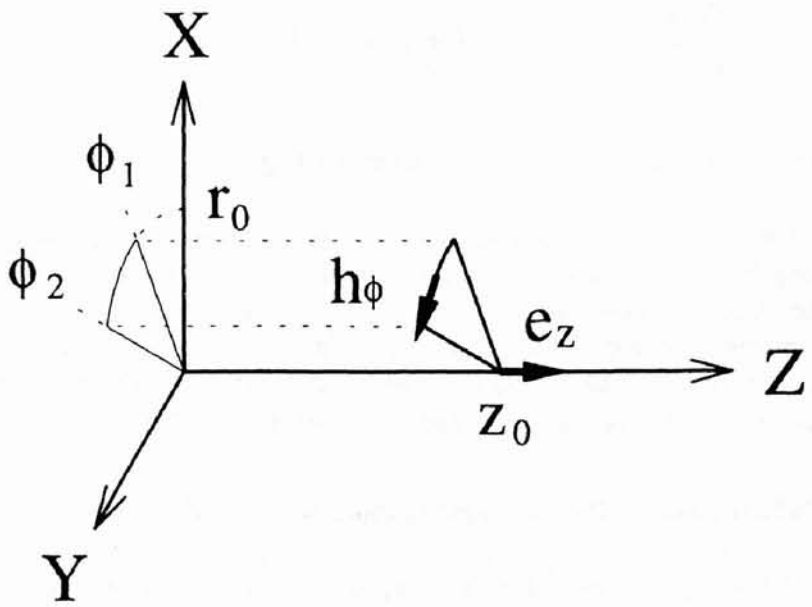


Fig. 12.7 Spatial relationship of the field components for calculating e_z on the z-axis.

$$\varepsilon \frac{\partial}{\partial t} \int_0^{r_0} \int_{\phi_1}^{\phi_2} \left[e_{z,u}(0,z,t)\cos m\phi + e_{z,v}(0,z,t)\sin m\phi \right] r_0 \, d\phi \, dr$$

$$= \int_0^{r_0} \left[h_{r,u}(0,z,t)\cos m\phi_1 + h_{r,v}(0,z,t)\sin m\phi_1 \right] dr$$

$$+ \int_{\phi_1}^{\phi_2} \left[h_{\phi,u}(r_0,z,t)\cos m\phi + h_{\phi,v}(r_0,z,t)\sin m\phi \right] r_0 \, d\phi$$

$$+ \int_0^{r_0} \left[h_{r,u}(0,z,t)\cos m\phi_2 + h_{r,v}(0,z,t)\sin m\phi_2 \right] dr \tag{12.39}$$

where $r_0 = \Delta r/2$. Now we proceed with integrating the above equations, obtaining

$$\frac{\varepsilon r_0^2}{2} \frac{\partial}{\partial t} \left[e_{z,u}(0,z,t)(\sin m\phi_2 - \sin m\phi_1) - e_{z,v}(0,z,t)(\cos m\phi_2 - \cos m\phi_1) \right]$$

$$= m r_0 \left[h_{r,u}(0,z,t)\cos m\phi_1 + h_{r,v}(0,z,t)\sin m\phi_1 \right]$$

$$+ r_0 \left[h_{\phi,u}(r_0,z,t)(\sin m\phi_2 - \sin m\phi_1) - h_{\phi,v}(r_0,z,t)(\cos m\phi_2 - \cos m\phi_1) \right]$$

$$- m r_0 \left[h_{r,u}(0,z,t)\cos m\phi_2 + h_{r,v}(0,z,t)\sin m\phi_2 \right] \tag{12.40}$$

Since ϕ_1 and ϕ_2 are arbitrary, the cosine and sine terms can be collected, resulting in the following set of equations:

$$\left[m r_0 \, h_{r,u}(0,z,t) + r_0 \, h_{\phi,v}(r_0,z,t) \right] \cos m\phi_1 = \left[\frac{\varepsilon r_0^2}{2} \frac{\partial}{\partial t} e_{z,v}(0,z,t) \right] \cos m\phi_1 \tag{12.41}$$

$$\left[-m r_0 \, h_{r,u}(0,z,t) - r_0 \, h_{\phi,v}(r_0,z,t) \right] \cos m\phi_2 = \left[-\frac{\varepsilon r_0^2}{2} \frac{\partial}{\partial t} e_{z,v}(0,z,t) \right] \cos m\phi_2 \tag{12.42}$$

$$\left[m r_0 \, h_{r,v}(0,z,t) - r_0 \, h_{\phi,u}(r_0,z,t) \right] \sin m\phi_1 = \left[-\frac{\varepsilon r_0^2}{2} \frac{\partial}{\partial t} e_{z,u}(0,z,t) \right] \sin m\phi_1 \tag{12.43}$$

$$\left[-m r_0 \, h_{r,v}(0,z,t) + r_0 \, h_{\phi,u}(r_0,z,t) \right] \sin m\phi_2 = \left[\frac{\varepsilon r_0^2}{2} \frac{\partial}{\partial t} e_{z,u}(0,z,t) \right] \sin m\phi_2 \tag{12.44}$$

As with Section 12.3.3, two equations result from the above set. The cosine equation has the azimuthal variation of the field component that matches the choice in Section 12.3.4, so that is the one we will proceed with. It is also worth noting that the area of this contour is less than one-half that of an off-axis contour.

Since this contour is of interest only for a constant ϕ field dependence, let $m = 0$. It follows that there is now no need for information about the h_r component, so its term can be dropped from the equation. From the remaining cosine equations, the following can be extracted:

$$\frac{\partial}{\partial t} e_{z,v}(0,z,t) = \frac{2}{\varepsilon r_0} h_{\phi,v}(r_0,z,t) \tag{12.45}$$

Using indexing consistent with Section 12.3.4, the following difference equation is written:

$$e_z\Big|_{0,k}^{n+1/2} = e_z\Big|_{0,k}^{n-1/2} + \frac{4\Delta t}{\varepsilon \Delta r} h_\phi\Big|_{1/2,k}^{n} \tag{12.46}$$

In the above equation, $\Delta r = 2r_0$.

12.4.2 Ampere's Law Contour Path Integral to Calculate e_ϕ on the z-Axis

Fig. 12.8 illustrates the geometry for computing an e_ϕ component on the z-axis using an Ampere's law contour path integral in the $r-z$ coordinate surface. Note that this component is zero for all mode numbers except $m = 1$.

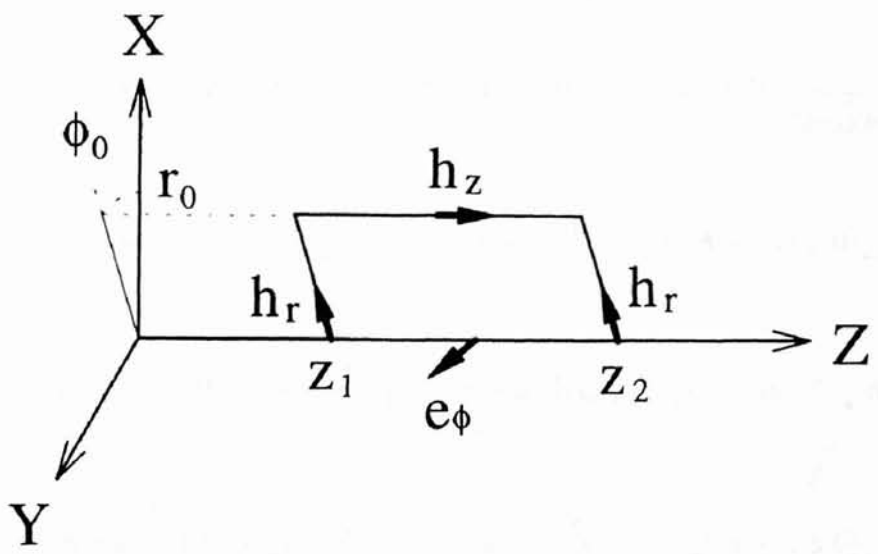


Fig. 12.8 Spatial relationship of the field components for calculating e_ϕ on the z-axis.

Upon applying (12.11), we obtain

$$\varepsilon \frac{\partial}{\partial t} \int_{z_1}^{z_2} \int_{0}^{r_0} \left[e_{\phi,u}(0,zz,t)\cos\phi + e_{\phi,v}(0,zz,t)\sin\phi \right] dr\, dz$$

$$= \int_{z_1}^{z_2} \left[h_{z,u}(0,zz,t)\cos\phi + h_{z,v}(0,zz,t)\sin\phi \right] dz$$

$$+ \int_{0}^{r_0} \left[h_{r,u}(0,z_2,t)\cos\phi + h_{r,v}(0,z_2,t)\sin\phi \right] dr$$

$$+ \int_{z_2}^{z_1} \left[h_{z,u}(r_0,zz,t)\cos\phi + h_{z,v}(r_0,zz,t)\sin\phi \right] dz$$

$$+ \int_{r_0}^{0} \left[h_{r,u}(0,z_1,t)\cos\phi + h_{r,v}(0,z_1,t)\sin\phi \right] dr \qquad (12.47)$$

where $r_0 = \Delta r/2$. The area of this contour is equal to one-half of the area of an off-axis contour. The next step is integrating the above equation:

$$\varepsilon\, \Delta z \frac{\Delta r}{2} \frac{\partial}{\partial t} \left[e_{\phi,u}(0,zz,t)\cos\phi + e_{\phi,v}(0,zz,t)\sin\phi \right]$$

$$= \Delta z \left[h_{z,u}(0,zz,t)\cos\phi + h_{z,v}(0,zz,t)\sin\phi \right]$$

$$+ \frac{\Delta r}{2} \left[h_{r,u}(0,z_2,t)\cos\phi + h_{r,v}(0,z_2,t)\sin\phi \right]$$

$$- \Delta z \left[h_{z,u}(r_0,zz,t)\cos\phi + h_{z,v}(r_0,zz,t)\sin\phi \right]$$

$$- \frac{\Delta r}{2} \left[h_{r,u}(0,z_1,t)\cos\phi + h_{r,v}(0,z_1,t)\sin\phi \right] \qquad (12.48)$$

The cosine and sine terms can be separated, since with $m = 1$, $h_z = 0$ on the axis. Upon collecting the sine and cosine terms into separate equations, we obtain

$$\left[\varepsilon\, \Delta z \frac{\Delta r}{2} \frac{\partial}{\partial t} e_{\phi,u}(0,zz,t) \right] \cos\phi$$

$$= \left\{ -\Delta z\, h_{z,u}(r_0,zz,t) + \frac{\Delta r}{2} \left[h_{r,u}(0,z_2,t) - h_{r,u}(0,z_1,t) \right] \right\} \cos\phi \qquad (12.49)$$

$$\left[\varepsilon\,\Delta z\frac{\Delta r}{2}\frac{\partial}{\partial t}e_{\phi,v}(0,zz,t)\right]\sin\phi$$

$$= \left\{-\Delta z\,h_{z,v}(r_0,zz,t) + \frac{\Delta r}{2}\left[h_{r,v}(0,z_2,t) - h_{r,v}(0,z_1,t)\right]\right\}\sin\phi \tag{12.50}$$

The information in the above equations is redundant, so either one can be used. The following can be written:

$$\frac{\partial}{\partial t}e_{\phi,u}(0,zz,t) = -\frac{2}{\varepsilon\,\Delta r}h_{z,u}(r_0,zz,t) + \frac{1}{\varepsilon\,\Delta z}\left[h_{r,u}(0,z_2,t) - h_{r,u}(0,z_1,t)\right] \tag{12.51}$$

Upon designating $\Delta r = 2r_0$, central-differencing the time derivative, and using indexing consistent with Section 12.3.4, (12.51) can be rewritten as

$$e_\phi\Big|_{0,k+1/2}^{n+1/2} = e_\phi\Big|_{0,k+1/2}^{n-1/2} - \frac{2\,\Delta t}{\varepsilon\,\Delta r}h_z\Big|_{1/2,k+1/2}^{n} + \frac{\Delta t}{\varepsilon\,\Delta z}\left(h_r\Big|_{0,k+1/2}^{n} - h_r\Big|_{0,k}^{n}\right) \tag{12.52}$$

12.4.3 Faraday's Law Calculation of h_r on the z-Axis

The h_r field component is zero on the coordinate axis for all mode numbers except $m = 1$. The starting point for the derivation of the difference equation used to calculate this component is the differential form of Faraday's law:

$$-\frac{\partial}{\partial t}\mu\boldsymbol{H} = \nabla\times\boldsymbol{E} \tag{12.53}$$

Now expand the fields in the ϕ-coordinate:

$$\boldsymbol{H} = \boldsymbol{h}_u\sin\phi + \boldsymbol{h}_v\cos\phi\ ; \qquad \boldsymbol{E} = \boldsymbol{e}_u\sin\phi + \boldsymbol{e}_v\cos\phi \tag{12.54a, b}$$

Substituting into (12.53) yields

$$-\frac{\partial}{\partial t}\mu\boldsymbol{h}_u = \nabla\times\boldsymbol{e}_u + \frac{1}{r}\hat{\boldsymbol{\phi}}\times\boldsymbol{e}_v \tag{12.55}$$

$$-\frac{\partial}{\partial t}\mu\boldsymbol{h}_v = \nabla\times\boldsymbol{e}_v - \frac{1}{r}\hat{\boldsymbol{\phi}}\times\boldsymbol{e}_u \tag{12.56}$$

To be consistent with Section 12.3.4, (12.55) is used. This vector equation contains the following scalar equation for the h_r component:

$$\frac{\partial}{\partial t}\mu h_{r,u} = \frac{\partial}{\partial z}e_{\phi,u} - \frac{1}{r}e_{z,v} \tag{12.57}$$

The next step is to discretize (12.57) at the axis. Computation of the time and space derivatives is straightforward. The e_z term is a measure of the derivative of e_z with respect to ϕ, and is not zero. However, the value of the e_z component is zero at the axis for $m > 0$. Therefore, a nearest-neighbor approximation is used: the e_z value at $r = \Delta r$. This yields

$$h_r\Big|_{0,k}^{n+1} = h_r\Big|_{0,k}^{n} - \frac{\Delta t}{\mu \Delta r} e_z\Big|_{1,k}^{n+1/2} + \frac{\Delta t}{\mu \Delta z}\left(e_\phi\Big|_{0,k+1/2}^{n+1/2} - e_\phi\Big|_{0,k-1/2}^{n+1/2}\right) \tag{12.58}$$

12.5 NUMERICAL STABILITY

Explicit finite-difference schemes have stability restrictions on the choices for the space and time increments. The numerical stability bound for the time-step used in the BOR-FDTD algorithm can be empirically represented as

$$\Delta t \leq \Delta x / sc \tag{12.59}$$

where Δx is the space increment, $s \sim m + 1$ for $m > 0$, and $s = \sqrt{2}$ for $m = 0$. Fig. 12.9 displays the stability factor s versus the mode number m. It has been observed that the stability of the algorithm is very sensitive to the way the field components near the axis are computed.

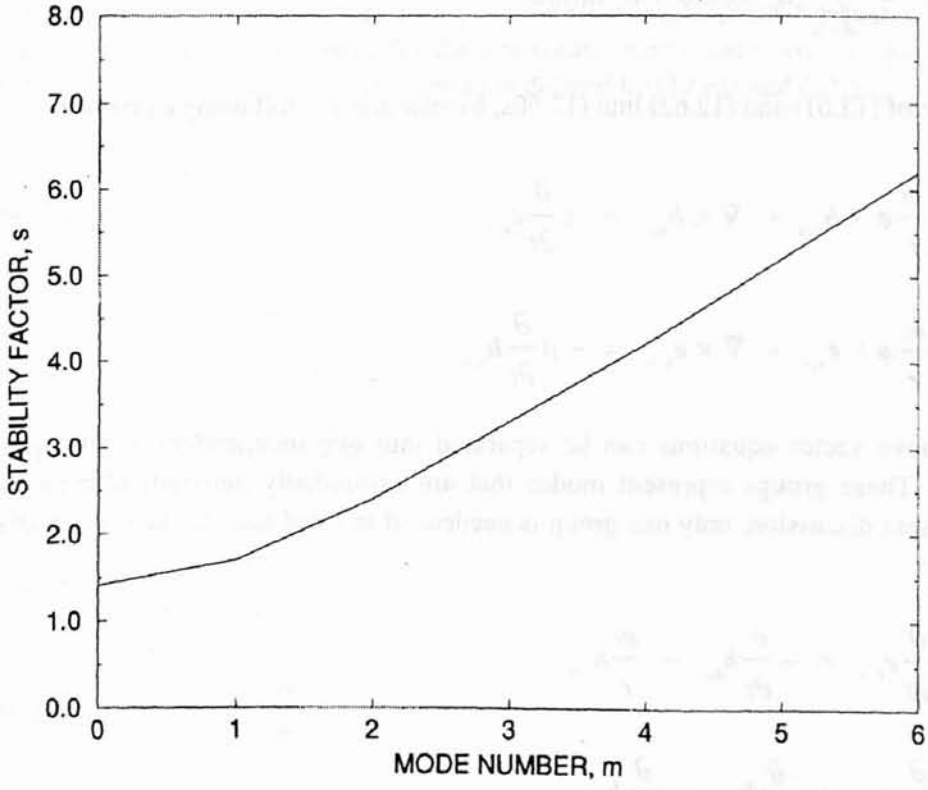


Fig. 12.9 Stability factor versus mode number for the BOR-FDTD algorithm.

12.6 PML ABSORBING BOUNDARY CONDITION

The PML theory developed by Berenger [12] is a major theoretical and practical advance in ABCs for FDTD methods. Chapter 7 discussed in detail PML ABCs for Cartesian space lattices. This section describes an extension of Berenger's PML ABC to terminate a BOR-FDTD grid in the axial *and* radial directions. Sample calculations for cylindrical waveguides are presented.

12.6.1 BOR-FDTD Background

The starting point for deriving the PML conditions for the BOR-FDTD method is Maxwell's curl equations written in cylindrical coordinates:

$$\nabla \times \boldsymbol{H} = \varepsilon \frac{\partial \boldsymbol{E}}{\partial t} \quad ; \qquad \nabla \times \boldsymbol{E} = -\mu \frac{\partial \boldsymbol{H}}{\partial t} \tag{12.60a, b}$$

As before, assume that the fields have a harmonic azimuthal dependence:

$$\boldsymbol{E} = \sum_{m=0}^{\infty} (\boldsymbol{e}_u \cos m\phi + \boldsymbol{e}_v \sin m\phi) \tag{12.61}$$

$$\boldsymbol{H} = \sum_{m=0}^{\infty} (\boldsymbol{h}_u \cos m\phi + \boldsymbol{h}_v \sin m\phi) \tag{12.62}$$

Substitution of (12.61) and (12.62) into (12.60a, b) results in the following equations:

$$\pm \frac{m}{r} \hat{\boldsymbol{\phi}} \times \boldsymbol{h}_{v,u} + \nabla \times \boldsymbol{h}_{u,v} = \varepsilon \frac{\partial}{\partial t} \boldsymbol{e}_{u,v} \tag{12.63}$$

$$\pm \frac{m}{r} \hat{\boldsymbol{\phi}} \times \boldsymbol{e}_{v,u} + \nabla \times \boldsymbol{e}_{u,v} = -\mu \frac{\partial}{\partial t} \boldsymbol{h}_{u,v} \tag{12.64}$$

The above vector equations can be separated into two independent groups of six scalar equations. These groups represent modes that are azimuthally perpendicular to each other. For the present discussion, only one group is needed. It is listed here for later comparison to the PML forms:

$$\varepsilon \frac{\partial}{\partial t} e_{r,v} = -\frac{\partial}{\partial z} h_{\phi,v} - \frac{m}{r} h_{z,u} \tag{12.65a}$$

$$\varepsilon \frac{\partial}{\partial t} e_{\phi,u} = \frac{\partial}{\partial z} h_{r,u} - \frac{\partial}{\partial r} h_{z,u} \tag{12.65b}$$

$$\varepsilon \frac{\partial}{\partial t} e_{z,v} = \frac{1}{r} \frac{\partial}{\partial r}(r h_{\phi,v}) + \frac{m}{r} h_{r,u} \tag{12.65c}$$

$$\mu \frac{\partial}{\partial t} h_{r,u} = \frac{\partial}{\partial z} e_{\phi,u} - \frac{m}{r} e_{z,v} \tag{12.66a}$$

$$\mu \frac{\partial}{\partial t} h_{\phi,v} = -\frac{\partial}{\partial z} e_{r,v} + \frac{\partial}{\partial r} e_{z,v} \tag{12.66b}$$

$$\mu \frac{\partial}{\partial t} h_{z,u} = -\frac{1}{r} \frac{\partial}{\partial r}(r e_{\phi,u}) + \frac{m}{r} e_{r,v} \tag{12.66c}$$

In the above equations, the first subscript refers to the field component's coordinate direction, while the second determines its azimuthal variation. The u and v subscripts will be omitted, since we are only working with one group of these equations.

12.6.2 Extension of PML to the General BOR Case

This section derives a PML absorbing boundary condition for the general body-of-revolution case, which applies and extends the techniques for cylindrical coordinates discussed in [13]. We start with the general BOR representation of Maxwell's time-dependent equations given by (12.65) and (12.66). Specializing these for the sinusoidal steady-state case, we have for each mode $m = 0, 1, \ldots, \infty$ of the harmonic representation defined by (12.61) and (12.62):

$$j\omega\varepsilon \breve{e}_r = -\frac{\partial \breve{h}_\phi}{\partial z} - \frac{m}{r} \breve{h}_z \tag{12.67a}$$

$$j\omega\varepsilon \breve{e}_\phi = \frac{\partial \breve{h}_r}{\partial z} - \frac{\partial \breve{h}_z}{\partial r} \tag{12.67b}$$

$$j\omega\varepsilon \breve{e}_z = \frac{1}{r} \frac{\partial}{\partial r}\left(r \breve{h}_\phi\right) + \frac{m}{r} \breve{h}_r \tag{12.67c}$$

$$j\omega\mu \breve{h}_r = \frac{\partial \breve{e}_\phi}{\partial z} - \frac{m}{r} \breve{e}_z \tag{12.68a}$$

$$j\omega\mu \breve{h}_\phi = -\frac{\partial \breve{e}_r}{\partial z} + \frac{\partial \breve{e}_z}{\partial r} \tag{12.68b}$$

$$j\omega\mu \breve{h}_z = -\frac{1}{r} \frac{\partial}{\partial r}\left(r \breve{e}_\phi\right) + \frac{m}{r} \breve{e}_r \tag{12.68c}$$

Assume that the region $r < r_1$ and $z_0 < z < z_1$ is comprised of ordinary media, and further assume that PML occupies the regions $z < z_0$, $z > z_1$, and $r > r_1$. We wish to map the BOR representation of Maxwell's equations into a complex stretched-coordinate space in the manner of Chapter 7, Section 7.4. To this end, we define the stretched-coordinate variables [14, 15]

$$\tilde{r} \rightarrow r_1 + \int_{r_1}^{r} s_r(r')\,dr' \tag{12.69a}$$

$$\tilde{z} \rightarrow z_0 - \int_{z}^{z_0} s_z(z')\,dz' \tag{12.69b}$$

$$\tilde{z} \rightarrow z_1 + \int_{z_1}^{z} s_z(z')\,dz' \tag{12.69c}$$

Note that in the ordinary media, $s_r = s_z = 1$ with $\tilde{r} \rightarrow r$ and $\tilde{z} \rightarrow z$. Further, this mapping leads to the differential relations

$$\frac{\partial}{\partial \tilde{r}} = \frac{1}{s_r}\frac{\partial}{\partial r} \quad ; \qquad \frac{\partial}{\partial \tilde{z}} = \frac{1}{s_z}\frac{\partial}{\partial z} \tag{12.70a, b}$$

Application of the stretched-coordinate mapping to (12.67) and (12.68) yields the BOR modal representations expressed in the stretched-coordinate space:

$$j\omega\varepsilon\breve{e}_r = -\frac{1}{s_z}\frac{\partial \breve{h}_\phi}{\partial z} - \frac{m}{\tilde{r}}\breve{h}_z \tag{12.71a}$$

$$j\omega\varepsilon\breve{e}_\phi = \frac{1}{s_z}\frac{\partial \breve{h}_r}{\partial z} - \frac{1}{s_r}\frac{\partial \breve{h}_z}{\partial r} \tag{12.71b}$$

$$j\omega\varepsilon\breve{e}_z = \frac{1}{\tilde{r}}\frac{1}{s_r}\frac{\partial}{\partial r}\left(\tilde{r}\breve{h}_\phi\right) + \frac{m}{\tilde{r}}\breve{h}_r \tag{12.71c}$$

$$j\omega\mu\breve{h}_r = \frac{1}{s_z}\frac{\partial \breve{e}_\phi}{\partial z} - \frac{m}{\tilde{r}}\breve{e}_z \tag{12.72a}$$

$$j\omega\mu\breve{h}_\phi = -\frac{1}{s_z}\frac{\partial \breve{e}_r}{\partial z} + \frac{1}{s_r}\frac{\partial \breve{e}_z}{\partial r} \tag{12.72b}$$

$$j\omega\mu\breve{h}_z = -\frac{1}{\tilde{r}}\frac{1}{s_r}\frac{\partial}{\partial r}\left(\tilde{r}\breve{e}_\phi\right) + \frac{m}{\tilde{r}}\breve{e}_r \tag{12.72c}$$

The stretched-coordinate formulation can be viewed as being equivalent to a uniaxial anisotropic medium as the result of additional scalings applied to the e and h fields [16, 17]. For example, let

$$\tilde{\tilde{e}}_r = s_r \tilde{e}_r \; ; \qquad \tilde{\tilde{e}}_\phi = \frac{\tilde{r}}{r}\tilde{e}_\phi \; ; \qquad \tilde{\tilde{e}}_z = s_z \tilde{e}_z \tag{12.73a, b, c}$$

$$\tilde{\tilde{h}}_r = s_r \tilde{h}_r \; ; \qquad \tilde{\tilde{h}}_\phi = \frac{\tilde{r}}{r}\tilde{h}_\phi \; ; \qquad \tilde{\tilde{h}}_z = s_z \tilde{h}_z \tag{12.74a, b, c}$$

The results of applying this scaling to (12.71) and (12.72) are

$$j\omega\varepsilon \frac{s_z \tilde{r}}{s_r r}\tilde{\tilde{e}}_r = -\frac{\partial \tilde{\tilde{h}}_\phi}{\partial z} - \frac{m}{r}\tilde{\tilde{h}}_z \tag{12.75a}$$

$$j\omega\varepsilon \frac{s_r s_z r}{\tilde{r}}\tilde{\tilde{e}}_\phi = \frac{\partial \tilde{\tilde{h}}_r}{\partial z} - \frac{\partial \tilde{\tilde{h}}_z}{\partial r} \tag{12.75b}$$

$$j\omega\varepsilon \frac{s_r \tilde{r}}{s_z r}\tilde{\tilde{e}}_z = \frac{1}{r}\frac{\partial}{\partial r}\left(r\tilde{\tilde{h}}_\phi\right) + \frac{m}{r}\tilde{\tilde{h}}_r \tag{12.75c}$$

$$j\omega\mu \frac{s_z \tilde{r}}{s_r r}\tilde{\tilde{h}}_r = \frac{\partial \tilde{\tilde{e}}_\phi}{\partial z} - \frac{m}{r}\tilde{\tilde{e}}_z \tag{12.76a}$$

$$j\omega\mu \frac{s_r s_z r}{\tilde{r}}\tilde{\tilde{h}}_\phi = -\frac{\partial \tilde{\tilde{e}}_r}{\partial z} + \frac{\partial \tilde{\tilde{e}}_z}{\partial r} \tag{12.76b}$$

$$j\omega\mu \frac{s_r \tilde{r}}{s_z r}\tilde{\tilde{h}}_z = -\frac{1}{r}\frac{\partial}{\partial r}\left(r\tilde{\tilde{e}}_\phi\right) + \frac{m}{r}\tilde{\tilde{e}}_r \tag{12.76c}$$

Equations (12.75) and (12.76) characterize a uniaxial anisotropic medium described by

$$\bar{\bar{s}} = \begin{bmatrix} s_r s_z r/\tilde{r} & 0 & 0 \\ 0 & s_r \tilde{r}/s_z r & 0 \\ 0 & 0 & s_z \tilde{r}/s_r r \end{bmatrix} \tag{12.77}$$

Therefore, the fields need *not* be represented as mapped quantities, and the "~" symbols over each field component can be dropped. In the PML region, s_r and s_z are complex-valued and spatially dependent with

$$s_r(r) = \kappa_r(r) - j\frac{\sigma_r(r)}{\omega\varepsilon_0} \; ; \qquad s_z(z) = \kappa_z(z) - j\frac{\sigma_z(z)}{\omega\varepsilon_0} \tag{12.78a, b}$$

The conductivity σ is used to attenuate propagating modes. Generally, σ varies from zero at the interface of the ordinary medium and the PML to some maximum value at the edge of the grid. The parameter κ provides for the attenuation of evanescent modes, and generally varies from a value of unity at the interface of the ordinary medium and the PML to some maximum value at the grid edge [18]. See also the discussion in Chapter 7, Sections 7.5.3 and 7.6.2.

The specification of s_r and s_z given by (12.78) is frequency-dependent, and thus introduces higher-order time derivatives in the time-domain representation of (12.75) and (12.76). To ensure stability for the FDTD solution, auxiliary variables are introduced that lead to a coupled set of first-order equations. We define the auxiliary variables

$$\breve{p}_r = \frac{\tilde{r}}{s_r r}\breve{e}_r \; ; \qquad \breve{p}_\phi = \frac{s_z r}{\tilde{r}}\breve{e}_\phi \; ; \qquad \breve{p}_z = \frac{\tilde{r}}{s_z r}\breve{e}_z \tag{12.79a, b, c}$$

$$\breve{q}_r = \frac{\tilde{r}}{s_r r}\breve{h}_r \; ; \qquad \breve{q}_\phi = \frac{s_z r}{\tilde{r}}\breve{h}_\phi \; ; \qquad \breve{q}_z = \frac{\tilde{r}}{s_z r}\breve{h}_z \tag{12.80a, b, c}$$

This yields from (12.75), (12.76), and (12.78)

$$j\omega\varepsilon\kappa_z(z)\breve{p}_r + \sigma_z(z)\breve{p}_r = -\frac{\partial\breve{h}_\phi}{\partial z} - \frac{m}{r}\breve{h}_z \tag{12.81a}$$

$$j\omega\varepsilon\kappa_r(r)\breve{p}_\phi + \sigma_r(r)\breve{p}_\phi = \frac{\partial\breve{h}_r}{\partial z} - \frac{\partial\breve{h}_z}{\partial r} \tag{12.81b}$$

$$j\omega\varepsilon\kappa_r(r)\breve{p}_z + \sigma_r(r)\breve{p}_z = \frac{1}{r}\frac{\partial}{\partial r}\left(r\breve{h}_\phi\right) + \frac{m}{r}\breve{h}_r \tag{12.81c}$$

$$j\omega\mu\kappa_z(z)\breve{q}_r + \sigma_z(z)\breve{q}_r = \frac{\partial\breve{e}_\phi}{\partial z} - \frac{m}{r}\breve{e}_z \tag{12.82a}$$

$$j\omega\mu\kappa_r(r)\breve{q}_\phi + \sigma_r(r)\breve{q}_\phi = -\frac{\partial\breve{e}_r}{\partial z} + \frac{\partial\breve{e}_z}{\partial r} \tag{12.82b}$$

$$j\omega\mu\kappa_r(r)\breve{q}_z + \sigma_r(r)\breve{q}_z = -\frac{1}{r}\frac{\partial}{\partial r}\left(r\breve{e}_\phi\right) + \frac{m}{r}\breve{e}_r \tag{12.82c}$$

Finally, the choice of s_r combined with (12.69a) yields the functional form

$$\tilde{r} = \begin{cases} r & r < r_1 \\ r_1 + \int_{r_1}^{r} \kappa_r(r')\,dr' - \dfrac{j}{\omega\varepsilon_0}\int_{r_1}^{r} \sigma_r(r')\,dr' & r > r_1 \end{cases} \tag{12.83}$$

Now, we have all of the information needed to develop a set of time-domain differential equations for the e and h updates suitable for implementation in a BOR-FDTD code. We shall next summarize these differential equations.

Updating of the e Components

First, we perform an inverse Fourier transformation of (12.81). This leads to

$$\varepsilon\kappa_z(z)\frac{\partial p_r}{\partial t} + \sigma_z(z)p_r = -\frac{\partial h_\phi}{\partial z} - \frac{m}{r}h_z \tag{12.84a}$$

$$\varepsilon\kappa_r(r)\frac{\partial p_\phi}{\partial t} + \sigma_r(r)p_\phi = \frac{\partial h_r}{\partial z} - \frac{\partial h_z}{\partial r} \tag{12.84b}$$

$$\varepsilon\kappa_r(r)\frac{\partial p_z}{\partial t} + \sigma_r(r)p_z = \frac{1}{r}\frac{\partial}{\partial r}\left(r h_\phi\right) + \frac{m}{r}h_r \tag{12.84c}$$

These can be time-stepped using the usual Yee leapfrogging to obtain $p^{n+1/2}$ from h^n.

Next, we apply (12.83) to (12.79) to expand the relationships between the components of $\tilde{p}$ and $\tilde{e}$. Inverse Fourier transformation of these relationships yields

$$\left(r_1 + K\right)\frac{\partial e_r}{\partial t} - \frac{S}{\varepsilon_0}e_r = r\kappa_r(r)\frac{\partial p_r}{\partial t} + \frac{r\sigma_r(r)}{\varepsilon_0}p_r \tag{12.85a}$$

$$r\kappa_z(z)\frac{\partial e_\phi}{\partial t} + \frac{r\sigma_z(z)}{\varepsilon_0}e_\phi = \left(r_1 + K\right)\frac{\partial p_\phi}{\partial t} - \frac{S}{\varepsilon_0}p_\phi \tag{12.85b}$$

$$\left(r_1 + K\right)\frac{\partial e_z}{\partial t} - \frac{S}{\varepsilon_0}e_z = r\kappa_z(z)\frac{\partial p_z}{\partial t} + \frac{r\sigma_z(z)}{\varepsilon_0}p_z \tag{12.85c}$$

where

$$K = \int_{r_1}^{r} \kappa_r(r')\,dr' \quad ; \qquad S = \int_{r_1}^{r} \sigma_r(r')\,dr' \tag{12.86a, b}$$

Equation (12.85) allows calculating $e^{n+1/2}$ using data for $p^{n+1/2}$ previously obtained in (12.84). Overall, we see that (12.84) and (12.85) comprise a two-step process for obtaining $e^{n+1/2}$ from h^n. This sets the stage for calculating h^{n+1} from $e^{n+1/2}$, which is summarized next.

Updating of the h Components

First, we perform an inverse Fourier transformation of (12.82). This leads to

$$\mu\kappa_z(z)\frac{\partial q_r}{\partial t} + \sigma_z(z)q_r = \frac{\partial e_\phi}{\partial z} - \frac{m}{r}e_z \tag{12.87a}$$

$$\mu\kappa_r(r)\frac{\partial q_\phi}{\partial t} + \sigma_r(r)q_\phi = -\frac{\partial e_r}{\partial z} + \frac{\partial e_z}{\partial r} \tag{12.87b}$$

$$\mu\kappa_r(r)\frac{\partial q_z}{\partial t} + \sigma_r(r)q_z = -\frac{1}{r}\frac{\partial}{\partial r}\left(re_\phi\right) + \frac{m}{r}e_r \tag{12.87c}$$

These can be time-stepped using the usual Yee leapfrogging to obtain q^{n+1} from $e^{n+1/2}$.

Next, we apply (12.83) to (12.80) to expand the relationships between the components of $\breve{q}$ and $\breve{h}$. Inverse Fourier transformation of these relationships yields

$$\left(r_1 + K\right)\frac{\partial h_r}{\partial t} - \frac{S}{\varepsilon_0}h_r = r\kappa_r(r)\frac{\partial q_r}{\partial t} + \frac{r\sigma_r(r)}{\varepsilon_0}q_r \tag{12.88a}$$

$$r\kappa_z(z)\frac{\partial h_\phi}{\partial t} + \frac{r\sigma_z(z)}{\varepsilon_0}h_\phi = \left(r_1 + K\right)\frac{\partial q_\phi}{\partial t} - \frac{S}{\varepsilon_0}q_\phi \tag{12.88b}$$

$$\left(r_1 + K\right)\frac{\partial h_z}{\partial t} - \frac{S}{\varepsilon_0}h_z = r\kappa_z(z)\frac{\partial q_z}{\partial t} + \frac{r\sigma_z(z)}{\varepsilon_0}q_z \tag{12.88c}$$

where K and S are defined in (12.86). Equation (12.88) allows calculating h^{n+1} using data for q^{n+1} previously obtained in (12.87). We see that (12.87) and (12.88) comprise a two-step calculation process for obtaining h^{n+1} from $e^{n+1/2}$. This completes a full time-step of the leapfrog algorithm.

Grading of the PML Loss Parameters

As discussed in Chapter 7, Section 7.6.2, the loss parameters in the PML must be graded to minimize the discretization error. One practical choice is polynomial grading. For a PML located between $r = r_1$ and $r = r_{max}$ (the maximum-r outer boundary of the grid), the following grading is found to be useful:

$$\sigma_r(r) = \left(\frac{r-r_1}{r_{\max}-r_1}\right)^{\ell} \sigma_{r,\max} \; ; \qquad \kappa_r(r) = 1 + \left(\frac{r-r_1}{r_{\max}-r_1}\right)^{\ell} \left(\kappa_{r,\max}-1\right) \qquad (12.89a, b)$$

This is analogous to the polynomial grading specified in (7.60a, b) for a rectangular space lattice. Similarly, for PML located between $z = z_1$ and $z = z_{\max}$ (the maximum-z outer boundary of the grid), we have

$$\sigma_z(z) = \left(\frac{z-z_1}{z_{\max}-z_1}\right)^{\ell} \sigma_{z,\max} \; ; \qquad \kappa_z(z) = 1 + \left(\frac{z-z_1}{z_{\max}-z_1}\right)^{\ell} \left(\kappa_{z,\max}-1\right) \qquad (12.90a, b)$$

An analogous expression holds for PML located between $z = 0$ and $z = z_0$.

In the corner regions of an r–z grid, wave attenuation is needed in both the r- and z-directions. Specifying the loss parameters of a PML medium as indicated in (12.89) and (12.90) accomplishes this objective.

12.6.3 Examples

We now show two examples of the effectiveness of the PML ABC for the important case of a BOR waveguide. Here, the waveguide is oriented parallel to the z-axis, and the PML is applied to absorb waves propagating in the $\pm z$-directions. No PML is needed at the outer r-boundary of the grid, which is outside of the waveguide.

The first example models a monochromatic, 6-GHz TM_{01} mode propagating in air in a circular, PEC waveguide having a 4.41-GHz cutoff frequency. The PML thickness is 20 cells, and the PML loss σ_z is graded quadratically with depth from the air boundary; that is, $\ell = 2$ in (12.90), for $\sigma_{z,\max} = 10$ S/m and $\kappa_{z,\max} = 1$.

Fig. 12.10 shows the superposition of the propagating waveforms of the TM_{01} mode both in an infinitely long waveguide and in the PML-terminated waveguide. The difference between the two waveforms is unobservable at the scale of the graph, being on the order of only 1 part per 14,000. The reflection coefficient due to the PML termination is calculated to be –83 dB [10].

The second example repeats the first (using the identical PML parameters), but here the circular PEC waveguide has a cutoff frequency of 3.38 GHz and is excited to propagate a 4-GHz TE_{11} mode. This mode illuminates the PML ABC with a substantially different field structure than the TM_{01} mode of the first example, and propagates much more slowly by virtue of its frequency being quite close to cutoff. Nevertheless, the difference between the infinitely long and PML-terminated guide waveforms is again unobservable at the scale of the graph of Fig. 12.11, being on the order of only 1 part per 22,000. The reflection coefficient due to the PML termination is calculated to be –87 dB [10].

12.7 APPLICATION TO PARTICLE ACCELERATOR PHYSICS

In high-energy physics, a charged particle beam is accelerated to nearly c, and collided with another relativistic beam or a fixed target. These collisions result in a shower of subnuclear particles. The study of these subnuclear particles enables the discovery of the innermost nature of matter and the forces that interact with it.

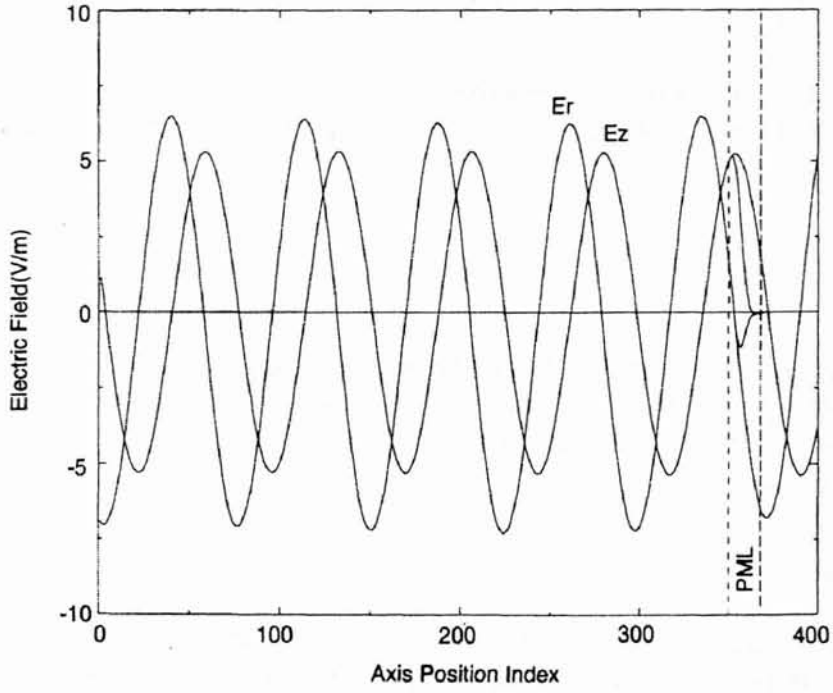


Fig. 12.10 Virtual congruence of the propagating waveforms of a TM_{01} mode in an infinitely long circular PEC waveguide, and the same waveguide terminated by 20-cell-thick PML.

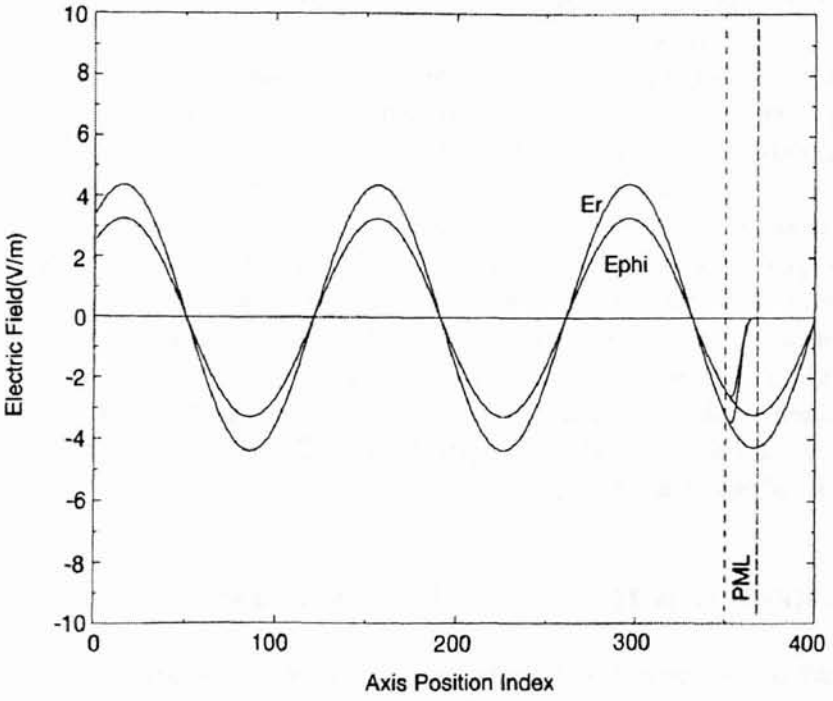


Fig. 12.11 Virtual congruence of the propagating waveforms of a TE_{11} mode in an infinitely long circular PEC waveguide, and the same waveguide terminated by 20-cell-thick PML.

Particle accelerators used in such studies consist of a series of devices that have an evacuated pathway through which the beam passes. Some of the possible devices in a particle accelerator include empty beam pipes, plates, cavities, and magnets. The beam itself is usually a series of particle bunches, rather than a continuous particle stream. Particles in these bunches typically are electrons or protons.

When a beam bunch passes through a device, electromagnetic fields are excited in that device. These excitations are called wake fields, and are analogous to the waves that make up the wake of a boat traveling through water. Wake fields that persist in the device after the beam bunch has exited can affect subsequent bunches. Given the right conditions, the wake fields can even affect the trailing portion of the bunch that excited them. As a result of these mechanisms, the beam could become unstable and break up.

This section discusses the application of FDTD methods to particle accelerator physics. In this technology area, the FDTD method is most commonly used for the calculation of wake fields and impedances. The section begins with a brief introduction to terms and concepts used in the wake-field analysis of particle accelerators.

12.7.1 Definitions and Concepts

Let a particle with charge q travel at essentially the free-space speed of light $v = c$ along the axis of a circular PEC pipe of radius b. The resulting electromagnetic field distribution inside the pipe is Lorentz-contracted into a flat disk. For example, the radial E-field distribution approaches [19–21]

$$E_r = \frac{q}{2\pi\varepsilon_0 r}\delta(z - ct) \tag{12.91}$$

Note that there are no fields left behind the particle; they are located only at $z = ct$. This situation does not change if the particle is moving parallel to, and offset from, the beam pipe axis; again, no wake fields are excited. However, wake fields are excited if the pipe is not a perfect conductor or if there is a change in the pipe's radius.

The particles that generate the wake field follow a prescribed path and are not affected by any surrounding fields. These kinds of particles are described as rigid or stiff. If the z-dependence of the particle distribution is not impulsive, then the fields generated by the front of the bunch can affect the tail of the bunch. Accounting for this type of interaction requires a self-consistent formulation of the problem where the beams are not rigid. *Particle-in-cell* (PIC) algorithms have this feature [22, 23].

In order to include charged-particle movement in the FDTD formalism, a current source term is added to Ampere's law:

$$\oint_C \boldsymbol{H}\cdot d\boldsymbol{l} = \varepsilon\frac{\partial}{\partial t}\iint_S \boldsymbol{E}\cdot d\boldsymbol{S} + \iint_S \boldsymbol{J}\cdot d\boldsymbol{S} \tag{12.92}$$

Let the current source in (12.92) be due to a bunch of charged particles traveling in the z-direction at velocity $v = c$. Following the BOR modal decomposition used in this chapter, the charged-particle current is expressed as a sum of azimuthal moments

$$J_z(z_0) = \sum_{m=0}^{\infty} \frac{c\rho(z_0)\,\delta(r-a)}{\pi a(1+\delta_{m0})} \cos m\phi \tag{12.93}$$

where

$$\delta_{m0} = \begin{cases} 1 & m = 0 \\ 0 & m \neq 0 \end{cases} \tag{12.94}$$

and $z_0 = ct - z$. The azimuthal moments of the bunch are rings of charge with a radius of a. The parameter ρ is the bunch charge density function in the z-direction.

Wake functions are the descriptions of the fields produced when the bunch is a point charge with a charge density profile of $\rho(z_0) = q\delta(z_0)$. They are proportional to the forces that would be encountered by a witness charge. The longitudinal and transverse wake functions of the m'th mode, $w_{m\|}$ and $w_{m\perp}$, are given as

$$w_{m\|}(s) = -\frac{e_{z,u}}{q a^m r_t^m} \quad ; \qquad w_{m\perp}(s) = \frac{e_{r,u} + c\mu h_{\phi,u}}{q a^m m r_t^{m-1}} \tag{12.95a, b}$$

In (12.95a, b), the witness charge travels behind the point bunch at a distance s with a velocity $v = c$, and along a path parallel to the pipe axis with an axial offset of r_t. The e- and h-field components are defined in Section 12.3. The transverse and longitudinal wake functions are related by [24]:

$$w_{m\|}(s) = \frac{\partial}{\partial z_0} w_{m\perp}(s) \tag{12.96}$$

Wake functions are time-domain descriptions of the fields. Since they originate from a point charge, they can be thought of as the Green function of the device. This description can be expressed in the frequency domain as longitudinal and transverse impedances. The longitudinal impedance is defined as the Fourier transform of the longitudinal wake function:

$$Z_{m\|}(\omega) = \int_{-\infty}^{\infty} w_{m\|}(\tau) e^{-j\omega\tau}\, d\tau \tag{12.97}$$

Likewise, the transverse impedance is defined as the Fourier transform of the transverse wake function:

$$Z_{m\perp}(\omega) = -\frac{1}{j} \int_{-\infty}^{\infty} w_{m\perp}(\tau) e^{-j\omega\tau}\, d\tau \tag{12.98}$$

From a computational point of view, (12.95a, b) are not calculable since the FDTD grid cannot support a delta distribution. Only bunches with a finite frequency content can be modeled.

The wake potential of a bunch (of finite length) is defined as the average effect on the test particle resulting from its traversal through a wake-field–filled device of length L. It is dependent upon the device's geometry and the charge distribution that created the wake. The longitudinal and transverse wake potential for a given mode are given by

$$W_{m\,\|}(s) = -\frac{1}{Q_0\, a^m\, r_t^m\, L} \int_0^L e_{z,u}(z, r_t, t)\Big|_{t\,=\,(s+z)/c} dz \tag{12.99}$$

$$W_{m\perp}(s) = -\frac{1}{Q_0\, a^m\, m r_t^{m-1}\, L} \int_0^L \left[e_{r,u}(z, r_t, t) + c\,\mu\, h_{\phi,u}\right]\Big|_{t\,=\,(s+z)/c} dz \tag{12.100}$$

where Q_0 is the total charge in the bunch. The Fourier transforms of the bunch wake potentials and the charge distribution are denoted as

$$\hat{Z}_m(\omega) = \int W_m(s) e^{-j\omega s}\, ds \tag{12.101}$$

$$\rho(\omega) = \frac{1}{2\pi} \int \rho(z_0) e^{-j\omega z_0}\, dz_0 \tag{12.102}$$

Therefore, the device impedance can be extracted through the equation

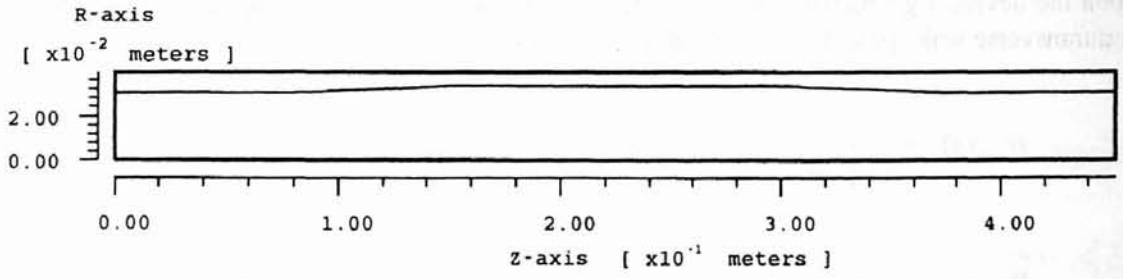
$$Z_m(\omega) = \frac{\hat{Z}_m(\omega)}{2\pi\rho(\omega)} \tag{12.103}$$

We note that since ρ has a finite bandwidth, the accuracy range of the impedance values is also bandlimited. In addition, for the impedance to have the correct phase, both $W(s)$ and $\rho(z_0)$ must be calculated from the same origin.

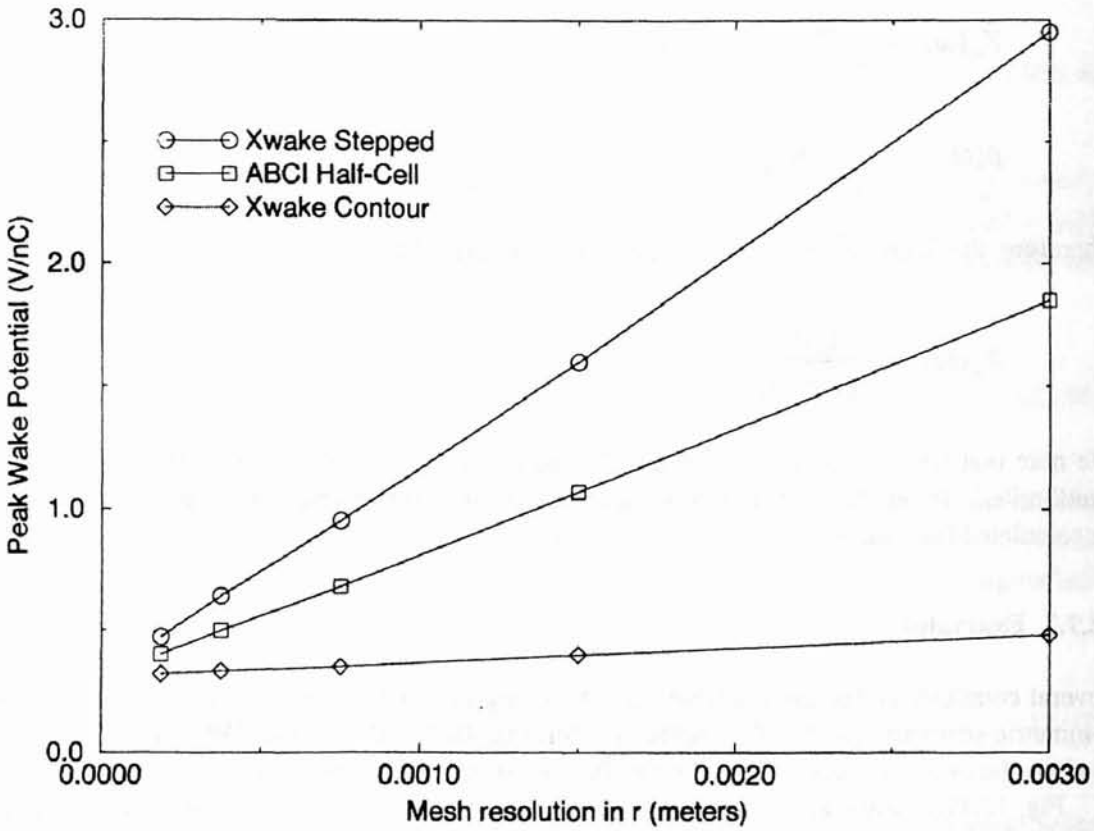
12.7.2 Examples

Several computer codes are available for calculating impedances and wake fields in azimuthally symmetric structures [4–8]. The surface-conforming BOR-FDTD code XWAKE [4] was used to produce the two examples presented here. In both examples, the mode number m equals zero.

Fig. 12.12(a) illustrates the geometry of the first example, a 14.8-cm-long section of circular PEC beam pipe of radius 3.3 cm. This pipe is provided with tapered input and output sections of angle 2.4° that narrow down to a 3.0-cm radius. Fig. 12.12(b) shows the results of a convergence study (versus grid resolution) for the peak longitudinal wake potential derived from the FDTD field data using (12.99). This figure compares three methods of approximating the surfaces of the tapered waveguide sections: ordinary staircasing, diagonal split-cells, and surface-conforming contour-path integrals (Section 12.3.5 and Fig. 12.5). While all three surface approximations converge to the same peak value of wake potential, the use of the surface-conforming contour-path integrals provides as much as a 64:1 advantage over staircasing in memory usage, and a 512:1 advantage in run time.

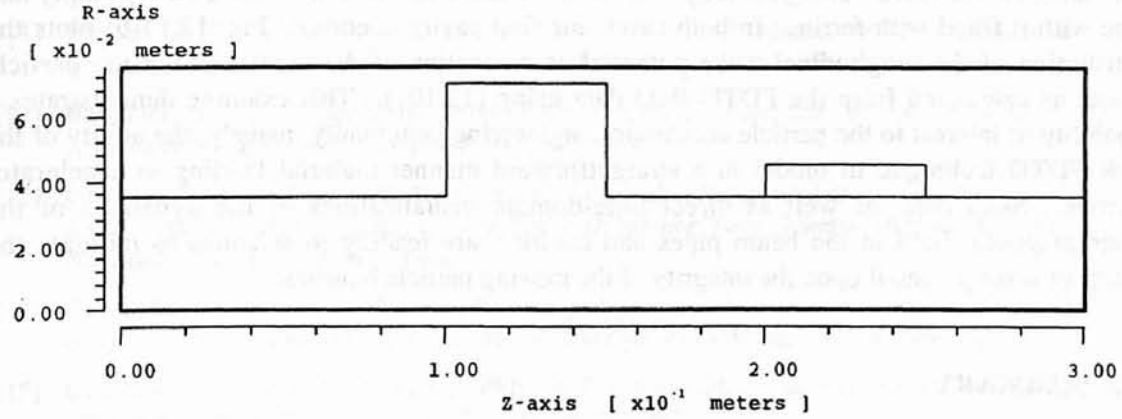


(a) Longitudinal cross section of tapered cylindrical beam pipe.

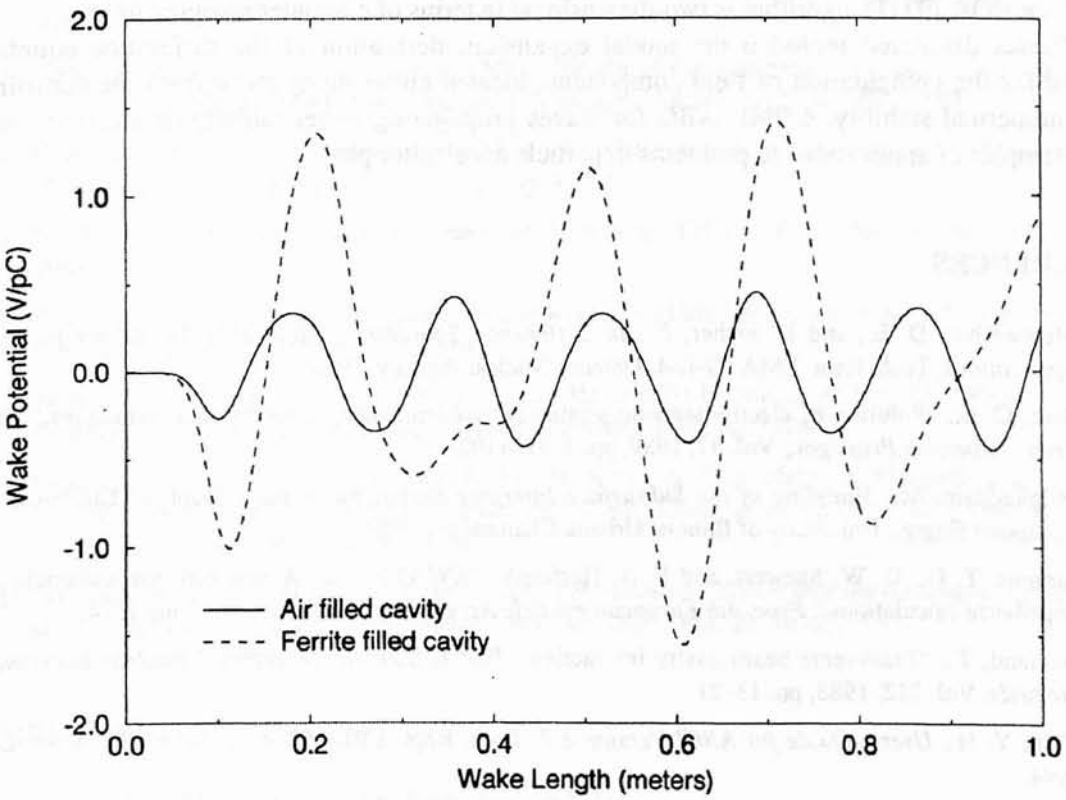


(b) Calculated peak wake potential versus grid resolution.

Fig. 12.12 Convergence study of a slowly tapered beam pipe comparing staircasing, diagonal split-cell, and surface-conforming contour-path models of the taper used in the BOR-FDTD code. For all cases, $\Delta z = 0.004$m.



(a) Upper half longitudinal cross section of ferrite-loaded cylindrical cavity. Mirror-image symmetry about the longitudinal axis is assumed.



(b) Calculated wake potential versus wake length.

Fig. 12.13 The wake potential of a ferrite-loaded particle accelerator cavity calculated by the BOR-FDTD code. Ferrite permeability $\mu_r = 10$; grid resolution $\Delta r = \Delta z = 0.001\,\text{m}$.

Fig. 12.13(a) illustrates the geometry of the second example, a section of PEC beam pipe containing two cavities. This geometry is modeled twice, once with the second cavity empty and once with it filled with ferrite. In both cases, the first cavity is empty. Fig. 12.13(b) plots the distribution of the longitudinal wake potential as a function of distance s behind the particle bunch, as calculated from the FDTD field data using (12.101). This example demonstrates a capability of interest to the particle accelerator engineering community, namely, the ability of the BOR-FDTD technique to model in a straightforward manner material loading in accelerator cavities. Such data, as well as direct time-domain visualizations of the dynamics of the electromagnetic field in the beam pipes and cavities, are leading to schemes to mitigate the effects of wake potential upon the integrity of the moving particle bunches.

12.8 SUMMARY

This chapter discussed FDTD modeling of bodies of revolution. Such objects are rotationally symmetric about an axis. This leads to the natural use of cylindrical coordinates and expression of the azimuthal dependence of the fields using a Fourier series. The algorithm is able to compute solutions for all the Fourier modes, provided it is rerun for each mode. Since the azimuthal field variation is analytically accounted, there is no gridding in the ϕ-direction. Thus, the BOR-FDTD algorithm is two-dimensional in terms of computer resource usage.

Topics discussed included the modal expansion, derivation of the difference equations needed for the computation of field components located either on or away from the coordinate axis, numerical stability, a PML ABC for waves propagating either radially or along the axis, and examples of applications to problems in particle accelerator physics.

REFERENCES

[1] Merewether, D. E., and R. Fisher, *Finite Difference Solution of Maxwell's Equations for EMP Applications*, Tech. Rept. EMA-79-R-4, Defense Nuclear Agency, 1980.

[2] Britt, C. L., "Solution of electromagnetic scattering problems using time domain techniques," *IEEE Trans. Antennas Propagat.*, Vol. 37, 1989, pp. 1181–1192.

[3] Moghaddam, M., *Modeling of the Subsurface Interface Radar*, Ph.D. diss., Dept. of Electrical and Computer Engrg., University of Illinois, Urbana-Champaign, 1990.

[4] Jurgens, T. G., G. W. Saewert, and F. A. Harfoush, "XWAKE 1.0: A new tool for wakefield and impedance calculations," *Proc. 4th European Particle Accelerator Conf.*, London, June 1994.

[5] Weiland, T., "Transverse beam cavity interaction. Part I: Short-range forces," *Nuclear Instruments Methods*, Vol. 212, 1983, pp. 13–21.

[6] Chin, Y. H., *User's Guide for ABCI Version 8.7*, Tech. Rept. LBL-35258, Lawrence Berkeley Lab., 1994.

[7] Shang, C. C., and J. F. DeFord, "Modified Yee solutions in the AMOS wakefield code," *Proc. 1990 Linear Accelerator Conf.*, Albuquerque, Sept. 1990, pp. 378–380.

[8] Aharonian, G., R. Meller, and R. H. Siemann, "Transverse wakefield calculations," *Nuclear Instruments Methods*, Vol. 212, 1983, pp. 23–35.

[9] Davidson, D. B., and R. W. Ziolkowski, "Body of revolution FDTD modeling of space time focusing by a three dimensional lens," *J. Opt. Soc. America*, June 1993.

[10] Jurgens, T. G., "A broadband absorbing boundary condition for the FDTD modeling of circular waveguides," *Proc. 1995 IEEE Intl. Microwave Symp.*, Orlando, FL, May 1995.

[11] Jurgens, T. G., A. Taflove, K. R. Umashankar, and T. G. Moore, "FDTD modeling of curved surfaces," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 357–366.

[12] Berenger, J. P., "A perfectly matched layer for the absorption of electromagnetic waves," *J. Computational Physics*, Vol. 114, 1994, pp. 185–200.

[13] Gedney, S. D., "The perfectly matched layer absorbing medium," Ch. 5 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, A. Taflove, (ed.), Norwood, MA: Artech House, 1998.

[14] Chew, W. C., and W. H. Weedon, "A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates," *Microwave Optical Tech. Lett.*, Vol. 7, 1994, pp. 599–604.

[15] Chew, W. C., and F. L. Teixeira, "PML-FDTD in cylindrical and spherical coordinates," *IEEE Microwave Guided Wave Lett.*, Vol. 7, 1997, pp. 285–287.

[16] 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 Propagat.*, Vol. 43, 1995, pp. 1460–1463.

[17] Maloney, J. M., M. Kesler, and G. S. Smith, "Generalization of PML to cylindrical geometries," *Proc. 13th Ann. Review Progress Applied Computational Electromagnetics*, Monterey, CA, 1997, Vol. 2, pp. 900–908.

[18] Berenger, J. P., "Evanescent waves in PMLs: Origin of the numerical reflection in wave-structure interaction problems," *IEEE Trans. Antennas Propagat.*, Vol. 47, 1999, pp. 1497–1503.

[19] Chao, A. W., "Coherent instabilities of a relativistic bunched beam," *AIP Conf. Proc. 105, The Physics of Particle Accelerators*, 1983, pp. 353–523.

[20] Ng, K.-Y., *Fields, Impedances and Structures*, Tech. Rept. FN-443, Fermi National Accelerator Lab., 1987.

[21] Wilson, P. B., "Introduction to wakefields and wake potentials," *AIP Conf. Proc. 184, The Physics of Particle Accelerators*, 1989, pp. 525–564.

[22] Weiland, T., "Solving Maxwell's equations in 3D and 2D by means of MAFIA," *Proc. Conf. on Computer Codes and the Linear Accelerator Community*, 1990, pp. 3–26.

[23] Chang, C. L., D. Chernin, A. Drobot, K. Ko, M. Kress, A. Mankofsky, A. Mondelli and J. Petillo, "Three dimensional modeling of accelerators," *Proc. Conf. on Computer Codes and the Linear Accelerator Community*, 1990, pp. 27–56.

[24] Panofsky, W. K. H., and W. A. Wenzel, "Some considerations concerning the transverse deflection of charged particles in radio frequency fields," *Rev. Sci. Instr.*, Vol. 27, 1956, p. 967.

PROBLEMS

12.1 Using Fig. 12.2, derive (12.5) from Ampere's law.

12.2 Integrate (12.5) to obtain (12.6).

12.3 Collect the sine and cosine terms of (12.6) to derive (12.7) to (12.10).

12.4 Derive (12.22) to (12.24) from Faraday's law.

12.5 Derive equations analogous to (12.31) to (12.33) for $\sigma \neq 0$.

12.6 Using the method of Section 12.3.5, obtain expressions analogous to (12.37) and (12.38) for h_r, given a smooth PEC contour in the $r-z$ plane that is to be represented conformally.

12.7 Derive (12.46) from (12.39), filling in all of the intermediate steps.

12.8 Derive (12.52) from (12.47), filling in all of the intermediate steps.

12.9 Derive (12.58) from (12.53) and (12.54), filling in all of the intermediate steps.

12.10 Derive the difference equations of a PML medium used to terminate a PEC circular waveguide.

PROJECTS

P12.1 Write a BOR-FDTD computer code that models the propagation of the TE_{11} mode in a circular PEC waveguide. Choose the waveguide radius and length carefully so that the proper mode can evolve from the excitation condition. Verify the accuracy of the model by comparing the computed transverse E-field distribution with the exact solution.

P12.2 Add a PML ABC to the waveguide model of Project P12.1. Replicate the results of Figs. 12.10 and 12.11.

Chapter 13

Periodic Structures

James Maloney and Morris Kesler

13.1 INTRODUCTION

Many structures of electromagnetic interest possess a periodicity in one or more dimensions. For example, a *frequency selective surface* (FSS) is commonly used in a radome to control the energy that reaches an antenna. A typical FSS consists of one or more layers of material, each of which is formed from an element periodically replicated in two dimensions. Another periodic structure that has received considerable attention in recent years is the *electromagnetic bandgap* (EBG) structure. An EBG structure is a periodic construct that has frequency regions in which electromagnetic wave propagation is forbidden. A third type of structure that can be considered periodic is an antenna array. If large enough, many of the important parameters of an array can be analyzed by assuming that the structure is periodic.

The FDTD technique can be applied to the analysis of periodic structures such as the ones mentioned above. Such structures often have fine details at the element level that must be accurately modeled in order to predict the correct electromagnetic behavior. Coupled with the fact that the overall structure may consist of many replicas of the basic element, this level of detail leads to computational problems that are unmanageable. One way to alleviate the computational burden is to model only the individual element, and use boundary conditions to simulate the effect of periodic replication.

As an example of a periodic structure, consider the two-dimensional, electromagnetic screen shown in Fig. 13.1. The screen consists of infinitely long, parallel conducting bars of width w, thickness d, and separated by a gap g. The illumination is an electromagnetic plane wave with the electric field parallel to the bars (z polarized), and the propagation direction at angle ϕ_i with respect to the x-axis, as shown in the figure. This is a two-dimensional electromagnetic problem, since only the E_z, H_x, and H_y fields are nonzero. Because of the periodicity inherent in the geometry, only the fields in the "unit cell" are unique and need to be determined. Dashed lines in the figure denote the edges of the unit cell.

Now, consider discretizing the solution space within the unit cell using a traditional Yee lattice. Field components that can be updated using the fields in the unit cell are denoted by solid black symbols, while those that cannot be updated are denoted by white symbols. The field components that cannot be updated instead can be determined using periodic boundary conditions that relate the fields on one side of the unit cell to those on the other side. Of course, this relationship depends on the angle at which the incident field impinges on the structure. For the geometry shown, the relationship is expressed in the phasor domain by

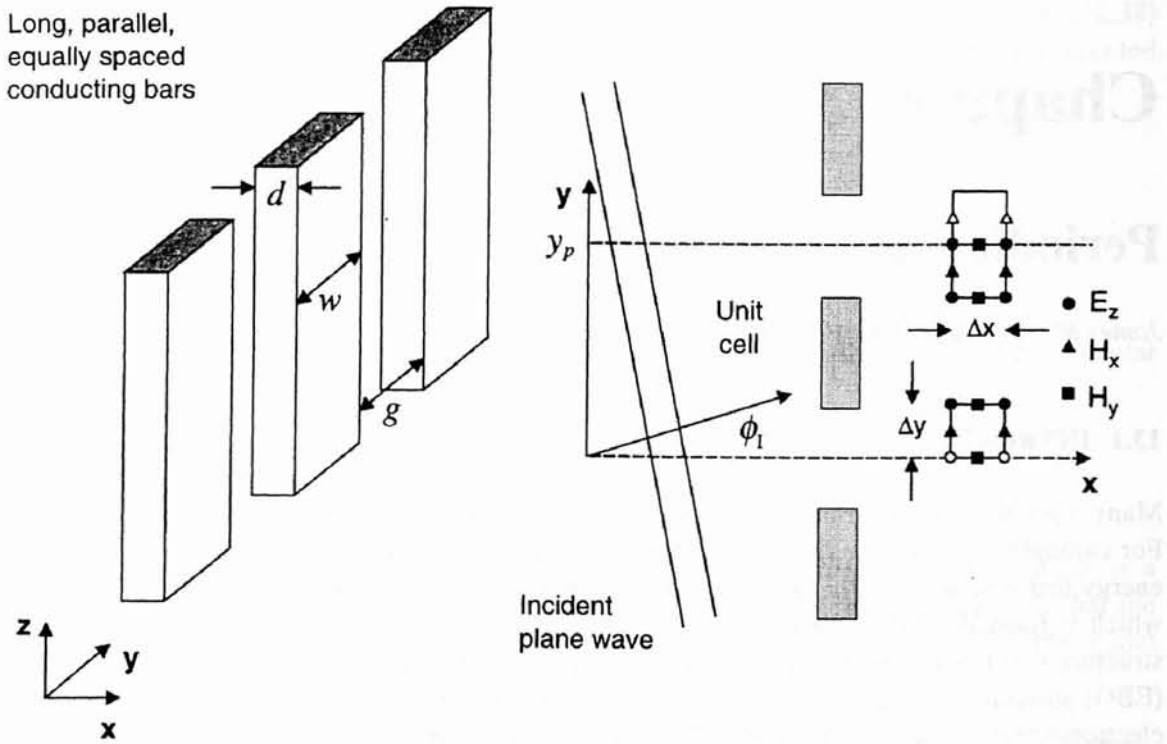


Fig. 13.1 Two-dimensional electromagnetic screen (left) and a top view of the structure showing the field components at the edges of the unit cell.

$$\breve{H}_x\left(x,\ y = y_p + \Delta y/2\right) = \breve{H}_x\left(x,\ y = \Delta y/2\right) \exp\left(-jk_y y_p\right) \tag{13.1a}$$

$$\breve{E}_z(x,\ y = 0) = \breve{E}_z\left(x,\ y = y_p\right) \exp\left(+jk_y y_p\right) \tag{13.1b}$$

where $k_y = k_0 \sin\phi_1$ and k_0 is the free-space wavenumber. FDTD requires the boundary conditions to be expressed in the time domain. In the time domain, (13.1) becomes

$$H_x\left(x,\ y = y_p + \Delta y/2,\ t\right) = H_x\left(x,\ y = \Delta y/2,\ t - y_p \sin\phi_1\right) \tag{13.2a}$$

$$E_z(x,\ y = 0,\ t) = E_z\left(x,\ y = y_p,\ t + y_p \sin\phi_1\right) \tag{13.2b}$$

Equation (13.2) shows that the periodic boundary condition involves using both previous as well as future field components. Specifically, the H_x values at $y = y_p + \Delta y/2$ are equal to previous values of H_x at $y = \Delta y/2$, while the E_z values at $y = 0$ are equal to future values of E_z at $y = y_p$. Previous values can simply be saved using additional computational storage. However, there is no simple solution for this need for future values of E_z.

This chapter discusses several of the techniques that have been proposed for dealing with this problem. We have divided the techniques into the two classes shown in Fig. 13.2. The techniques listed on the right side of this figure all use a field transformation to eliminate the need for time-advanced data. The transformed field equations are then discretized and solved using FDTD techniques. Each technique uses a different approach in handling the extra terms introduced by the field transformation. Numerical stability is an important consideration in each of these approaches. All of the other techniques are listed on the left under "Direct Field Methods." These methods work directly with Maxwell's equations with periodic boundary conditions.

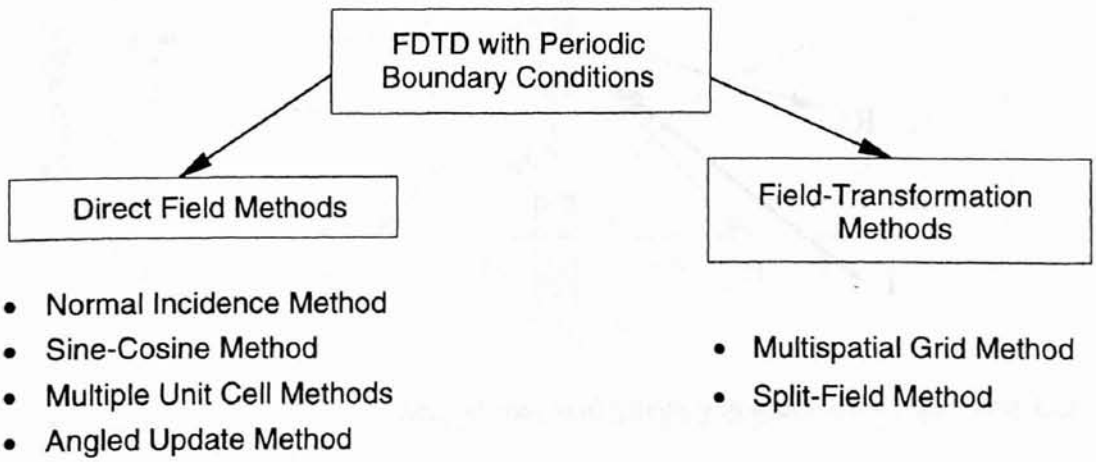


Fig. 13.2 The techniques used to perform FDTD with periodic boundary conditions are divided into two classes. Each technique is discussed in this chapter.

Another type of periodic problem is one in which the periodicity is in the direction of propagation. One example of this type is a Bragg filter. Several authors have investigated this class of problem using an FDTD approach [1–3]. In this chapter, we will focus on geometries that are periodic in the transverse directions, and will not consider longitudinal periodicity.

13.2 REVIEW OF SCATTERING FROM PERIODIC STRUCTURES

Before we delve into the details of FDTD modeling of scattering from periodic structures, a brief general discussion of this topic is appropriate. For simplicity, consider the case of a planar scatterer having one dimension of periodicity, as shown in Fig. 13.3. This geometry is periodic in the y-direction with period y_p. The illumination, denoted I in this figure, is a plane wave propagating at angle ϕ_I relative to the x-axis. The traditional far-field transmitted and reflected waves, denoted as T_0 and R_0, propagate at angles $\phi_{T_0} = \phi_I$ and $\phi_{R_0} = 180° - \phi_I$, respectively.

The presence of periodicity in the scatterer can lead to the appearance of additional angles of far-field transmission and reflection. These are often referred to as Floquet modes or grating lobes. The set of angles $\{\phi_{T_m}\}$ are the roots of the following equation:

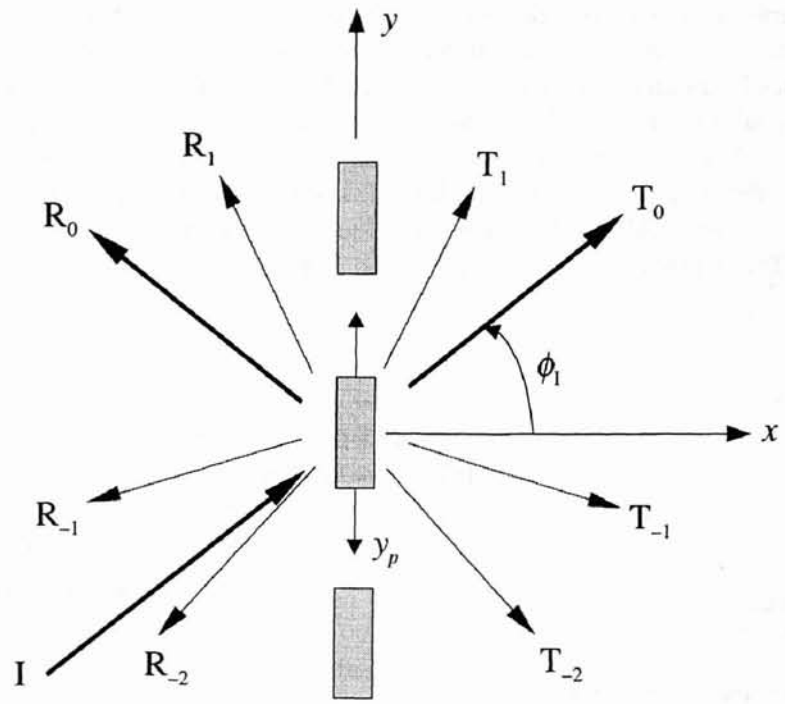


Fig. 13.3 Schematic representation of scattering from periodic media.

$$\sin \phi_{T_m} = \left(\frac{c}{y_p f}\right) m + \sin \phi_I \tag{13.3}$$

The angles of reflection ϕ_{R_m} are related to the angles of transmission by

$$\phi_{R_m} = 180^\circ - \phi_{T_m} \tag{13.4}$$

To illustrate the appearance and behavior of the additional angles of transmission and reflection, consider the case $\phi_I = 30^\circ$. Fig. 13.4 shows $\{\phi_{T_m}\}$ as a function of the normalized frequency $y_p f/c$. As the frequency increases, new solutions appear with angles beginning at $\pm 90^\circ$. Note that the number of additional angles is dependent on the frequency and the angle of incidence. For positive ϕ_I, the following are the "turn-on" frequencies for the first few orders:

$$f_{\text{turn-on}} = 0 \qquad m = 0 \tag{13.5a}$$

$$f_{\text{turn-on}} = \frac{c}{y_p(1 \mp \sin \phi_I)} \qquad m = \pm 1 \tag{13.5b}$$

$$f_{\text{turn-on}} = \frac{2c}{y_p(1 \mp \sin \phi_I)} \qquad m = \pm 2 \tag{13.5c}$$

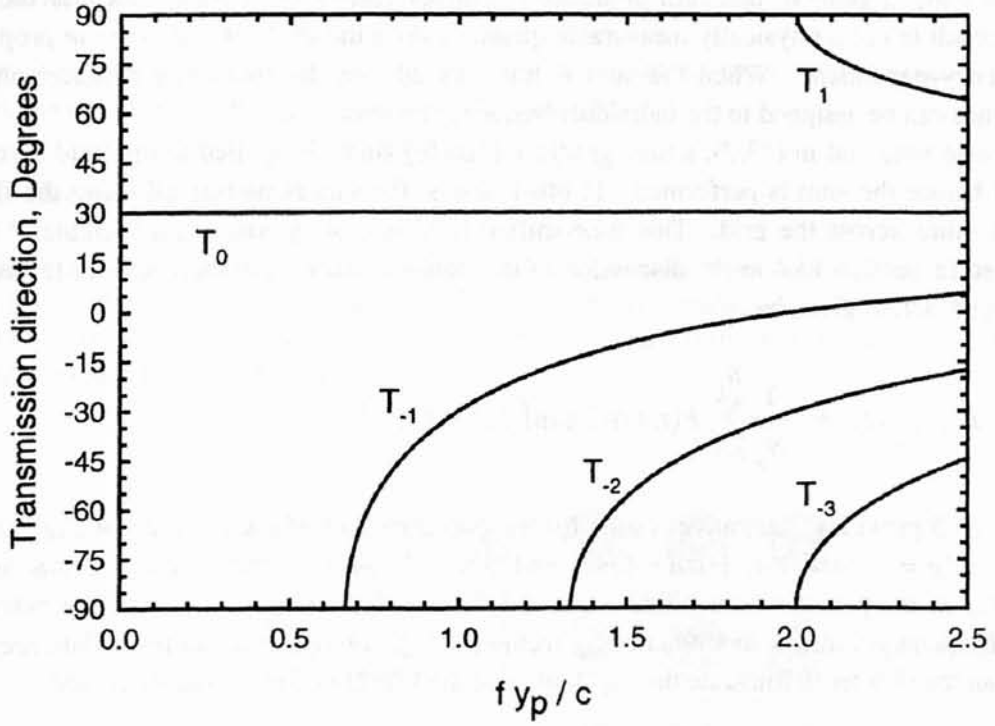


Fig. 13.4 Transmission directions ϕ_{T_m} of the periodic structure as a function of frequency for $\phi_I = 30°$.

We see that for low frequencies $f < c/[y_p(1 + \sin\phi_I)]$, only one root exists, namely $\phi_{T_0} = \phi_I$. This is in the traditional transmitted direction.

The fields on a planar surface in front of and behind the scatterer can be used to determine the amplitude and phase of the signals reflected and transmitted in each order m. Because of the periodic nature of the geometry, only the fields in the unit cell need be considered in performing this calculation. In the phasor domain, the far field for order m is given by

$$\breve{E}_{\text{far field}_m}(\omega) = \frac{1}{y_p}\int_0^{y_p} \breve{E}(\omega, y)\exp\left[j\left(2\pi m y/y_p + k_0 y\sin\phi_I\right)\right] dy \tag{13.6}$$

Thus, the far field for a given ray is found by summing up the fields across the unit cell with the appropriate phase taper. Although (13.6) is written in terms of the E-field in the phasor domain, this sum can also be performed using the time-domain fields directly. Transforming (13.6) to the time-domain and discretizing yields

$$E_{\text{far field}_m}(t) = \frac{1}{N_y}\sum_{\ell=1}^{N_y} E\left(t + \frac{\ell\Delta y\sin\phi_I}{c},\ \ell\Delta y\right)\exp\left(j2\pi m\ell/N_y\right) \tag{13.7}$$

where N_y is the number of FDTD grid cells across the unit cell of the scatterer.

Note that, in general, this sum produces a complex result. This occurs because the time-domain result is *not* a physically measurable quantity, since the angle at which ray m propagates is frequency-dependent. When the sum is transformed into the frequency domain, physical significance can be assigned to the individual frequency components.

We also note that in (13.7), a time gradient $(\ell \Delta y / c) \sin \phi_I$ is applied to the field across the unit cell before the sum is performed. In other words, the sum is performed using the field at different times across the grid. This time-shifted field is exactly equal to a variable P that is introduced in Section 13.4 in the discussion of the field-transformation methods. In terms of P, the sum of (13.7) is given by

$$E_{\text{far field}_m}(t) = \frac{1}{N_y} \sum_{\ell=1}^{N_y} P(t, \ell \Delta y) \exp\left(j 2\pi m \ell / N_y\right) \tag{13.8}$$

Fig. 13.5 provides illustrative results for the two-dimensional lossy screen of Fig. 13.1 for the case $w/g = 1$, $w/d = 4$, $1/\sigma d = 63\Omega$, and $\phi_I = 30°$. These results are for the relative amplitudes of the primary transmitted wave and the first few grating lobes, and are calculated using a frequency-domain, mode-matching technique [4]. Comparisons to these data are made throughout the chapter to illustrate the capabilities of the FDTD methods to be discussed.

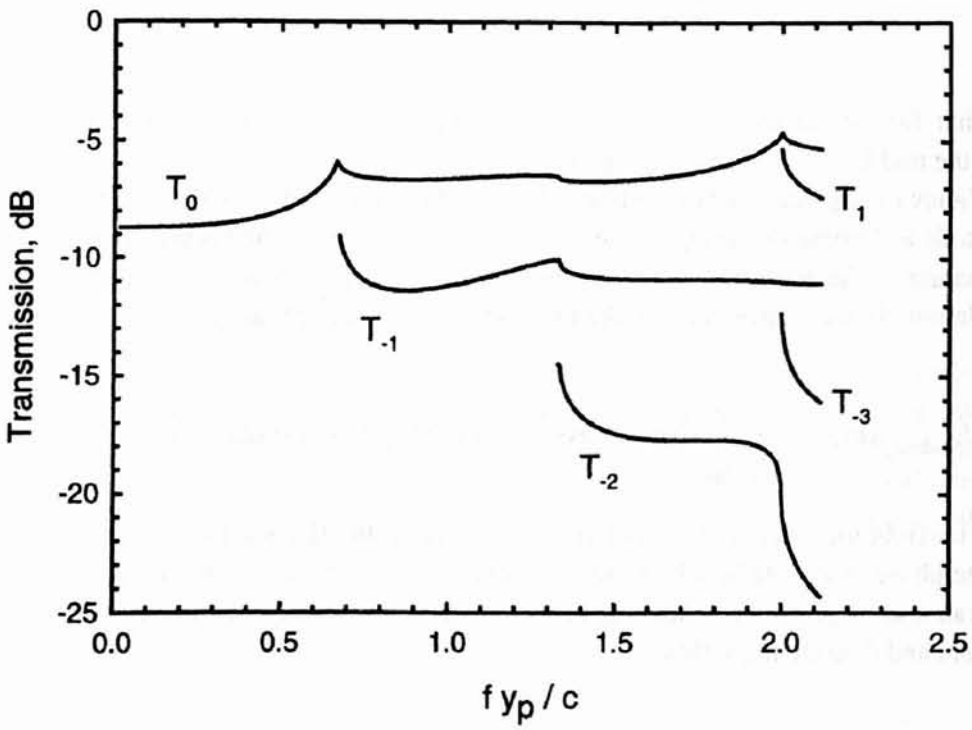


Fig. 13.5 Amplitudes of the primary transmitted wave and the first few grating lobes for the lossy screen of Fig. 13.1. Problem parameters: $w/g = 1$, $w/d = 4$, $1/\sigma d = 63\Omega$, and $\phi_I = 30°$.

13.3 DIRECT FIELD METHODS

There have been several direct field methods proposed to implement periodic boundary conditions in FDTD. This section describes these techniques.

13.3.1 Normal Incidence Case

The difficulty associated with the periodic boundary condition in the time domain does not exist when the excitation is normally incident. Consider the two-dimensional configuration shown in Fig. 13.6. Here, the wavefront has no spatial tilt, and therefore, no time delay or advance is required in the periodic boundary condition.

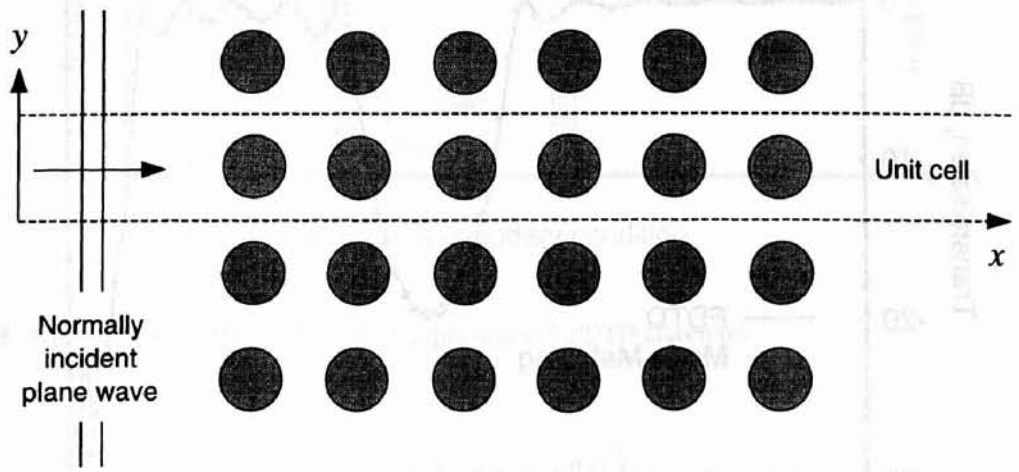


Fig. 13.6 Cross section view of a plane wave normally incident on a periodic structure.

For this case, (13.2) for $\phi_i = 0°$ becomes

$$H_x\left(x,\ y = y_p + \Delta y/2,\ t\right) = H_x\left(x,\ y = \Delta y/2,\ t\right) \tag{13.9a}$$

$$E_z\left(x,\ y = 0,\ t\right) = E_z\left(x,\ y = y_p,\ t\right) \tag{13.9b}$$

Because the normal FDTD technique is used and boundary condition (13.9) relies only on field values at the current time level, the stability of the method is unchanged; that is,

$$\frac{c\,\Delta t}{\Delta} \le \frac{1}{\sqrt{N}} \tag{13.10}$$

for N-dimensional problems. Similarly, the grid dispersion is the same as for the traditional FDTD method.

The normal-incidence periodic boundary condition of (13.9) has been successfully applied to FDTD modeling of a number of scattering and radiation problems. These include scattering from parallel strips and radiation from parallel slots [5, 6]; a multilayer, quasi-optical rotator [7]; and rough-surface scattering [8].

An emerging important application for the periodic boundary condition in FDTD modeling involves EBG materials [9, 10]. For example, consider an EBG material composed of six rows of 4-mm-diameter Pyrex rods ($\varepsilon_r = 4.2$) on a 9-mm square lattice, as shown in Fig. 13.6. For the FDTD grid, 0.25-mm spatial cells are used. Thus, the rods are 16 cells in diameter, and the unit cell is 36 cells wide.

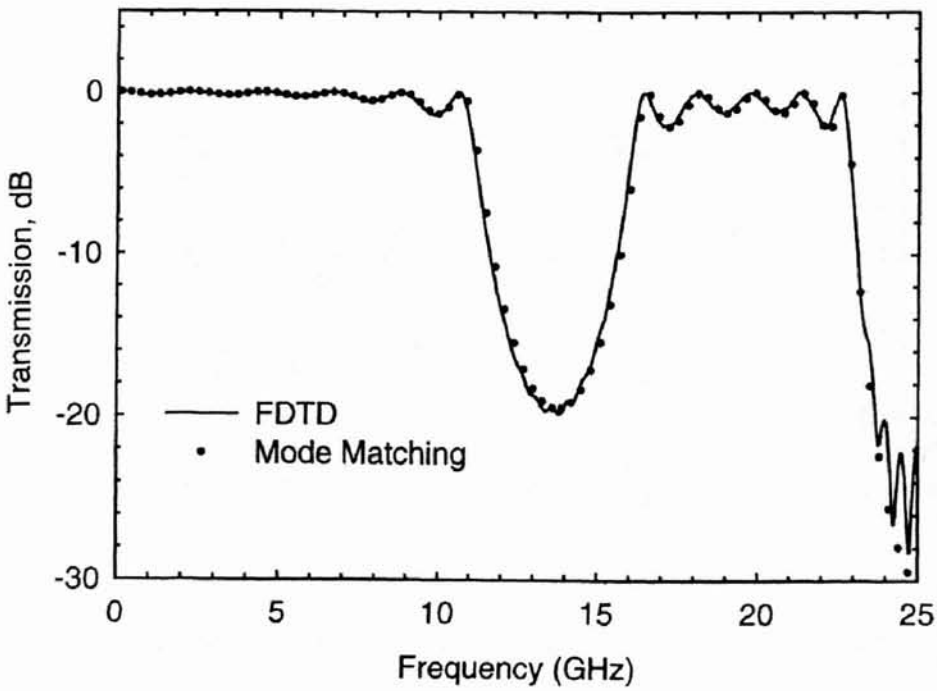


Fig. 13.7 Magnitude of transmission through the 6-row EBG crystal of Fig. 13.6 composed of 4-mm-diameter Pyrex rods (permittivity $\varepsilon_r = 4.2$) on a 9-mm square lattice.

Fig. 13.7 shows the magnitude of the transmission coefficient for the periodic array of Pyrex rods. The FDTD results for normal incidence (solid line) and the frequency-domain mode-matching results (dots) are in excellent agreement. Comparisons of FDTD and experimental data for EBG materials also indicate good agreement [9, 10].

13.3.2 Multiple Unit Cells for Oblique Incidence

The use of multiple unit cells is one means to solve the problem of needing future fields to implement the periodic boundary condition for the oblique-incidence case. Fig. 13.8 illustrates the basic strategy for a two-dimensional problem; the extension to three dimensions is straightforward. The portion of the periodic boundary condition involving time delays is used to truncate one edge of the grid; that is,

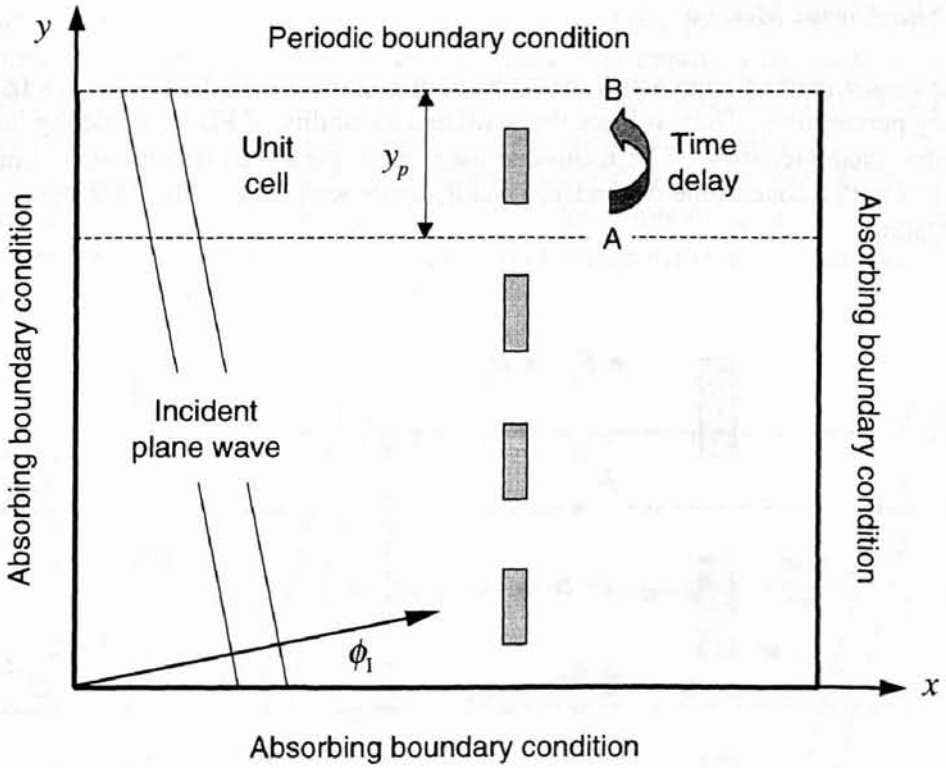


Fig. 13.8 Cross section of the grid for the multiple-unit-cell FDTD algorithm.

$$H_x\left(x,\ y=B,\ t\right) \ = \ H_x\left(x,\ y=A,\ t-y_p \sin\phi_I\right) \tag{13.11}$$

where $\phi_I \geq 0$. The other three edges of the solution space are truncated with ABCs. The plane wave is injected at a surface on the left.

A key difficulty with this method is that the ABC along the bottom of the grid interacts strongly with the incident wave that is propagating along and away from the bottom boundary. This source of error is apparently mitigated through the use of additional unit cells. At the completion of the simulation, only the fields in the unit cell at the top of Fig. 13.8 are considered the solution.

Numerical experiments using this approach indicate that scattering from a singly periodic array of strips using 3 unit cells compares well with a large-array simulation of 11 unit cells. Radiation from a doubly periodic array of rectangular apertures using 3×3 unit cells also compares well with a modal analysis for angles out to 80° from normal [11]. Input impedance calculations of a doubly periodic array of patch antennas using 5×5 unit cells compare reasonably with moment-method results [12].

This method appears to retain the usual FDTD stability limit (13.10) and dispersion characteristics. The drawbacks are the increased computational storage requirements of multiple unit cells, and the inherent finite-array approximation that is utilized to avoid the time-advanced field conundrum. We expect the level of accuracy to be comparable to assuming that the result from a unit cell in the center of a small finite array is equivalent to the infinite-array result.

13.3.3 Sine-Cosine Method

The sine-cosine method approaches the problem of modeling periodic boundaries from a single-frequency perspective. Thus, it loses the wideband capability of FDTD modeling but retains all of its other characteristics. The technique uses two separate grids computed simultaneously, one excited with a $\cos\omega t$ time-dependence and the other with $\sin\omega t$. Fig. 13.9 shows a schematic representation.

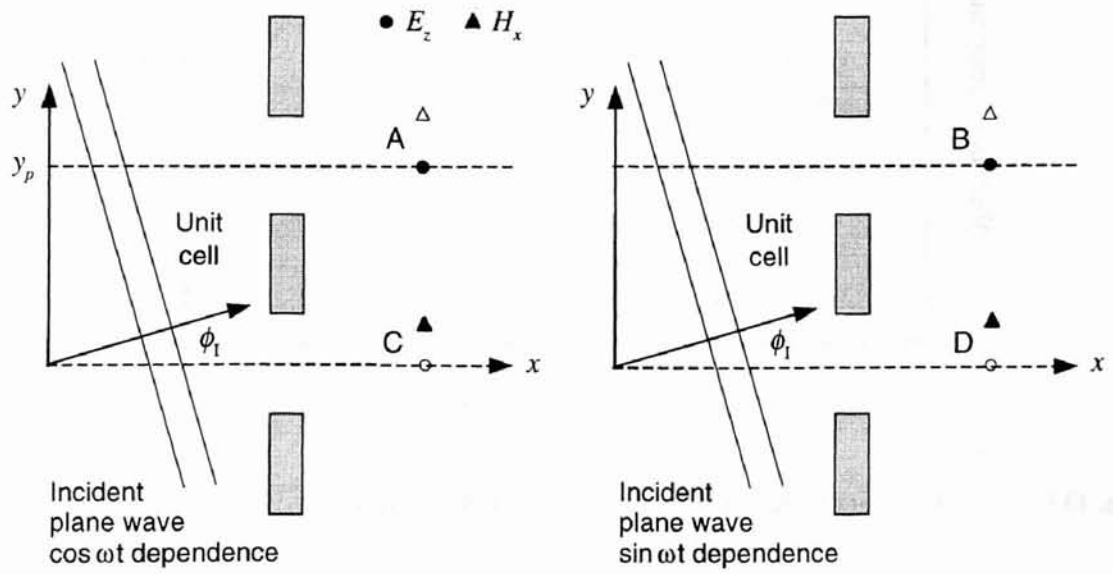


Fig. 13.9 Periodic unit cell for the sine-cosine single-frequency technique in two dimensions.

With the sine-cosine method, the instantaneous fields at the known boundary points are first combined to form a complex number, for example, $E_z(A) + jE_z(B)$. Next, the phase shifts in periodic boundary condition (13.1) are used to compute the complex field at the unknown locations. Finally, the real part is used to truncate the $\cos\omega t$ grid, and the imaginary part is used to truncate the $\sin\omega t$ grid. In Fig. 13.9, this yields

$$E_z(C) = \mathrm{Re}\left\{\left[E_z(A) + jE_z(B)\right]\exp\left(jk_y y_p\right)\right\} \tag{13.12a}$$

$$E_z(D) = \mathrm{Im}\left\{\left[E_z(A) + jE_z(B)\right]\exp\left(jk_y y_p\right)\right\} \tag{13.12b}$$

and

$$H_x(A) = \mathrm{Re}\left\{\left[H_x(C) + jH_x(D)\right]\exp\left(-jk_y y_p\right)\right\} \tag{13.13a}$$

$$H_x(B) = \mathrm{Im}\left\{\left[H_x(C) + jH_x(D)\right]\exp\left(-jk_y y_p\right)\right\} \tag{13.13b}$$

This approach has been successfully employed to model various periodic structures [13–16]. A finite-volume time-domain technique incorporated a similar approach for modeling periodic boundaries in [17]. The stability of this approach is the same as for the usual FDTD algorithm. The major drawback is that only a single frequency is calculated per simulation.

To illustrate the sine-cosine technique, the transmission through the lossy electromagnetic screen discussed earlier is calculated. Fig. 13.10 shows the results for an angle of incidence of 30°, along with the results from the frequency-domain mode-matching technique. There is excellent agreement between the two techniques.

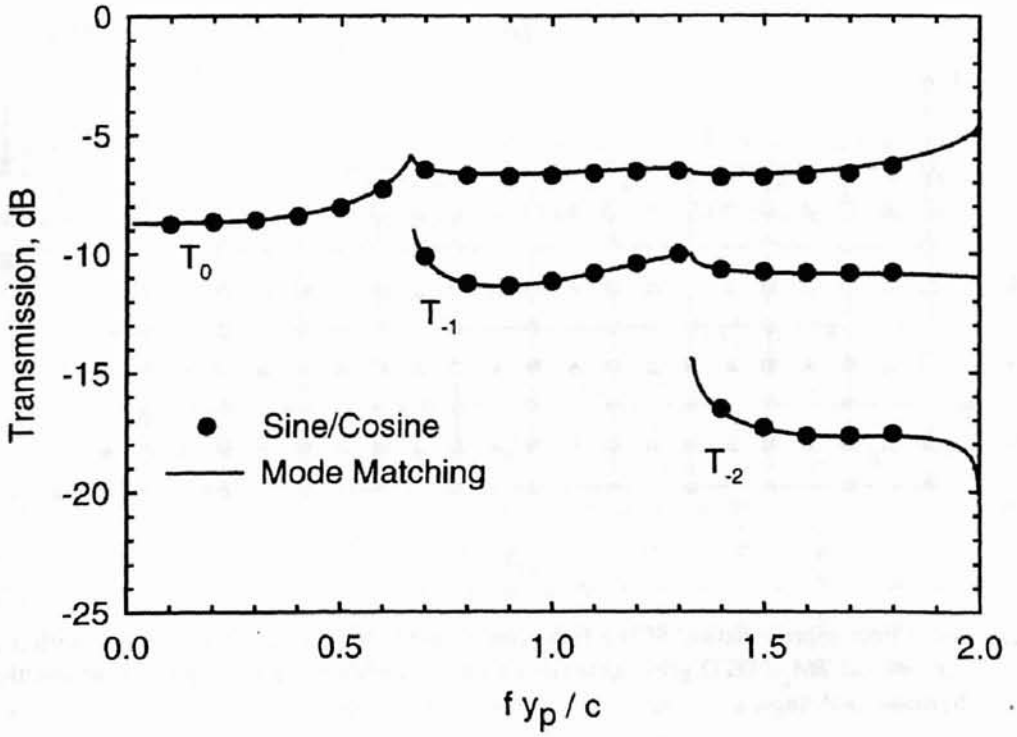


Fig. 13.10 Transmission through the lossy screen of Figs. 13.1 and 13.5 computed using the sine-cosine FDTD technique. The angle of incidence is 30°.

13.3.4 Angled-Update Method

The last direct field method to be discussed is the angled-update method. The formulation of this method is based on the observation that field components in the solution space do not all have to be at the same time level. In fact, this same observation is the basis for the efficient out-of-core FDTD solvers [18].

Fig. 13.11 shows the time evolution of the field components along a cut of constant x within a two-dimensional TM_z grid. This figure illustrates the development of a time gradient across a numerical grid, which is the basis of the angled-update method.

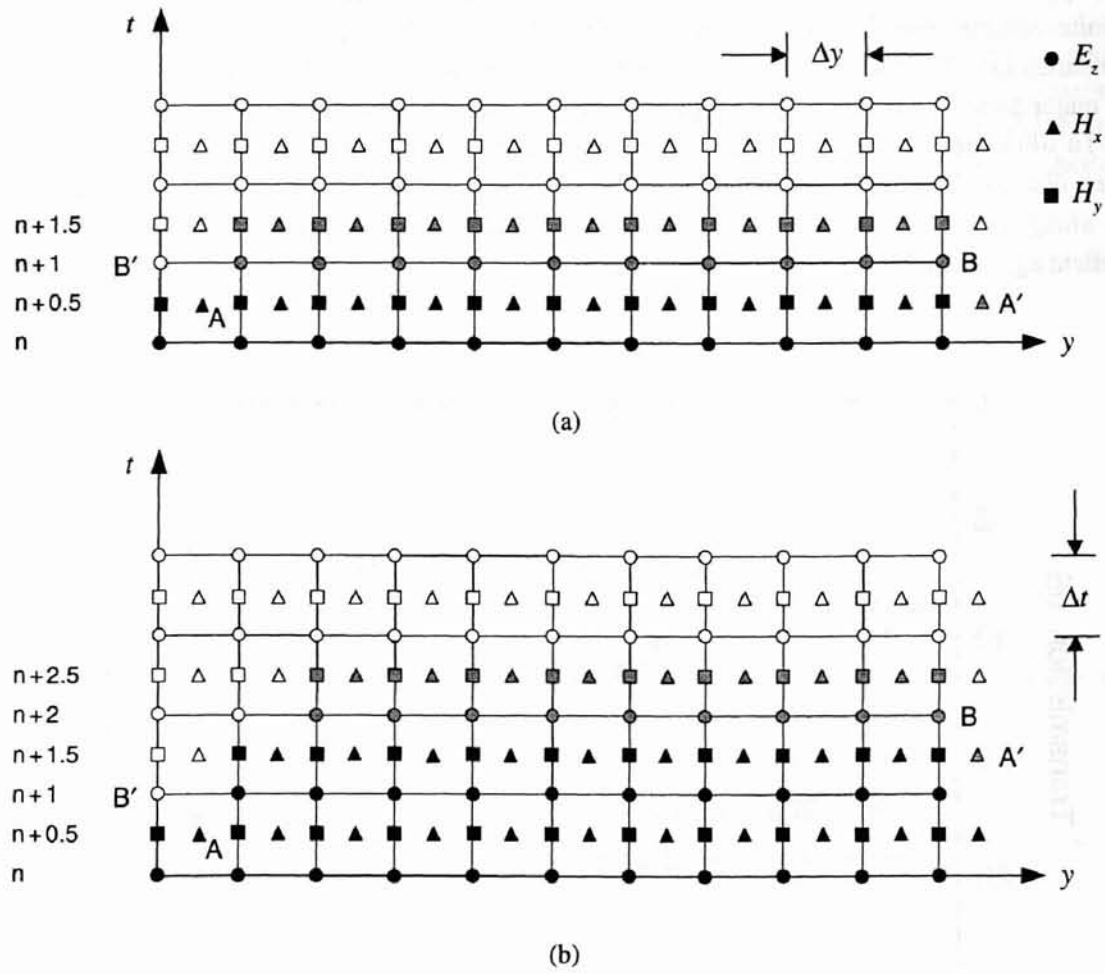


Fig. 13.11 Space-time representation of the field components along a cut of constant x within a two-dimensional TM_z FDTD grid: (a) initial stage of building the time gradient across the grid; (b) the second stage.

In Fig. 13.11(a), the black symbols denote E_z components at time n, and H_x and H_y components at time $n + 0.5$, which are assumed to be initially known. The H_x component at A′, not "updateable" using the Yee algorithm, is next assumed to be determined from saved, previous values of H_x at A by applying periodic boundary condition (13.2a). Now, the E_z components at time $n + 1$ can be calculated, with the exception of B′, which cannot be updated directly nor computed yet using periodic boundary condition (13.2b). Despite this, most of the H_x and H_y components at $n + 1.5$ can be updated using the recently updated E_z fields at $n + 1$. The components computed in this sequence are denoted by gray symbols in Fig. 13.11(a).

This sequence can be repeated as shown in Fig. 13.11(b). Again, the H_x field at A′ is determined from the saved previous values of H_x at A. This allows most of the E_z fields at time $n + 2$ to be updated. Now, most of the H_x and H_y components at $n + 2.5$ can be updated using the recently updated E_z fields at $n + 2$. As before, the components computed in this sequence are denoted by gray symbols.

This process can be repeated up to $N_y - 1$ times, where N_y is the number of cells across the grid in the y-direction. Eventually, the E_z field at B is known $N_y - 1$ time-steps in the future, relative to the time level of the unknown E_z field at B′. For moderate plane-wave steering angles, this sequence is stopped after

$$N_{\text{slant}} = \text{ceil}\left[(N_y - 1)\sqrt{2}\sin\phi_1\right] \leq N_y - 1 \tag{13.14}$$

stages, where "ceil" is a function that rounds its argument to the next higher integer. The sequence is stopped at this point so that the largest possible Δt can be used. Fig. 13.12 shows an example of this general case. Now, the time-step is chosen as

$$\frac{c\Delta t}{\Delta y} = \frac{N_y \sin\phi_1}{N_{\text{slant}}} \leq \frac{1}{\sqrt{2}} \tag{13.15}$$

so that the time differential between points A–A′ and between B–B′ is identical to the time delay and time advance required in the periodic boundary condition (13.2). We note that continuing the above sequence beyond the point indicated in (13.14) would require a smaller Δt according to (13.15).

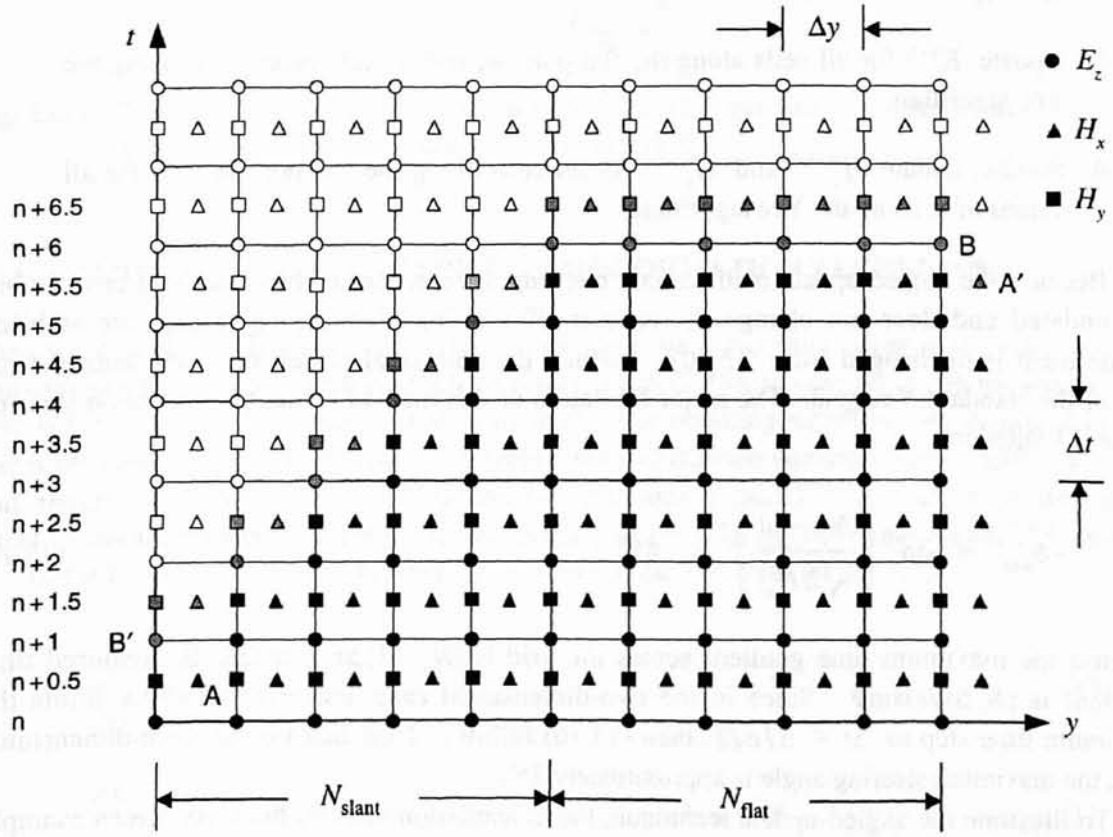


Fig. 13.12 Space-time representation of the field components of Fig. 13.11 after time-stepping has proceeded sufficiently to allow determination of $E_z(B')$ from the periodic boundary condition. Black symbols denote field components that are known, gray symbols indicate those that will be updated in the next cycle, and white symbols are unknown.

Because all fields in the FDTD grid are zero at the start of the simulation, the angled-update method can begin the simulation with the time gradient already established. Only computer storage for the most recent E- and H-field components are required. Referring to the specific example in Fig. 13.12, a single cycle of updating of the field components (shown as gray symbols) is as follows:

1. Assume N_{slant} is set by (13.14) and Δt is set by (13.15). Then, from periodic boundary condition (13.2a), calculate $H_x^{n+5.5}(\text{A}') = H_x^{n+0.5}(\text{A})$ for all values of x.

2. Update $E_z^{n+6}(\text{B})$ for all values of x using the standard Yee algorithm.

3. From periodic boundary condition (13.2b), calculate $E_z^{n+1}(\text{B}') = E_z^{n+6}(\text{B})$ for all values of x.

4. Update the fields along the slant portion of the grid:

   (a) Update $H_x^{n+1.5}$ and $H_y^{n+1.5}$ for all values of x using the Yee algorithm.

   (b) Update E_z^{n+2} for all values of x using the Yee algorithm.

   (c) Repeat (a) and (b) sequentially.

5. Update E_z^{n+6} for all cells along the flat portion, and for all values of x, using the Yee algorithm.

6. Finally, update $H_x^{n+6.5}$ and $H_y^{n+6.5}$ for all cells along the flat portion, and for all values of x, using the Yee algorithm.

Because the angled-update method only rearranges the order in which the field components are updated and does not change the regular FDTD time-stepping equations, the stability requirement is unchanged from (13.10). Further, the numerical dispersion is unchanged from that of the standard Yee grid. The major limitation of this method is that the maximum steering angle is limited to

$$\phi_{\max} = \sin^{-1}\left(\frac{N_y - 1}{\sqrt{2}\,N_y}\right) < 45^\circ \tag{13.16}$$

because the maximum time gradient across the grid is $(N_y - 1)\Delta t$, whereas the required time gradient is $(N_y \Delta y/c)\sin\phi$. Since in the two-dimensional case, numerical stability limits the maximum time-step to $\Delta t \le \Delta / c\sqrt{2}$, then (13.16) follows. Note that for the three-dimensional case, the maximum steering angle is approximately 35°.

To illustrate the angled-update technique, the transmission through the lossy screen example discussed earlier is computed. Fig. 13.13 shows the results for an angle of incidence of 30°, along with the results from the frequency-domain mode-matching technique. There is excellent agreement between the two techniques.

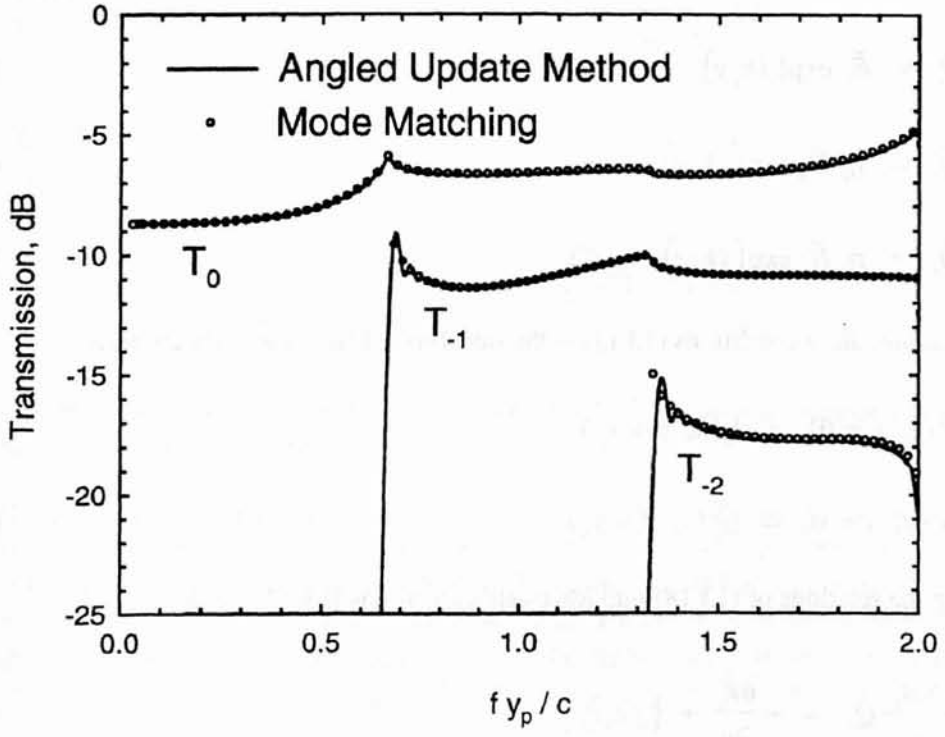


Fig. 13.13 Transmission through the lossy screen of Figs. 13.1 and 13.5 calculated using the angled-update FDTD technique. The angle of incidence is 30°.

13.4 INTRODUCTION TO THE FIELD-TRANSFORMATION TECHNIQUE

The second class of FDTD techniques for modeling periodic structures uses a field transformation to remove the time gradient across the grid [19], in contrast to the numerical time gradient introduced in the angled-update method. The resulting equations are then discretized and solved using an FDTD approach. However, the transformed equations have additional terms that require special treatment. This section provides a detailed derivation of the field-transformation approach. For simplicity and clarity, we consider the two-dimensional TM_z case.

In the frequency domain, Maxwell's equations for the TM_z case are given by

$$\frac{\partial \breve{E}_z}{\partial y} = -j\omega\mu \breve{H}_x \tag{13.17a}$$

$$\frac{\partial \breve{E}_z}{\partial x} = j\omega\mu \breve{H}_y \tag{13.17b}$$

$$\frac{\partial \breve{H}_y}{\partial x} - \frac{\partial \breve{H}_x}{\partial y} = j\omega\varepsilon \breve{E}_z \tag{13.17c}$$

We introduce a new set of field variables to remove the phase that is accumulated across the grid:

$$\breve{P}_z = \breve{E}_z \exp\left(jk_y y\right) \tag{13.18a}$$

$$\breve{Q}_x = \eta_0 \breve{H}_x \exp\left(jk_y y\right) \tag{13.18b}$$

$$\breve{Q}_y = \eta_0 \breve{H}_y \exp\left(jk_y y\right) \tag{13.18c}$$

The periodic boundary conditions (13.1) for the transformed field variables are now

$$\breve{P}_z(x,\ y=0) = \breve{P}_z(x,\ y=y_p) \tag{13.19a}$$

$$\breve{Q}_x(x,\ y=0) = \breve{Q}_x(x,\ y=y_p) \tag{13.19b}$$

Substituting the relations of (13.18) into Maxwell's equations (13.17) yields

$$\frac{j\omega\mu_r}{c}\breve{Q}_x = -\frac{\partial \breve{P}_z}{\partial y} + \left\{jk_y \breve{P}_z\right\} \tag{13.20a}$$

$$\frac{j\omega\mu_r}{c}\breve{Q}_y = \frac{\partial \breve{P}_z}{\partial x} \tag{13.20b}$$

$$\frac{j\omega\varepsilon_r}{c}\breve{P}_z = \frac{\partial \breve{Q}_y}{\partial x} - \frac{\partial \breve{Q}_x}{\partial y} + \left\{jk_y \breve{Q}_x\right\} \tag{13.20c}$$

In system (13.20), the terms in brackets are extra terms resulting from transformation (13.18). Expressing (13.20) in the time domain, we have

$$\frac{\mu_r}{c}\frac{\partial Q_x}{\partial t} = -\frac{\partial P_z}{\partial y} + \left\{\frac{\sin\phi}{c}\frac{\partial P_z}{\partial t}\right\} \tag{13.21a}$$

$$\frac{\mu_r}{c}\frac{\partial Q_y}{\partial t} = \frac{\partial P_z}{\partial x} \tag{13.21b}$$

$$\frac{\varepsilon_r}{c}\frac{\partial P_z}{\partial t} = \frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y} + \left\{\frac{\sin\phi}{c}\frac{\partial Q_x}{\partial t}\right\} \tag{13.21c}$$

where the wave is assumed to be incident on the periodic structure from free space. Combining (13.21a) and (13.21c) to eliminate the time derivatives on the right-hand side results in the basic equations to be discretized:

$$\left(\frac{\varepsilon_r \mu_r - \sin^2\phi}{c}\right)\frac{\partial Q_x}{\partial t} = -\varepsilon_r \frac{\partial P_z}{\partial y} + \left(\frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y}\right)\sin\phi \tag{13.22a}$$

$$\frac{\mu_r}{c}\frac{\partial Q_y}{\partial t} = \frac{\partial P_z}{\partial x} \tag{13.22b}$$

$$\left(\frac{\varepsilon_r \mu_r - \sin^2\phi}{c}\right)\frac{\partial P_z}{\partial t} = -\frac{\partial P_z}{\partial y}\sin\phi + \mu_r\left(\frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y}\right) \tag{13.22c}$$

It is important to note the differences between (13.22) and the normal FDTD update equations, which are found by letting $\phi = 0$.

1. The presence of a space and time derivative of the same variable in an equation means that the usual Yee leapfrog updating must be modified. For example, in (13.22a) with $\phi = 0$, Q_x at time-step $n+1/2$ is determined by the derivative of P_z at n and the previous value of Q_x at $n-1/2$. However, with $\phi \neq 0$, there are spatial derivatives of Q_x and Q_y present on the right-hand side as well. These derivatives must be handled properly before this technique can be made stable.

2. The Courant stability criterion is impacted by the extra terms and the factor multiplying the time derivatives in (13.22a) and (13.22c).

Before discussing techniques for discretizing (13.22), it is instructive to examine the properties of the continuous equations. In particular, the dispersion relation is of interest, since we will be looking at the numerical dispersion of the different techniques used to implement (13.22). To this end, we assume that the fields are of the form

$$Q_x,\ Q_y,\ P_z \propto \exp\left(j\omega t - jk_{x,s}x - jk_{y,s}y\right) \tag{13.23}$$

and substitute into (13.22). Here, the subscript s denotes the scattered field, which can be at an arbitrary direction with respect to the angle of incidence. The result is

$$\begin{bmatrix} \left(\frac{j\omega}{c}\cos^2\phi - jk_{y,s}\sin\phi\right) & jk_{x,s}\sin\phi & -jk_{y,s}\varepsilon_r \\ 0 & \frac{j\omega\mu_r}{c} & jk_{x,s} \\ -jk_{y,s}\mu_r & jk_{x,s}\mu_r & \left(\frac{j\omega}{c}\cos^2\phi - jk_{y,s}\sin\phi\right) \end{bmatrix} \begin{bmatrix} Q_x \\ Q_y \\ P_z \end{bmatrix} = 0 \tag{13.24}$$

The dispersion relation is found by setting the determinant of the matrix in (13.24) equal to zero. After some manipulation, this yields

$$\left(k_{x,s}\right)^2 + \left(k_{y,s} + \frac{\omega \sin\phi}{c}\right)^2 = \left(\frac{\omega}{c}\right)^2 \tag{13.25}$$

If we define angle α such that $k_{x,s} = (\omega/v_p)\cos\alpha$ and $k_{y,s} = (\omega/v_p)\sin\alpha$, then (13.25) yields

$$\cos^2\phi \left(\frac{v_p}{c}\right)^2 - 2\sin\alpha \sin\phi \left(\frac{v_p}{c}\right) - 1 = 0 \tag{13.26}$$

where we have introduced the phase velocity v_p as a function of α. The dispersion is characterized by finding $v_p(\alpha)$. The solution to (13.26) is simply

$$\frac{v_p}{c} = \frac{\sin\alpha \sin\phi + \sqrt{(\sin\alpha \sin\phi)^2 + \cos^2\phi}}{\cos^2\phi} \tag{13.27}$$

A comment about notation is necessary to avoid confusion at this point. ϕ is the angle at which an incident plane wave impinges on the periodic structure, while α represents a direction of propagation in the structure. Note that the solution to a particular problem involves a single ϕ, but there are many possible values of α for a given ϕ. Any scattering in the structure can create propagation at different angles.

Fig. 13.14 shows plots of the phase velocity for different angles of incidence. The phase velocities along the principal directions are given by

$$\frac{v_p}{c} = \frac{1}{\cos\phi} \qquad \alpha = 0° \tag{13.28a}$$

$$\frac{v_p}{c} = \frac{1}{1 - \sin\phi} \qquad \alpha = 90° \tag{13.28b}$$

$$\frac{v_p}{c} = \frac{1}{1 + \sin\phi} \qquad \alpha = -90° \tag{13.28c}$$

The phase velocities for the continuous case (or equivalently, the wavelength) will be used to normalize the numerical results presented later in this section.

The minimum phase velocity determines the spatial discretization required to achieve an accurate solution. This minimum velocity occurs for $\alpha = -90°$, and is given by (13.28c). This equation shows that, as the angle of incidence increases, a smaller cell size is required to maintain the same effective number of cells per wavelength. In general, the cell size needs to be reduced by as much as a factor of two as the angle of incidence approaches 90°.

A number of approaches for discretizing the system of (13.22) have been studied, and numerical stability has always been a key concern. As a result, for each such approach discussed in the following sections, a stability analysis will be a main focus. A simple estimate of the stability can be made from the maximum velocity in the grid, given by (13.28b). Using this value, we might expect the stability limit to be given by

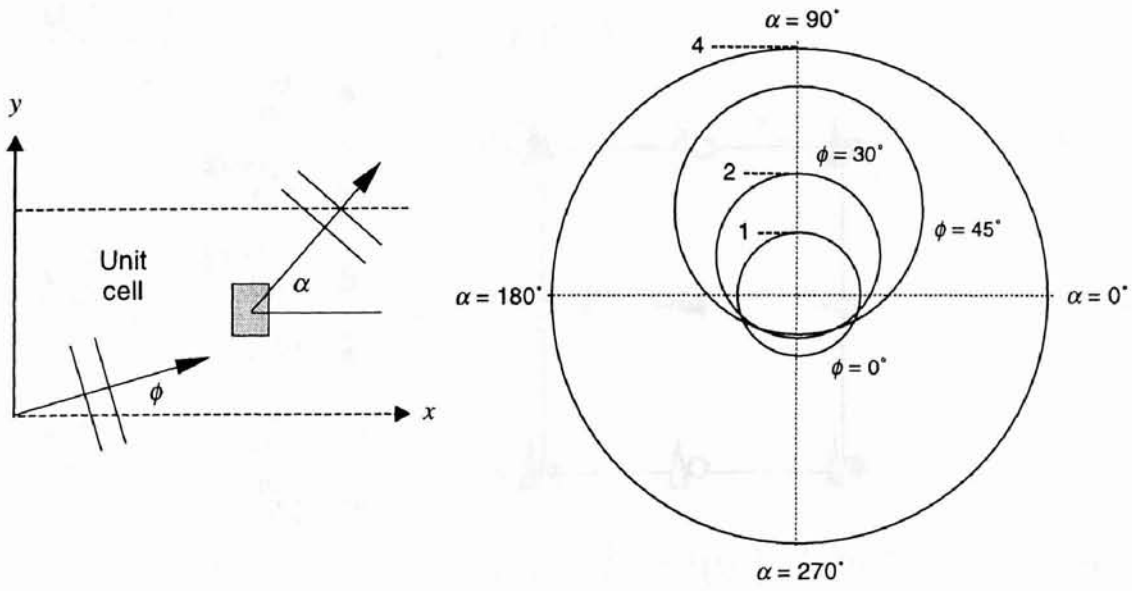


Fig. 13.14 Normalized phase velocity v_p/c as a function of the propagation angle α for three angles of incidence ϕ.

$$\frac{c\Delta t}{\Delta x} \leq \frac{1 - \sin\phi}{\sqrt{D}} \tag{13.29}$$

where D is the dimensionality of the problem. Stability analyses to be presented for various discretization schemes will show that this limit is conservative.

13.5 MULTIPLE-GRID APPROACH

This section discusses the multiple-grid approach [20] to discretize the transformed field equations of (13.22). Along with the split-field method, to be discussed in Section 13.6, this is one of two successful approaches for this purpose reported in the literature.

13.5.1 Formulation

Fig. 13.15 shows the two grids needed for the two-dimensional case. Here, the solid black symbols denote field components at time levels n, $n+1$, and so forth, while the white symbols denote field components at time levels $n-1/2$, $n+1/2$, and so forth. The additional subscript indicates in which grid the field component resides. For example, P_{z1} is the P_z component in grid 1. Overall, this figure depicts the second grid as being spatially shifted one-half cell in the direction of periodicity (the y-direction in this case), and temporally shifted one-half time-step from the first grid. (The more general three-dimensional case involves four spatial grids [20].) This arrangement permits centered differences in space and time to be applied to the system of (13.22), with the space derivatives involving quantities located, at most, one-half cell away.

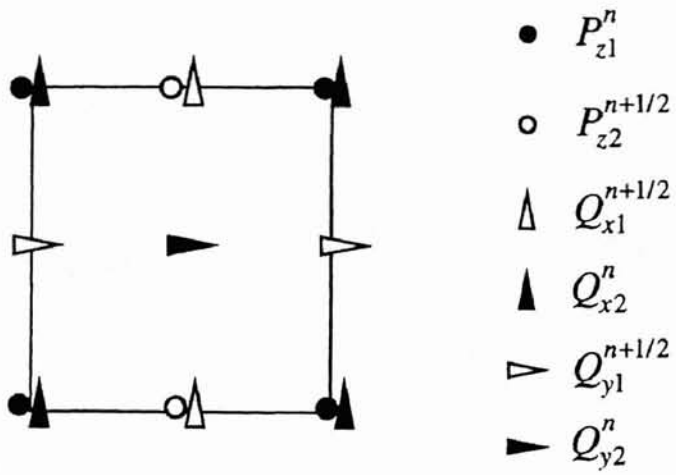


Fig. 13.15 Schematic of the two spatial grids required for the multiple-grid approach in two dimensions.

The equations to update the white-symbol field components of Fig. 13.15 to time level $n+1/2$ are

$$\frac{Q_{x1}\big|_{i,j}^{n+1/2} - Q_{x1}\big|_{i,j}^{n-1/2}}{T_y} = -\varepsilon_r A\left(P_{z1}\big|_{i,j}^{n} - P_{z1}\big|_{i,j-1}^{n}\right) + aB\left(Q_{y2}\big|_{i,j-1}^{n} - Q_{y2}\big|_{i-1,j-1}^{n}\right) - B\left(Q_{x2}\big|_{i,j}^{n} - Q_{x2}\big|_{i,j-1}^{n}\right) \tag{13.30a}$$

$$\frac{Q_{y1}\big|_{i,j}^{n+1/2} - Q_{y1}\big|_{i,j}^{n-1/2}}{T_y} = \frac{a}{\mu_r}\left(P_{z1}\big|_{i+1,j}^{n} - P_{z1}\big|_{i,j}^{n}\right) \tag{13.30b}$$

$$\frac{P_{z2}\big|_{i,j}^{n+1/2} - P_{z2}\big|_{i,j}^{n-1/2}}{T_y} = -B\left(P_{z1}\big|_{i,j}^{n} - P_{z1}\big|_{i,j-1}^{n}\right) + \mu_r aA\left(Q_{y2}\big|_{i,j-1}^{n} - Q_{y2}\big|_{i-1,j-1}^{n}\right) - \mu_r A\left(Q_{x2}\big|_{i,j}^{n} - Q_{x2}\big|_{i,j-1}^{n}\right) \tag{13.30c}$$

The equations to update the black-symbol field components to time level $n+1$ are

$$\frac{Q_{x2}\big|_{i,j}^{n+1} - Q_{x2}\big|_{i,j}^{n}}{T_y} = -\varepsilon_r A\left(P_{z2}\big|_{i,j}^{n+1/2} - P_{z2}\big|_{i,j-1}^{n+1/2}\right) + aB\left(Q_{y1}\big|_{i,j-1}^{n+1/2} - Q_{y1}\big|_{i-1,j-1}^{n+1/2}\right) - B\left(Q_{x1}\big|_{i,j}^{n+1/2} - Q_{x1}\big|_{i,j-1}^{n+1/2}\right) \tag{13.31a}$$

$$\frac{Q_{y2}\big|_{i,j}^{n+1} - Q_{y2}\big|_{i,j}^{n}}{T_y} = \frac{a}{\mu_r}\left(P_{z2}\big|_{i+1,j}^{n+1/2} - P_{z2}\big|_{i,j}^{n+1/2}\right) \tag{13.31b}$$

$$\frac{P_{z1}\big|_{i,j}^{n+1} - P_{z1}\big|_{i,j}^{n}}{T_y} = -B\left(P_{z2}\big|_{i,j}^{n+1/2} - P_{z2}\big|_{i,j-1}^{n+1/2}\right) + \mu_r aA\left(Q_{y1}\big|_{i,j-1}^{n+1/2} - Q_{y1}\big|_{i-1,j-1}^{n+1/2}\right) - \mu_r A\left(Q_{x1}\big|_{i,j}^{n+1/2} - Q_{x1}\big|_{i,j-1}^{n+1/2}\right) \tag{13.31c}$$

where $a = \Delta y/\Delta x$, $T_y = c\Delta t/\Delta y$, $A = 1/(\mu_r\varepsilon_r - \sin^2\phi)$, and $B = A\sin\phi$.

13.5.2 Numerical Stability Analysis

Assuming a generalized plane wave propagating in the grid, it can be shown that the von Neumann local stability matrix determinant expression is given by

$$\begin{vmatrix} \Lambda & 0 & 0 & -BS_y & aBS_x & -\varepsilon_r AS_y \\ 0 & \Lambda & 0 & 0 & 0 & aS_x/\mu_r \\ 0 & 0 & \Lambda & -\mu_r AS_y & \mu_r aAS_x & -BS_y \\ -BS_y & aBS_x & -\varepsilon_r AS_y & \Lambda & 0 & 0 \\ 0 & 0 & aS_x/\mu_r & 0 & \Lambda & 0 \\ -\mu_r AS_y & \mu_r aAS_x & -BS_y & 0 & 0 & \Lambda \end{vmatrix} = 0 \tag{13.32}$$

where $S_x = \sin(k_{x,s}\Delta x/2)$ and $S_y = \sin(k_{y,s}\Delta y/2)$. Solving (13.32) for nonzero values of Λ yields

$$\Lambda = \pm BS_y \pm \sqrt{aA(S_x)^2 + \mu_r\varepsilon_r A^2(S_y)^2} \tag{13.33}$$

Since S_x and S_y lie between ± 1, the stability limit for free space is found to be

$$T_y \le \frac{1}{\Lambda_{\max}} = \frac{\cos^2\phi}{\sin\phi + \sqrt{1 + a\cos^2\phi}} \tag{13.34}$$

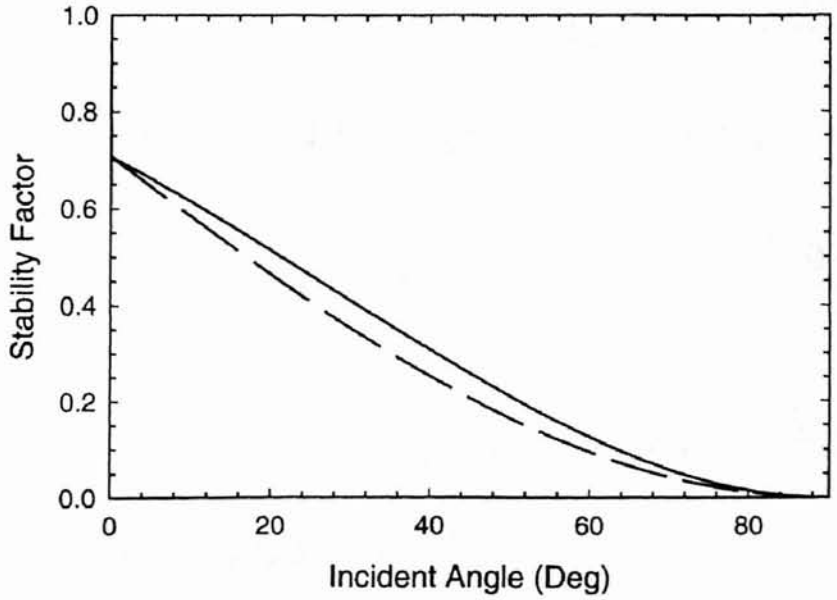


Fig. 13.16 Stability limit for the two-dimensional multiple-spatial-grid technique (solid line). Shown also is the conservative stability limit derived from the maximum velocity in the grid (dashed line).

Fig. 13.16 is a plot of the stability limit as a function of the angle of incidence. For comparison, also shown in the figure is a heuristic stability criterion derived from the fact that the fastest velocity in the grid is given by $c/(1-\sin\phi)$, yielding

$$T_y \leq (1 - \sin\phi)/\sqrt{2} \tag{13.35}$$

The stability limit is less restrictive by a factor of 1.22 at 40°, and by a factor of 1.4 at 88°.

13.5.3 Numerical Dispersion Analysis

The numerical dispersion for the multiple-grid technique is readily found from (13.33). For free space, the dispersion equation is of the form

$$\left(\Lambda\cos^2\phi - S_y\sin\phi\right)^2 = a\left(S_x\right)^2\cos^2\phi + \left(S_y\right)^2 \tag{13.36}$$

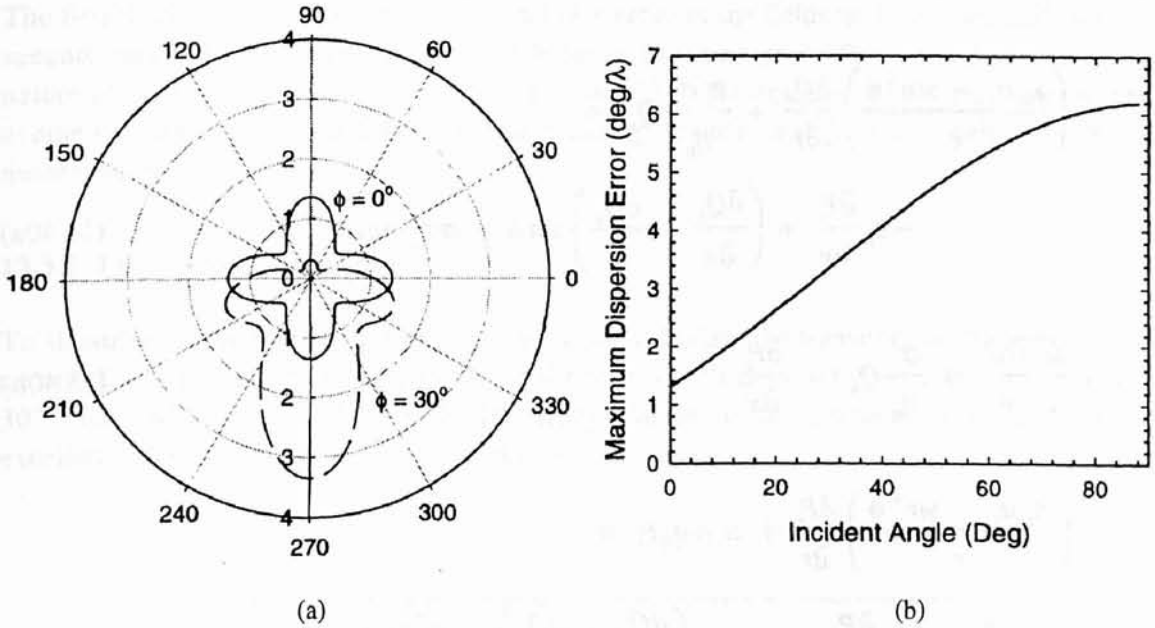
where

$$\Lambda = \frac{1}{T_y}\sin\left(\frac{\omega\Delta t}{2}\right) = \frac{1}{T_y}\sin\left(\frac{\pi\Delta y T_y}{\lambda_0}\right) \tag{13.37a}$$

$$S_x = \sin\left(\frac{\pi\Delta x\cos\alpha}{\tilde{\lambda}}\right) \quad ; \qquad S_y = \sin\left(\frac{\pi\Delta y\sin\alpha}{\tilde{\lambda}}\right) \tag{13.37b}$$

The dispersion is characterized by finding a numerical wavelength $\tilde{\lambda}(\alpha)$ that satisfies (13.36) for a given angle of incidence ϕ and discretization parameters. This numerical wavelength is then compared with the actual wavelength, and a dispersion error per wavelength is calculated using

$$\Psi_{\text{error}} = \left(\frac{\lambda(\alpha)}{\tilde{\lambda}(\alpha)} - 1 \right) 360^{\circ} \tag{13.38}$$



(a) (b)

Fig. 13.17 Dispersion error per wavelength for the multiple-grid technique: (a) dispersion error as a function of scattering direction for $\phi = 0^{\circ}, 30^{\circ}$; (b) maximum dispersion error as a function of the angle of incidence.

Fig. 13.17(a) shows results for Ψ_{error} for 0° and 30° angles of incidence. In each case, the discretization is at 20 cells/λ_0 with a time-step of one-half the stability limit. To characterize the dispersion for all angles, we look at the maximum dispersion error as a function of the angle of incidence. This result is shown in Fig. 13.17(b).

13.5.4 Lossy Materials

Introducing lossy materials into the grid makes the update equations more complex. We start with the transformed field equations, assuming the presence of both electric conductivity σ and magnetic loss σ^*:

$$\frac{\mu_r}{c}\frac{\partial Q_x}{\partial t} + \frac{\sigma^*}{\eta_0} Q_x = -\frac{\partial P_z}{\partial y} + \left\{ \frac{\sin\phi}{c}\frac{\partial P_z}{\partial t} \right\} \tag{13.39a}$$

$$\frac{\mu_r}{c}\frac{\partial Q_y}{\partial t} + \frac{\sigma^*}{\eta_0}Q_y = \frac{\partial P_z}{\partial x} \tag{13.39b}$$

$$\frac{\varepsilon_r}{c}\frac{\partial P_z}{\partial t} + \sigma\eta_0 P_z = \frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y} + \left\{\frac{\sin\phi}{c}\frac{\partial Q_x}{\partial t}\right\} \tag{13.39c}$$

The time derivatives on the right-hand sides of (13.39a) and (13.39c) are eliminated by substituting (13.39c) into (13.39a) and vice versa. This yields

$$\left(\frac{\varepsilon_r\mu_r - \sin^2\phi}{c}\right)\frac{\partial Q_x}{\partial t} + \frac{\varepsilon_r\sigma^*}{\eta_0}Q_x = -\varepsilon_r\frac{\partial P_z}{\partial y} + \left(\frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y}\right)\sin\phi - \sigma\eta_0\sin\phi\, P_z \tag{13.40a}$$

$$\frac{\mu_r}{c}\frac{\partial Q_y}{\partial t} + \frac{\sigma^*}{\eta_0}Q_y = \frac{\partial P_z}{\partial x} \tag{13.40b}$$

$$\left(\frac{\varepsilon_r\mu_r - \sin^2\phi}{c}\right)\frac{\partial P_z}{\partial t} + \mu_r\sigma\eta_0 P_z = -\frac{\partial P_z}{\partial y}\sin\phi + \mu_r\left(\frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y}\right) - \frac{\sigma^*\sin\phi}{\eta_0}Q_x \tag{13.40c}$$

In discretizing the system of (13.40), there are several options in dealing with the loss terms on the left-hand sides. Consider the space-time diagram in Fig. 13.18.

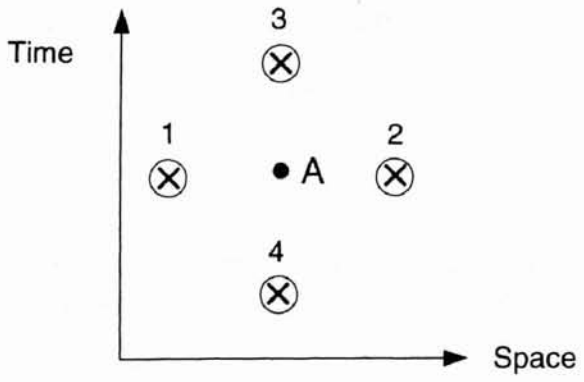


Fig. 13.18 Space-time diagram of the fields in the multiple-grid technique. The field is needed at point A, and must be estimated from the fields at points 1, 2, 3, and 4.

To maintain a centered approach, we need to have the field at point A, but the field is known only at points 1 to 4. Points 1 and 2 are at the correct time, but not at the desired space point, while points 3 and 4 are correct spatially but not temporally. Consider arriving at an estimate of the field at A using the following general average:

$$\frac{\varepsilon_r \sigma^*}{\eta_0} Q_x \rightarrow \left(\frac{1-2\beta}{2}\right)\left[Q_x(1)+Q_x(2)\right] + \beta\left[Q_x(3)+Q_x(4)\right] \tag{13.41}$$

The first term on the right-hand side is a spatial average of the fields in the second grid, while the second term is a temporal average of the fields in the same grid. Parameter β determines the nature of the averaging and impacts numerical stability. For example, $\beta = 0$ implements spatial averaging only, $\beta = 0.5$ implements temporal averaging only, and $\beta \geq 0.5$ is helpful in ensuring numerical stability.

13.5.5 Lossy Screen Example

To illustrate the multiple-grid technique, we again calculate the transmission through the lossy screen discussed earlier. Fig. 13.19 shows the results with $\beta = 0.5$ for an angle of incidence of 30°, along with the results from the frequency-domain mode-matching technique. There is excellent agreement between the two techniques.

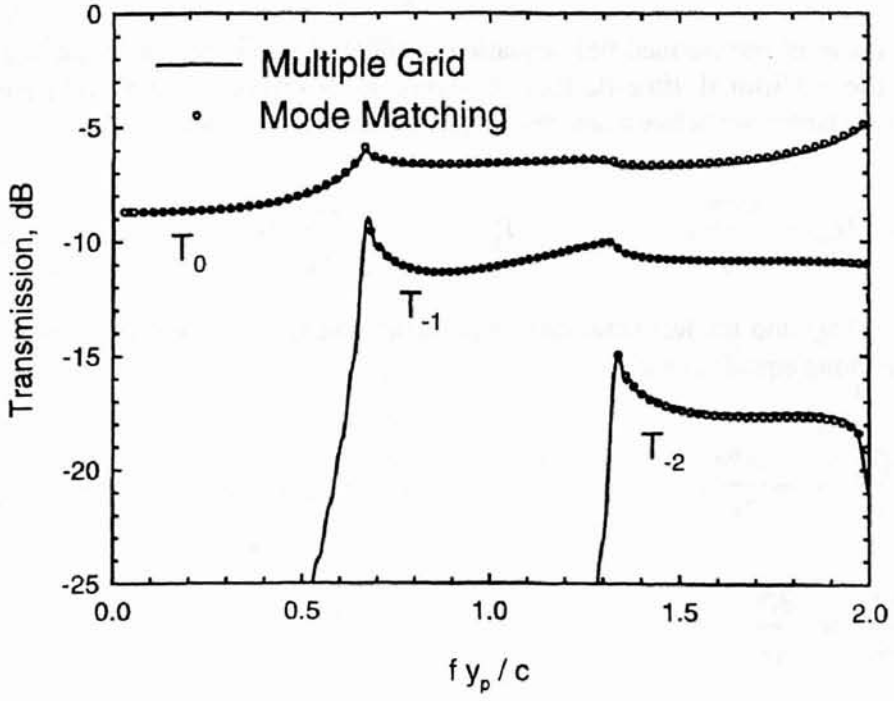


Fig. 13.19 Transmission through the lossy screen of Figs. 13.1 and 13.5 calculated using the multiple-grid FDTD technique. The angle of incidence is 30°.

The multiple-grid technique appears to be a viable means to handle the transformed field equations. Numerical stability problems are abated at the cost of introducing additional field components. Compared to the traditional three-dimensional Yee algorithm that has 6 field components per unit cell, the multiple-grid technique has 24 per cell. Using Berenger's PML ABC as in [20], there are additional variables introduced that bring the total to 40 per cell. However, because of the periodic boundary conditions, this large number of variables per cell does not cause a significant problem. (The transverse directions usually contain far fewer cells than the longitudinal direction.) Usually running-time, not memory, is the limiting factor if data at many angles is desired.

A potentially troubling issue with this technique is that the multiple spatial grids can be decoupled from each other [21]. That is, it is possible to excite only the fields associated with one of the grids, while the fields in the other grids remain zero. In addition, certain types of scattering objects affect the fields in one grid only. For example, a thin wire along the edge of a cell scatters into only one of the grids. Other scattering objects impact the separate grids differently. For objects covering many cells, the level of error caused by the grid decoupling should be small, but it is something to consider when using this technique.

13.6 SPLIT-FIELD METHOD, TWO DIMENSIONS

The second successful approach to discretize the transformed field equations is called the split-field method [22–24]. This section discusses the two-dimensional version to introduce the technique, leaving the three-dimensional case for Section 13.7.

13.6.1 Formulation

Consider the system of transformed field equations of (13.21) before any substitutions are made to eliminate the additional time-derivative terms in brackets on the right-hand sides. To eliminate these terms, we define a new set of split variables as follows:

$$Q_x = Q_{xa} + \frac{\sin\phi}{\mu_r} P_z \quad ; \qquad P_z = P_{za} + \frac{\sin\phi}{\varepsilon_r} Q_x \tag{13.42a, b}$$

Substituting (13.42a) into the left-hand side of (13.21a), and (13.42b) into the left-hand side of (13.21c), the resulting equations are

$$\frac{\mu_r}{c}\frac{\partial Q_{xa}}{\partial t} = -\frac{\partial P_z}{\partial y} \tag{13.43a}$$

$$\frac{\mu_r}{c}\frac{\partial Q_y}{\partial t} = \frac{\partial P_z}{\partial x} \tag{13.43b}$$

$$\frac{\varepsilon_r}{c}\frac{\partial P_{za}}{\partial t} = \frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y} \tag{13.43c}$$

Equations (13.43a) and (13.43c) show that the new variables P_{za} and Q_{xa} defined by (13.42) satisfy the standard Maxwell's equations, with the caveat that the curl terms on the right involve the total fields. Once P_{za} and Q_{xa} are known, the total fields can be found using (13.42). We note also that (13.42) requires that the field components be collocated in time.

Therefore, the systems of equations to be solved are first (13.43) and then (13.42). System (13.43) allows updating the "a" portions of the split fields in terms of the total fields. However, (13.42) still involves the total fields on the right-hand sides, and must be manipulated to eliminate this dependence. (The total fields are not yet known for the updates.) There are two useful strategies for solving the systems in question. The solution resulting from each approach has distinct numerical stability and dispersion characteristics. The two possibilities are:

Strategy 1: Substitute (13.42b) into (13.42a) to eliminate P_z to form a new equation to use with (13.42b). The two equations to use are thus

$$Q_x\left(1 - \frac{\sin^2\phi}{\mu_r \varepsilon_r}\right) = Q_{xa} + \frac{\sin\phi}{\mu_r} P_{za} \tag{13.44a}$$

$$\varepsilon_r P_z = \varepsilon_r P_{za} + Q_x \sin\phi \tag{13.44b}$$

Equation (13.44a) determines Q_x at time-step n from Q_{xa} and P_{za} at the same time-step. This value for Q_x can then be used to determine P_z using (13.44b). Spatial averaging is needed for the P_{za} term in (13.44a) and the Q_x term in (13.44b), since the fields are not collocated in the grid.

Strategy 2: Substitute (13.42a) into (13.42b) to form an equation for P_z in terms of Q_{xa} and P_{za}. In this case, the two equations are

$$P_z\left(1 - \frac{\sin^2\phi}{\mu_r \varepsilon_r}\right) = P_{za} + \frac{\sin\phi}{\varepsilon_r} Q_{xa} \tag{13.45a}$$

$$\mu_r Q_x = \mu_r Q_{xa} + P_z \sin\phi \tag{13.45b}$$

The comments about spatial averaging made for Strategy 1 apply here as well.

These strategies result in solutions having similar, but not identical, dispersion characteristics. However, we choose to follow the first strategy, since it results in a slightly more relaxed stability limit. This results in the following discretized system for (13.43):

$$Q_{xa}\Big|_{i,j}^{n+1/2} = Q_{xa}\Big|_{i,j}^{n-1/2} - \frac{T_y}{\mu_r}\left(P_z\Big|_{i,j}^{n} - P_z\Big|_{i,j-1}^{n}\right) \tag{13.46a}$$

$$Q_y\Big|_{i,j}^{n+1/2} = Q_y\Big|_{i,j}^{n-1/2} + \frac{a T_y}{\mu_r}\left(P_z\Big|_{i,j}^{n} - P_z\Big|_{i-1,j}^{n}\right) \tag{13.46b}$$

$$P_{za}\Big|_{i,j}^{n+1/2} = P_{za}\Big|_{i,j}^{n-1/2} + \frac{aT_y}{\varepsilon_r}\left(Q_y\Big|_{i,j}^{n} - Q_y\Big|_{i-1,j}^{n}\right) - \frac{T_y}{\varepsilon_r}\left(Q_x\Big|_{i,j}^{n} - Q_x\Big|_{i,j-1}^{n}\right) \tag{13.46c}$$

For (13.44), we have

$$Q_x\Big|_{i,j}^{n+1/2} = \mu_r \varepsilon_r A Q_{xa}\Big|_{i,j}^{n+1/2} + \frac{\varepsilon_r A \sin\phi}{2}\left(P_{za}\Big|_{i,j+1}^{n+1/2} + P_{za}\Big|_{i,j}^{n+1/2}\right) \tag{13.47a}$$

$$P_z\Big|_{i,j}^{n+1/2} = P_{za}\Big|_{i,j}^{n+1/2} + \frac{\sin\phi}{2\varepsilon_r}\left(Q_x\Big|_{i,j}^{n+1/2} + Q_x\Big|_{i,j-1}^{n+1/2}\right) \tag{13.47b}$$

where $a = \Delta y/\Delta x$, $T_y = c\Delta t/\Delta y$, and $A = 1/(\varepsilon_r\mu_r - \sin^2\phi)$. These two systems of equations are the basic update relations for the split-field technique in the two-dimensional case.

13.6.2 Numerical Stability Analysis

Assuming a generalized plane wave propagating in the grid, it can be shown that the determinant expression for von Neumann stability of the combined systems of (13.46) and (13.47) is given by

$$\begin{vmatrix} \Lambda & 0 & 0 & 0 & -S_y/\mu_r \\ 0 & \Lambda & 0 & 0 & aS_x/\mu_r \\ 0 & aS_x/\varepsilon_r & \Lambda & -S_y/\varepsilon_r & 0 \\ \mu_r\varepsilon_r A & 0 & \varepsilon_r A C_y \sin\phi & -1 & 0 \\ 0 & 0 & 1 & C_y \sin\phi/\varepsilon_r & -1 \end{vmatrix} = 0 \tag{13.48}$$

where $[Q_{xa}, Q_y, P_{za}, Q_x, P_z]^T$ is the order of the field components in the column vector. Solving (13.48) for Λ yields

$$\Lambda = AS_yC_y\sin\phi \pm \sqrt{\left(AS_y^2 + \frac{a^2}{\mu_r\varepsilon_r}S_x^2\right)\left(1 + AC_y^2\sin^2\phi\right)} \tag{13.49}$$

Stability requires that the largest value of $|\Lambda T_y|$ be no greater than 1 for any value of S_x and S_y between ± 1. It is easy to see that the maximum Λ will occur for $S_x = \pm 1$, but it is less evident which value of S_y should be used, since both C_y and S_y are functions of $k_{y,s}$. After considerable manipulation, the stability limit for square cells and free space is derived as

$$T_y \le \frac{\cos^2\phi}{\sqrt{1 + \cos^2\phi}} \tag{13.50}$$

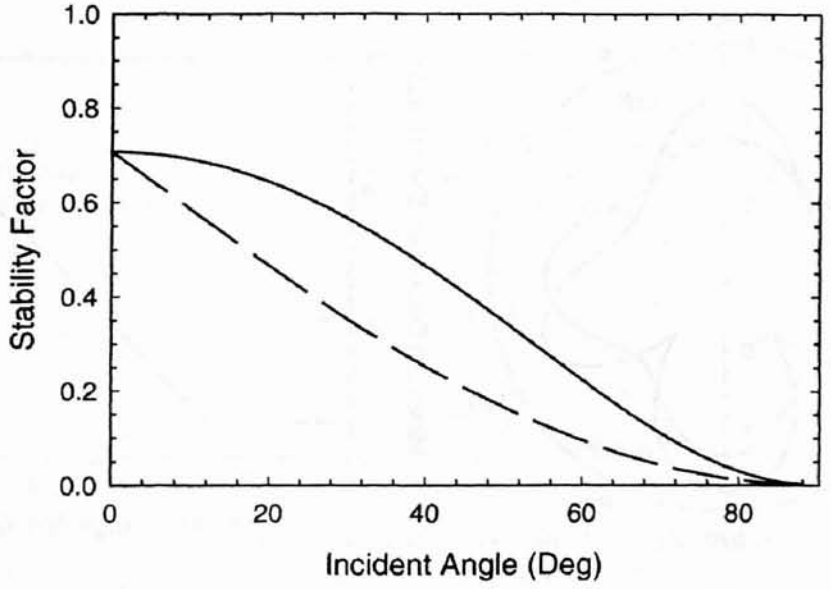


Fig. 13.20 Stability limit for the split-field method (solid line). Shown also is the conservative stability limit (dashed line) derived from (13.35), using the maximum velocity in the grid.

Fig. 13.20 is a plot of this stability limit as a function of the angle of incidence. Shown also is the heuristic stability criterion of (13.35) obtained by considering the maximum velocity in the grid. Upon comparing with Fig. 13.16, we see that the split-field method has a more relaxed stability requirement than the multiple-grid approach.

13.6.3 Numerical Dispersion Analysis

The numerical dispersion of the split-field technique is found by rewriting (13.49) as

$$\left(\Lambda \cos^2\phi - S_y C_y \sin\phi\right)^2 = \left(S_y^2 + a^2 S_x^2 \cos^2\phi\right)\left(\cos^2\phi + C_y^2 \sin^2\phi\right) \tag{13.51}$$

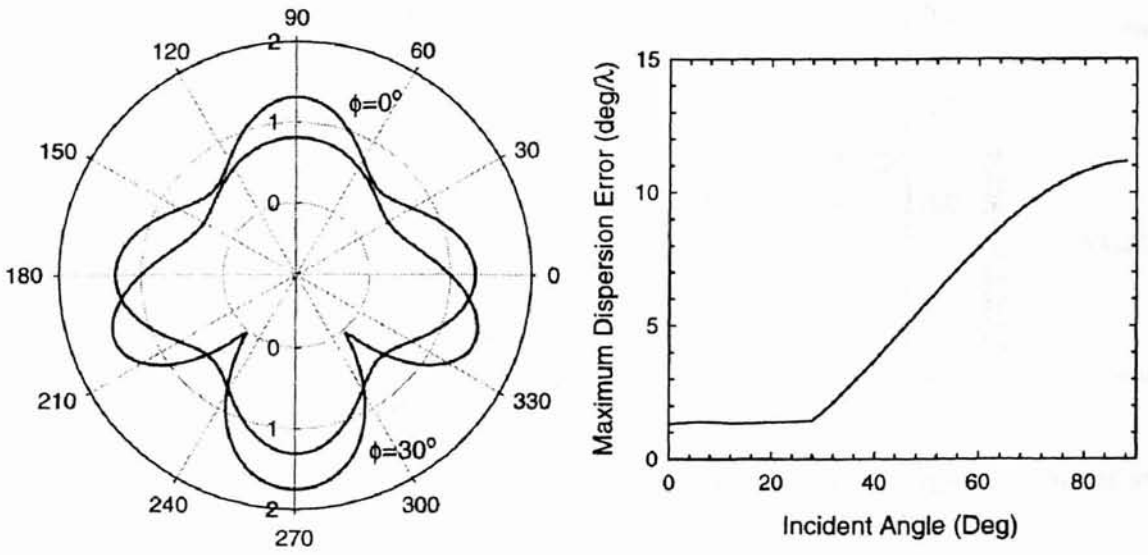
where

$$\Lambda = \frac{1}{T_y}\sin\left(\frac{\omega \Delta t}{2}\right) = \frac{1}{T_y}\sin\left(\frac{\pi \Delta y T_y}{\lambda_0}\right) \tag{13.52a}$$

$$S_x = \sin\left(\frac{\pi \Delta x \cos\alpha}{\tilde{\lambda}}\right) \quad ; \qquad S_y = \sin\left(\frac{\pi \Delta y \sin\alpha}{\tilde{\lambda}}\right) \tag{13.52b}$$

$$C_y = \cos\left(\frac{\pi \Delta y \sin\alpha}{\tilde{\lambda}}\right) \tag{13.52c}$$

Equation (13.51) is solved to determine the numerical wavelength as a function of the scattered direction α for a given angle of incidence and discretization parameters.



(a) (b)

Fig. 13.21 Dispersion error per wavelength for the split-field method: (a) error as a function of scattering direction for $\phi = 0°, 30°$; (b) maximum error as a function of the angle of incidence.

Fig. 13.21(a) shows the dispersion error Ψ_{error} for 0° and 30° angles of incidence. In each case, the discretization is 20 cells $/\lambda_0$, and Δt is set at one-half the stability limit. To characterize the dispersion for all angles, we look at the maximum dispersion error as a function of the angle of incidence. This result is shown in Fig. 13.21(b). Comparing Fig. 13.21 with Fig. 13.17, we see that the dispersion error for the split-field method exceeds that for the multiple-grid technique at incident angles larger than 40°.

13.6.4 Lossy Materials

The equations for the split-field method must be modified to allow the inclusion of lossy materials in the grid. In fact, the introduction of lossy materials requires additional stability considerations. Whereas modeling lossy materials using the Yee cell with the standard FDTD equations does not impact the stability criterion, the usual approach fails when applied to the split-field equations.

The transformed equations with both electric conductivity σ and magnetic loss σ^* are given by the system of (13.39). Substituting (13.42a) and (13.42b) into the left-hand sides of (13.39a) and (13.39c), respectively, yields

$$\frac{\mu_r}{c}\frac{\partial Q_{xa}}{\partial t} + \frac{\sigma^*}{\eta_0} Q_{xa} = -\frac{\partial P_z}{\partial y} - \frac{\sigma^* \sin\phi}{\eta_0 \mu_r} P_z \tag{13.53a}$$

$$\frac{\mu_r}{c}\frac{\partial Q_y}{\partial t} + \frac{\sigma^*}{\eta_0} Q_y = \frac{\partial P_z}{\partial x} \tag{13.53b}$$

$$\frac{\varepsilon_r}{c}\frac{\partial P_{za}}{\partial t} + \sigma\eta_0 P_{za} = \frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y} + \frac{\sigma\eta_0 \sin\phi}{\varepsilon_r} Q_x \tag{13.53c}$$

The system of (13.53), along with the system of (13.44), are the equations to be solved. The usual approach in FDTD to discretize a field update with loss [for example, (13.53a)] is to time-average the loss term. With that approach, (13.53a) becomes

$$\frac{\mu_r}{c\Delta t}\left(Q_{xa}\big|_{i,j}^{n+1} - Q_{xa}\big|_{i,j}^{n}\right) + \frac{\sigma^*}{2\eta_0}\left(Q_{xa}\big|_{i,j}^{n+1} + Q_{xa}\big|_{i,j}^{n}\right) =$$

$$-\frac{1}{\Delta y}\left(P_z\big|_{i,j+1}^{n+1/2} - P_z\big|_{i,j}^{n+1/2}\right) - \frac{\sigma^* \sin\phi}{2\eta_0 \mu_r}\left(P_z\big|_{i,j+1}^{n+1/2} + P_z\big|_{i,j}^{n+1/2}\right) \tag{13.54}$$

Unfortunately, with this formalism, the approach is unstable for angles larger than approximately 30°. A more general approach for the time average is to use [25, 26]

$$\frac{\sigma^*}{\eta_0} Q_{xa} \rightarrow \frac{\sigma^*}{\eta_0}\left[\beta Q_{xa}\big|_{i,j}^{n+1} + (1-2\beta) Q_{xa}\big|_{i,j}^{n+1/2} + \beta Q_{xa}\big|_{i,j}^{n}\right] \tag{13.55}$$

By allowing β to be a function of the angle of incidence, the technique can be made stable. Specifically, there is a minimum value of β required for stability at each angle of incidence. However, the minimum β for stability does not have the largest allowable time-step, and is therefore not optimal. The choice for β that gives the largest allowable time-step for stability is empirically determined to be

$$\beta_{\text{opt}} = \frac{0.5}{\cos^2\phi} \tag{13.56}$$

With this choice of β, the stability criterion is well approximated by (13.50), the stability limit for the lossless case.

13.6.5 Lossy Screen Example

To illustrate the split-field method, we again calculate the transmission through the lossy screen discussed earlier. Fig. 13.22 shows the results for an angle of incidence of 30°, along with the results from the frequency-domain mode-matching technique. There is excellent agreement between the two methods.

13.7 SPLIT-FIELD METHOD, THREE DIMENSIONS

The split-field method can be readily generalized to model problems involving three-dimensional periodic structures. This section discusses the technique as applied to such structures that may contain lossy, anisotropic materials.

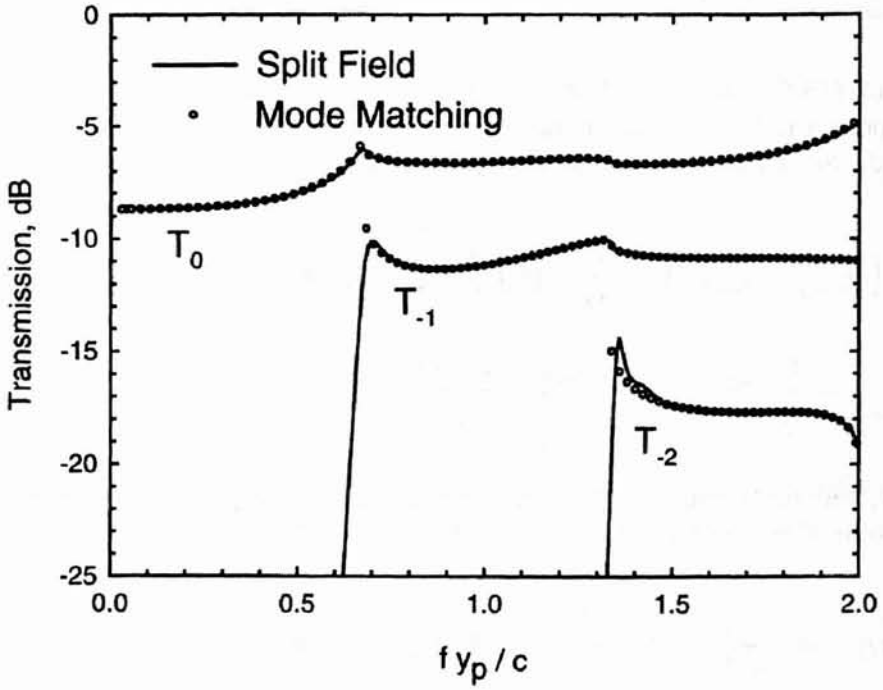


Fig. 13.22 Transmission through the lossy screen of Figs. 13.1 and 13.5 calculated using the split-field FDTD method. The angle of incidence is 30°.

13.7.1 Formulation

Fig. 13.23 shows the geometry that is considered. The periodic boundaries are the y- and z-sides, while ABCs are used to truncate the space lattice in the x-direction. We start with Maxwell's equations for the transformed field variables, given by

$$\frac{\varepsilon_{xr}}{c}\frac{\partial P_x}{\partial t} + \sigma_x \eta_0 P_x = \frac{\partial Q_z}{\partial y} - \frac{\partial Q_y}{\partial z} - \frac{\bar{k}_y}{c}\frac{\partial Q_z}{\partial t} + \frac{\bar{k}_z}{c}\frac{\partial Q_y}{\partial t} \tag{13.57a}$$

$$\frac{\varepsilon_{yr}}{c}\frac{\partial P_y}{\partial t} + \sigma_y \eta_0 P_y = -\frac{\partial Q_z}{\partial x} + \frac{\partial Q_x}{\partial z} - \frac{\bar{k}_z}{c}\frac{\partial Q_x}{\partial t} \tag{13.57b}$$

$$\frac{\varepsilon_{zr}}{c}\frac{\partial P_z}{\partial t} + \sigma_z \eta_0 P_z = \frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y} + \frac{\bar{k}_y}{c}\frac{\partial Q_x}{\partial t} \tag{13.57c}$$

$$\frac{\mu_{xr}}{c}\frac{\partial Q_x}{\partial t} + \frac{\sigma_x^*}{\eta_0} Q_x = -\frac{\partial P_z}{\partial y} + \frac{\partial P_y}{\partial z} + \frac{\bar{k}_y}{c}\frac{\partial P_z}{\partial t} - \frac{\bar{k}_z}{c}\frac{\partial P_y}{\partial t} \tag{13.58a}$$

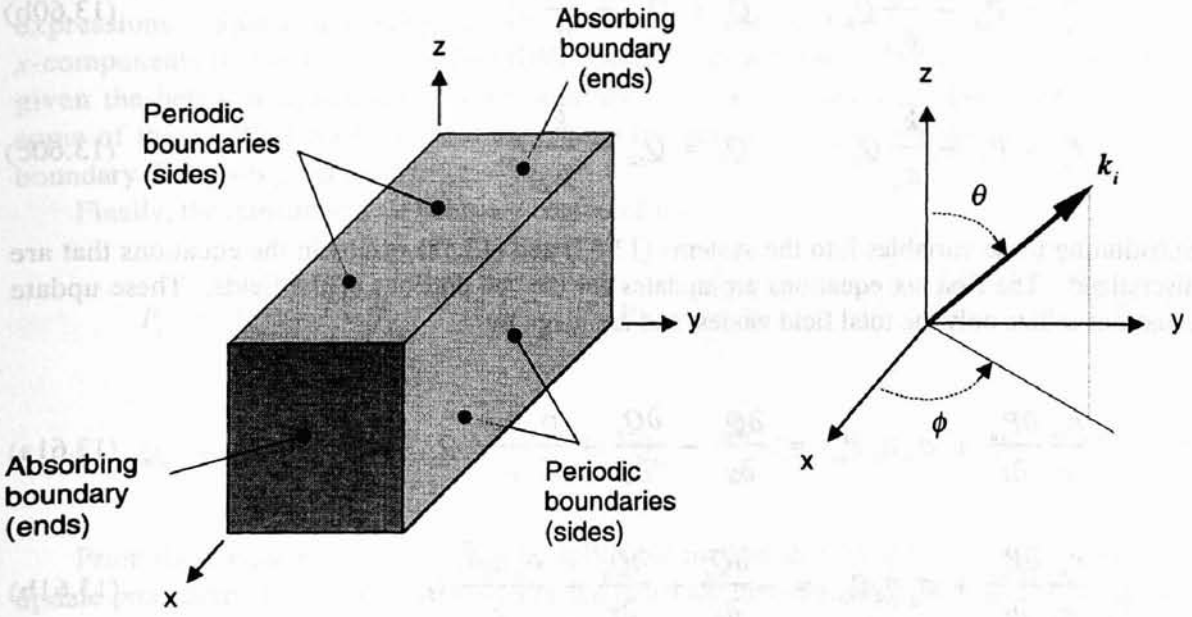


Fig. 13.23 Geometry used in the analysis of the split-field method in three dimensions.

$$\frac{\mu_{yr}}{c}\frac{\partial Q_y}{\partial t} + \frac{\sigma_y^*}{\eta_0}Q_y = \frac{\partial P_z}{\partial x} - \frac{\partial P_x}{\partial z} + \frac{\bar{k}_z}{c}\frac{\partial P_x}{\partial t} \tag{13.58b}$$

$$\frac{\mu_{zr}}{c}\frac{\partial Q_z}{\partial t} + \frac{\sigma_z^*}{\eta_0}Q_z = -\frac{\partial P_y}{\partial x} + \frac{\partial P_x}{\partial y} - \frac{\bar{k}_y}{c}\frac{\partial P_x}{\partial t} \tag{13.58c}$$

where we have introduced the notation

$$\hat{k}_i = \sin\theta\cos\phi\,\hat{x} + \sin\theta\sin\phi\,\hat{y} + \cos\theta\hat{z} \tag{13.59a}$$

$$\bar{k}_x = \hat{k}_i\cdot\hat{x} \quad ; \qquad \bar{k}_y = \hat{k}_i\cdot\hat{y} \quad ; \qquad \bar{k}_z = \hat{k}_i\cdot\hat{z} \tag{13.59b}$$

Defining new variables as in the two-dimensional case, the extra time derivatives on the right-hand sides can be eliminated. Because the angle of incidence is arbitrary, all of the field components are split. The new "a" variables are given by

$$P_x = P_{xa} + \frac{\bar{k}_z}{\varepsilon_{xr}}Q_y - \frac{\bar{k}_y}{\varepsilon_{xr}}Q_z \quad ; \qquad Q_x = Q_{xa} - \frac{\bar{k}_z}{\mu_{xr}}P_y + \frac{\bar{k}_y}{\mu_{xr}}P_z \tag{13.60a}$$

$$P_y = P_{ya} - \frac{\bar{k}_z}{\varepsilon_{yr}} Q_x \; ; \qquad Q_y = Q_{ya} + \frac{\bar{k}_z}{\mu_{yr}} P_x \tag{13.60b}$$

$$P_z = P_{za} + \frac{\bar{k}_y}{\varepsilon_{zr}} Q_x \; ; \qquad Q_z = Q_{za} - \frac{\bar{k}_y}{\mu_{zr}} P_x \tag{13.60c}$$

Substituting these variables into the systems (13.57) and (13.58) results in the equations that are discretized. The first six equations are updates for the "a" portions of the fields. These update equations utilize only the total field values, and are given by

$$\frac{\varepsilon_{xr}}{c}\frac{\partial P_{xa}}{\partial t} + \sigma_x \eta_0 P_{xa} = \frac{\partial Q_z}{\partial y} - \frac{\partial Q_y}{\partial z} + \frac{\sigma_x \eta_0 \bar{k}_y}{c} Q_z - \frac{\sigma_x \eta_0 \bar{k}_z}{c} Q_y \tag{13.61a}$$

$$\frac{\varepsilon_{yr}}{c}\frac{\partial P_{ya}}{\partial t} + \sigma_y \eta_0 P_{ya} = -\frac{\partial Q_z}{\partial x} + \frac{\partial Q_x}{\partial z} + \frac{\sigma_y \eta_0 \bar{k}_z}{c} Q_x \tag{13.61b}$$

$$\frac{\varepsilon_{zr}}{c}\frac{\partial P_{za}}{\partial t} + \sigma_z \eta_0 P_{za} = \frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y} - \frac{\sigma_z \eta_0 \bar{k}_y}{c} Q_x \tag{13.61c}$$

$$\frac{\mu_{xr}}{c}\frac{\partial Q_{xa}}{\partial t} + \frac{\sigma_x^*}{\eta_0} Q_{xa} = -\frac{\partial P_z}{\partial y} + \frac{\partial P_y}{\partial z} - \frac{\sigma_x^* \bar{k}_y}{\eta_0 c} P_z + \frac{\sigma_x^* \bar{k}_z}{\eta_0 c} P_y \tag{13.62a}$$

$$\frac{\mu_{yr}}{c}\frac{\partial Q_{ya}}{\partial t} + \frac{\sigma_y^*}{\eta_0} Q_{ya} = \frac{\partial P_z}{\partial x} - \frac{\partial P_x}{\partial z} - \frac{\sigma_y^* \bar{k}_z}{\eta_0 c} P_x \tag{13.62b}$$

$$\frac{\mu_{zr}}{c}\frac{\partial Q_{za}}{\partial t} + \frac{\sigma_z^*}{\eta_0} Q_{za} = -\frac{\partial P_y}{\partial x} + \frac{\partial P_x}{\partial y} + \frac{\sigma_z^* \bar{k}_y}{\eta_0 c} P_x \tag{13.62c}$$

Spatial averaging of the nonderivative terms on the right-hand sides is needed to retain the centered nature of the method.

Once the "a" portions are updated, the new values of the total fields P_x and Q_x are found by algebraically manipulating system (13.60) to obtain

$$\left(1 - \frac{\bar{k}_y^2}{\mu_{zr}\varepsilon_{xr}} - \frac{\bar{k}_z^2}{\mu_{yr}\varepsilon_{xr}}\right) P_x = -\frac{\bar{k}_y}{\varepsilon_{xr}} Q_{za} + \frac{\bar{k}_z}{\varepsilon_{xr}} Q_{ya} + P_{xa} \tag{13.63a}$$

$$\left(1 - \frac{\bar{k}_y^2}{\mu_{xr}\varepsilon_{zr}} - \frac{\bar{k}_z^2}{\mu_{xr}\varepsilon_{yr}}\right) Q_x = \frac{\bar{k}_y}{\mu_{xr}} P_{za} - \frac{\bar{k}_z}{\mu_{xr}} P_{ya} + Q_{xa} \tag{13.63b}$$

As before in the two-dimensional case, these relations are equalities and not standard update expressions. Spatial averaging is also needed for these equations. We have chosen the x-components of the fields to update first. Other combinations are also possible. However, given the behavior observed in the two-dimensional case, we expect stability problems with some of them. We have found that computing the field components normal to the absorbing boundary first gives good results.

Finally, the remaining total fields are computed using

$$P_y = P_{ya} - \frac{\bar{k}_z}{\varepsilon_{yr}} Q_x \quad ; \qquad P_z = P_{za} + \frac{\bar{k}_y}{\varepsilon_{zr}} Q_x \tag{13.64a, b}$$

$$Q_y = Q_{ya} + \frac{\bar{k}_z}{\mu_{yr}} P_x \quad ; \qquad Q_z = Q_{za} - \frac{\bar{k}_y}{\mu_{zr}} P_x \tag{13.65a, b}$$

From these equations, we see that the split-field method in three dimensions is a three-step update procedure. First, the "a" portions of the fields are updated using the total fields. Second, the normal components are found. Third, the remaining tangential components are determined. Each of the equations is spatially discretized with the usual Yee cell, as shown in Fig. 13.24. However, spatial and temporal averaging is needed to properly deal with some of the terms. We include the discretized version of one equation from each group to illustrate the procedure. For example, the discretized version of (13.61a) is given by

$$\begin{aligned} P_{xa}\big|_{i,j,k}^{n+1} &= \left[\frac{-\xi_x(1-2\beta)}{1+\beta\xi_x}\right] P_{xa}\big|_{i,j,k}^{n+1/2} + \left(\frac{1-\beta\xi_x}{1+\beta\xi_x}\right) P_{xa}\big|_{i,j,k}^{n} \\ &+ \left[\frac{T_y}{\varepsilon_{xr}(1+\beta\xi_x)}\right]\left(Q_z\big|_{i,j,k}^{n+1/2} - Q_z\big|_{i,j-1,k}^{n+1/2}\right) + \left[\frac{-T_z}{\varepsilon_{xr}(1+\beta\xi_x)}\right]\left(Q_y\big|_{i,j,k}^{n+1/2} - Q_y\big|_{i,j,k-1}^{n+1/2}\right) \\ &+ \left[\frac{\bar{k}_y\,\xi_x}{2\varepsilon_{xr}(1+\beta\xi_x)}\right]\left(Q_z\big|_{i,j,k}^{n+1/2} + Q_z\big|_{i,j-1,k}^{n+1/2}\right) + \left[\frac{-\bar{k}_z\,\xi_x}{2\varepsilon_{xr}(1+\beta\xi_x)}\right]\left(Q_y\big|_{i,j,k}^{n+1/2} + Q_y\big|_{i,j,k-1}^{n+1/2}\right) \end{aligned} \tag{13.66}$$

where $\xi_x = \sigma_x \Delta t/\varepsilon_{xr}\varepsilon_0$, $T_y = c\Delta t/\Delta y$, and $T_z = c\Delta t/\Delta z$. The modified time-averaging of the loss term is evident by the presence of the parameter β. In the three-dimensional case, an optimal choice for β is found to be

$$\beta_{\text{opt}} = \frac{0.5}{(\bar{k}_x)^2} = \frac{0.5}{\sin^2\theta\cos^2\phi} \tag{13.67}$$

Using this value for β results in a maximum time-step that is the same as for the lossless case. The lossless stability case is discussed later in this section.

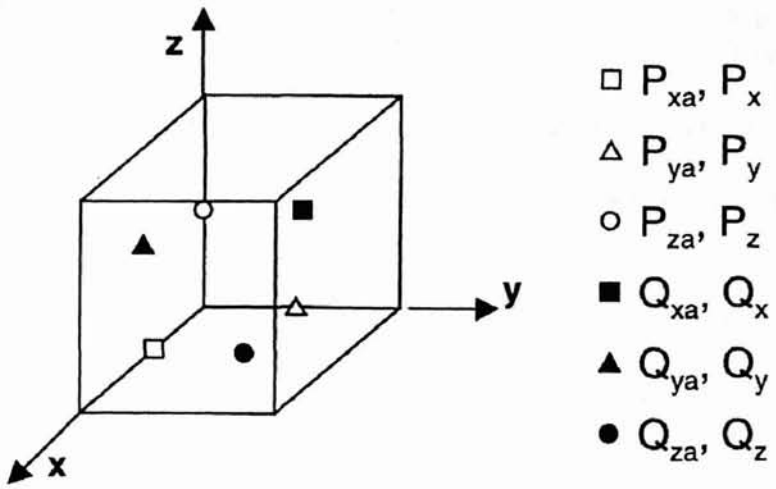


Fig. 13.24 Three-dimensional Yee cell used in the discretization procedure.

The discretized version of (13.63a) is straightforward and is given by

$$P_x\Big|_{i,j,k}^{n+1} = \left(1 - \frac{\bar{k}_y^2}{\mu_{zr}\,\varepsilon_{xr}} - \frac{\bar{k}_z^2}{\mu_{yr}\,\varepsilon_{xr}}\right)^{-1} P_{xa}\Big|_{i,j,k}^{n+1}$$

$$- \left(1 - \frac{\bar{k}_y^2}{\mu_{zr}\,\varepsilon_{xr}} - \frac{\bar{k}_z^2}{\mu_{yr}\,\varepsilon_{xr}}\right)^{-1} \left(\frac{\bar{k}_y}{2\,\varepsilon_{xr}}\right) \left(Q_{za}\Big|_{i,j,k}^{n+1} + Q_{za}\Big|_{i,j-1,k}^{n+1}\right)$$

$$+ \left(1 - \frac{\bar{k}_y^2}{\mu_{zr}\,\varepsilon_{xr}} - \frac{\bar{k}_z^2}{\mu_{yr}\,\varepsilon_{xr}}\right)^{-1} \left(\frac{\bar{k}_z}{2\,\varepsilon_{xr}}\right) \left(Q_{ya}\Big|_{i,j,k}^{n+1} + Q_{ya}\Big|_{i,j,k-1}^{n+1}\right) \tag{13.68}$$

Here, spatial averaging is used to preserve the centered nature of the algorithm. Finally, the discretized version of (13.64a) is given by

$$P_y\Big|_{i,j,k}^{n+1} = P_{ya}\Big|_{i,j,k}^{n+1} - \frac{\bar{k}_z}{2\,\varepsilon_{yr}} \left(Q_x\Big|_{i,j,k}^{n+1} + Q_x\Big|_{i,j,k-1}^{n+1}\right) \tag{13.69}$$

These equations give a full prescription of the split-field technique in three dimensions in the main portion of the lattice. The only remaining issues are the stability relation and the truncation of the lattice with an absorbing boundary condition.

13.7.2 Numerical Stability Analysis

The stability relation for the three-dimensional split-field technique is derived by implementing the von Neumann analysis, as illustrated for the two-dimensional cases. This results in a 12×12 stability matrix for the general case with loss. Setting the determinant of this matrix to zero results in a twelfth-order polynomial that must be solved to determine the stability limit. Upon performing this evaluation, we determine that the use of the optimum value of β from (13.67) results in a stability limit that is essentially the same as for the lossless case.

The roots of the characteristic equation for the lossless case are determined analytically with the largest one given by

$$\Lambda = \frac{1}{\bar{k}_x^2}\left\{\bar{k}_y S_y C_y + \bar{k}_z S_z C_z + \sqrt{\begin{array}{c}\left(\bar{k}_y S_y C_y + \bar{k}_z S_z C_z\right)^2 + \\ \bar{k}_x^2\left[S_y^2 + S_z^2 + S_x^2\left(\bar{k}_x^2 + \bar{k}_y^2 C_y^2 + \bar{k}_z^2 C_z^2\right)\right]\end{array}}\right\} \tag{13.70}$$

The stability limit is determined by maximizing (13.70) over all possible values of S_x, S_y, and S_z (that are limited to lie between ± 1). However, since C_y and C_z are not independent of S_y and S_z, respectively, the values of S_y and S_z that maximize (13.70) are not obvious.

Analytical solutions to (13.70) for the cases $k_y = 0$ and $k_z = 0$ have been derived for the general case of nonsquare cells [24], but the results are quite complicated. The results are somewhat simpler for square cells. Subject to this restriction, we now present analytical solutions for the special cases of propagation along a coordinate axis ($k_y = 0$ or $k_z = 0$) and along the direction $k_y = k_z$. These two cases bound the stability limit for a general direction. The simplest and most stringent result occurs for $k_y = k_z$. Here, we have

$$T_y \leq \frac{\bar{k}_x^2}{\sqrt{2 + \bar{k}_x^2}} \tag{13.71}$$

Along a coordinate axis ($k_z = 0$ in this case), the stability limit is given by

$$T_y \leq \frac{\bar{k}_x^2}{\bar{k}_y F_2\sqrt{1 - F_2^2} + \sqrt{\left(\bar{k}_x^2 + \dfrac{\bar{k}_y^4}{F_1}\right)\left(2 + F_2^2 \bar{k}_x^2\right)}} \tag{13.72}$$

where

$$F_1 = 8\bar{k}_y^2 + \bar{k}_x^6 + 2\bar{k}_x^2 \bar{k}_y^2 \tag{13.73a}$$

$$F_2^2 = \frac{\left(F_1 - \bar{k}_x^2 \bar{k}_y^2\right) + \sqrt{\left(F_1 - \bar{k}_x^2 \bar{k}_y^2\right)^2 - 8F_1 \bar{k}_y^2}}{2F_1} \tag{13.73b}$$

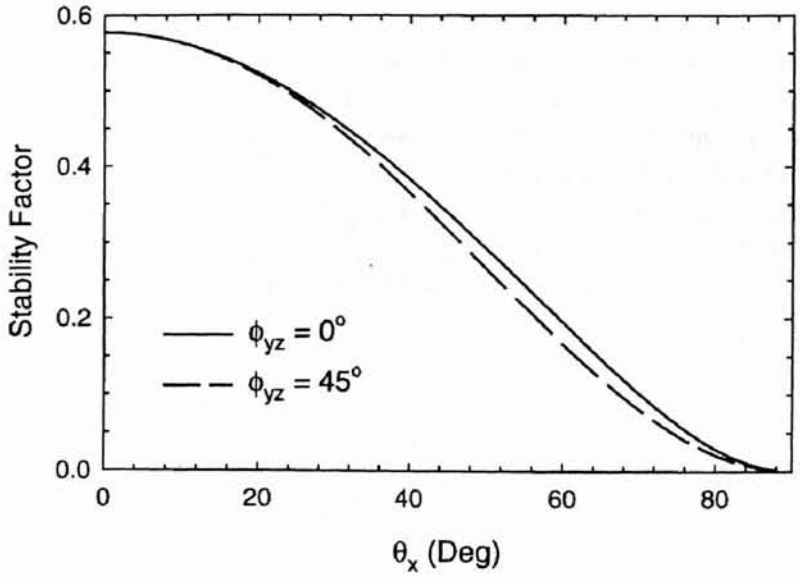


Fig. 13.25 Stability limit for the three-dimensional split-field technique for two special cases.

Fig. 13.25 shows plots of the stability limits for these two cases. To present the data, these plots use polar angles with respect to the x-axis:

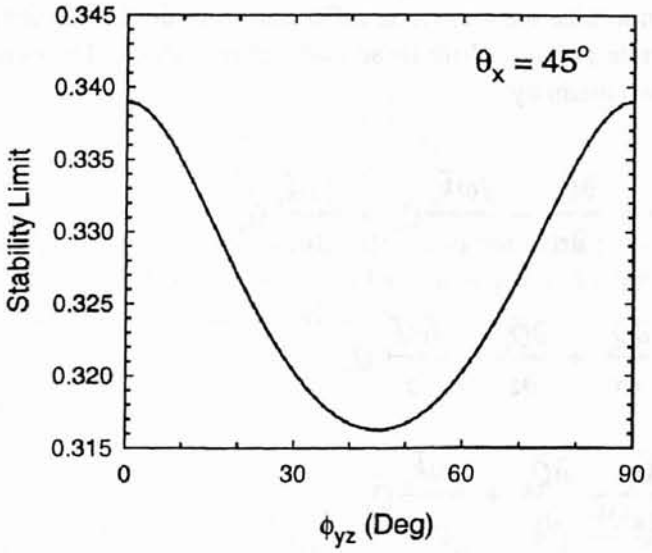
$$\bar{k}_x = \cos\theta_x \; ; \qquad \bar{k}_y = \sin\theta_x \cos\phi_{yz} \; ; \qquad \bar{k}_z = \sin\theta_x \sin\phi_{yz} \tag{13.74}$$

With this notation, stability limit (13.71) is for $\phi_{yz} = 45°$, while stability limit (13.72) is for $\phi_{yz} = 0°$. The curves are similar, and it appears that using just the lower curve would be suitable. However, for large steering angles where the stability limit requires small time-steps, stability limit (13.71) is considerably more stringent (for example, about 25% lower at 76°). To reduce the long run-times at these angles, it is desirable to use the largest possible time-step. Thus, it is beneficial to know the behavior of the stability limit as a function of ϕ_{yz} for a fixed θ_x.

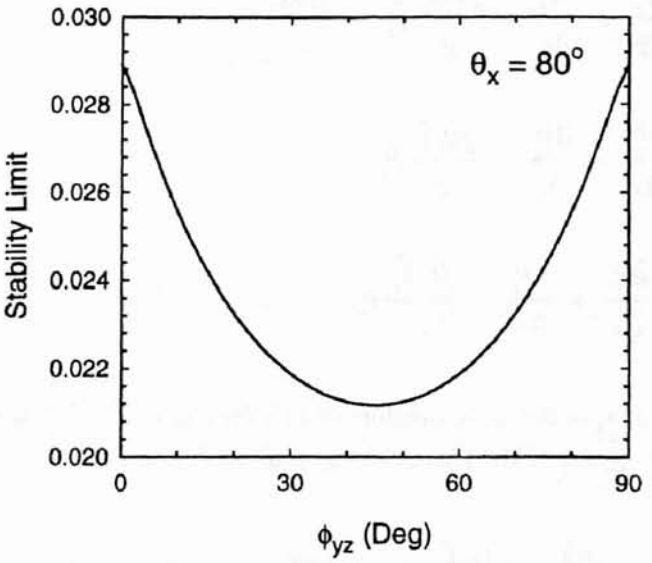
The behavior of the stability limit is investigated by numerically maximizing (13.70) as a function of ϕ_{yz} for different values of θ_x. Fig. 13.26 shows results for $\theta_x = 45°$ and 80°. This figure shows that the variation of the stability limit with ϕ_{yz} depends on θ_x. One approach to determining an approximate stability limit for a general angle of incidence is to fit a quadratic function of ϕ_{yz} through the minimum and maximum points given by (13.71) and (13.72), respectively. We find that using 98% of this approximate limit is adequate to ensure that the time-step is below the true stability limit.

13.7.3 UPML Absorbing Boundary Condition

An important part of any FDTD simulation is the ABC used to truncate the space lattice. Since 1994, the PML concept has proven to be very effective for this purpose (see Chapter 7). This section discusses how the *uniaxial PML* (UPML) ABC can be applied to the split-field FDTD modeling of periodic structures.



(a)



(b)

Fig. 13.26 Stability limit of the three-dimensional split-field method for two sets of incidence angles: (a) $\theta_x = 45°$; (b) $\theta_x = 80°$.

Since we have absorbing boundaries only at the ends of the space lattice in a periodic problem, the UPML is a material having the properties

$$\frac{\varepsilon}{\varepsilon_0} = \frac{\mu}{\mu_0} = \begin{bmatrix} 1/s_x & 0 & 0 \\ 0 & s_x & 0 \\ 0 & 0 & s_x \end{bmatrix} \tag{13.75}$$

where $s_x = \kappa + \chi c/j\omega$. The attenuation in the UPML is determined by χ, and the usual tapering of χ should be used to minimize the numerical reflection from the free-space / UPML interface and the steps in the discrete profile. With these material properties, the transformed equations are given in the frequency domain by

$$\frac{j\omega}{s_x c}\breve{P}_x = \frac{\partial \breve{Q}_z}{\partial y} - \frac{\partial \breve{Q}_y}{\partial z} - \frac{j\omega \bar{k}_y}{c}\breve{Q}_z + \frac{j\omega \bar{k}_z}{c}\breve{Q}_y \tag{13.76a}$$

$$\frac{j\omega s_x}{c}\breve{P}_y = -\frac{\partial \breve{Q}_z}{\partial x} + \frac{\partial \breve{Q}_x}{\partial z} - \frac{j\omega \bar{k}_z}{c}\breve{Q}_x \tag{13.76b}$$

$$\frac{j\omega s_x}{c}\breve{P}_z = \frac{\partial \breve{Q}_y}{\partial x} - \frac{\partial \breve{Q}_x}{\partial y} + \frac{j\omega \bar{k}_y}{c}\breve{Q}_x \tag{13.76c}$$

$$\frac{j\omega}{s_x c}\breve{Q}_x = -\frac{\partial \breve{P}_z}{\partial y} + \frac{\partial \breve{P}_y}{\partial z} + \frac{j\omega \bar{k}_y}{c}\breve{P}_z - \frac{j\omega \bar{k}_z}{c}\breve{P}_y \tag{13.77a}$$

$$\frac{j\omega s_x}{c}\breve{Q}_y = \frac{\partial \breve{P}_z}{\partial x} - \frac{\partial \breve{P}_x}{\partial z} + \frac{j\omega \bar{k}_z}{c}\breve{P}_x \tag{13.77b}$$

$$\frac{j\omega s_x}{c}\breve{Q}_z = -\frac{\partial \breve{P}_y}{\partial x} + \frac{\partial \breve{P}_x}{\partial y} - \frac{j\omega \bar{k}_y}{c}\breve{P}_x \tag{13.77c}$$

Because of the presence of s_x in the denominators of (13.76a) and (13.77a), we change variables to $\hat{P}_x = \breve{P}_x / s_x$ and $\hat{Q}_x = \breve{Q}_x / s_x$ [27]. Then, (13.76) and (13.77) become

$$\frac{j\omega}{c}\hat{P}_x = \frac{\partial \breve{Q}_z}{\partial y} - \frac{\partial \breve{Q}_y}{\partial z} - \frac{j\omega \bar{k}_y}{c}\breve{Q}_z + \frac{j\omega \bar{k}_z}{c}\breve{Q}_y \tag{13.78a}$$

$$\frac{j\omega s_x}{c}\breve{P}_y = -\frac{\partial \breve{Q}_z}{\partial x} + s_x\frac{\partial \hat{Q}_x}{\partial z} - \frac{j\omega s_x \bar{k}_z}{c}\hat{Q}_x \tag{13.78b}$$

$$\frac{j\omega s_x}{c}\breve{P}_z = \frac{\partial \breve{Q}_y}{\partial x} - s_x\frac{\partial \hat{Q}_x}{\partial y} + \frac{j\omega s_x \bar{k}_y}{c}\hat{Q}_x \tag{13.78c}$$

$$\frac{j\omega}{c}\hat{Q}_x = -\frac{\partial \breve{P}_z}{\partial y} + \frac{\partial \breve{P}_y}{\partial z} + \frac{j\omega \bar{k}_y}{c}\breve{P}_z - \frac{j\omega \bar{k}_z}{c}\breve{P}_y \tag{13.79a}$$

$$\frac{j\omega s_x}{c}\breve{Q}_y = \frac{\partial \breve{P}_z}{\partial x} - s_x \frac{\partial \hat{P}_x}{\partial z} + \frac{j\omega s_x \bar{k}_z}{c}\hat{P}_x \tag{13.79b}$$

$$\frac{j\omega s_x}{c}\breve{Q}_z = -\frac{\partial \breve{P}_y}{\partial x} + s_x \frac{\partial \hat{P}_x}{\partial y} - \frac{j\omega s_x \bar{k}_y}{c}\hat{P}_x \tag{13.79c}$$

Defining new variables as in the non-UPML case, the extra time derivatives on the right-hand side are eliminated. These new variables are slightly different than before, with the tangential components now split into "a" and "b" parts:

$$\hat{P}_x = \hat{P}_{xa} + \bar{k}_z \breve{Q}_y - \bar{k}_y \breve{Q}_z \tag{13.80a}$$

$$\breve{P}_y = \breve{P}_{ya} + c\int_{-\infty}^{t} \frac{\partial \hat{Q}_x}{\partial z} dt - \bar{k}_z \hat{Q}_x \; ; \qquad \breve{P}_{yb} = c\int_{-\infty}^{t} \frac{\partial \hat{Q}_x}{\partial z} dt \tag{13.80b}$$

$$\breve{P}_z = \breve{P}_{za} - c\int_{-\infty}^{t} \frac{\partial \hat{Q}_x}{\partial y} dt + \bar{k}_y \hat{Q}_x \; ; \qquad \breve{P}_{zb} = -c\int_{-\infty}^{t} \frac{\partial \hat{Q}_x}{\partial y} dt \tag{13.80c}$$

$$\hat{Q}_x = \hat{Q}_{xa} - \bar{k}_z \breve{P}_y + \bar{k}_y \breve{P}_z \tag{13.81a}$$

$$\breve{Q}_y = \breve{Q}_{ya} - c\int_{-\infty}^{t} \frac{\partial \hat{P}_x}{\partial z} dt + \bar{k}_z \hat{P}_x \; ; \qquad \breve{Q}_{yb} = -c\int_{-\infty}^{t} \frac{\partial \hat{P}_x}{\partial z} dt \tag{13.81b}$$

$$\breve{Q}_z = \breve{Q}_{za} + c\int_{-\infty}^{t} \frac{\partial \hat{P}_x}{\partial y} dt - \bar{k}_y \hat{P}_x \; ; \qquad \breve{Q}_{zb} = c\int_{-\infty}^{t} \frac{\partial \hat{P}_x}{\partial y} dt \tag{13.81c}$$

Substituting these into the left-hand sides of (13.78) and (13.79) and transforming to the time domain yields

$$\frac{1}{c}\frac{\partial P_{xa}}{\partial t} = \frac{\partial Q_z}{\partial y} - \frac{\partial Q_y}{\partial z} \tag{13.82a}$$

$$\frac{\kappa}{c}\frac{\partial P_{ya}}{\partial t} + \chi P_{ya} = -\frac{\partial Q_z}{\partial x} \; ; \qquad \frac{1}{c}\frac{\partial P_{yb}}{\partial t} = \frac{\partial \hat{Q}_x}{\partial z} \tag{13.82b}$$

$$\frac{\kappa}{c}\frac{\partial P_{za}}{\partial t} + \chi P_{za} = \frac{\partial Q_y}{\partial x} \; ; \qquad \frac{1}{c}\frac{\partial P_{zb}}{\partial t} = -\frac{\partial \hat{Q}_x}{\partial y} \tag{13.82c}$$

$$\frac{1}{c}\frac{\partial Q_{xa}}{\partial t} = -\frac{\partial P_z}{\partial y} + \frac{\partial P_y}{\partial z} \tag{13.83a}$$

$$\frac{\kappa}{c}\frac{\partial Q_{ya}}{\partial t} + \chi Q_{ya} = \frac{\partial P_z}{\partial x}\ ; \qquad \frac{1}{c}\frac{\partial Q_{yb}}{\partial t} = -\frac{\partial \hat{P}_x}{\partial z} \tag{13.83b}$$

$$\frac{\kappa}{c}\frac{\partial Q_{za}}{\partial t} + \chi Q_{za} = -\frac{\partial P_y}{\partial x}\ ; \qquad \frac{1}{c}\frac{\partial Q_{zb}}{\partial t} = \frac{\partial \hat{P}_x}{\partial y} \tag{13.83c}$$

The systems (13.82) and (13.83) show how the "a" and "b" portions of the fields are updated in terms of the total-field values. This is similar to the split field in the non-UPML regions, but the additional splitting eliminates the nonderivative terms from the right-hand sides. As a result, no spatial averaging is needed in these equations. The loss terms (terms multiplied by χ) are handled in the standard FDTD fashion. This is equivalent to using $\beta_{\text{UPML}} = 0.5$ in the modified temporal averaging in (13.55).

Once the "a" and "b" portions of the fields are known, then the components normal to the UPML boundary are found using the relations

$$\left(1 - \bar{k}_y^2 - \bar{k}_z^2\right)\hat{P}_x = P_{xa} - \bar{k}_y\left(Q_{za} + Q_{zb}\right) + \bar{k}_z\left(Q_{ya} + Q_{yb}\right) \tag{13.84a}$$

$$\left(1 - \bar{k}_y^2 - \bar{k}_z^2\right)\hat{Q}_x = Q_{xa} + \bar{k}_y\left(P_{za} + P_{zb}\right) - \bar{k}_z\left(P_{ya} + P_{yb}\right) \tag{13.84b}$$

Finally, the remaining field components are found from

$$P_y = P_{ya} + P_{yb} - \bar{k}_z \hat{Q}_x\ ; \qquad P_z = P_{za} + P_{zb} + \bar{k}_y \hat{Q}_x \tag{13.85}$$

$$Q_y = Q_{ya} + Q_{yb} + \bar{k}_z \hat{P}_x\ ; \qquad Q_z = Q_{za} + Q_{zb} - \bar{k}_y \hat{P}_x \tag{13.86}$$

Note that it is not necessary to find P_x from $\hat{P}_x$, since the fields within the UPML are irrelevant to the solution in the non-UPML region. The only requirement is that the reflection from the UPML region be sufficiently small.

Although this procedure is more complex than for the non-UPML regions, it is necessary to perform the additional splitting of the tangential components in the UPML region to make the technique stable. This has been verified by both a local stability analysis and simulations involving many time-steps.

13.8 APPLICATION OF THE PERIODIC FDTD METHOD

The periodic FDTD method has been used to model several types of periodic structures [10, 24]. This section discusses three such applications: EBG structures, FSS, and antenna arrays.

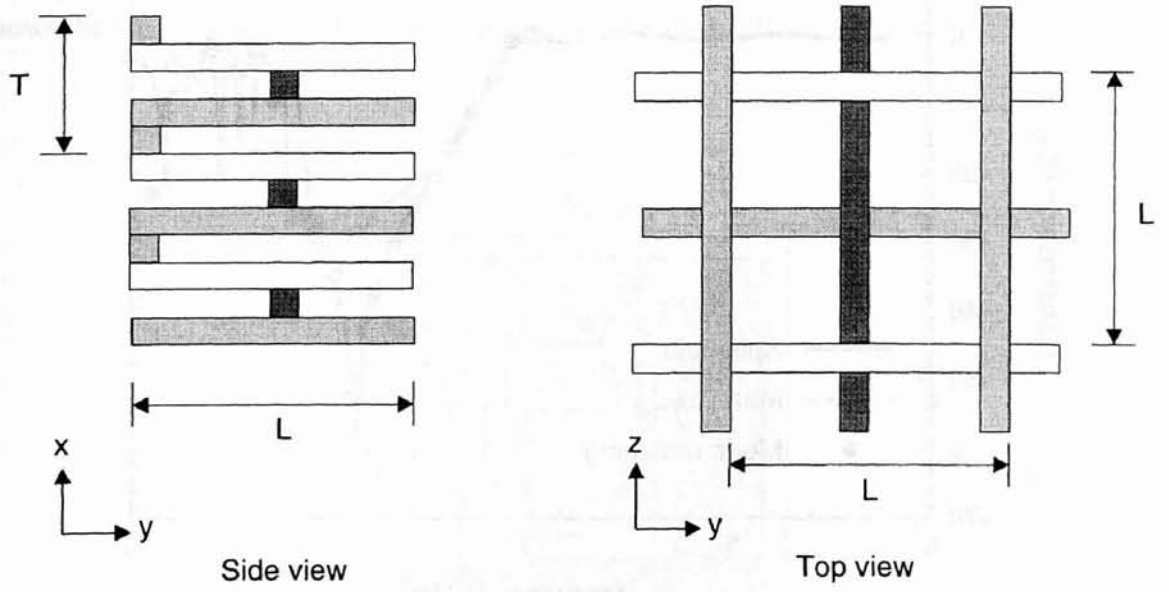


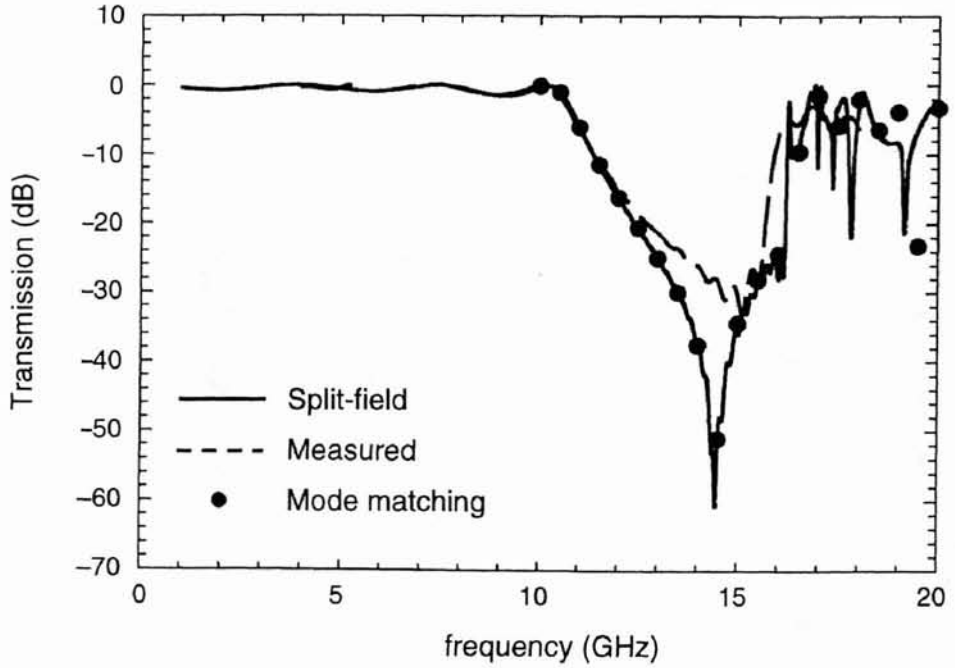
Fig. 13.27 Side and top views of the "woodpile" EBG structure comprised of dielectric rods of square cross section W = T/4 = 3.175 mm on a side, with $\varepsilon_r = 8$. The rod spacing is L = 11.11 mm, and the periodic cell is L×L.

13.8.1 Electromagnetic Bandgap Structures

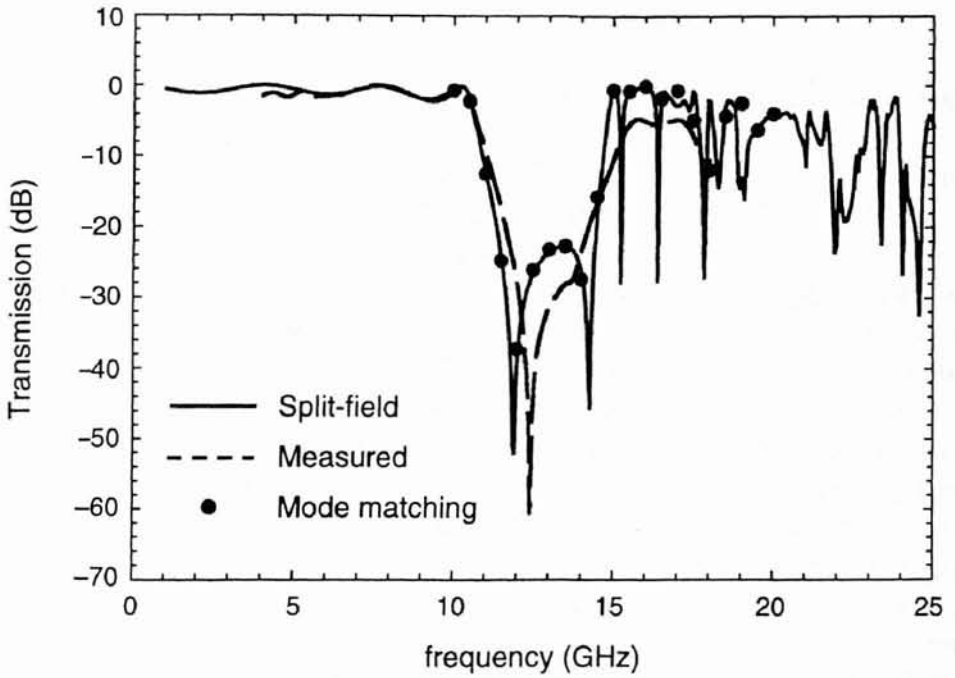
We first discuss the split-field periodic FDTD modeling of a particular three-dimensional EBG structure, the "woodpile" geometry shown in Fig. 13.27. This structure is formed by stacking dielectric rods [28]. Each dielectric rod has a square cross section 1/8 in (3.175 mm) on a side, and is characterized by the relative permittivity $\varepsilon_r = 8$. The spacing between adjacent rods in each layer is L = 0.4375 in (6.11 mm), and the rod orientation is alternated from layer to layer. Every fourth layer is located directly beneath each other.

Fig. 13.28(a) shows the transmission through three periods of this woodpile structure for a normally incident plane wave. The split-field FDTD results agree very well with the calculations of a frequency-domain mode-matching code, and reasonably well with experimental data. Characteristic of an EBG structure, there is a bandgap frequency range in which electromagnetic propagation is not allowed. In the bandgap, the transmission through the structure decreases sharply. Since the structure is lossless, the magnitude of its reflection coefficient in the bandgap is very close to unity, and the EBG device behaves much like a metal reflector. However, the phase of the reflection coefficient is usually a value other than 180°.

Fig. 13.28(b) shows the transmission versus frequency of the woodpile EBG structure for an incident angle of 30°. The shape of the bandgap is different in this case than in Fig. 13.28(a), and the bandgap is shifted to lower frequencies. Again, there is good agreement between the split-field FDTD and mode-matching methods.



(a) Normal incidence.



(b) Incident angle = 30°.

Fig. 13.28 Comparison of split-field periodic FDTD, frequency-domain mode-matching, and measured results for transmission through the woodpile EBG structure of Fig. 13.27.

13.8.2 Frequency-Selective Surfaces

We next discuss the application of the split-field periodic FDTD method to an FSS. Fig. 13.29 shows the top view of one cell of the doubly periodic FSS [14, 29].

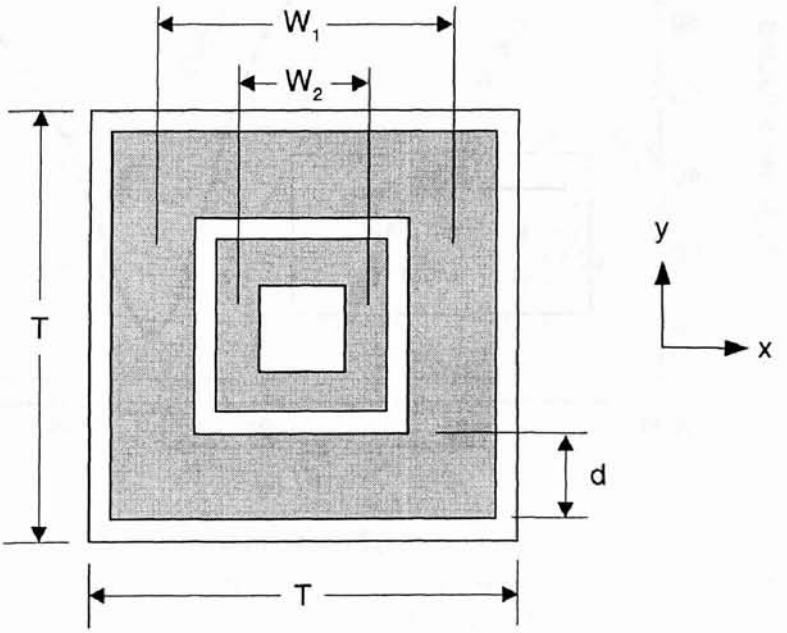


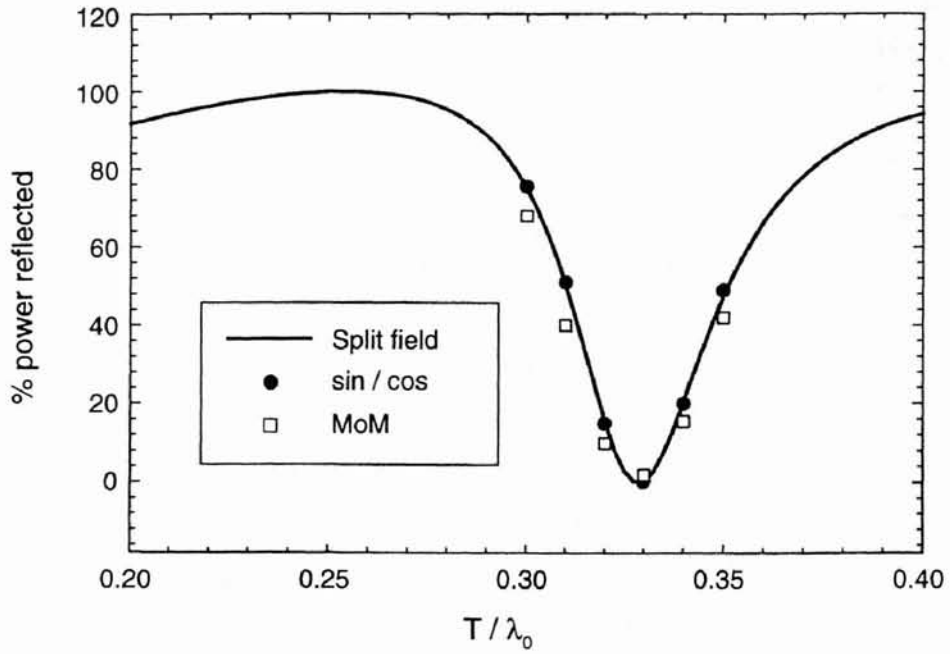
Fig. 13.29 Top view of a double-concentric-loop FSS that is doubly periodic in x and y. The conductors forming the loops have a thickness of T/16, where T is the periodic cell dimension. Also, $d = T/16$, $W_1/T = 0.875$, and $W_2/T = 0.6875$.

Fig. 13.30(a) shows data for the percentage of reflected power for the FSS of Fig. 13.29 for a normally incident plane wave. The split-field periodic FDTD results agree well with calculations resulting from sine-cosine periodic FDTD and frequency-domain method-of-moments models [14, 29]. Fig. 13.30(b) shows the corresponding results for a plane wave impinging at 60°. Good agreement is again observed.

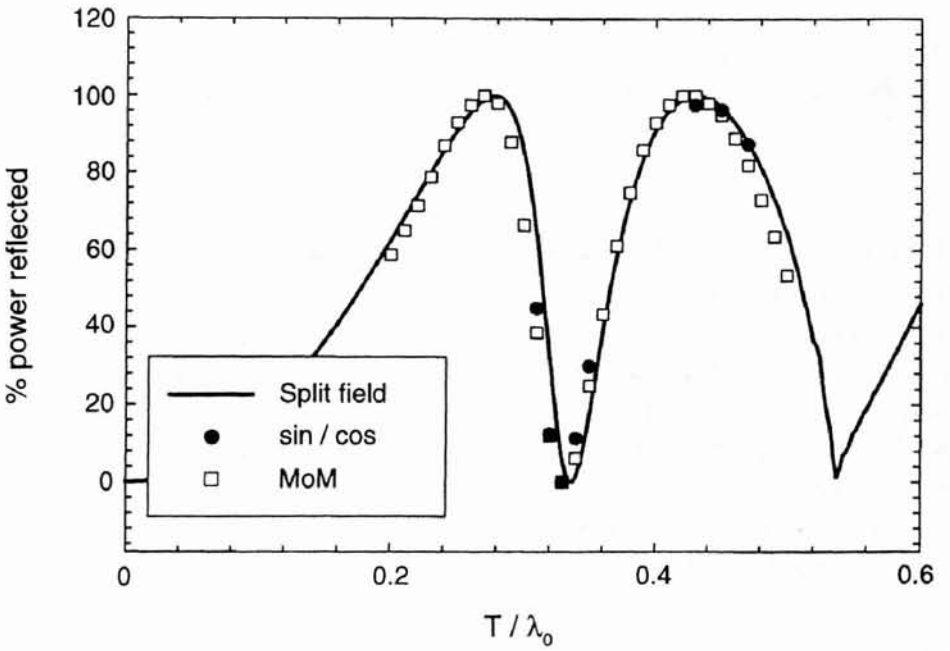
We note that the wideband, split-field FDTD method is much more efficient than either the single-frequency sine-cosine FDTD or moment-method approaches. In either Fig. 13.30(a) or 13.30(b), *all* of the split-field data are obtained from a single computer run, whereas each symbol represents a separate run for the other methods.

13.8.3 Antenna Arrays

The third and final class of application discussed is the analysis of a large antenna array. A single element in the array is modeled along with a simple transmission line feed structure, and periodic boundary conditions are used to model the effect of the element being in an array.



(a) Normal incidence.



(b) Incident angle = 60°.

Fig. 13.30 Comparison of split-field periodic FDTD, sin-cos periodic FDTD, and frequency-domain method-of-moments results for the percentage reflected power for the FSS of Fig. 13.29.

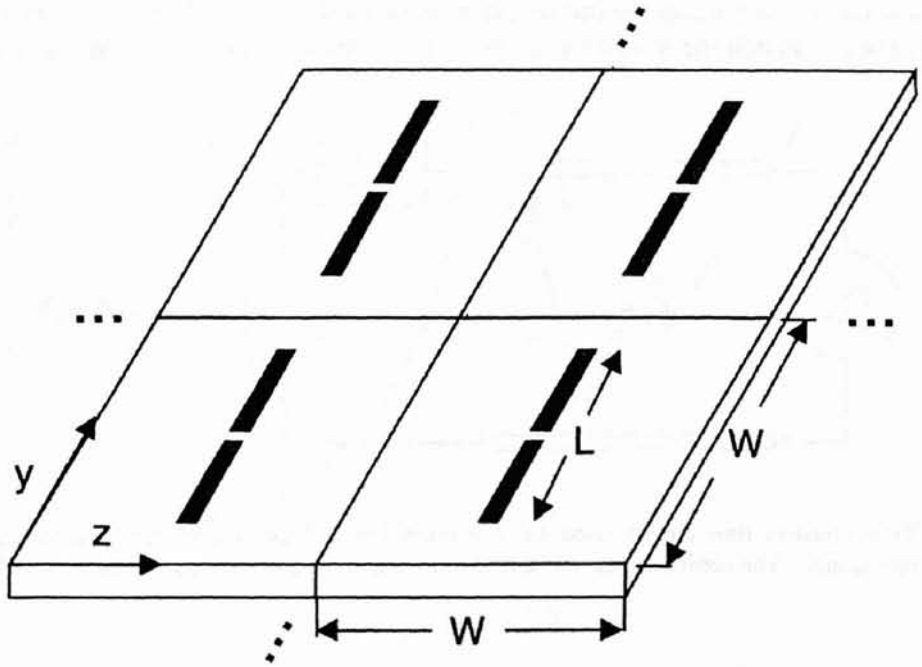


Fig. 13.31 Four unit cells of the infinite dipole array. The dipole spacing W is 0.5λ and the dipole length L is 0.39λ, with a gap of 0.01λ. The substrate is 0.19λ-thick with $\varepsilon_r = 2.55$, and is backed by a metal layer.

Fig. 13.31 illustrates the geometry chosen for the simulation, an array of dipole antennas. The dimensions of the problem match those given in [30] where a moment-method code is used to analyze the structure.

Although the split-field FDTD calculation uses a pulsed excitation and thus contains broadband data, we can compare its results only to the specific frequency used in [30]. At this frequency, the dipole length is $L = 0.39\lambda$, including a 0.01λ gap between the arms, while the dipole width is 0.01λ. The size of a unit cell is 0.5λ in both the y- and z-directions. Beneath the dipole array is a 0.19λ-thick dielectric substrate ($\varepsilon_r = 2.55$) above a ground plane. When discretized, the fundamental unit cell consists of $50 \times 50 \times 20$ FDTD space-lattice cells, with an additional 20 free-space lattice cells above the antenna, for a total grid (excluding the UPML region) of $50 \times 50 \times 40$ FDTD space-lattice cells.

Each dipole element is assumed to be excited at its terminals using a balanced 50Ω transmission line. The element's impedance is calculated from the reflected voltage wave on this line. (See Chapter 14, Section 14.3.3 for details of the transmission line feed.) We note that, because of the periodic boundary conditions, all of the antenna elements are driven simultaneously and phased to radiate in the scan direction. Thus, the reflected voltage wave on the feedline includes not only the reflection from the antenna terminals, which represents the isolated antenna impedance, but also any signal received because the other elements in the array are radiating as well. Therefore, the impedance calculated from the reflected voltage is the so-called "active impedance" of the element when the array is fully excited.

To compare the results of the split-field FDTD method with those in [30], we use the FDTD-calculated antenna impedance in the transmission line circuit shown in Fig. 13.32. Here, a series reactance is used to match the antenna impedance to the transmission line feed for a scan angle of 0°.

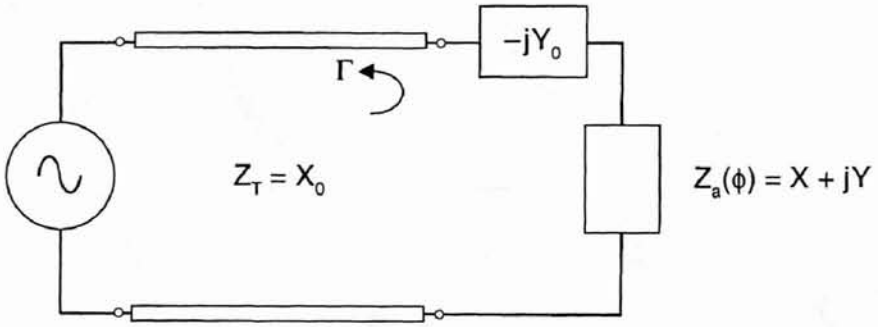


Fig. 13.32 Transmission line circuit used to determine the reflection coefficient from the antenna impedance. The conjugate reactance is used to match the antenna impedance at $\phi = 0°$.

Since the antenna impedance varies with scan angle, the match is no longer perfect as the array is steered away from 0°. With this matching arrangement, the reflection coefficient is given by

$$\Gamma = \frac{Z_A(\phi) - Z_A(0)}{Z_A(\phi) + Z_A^*(0)} \tag{13.87}$$

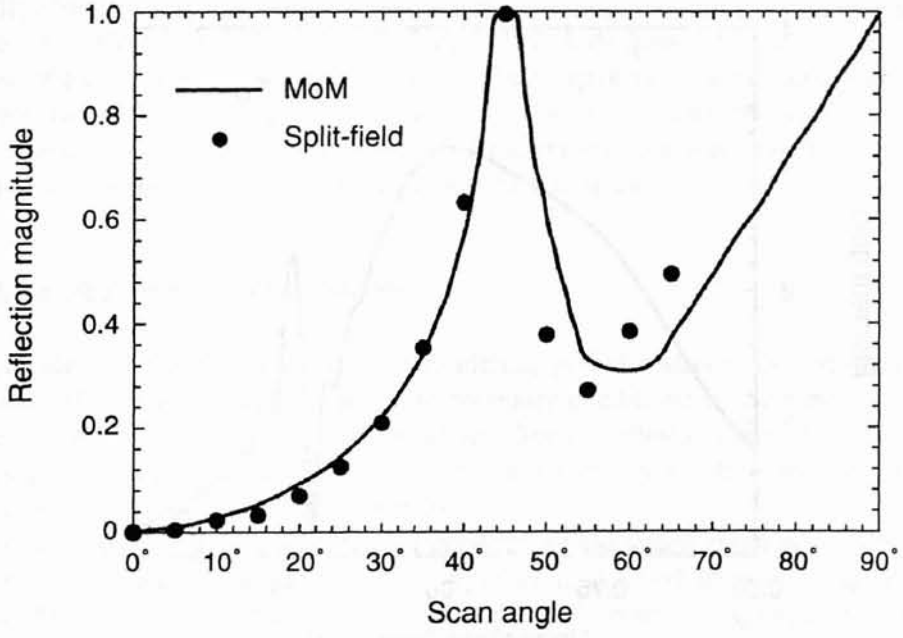
where $Z_A(0) = X_0 + jY_0$ is the active impedance of the dipole element at 0°, and $Z_A(\phi)$ is the active impedance when the array is scanned to the angle ϕ.

Fig. 13.33(a) compares, as a function of scan angle, $|\Gamma|$ calculated using split-field FDTD and (13.87) with MoM results for the same problem [30]. The agreement between the two techniques is good, especially for scan angles below 45°, considering that there are slight differences in the physical models used. Note that the magnitude of the reflection coefficient approaches unity near 45°. This is the array *scan-blindness* phenomenon that occurs because of the excitation of surface waves in the dielectric substrate.

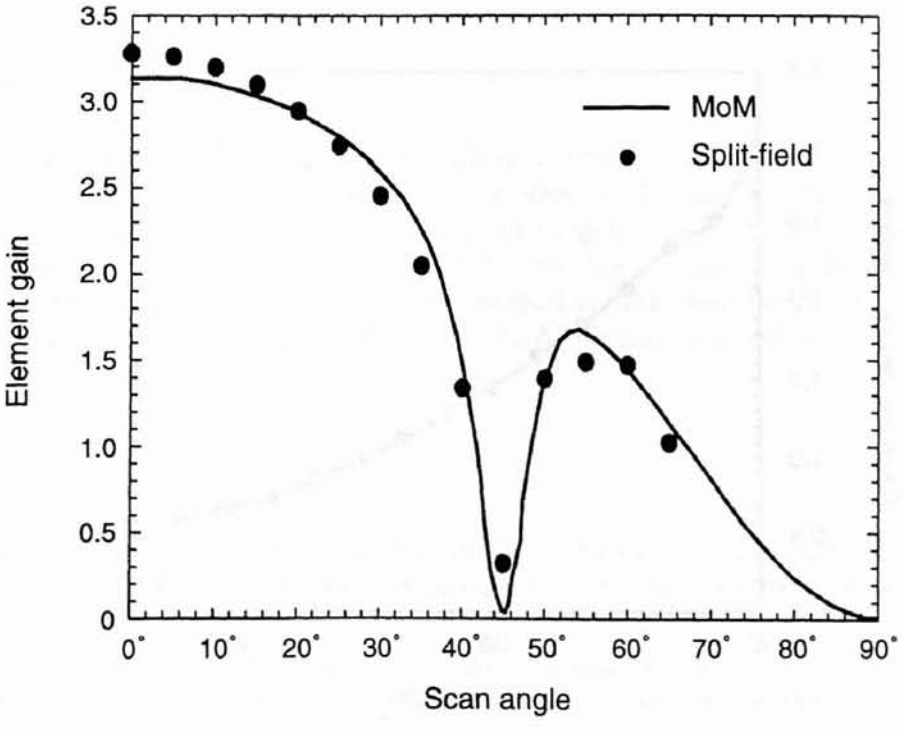
The radiated far-field is also computed with the split-field FDTD code. Since the element spacing is $\lambda/2$ at the comparison frequency, only the lowest order Floquet mode is computed. This result is then used to determine the gain of the element for comparison to the results given in [30]. The conversion from far-field to gain is given by

$$G(\phi) = T_0(\phi)\cos\phi \tag{13.88}$$

Fig. 13.33(b) compares, as a function of scan angle, split-field FDTD results for the gain of a single dipole element embedded in the array of Fig. 13.31, with MoM results for the same problem [30]. Again, there is good agreement between the two computational approaches. Here, the scan blindness shows up as a reduction in the active gain of the element. The angle at which scan blindness occurs is a function of the operating frequency. Since the FDTD results cover a broad frequency range, these data can be used to track the angle at which scan blindness occurs.

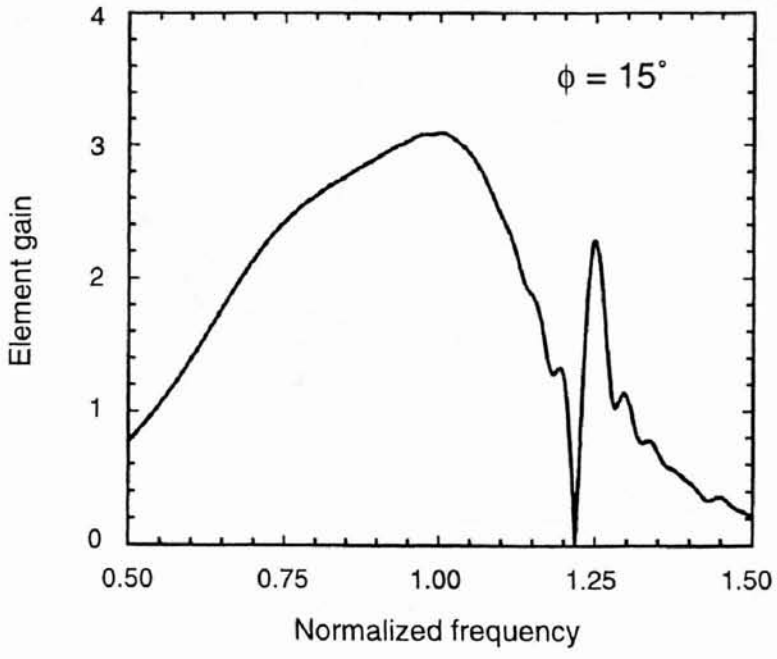


(a)

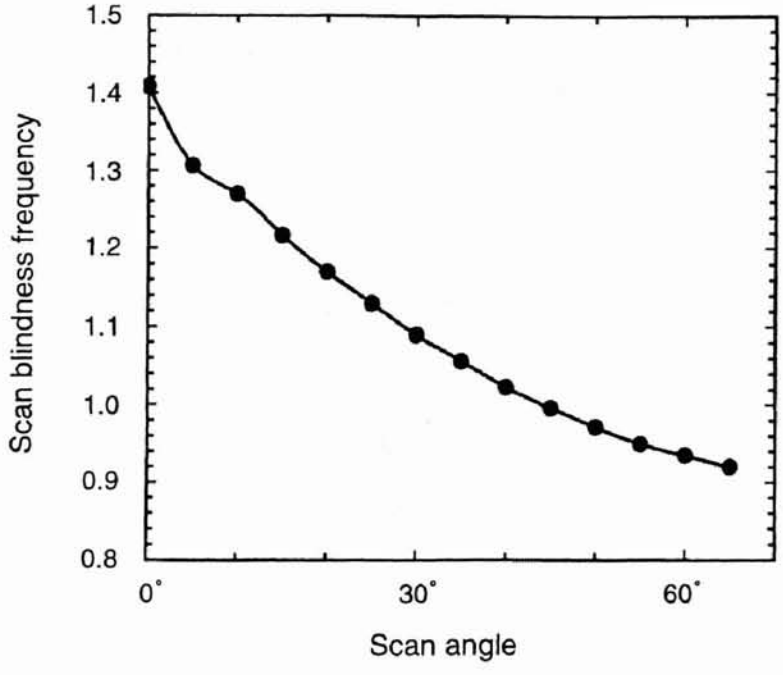


(b)

Fig. 13.33 Comparison of split-field FDTD and MoM results for the infinite dipole array of Fig. 13.31: (a) magnitude of the reflection coefficient as a function of the scan angle; (b) gain of a single dipole element embedded in the infinite array as a function of the scan angle.



(a)



(b)

Fig. 13.34 Further examination of the phenomenon of scan blindness for the antenna array of Fig. 13.31: (a) element gain as a function of normalized frequency when the array is scanned to 15°; (b) variation of the normalized scan-blindness frequency with scan angle.

Fig. 13.34 provides a further examination of the phenomenon of scan blindness. Specifically, Fig. 13.34(a) shows the element gain as a function of frequency when the array is steered to 15°. Here, a normalized frequency of 1.0 corresponds to the λ used earlier in the description of the array geometry of Fig. 13.31. The sharp dip near 1.2 indicates the frequency at which scan blindness occurs for this angle. Fig. 13.34(b) shows the variation of the scan-blindness frequency with scan angle. Overall, we comment that scan blindness can be a strong limitation on the useable scan range of broadband antenna arrays.

13.9 SUMMARY AND CONCLUSIONS

The FDTD method has strong appeal for modeling periodic structures. However, there are fundamental difficulties in applying periodic boundary conditions in the time domain. Several techniques have been proposed for this problem. Some methods approximate the problem by using multiple unit cells, while others use a single frequency to allow the periodic boundary condition to be applied in the frequency domain.

The most promising approaches involve applying a field transformation to eliminate the time delay across the FDTD grid. Most of this chapter is devoted to discussing techniques for discretizing Maxwell's equations when such a field transformation is applied. Of the techniques investigated, two stand out as the best: the multiple-spatial-grid method and the split-field method. A computer program utilizing either of these approaches provides the researcher with a valuable tool for analyzing electromagnetic wave interactions with periodic structures.

ACKNOWLEDGMENTS

The authors specifically acknowledge the significant contributions of Dr. Paul Harms, Dr. Alan Roden, and Dr. Eric Kuster to this chapter. Dr. Harms and Dr. Roden performed most of the initial development of the periodic-FDTD computer programs, and were the originators of the split-field concept. Dr. Kuster developed the mode-matching computer program and participated in many fruitful discussions. Without their contributions, the work discussed in this chapter would not have been possible. The authors also thank Professor Stephen Gedney for his many helpful discussions on numerical stability.

REFERENCES

[1] Cangellaris, A. C., M. Gribbons, and G. Sohos, "A hybrid spectral / FDTD method for the electromagnetic analysis of guided waves in periodic structures," *IEEE Microwave Guided Wave Lett.*, Vol. 3, 1993, pp. 375–377.

[2] Celuch-Marcysiak, M., and W. K. Gwarek, "Spatially looped algorithms for time-domain analysis of periodic structures," *IEEE Trans. Microwave Theory Tech.*, Vol. 43, 1995, pp. 860–865.

[3] Okoniewski, M., "Vector wave equation 2D FDTD method for guided wave problems," *IEEE Microwave Guided Wave Lett.*, Vol. 3, 1993, pp. 307–309.

[4] Kuster, E. J., private communication to the authors.

[5] Tsay, W. J., and D. M. Pozar, "Application of the FDTD technique to periodic problems in scattering and radiation," *IEEE Microwave Guided Wave Lett.*, Vol. 3, 1993, pp. 250–252.

[6] Navarro, E. A., B. Gimeno, and J. L. Cruz, "Modeling of periodic structures using the finite difference time domain method combined with the Floquet theorem," *Electronics Lett.*, Vol. 29, 1993, pp. 446–447.

[7] Alexanian, N. J. Kolias, R. C. Compton, and R. A. York, "Three-dimensional FDTD analysis of quasi-optical arrays using Floquet boundary conditions and Berenger's PML," *IEEE Microwave Guided Wave Lett.*, Vol. 6, 1996, pp. 138–140.

[8] Chan, H., S. H. Lou, L. Tsang, and J. A. Kong, "Electromagnetic scattering of waves by random rough surface: a finite-difference time-domain approach," *Microwave Optical Tech. Lett.*, Vol. 4, 1991, pp. 355–359.

[9] Kelly, P. K., J. G. Maloney, B. L. Shirley, and R. L. Moore, "Photonic bandgap structures of finite thickness: theory and experiment," *Proc. 1994 IEEE Antennas Propagat. Soc. Intl. Symp.*, Seattle, WA, 1994, Vol. 2, pp. 718–721.

[10] Kesler, M. P., J. G. Maloney, and G. Smith, "Antenna design with the use of photonic bandgap materials as all-dielectric planar reflectors," *Microwave Optical Tech. Lett.*, Vol. 11, 1996, pp. 169–174.

[11] Ren, J., O. P. Gandhi, L. R. Walker, J. Fraschilla, and C. R. Boerman, "Floquet-based FDTD analysis of two-dimensional phased array antennas," *IEEE Microwave Guided Wave Lett.*, Vol. 4, 1994, pp. 109–112.

[12] Lee, K. H., and S. R. Laxpati, "FDTD analysis of an infinite array of microstrip patches," *Proc. 1996 IEEE Antennas Propagat. Soc. Intl. Symp.*, Baltimore, MD, July 1996, Vol. 2, pp. 1284–1287.

[13] Ko, W. L., and R. Mittra, "Implementation of Floquet boundary condition in FDTD for FSS analysis," *Proc. 1993 IEEE Antennas Propagat. Soc. Intl. Symp.*, Ann Arbor, MI, 1993, Vol. 1, pp. 14–17.

[14] Harms, P., R. Mittra, and W. Ko, "Implementation of the periodic boundary condition in the finite-difference time-domain algorithm for FSS structures," *IEEE Trans. Antennas Propagat.*, Vol. 42, 1994, pp. 1317–1324.

[15] Prescott, D. T., and N. V. Shuley, "Extensions to the FDTD method for the analysis of infinitely periodic arrays," *IEEE Microwave Guided Wave Lett.*, Vol. 4, 1994, pp. 352–354.

[16] Davidson, B., A. G. Smith, and J. J. van Tonder, "The analysis, measurement and design of frequency selective surfaces," *10th Intl. Conf. Antennas Propagat.*, 1997, IEE conf. pub. No. 436, pp. 156–159.

[17] Noda, T., T. Kanetani, and K. Uchidat, "FVTD analysis of metallic grating," *IEICE Trans. Electronics*, Vol. E79-C, 1996, pp. 1772–1775.

[18] Marek, J. R., and J. MacGillivray, "A method for reducing run-times of out-of-core FDTD problems," *Proc. 9th Ann. Review Progress Applied Computational Electromagn.*, Monterey, CA, 1993, pp. 344–351.

[19] Veysoglu, M. E., R. T. Shin, and J. A. Kong, "A finite-difference time-domain analysis of wave scattering from periodic surfaces: Oblique incidence case," *J. Electromagn. Waves Apps.*, Vol. 7, 1993, pp. 1595–1607.

[20] Kao, Y-C., and R. G. Atkins, "A finite-difference time-domain approach for frequency selective surfaces at oblique incidence," *Proc. 1996 IEEE Antennas Propagat. Soc. Intl. Symp.*, Baltimore, MD, 1996, pp. 1432–1435.

[21] Ray, S. L., "Grid decoupling in finite element solutions of Maxwell's equations," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 443–445.

[22] Roden, J. A., *Electromagnetic Analysis of Complex Structures Using the FDTD Technique in General Curvilinear Coordinates*, Ph.D. thesis, Univ. Kentucky, Lexington, KY, 1997.

[23] Roden, J. A., S. D. Gedney, M. P. Kesler, J. G. Maloney, P. H. Harms, "Time-domain analysis of periodic structures at oblique incidence: Orthogonal and nonorthogonal FDTD implementations," *IEEE Trans. Microwave Theory Tech.*, Vol. 46, 1998, pp. 420–427.

[24] Harms, P. H., J. A. Roden, J. G. Maloney, M. P. Kesler, E. J. Kuster, and S. D. Gedney, "Numerical analysis of periodic structures using the split-field algorithm," *Proc. 13th Ann. Review Progress Applied Computational Electromagn.*, Monterey, CA, 1997, pp. 104–111.

[25] Gedney, S. D., and U. Navsariwala, "An unconditionally stable finite-element time-domain solution of the vector wave equation," *IEEE Microwave Guided Wave Lett.*, Vol. 5, 1995, pp. 332–334.

[26] Zienkiewicz, O. C., "A new look at the Newmark, HouBolt, and other time-stepping formulas: A weighted residual approach," *Earthquake Engineering and Structural Dynamics*, Vol. 5, 1977, pp. 413–418.

[27] Roden, J. A., and S. D. Gedney, "Efficient implementation of the uniaxial-based PML media in three-dimensional nonorthogonal coordinates with the use of the FDTD technique," *Microwave Optical Tech. Lett.*, Vol. 14, 1997, pp. 71–75.

[28] Ho, K. M., C. T. Chan, C. M. Soukoulis, R. Biswas, and M. Sigalis, "Photonic band gaps in three dimensions: New layer-by-layer periodic structures," *Solid State Comm.*, Vol. 89, 1994, pp. 413–416.

[29] Webb, D. B., *Design Techniques for Bandpass Frequency Selective Surfaces*, M.S. thesis, Univ. Illinois, Urbana-Champaign, IL, 1992.

[30] Pozar, D. M., "Analysis of finite phased arrays of printed dipoles," *IEEE Trans. Antennas Propagat.*, Vol. 33, 1985, pp. 1045–1053.

PROJECTS

P13.1 Write a direct-field periodic-boundary FDTD program for the normal-incidence case. Verify the program by replicating the results of Fig. 13.7.

P13.2 Write a periodic-boundary FDTD program using the sine-cosine method. Verify the program by replicating the results of Fig. 13.10.

P13.3 Write a periodic-boundary FDTD program using the angled-update method. Verify the program by replicating the results of Fig. 13.13.

P13.4 Write a periodic-boundary FDTD program using the multiple-grid method. Verify the program by replicating the results of Fig. 13.19.

P13.5 Write a periodic-boundary FDTD program using the split-field method. Verify the program by replicating the results of Fig. 13.22.

Chapter 14

Modeling of Antennas

James Maloney, Glenn Smith, Eric Thiele, Om Gandhi, Nicolas Chavannes, and Susan Hagness

14.1 INTRODUCTION

The FDTD method was initially used primarily to model the scattering of electromagnetic waves from objects. Later, after techniques were developed for including sources within the computation lattice, the method was used to model radiating structures. The first antennas analyzed with the FDTD method were simple structures such as rotationally symmetric, perfectly conducting, cylindrical and conical monopoles in free space [1]. The excellent agreement between the FDTD results and measured data for these simple antennas showed the potential of the method for modeling antennas having realistic complexity.

Since these early investigations, the complexity of the antennas modeled with FDTD has steadily increased. A partial list of the different types of antennas analyzed includes loaded linear elements [2, 3] and arrays of linear elements [4–6], waveguide and TEM horns [7–12], planar designs (microstrip patches and spirals) [13–16], dielectric resonators and lenses [17–19], and active antennas [20]. A more complete accounting is in the detailed review in [21].

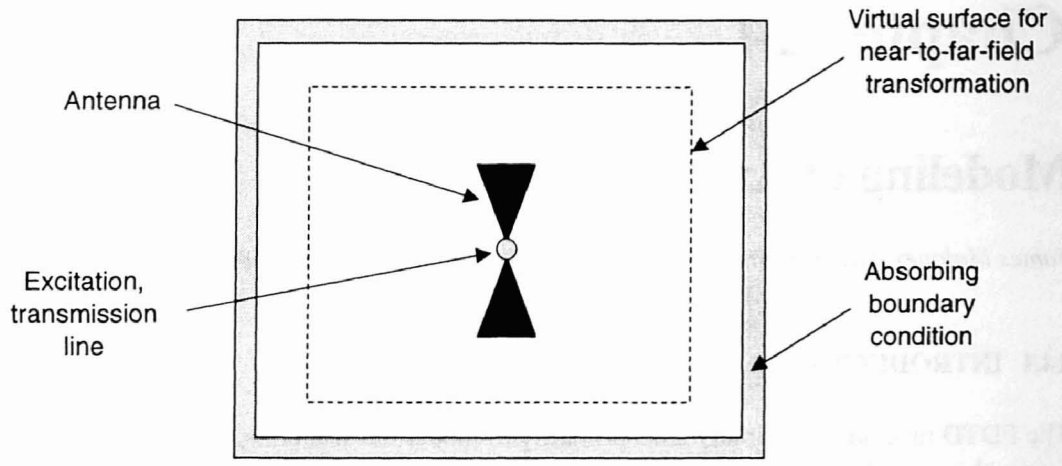
Today, the FDTD method is routinely used to analyze practical antennas transmitting or receiving energy in the presence of realistic material structures. The ability of this method to accurately determine wideband near-field and far-field antenna characteristics in the presence of complex materials (inhomogeneous, anisotropic, and frequency dispersive) allows antenna designers great flexibility, and opens possibilities for novel antenna concepts.

This chapter discusses FDTD modeling theory and procedures for antennas. It presents case studies that indicate the potentially high accuracy of the method, and reviews selected recent applications that point toward future modeling possibilities of technological importance.

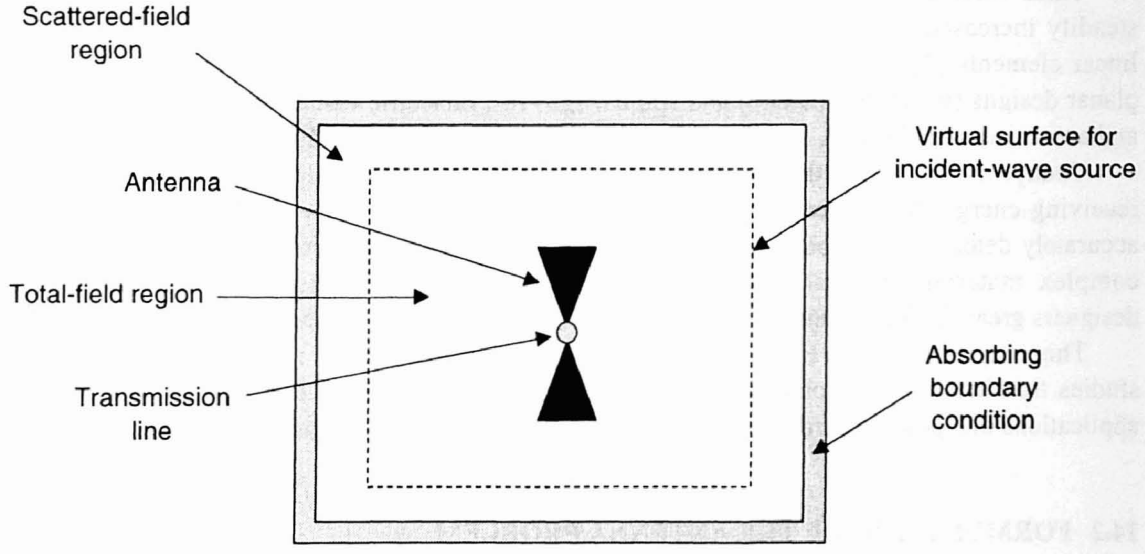
14.2 FORMULATION OF THE ANTENNA PROBLEM

14.2.1 Transmitting Antenna

Fig. 14.1(a) shows the basic elements for the FDTD analysis of a transmitting antenna. This figure is for a cross section through the computational volume. The antenna is placed near the center of the volume, and is fed by a transmission line or waveguide. As we discuss later, the details of the transmission line can be completely modeled within the FDTD lattice, or the line may be simply modeled as a one-dimensional FDTD grid attached to the antenna at the feed



(a) Transmitting antenna case.



(b) Receiving antenna case.

Fig. 14.1 Basic elements in the FDTD analysis of antennas.

point. The computational volume is surrounded by an ABC to reduce the reflection of the radiated energy at the boundaries of the volume. At this time, the PML ABC discussed in Chapter 7 is the most effective for this purpose.

The FDTD method provides the electromagnetic field for all space lattice points within the computational volume. However, for many antenna applications, we want to know the electromagnetic field at a large distance from the antenna — the radiated or far-zone field. This field can be obtained by applying a *near-to-far-field* (NTFF) transformation, as discussed in Chapter 8. To perform this transformation, a virtual surface S' is placed around the antenna, as indicated by the dashed line in Fig. 14.1(a). The field on this surface is obtained for the time period of interest. This field is then transformed to obtain the far field using a version of Huygens' principle for electromagnetic fields. One analytic form for this principle is

$$\boldsymbol{E}^r(\boldsymbol{r},t) = \frac{\mu_0}{4\pi r}\oiint_{S'}\left\{\hat{\boldsymbol{r}}\times\hat{\boldsymbol{r}}\times\frac{\partial}{\partial t'}[\hat{\boldsymbol{n}}'\times\boldsymbol{H}(\boldsymbol{r}',t')] - \frac{1}{\eta_0}\hat{\boldsymbol{r}}\times\frac{\partial}{\partial t'}[\hat{\boldsymbol{n}}'\times\boldsymbol{E}(\boldsymbol{r}',t')]\right\}_{t'=t_r} dS' \tag{14.1}$$

where the retarded time is

$$t_r = t - (r - \hat{\boldsymbol{r}}\cdot\boldsymbol{r}')/c \tag{14.2}$$

Here, $\boldsymbol{r}'$ is a source point on surface S', which has the outward-pointing unit normal vector $\hat{\boldsymbol{n}}'$, and $\boldsymbol{r}$ is an observation point in the far zone where the radiated field $\boldsymbol{E}^r$ is to be calculated.

In some situations, the near field is needed at points so far from the antenna that it is impractical to extend the computational volume to include these points. A near-to-near-field transformation may be used to obtain these results. For this case, the FDTD analysis is performed for a volume such as that shown in Fig. 14.1(a), and the field on the surface of the volume is transformed to obtain the near field outside the volume. Details of this transformation are found in [22, 23].

Note that in certain situations (such as for a microstrip), more than one material region can intersect the transformation surface. The modifications to the NTFF transformation required for this case are discussed in [24].

14.2.2 Receiving Antenna

Fig. 14.1(b) shows the basic elements for the FDTD analysis of a receiving antenna. Again, the figure is for a cross section through the computational volume. The response of the antenna, usually the signal in the attached transmission line, is determined when the antenna is irradiated by an incident plane wave having the fields $[\boldsymbol{E}_{\text{inc}}(\boldsymbol{r}, t), \boldsymbol{H}_{\text{inc}}(\boldsymbol{r}, t)]$.

In the FDTD formulation, a source is placed on the virtual surface of a volume containing the antenna; the dashed line in Fig. 14.1(b) indicates this virtual surface. This source produces the incident field within the volume and a null field outside the volume. An analytic form for this source specifies electric and magnetic surface currents with the values

$$\boldsymbol{J}_s(\boldsymbol{r}, t) = -\hat{\boldsymbol{n}}\times\boldsymbol{H}_{\text{inc}}(\boldsymbol{r}, t)\ ; \qquad \boldsymbol{M}_s(\boldsymbol{r}, t) = \hat{\boldsymbol{n}}\times\boldsymbol{E}_{\text{inc}}(\boldsymbol{r}, t) \tag{14.3a, b}$$

We note that the electromagnetic field within this volume is the total field (the incident field plus the scattered field), while outside the volume, it is only the scattered field.

14.2.3 Symmetry

The symmetry of an antenna and its excitation can be used to reduce the size of the FDTD modeling volume. This is illustrated for a bow-tie transmitting antenna in Figs. 14.2(a, b). Here, the antenna is located in the y-z plane. It is formed from a thin PEC and fed at its center. First, the antenna is replaced by its monopole configuration. That is, the top half of the antenna is placed over a PEC image plane, here the x-y plane as shown in Fig. 14.2(b). Next, the x-z plane and the y-z plane are made PMCs to exploit the remaining symmetry. For this example, the use of symmetry reduces the modeling volume, and thus the computer storage and running time, by a factor of 8:1.

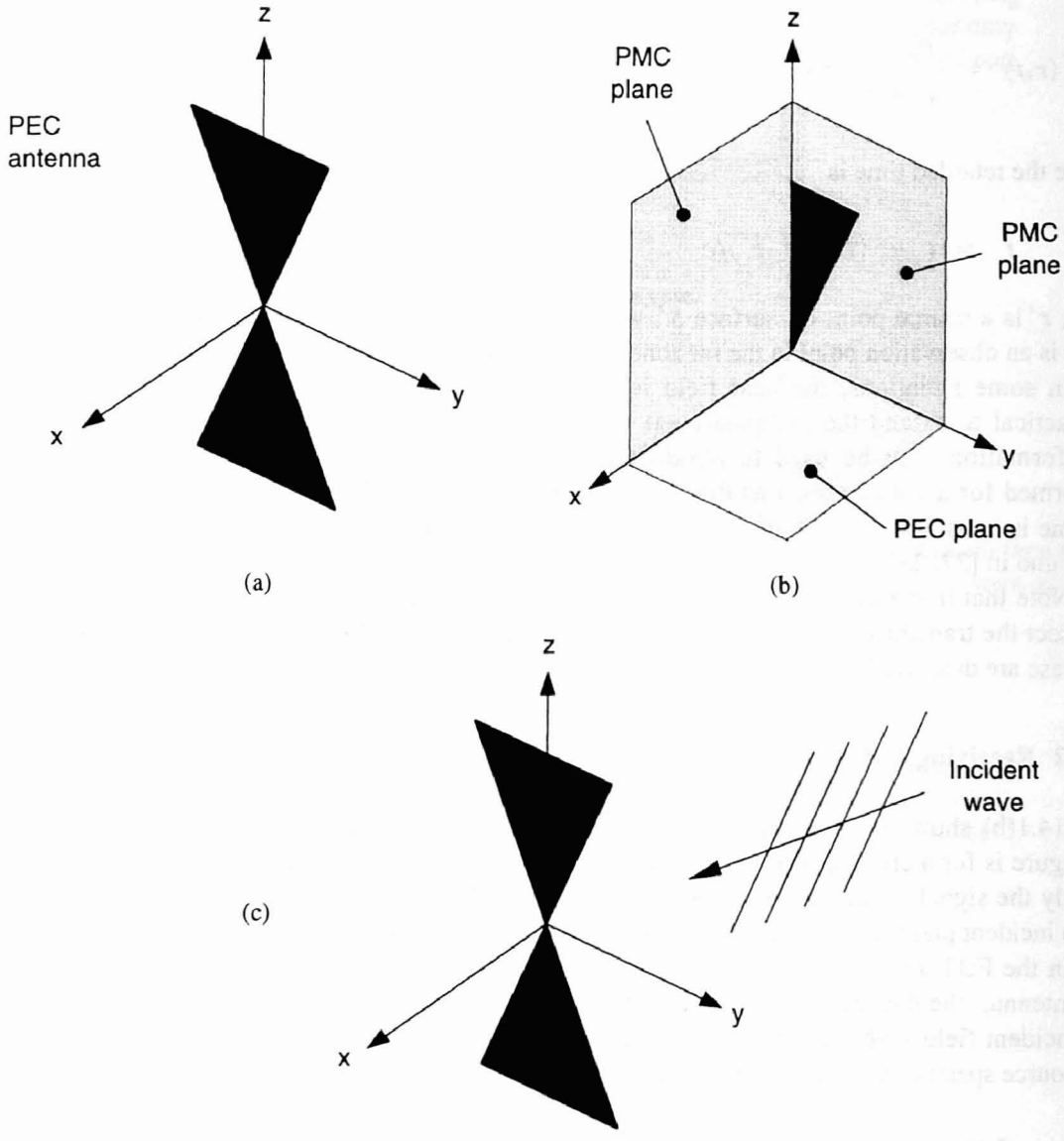


Fig. 14.2 Potential use of symmetry in the modeling of antennas: (a) bow-tie antenna example; (b) use of symmetry for a transmitting bow-tie antenna to reduce the modeling volume by 8:1; (c) receiving antenna with plane wave incident from an arbitrary direction.

Fig. 14.2(c) shows the same antenna as a receiver for a plane wave incident from an arbitrary direction. In this case, none of the symmetry of the antenna can be exploited to reduce the modeling volume. However, if the incident wave possesses some of the same symmetry as the antenna, a reduction in the modeling volume is possible. For example, if a plane wave is incident in the x direction with $\boldsymbol{E}_{\text{inc}}$ oriented in the z direction, the PEC x-y plane and PMC x-z plane shown in Fig. 14.2(b) are used to decrease the computational volume by a factor of 4:1.

14.2.4 Excitation

The excitation to be modeled for an antenna depends upon the particular application. For example, when we are determining the locations on an antenna at which reflections or radiation originate, a smooth, narrow pulse such as the Gaussian shown in Fig. 14.3(a) may be an appropriate excitation. If τ_p denotes the characteristic time of this pulse, then the incident voltage in the feeding transmission line of the transmitting antenna would be

$$V_{\text{inc}}(t) = V_0 \exp\left[-(t/\tau_p)^2/2\right] \tag{14.4}$$

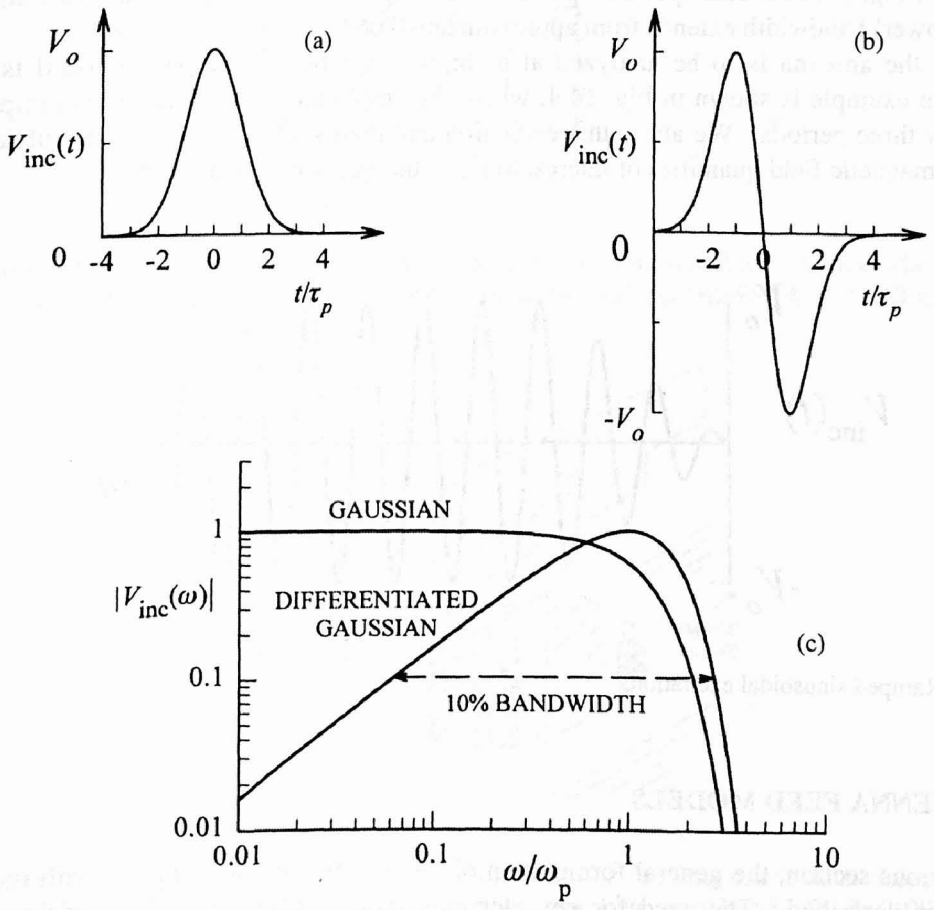


Fig. 14.3 Potential antenna excitations: (a) Gaussian pulse; (b) differentiated Gaussian pulse; (c) spectra of the Gaussian pulse and the differentiated Gaussian pulse.

Fig. 14.3(c) shows the Fourier transform of this signal, which is given by

$$V_{\text{inc}}(\omega) = \sqrt{2\pi}\,\tau_p\,V_0\,\exp\left[-(\omega\tau_p)^2/2\right] \tag{14.5}$$

When the antenna is to be analyzed over a band of frequencies, it is efficient to excite the antenna with a pulsed signal and then use the Fourier transform to obtain the desired frequency-domain response. The Gaussian pulse can be used for this purpose; however, its spectrum has significant low-frequency content, as seen in Fig. 14.3(c). For some problems, this can produce an unacceptably long settling time for the solution. An alternative is to use the differentiated Gaussian pulse shown in Fig. 14.3(b). For the transmitting antenna, the incident voltage and its corresponding Fourier spectrum in the feeding transmission line would then be

$$V_{\text{inc}}(t) = -V_0\left(\frac{t}{\tau_p}\right)\exp\left\{-\left[(t/\tau_p)^2-1\right]/2\right\} \tag{14.6}$$

$$V_{\text{inc}}(\omega) = -j\omega\sqrt{2\pi}\,\tau_p^2\,V_0\,\exp\left\{-\left[(\omega\tau_p)^2-1\right]/2\right\} \tag{14.7}$$

as shown in Fig. 14.3(c). This spectrum peaks at the frequency $\omega_p = 1/\tau_p$, and the 10%-amplitude (–20 dB power) bandwidth extends from approximately 0.06 to $2.8\omega_p$.

When the antenna is to be analyzed at a single frequency, a ramped sinusoid is a good choice. An example is shown in Fig. 14.4, where the amplitude of the sinusoid is ramped from 0 to 1 over three periods. We apply this excitation and then wait a length of time sufficient for the electromagnetic field quantities of interest to reach the sinusoidal steady-state.

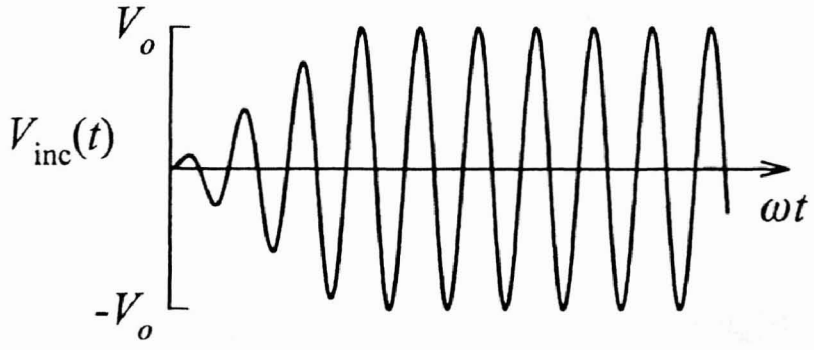


Fig. 14.4 Ramped sinusoidal excitation.

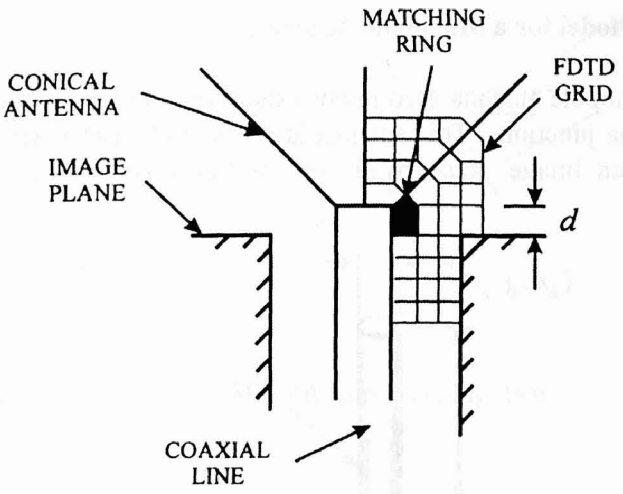
14.3 ANTENNA FEED MODELS

In the previous section, the general formulation of the antenna problem for use with the FDTD method was described. The need for key elements such as antenna feed models and NFFF transformations was introduced. This section and those that follow discuss the details of the implementation of these elements in the FDTD method. We begin with the antenna feed models.

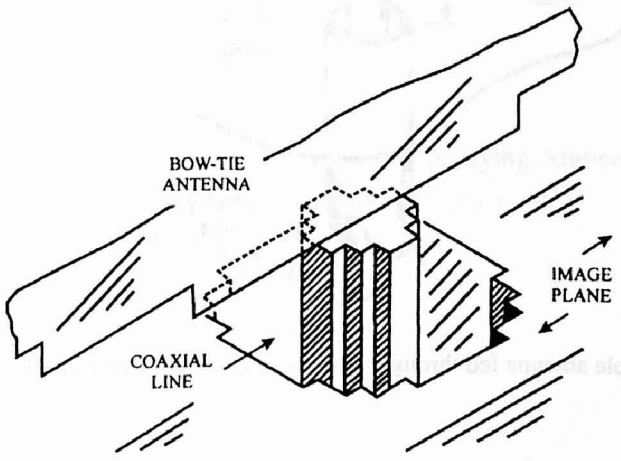
14.3.1 Detailed Modeling of the Feed

The antenna is connected at its terminals to a transmission line or waveguide, into which a signal is either introduced (transmitting) or extracted (receiving). When necessary, the full details of the transmission line and transmission line / antenna junction are included in the FDTD model. This may be desirable, for example, when designing an antenna for very low internal reflections.

Fig. 14.5(a) illustrates an example of a detailed antenna feed model, the junction between a coaxial line and a metallic conical antenna [25]. Using FDTD modeling, the details of the junction (distance d and the size of the matching ring) were adjusted to make the peak reflected voltage from this junction –46 dB relative to the peak of the incident voltage.



(a) Rotationally symmetric conical monopole antenna fed from coaxial line. *Source:* Maloney and Smith, *IEEE Trans. Antennas and Propagation*, 1993, pp. 940-957, © 1993 IEEE.



(b) Staircasing of geometry for bow-tie monopole antenna fed from coaxial line. *Source:* Shlager et al., *IEEE Trans. Antennas and Propagation*, 1994, pp. 975–982, © 1994 IEEE.

Fig. 14.5 Details of two antenna feeds directly modeled using the FDTD method.

The antenna shown in Fig. 14.5(a) is rotationally symmetric, allowing conformal modeling using a rotationally symmetric FDTD space lattice (see Chapter 12). For more complicated antennas that do not possess such symmetry, a staircased approximation to the geometry of the antenna and transmission line can be used. An example of this is shown in Fig. 14.5(b), the junction between a bow-tie antenna and a coaxial line [26].

The dimensions of the transmission line and antenna feed region are usually smaller than the other key antenna dimensions. Thus, the FDTD cell size required to model the antenna feed is often smaller than that needed to model only the antenna. If a small cell is used for the entire lattice, the computer storage and run-time may be very large. Alternatives are to use a graded mesh with the cell size increasing away from the antenna feedpoint, or an approximate model for the feedline and antenna junction. Two approaches to the latter are examined below.

14.3.2 Simple Gap Feed Model for a Monopole Antenna

Fig. 14.6 illustrates the monopole antenna used in our examples of an approximate FDTD model for the feedline and antenna junction. This antenna is a PEC cylindrical wire of radius a and length h driven through an image plane by a coaxial transmission line of characteristic impedance Z_0.

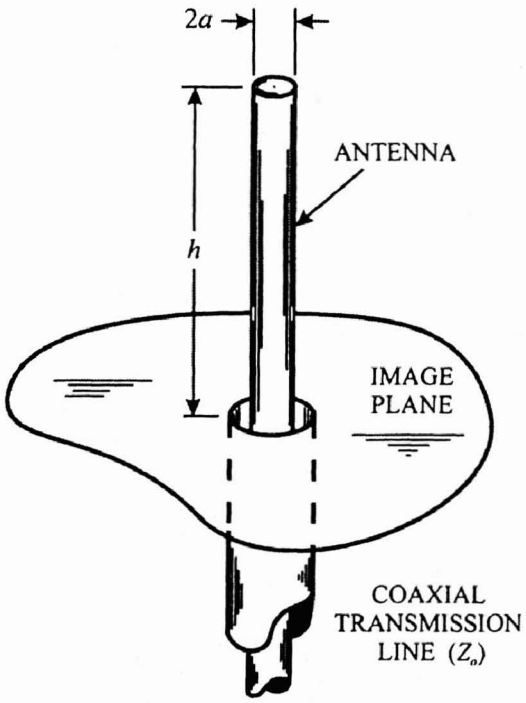


Fig. 14.6 Cylindrical monopole antenna fed through an image plane from a coaxial line.

Fig. 14.7 illustrates the simplest method to feed this antenna. In the feed region, a gap of length Δz exists between the wire and the image plane. The lower end of the gap is on the image plane at the point $(x = i_a\Delta x,\ y = j_a\Delta y,\ z = k_a\Delta z)$, for which we use the notation (i_a, j_a, k_a). In the wire, the E_z component of the electric field is set to zero, and on the image plane, the E_x and E_y components are set to zero.

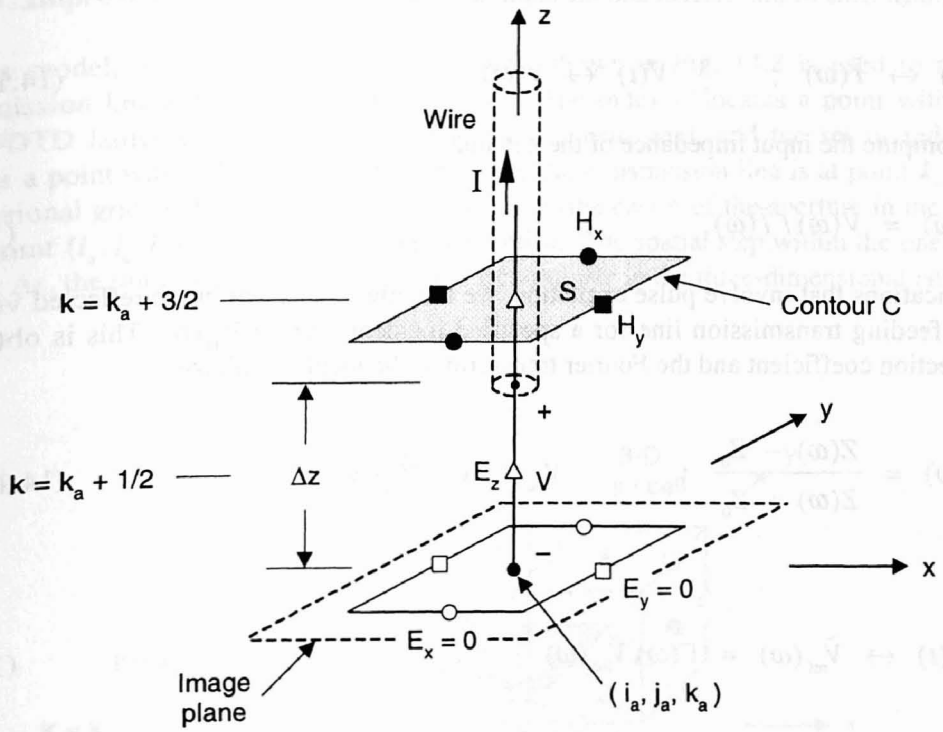


Fig. 14.7 Simple model for feeding the monopole antenna of Fig. 14.6.

Transmitting Antenna

For a transmitting antenna, E_z in the gap is related to the specified voltage V by

$$E_z\Big|^{n}_{i_a,\, j_a,\, k_a+1/2} = -\frac{V(n\Delta t)}{\Delta z} \tag{14.8}$$

The current in the wire at the feedpoint is obtained by applying Ampere's law [27] to the surface S with the bounding contour C centered on the wire at $(i_a, j_a, k_a+3/2)$:

$$\oint_C \boldsymbol{H}\cdot d\boldsymbol{L} = \iint_S \boldsymbol{J}\cdot d\boldsymbol{S} + \varepsilon_0 \iint_S \frac{\partial \boldsymbol{E}}{\partial t}\cdot d\boldsymbol{S} \tag{14.9}$$

This gives the current:

$$\begin{aligned} I\Big|^{n+1/2} &= \Delta x\left(H_x\Big|^{n+1/2}_{i_a,\, j_a-1/2,\, k_a+3/2} - H_x\Big|^{n+1/2}_{i_a,\, j_a+1/2,\, k_a+3/2}\right) \\ &+ \Delta y\left(H_y\Big|^{n+1/2}_{i_a+1/2,\, j_a,\, k_a+3/2} - H_y\Big|^{n+1/2}_{i_a-1/2,\, j_a,\, k_a+3/2}\right) \end{aligned} \tag{14.10}$$

The Fourier transforms of this current and the applied voltage

$$I(t) \leftrightarrow \breve{I}(\omega) \quad ; \qquad V(t) \leftrightarrow \breve{V}(\omega) \tag{14.11a, b}$$

are used to compute the input impedance of the antenna:

$$Z(\omega) = \breve{V}(\omega) / \breve{I}(\omega) \tag{14.12}$$

For applications that involve pulse excitation, we are often interested in the reflected voltage $V_{\text{ref}}(t)$ in the feeding transmission line for a specified incident voltage $V_{\text{inc}}(t)$. This is obtained from the reflection coefficient and the Fourier transform of the incident voltage:

$$\Gamma(\omega) = \frac{Z(\omega) - Z_0}{Z(\omega) + Z_0} \quad ; \qquad V_{\text{inc}}(t) \leftrightarrow \breve{V}_{\text{inc}}(\omega) \tag{14.13a, b}$$

as follows:

$$V_{\text{ref}}(t) \leftrightarrow \breve{V}_{\text{ref}}(\omega) = \Gamma(\omega)\, \breve{V}_{\text{inc}}(\omega) \tag{14.14}$$

Receiving Antenna

For a receiving antenna, the open-circuit voltage is obtained from E_z in the gap:

$$V_{\text{oc}}\Big|^n = -\Delta z\, E_z\Big|^n_{i_a,\, j_a,\, k_a+1/2} \tag{14.15}$$

The voltage into a section of transmission line matched at the far end is

$$V_{\text{rec}}(t) \leftrightarrow \breve{V}_{\text{rec}}(\omega) \tag{14.16}$$

where

$$\breve{V}_{\text{rec}}(\omega) = \left[\frac{Z_0}{Z(\omega) + Z_0}\right] \breve{V}_{\text{oc}}(\omega) \tag{14.17}$$

$$V_{\text{oc}}(t) \leftrightarrow \breve{V}_{\text{oc}}(\omega) \tag{14.18}$$

Due to the required Fourier and inverse Fourier transformation of several quantities, the above procedures are awkward to use when only time-domain values are of interest. Such data include the reflected voltage in the feedline of a transmitting antenna (14.14) and the voltage in the matched transmission line of a receiving antenna (14.16). Since FDTD is inherently a time-domain approach, it is advantageous, when possible, to carry out all computations in the time domain. This can be accomplished by using the improved, simple feed model described below.

14.3.3 Improved Simple Feed Model

In this model, the one-dimensional FDTD grid shown in Fig. 14.8 is used to represent the transmission line attached to the antenna [28]. The index k' locates a point within this grid. The FDTD lattice used for the antenna is three-dimensional, and the set of indexes (i, j, k) locates a point within this lattice. The top end of the transmission line is at point k'_{top} in the one-dimensional grid, and the feedpoint for the antenna (the center of the aperture in the coaxial line) is at point (i_a, j_a, k_a) in the three-dimensional lattice. The spatial step within the one-dimensional grid is Δz, the same as the spatial step for the z-coordinate in the three-dimensional lattice.

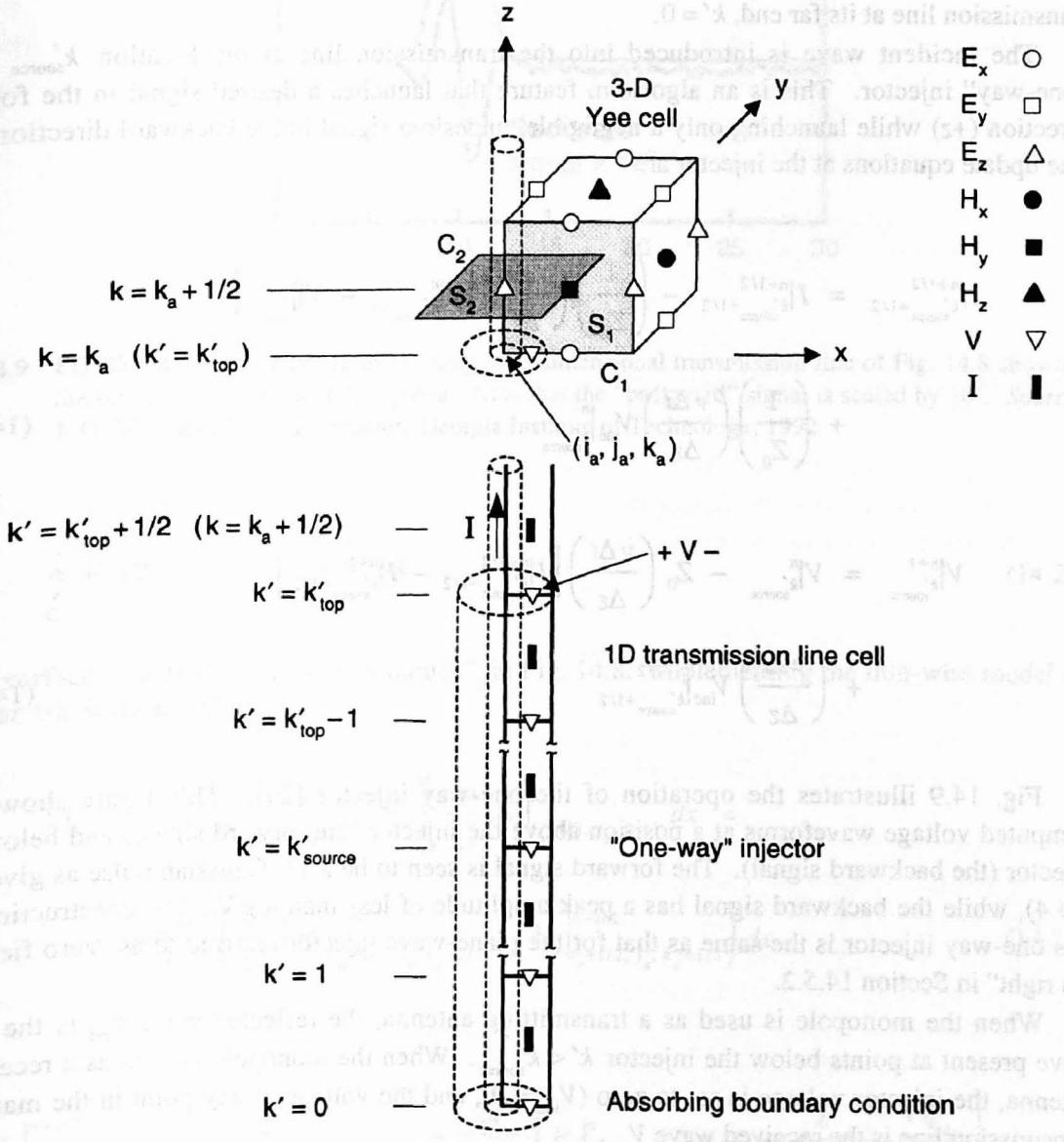


Fig. 14.8 Improved simple model for feeding the monopole antenna of Fig. 14.6 from a coaxial line, showing details of the one-dimensional FDTD grid for the transmission line and the three-dimensional FDTD lattice at the antenna feedpoint.

For this geometry, the update equations for the current and voltage in the transmission line are

$$I\big|_{k'+1/2}^{n+1/2} = I\big|_{k'+1/2}^{n-1/2} - \left(\frac{1}{Z_0}\right)\left(\frac{v\,\Delta t}{\Delta z}\right)\left[V\big|_{k'+1}^{n} - V\big|_{k'}^{n}\right] \tag{14.19a}$$

$$V\big|_{k'}^{n+1} = V\big|_{k'}^{n} - Z_0\left(\frac{v\,\Delta t}{\Delta z}\right)\left[I\big|_{k'+1/2}^{n+1/2} - I\big|_{k'-1/2}^{n+1/2}\right] \tag{14.19b}$$

where v is the phase velocity within the transmission line. An ABC is used to match the transmission line at its far end, $k' = 0$.

The incident wave is introduced into the transmission line at the location k'_{source} by a "one-way" injector. This is an algorithm feature that launches a desired signal in the forward direction ($+z$) while launching only a negligible, undesired signal in the backward direction ($-z$). The update equations at the injector are

$$I\big|_{k'_{\text{source}}+1/2}^{n+1/2} = I\big|_{k'_{\text{source}}+1/2}^{n-1/2} - \left(\frac{1}{Z_0}\right)\left(\frac{v\,\Delta t}{\Delta z}\right)\left[V\big|_{k'_{\text{source}}+1}^{n} - V\big|_{k'_{\text{source}}}^{n}\right] + \left(\frac{1}{Z_0}\right)\left(\frac{v\,\Delta t}{\Delta z}\right) V_{\text{inc}}\big|_{k'_{\text{source}}}^{n} \tag{14.20a}$$

$$V\big|_{k'_{\text{source}}}^{n+1} = V\big|_{k'_{\text{source}}}^{n} - Z_0\left(\frac{v\,\Delta t}{\Delta z}\right)\left[I\big|_{k'_{\text{source}}+1/2}^{n+1/2} - I\big|_{k'_{\text{source}}-1/2}^{n+1/2}\right] + \left(\frac{v\,\Delta t}{\Delta z}\right) V_{\text{inc}}\big|_{k'_{\text{source}}+1/2}^{n+1/2} \tag{14.20b}$$

Fig. 14.9 illustrates the operation of the one-way injector [29]. This figure shows the computed voltage waveforms at a position above the injector (the forward signal) and below the injector (the backward signal). The forward signal is seen to be a 1V Gaussian pulse as given by (14.4), while the backward signal has a peak amplitude of less than 4 μV. The construction for this one-way injector is the same as that for the plane-wave injector referred to as "zero field on the right" in Section 14.5.2.

When the monopole is used as a transmitting antenna, the reflected wave V_{ref} is the only wave present at points below the injector $k' < k'_{\text{source}}$. When the monopole is used as a receiving antenna, the injector voltage is set to zero ($V_{\text{inc}} = 0$), and the voltage at any point in the matched transmission line is the received wave V_{rec}.

In the FDTD model, the one-dimensional transmission-line grid and the three-dimensional antenna lattice are attached in the following manner. Location $k' = k'_{\text{top}}$ in the one-dimensional grid is assumed to be coincident with point (i_a, j_a, k_a) in the three-dimensional lattice. An update equation for H_y is then obtained by applying Faraday's law [27]

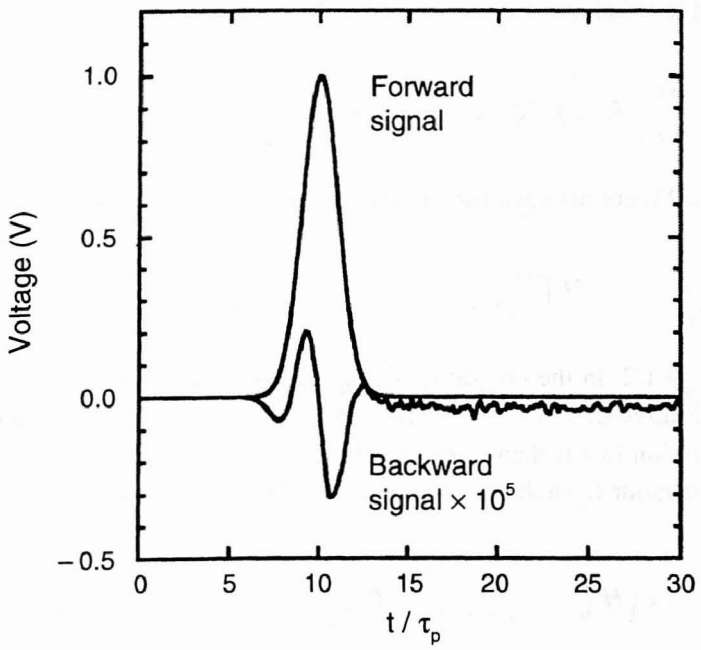


Fig. 14.9 FDTD-computed signals in the feeding, one-dimensional transmission line of Fig. 14.8 showing the operation of the one-way injector. Note that the "backward" signal is scaled by 10^5. *Source:* J. G. Maloney, Ph.D. dissertation, Georgia Institute of Technology, 1992.

$$\oint_{C_1} \boldsymbol{E} \cdot d\boldsymbol{L} = -\mu_0 \iint_{S_1} \frac{\partial \boldsymbol{H}}{\partial t} \cdot d\boldsymbol{S} \tag{14.21}$$

to the surface S_1 with the bounding contour C_1 in Fig. 14.8. Implementing the thin-wire model of Chapter 10, Section 10.5 yields

$$V\big|^n_{k'_{\text{top}}} + \Delta z\, E_z\big|^n_{i_a+1,\, j_a,\, k_a+1/2} + \int_{\Delta x}^{a} \frac{\Delta x}{2x} E_x\big|^n_{i_a+1/2,\, j_a,\, k_a+1}\, dx =$$

$$\frac{\mu_0 \Delta z}{\Delta t} \int_{a}^{\Delta x} \frac{\Delta x}{2x} \left(H_y\big|^{n+1/2}_{i_a+1/2,\, j_a,\, k_a+1/2} - H_y\big|^{n-1/2}_{i_a+1/2,\, j_a,\, k_a+1/2} \right) dx \tag{14.22}$$

or

$$H_y\big|^{n+1/2}_{i_a+1/2,\, j_a,\, k_a+1/2} = H_y\big|^{n-1/2}_{i_a+1/2,\, j_a,\, k_a+1/2} - \frac{\Delta t}{\mu_0 \Delta z} \cdot \left[E_x\big|^n_{i_a+1/2,\, j_a,\, k_a+1} - \frac{2}{\ln(\Delta x / a)} \cdot \frac{V\big|^n_{k'_{\text{top}}}}{\Delta x} \right]$$

$$+ \frac{\Delta t}{\mu_0 \Delta x} \cdot \frac{2}{\ln(\Delta x / a)} \cdot E_z\big|^n_{i_a+1,\, j_a,\, k_a+1/2} \tag{14.23}$$

Here, we assume that E_x and H_y are zero within the wire, and that close to the wire, they vary as $1/\rho$, where ρ is the radial distance from the center of the wire:

$$E_x(x) = \left(\frac{\Delta x}{2x}\right) E_x(\Delta x/2) \quad ; \qquad H_y(x) = \left(\frac{\Delta x}{2x}\right) H_y(\Delta x/2) \tag{14.24}$$

Updates similar to (14.23) are obtained for the other components of $\boldsymbol{H}$ that loop around the wire:

$$H_y\Big|^{n+1/2}_{i_a-1/2,\, j_a,\, k_a+1/2} \quad ; \quad H_x\Big|^{n+1/2}_{i_a,\, j_a+1/2,\, k_a+1/2} \quad ; \quad H_x\Big|^{n+1/2}_{i_a,\, j_a-1/2,\, k_a+1/2}$$

Location $k' = k'_{\text{top}} + 1/2$ in the one-dimensional grid is assumed to be coincident with point $(i_a, j_a, k_a+1/2)$ in the three-dimensional FDTD space lattice. The update equation for current $I|_{k'_{\text{top}}+1/2}$ in the transmission line is then calculated by applying Ampere's law (14.9) to the surface S_2 with the bounding contour C_2 in the three-dimensional lattice, as shown in Fig. 14.8:

$$\begin{aligned} I\Big|^{n+1/2}_{k'_{\text{top}}+1/2} &= \Delta x \left(H_x\Big|^{n+1/2}_{i_a,\, j_a-1/2,\, k_a+1/2} - H_x\Big|^{n+1/2}_{i_a,\, j_a+1/2,\, k_a+1/2} \right) \\ &+ \Delta y \left(H_y\Big|^{n+1/2}_{i_a+1/2,\, j_a,\, k_a+1/2} - H_y\Big|^{n+1/2}_{i_a-1/2,\, j_a,\, k_a+1/2} \right) \end{aligned} \tag{14.25}$$

Equations (14.23) and (14.25) couple the transmission line to the antenna and vice versa.

At other points along the monopole $k > k_a$, the subcell monopole radius is easily accounted for in the update equations for the H-field, using the thin-wire model [30]. For example, we use

$$\begin{aligned} H_y\Big|^{n+1/2}_{i_a+1/2,\, j_a,\, k+1/2} &= H_y\Big|^{n-1/2}_{i_a+1/2,\, j_a,\, k+1/2} - \frac{\Delta t}{\mu_0 \Delta z} \cdot \left(E_x\Big|^{n}_{i_a+1/2,\, j_a,\, k+1} - E_x\Big|^{n}_{i_a+1/2,\, j_a,\, k} \right) \\ &+ \frac{\Delta t}{\mu_0 \Delta x} \cdot \frac{2}{\ln(\Delta x/a)} \cdot E_z\Big|^{n}_{i_a+1,\, j_a,\, k+1/2} \end{aligned} \tag{14.26}$$

This equation is included to complete the description of the monopole antenna. It is not part of the feed model.

Fig. 14.10 compares FDTD results with measured data for the reflected voltage waveform in the transmission line feeding the transmitting monopole antenna of Fig. 14.6. Here, time is normalized as t/τ_a, where $\tau_a = h/c$ is the time for an electromagnetic wave propagating in free space to travel the length of the monopole. Other parameters are $a/h = 0.0304$ and $Z_0 = 50\Omega$, and the excitation is a 1V Gaussian pulse characterized by $\tau_p/\tau_a = 0.161$. Despite the complexity of the reflected waveform, the FDTD calculations and the measurements agree very well.

In the above discussion, the improved simple feed model is applied to a wire antenna connected to a coaxial transmission line (an unbalanced feed). This model is also adaptable to wire antennas connected to a two-wire transmission line (a balanced feed) [31, 32]. Additional discussion regarding FDTD modeling of antenna feeds is provided in [33–35].

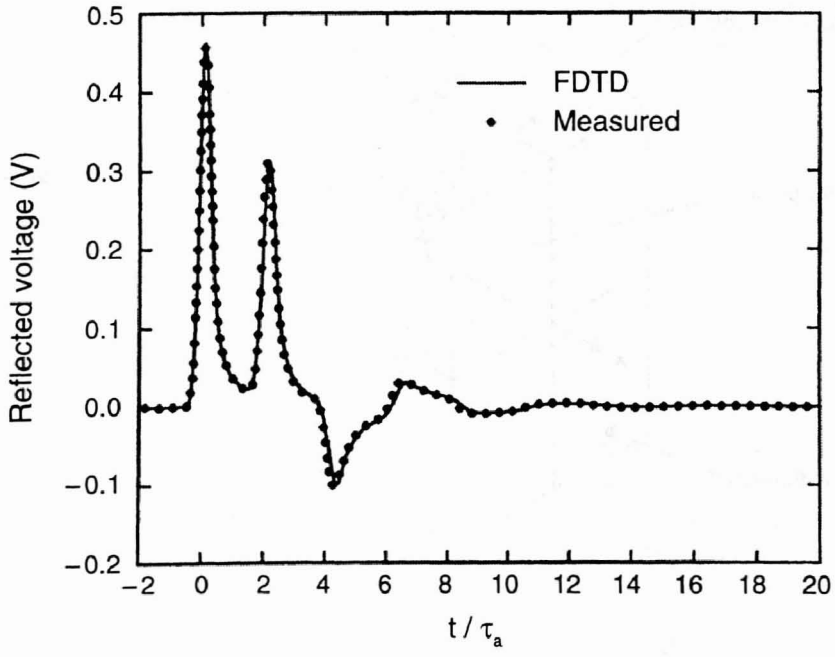


Fig. 14.10 Comparison of FDTD results with measured data for the reflected voltage in a coaxial transmission line feeding the cylindrical monopole antenna of Fig. 14.6. The improved simple feed is used in the FDTD model. *Source:* Maloney et al., *IEEE Trans. Antennas and Propagation*, 1994, pp. 289–292, © 1994 IEEE.

14.4 NEAR-TO-FAR-FIELD TRANSFORMATIONS

The FDTD method is inherently a near-field technique. It accurately computes the electromagnetic field within a computational volume surrounding the antenna as a function of time. Usually, it is impractical for the computational volume to extend from the antenna to the far-field zone. Therefore, an NTFF transformation is needed to compute far-field antenna parameters, such as radiation patterns and gain.

Since the standard FDTD method is based on an ordered, rectilinear gridding of the solution space, it is natural to consider a rectangular volume for the evaluation of an NTFF transformation. This volume is placed in the computational space so that its surface (the transformation surface) is located between the exterior of the antenna and the beginning of the ABC, as shown in Fig. 14.1(a). In practice, the rectangular faces of the transformation surface are chosen to coincide with the faces of individual Yee cells.

14.4.1 Use of Symmetry

The far-field transformation surface must enclose the whole antenna. When PEC and PMC image planes are used to reduce the computational volume, as illustrated in Fig. 14.2(b), image theory must be used to construct the field on the remainder of the transformation surface from the field on the portions of the surface inside the computational volume.

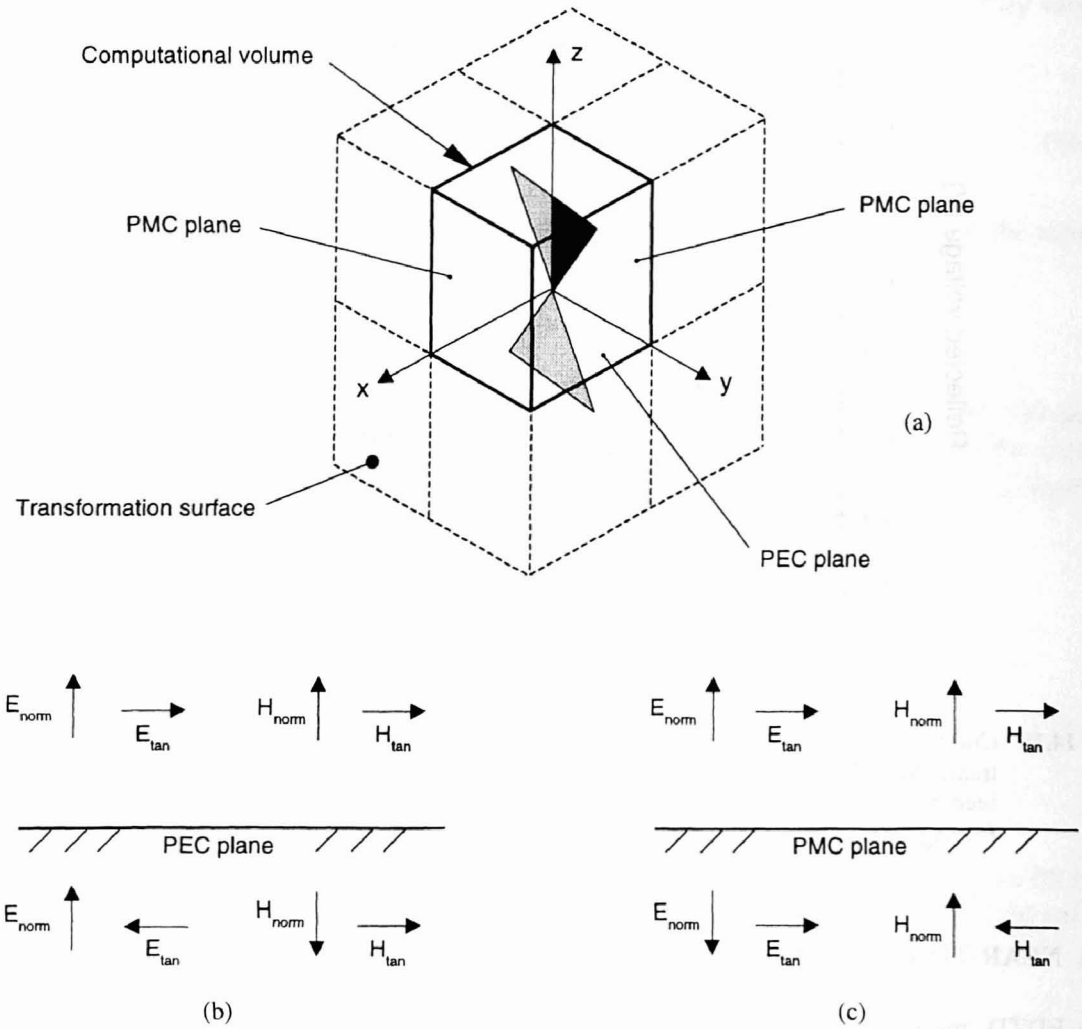


Fig. 14.11 Near-to-far-field transformation geometry: (a) transformation surface developed from the surface of a smaller computational volume that exploits the symmetry of the antenna; (b), (c) relationships for the field components at the PEC and PMC image planes.

Fig. 14.11(a) illustrates this case for a bow-tie transmitting antenna. Here, field components on the three faces inside the modeling volume are used, along with image theory, to compute the field on the rest of the transformation surface. The relationships between the field components at the PEC and PMC image planes are summarized in Figs. 14.11(b) and 14.11(c), respectively.

14.4.2 Time-Domain Near-to-Far-Field Transformation

Two types of near-to-far-field transformations are routinely used in FDTD antenna modeling: the *time-domain NTFF* (TD-NTFF) transformation, and the *frequency-domain NTFF* (FD-NTFF) transformation. The TD-NTFF transformation computes the transient radiation response in a specific far-field direction. Upon Fourier transformation, this yields the broadband gain of the antenna in that direction. This section briefly reviews the TD-NTFF transformation.

The TD-NTFF transformation is based upon the general integral relation of (14.1). From Chapter 8, Section 8.6, the theoretical basis of this transformation is given in (8.44) to (8.47) for purposes of efficient numerical implementation. These equations are repeated here for convenience. First, (8.44) and (8.45) define the pair of vector potentials

$$\boldsymbol{W}(\boldsymbol{r}, t) = \frac{1}{4\pi rc} \frac{\partial}{\partial t} \left[\iint_S \boldsymbol{J}_s \left(t - \frac{r - \boldsymbol{r}' \cdot \hat{\boldsymbol{r}}}{c} \right) dS' \right]$$

$$\boldsymbol{U}(\boldsymbol{r}, t) = \frac{1}{4\pi rc} \frac{\partial}{\partial t} \left[\iint_S \boldsymbol{M}_s \left(t - \frac{r - \boldsymbol{r}' \cdot \hat{\boldsymbol{r}}}{c} \right) dS' \right]$$

where $\boldsymbol{J}_s = \hat{\boldsymbol{n}} \times \boldsymbol{H}$ and $\boldsymbol{M}_s = -\hat{\boldsymbol{n}} \times \boldsymbol{E}$ are equivalent electric and magnetic currents on S, the rectangular-box virtual surface enclosing the antenna at which the TD-NTFF transformation is implemented, and $\hat{\boldsymbol{n}}$ is the unit outward normal vector to S. Next, (8.46) and (8.47) define the E-field radiated by these vector potentials, and hence the antenna contained within S:

$$E_\theta(\boldsymbol{r}, t) \equiv -\eta_0 W_\theta(\boldsymbol{r}, t) - U_\phi(\boldsymbol{r}, t) \; ; \quad E_\phi(\boldsymbol{r}, t) \equiv -\eta_0 W_\phi(\boldsymbol{r}, t) + U_\theta(\boldsymbol{r}, t)$$

Section 8.6 discusses an efficient concurrent-processing approach to calculate $\boldsymbol{W}$ and $\boldsymbol{U}$ "on-the-fly" in step with the FDTD simulation. Here, a total of 24 two-dimensional arrays are required to store the Cartesian components of $\boldsymbol{J}_S$ and $\boldsymbol{M}_S$ on S at the current time-step. In addition, 6 one-dimensional arrays are required to store the time variation of the Cartesian components of $\boldsymbol{W}$ and $\boldsymbol{U}$ for each (θ, ϕ) in the radiation zone at which a field-versus-time waveform is needed. Each array of this type contains $NMAX + 2s$ storage locations, where $NMAX$ is the number of time-steps in the FDTD run, s is the maximum number of space lattice cells between any two points on S, and $2c\Delta t = \Delta$. This allows accounting for delayed contributions arriving from all parts of S at the conclusion of time-stepping, regardless of the (θ, ϕ) orientation of the far-field observation point $\boldsymbol{r}$. We note that the one- and two-dimensional storage arrays required for the TD-NTFF transformation represent, in total, a dimensionally reduced computer storage burden relative to the three-dimensional E- and H-field arrays required for the normal FDTD operations.

Upon completion of the filling of the storage arrays for $\boldsymbol{W}$ and $\boldsymbol{U}$, the rectangular-to-spherical vector-component conversion of (8.61) and (8.62) is performed. This yields time waveforms of the spherical vector components of $\boldsymbol{W}$ and $\boldsymbol{U}$, which are inserted into (8.46) and (8.47) to directly obtain the desired time waveforms $E_\theta(\boldsymbol{r}, t)$ and $E_\phi(\boldsymbol{r}, t)$ at observation point $\boldsymbol{r}$ in the far-field zone. From (8.44) and (8.45), we note that the amplitudes of these waveforms diminish as $1/r$, where r is the distance from the origin. Therefore, it is possible to obtain a normalized far-field response that is independent of the distance from the origin by simply multiplying (8.46) and (8.47) by r.

This procedure can be used to calculate $E_\theta(\boldsymbol{r}, t)$ and $E_\phi(\boldsymbol{r}, t)$ in arbitrary (θ, ϕ) directions. Straightforward Fourier transformation then yields $\breve{E}_\theta(\boldsymbol{r}, \omega)$ and $\breve{E}_\phi(\boldsymbol{r}, \omega)$, respectively. This permits the antenna gain in specific (θ, ϕ) directions to be calculated over a wide bandwidth in a single FDTD run. Since the radiated field contains both $\hat{\theta}$ and $\hat{\phi}$ components, it is convenient to define separate gains for each component:

$$G_\theta(\omega) = 4\pi \left| r \breve{E}_\theta(\boldsymbol{r}, \omega) \right|^2 / 2\eta_0 P_{\text{ant}}(\omega) \tag{14.27a}$$

$$G_\phi(\omega) = 4\pi \left| r \breve{E}_\phi(\boldsymbol{r}, \omega) \right|^2 / 2\eta_0 P_{\text{ant}}(\omega) \tag{14.27b}$$

where $P_{\text{ant}}(\omega)$ is the net power (incident power minus reflected power in a single mode) that enters the transmission line or waveguide feeding the antenna. $P_{\text{ant}}(\omega)$ is computed from

$$P_{\text{ant}}(\omega) = \left[1 - \left| \Gamma(\omega) \right|^2 \right] \left[\left| \breve{V}_{\text{inc}}(\omega) \right|^2 / 2Z_0 \right] \tag{14.28}$$

where we define the Fourier transform pairs $\breve{V}_{\text{inc}}(\omega) \leftrightarrow V_{\text{inc}}(t)$ and $\breve{V}_{\text{refl}}(\omega) \leftrightarrow V_{\text{refl}}(t)$, and the reflection coefficient $\Gamma(\omega) = \breve{V}_{\text{refl}}(\omega) / \breve{V}_{\text{inc}}(\omega)$.

The TD-NTFF transformation is an efficient approach for computing the wideband, frequency-domain gain in a number of discrete (θ, ϕ) directions. However, this transformation can be inefficient for computing frequency-domain antenna patterns because of the fine angular spacing often required. The more efficient approach for computing frequency-domain antenna patterns is the FD-NTFF transformation, which is briefly reviewed next.

14.4.3 Frequency-Domain Near-to-Far-Field Transformation

From Chapter 8, Section 8.5, the theoretical basis of the FD-NTFF transformation is given by (8.33), (8.34), and (8.29) for purposes of efficient numerical implementation. These equations are repeated here for convenience. First, (8.33) and (8.34) define spherical-coordinate components of the vector potentials

$$\breve{N}_\theta = \iint_S \left(\breve{J}_x \cos\theta \cos\phi + \breve{J}_y \cos\theta \sin\phi - \breve{J}_z \sin\theta \right) e^{+jkr' \cos\psi} \, ds'$$

$$\breve{N}_\phi = \iint_S \left(-\breve{J}_x \sin\phi + \breve{J}_y \cos\phi \right) e^{+jkr' \cos\psi} \, ds'$$

$$\breve{L}_\theta = \iint_S \left(\breve{M}_x \cos\theta \cos\phi + \breve{M}_y \cos\theta \sin\phi - \breve{M}_z \sin\theta \right) e^{+jkr' \cos\psi} \, ds'$$

$$\breve{L}_\phi = \iint_S \left(-\breve{M}_x \sin\phi + \breve{M}_y \cos\phi \right) e^{+jkr' \cos\psi} \, ds'$$

where $\breve{J}_x$, $\breve{J}_y$, and $\breve{J}_z$ are Cartesian components of $\breve{\boldsymbol{J}}_S = \hat{\boldsymbol{n}} \times \breve{\boldsymbol{H}}$, the equivalent phasor electric current on the rectangular-box NTFF-transformation virtual surface S; $\breve{M}_x$, $\breve{M}_y$, and $\breve{M}_z$ are Cartesian components of $\breve{\boldsymbol{M}}_S = -\hat{\boldsymbol{n}} \times \breve{\boldsymbol{E}}$, the equivalent phasor magnetic current on S; $\hat{\boldsymbol{n}}$ is the unit outward normal to S; and ψ is the angle between $\boldsymbol{r}$ and $\boldsymbol{r}'$. Next, (8.29b) and (8.29c) define the phasor E-field radiated by these vector potentials, and hence the antenna contained within S:

$$\breve{E}_\theta = -\frac{jke^{-jkr}}{4\pi r}\left(\breve{L}_\phi + \eta_0 \breve{N}_\theta\right) \; ; \quad \breve{E}_\phi = +\frac{jke^{-jkr}}{4\pi r}\left(\breve{L}_\theta - \eta_0 \breve{N}_\phi\right)$$

Section 8.3 discusses an efficient discrete Fourier transform method to calculate $\breve{\boldsymbol{J}}_s$ and $\breve{\boldsymbol{M}}_s$ "on-the-fly" in step with the FDTD simulation. This requires a total of 24 complex-valued, two-dimensional arrays to store the Cartesian components of the equivalent currents at each frequency of interest. With these data in hand at the end of the FDTD run, $\breve{N}_\theta$, $\breve{N}_\phi$, $\breve{L}_\theta$, and $\breve{L}_\phi$ are computed from (8.33) and (8.34) in a postprocessing step at the desired (θ, ϕ) orientations in the far field at each frequency of interest. Then, $\breve{E}_\theta$ and $\breve{E}_\phi$ are obtained from (8.29). We see that the complete radiation pattern of the antenna at selected frequencies can be obtained in this manner during the postprocessing of the $\breve{\boldsymbol{J}}_s$ and $\breve{\boldsymbol{M}}_s$ data resulting from a single FDTD run.

In practice, both TD-NTFF and FD-NTFF transformations are commonly used in FDTD models of transmitting antennas. The selection of one NTFF technique or the other is usually determined by whether broadband gain characteristics are desired at just a few angles of interest (TD-NTFF preferred), or whether wide-angle gain patterns are desired at just a few frequencies of interest (FD-NTFF preferred).

14.5 PLANE-WAVE SOURCE

FDTD models of receiving antennas use a plane-wave source condition to create an incident electromagnetic wave. Here, the plane wave is produced by electric and magnetic current sources placed tangentially on a virtual surface surrounding the antenna. As shown in Fig. 14.1(b), the plane-wave source surface is located between the ABC and the antenna.

Ideally, the tangential electric and magnetic current sources are placed on the same virtual surface. However, in the standard FDTD method, these sources cannot be coplanar because of the spatially staggered nature of the Yee cell. Hence, a straightforward application of (14.3) can lead to nonideal behavior. That is, the field inside the computation volume may only approximate the desired plane wave, and the field outside may be small, but not zero. This section discusses the subtleties involved in the plane-wave source condition due to the nature of the Yee algorithm.

14.5.1 Effect of an Incremental Displacement of the Surface Currents

Consider the $-\hat{\boldsymbol{x}}$-directed plane wave $[\boldsymbol{E}_{\text{inc}} = f(t + x/c)\boldsymbol{E}_0, \; \boldsymbol{H}_{\text{inc}} = -\hat{\boldsymbol{x}} \times \boldsymbol{E}_{\text{inc}}/\eta_0]$ generated by the planar electric and magnetic surface currents defined by

$$\boldsymbol{J}_s = f(t)\, \hat{\boldsymbol{x}} \times \hat{\boldsymbol{x}} \times \boldsymbol{E}_0 / \eta_0 \; ; \qquad \boldsymbol{M}_s = f(t)\, \hat{\boldsymbol{x}} \times \boldsymbol{E}_0 \tag{14.29a, b}$$

Fig. 14.12(a) illustrates the ideal case where the source currents are collocated at the same surface. Here, the superposition of the radiation from these currents generates the desired plane wave propagating to the left of the source plane, and zero field to the right of the source plane.

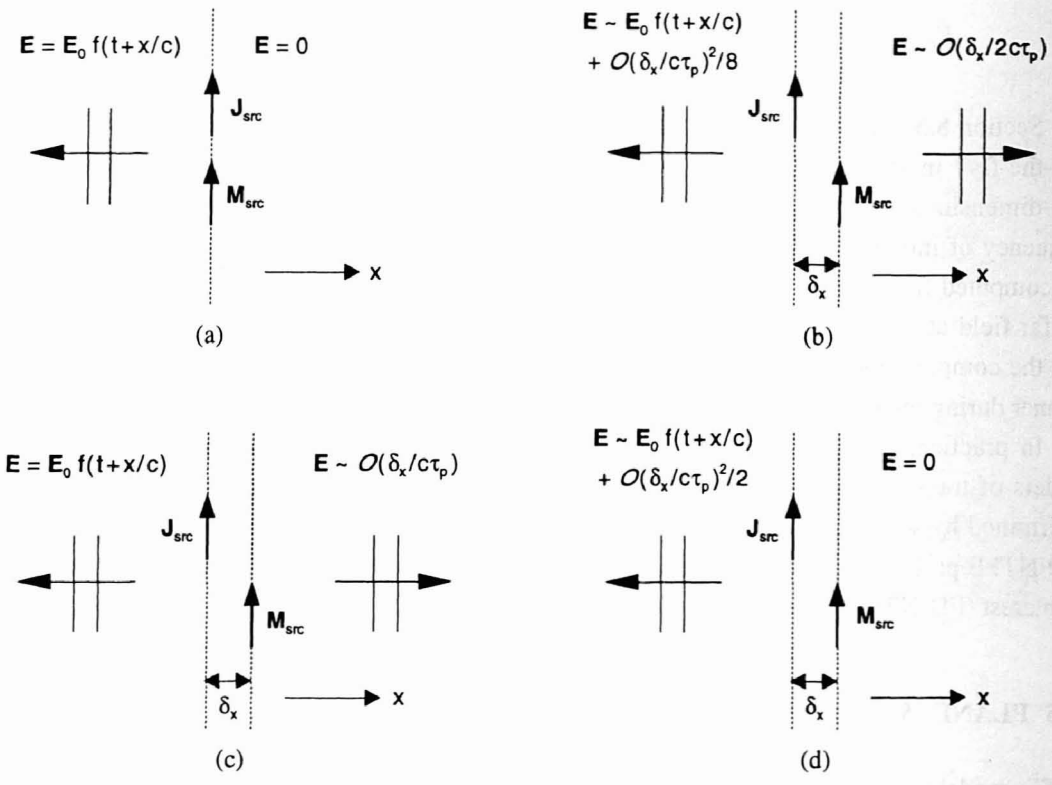


Fig. 14.12 Fields produced to the right and left of plane sheets of uniform electric and magnetic current: (a) ideal case of collocated currents; (b) currents displaced by δ_x; (c) time-shifted currents generating exact field on the left; (d) time-shifted currents generating zero field on the right.

Fig. 14.12(b) illustrates the nonideal case corresponding to the Yee space lattice where the surface currents are displaced by the small distance δ_x. Here, the net field to the left and right of the source planes is given by, respectively:

$$\boldsymbol{E}_{\text{left}} \approx f\left(t - \delta_x/2c + x/c\right)\boldsymbol{E}_0 + \frac{1}{8}\left(\frac{\delta_x}{c\tau_p}\right)^2\left[\tau_p^2 f''(t - \delta_x/2c + x/c)\right]\boldsymbol{E}_0 + O\left(\frac{\delta_x}{2c\tau_p}\right)^4 \tag{14.30a}$$

$$\boldsymbol{E}_{\text{right}} \approx -\left(\frac{\delta_x}{2c\tau_p}\right)\left[\tau_p f'(t + \delta_x/2c - x/c)\right]\boldsymbol{E}_0 + O\left(\frac{\delta_x}{2c\tau_p}\right)^3 \tag{14.30b}$$

where τ_p is a characteristic time associated with $f(t)$, such as τ_p for the Gaussian pulse of (14.4); and $c\tau_p >> \delta x$. We see that the field radiated to the left of the source is no longer the desired field; it is the desired signal with a slight time shift and an additional small replica of the second derivative of the source signal. Further, the field to the right of the source is no longer zero; it is now a small replica of the derivative of the source signal.

14.5.2 Effect of an Incremental Time Shift

A modification can be introduced that improves the situation of the noncollocated source currents in Fig. 14.12(b). Namely, the source currents are time shifted to generate either the desired signal propagating to the left of the source planes or zero signal on the right, but not both conditions simultaneously. Fig. 14.12(c) illustrates the case of time shifting of the source currents to generate the exact field on the left. Here, the currents are given by

$$\boldsymbol{J}_s = f(t)\,\hat{\boldsymbol{x}} \times \hat{\boldsymbol{x}} \times \boldsymbol{E}_0/\eta_0 \;; \quad \boldsymbol{M}_s = f(t + \delta_x/c)\,\hat{\boldsymbol{x}} \times \boldsymbol{E}_0 \tag{14.31a, b}$$

with the resulting radiated fields

$$\boldsymbol{E}_{\text{left}} = f(t + x/c)\,\boldsymbol{E}_0 \tag{14.32a}$$

$$\boldsymbol{E}_{\text{right}} \approx \left(\frac{\delta_x}{c\tau_p}\right)\left[\tau_p f'(t + \delta_x/c - x/c)\right]\boldsymbol{E}_0 + O\left(\frac{\delta_x}{c\tau_p}\right)^3 \tag{14.32b}$$

For this case, the field propagating to the left of the source planes is the desired field, while the field appearing to the right of the source planes is again a small replica of the derivative of the source signal. While small, the amplitude of the signal propagating to the right is doubled relative to that in (14.30b); that is, all of the error is shifted onto the unwanted signal propagating to the right.

Fig. 14.12(d) illustrates the case of time shifting of the source currents for zero field on the right. Here, the currents are given by

$$\boldsymbol{J}_s = f(t + \delta_x/c)\,\hat{\boldsymbol{x}} \times \hat{\boldsymbol{x}} \times \boldsymbol{E}_0/\eta_0 \;; \quad \boldsymbol{M}_s = f(t)\,\hat{\boldsymbol{x}} \times \boldsymbol{E}_0 \tag{14.33a, b}$$

with the resulting radiated fields

$$\boldsymbol{E}_{\text{left}} \approx f(t + x/c)\,\boldsymbol{E}_0 + \frac{1}{2}\left(\frac{\delta_x}{c\tau_p}\right)^2\left[\tau_p^2 f''(t + x/c)\right]\boldsymbol{E}_0 + O\left(\frac{\delta_x}{c\tau_p}\right)^4 \tag{14.34a}$$

$$\boldsymbol{E}_{\text{right}} = 0 \tag{14.34b}$$

For this case, the field to the right of the source planes is identically zero, while the field to the left of the source planes is the desired signal plus an additional small replica of the second derivative of the source signal. All the error is shifted onto the signal propagating to the left of the source planes. We see that the amplitude of this error is increased by a factor of 4:1 over that in (14.30a).

In practice, the external field to the right of the source planes in Fig. 14.12(d) is never exactly zero because of the finite numerical dispersion inherent in the Yee algorithm. This introduces a small error that negates the perfect cancellation in the external region.

14.5.3 Relation to Total-Field / Scattered-Field Lattice Zoning

As discussed in detail in Chapter 5, Sections 5.6 to 5.9, the traditional plane-wave source condition for FDTD is based on *total-field / scattered-field* (TF/SF) zoning of the space lattice. In fact, it can be shown that the TF/SF technique is equivalent to the zero-external-field plane-wave source case of Fig. 14.12(d). We note that this is the preferred situation for computation of the scattered field, for example, when calculating radar cross section. However, for modeling receiving antennas, it appears that the exact-signal case of Fig. 14.12(c) is more appropriate.

In practice, one usually constructs a single FDTD code suitable for a wide range of modeling problems involving scattering, antenna radiation, and antenna reception. Therefore, it makes sense to use only one form of the plane-wave source. The TF/SF technique, equivalent to the zero-external-field configuration of Fig. 14.12(d), is generally used for this purpose. From (14.34a), we note that this causes a small error of $O(\delta_x / c\tau_p)^2$ in the generated plane-wave field.

14.6 CASE STUDY I: THE STANDARD-GAIN HORN

Now that the elements of the FDTD analysis of antennas have been discussed, they will be applied to a particular antenna. The purpose of this example is twofold: (1) to show how the elements are combined in a practical situation, and (2) to show the high level of accuracy that can be expected from the method when properly applied. The antenna selected for this example is the Flann Microwave Instruments Ltd. Model 1624-20, an X-band (8.2 to 12.4 GHz) metal pyramidal horn fed by rectangular WR-90 waveguide. Antennas like this are used in many microwave applications and may serve as gain standards (standard-gain horns).

Fig. 14.13 shows the FDTD model for this horn antenna. The small inset drawings at the left of the figure define the lengths and angles that characterize the original antenna geometry. The numerical values of these parameters are: $a = 109.5$ mm, $b = 78.5$ mm, $D = 228.4$ mm, $l_w = 50.8$, $\alpha = 10.74°$, and $\beta = 8.508°$.

The FDTD model uses a uniform space lattice composed of $519 \times 116 \times 183$ cubic Yee cells having $\Delta x = \Delta y = \Delta z = \Delta = 0.635$ mm ($\cong \lambda_0/38$ at 12.4 GHz). The even symmetry of the horn geometry is used to define the x-z plane at $y = 0$ as a PMC. Looking closely at the perspective view of the model in Fig. 14.13, we see the faces of the individual cells that model the horn. In fact, the cells are drawn twice their actual size for clarity. This fine spatial resolution permits direct modeling of the 1.35-mm wall thickness of the metallic components, assumed to be PECs. A 10Δ-thick PML ABC is used on all sides of the FDTD lattice. This is spaced 20Δ from the closest surface of the antenna, except directly in front of the aperture, where the spacing is 40Δ.

The antenna is assumed to be fed from a coaxial transmission line having the characteristic impedance $Z_0 = 50\Omega$. The center conductor of the coaxial line (the probe) has a radius $a_{\text{probe}} \cong 0.2$ mm and extends into the waveguide a distance $h_{\text{probe}} = 6.35$ mm. The probe is centered laterally in the waveguide, and is spaced at a distance $l_{\text{probe}} = 5.08$ mm from the waveguide's back wall to minimize reflection at the coaxial line / waveguide transition at midband, 10 GHz. In the FDTD analysis, the improved simple feed model of Section 14.3.3 is used for the transmission line and probe. For the transmitting case, the antenna is excited with the differentiated Gaussian pulse of (14.6) with $\tau_p = 1.59 \times 10^{-11}$ sec.

Fig. 14.14 compares FDTD calculations and measurements [36, 37] of the far-field E-plane and H-plane radiation patterns of the pyramidal horn antenna of Fig. 14.13 at 10 GHz. The FDTD results are in excellent agreement with the measurements.

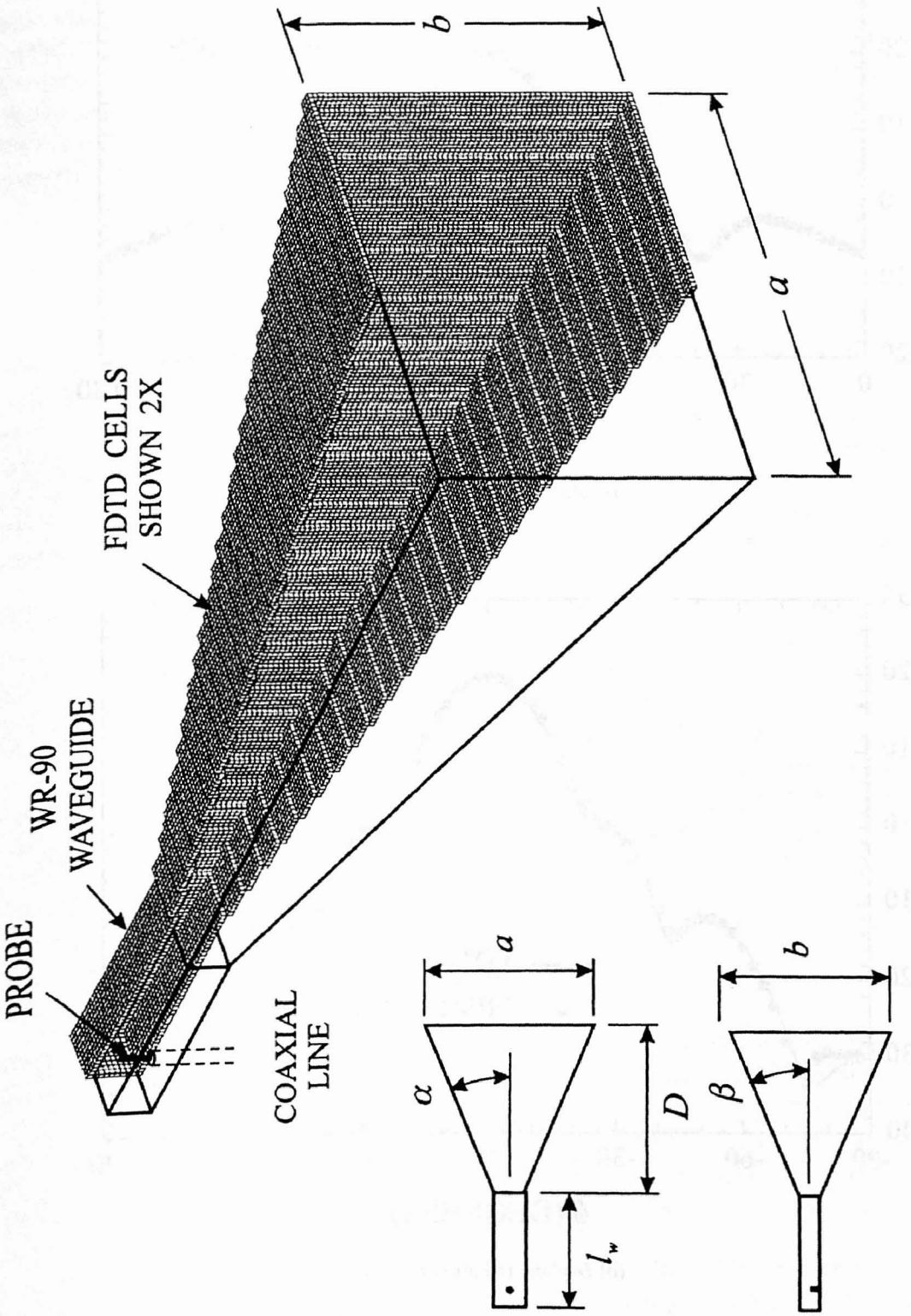
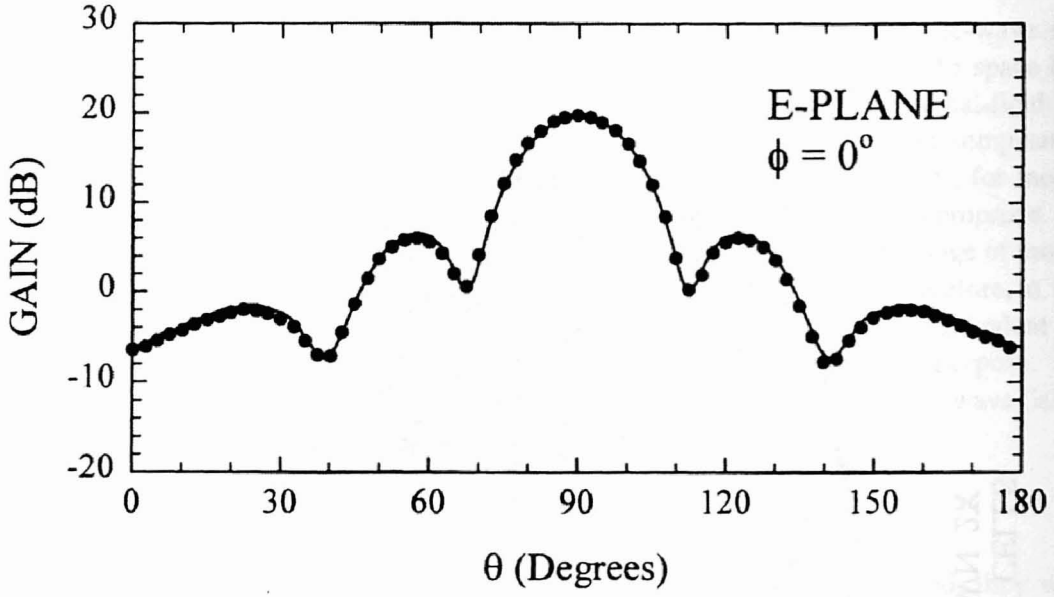
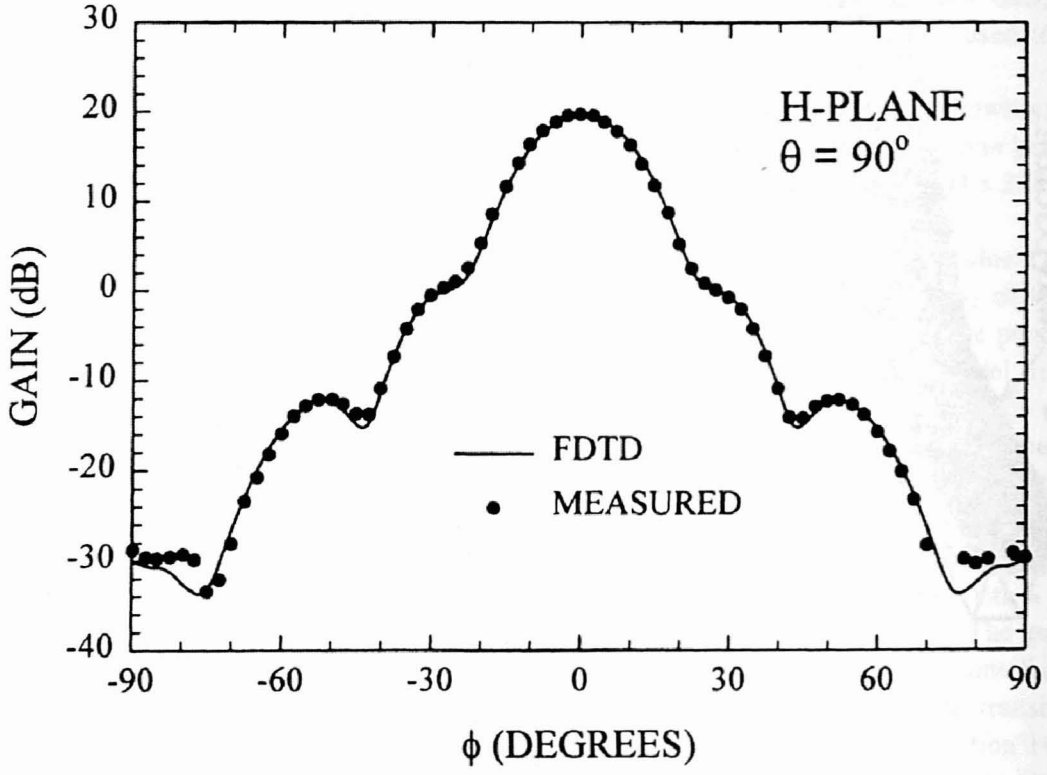


Fig. 14.13 Details of the geometry for the FDTD simulation of the Flann Model 1624-20 pyramidal horn antenna.



(a) *E*-plane radiation pattern.



(b) *H*-plane radiation pattern.

Fig. 14.14 Comparison of FDTD-calculated and measured far-field patterns at 10 GHz for the Flann pyramidal horn antenna of Fig. 14.13.

Fig. 14.15 compares FDTD calculations and measurements of the boresight properties of the horn antenna of Fig. 14.13. The first comparison, shown in Fig. 14.15(a), is for the boresight gain versus frequency. There is again excellent agreement of the FDTD model results and the test data. The second comparison, shown in Fig. 14.15(b), is for the boresight effective area versus frequency. Here, the FDTD analysis is performed for the horn as a receiving antenna illuminated by a z-polarized plane wave incident in the boresight direction. The time waveform of the impinging signal is the same differentiated Gaussian pulse used for the transmitting case. The effective area for the horn is derived from the FDTD-computed received voltage in the coaxial transmission line as

$$A_e(\omega) = \frac{\eta_0 \left| Z(\omega)/Z_0 + 1 \right|^2 \left| \breve{V}_{\text{rec}}(\omega) \right|^2}{4\,R(\omega) \left| \breve{E}_{\text{inc}}(\omega) \right|^2} \tag{14.35}$$

where $Z(\omega) = R(\omega) + jX(\omega)$ is the input impedance of the transmitting antenna. Fig. 14.15(b) compares the FDTD results for A_e obtained using (14.35), with the effective area $A_e = \lambda_0^2 G/4\pi$ derived from the measured gain G. Note that these results are shown in decibels relative to one square meter. Again, the agreement is excellent. Overall, these data show that, when properly used, the FDTD analysis produces equally good results for the transmitting and receiving cases. Further, Fig. 14.15 demonstrates reciprocity for the FDTD results for transmitting and receiving antennas in the frequency domain. A demonstration of reciprocity for the time domain is in [38].

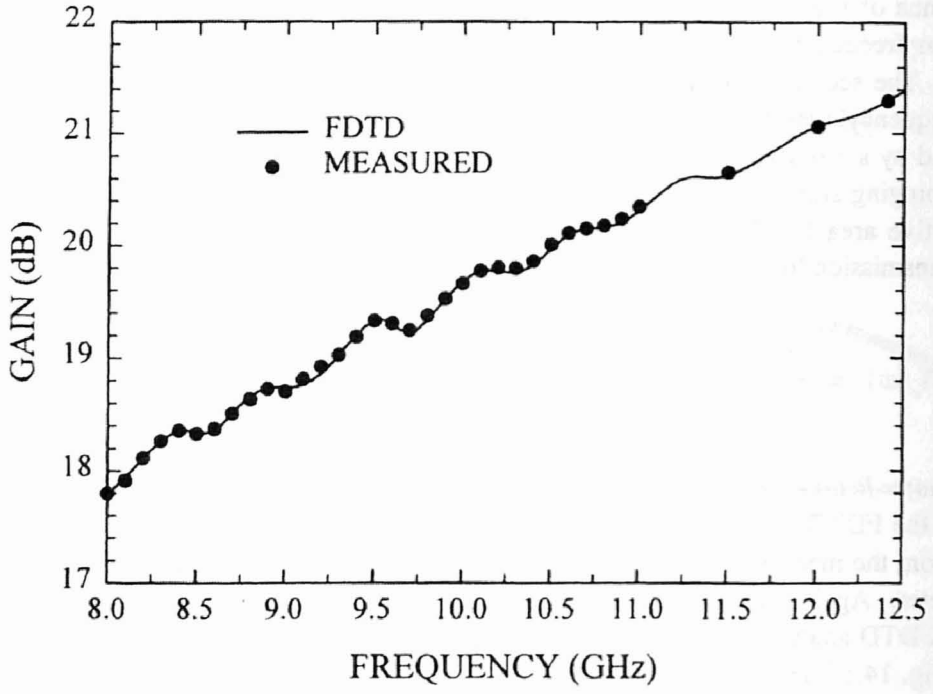
The FDTD method inherently provides information about the dynamics of the electromagnetic field within the computational volume over the period of the simulation. Only a small fraction of this information is used when investigating conventional antenna-performance metrics such as radiation pattern, gain, and effective area. Mapping or visualizing the computed space-time evolution of the electromagnetic field can provide useful additional information, helping to improve our understanding of the radiation process for an antenna.

Fig. 14.16 illustrates how space-time visualization is used to help develop an understanding of how an antenna radiates. Here, the magnitude of the E-field on the vertical symmetry plane of the horn antenna of Fig. 14.13 is plotted on a grayscale for two different times. The incident voltage in the coaxial line feeding the horn is a sinusoid that is amplitude modulated by a Gaussian pulse:

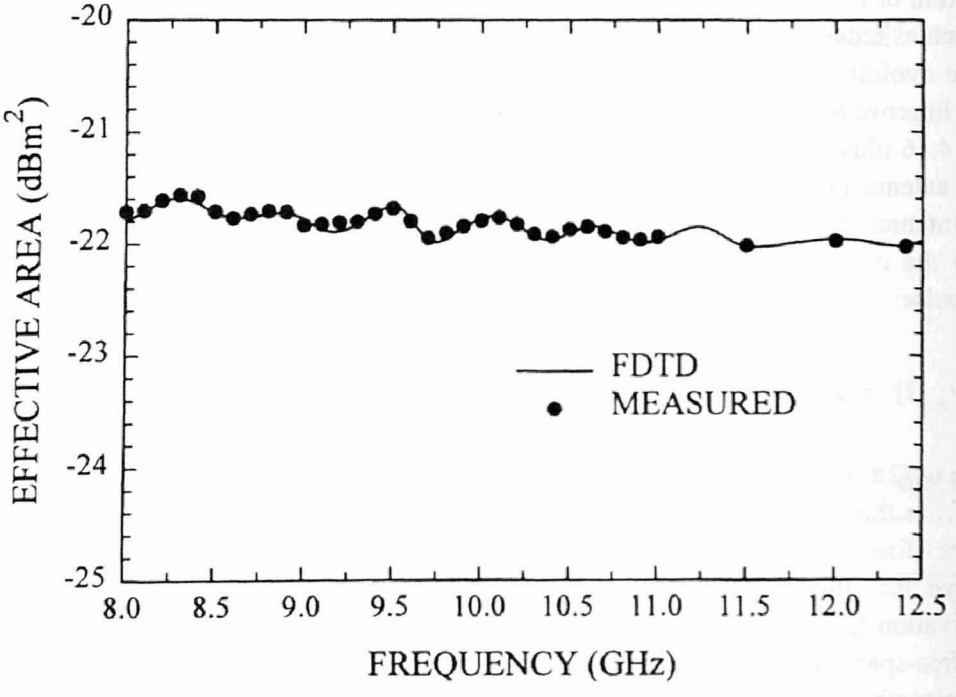
$$V_{\text{inc}}(t) = V_0 \exp\left[-(t/\tau_p)^2/2 \right] \sin(\omega_0 t) \tag{14.36}$$

where $f_0 = \omega_0/2\pi = 10.0$ GHz and $\tau_p = 7.96 \times 10^{-11}$ sec. In Fig. 14.16(a), the pulse has entered the horn from the waveguide but has not yet reached the aperture. The spacing between the white lines (lines of null field) roughly corresponds to one-half of the guide wavelength. We observe that this spacing decreases between the throat of the horn and the aperture. This observation is reasonable in that the wavelength in the rectangular waveguide is about 1.3 times the free-space wavelength, whereas at the aperture of the horn it is much closer to the free-space wavelength.

In Fig. 14.16(b), the pulse has reached the aperture. The white lines in the horn near the aperture are distorted; there is a small segment that is concave to the right. This is caused by the reflection from the aperture that is traveling back toward the throat of the horn. Directly in front of the aperture, the radiated wave is roughly planar. It becomes spherical at greater distances.

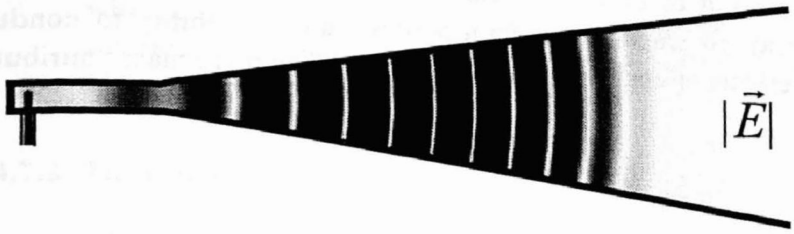


(a) Boresight gain versus frequency.

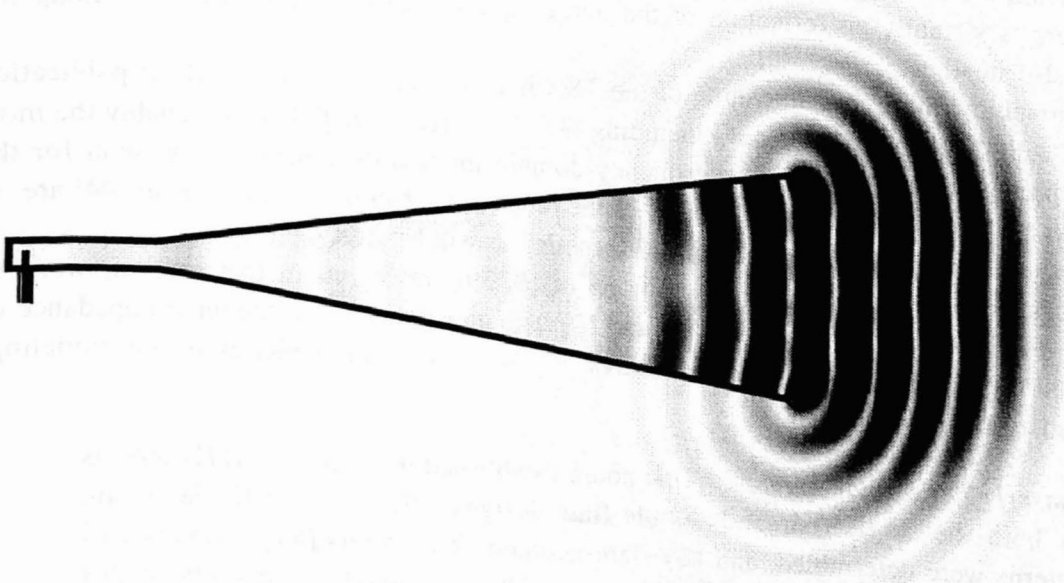


(b) Boresight effective area versus frequency.

Fig. 14.15 Comparison of FDTD-calculated and measured boresight properties versus frequency for the Flann pyramidal horn antenna of Fig. 14.13.



(a) The pulse has entered the horn from the waveguide.



(b) The pulse has reached the aperture of the horn.

Fig. 14.16 Grayscale visualization showing the magnitude of the FDTD-calculated electric field on the vertical symmetry plane of the Flann pyramidal horn antenna of Fig. 14.13 at two early times.

14.7 CASE STUDY II: THE VIVALDI SLOTLINE ARRAY

This section focuses on the FDTD analysis of Vivaldi slotline antennas including the single and double element, the crossed-pair subarray element, and linear arrays of crossed-pair elements [9, 39]. It reveals a number of important features of FDTD modeling of antennas and antenna arrays: (1) calculation of radiation patterns; (2) calculation of driving-point impedance, especially as a function of array phasing (*active impedance*); and (3) ability to conduct sensitivity studies probing the effects of engineering tolerances upon key performance attributes such as cross polarization.

14.7.1 Background

Slotline antennas are known to be traveling-wave antennas with the capability of producing broadband endfire radiation. A recent implementation of slotline antennas was introduced in [40]. In this structure, a microstrip slotline is flared outward to some width at which the desired radiation occurs. This is referred to as a *tapered slot antenna* (TSA). It is this type of slot antenna that is under consideration in this section.

Derived from the slotline, the TSA is a traveling-wave antenna; specifically, a surface-wave antenna. An advantage of this antenna over resonant antennas is its ability to work over a large bandwidth. On a fundamental level, TSAs work because the fields confined in a slotline can decouple from the slot edges and radiate as slot width is increased. Since the structure is open-ended, there is virtually no reflection of the outgoing wave, and current reflections along the edges are also small.

Analytical and numerical modeling of the TSA had not been extensive until the publication of [9, 39], with the primary references including [41–44]. Reference [44] was arguably the most advanced previous model, using the frequency-domain method of moments to account for the finite width of the antenna and the presence of the dielectric substrate. The results of [44] are, in fact, used in one of the validations of the FDTD model, as will be seen.

The primary goals of the study reported in [9, 39], reviewed in this section, were to investigate FDTD modeling of the E-plane and H-plane gain patterns and the input impedance of stripline-fed Vivaldi TSA elements and arrays. There were several phases to the modeling, summarized as follows.

- *Phase 1: The Planar Element.* This phase established the baseline FDTD models for both the single-flair and double-flair designs. The E- and H-plane gain patterns were determined, and key flare-geometry parameters (length and rate of the taper) were examined. To validate the FDTD method in the study of this general antenna type, the *linear tapered slot antenna* (LTSA) of [44] was also modeled.

- *Phase 2: The Quad Pair.* This phase developed and validated an FDTD model of the quad element. The quad element consisted of two perpendicular planar double flares constructed with collocation slots to allow for an interleaved mechanical assembly. Such a design avoided shorting either the horizontal or vertical ground planes upon assembly.

- *Phase 3: The Linear Phased Array.* The final phase developed and validated an FDTD model of a linear phased array comprised of the quad elements studied in Phase 2. The array was capable of beam steering and selectable polarization in either primary plane, and exhibited the broadband behavior characteristic of the Vivaldi TSA. Radiation patterns were computed for beam-steer angles of 0°, 20°, 45°, and 60° at selected frequencies across the bandwidth, and then compared with measurements. Calculations were made of the driving-point impedance as a function of the beam-steer angle and the operating frequency, to address the issue of the array active impedance.

14.7.2 The Planar Element

Fig. 14.17 illustrates the geometry of the baseline planar Vivaldi element. This element was comprised of two 0.020-in-thick, 1 oz/ft^2-copper-clad circuit boards forming a sandwich structure around a 0.032-in-wide copper stripline. Each of the outer copper planes was etched to have a flared slot extending from a 0.130-in-diameter matching hole. The stripline passed under (over) each slot at the point where the slot connected to its matching hole; the stripline was then terminated in a 0.105-in-radius, 70° circular sector. The dielectric substrate was Duroid 5880 ($\varepsilon_r = 2.2$, $\tan\delta = 0.0009$).

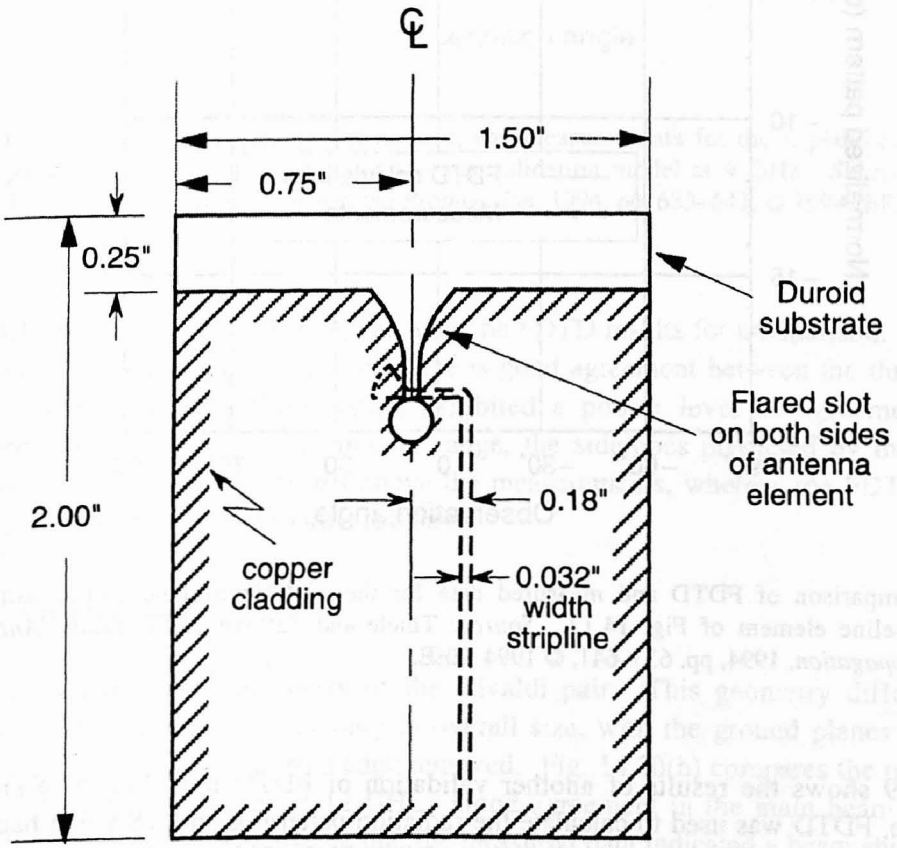


Fig. 14.17 Geometry of the Vivaldi single-flare baseline element. *Source:* Thiele and Taflove, *IEEE Trans. Antennas and Propagation*, 1994, pp. 633–641, © 1994 IEEE.

In the FDTD model, the cubic cell size was $\Delta = 0.5$ mm (0.020 in). This was based on the smallest physical dimension characterizing the antenna element: the gap across a slot at the point of its connection to its matching hole. The spatial resolution of $\Delta = 0.5$ mm corresponded to $\lambda_0/33$ to $\lambda_0/99$ over the 6 to 18 GHz bandwidth of the element. This high resolution permitted a simple stepped-edge (staircase) model to simulate the antenna radiation characteristics nearly as accurately as a more elaborate conformal contour-path model. Therefore, all modeling was performed using stepped edges. The resulting space-lattice size for the single- and double-flare Vivaldi models was $42 \times 116 \times 142$ cells, corresponding to 4.2 million field unknowns.

Fig. 14.18 compares the FDTD and measured results for the 12-GHz co-polarization (co-pol) pattern of the single Vivaldi element of Fig. 14.17. Here the radiation pattern was calculated in the plane of the antenna (the E-plane), with 0° denoting the forward direction. For the co-pol pattern, the E-field was polarized in this plane. From Fig. 14.18, we see a ±2.5-dB ripple in the main beam predicted by FDTD that was borne out by the measurements.

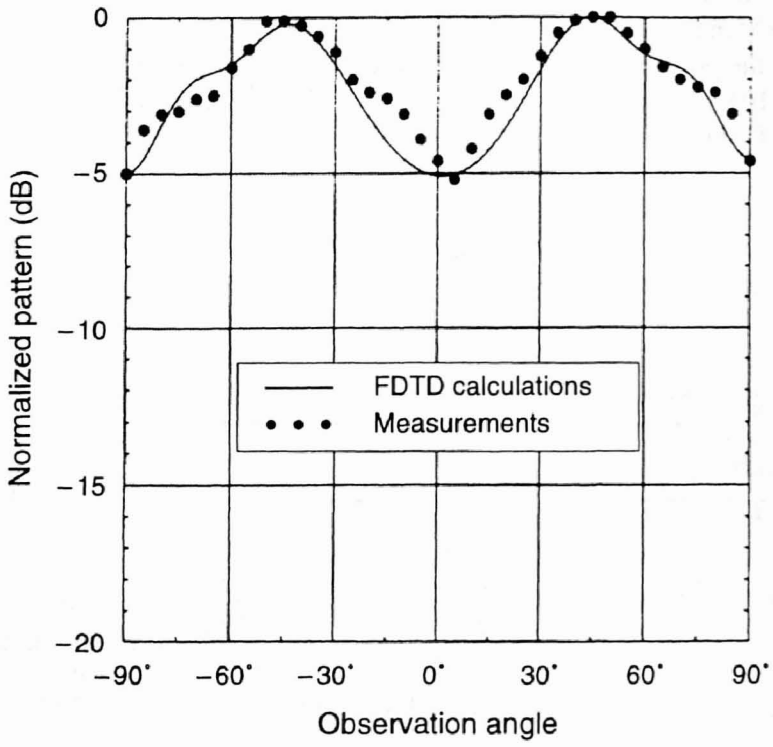


Fig. 14.18 Comparison of FDTD and measured data for the 12-GHz E-plane co-pol pattern of the baseline element of Fig. 14.17. *Source:* Thiele and Taflove, *IEEE Trans. Antennas and Propagation*, 1994, pp. 633–641, © 1994 IEEE.

Fig. 14.19 shows the results of another validation of FDTD modeling of a single-flare antenna. Here, FDTD was used to calculate the radiation pattern of an LTSA that had been the subject of measurements and moment-method modeling [44]. This antenna had the following physical parameters at 9 GHz: length $L/\lambda_0 = 3.0$, ground-plane width $H/\lambda_0 = 0.9$, substrate thickness $d/\lambda_0 = 0.021$, substrate permittivity $\varepsilon_r = 2.33$, and flare angle $2\alpha = 12°$. For this case, the FDTD space lattice resolution Δ was defined by the substrate thickness.

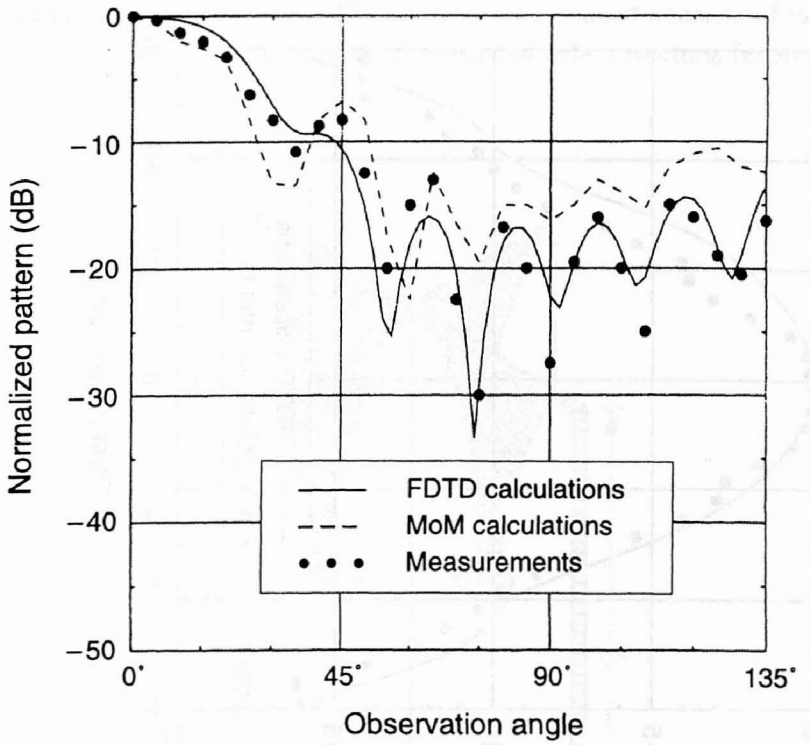


Fig. 14.19 Comparison of FDTD, moment-method, and measured data for the *E*-plane co-pol radiation pattern of a linearly tapered slot antenna validation model at 9 GHz. *Source:* Thiele and Taflove, *IEEE Trans. Antennas and Propagation*, 1994, pp. 633–641, © 1994 IEEE.

Fig. 14.19 replicates Fig. 6 of [44] and adds the FDTD results for comparison. Through the first 60° relative to the forward direction, there is good agreement between the three data sets. Beyond 60°, the moment-method results exhibited a poorer level of agreement with the measurements than FDTD. In this angular range, the sidelobes predicted by the method of moments were as much as 6 to 10 dB above the measurements, whereas the FDTD-computed sidelobes were very close to the reported test data.

14.7.3 The Vivaldi Pair

Fig. 14.20(a) illustrates the geometry of the Vivaldi pair. This geometry differs from the baseline configuration of Fig. 14.17 only in overall size, with the ground planes narrowed to 1 in, and the dielectric along the front edge removed. Fig. 14.20(b) compares the measured and FDTD-calculated co-pol patterns at 12 GHz. Good agreement in the main beam can be seen throughout the 180° range. We observe that the measured data indicated a beam shift (squint) of approximately 5°. The probable physical basis for this squint was a phase offset in the excitation of the elements. With the elements fed by a pair of coaxial lines, it can be shown that this level of squint arises from as little as a 0.5-mm difference in line length.

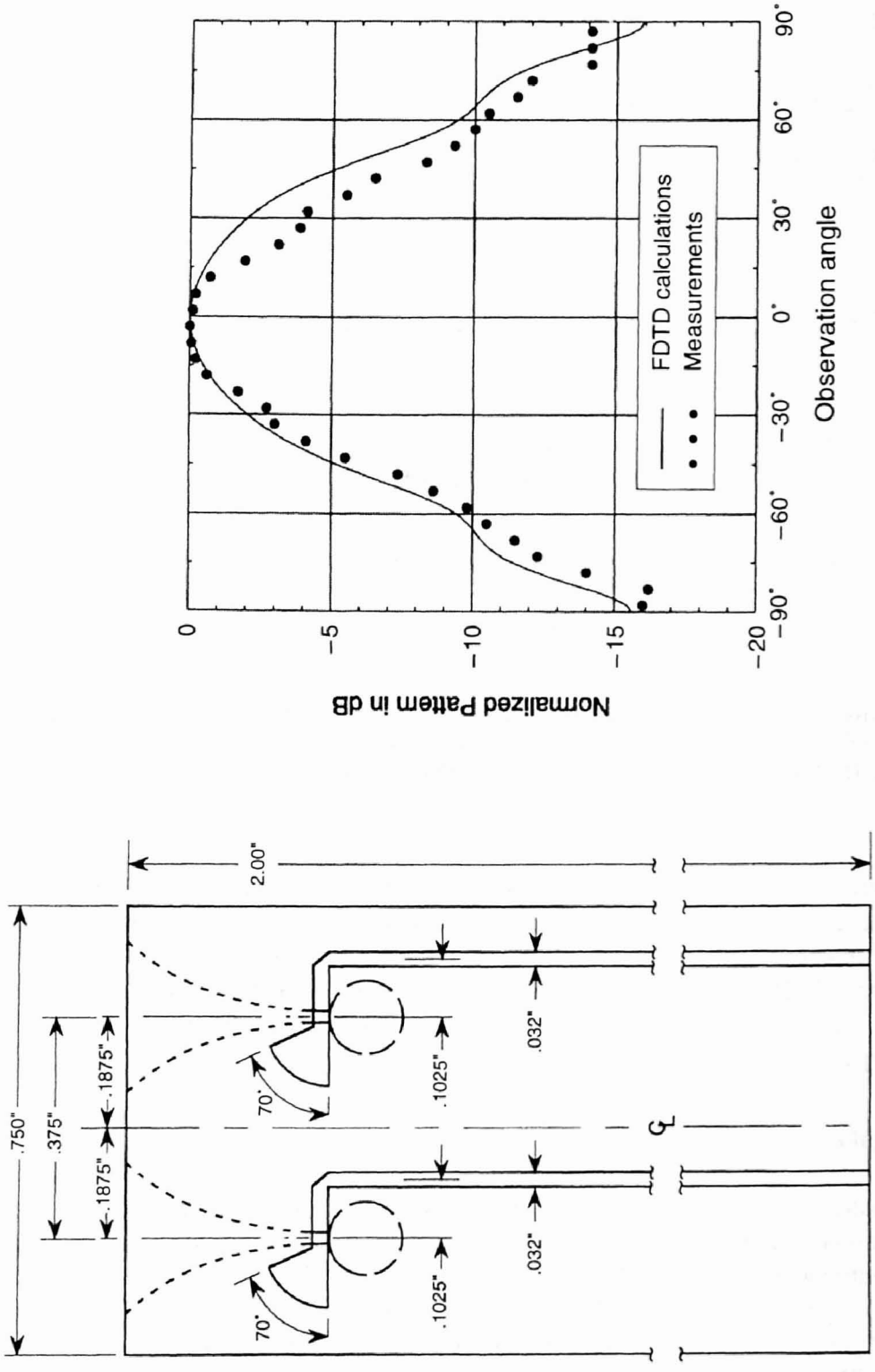


Fig. 14.20 The Vivaldi pair: (a) geometry; (b) comparison of FDTD calculations and measured E-plane co-pol pattern at 12 GHz. *Source:* Thiele and Taflove, *IEEE Trans. Antennas and Propagation*, 1994, pp. 633–641, © 1994 IEEE.

14.7.4 The Vivaldi Quad

Fig. 14.21 illustrates the geometry of the Vivaldi quad. Each quad consists of two Vivaldi pairs, discussed in the previous section, that are mutually perpendicular bisectors, forming a "+" shape.

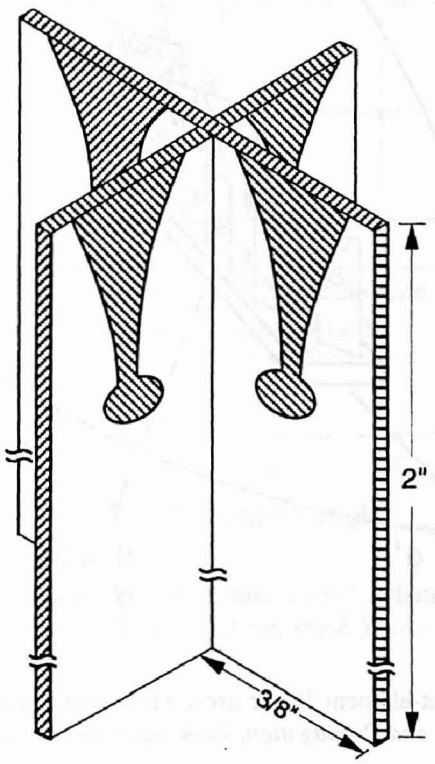


Fig. 14.21 Geometry of the Vivaldi quad antenna. *Source:* Thiele and Taflove, *IEEE Trans. Antennas and Propagation*, 1994, pp. 633–641, © 1994 IEEE.

The Vivaldi quad is the basic repetitive element used in the linear array of Fig. 14.22. Clearly, the Cartesian nature of the FDTD space lattice is well suited for geometry definition in this model. However, this arrangement requires a change in the usual definition of a far-field pattern "cut," since the axis of the linear array of Vivaldi quads (denoted as the x-axis in Fig. 14.22) lies along a diagonal in the FDTD lattice. To account for this, a great-circle cut is made. This involves defining two rotated axes x' and z', determining their unit vectors, and using these unit vectors to define a rotation matrix through which θ' and ϕ' can be mapped to the standard θ and ϕ spherical coordinates. This allows the provision of an E-plane scan with respect to the diagonal orientation of the array. Further, the far-field radiation components E_θ and E_ϕ are rotated to obtain the correct correspondence to co-pol and cross-pol radiation. These great-circle pattern cuts are used for both the single Vivaldi quad of Fig. 14.21 and the eight-element linear array of Vivaldi quads shown in Fig. 14.22.

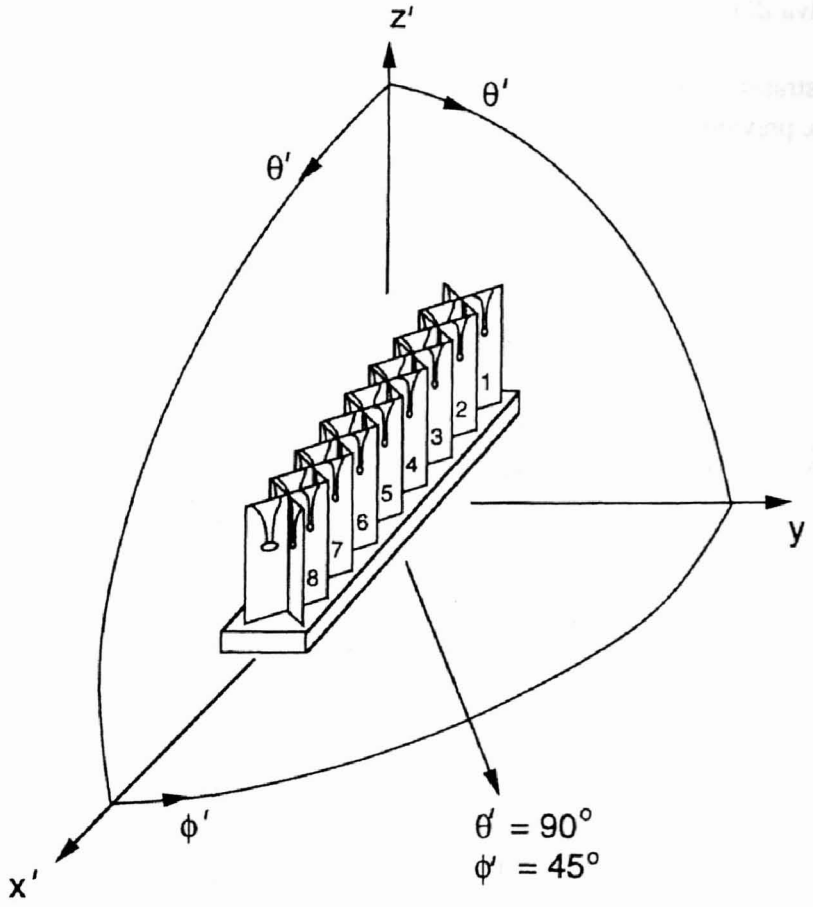


Fig. 14.22 Geometry of the eight-element linear array of Vivaldi quads. *Source:* Thiele and Taflove, *IEEE Trans. Antennas and Propagation*, 1994, pp. 633–641, © 1994 IEEE.

Fig. 14.23 compares FDTD modeling results with measured *E*-plane co-pol radiation patterns at 18 GHz for a single prototype Vivaldi quad having the geometry of Fig. 14.21. In general, good agreement is seen. However, while the FDTD model properly predicts a symmetric pattern, the measured data are not symmetric. For example, the measurements indicate a pattern "shoulder" at −40°, but not at +40°. Relative to the first sidelobes at ±70°, the measured levels are approximately 7 dB out-of-symmetry. In fact, the FDTD predictions of these sidelobes are approximately the geometric mean of the asymmetric measurements.

14.7.5 The Linear Phased Array

This section presents FDTD and measurement data for the complete eight-element linear array of Vivaldi quads shown in Fig. 14.22. The FDTD model, implemented on a 222 × 222 × 140-cell cubic lattice, uses an impulsive excitation along with an on-the-fly DFT of the fields at the near-to-far-field transformation surface. In this manner, a single run provides the complete radiation pattern at multiple key frequencies spanning the desired 6 to 18 GHz operating band. Applying an appropriate time-delay taper to the feeding striplines across the array provides for the desired beam-steering angle.

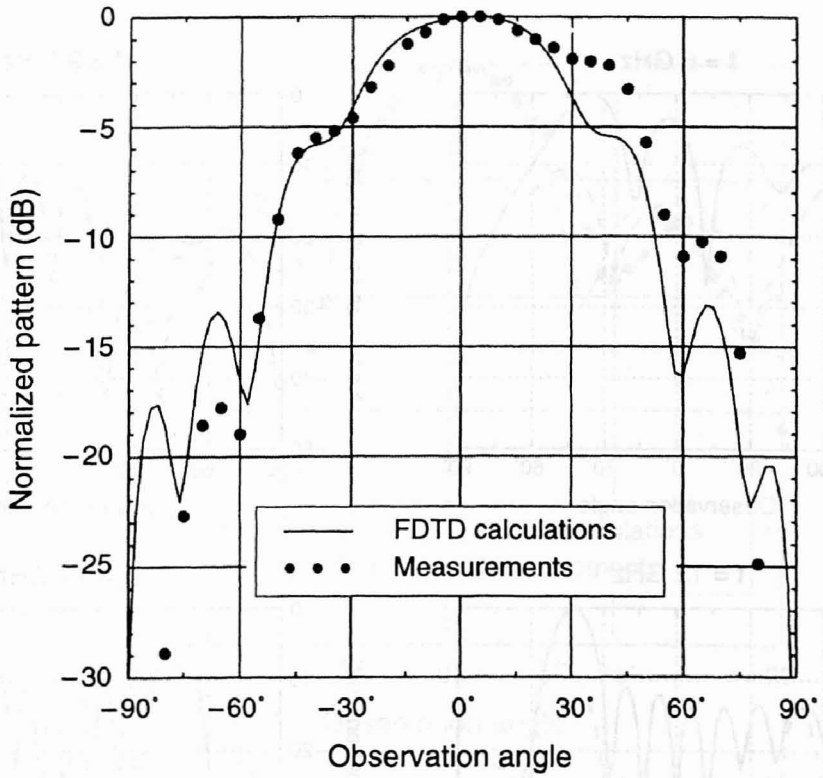


Fig. 14.23 Comparison of FDTD and measured *E*-plane co-pol radiation patterns for the single Vivaldi quad at 18 GHz. *Source:* Thiele and Taflove, *IEEE Trans. Antennas and Propagation*, 1994, pp. 633–641, © 1994 IEEE.

Fig. 14.24 graphs the FDTD modeling results at 6, 9, 12, 15, and 18 GHz for the *E*-plane co-pol and cross-pol radiation patterns for the eight-element array, assuming a nominal 45° beam steer. We see that grating lobes evolve as the operating frequency increases. In fact, the strength of the principal grating lobe equals or exceeds that of the nominal main beam for frequencies greater than 15 GHz. Further, the cross-pol levels are quite high, rising to within 10 dB of the co-pol levels in the main beam at all of the frequencies modeled.

Fig. 14.25 compares the FDTD calculations for the *E*-plane co-pol radiation pattern with measured results at 6 GHz (0° beam steer) and 12 GHz (45° beam steer). Good agreement is seen at most points in the radiation patterns, even the sidelobes, over a dynamic range approaching 30 dB.

14.7.6 Phased-Array Radiation Characteristics Indicated by the FDTD Modeling

Simple array theory can be used to qualitatively assess the above results. We note that in both the computed and measured patterns, the nominal beam-steer angle is not quite reached. For example, in Fig. 14.25(b), a desired beam-steer angle of 45° results in an actual angle of approximately 40°. Array theory predicts the overall radiation pattern to be the product of the element pattern and the array factor. The deviation of the beam-steer angle can logically be attributed to this pattern multiplication.

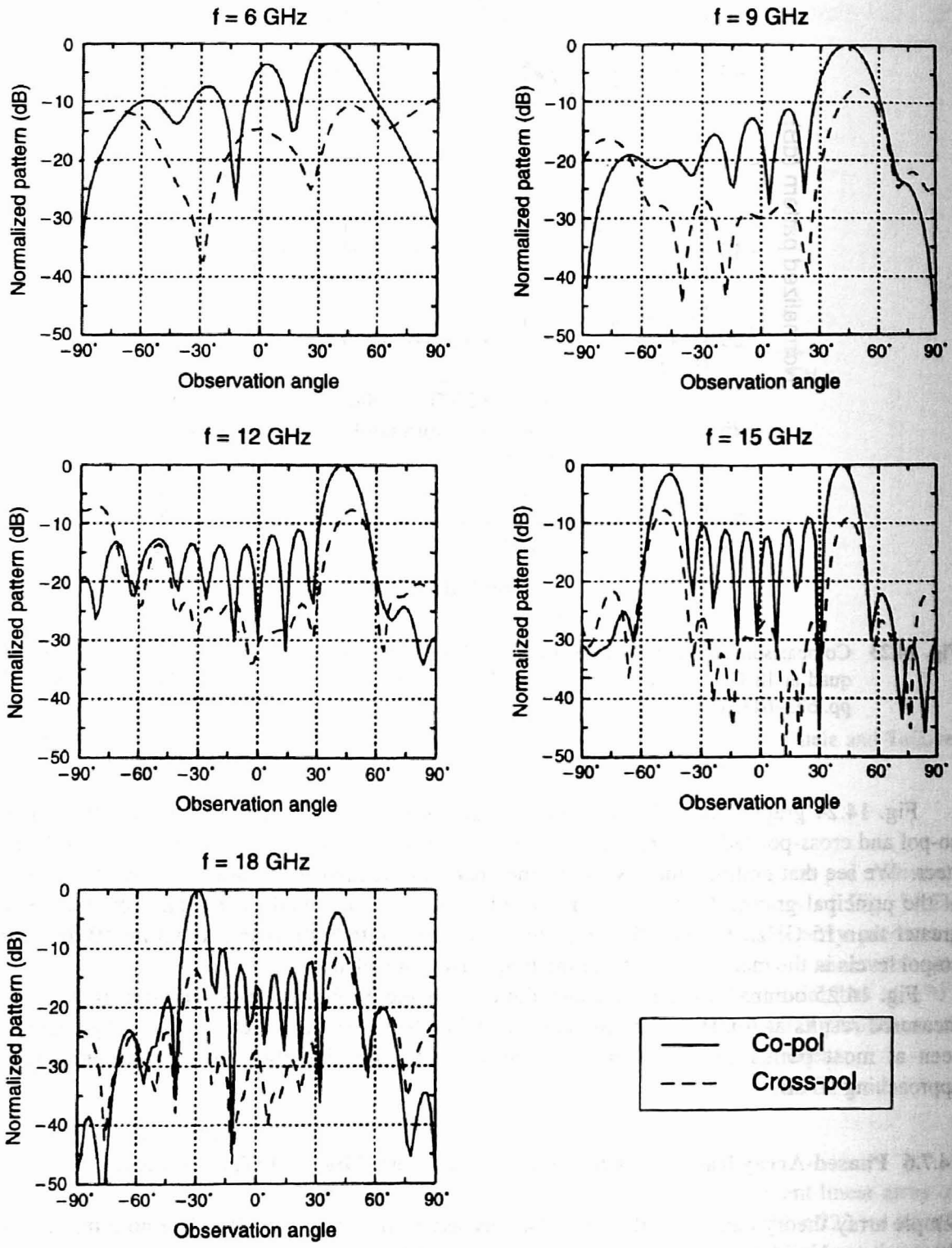


Fig. 14.24 FDTD-computed E-plane co-pol and cross-pol radiation patterns for the eight-element linear array of Vivaldi quads of Fig. 14.22, assuming a nominal 45° beam steer. Note the evolution of grating lobes as the operating frequency increases.

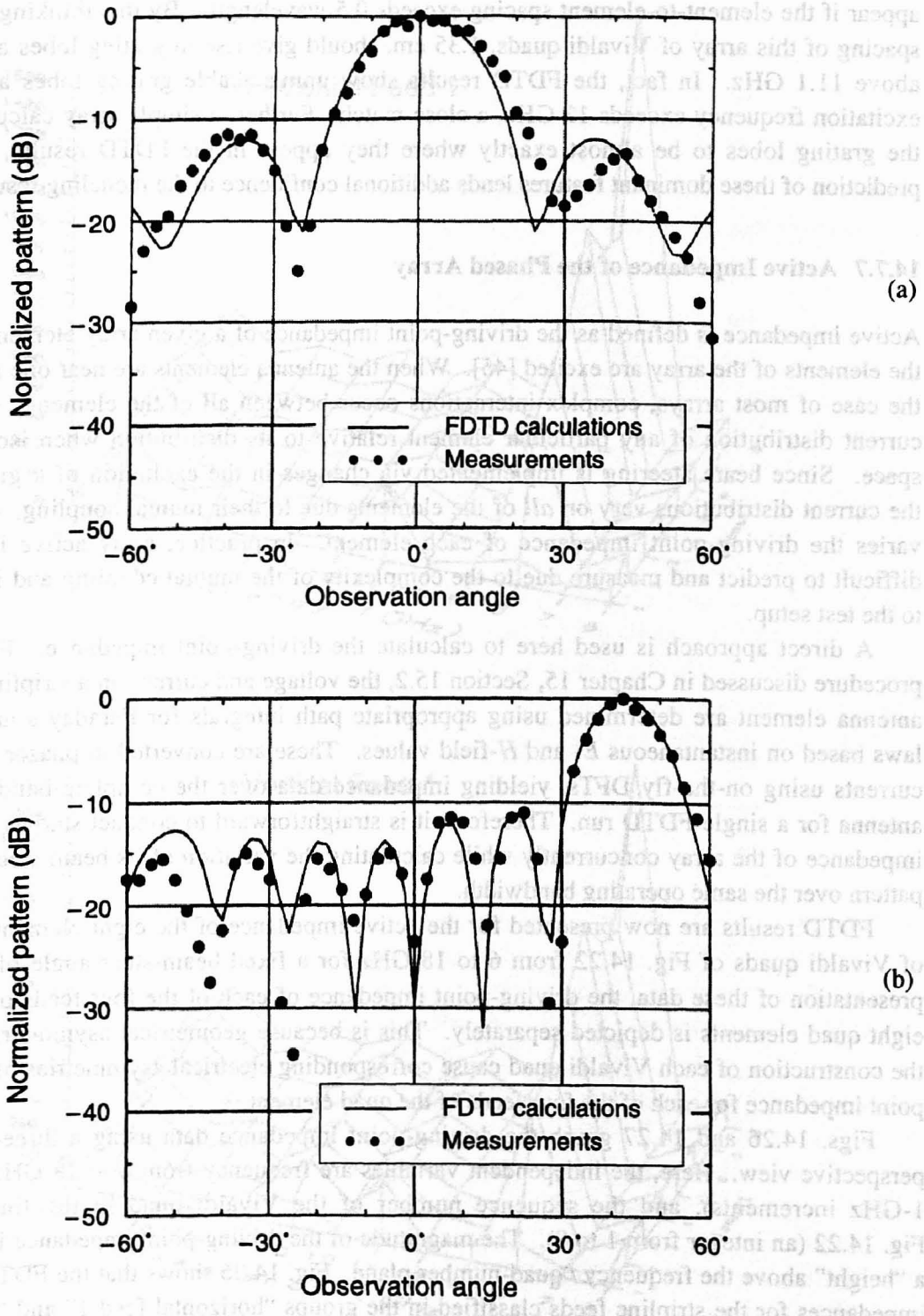


Fig. 14.25 Comparison of FDTD and measured *E*-plane co-pol radiation patterns for the eight-element linear array of Vivaldi quads of Fig. 14.22: (a) 6 GHz, 0° beam steer; (b) 12 GHz, 45° beam steer. *Source:* Thiele and Taflove, *IEEE Trans. Antennas and Propagation*, 1994, pp. 633–641, © 1994 IEEE.

Array theory can also be used to predict one of the most notable features of the FDTD modeling results, the grating lobes. For a linear array, array theory predicts that grating lobes appear if the element-to-element spacing exceeds 0.5 wavelength. By this thinking, the element spacing of this array of Vivaldi quads, 1.35 cm, should give rise to grating lobes at frequencies above 11.1 GHz. In fact, the FDTD results show unmistakable grating lobes arising as the excitation frequency exceeds 12 GHz, a close match. Further, a simple array calculation shows the grating lobes to be almost exactly where they appear in the FDTD results. The FDTD prediction of these dominant features lends additional confidence to the modeling results.

14.7.7 Active Impedance of the Phased Array

Active impedance is defined as the driving-point impedance of a given array element when all of the elements of the array are excited [45]. When the antenna elements are near one another, as in the case of most arrays, complex interactions occur between all of the elements, changing the current distribution of any particular element relative to its distribution when isolated in free space. Since beam steering is implemented via changes in the excitation of a given element, the current distributions vary on *all* of the elements due to their mutual coupling, which in turn varies the driving-point impedance of each element. In practice, array active impedance is difficult to predict and measure due to the complexity of the mutual coupling and its sensitivity to the test setup.

A direct approach is used here to calculate the driving-point impedance. Following the procedure discussed in Chapter 15, Section 15.2, the voltage and current on a stripline feeding an antenna element are determined using appropriate path integrals for Faraday's and Ampere's laws based on instantaneous E- and H-field values. These are converted to phasor voltages and currents using on-the-fly DFTs, yielding impedance data over the complete bandwidth of the antenna for a single FDTD run. Therefore, it is straightforward to conduct studies of the active impedance of the array concurrently while calculating the variation of its beam-steered radiation pattern over the same operating bandwidth.

FDTD results are now presented for the active impedance of the eight-element linear array of Vivaldi quads of Fig. 14.22 from 6 to 18 GHz for a fixed beam-steer angle of 45°. In the presentation of these data, the driving-point impedance of each of the four feeds of each of the eight quad elements is depicted separately. This is because geometrical asymmetries arising in the construction of each Vivaldi quad cause corresponding electrical asymmetries of the driving-point impedance for each of the four feeds of the quad element.

Figs. 14.26 and 14.27 graph the driving-point impedance data using a three-dimensional perspective view. Here, the independent variables are frequency from 6 to 18 GHz (in discrete 1-GHz increments), and the sequence number of the Vivaldi quad in the linear array of Fig. 14.22 (an integer from 1 to 8). The magnitude of the driving-point impedance is depicted as a "height" above the frequency / quad-number plane. Fig. 14.26 shows that the FDTD-calculated impedances for the stripline feeds classified in the groups "horizontal feed 1" and "vertical feed 1" are extremely high at 8 GHz, reaching a peak of nearly 1,300Ω at Vivaldi quad #8. Less substantial peaks in impedance are calculated at 11 and 17 GHz. Fig. 14.27 shows that the FDTD-calculated impedances for the feeds classified in the groups "horizontal feed 2" and "vertical feed 2" attain large values at 11 GHz, with lesser peaks at 13 and 17 GHz. Here, Vivaldi quads #4 and #5 experience the largest values of driving-point impedance.

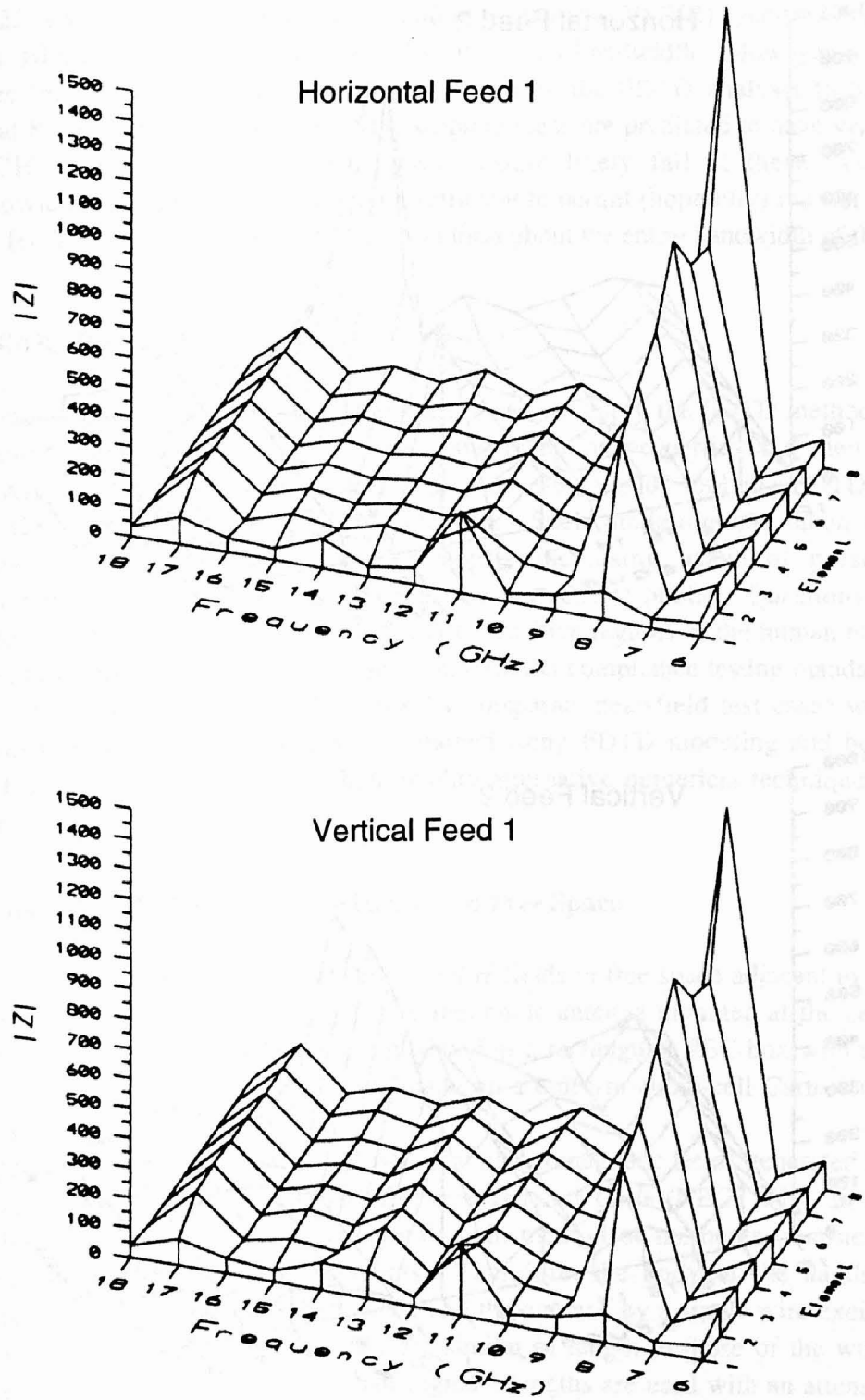


Fig. 14.26 FDTD-computed active impedance of the eight-element linear array of Vivaldi quads of Fig. 14.22 for a fixed 45° beam steer and for Type-1 stripline feeds. *Source:* Thiele, Ph.D. dissertation, 1994, p. 99.

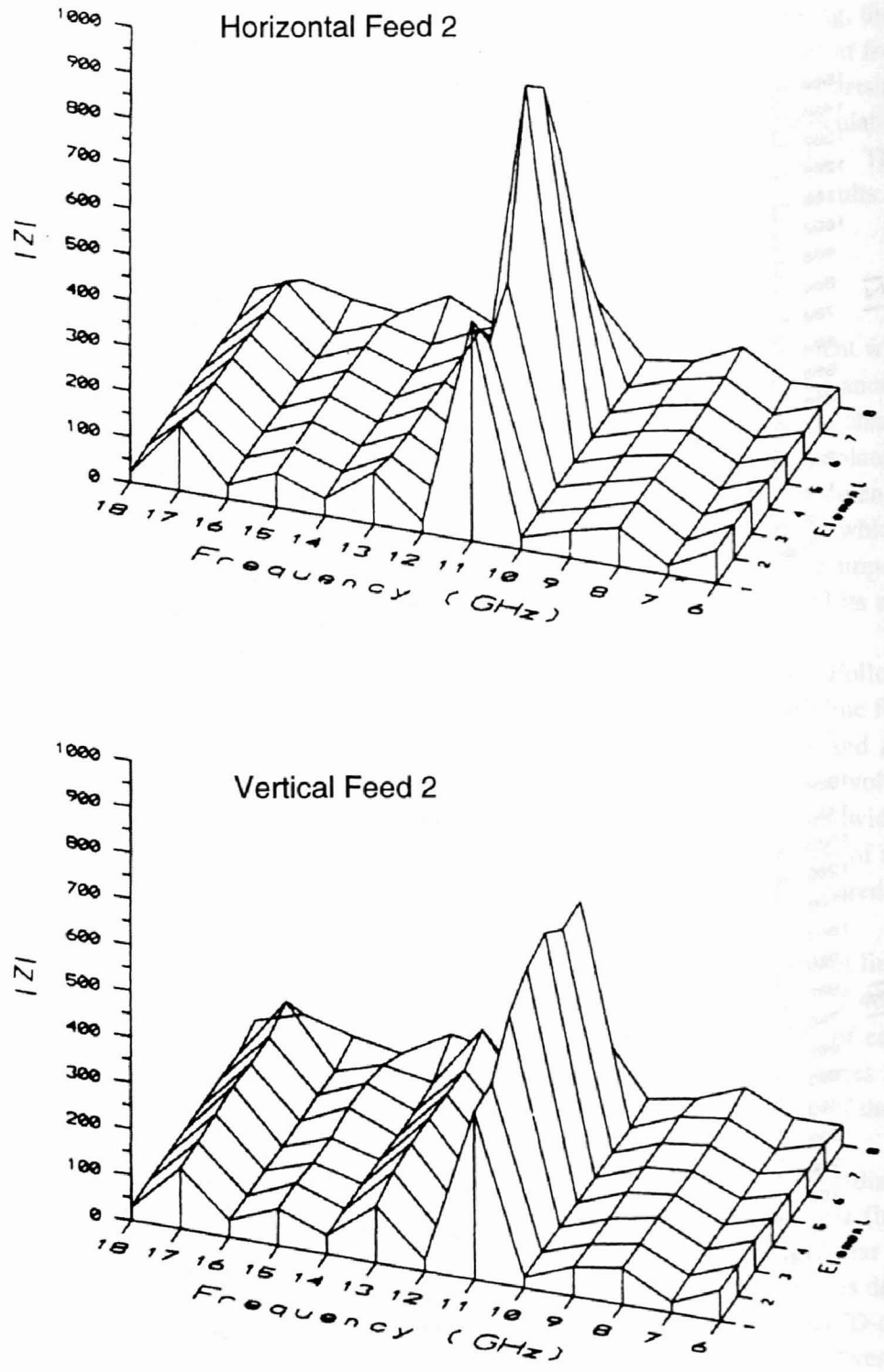


Fig. 14.27 FDTD-computed active impedance of the eight-element linear array of Vivaldi quads of Fig. 14.22 for a fixed 45° beam steer and for Type-2 stripline feeds. *Source:* Thiele, Ph.D. dissertation, 1994, p. 100.

Overall, these FDTD calculations indicate that, for a fixed beam-steer angle of 45°, the driving-point impedance behavior of the eight-element linear array of Vivaldi quads of Fig. 14.22 is within the desired *voltage standing-wave ratio* (VSWR) specification of 2:1 or less (25Ω to 100Ω), for much of the 6 to 18 GHz design bandwidth. However, one-half of the striplines feeding the Vivaldi quads are predicted by the FDTD analysis to have very high VSWR at 8 GHz, and the other half of the stripline feeds are predicted to have very high VSWR at 11 GHz. Therefore, the antenna array would likely fail at these two frequencies. Prior knowledge of this problem should be sufficient to permit (hopefully) modest changes in the stripline feeds to meet the VSWR specification throughout the entire bandwidth of the array.

14.8 NEAR-FIELD SIMULATIONS

The previous sections in this chapter discussed means to apply the FDTD method to accurately model antennas for their far-field radiation and reception properties and their driving-point characteristics. In contrast, this section focuses on examples of applying FDTD techniques to analyze the near fields generated by antennas. Such modeling has taken on increasing importance, even urgency, due to the rapidly increasing usage of personal wireless communications devices, such as cellphones, by the general public. Questions regarding the near-field coupling of such devices to potentially sensitive regions of the human body, especially the brain, have prompted *specific absorption rate* (SAR) compliance testing mandated by various agencies [46–49]. This section describes five disparate near-field test cases where excellent agreement is found between the results obtained using FDTD modeling and benchmark data derived from either exact analyses, high-quality alternative numerical techniques, or carefully performed experiments.

14.8.1 Generic 900-MHz Cellphone Handset in Free Space

This example involves calculation of the E- and H-fields in free space adjacent to a generic 900-MHz cellphone handset having a $\lambda_0/4$ wire monopole antenna mounted at the center of its top face [50]. For simplicity, the handset is modeled as a rectangular PEC box with the dimensions $15 \times 6 \times 2.4$ cm. FDTD calculations are made on a uniform cubic-cell Cartesian mesh having $\Delta = 2$ mm.

To provide a validation benchmark, the near electromagnetic fields generated by the handset are also calculated using the *Numerical Electromagnetics Code* (NEC) [51]. In modeling with NEC, which is based upon the frequency-domain method of moments, a structure has to be described in terms of thin wire segments. Therefore, the body of the handset in NEC is represented by a cage of intermeshed wires, and the antenna by a single wire excited at its base. The segments at the base of the antenna are similar in length to those of the wire mesh at the point it is connected. For accuracy, small segment lengths are used with an attendant scheme to transition from the fine segmentation in the antenna to the relatively coarse segmentation required to describe the surface of the handset with a reasonably small number of segments. A 16-wire radial segmentation scheme is used at the connection of the monopole wire to the top of the handset box. In all, there are 1,150 segments in the NEC model.

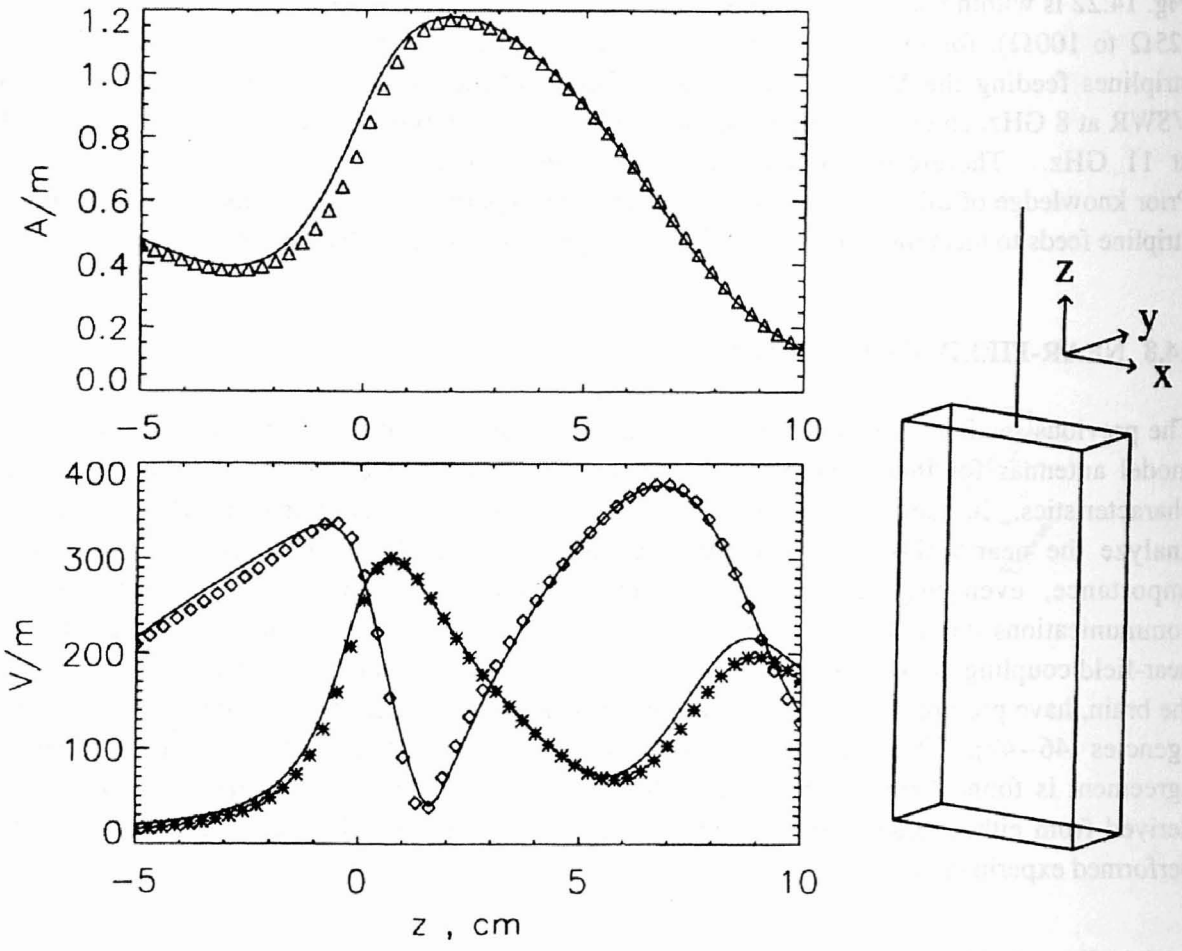


Fig. 14.28 Comparison of FDTD and NEC values of the near fields adjacent to a generic 900-MHz cellphone handset in free space for a transmitter output power of 1W into a $\lambda_0/4$ wire monopole antenna centered at the top of the handset. The observation locus is a vertical cut located 1 cm from the side of the handset box in the x-direction, where the top of the box is at the plane $z = 0$. Legend: FDTD values = solid line; NEC values of H_y = triangles; NEC values of E_x = diamonds; NEC values of E_z = asterisks. *Source:* Dimbylow and Mann, *Phys. Med. Biol.*, 1994, pp. 1537–1553.

Fig. 14.28 compares the calculated FDTD and NEC field values for an assumed transmitter output power of 1W. Here, the vertical electric field E_z, the radial electric field E_x, and the circumferential magnetic field H_y are graphed as a function of height z along a vertical cut located a distance of 1 cm from the side of the handset box in the x-direction, where the top of the box is at the plane $z = 0$. We observe very good agreement between the two sets of data. This high level of agreement of predictions arising from two codes with completely different theoretical and numerical bases indicates reliably that both techniques are valid for such free-space near-field models.

14.8.2 900-MHz Dipole Antenna Near a Layered Bone-Brain Half-Space

The second example involves calculation of the E-field generated by a 12-cm-long, 900-MHz dipole antenna and penetrating into a layered bone-brain half-space, where the dipole is located in air 2 cm away from the surface of the half-space. An exact analytical solution based on expansion of the dipole fields and enforcement of the boundary conditions at the planar material interfaces is given in [52]. For this specific case, the bone layer is 1-cm thick with $\varepsilon_r = 5.98$ and $\sigma = 0.099$ S/m. The underlying brain is infinitely thick with $\varepsilon_r = 50.3$ and $\sigma = 1.34$ S/m.

The FDTD space lattice uses a uniform cell size of $\Delta x = \Delta y = 5$ mm and $\Delta z = 9.23$ mm. This gives an antenna length of $13\Delta z$ with the feedpoint in the center, and $4\Delta y$ of free space between the antenna and the bone layer, which is $2\Delta y$ thick. The underlying brain layer is 24.5 cm ($49\Delta y$) thick. This is sufficient to attenuate the fields to well below 1% of their peak values before the end of the brain layer is reached. The infinite nature of the layered half-space in the transverse x-z directions is simulated by using a sufficiently large grid (12×33.2 cm) in these directions, such that the fields are very small at the grid edges. Convergence to the sinusoidal steady state is observed at about eight periods of the 900-MHz excitation.

Fig. 14.29 compares the magnitude of the FDTD-computed E-field along the y-axis (penetrating into the half-space) with the analytical solution (Table 1 of [52]). The field values are normalized to an antenna feedpoint current of 100 mA rms. Agreement is seen to be good.

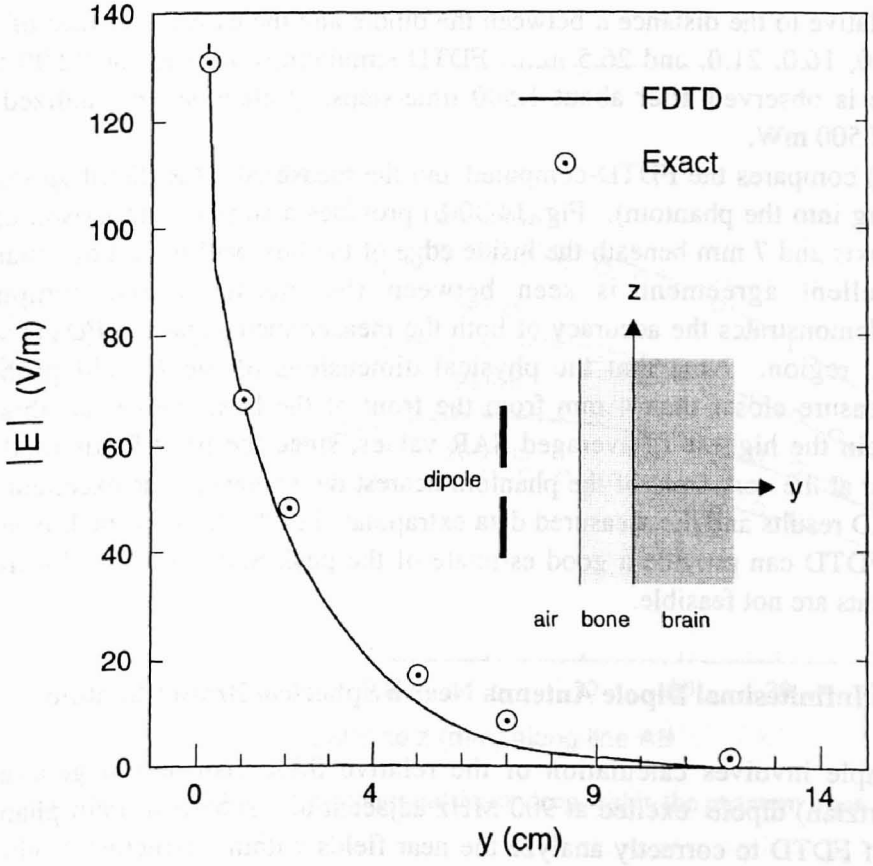


Fig. 14.29 Comparison of FDTD and exact results for the E-field penetrating along the y-axis into a layered bone-brain half-space excited by a 12-cm-long dipole antenna operating at 900 MHz.

14.8.3 840-MHz Dipole Antenna Near a Rectangular Brain Phantom

The third example involves calculation of the SAR generated by a half-wavelength (17.4-cm-long), 840-MHz dipole antenna within an acrylic box filled with brain-equivalent phantom material. Here, comparison is made to measured data. The dipole is placed at several distances from the box, but always with the feedpoint centered in front of it. The box dimensions are $30 \times 15 \times 50$ cm in the x-, y-, and z- directions, respectively.

The walls of the box are made of 6.35-mm-thick acrylic, which is assumed to have the electrical properties of polystyrene ($\varepsilon_r = 2.55$). Since the FDTD cell size used in this simulation is slightly larger than the wall thickness, the effective dielectric constant for the acrylic is calculated from the formula given in [53]

$$\varepsilon_{\text{eff}} = \frac{\varepsilon_r \Delta}{\varepsilon_r(\Delta - w) + w} \tag{14.37}$$

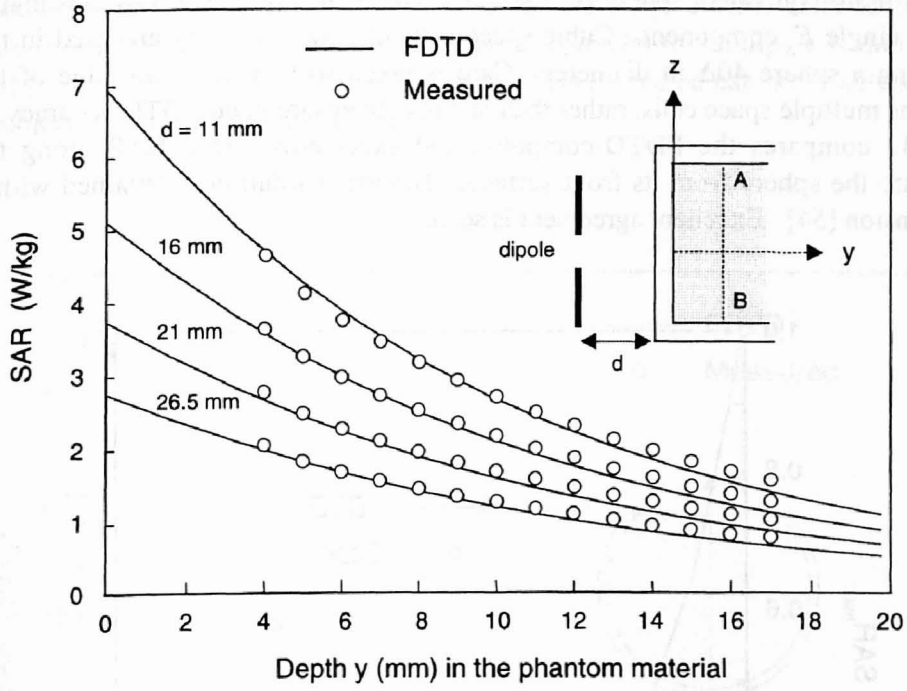
where Δ is the size of the FDTD cell, w is the thickness of the acrylic, and ε_r is the dielectric constant of polystyrene. In [53], this formula is derived by finding the average field at the air-acrylic interface and setting this equal to the field in the larger-sized FDTD cell, thereby giving the smaller effective dielectric constant of (14.37). The acrylic box is filled with brain-simulating material with $\varepsilon_r = 41.1$ and $\sigma = 1.06$ S/m.

SAR measurements are taken with a Narda Model 8021B E-field probe. All data are parameterized relative to the distance d between the dipole and the exterior surface of the acrylic box for d = 11.0, 16.0, 21.0, and 26.5 mm. FDTD simulations are run for 2,500 time-steps, and convergence is observed after about 1,500 time-steps. Values are normalized to a total radiated power of 500 mW.

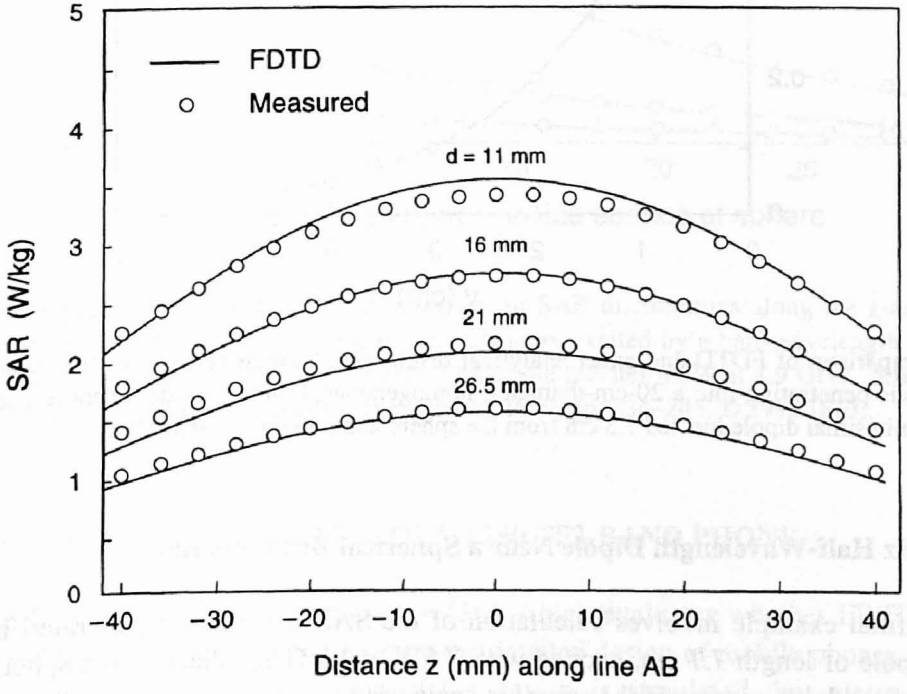
Fig. 14.30(a) compares the FDTD-computed and the measured SAR distributions along the y-axis (penetrating into the phantom). Fig. 14.30(b) provides a similar comparison along a line parallel to the z-axis and 7 mm beneath the inside edge of the box wall within the phantom brain material. Excellent agreement is seen between the measured and computed data. This agreement demonstrates the accuracy of both the measurements and the FDTD simulations in this near-field region. Note that the physical dimensions of the E-field probe make it impossible to measure closer than 4 mm from the front of the box. However, this would be necessary to obtain the highest 1g-averaged SAR values, since the peak fields (and hence the peak SARs) occur at the very front of the phantom nearest the antenna. The excellent agreement between the FDTD results and the measured data extrapolated to the front of the box supports the conclusion that FDTD can provide a good estimate of the peak SARs, even in locations where direct measurements are not feasible.

14.8.4 900-MHz Infinitesimal Dipole Antenna Near a Spherical Brain Phantom

The fourth example involves calculation of the relative SAR distribution generated by an infinitesimal (Hertzian) dipole excited at 900 MHz adjacent to a spherical brain phantom. This tests the ability of FDTD to correctly analyze the near fields within a structure having a curved surface. Here, simple staircasing is used (i.e., the best fit to the sphere's surface by the rectangular unit cells of the space lattice). Note that the smaller the sphere is relative to the cell size, the coarser is its staircased model.



(a) Along the y-axis penetrating into the phantom brain material.



(b) Along line AB parallel to the dipole and 7 mm deep within the phantom brain material.

Fig. 14.30 Comparison of FDTD-computed and measured SAR distributions generated by a half-wavelength 840-MHz dipole antenna within an acrylic box filled with brain-equivalent phantom material. The total radiated power is 500 mW.

For this case, the infinitesimal dipole is located 1.5 cm from a 20-cm-diameter, homogeneous brain-equivalent sphere ($\varepsilon_r = 43$, $\sigma = 0.83$ S/m) [54]. The infinitesimal dipole is modeled as a single E_z component. Cubic space cells of size $\Delta = 5$ mm are used in the FDTD model, yielding a sphere 40Δ in diameter. Care is taken so that the front edge of the model sphere contains multiple space cells, rather than just one, to ensure good FDTD accuracy.

Fig. 14.31 compares the FDTD-computed and exact normalized SAR along the y-axis penetrating into the sphere from its front surface. The exact solution is obtained with a Bessel function expansion [54]. Excellent agreement is seen.

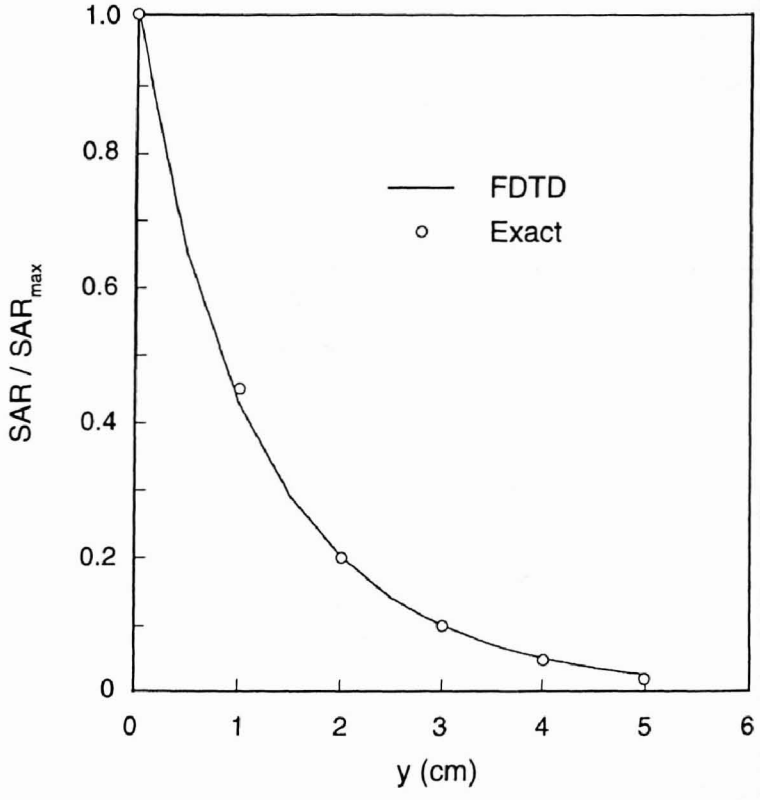


Fig. 14.31 Comparison of FDTD and exact analytical results for the relative SAR distribution along the y-axis penetrating into a 20-cm-diameter, homogeneous, brain-equivalent sphere excited by an infinitesimal dipole located 1.5 cm from the sphere and radiating at 900 MHz.

14.8.5 1.9-GHz Half-Wavelength Dipole Near a Spherical Brain Phantom

The fifth and final example involves calculation of the SAR distribution generated by a half-wavelength dipole of length 7.7 cm, excited with 0.5W at 1.9 GHz adjacent to a spherical brain phantom [55]. Simple staircasing of the sphere is again used.

For this case, the dipole is located at $d = 5$, 15, or 25 mm from the bottom of a 223-mm outside-diameter spherical glass container having a wall thickness of 5 mm and wall permittivity $\varepsilon_r = 4$. To simulate the average dielectric properties of the brain at 1.9 GHz, this container is filled with a homogeneous fluid of $\varepsilon_r = 45.5$ and $\sigma = 1.31$ S/m. Cubic space cells of size $\Delta = 2.5$ mm are used in the FDTD model.

Fig. 14.32 compares the FDTD-calculated and measured SAR distributions along the vertical z-axis penetrating into the phantom for the three spacings d [55]. Excellent agreement is seen. There is similar agreement along horizontal cuts parallel to the dipole within the phantom medium [55]. These results give confidence that the FDTD method can accurately model curved absorbing objects for near-field simulations.

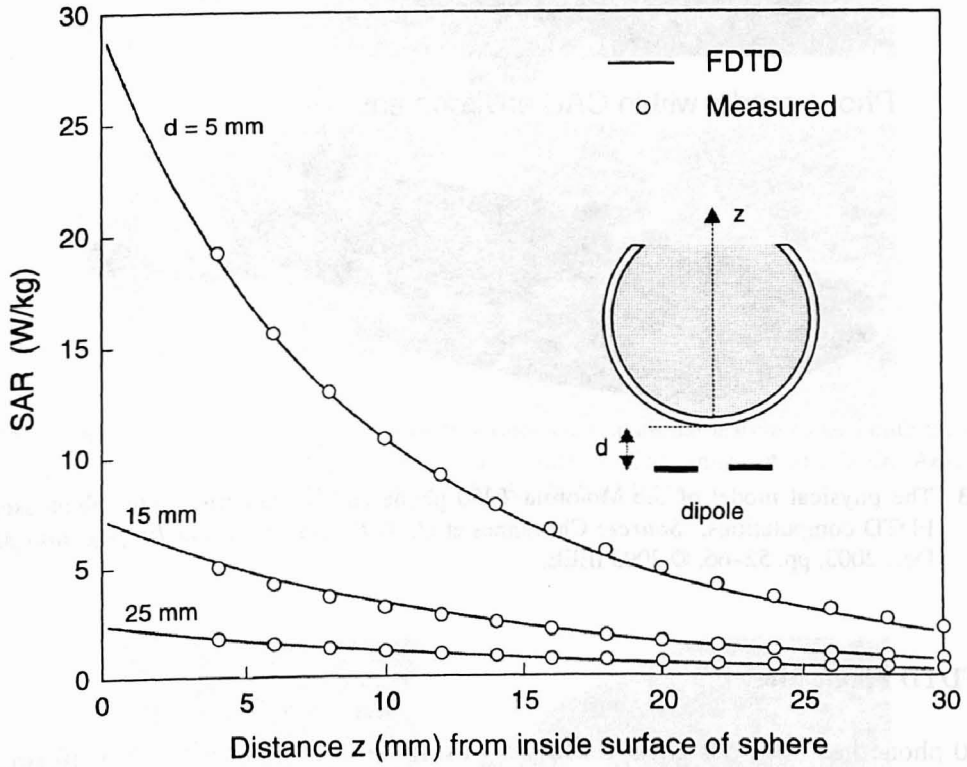


Fig. 14.32 Comparison of FDTD and measured data for SAR distributions along the z-axis penetrating into a 223-mm-diameter spherical brain phantom excited by a half-wavelength dipole located $d = 5$, 15, or 25 mm from the sphere. The dipole radiates 0.5W at 1.9 GHz. *Source:* Yu et al., *IEEE Trans. Electromagnetic Compatibility*, 1999, pp. 234–245, © 1999 IEEE.

14.9 CASE STUDY III: THE MOTOROLA T250 TRI-BAND PHONE

This section focuses on the study reported in [56], which evaluates whether FDTD-based tools can be utilized for supporting RF engineers in the detailed design of mobile phones. This clearly demands that not only the outer shape of the device is simulated, but also all embedded electromechanical components that are relevant to the RF characteristics of the phone, as shown in Fig. 14.33. This requires spatial resolutions down to 100 μm or less.

The quality of the FDTD simulation is validated by measurements utilizing the DASY V4 dosimetric assessment system [57] near-field scanners. The phone selected for this study is the Motorola T250, a commercially available tri-band phone (GSM, DCS, PCS).

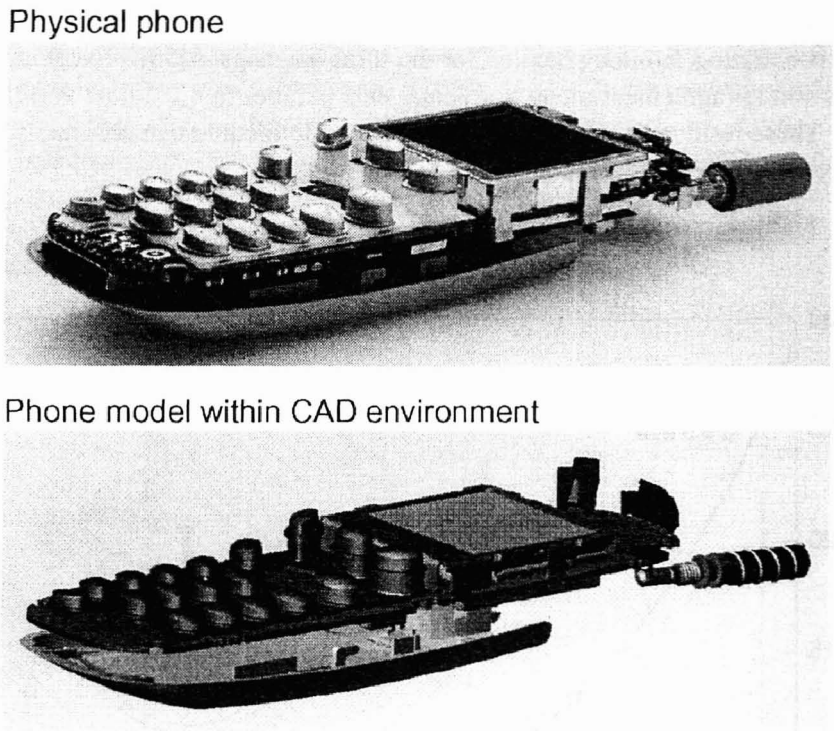


Fig. 14.33 The physical model of the Motorola T250 phone and its numerical equivalent used for the FDTD computations. *Source:* Chavannes et al., *IEEE Antennas and Propagation Magazine*, Dec. 2003, pp. 52–66, © 2003 IEEE.

14.9.1 FDTD Phone Model

The T250 phone had been designed at Motorola using the ProEngineer CAD platform, thereby allowing its CAD dataset to be exported in stereolithography format to the SEMCAD FDTD software [58]. In the FDTD model, the phone is oriented with its *printed circuit board* (PCB) aligned with a primary space lattice plane to reduce staircasing. Whereas basic dielectric parts such as the housing and pushbuttons are maintained as imported into the FDTD software, special attention must be paid to parts which significantly influence electromagnetic field behavior.

The PCB is modeled as two 450-μm-thick dielectric layers embedded into three 110-μm-thick PEC ground layers. PCB grounds are connected to each other by 50 vias that are uniformly distributed over the entire board area. All metallic parts are modeled as PEC. Dielectric components are represented within the FDTD model using material parameters obtained from the manufacturers of the various parts.

As illustrated in Fig. 14.34, to effectively mesh the phone geometry with acceptable memory and running-time requirements, a combination of graded meshes and local FDTD mesh refinement is applied. In particular, the mesh-refinement scheme outlined in Chapter 11, Section 11.8, is applied to model certain regions of the phone with a fine spatial resolution. The most crucial parts to model with fine resolution are the antenna, including its wire helix and monopole (0.65-mm wire diameter), as well as the PCB with its thin PEC ground layers. Since the helix is completely nonconformal in the FDTD lattice, it must be resolved by multiple cells in its wire diameter to permit accurate calculation of its field distribution.

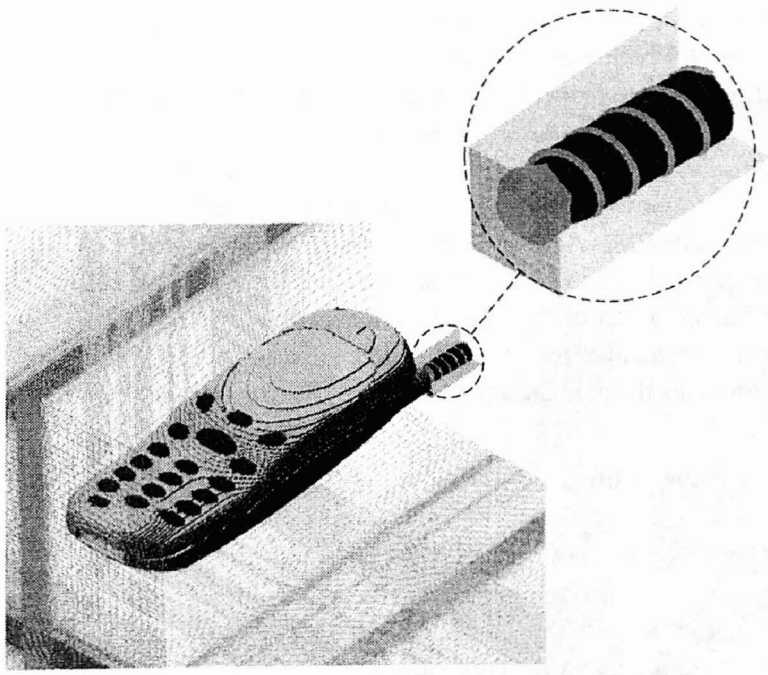


Fig. 14.34 Enhanced FDTD meshing approach to reduce computational resources by combining a graded mesh and a subgrid for the antenna. *Source:* Chavannes et al., *IEEE Antennas and Propagation Magazine*, Dec. 2003, pp. 52–66, © 2003 IEEE.

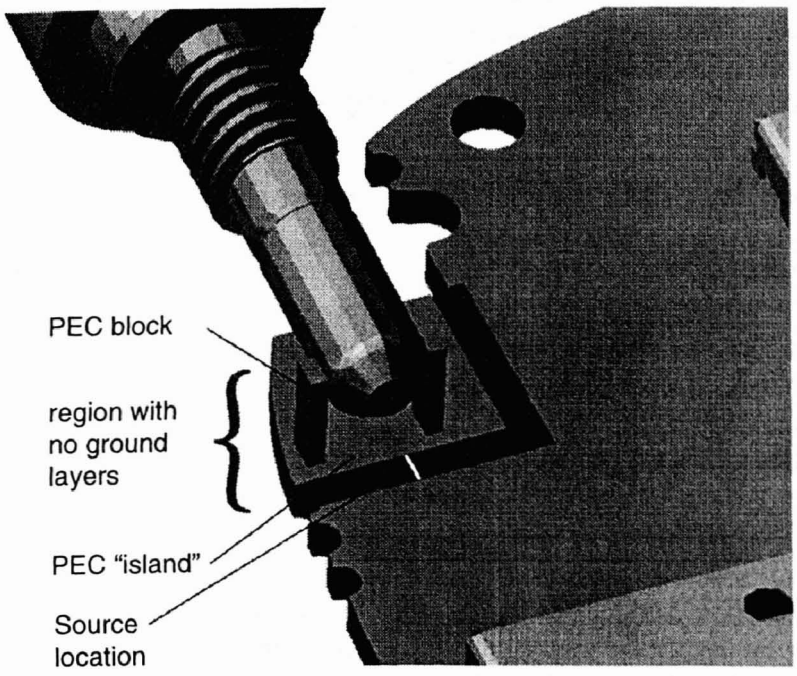


Fig. 14.35 Excitation of the antenna-PCB structure in the modified source region. *Source:* Chavannes et al., *IEEE Antennas and Propagation Magazine*, Dec. 2003, pp. 52–66, © 2003 IEEE.

A spatial resolution of $0.2 \times 0.2 \times 0.1$ mm in the x-, y, and z-directions is used within the 500,000-cell subgrid (2:1 refinement factor) surrounding the antenna. For the outer grid in the free-space simulation, an inhomogeneous mesh having an increasing cell size from 0.2 to 7 mm (grading ratio = 1.5) is used. All simulations including head models consist of the same minimum and maximum cell sizes. For these models, the human ear region is rendered with a $0.4 \times 0.4 \times 0.2$-mm resolution, leading to a total of about 6 million cells. Finally, all models employ an eight-cell-thick PML ABC to terminate the space lattice.

To achieve proper excitation of the simulated phone structure, a source region is modeled that closely corresponds to that of the physical phone. Figure 14.35 illustrates the details of this setup in the CAD environment. Here, the RF source is placed across a 1-mm air gap between the PEC island that terminates the antenna and the upper PCB ground layer.

14.9.2 Measurement Procedures

The near-field scanning system DASY4 [57] that is used for the measurements in this case study is equipped with the latest probes providing the required isotropy, sensitivity, and spatial resolution. The phantom used is the Twin SAM [59], which has been standardized for antenna performance characterization by various organizations.

We note that the T250 design does not permit an accurate, direct measurement of the antenna input power. Therefore, the antenna input power is determined indirectly by matching calculated radiated H-field distributions with measured ones. In order to minimize uncertainty, a global least-squares fit is conducted on the H-field distributions in free space at a variety of distances from the phone, and inside a flat phantom. This procedure includes information regarding the reactive near-field, thereby providing a more reliable determination of the antenna input power than using either conductive power measurements or far-field radiated power measurements alone. Following this approach, the values for the antenna input power (to which all reported simulation results are normalized) are found to be +30.6 dBm at 902 MHz (expressed as GSM900), and +29.1 dBm at 1,747 MHz (expressed as DCS1800).

14.9.3 Free Space Near-Field Investigations and Assessment of Design Capabilities

Fig. 14.36 provides grayscale visualizations of the measured and FDTD-calculated E- and H-field distributions immediately adjacent to the T250 at the DCS1800 frequency. The fields are shown in a plane located 10 mm above the phone with the PCB aligned horizontally, and the phone placed on a Styrofoam block. While there are slight deviations in the region of the field maxima close to the antenna, the agreement is seen to be very good.

One benefit of the high level of modeling detail afforded by FDTD is that we can assess how the non-RF electromechanical parts in the phone might influence the phone's electromagnetic field behavior. An illustrative example of this benefit involves a modeling study of how the *liquid crystal display* (LCD) is connected to the PCB. FDTD modeling has revealed that major differences in the phone's electromagnetic fields in the vicinity of the LCD arise, dependent upon whether or not the four clamp-like electrical connections between the LCD and the PCB constitute a proper RF short-circuit to the PCB ground. These modeling results have been verified by measurements for both possibilities (i.e., poor and good RF ground), with an excellent level of agreement of the type shown in Fig. 14.36.

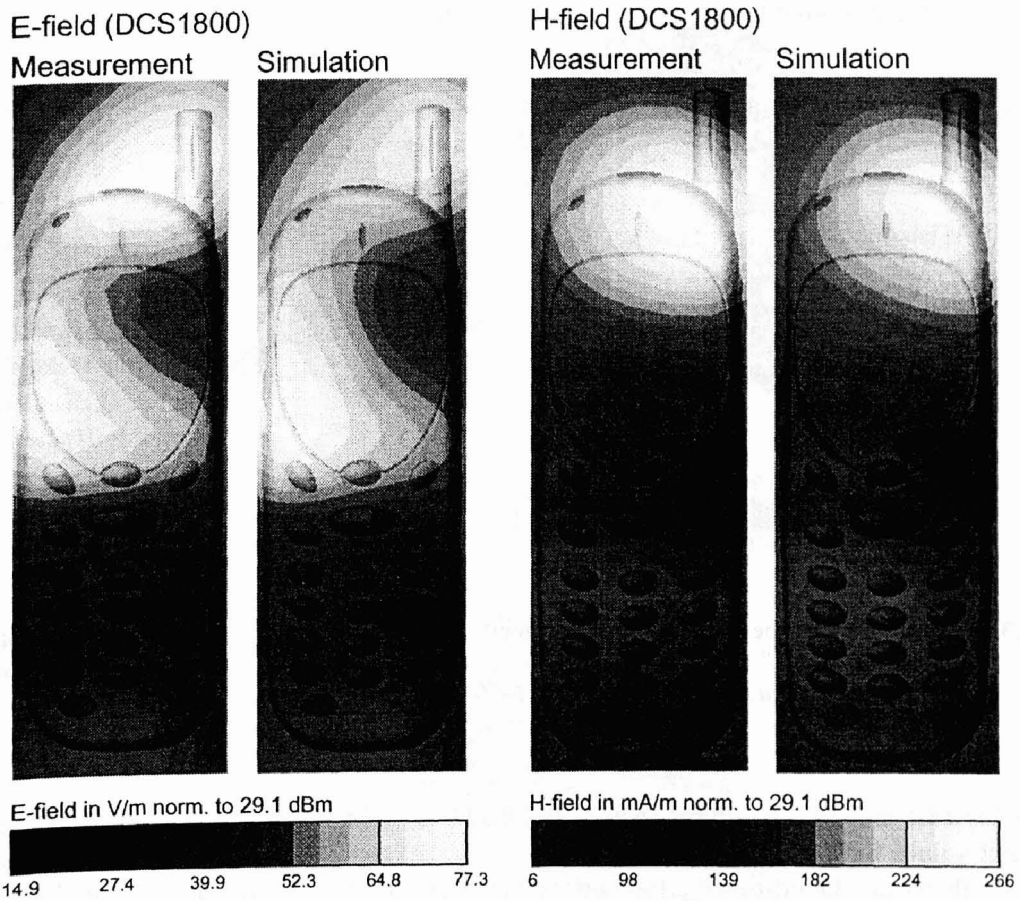


Fig. 14.36 Comparison of FDTD-computed and measured *E*-field (left) and *H*-field (right) (DCS1800), shown in a plane located 10 mm above the phone. All fields are normalized to an antenna input power of +29.1 dBm. *Source:* Chavannes et al., *IEEE Antennas and Propagation Magazine*, Dec. 2003, pp. 52–66, © 2003 IEEE.

14.9.4 Performance in Loaded Conditions (SAM and MRI-Based Human Head Model)

For this case study, the T250 phone is placed adjacent to the SAM standard phantom [59] in a manner equivalent to that performed during SAR compliance testing. The dielectric parameters for the liquid and shell used in the measurements and simulations at both the GSM900 and DCS1800 frequencies are consistent with the values proposed in [60], which have been derived from worst-case considerations. SAR distributions and averaged spatial peak SAR are assessed and compared for the left application side "touch" and 15° "tilted" standard test positions, as described in [47, 59].

Fig. 14.37 provides grayscale visualizations of the measured and simulated SAR distributions in the SAM phantom at a distance of 4.7 mm normal to the inner shape of the shell. The scans are performed for the left-side "touch" position at $f = 1{,}747$ MHz, and are logarithmically scaled (i.e., 0 dB equals maximum peak experimental SAR value) for an antenna input power of +29.1 dBm. Good agreement is obtained between the measured and simulated distributions over the entire observation area.

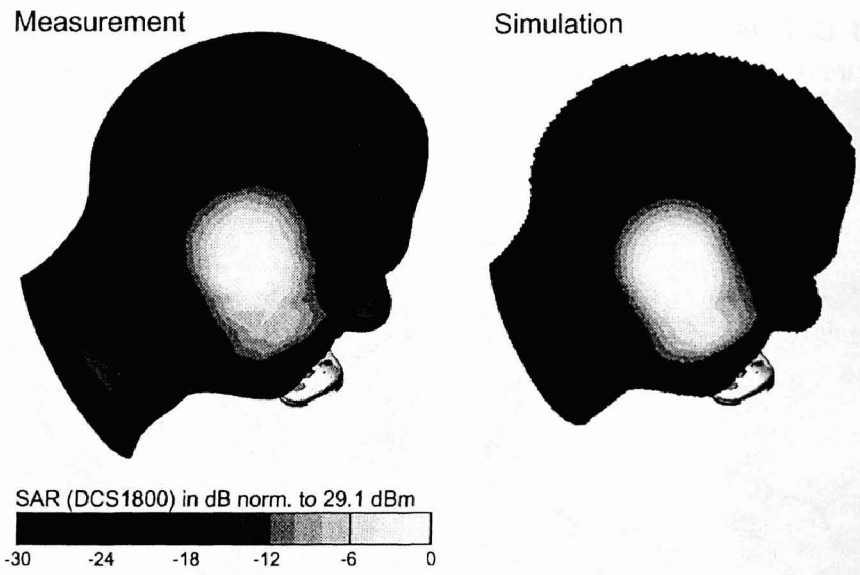


Fig. 14.37 Comparison of measured and FDTD-simulated SAR distributions within the SAM phantom (left-side "touch" position, GSM1800). *Source:* Chavannes et al., *IEEE Antennas and Propagation Magazine*, Dec. 2003, pp. 52–66, © 2003 IEEE.

Tables 14.1 and 14.2 list the measured and FDTD-calculated 1g-averaged and 10g-averaged peak SAR values for the T250 phone in the left-side "touch" and 15° "tilted" positions relative to the SAM phantom. Good-to-excellent agreement between the experimental and FDTD data is observed. In the "touch" position, the SAR values show little frequency dependence, whereas there is a strong frequency dependence of the SAR in the "tilted" position. The latter reveals the influence of the large currents close to the antenna.

Tables 14.1 and 14.2 also list FDTD calculations of the SAR for the T250 phone positioned adjacent to the HR-EF-1 MRI-based inhomogeneous human head model [61], visualized in Fig. 14.38. This model consists of a total of 121 slices (1 mm thick in the ear region; 3 mm thick elsewhere). Each slice has a transverse spatial resolution of 0.2 mm. A total of 15 tissue types are distinguished, with dielectric parameters at mobile phone frequencies taken from [62].

TABLE 14.1

1g-Averaged and 10g-Averaged SAR Values (in W/kg) for the T250 Phone in the Left-Side "Touch" Position for the SAM and MRI-Based Human-Head Phantoms. *Source:* Chavannes et al., *IEEE Antennas and Propagation Magazine*, Dec. 2003, pp. 52–66, © 2003 IEEE.

SAR Type	*SAM (Measured)*	*SAM (FDTD)*	*% Difference*	*Head Model (FDTD)*
902 MHz (1 g)	0.86	0.91	5.8%	0.77
1,747 MHz (1 g)	0.91	0.98	7.7%	0.80
902 MHz (10 g)	0.58	0.66	13.7%	0.58
1,747 MHz (10 g)	0.57	0.61	7.0%	0.35

TABLE 14.2

1g-Averaged and 10g-Averaged SAR Values (in W/kg) for the T250 Phone in the Left-Side "Tilted" Position for the SAM and MRI-Based Human-Head Phantoms. *Source:* Chavannes et al., *IEEE Antennas and Propagation Magazine*, Dec. 2003, pp. 52–66, © 2003 IEEE.

SAR Type	*SAM (Measured)*	*SAM (FDTD)*	*% Difference*	*Head Model (FDTD)*
902 MHz (1 g)	0.40	0.43	7.5%	0.40
1,747 MHz (1 g)	1.10	1.08	1.8%	0.73
902 MHz (10 g)	0.29	0.31	6.9%	0.36
1,747 MHz (10 g)	0.66	0.65	1.5%	0.32

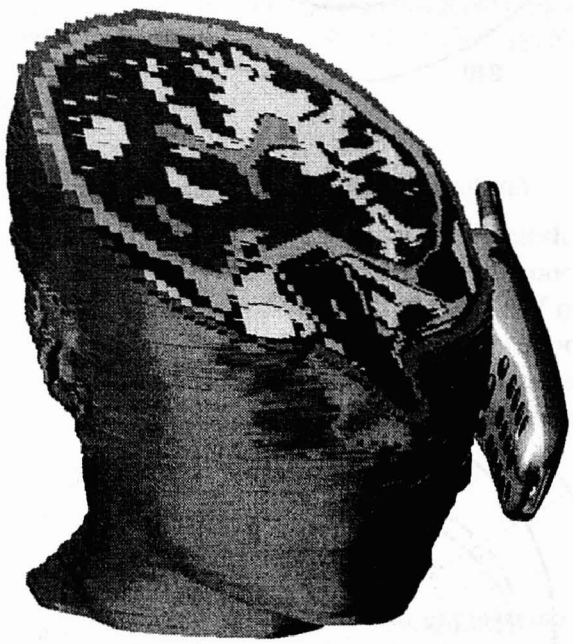
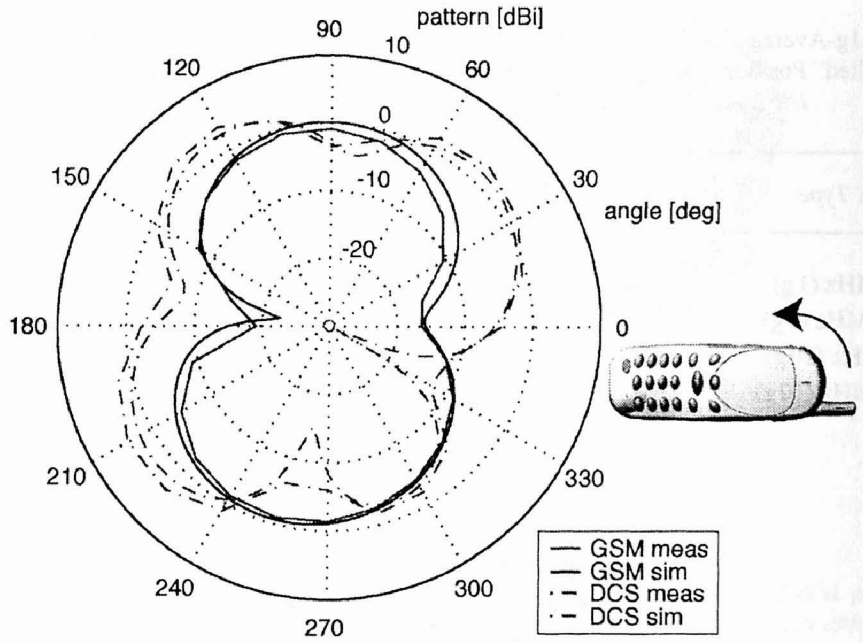


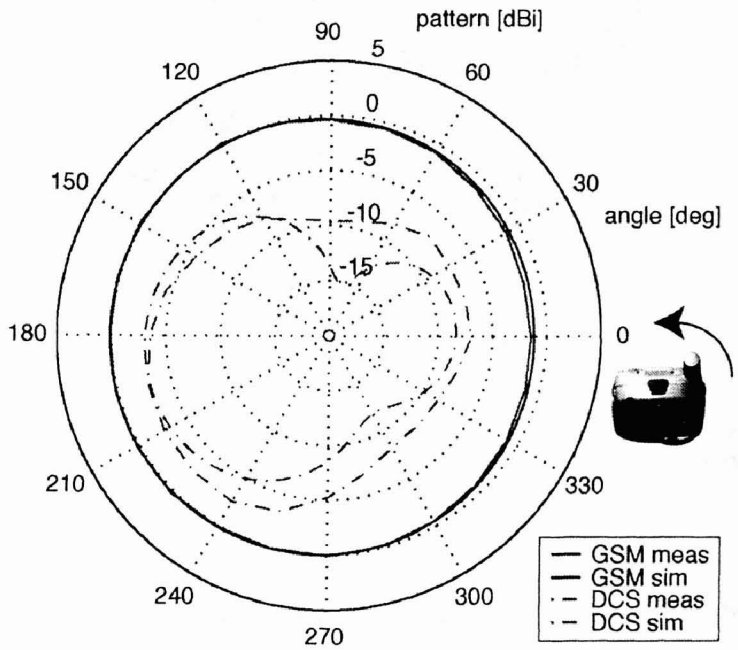
Fig. 14.38 T250 phone positioned adjacent to the HR-EF-1 MRI-based high-resolution European female inhomogeneous head model. *Source:* Chavannes et al., *IEEE Antennas and Propagation Magazine*, Dec. 2003, pp. 52–66, © 2003 IEEE.

14.9.5 Radiation Performance in Free Space and Adjacent to the SAM Head

Fig. 14.39 plots the measured and FDTD-calculated two-dimensional radiation patterns of the T250 phone located in free space. Good agreement is seen for both polarizations in the two observation cuts. Due to the smaller number of experimental data points, the field minima are not determined properly, leading to deviations in these regions. Fig. 14.40 displays the three-dimensional far-field patterns of the T250 phone placed in the "touch" position with the SAM phantom at the GSM900 and the DCS1800 frequencies.



(a) Horizontal polarization in the x-y plane.



(b) Vertical polarization in the x-z plane.

Fig. 14.39 Comparison of measured and FDTD-calculated free-space radiation patterns (in dBi) for the T250 phone in the GSM900 and DCS1800 bands. *Source:* Chavannes et al., *IEEE Antennas and Propagation Magazine*, Dec. 2003, pp. 52–66, © 2003 IEEE.

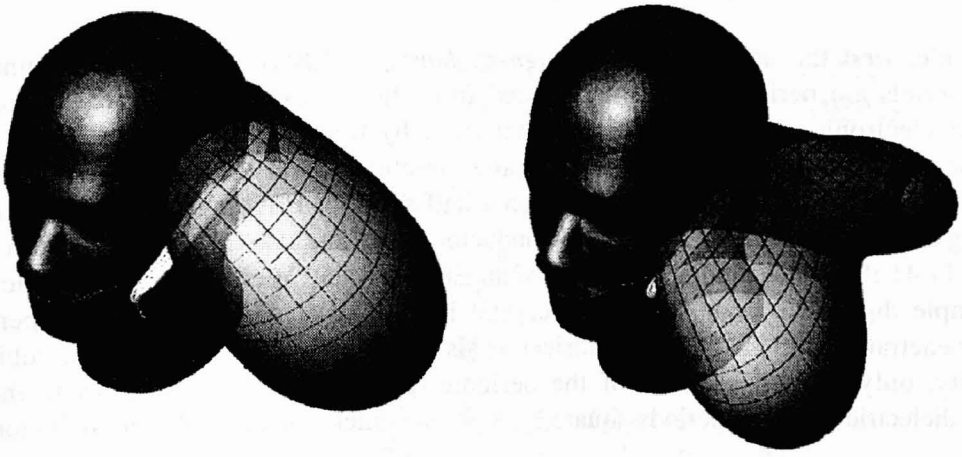


Fig. 14.40 Three-dimensional radiation pattern of the T250 phone when positioned adjacent to the SAM phantom operating in the GSM900 (left) and in the DCS1800 (right) band. *Source:* Chavannes et al., *IEEE Antennas and Propagation Magazine*, Dec. 2003, pp. 52–66, © 2003 IEEE.

14.9.6 Computational Requirements

Using the latest SSE-optimized SEMCAD FDTD solver with mesh grading and local mesh refinement on a 3.6-GHz Intel Pentium 4, the entire phone + head model is computed in less than two hours with a memory requirement of approximately 500 MB. If only mesh grading is used, the running time lengthens to 15 hours with more than 1 GB of memory needed. Finally, if one were to apply a simple uniform mesh, the running time extrapolates to weeks with hundreds of gigabytes of memory needed. The computational-resource advantages of mesh grading and local mesh refinement for this problem are very clear, indeed.

14.9.7 Overall Assessment

This case study demonstrates that FDTD is a very useful engineering tool for analysis, design, and optimization of complex mobile phones and similar personal wireless communications devices operating in the vicinity of the human body. This study also reveals that graded FDTD meshes combined with local mesh-refinement schemes permit a computationally efficient and accurate means to resolve small geometric features in the order of 100 μm located within compact subsets of an overall computational domain that can span in the order of 1 m.

14.10 SELECTED ADDITIONAL APPLICATIONS

This section reviews several additional applications of FDTD modeling in the analysis and design of antennas and antenna systems used in complex environments. These include: (1) a reflector antenna constructed using a dielectric electromagnetic bandgap structure instead of a metal sheet, (2) a ground-penetrating radar for detection of buried all-dielectric land mines, (3) analysis of antenna-radome interactions, and (4) analysis and design of antennas and antenna systems for diagnostic and therapeutic biomedical applications.

14.10.1 Use of Electromagnetic Bandgap Materials

We consider first the usage of *electromagnetic bandgap* (EBG) materials in antenna design. EBG materials are periodic structures formed from dielectrics. They are the electromagnetic analog of electronic crystals, and are characterized by frequency bands (bandgaps) in which electromagnetic waves cannot propagate in any direction within the material. At frequencies within the bandgap, a plane wave incident on a half-space of EBG material is totally reflected. In this regard, the material acts like a good conductor [63]. (See also Chapter 16, Section 16.8.)

Fig. 14.41 illustrates the FDTD model of an EBG "crystal" used as an all-dielectric reflector for a simple dipole antenna. The EBG crystal is a block of dielectric (relative permittivity $\varepsilon_r = 13$) penetrated with air-filled cylindrical voids at the sites of a body-centered cubic lattice. For clarity, only a small segment of the periodic reflector structure is shown in the figure. In fact, a dielectric block 21-periods-square by 3-periods-thick is used for the full reflector.

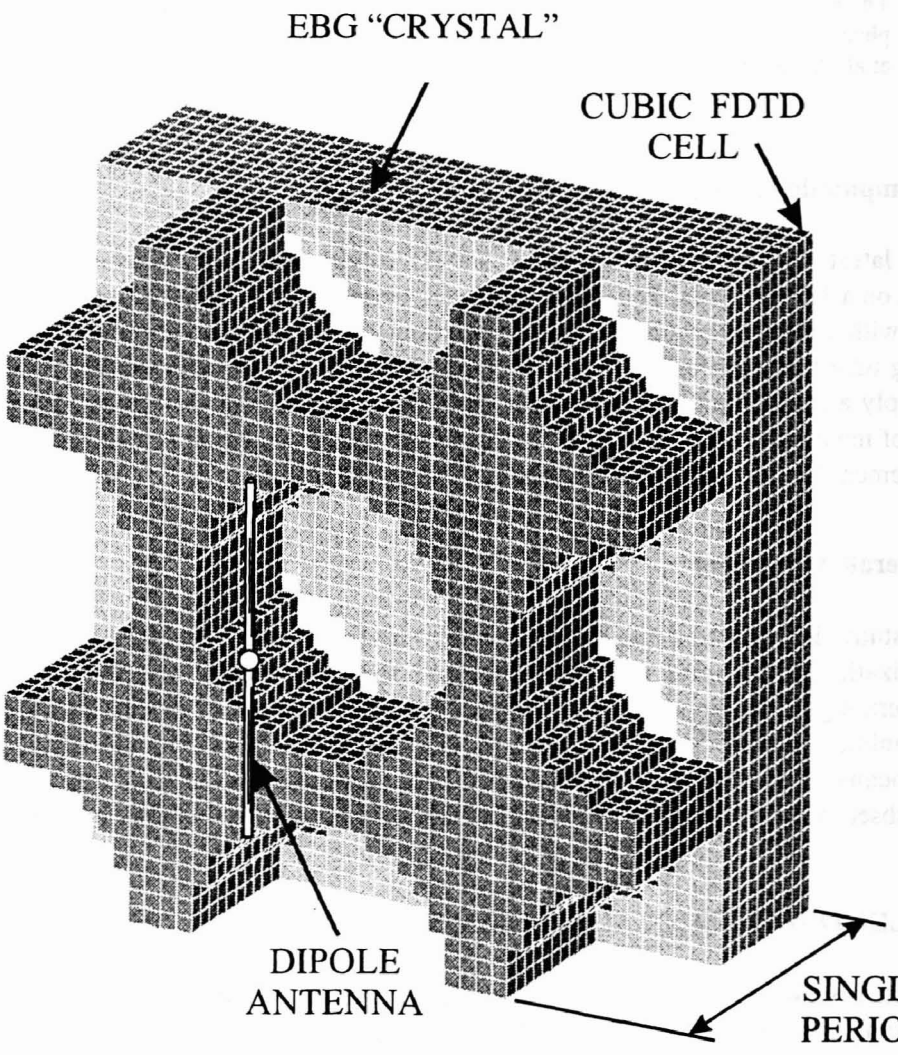


Fig. 14.41 FDTD model of an electromagnetic bandgap crystal used as an all-dielectric reflector for a dipole antenna.

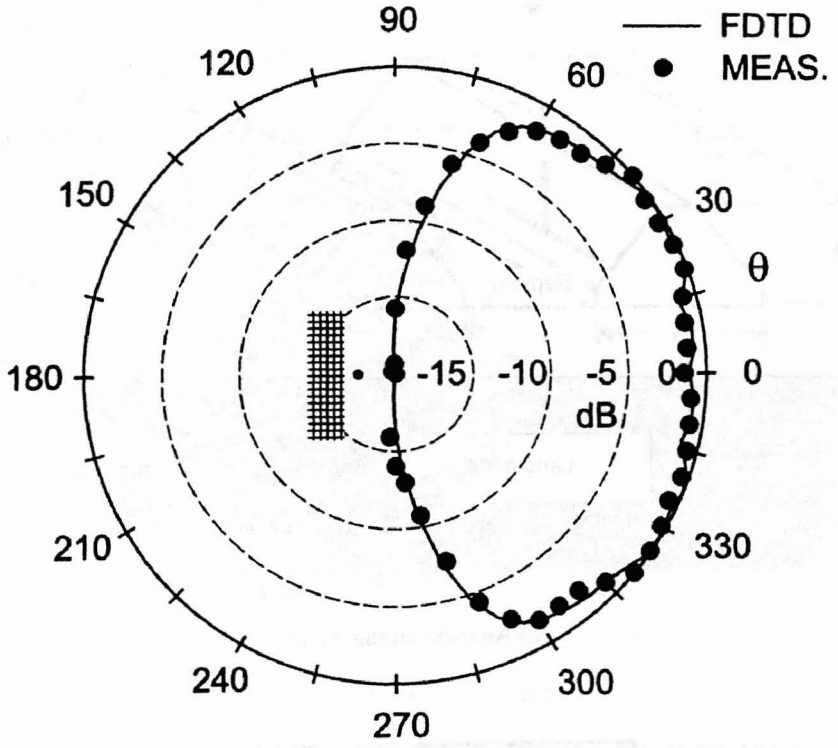
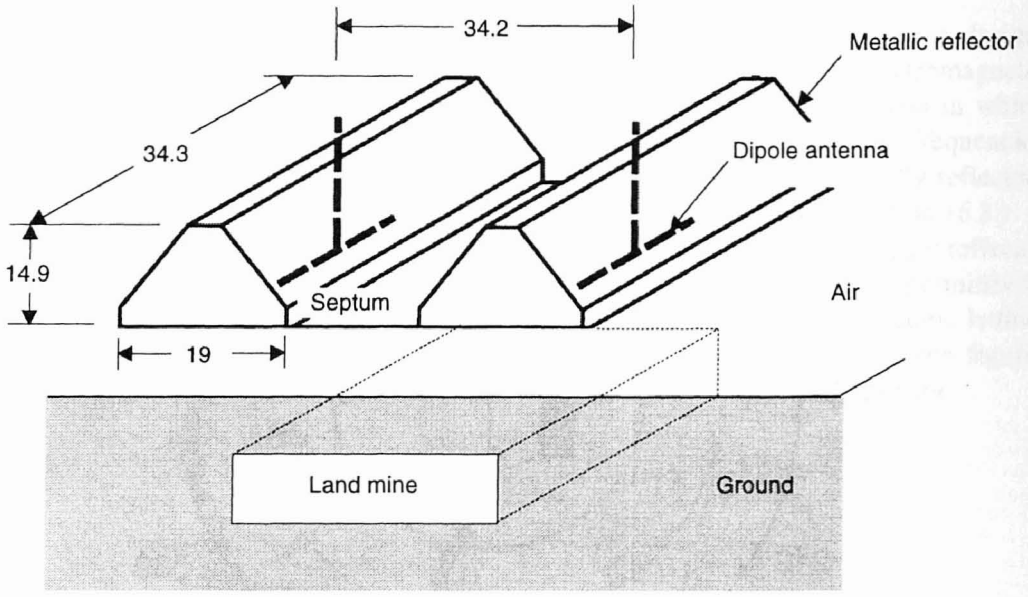


Fig. 14.42 Comparison of FDTD-calculated and measured far-field radiation patterns in the horizontal plane for the EBG reflector antenna of Fig. 14.41.

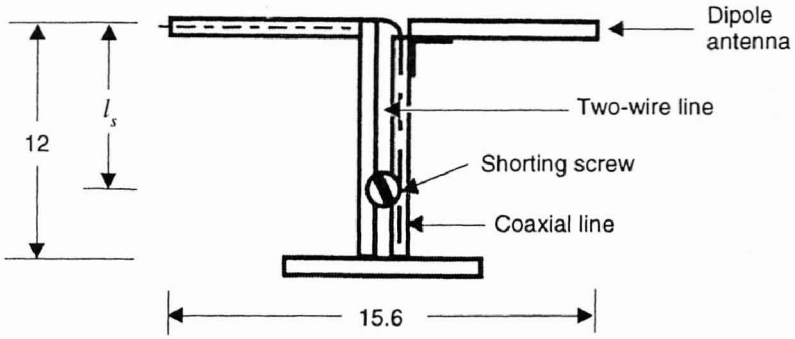
Fig. 14.42 compares the FDTD-computed and measured far-field radiation patterns in the horizontal plane of the antenna of Fig. 14.41 for an excitation frequency within the bandgap of the EBG crystal. The antenna pattern is maximum for azimuth angles pointing away from the EBG crystal, just as if it were metal. Very good agreement is noted between the FDTD predictions and the measured data.

14.10.2 Ground-Penetrating Radar

We next discuss a continuous-wave ground-penetrating radar designed for the detection of buried, all-dielectric land mines. Fig. 14.43(a) shows the arrangement of antennas for signal transmission and reception, which is designated as the separated-aperture sensor [31, 32, 64]. This sensor consists of two parallel dipole antennas, each contained in a metallic reflector. The dipoles are fed in the manner shown in Fig. 14.43(b), with the length of the dipoles and the position of the shorting screws on the tuning stubs adjusted to match the antennas at a specified frequency (typically 790 MHz). The coupling $|S_{21}|$ between the two antennas is measured as the detector is moved over the surface of the ground, and an increase in the coupling indicates the presence of a buried anomaly, such as a land mine.



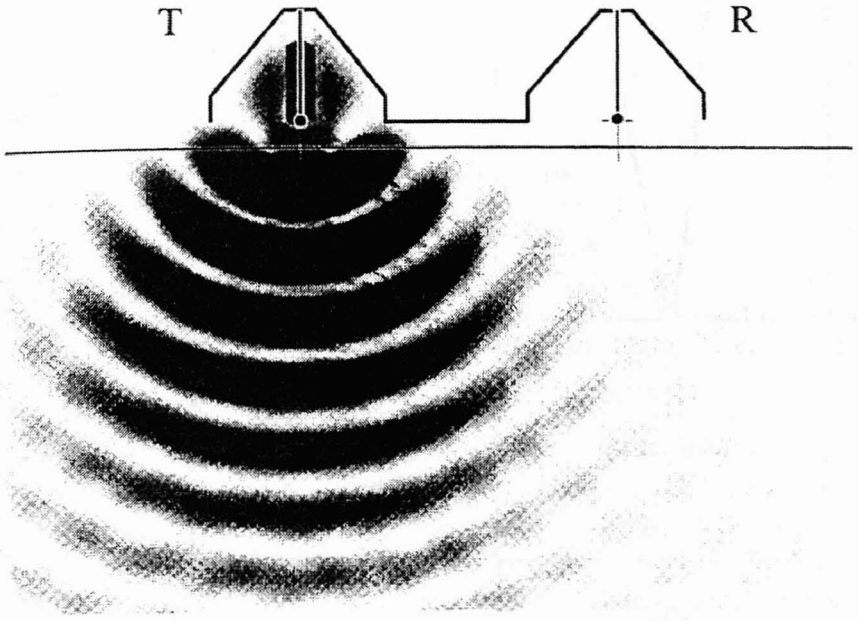
(a) Antenna arrangement.



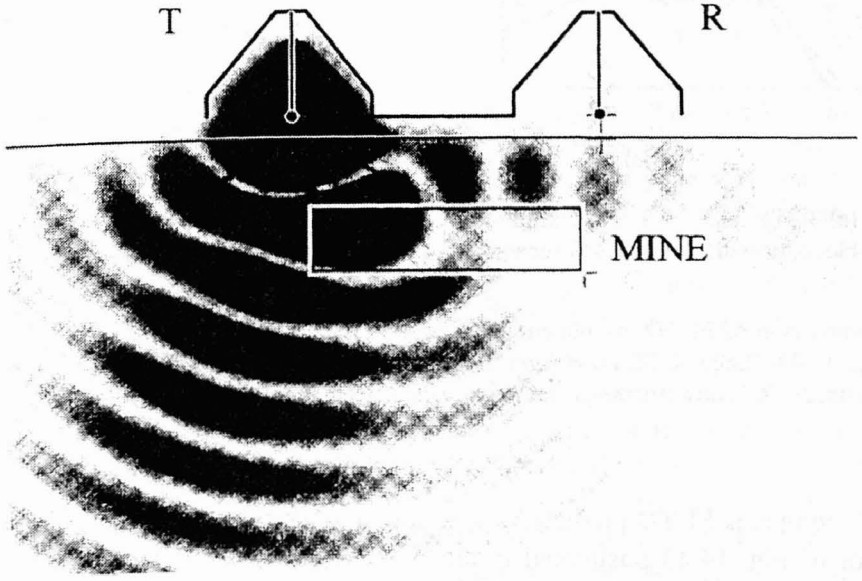
(b) Details of the feed for the dipole antennas.

Fig. 14.43 Separated-aperture sensor for detecting buried land mines. The sensor is designed for 790 MHz, and all dimensions are in centimeters. *Source:* Bourgeois and Smith, *IEEE Trans. Antennas and Propagation*, 1998, pp. 1419–1426, © 1998 IEEE.

The complete system, including two antennas with tuning stubs and reflectors, the half-space of soil, and a buried land mine, has been modeled using the FDTD method [31, 32, 64]. Fig. 14.44 uses grayscale plots to visualize the magnitude of the E-field on a plane through the center of the detector of Fig. 14.43. In Fig. 14.44(a), the detector is over empty ground. In Fig. 14.44(b), the detector is over ground containing an all-dielectric mine. The increase in the coupling between the transmitting antenna T and the receiving antenna R caused by the presence of the mine is clearly shown.

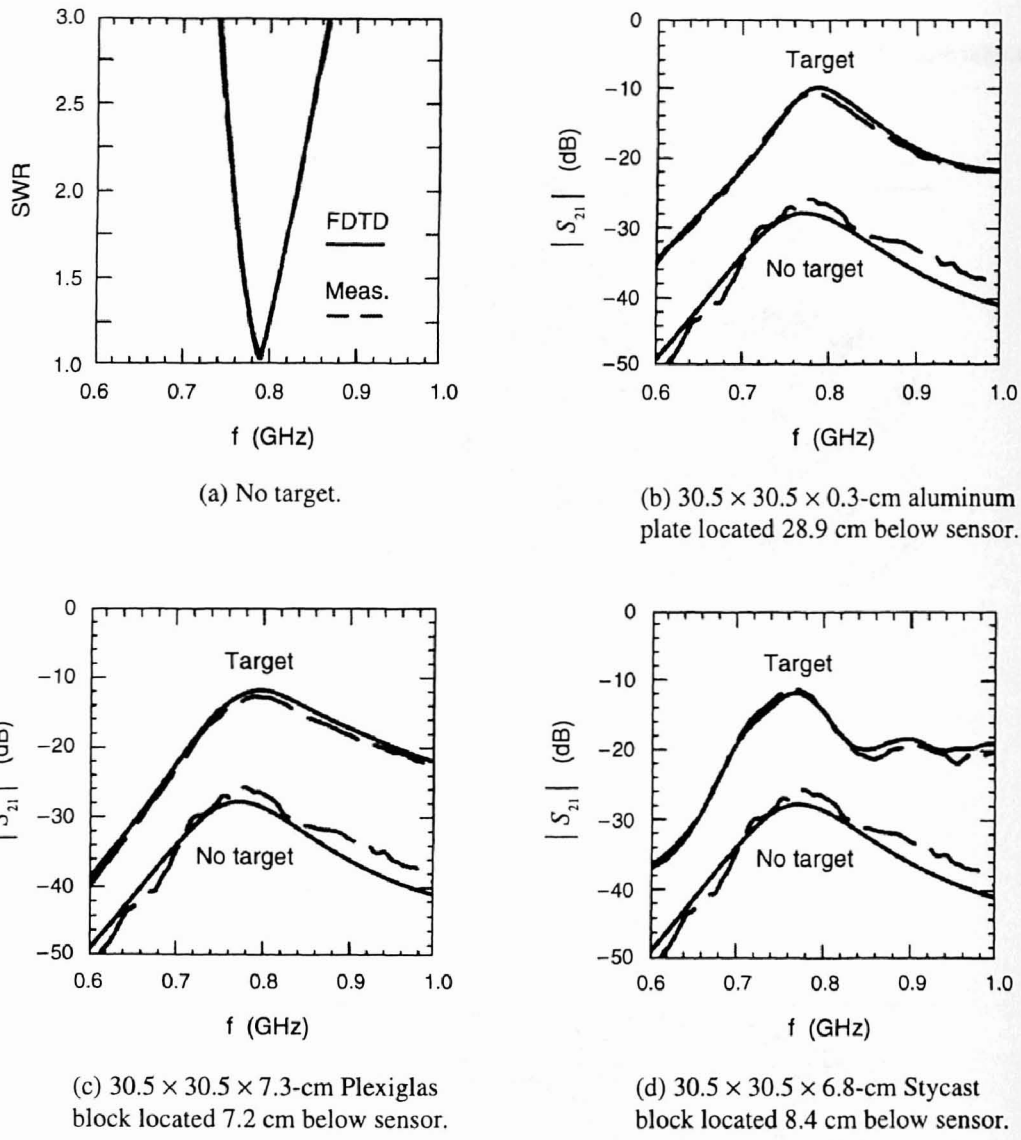


(a) Sensor over empty ground.



(b) Sensor centered over an all-dielectric land mine.

Fig. 14.44 Grayscale visualizations showing the magnitude of the *E*-field on the transverse symmetry plane of the separated-aperture sensor of Fig. 14.43. *Source:* J. M. Bourgeois, Ph.D. dissertation, Georgia Institute of Technology, Atlanta, GA, 1997.



(a) No target.

(b) 30.5 × 30.5 × 0.3-cm aluminum plate located 28.9 cm below sensor.

(c) 30.5 × 30.5 × 7.3-cm Plexiglas block located 7.2 cm below sensor.

(d) 30.5 × 30.5 × 6.8-cm Stycast block located 8.4 cm below sensor.

Fig. 14.45 Comparison of FDTD-calculated and measured responses of the separated-aperture sensor of Fig. 14.43. Legend: FDTD = solid line; measurements = dashed line. *Source:* Bourgeois and Smith, *IEEE Trans. Antennas and Propagation*, 1998, pp. 1419–1426, © 1998 IEEE.

Fig. 14.45 compares FDTD predictions with measurements for four cases of the separated-aperture sensor of Fig. 14.43 positioned in air 29.2 cm above a low-frequency RF absorber: (1) SWR for no target, (2) $|S_{21}|$ for a 30.5 × 30.5 × 0.3-cm aluminum plate located 28.9 cm below the sensor, (3) $|S_{21}|$ for a 30.5 × 30.5 × 7.3-cm Plexiglas block ($\varepsilon_r = 2.6$) located 7.2 cm below the sensor, and (4) $|S_{21}|$ for a 30.5 × 30.5 × 6.8-cm Stycast block ($\varepsilon_r = 7.4$) located 8.4 cm below the sensor. Each target is centered below the sensor, and positioned at a distance that yields the maximum $|S_{21}|$ at 790 MHz. Data are obtained over the 0.6 to 1.0 GHz range without retuning the sensor.

We see from Fig. 14.45 that the FDTD results are in good agreement with the measurements in all of the cases considered. More quantitatively, at 790 MHz, there is less than ±2 dB difference in the FDTD and test values of $|S_{21}|$. The overall good agreement shows that the FDTD simulation properly models key features of the sensor (including the dipoles, dipole feeds, stubs, and corner reflectors), as well as the target and the free-space medium beyond.

14.10.3 Antenna-Radome Interaction

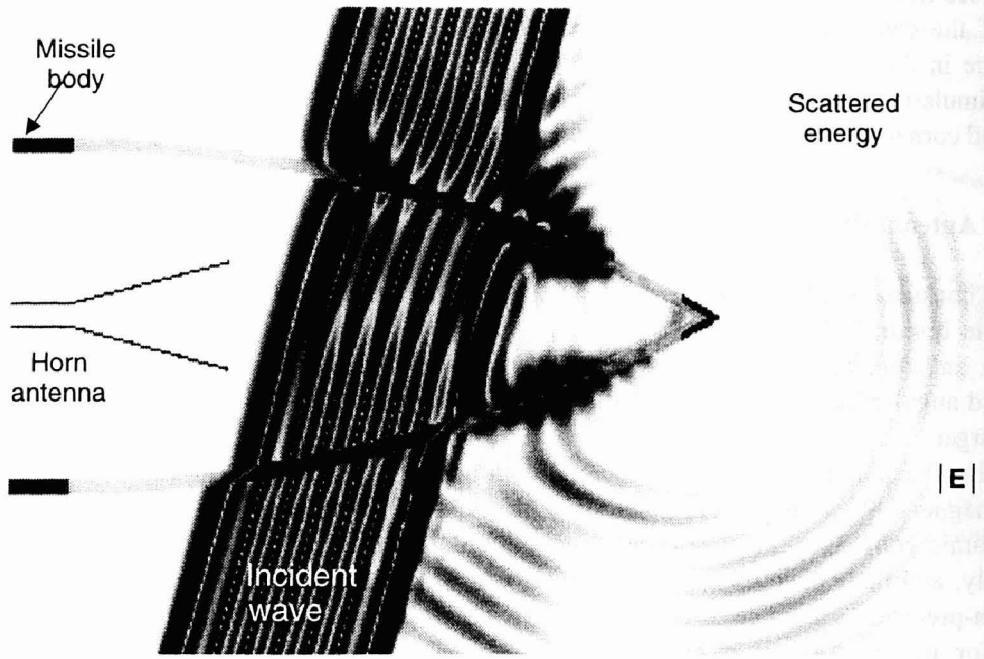
We next consider FDTD modeling of a long-standing defense-technology problem that can cause unreliable homing of radar-guided missiles: the complex electromagnetic wave interactions between an antenna and its protective radome. These interactions generate errors in the perceived angular location of a 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.

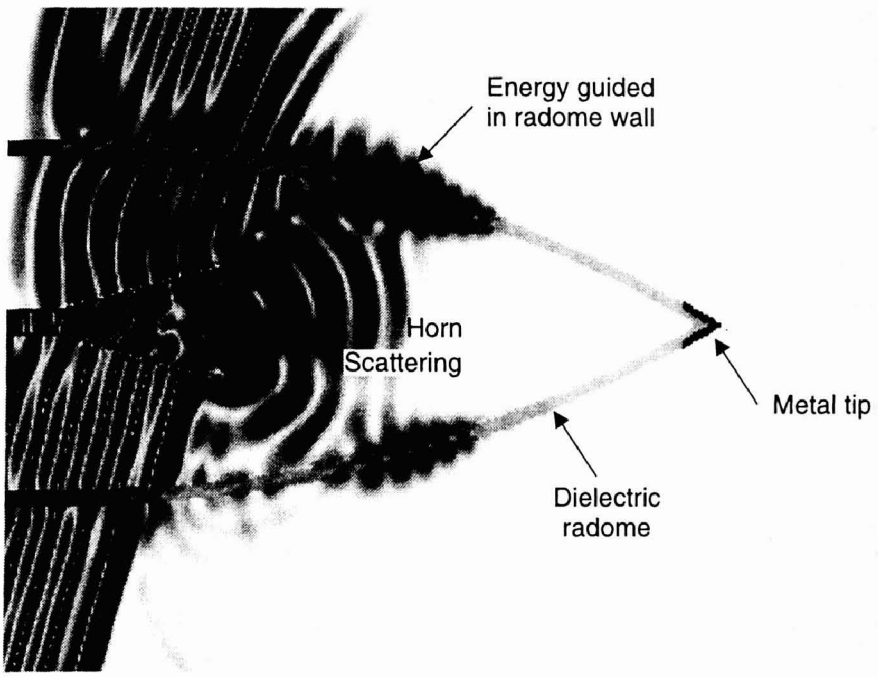
Referring to Figs. 14.13 and 14.46(a), we consider an X-band pyramidal horn antenna having the characteristic dimensions $a = 7.30$ cm, $b = 5.32$ cm, and $D = 8$ cm. The missile is modeled as a 17.8-cm-diameter, hollow, circular, PEC tube of wall thickness 7.5 mm. Its 33.0-cm-long nose section (the radome) is an ogive body of revolution formed from a lossless dielectric of permittivity $\varepsilon_r = 4$. 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 oblique incident plane-wave pulse propagating at 15° from boresight. The pulse consists of six sinusoidal cycles of a 10-GHz carrier. The FDTD model uses a uniform space lattice composed of cubic Yee cells, each spanning 1.27 mm ($\cong \lambda_0/24$ at 10 GHz). Figs. 14.46(a, b) visualize the magnitude of the E-field within and near the missile radome at two instants in time. (The corresponding color visualizations are shown in Fig. 1.4 of Chapter 1.) These graphics provide valuable insights into the complex interactions of the missile body, radome, and internal horn antenna.

Fig. 14.46(a) shows the incident pulse, propagating from right to left, after it has entered the radome. A quasispherical, 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 c inside the radome's dielectric wall is evident. In Fig. 14.46(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 that has been trapped inside the wall of the dielectric radome. The strong evanescent fields near the metal tip and the curvature of the radome allow the excitation of this trapped, guided wave. Subsequently, after the illumination passes the nose section of the missile, the trapped, guided mode identified in Fig. 14.46(b) encounters the junction between the dielectric radome and the PEC missile body. The trapped energy then reflects and reradiates. In addition, a new guided wave forms in the dielectric radome due to structural scattering from the horn antenna.



(a) Incident pulse has just entered the radome.



(b) Incident pulse has just encountered the horn antenna.

Fig. 14.46 Snapshot visualizations of the FDTD-computed interaction of a 10-GHz microwave pulse with a missile radome containing a horn antenna. The illuminating plane-wave pulse propagates from right to left at 15° from boresight.

14.10.4 Biomedical Applications of Antennas

Often in biomedical applications, the goal of the antenna system is to radiate into the human body. These applications fall into three broad categories: diagnostic imaging, medical implants, and thermal therapy. An excellent compilation of recent advances in these areas is provided in [65]. The ability of FDTD to model essentially arbitrary configurations of inhomogeneous, lossy dielectric materials makes this method broadly applicable in each of these three areas.

Diagnostic Imaging

Significant progress has been made over the past decade towards the clinical realization of noninvasive nonionizing microwave imaging techniques employing large arrays of sensing antennas and sophisticated inversion algorithms. In particular, the *ultrawideband* (UWB) microwave technology discussed in [66–71] has shown promise for early-stage breast cancer detection. As illustrated in Fig. 14.47 (and in the corresponding color visualizations of Fig. 1.3 of Chapter 1), detailed FDTD modeling indicates that UWB technology such as microwave imaging via space-time beamforming has the potential to detect malignant tumors as small as 2 mm embedded several centimeters deep within normal, inhomogeneous breast tissues [70]. These promising results have currently brought UWB technology to the point of initial preclinical investigations [71].

Antennas capable of transmitting and receiving short microwave pulses with high fidelity are a vital component of UWB microwave imaging systems for breast cancer detection. Recent work has used FDTD modeling to assist in the design of a variety of UWB antennas, including resistively loaded bowties and dipoles [67, 68], and ridged pyramidal horn antennas [71].

In the area of MRI, the frequency dependence of RF power deposition into patients being scanned is drawing increased attention with the use of systems operating at very high magnetic field strengths. Recent work has used FDTD modeling to analyze the relationship between the MRI frequency of operation and the RF power dissipated in the body for field strengths up to 8.5T [72]. Unlike previous studies, this analysis treats the coil and human head as a single system, models the actual excitation rather than an idealized current source, and utilizes a six-tissue anatomically detailed head model.

Medical Implants

In this area, novel antennas are being developed for telemetry and power links for retinal prostheses [73] and other medical implants such as cardiac pacemakers [74, 75]. Detailed FDTD modeling permits design of such antennas for very small size and optimized efficiency subject to the constraint of compatibility of the antenna materials with the surrounding biological tissues.

Thermal Therapy

The goal in radio-frequency and microwave hyperthermia cancer therapy is to configure a set of antennas around the human body so that they irradiate and focus electromagnetic energy onto a malignant tumor. The resulting tissue heating induces a tumor temperature that is substantially greater than the normal physiologic range. This heat either directly damages the malignant cells or renders them more vulnerable to damage by a concurrent program of chemotherapy or ionizing radiotherapy.

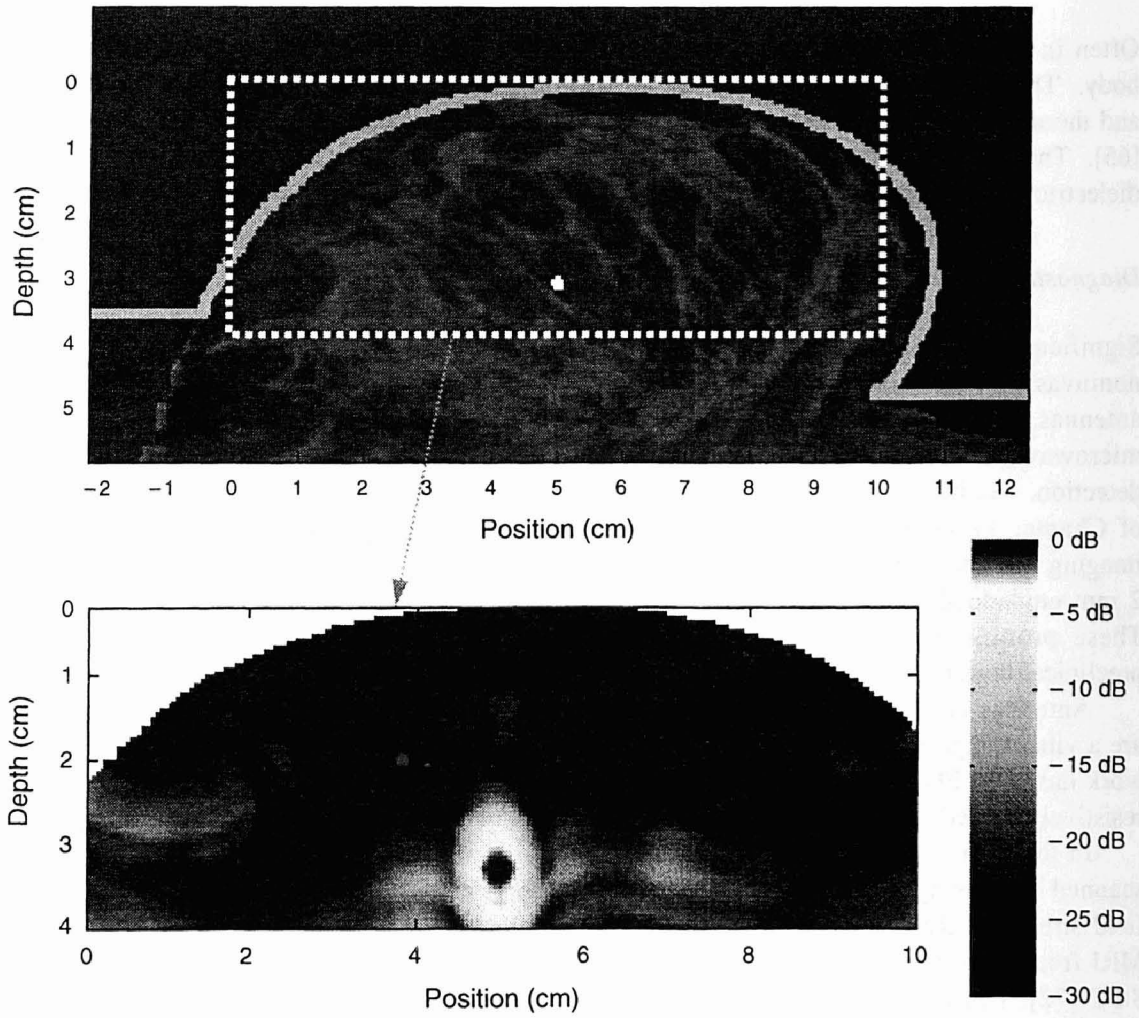


Fig. 14.47 Top: MRI-derived breast model shows the location of a simulated 2-mm diameter malignant breast tumor at a depth of approximately 3 cm. Bottom: Gray-scale rendition of the color image of the backscattered signal energy (decibel scale) created by the scanned output of a space-time beamformer applied to the FDTD-computed backscattered waveforms. Note that the tumor signature is 17 dB stronger than the backscattering clutter due to the surrounding normal tissues. *Adapted from:* Bond et al., *IEEE Trans. Antennas and Propagation*, 2003, pp. 1690–1705, © 2003 IEEE.

In conjunction with ionizing radiotherapy, hyperthermia has been shown to result in the regression of superficially located malignant tumors [76]. Annular phased arrays of multiple aperture antennas or dipoles placed symmetrically around the body have been designed to heat deeply embedded tumors by controlling the phases of the fields radiated from each of the antennas [77]. FDTD modeling has been used to optimize the phases and magnitudes of the excitations of the various radiating elements, using either a generic model of the human body [78] or patient-specific models [78–81]. FDTD modeling has also been used to explore and demonstrate the feasibility of UWB microwave hyperthermia treatment for breast cancer [82].

14.11 SUMMARY AND CONCLUSIONS

This chapter summarized the primary elements involved in FDTD modeling of transmitting and receiving antennas, namely:

- Formulation of the antenna problem, including the use of symmetry;
- Antenna feed models;
- Near-field to far-field transformations;
- The plane-wave source for the receiving antenna;
- Near-field simulations;
- Detailed case studies with experimental validations;
- Selected recent applications.

A key conclusion is that, when properly implemented, FDTD analyses of antennas produce results for near fields, far-field radiation patterns, receiving apertures, and driving-point impedances that agree very well with carefully performed experiments. FDTD has a powerful ability to provide, in a straightforward manner, wideband results for complex antenna structures comprised of, adjacent to, or even embedded within essentially arbitrary configurations of inhomogeneous materials. This robustness allows the use of the FDTD method to confidently test proposed novel antenna designs on the computer before they are built.

REFERENCES

[1] 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 Propagat.*, Vol. 38, 1990, pp. 1059–1068.

[2] Maloney, J. G., and G. S. Smith, "A study of transient radiation from the Wu-King resistive monopole — FDTD analysis and experimental measurements," *IEEE Trans. Antennas Propagat.*, Vol. 41, 1993, pp. 668–676.

[3] Montoya, T. P., and G. S. Smith, "A study of pulse radiation from several broad-band loaded monopoles," *IEEE Trans. Antennas Propagat.*, Vol. 44, 1996, pp. 1172–1182.

[4] Boonzaaier, J. J., and C. W. Pistorius, "Thin-wire Yagi antenna radiation patterns using the finite-difference time-domain method," *Microwave Optical Tech. Lett.*, Vol. 4, 1991, pp. 311–313.

[5] Luebbers, R., and K. Kunz, "Finite difference time domain calculations of antenna mutual coupling," *IEEE Trans. Electromagn. Compat.*, Vol. 34, 1992, pp. 357–359.

[6] Kashiwa, T., S. Tanaka, and I. Fukai, "Time-domain analysis of Yagi-Uda antennas using the FDTD method," *Electronics Comm. Japan, Part 1: Communications*, Vol. 77, 1994, pp. 96–105.

[7] Katz, D. S., M. J. Piket-May, A. Taflove, and K. R. Umashankar, "FDTD analysis of electromagnetic wave radiation from systems containing horn antennas," *IEEE Trans. Antennas Propagat.*, Vol. 39, 1991, pp. 1203–1212.

[8] Tirkas, P. A., and C. A. Balanis, "Contour path FDTD methods for analysis of pyramidal horns with composite inner *E*-plane walls," *IEEE Trans. Antennas Propagat.*, Vol. 42, 1994, pp. 1476–1483.

[9] Thiele, E., and A. Taflove, "FDTD analysis of Vivaldi flared horn antennas and arrays," *IEEE Trans. Antennas Propagat.*, Vol. 42, 1994, pp. 633–641.

[10] Shlager, K. L., G. S. Smith, and J. G. Maloney, "TEM horn antenna for pulse radiation: An improved design," *Microwave Optical Tech. Lett.*, Vol. 12, 1996, pp. 86–90.

[11] Shlager, K. L., G. S. Smith, and J. G. Maloney, "Accurate analysis of TEM horn antennas for pulse radiation," *IEEE Trans. Electromagn. Compat.*, Vol. 38, 1996, pp. 414–423.

[12] Reig, C., E. A. Navarro, and V. Such, "FDTD analysis of *E*-sectoral horn antennas for broad-band applications," *IEEE Trans. Antennas Propagat.*, Vol. 45, 1997, pp. 1484–1487.

[13] Reineix, A., and B. Jecko, "Analysis of microstrip patch antennas using finite difference time domain method," *IEEE Trans. Antennas Propagat.*, Vol. 37, 1989, pp. 1361–1369.

[14] Wu, C., K.-L. Wu, Z.-Q. Bi, and J. Litva, "Accurate characterization of planar printed antennas using finite-difference time-domain method," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 526–533.

[15] Reineix, A., and B. Jecko, "A time domain theoretical method for the analysis of microstrip antennas composed by slots," *Annales des Telecommunications*, Vol. 48, 1993, pp. 29–34.

[16] Penney, C. W., and R. J. Luebbers, "Input impedance, radiation pattern, and radar cross-section of spiral antennas using FDTD," *IEEE Trans. Antennas Propagat.*, Vol. 42, 1994, pp. 1328–1332.

[17] Shum, S. M., and K. M. Luk, "FDTD analysis of probe-fed cylindrical dielectric resonator antenna operating in fundamental broadside mode," *Electronics Lett.*, Vol. 31, 1995, pp. 1210–1212.

[18] Shum, S. M., and K. M. Luk, "Stacked annular ring dielectric resonator antenna excited by axi-symmetric coaxial probe," *IEEE Trans. Antennas Propagat.*, Vol. 43, 1995, pp. 889–892.

[19] Toland, B., C. C. Liu, and P. G. Ingerson, "Design and analysis of arbitrarily shaped dielectric antennas," *Microwave J.*, Vol. 40, 1997, pp. 278–286.

[20] Thomas, V. A., K. M. Ling, M. E. Jones, B. Toland, J. Lin, and T. Itoh, "FDTD analysis of an active antenna," *IEEE Microwave Guided Wave Lett.*, Vol. 4, 1994, pp. 296–298.

[21] Shlager, K. L., and J. B. Schneider, "A survey of the finite-difference time-domain literature," Ch. 1 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, A. Taflove, (ed.), Norwood, MA: Artech House, 1998.

[22] Shlager, K. L., and G. S. Smith, "Near-field to near-field transformation for use with FDTD method and its application to pulsed antenna problems," *Electronics Lett.*, Vol. 30, 1994, pp. 1262–1264.

[23] Shlager, K. L., and G. S. Smith, "Comparison of two near-field to near-field transformations applied to pulsed antenna problems," *Electronics Lett.*, Vol. 30, 1995, pp. 936–938.

[24] Reineix, A., J. Paillol, and B. Jecko, "FDTD method applied to the study of radar cross section of microstrip patch antennas," *Annales des Telecommunications*, Vol. 48, 1993, pp. 589–593.

[25] Maloney, J. G., and G. S. Smith, "Optimization of a conical antenna for pulse radiation: an efficient design using resistive loading," *IEEE Trans. Antennas Propagat.*, Vol. 41, 1993, pp. 940–947.

[26] Shlager, K. L., G. S. Smith, and J. G. Maloney, "Optimization of bow-tie antennas for pulse radiation," *IEEE Trans. Antennas Propagat.*, Vol. 42, 1994, pp. 975–982.

[27] Smith, G. S., *An Introduction to Classical Electromagnetic Radiation*, New York: Cambridge University Press, 1997.

[28] Maloney, J. G., K. L. Shlager, and G. S. Smith, "A simple FDTD model for transient excitation of antennas by transmission lines," *IEEE Trans. Antennas Propagat.*, Vol. 42, 1994, pp. 289–292.

[29] Maloney, J. G., *Analysis and Synthesis of Transient Antennas Using the Finite-Difference Time-Domain (FDTD) Method*, Ph.D. diss., Georgia Institute of Technology, Atlanta, GA, 1992.

[30] 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 Propagat.*, Vol. 35, 1987, pp. 1248–1257.

[31] Bourgeois, J. M., and G. S. Smith, "A fully three-dimensional simulation of a ground-penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geoscience Remote Sensing*, Vol. 34, 1996, pp. 36–44.

[32] Bourgeois, J. M., *A Complete Three-Dimensional Electromagnetic Simulation of Ground-Penetrating Radars Using the Finite-Difference Time-Domain Method*, Ph.D. diss., Georgia Institute of Technology, Atlanta, GA, 1997.

[33] Tirkas, P. A., and C. A. Balanis, "Finite-difference time-domain method for antenna radiation," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 334–340.

[34] Luebbers, R., L. Chen, T. Uno, and S. Adachi, "FDTD calculation of radiation patterns, impedance, and gain for a monopole antenna on a conducting box," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 1577–1583.

[35] Hertel, T. W., and G. S. Smith, "On the convergence of common FDTD feed models for antennas," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 1771–1779.

[36] Nye, J. F., "Theory of the directional pattern of a pyramidal horn," *IEEE Trans. Antennas Propagat.*, Vol. 45, 1997, pp. 1357–1365.

[37] Gentle, D., "Horn antenna measurements," private communication, National Phys. Lab., Teddington, Middlesex, U.K.

[38] Maloney, J. G., and G. S. Smith, "The reception of short pulses by antennas: FDTD results and reciprocity," in *Ultra-Wideband, Short-Pulse Electromagnetics 2*, L. Carin and L.B. Felsen, (eds.), Plenum Press, NY, 1995.

[39] Thiele, E., *FDTD Analysis of Vivaldi Flared Horn Antennas and Arrays*, Ph.D. diss., Northwestern University, Evanston, IL, 1994.

[40] Gibson, P. J., "The Vivaldi aerial," *Proc. 9th European Microwave Conf.*, 1979, pp. 101–105.

[41] Prasad, S. N., and S. Mahapatra, "A new MIC slot-line aerial," *IEEE Trans. Antennas Propagat.*, Vol. 31, 1983, pp. 525–527.

[42] Janaswamy, R., D. H. Schaubert, and D. M. Pozar, "Analysis of the transverse electromagnetic mode linearly tapered slot antenna," *Radio Science*, Vol. 21, 1986, pp. 797–804.

[43] Janaswamy, R., and D. H. Schaubert, "Analysis of the tapered slot antenna," *IEEE Trans. Antennas Propagat.*, Vol. 35, 1987, pp. 1058–1064.

[44] Janaswamy, R., "An accurate moment method model for the tapered slot antenna," *IEEE Trans. Antennas Propagat.*, Vol. 37, 1989, pp. 1523–1528.

[45] Stutzman, W. L., and G. A. Thiele, *Antenna Theory and Design*, New York: Wiley, 1981.

[46] U. S. Federal Communications Commission, *Evaluating Compliance with FCC Guidelines for Human Exposure to Radiofrequency Electromagnetic Fields*, Tech. Rep. Supplement C to OET Bulletin 65, Washington, D.C., 2001.

[47] CENELEC, EN 50360, *Product Standard to Demonstrate the Compliance of Mobile Telephones with the Basic Restrictions Related to Human Exposure to Electromagnetic Fields (300 MHz – 3 GHz)*, Brussels, 2001.

[48] CENELEC, EN 50361, *Basic Standard for the Measurement of Specific Absorption Rate Related to Human Exposure to Electromagnetic Fields from Mobile Phones (300 MHz – 3 GHz)*, Brussels, 2001.

[49] IEC PT 62209, *Procedure to Measure the Specific Absorption Rate (SAR) for Hand-Held Mobile Wireless Devices in the Frequency Range of 300 MHz to 3 GHz*, 2002.

[50] Dimbylow, P. J., and S. M. Mann, "SAR calculations in an anatomically based realistic model of the head for mobile communication transceivers at 900 MHz and 1.8 GHz," *Physics Medicine Biology*, Vol. 39, 1994, pp. 1537–1553.

[51] Burke, G. J., and A. J. Poggio, *Numerical Electromagnetics Code (NEC) – Method of Moments*, Lawrence Livermore Nat. Lab. Rpt. UCID–18834, Livermore, CA, 1981.

[52] King, R. W. P., "The electromagnetic field of a horizontal electric dipole in the presence of a three-layered region: Supplement," *J. Applied Physics*, Vol. 74, 1993, pp. 4845–4848.

[53] Gandhi, O. P., G. Lazzi, and C. M. Furse, "Electromagnetic absorption in the human head and neck for mobile telephones at 835 and 1,900 MHz," *IEEE Trans. Microwave Theory Tech.*, Vol. 44, 1996, pp. 1884–1897.

[54] Dhondt, G., and L. Martens, "A canonical case with an analytical solution for the comparison of electromagnetic field solvers," *Proc. COST 244 Meeting on Reference Models for Bioelectromagnetic Test of Mobile Communication Systems*, D. Simunic, (ed.), Rome, Italy, 1995, pp. 98–104.

[55] Yu, Q., O. P. Gandhi, M. Aronsson, and D. Wu, "An automated SAR measurement system for compliance testing of personal wireless devices," *IEEE Trans. Electromagn. Compat.*, Vol. 41, 1999, pp. 234–245.

[56] Chavannes, N., R. Tay, N. Nikoloski, and N. Kuster, "Suitability of FDTD-based TCAD tools for RF design of mobile phones," *IEEE Antennas Propagat. Magazine*, Vol. 45, Dec. 2003, pp. 52–66.

[57] Schmid, T., O. Egger, and N. Kuster, "Automated E-field scanning system for dosimetric assessments," *IEEE Trans. Microwave Theory and Techniques*, Vol. 44, 1996, pp. 105–113 (see also DASY Web site [online]: www.dasy4.com).

[58] SEMCAD Web site [online]: www.semcad.com.

[59] IEEE P1528/D1.2, *Recommended Practice for Determining the Spatial-Peak Specific Absorption Rate (SAR) in the Human Body Due to Wireless Communications Devices: Measurement Techniques*, 2003.

[60] Drossos, A., V. Santomaa, and N. Kuster, "The dependence of electromagnetic energy absorption upon human head tissue composition in the frequency range of 300–3000 MHz," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 1988–1995.

[61] Burkhardt, M., *Contributions Toward Uncertainty Assessments and Error Minimization of FDTD Simulations Involving Complex Dielectric Bodies*, Ph.D. diss, Diss. ETH Nr.13176, Zurich, 1999.

[62] Gabriel, C., *Compilation of the Dielectric Properties of Body Tissues at RF and Microwave Frequencies*, Tech. Rept. N.AL/OETR-1996-0037, Occupational and Environmental Health Directorate, Radiofrequency Radiation Division, Brooks AFB, TX, 1996.

[63] Kesler, M. P., J. G. Maloney, B. L. Shirley, and G. S. Smith, "Antenna design with the use of photonic bandgap materials as all-dielectric reflectors," *Microwave Optical Tech. Lett.*, Vol. 11, 1996, pp. 169–174.

[64] Bourgeois, J. M., and G. S. Smith, "A complete electromagnetic simulation of the separated-aperture sensor for detecting buried land mines," *IEEE Trans. Antennas Propagat.*, Vol. 46, 1998, pp. 1419–1426.

[65] Lazzi, G., O. P. Gandhi, and S. Ueno, (eds.), *IEEE Trans. Microwave Theory Tech.: Special Issue on Medical Applications and Biological Effects of RF/Microwaves*, Vol. 52, 2004, pp. 1853–2083.

[66] Hagness, S. C., A. Taflove, and J. E. Bridges, "Two-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: Fixed-focus and antenna-array sensors," *IEEE Trans. Biomedical Engrg.*, Vol. 45, 1998, pp. 1470–1479.

[67] Hagness, S. C., A. Taflove, and J. E. Bridges, "Three-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: Design of an antenna-array element," *IEEE Trans. Antennas Propagat.*, Vol. 47, 1999, pp. 783–791.

[68] Fear, E. C. and M. A. Stuchly, "Microwave detection of breast cancer," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 1854–1863.

[69] Li, X. and S. C. Hagness, "A confocal microwave imaging algorithm for breast cancer detection," *IEEE Microwave Wireless Comp. Lett.*, Vol. 11, 2001, pp. 130–132.

[70] Bond, E. J., X. Li, S. C. Hagness, and B. D. Van Veen, "Microwave imaging via space-time beamforming for early detection of breast cancer," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 1690–1705.

[71] Li, X., S. K. Davis, S. C. Hagness, D. W. van der Weide, and B. D. Van Veen, "Microwave imaging via space-time beamforming: Experimental investigation of tumor detection in multilayer breast phantoms," *IEEE Trans. Microwave Theory Tech.: Special Issue on Medical Applications and Biological Effects of RF/Microwaves*, Vol. 52, 2004, pp. 1856–1865.

[72] Ibrahim, T. S., "A numerical analysis of radio-frequency power requirements in magnetic resonance imaging experiment," *IEEE Trans. Microwave Theory Tech.: Special Issue on Medical Applications and Biological Effects of RF/Microwaves*, Vol. 52, 2004, pp. 1999–2003.

[73] Gosalia, K., G. Lazzi, and M. Humayun, "Investigation of a microwave data telemetry link for a retinal prosthesis," *IEEE Trans. Microwave Theory Tech.: Special Issue on Medical Applications and Biological Effects of RF/Microwaves*, Vol. 52, 2004, pp. 1925–1933.

[74] Soontornpipit, P., C. M. Furse, and Y. C. Chung, "Design of implantable microstrip antenna for communication with medical implants," *IEEE Trans. Microwave Theory Tech.: Special Issue on Medical Applications and Biological Effects of RF/Microwaves*, Vol. 52, 2004, pp. 1944–1951.

[75] Schmidt, S., and G. Lazzi, "Use of the FDTD thin-strut formalism for biomedical telemetry coil designs," *IEEE Trans. Microwave Theory Tech.: Special Issue on Medical Applications and Biological Effects of RF/Microwaves*, Vol. 52, 2004, pp. 1952–1956.

[76] Meyer, J. L., D. S. Kapp, P. Fessenden, and G. H. Hahn, "Hyperthermia oncology: Current biology, physics, and clinical results," *Pharmac. Ther.*, Vol. 42, 1989, pp. 251–288.

[77] Samulski, T. V., D. S. Kapp, P. Fessenden, and A. Lohrbach, "Heating deep-seated eccentrically located tumors with an annular phased array system: A comparative clinical study using two annular array operating configurations," *Intl. J. Radiation Oncology, Biology, Physics*, Vol. 13, 1987, pp. 83–94.

[78] Chen, J. Y., and O. P. Gandhi, "Numerical simulation of annular phased arrays of dipoles for hyperthermia of deep-seated tumors," *IEEE Trans. Biomedical Engrg.*, Vol. 39, 1992, pp. 209–216.

[79] James, B. J., and D. M. Sullivan, "Direct use of CT scans for hyperthermia treatment planning," *IEEE Trans. Biomedical Engrg.*, Vol. 39, 1992, pp. 845–851.

[80] Sullivan, D. M., R. Ben-Yosef, and D. S. Kapp, "Stanford 3D hyperthermia treatment planning system: technical review and clinical summary," *Int. J. Hyperthermia*, Vol. 9, 1993, pp. 627–643.

[81] Kowalski, M. E., and J.-M. Jin, "Model-based optimization of phased arrays for electromagnetic hyperthermia," *IEEE Trans. Microwave Theory Tech.: Special Issue on Medical Applications and Biological Effects of RF/Microwaves*, Vol. 52, 2004, pp. 1964–1977.

[82] Converse, M., E. J. Bond, S. C. Hagness, and B. D. Van Veen, "Ultrawideband microwave space-time beamforming for hyperthermia treatment of breast cancer: A computational feasibility study," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, 2004, pp. 1876–1889.

PROJECTS

P14.1 Implement the improved simple feed model of Section 14.3.3 to excite a monopole over a ground plane in a three-dimensional Cartesian FDTD space lattice. Replicate the feedline time-domain waveforms of Figs. 14.9 and 14.10.

P14.2 Model the standard-gain horn antenna of Section 14.6 using a space increment as coarse as $\lambda_0/10$ to keep the computer-resource burden tractable. Replicate the results for the far-field radiation patterns in Fig. 14.14, and for the boresight gain and boresight effective area versus frequency in Fig. 14.15.

P14.3 Model the Vivaldi single-flare baseline element of Fig. 14.17. Replicate the far-field radiation pattern in Fig. 14.18.

P14.4 Model the linearly tapered slot antenna considered in Fig. 14.19, and replicate the far-field radiation pattern shown in this figure.

P14.5 Model the generic cellphone geometry of Fig. 14.28, and replicate the near-field distributions shown in this figure.

P14.6 Model the dipole and layered half-space geometry shown in Fig. 14.29, and replicate the near-field distribution shown in this figure.

P14.7 Model the dipole and spherical brain phantom geometry considered in Fig. 14.32, and replicate the SAR distributions shown in this figure.

Chapter 15

High-Speed Electronic Circuits with Active and Nonlinear Components

Melinda Piket-May, Wojciech Gwarek, Tzong-Lin Wu, Bijan Houshmand, Tatsuo Itoh, and Jamesina Simpson

15.1 INTRODUCTION

High-speed electronic circuits are traditionally grouped into two classes: analog microwave circuits and digital logic circuits.

1. Microwave circuits typically process bandpass signals above 3 GHz. Common passive circuit features include waveguides, microstrips, couplers, circulators, and filters. Active devices such as transistors perform as amplifiers or oscillators. Small-signal (linear) operation is usually characterized by S parameters, whereas large-signal (nonlinear) operation is usually characterized by specifications for gain compression and generation of harmonic and intermodulation products. Designers of microwave circuits account for electromagnetic wave effects as a matter of course, since the circuit operation is based upon these effects.

2. Digital circuits typically process lowpass pulses having clock rates below 3 GHz. Typical circuits include densely packed, multiple planes of metal traces providing flow paths for the signals, dc power feeds, and ground returns. Via pins provide electrical connections between the planes. Designers may not always account for electromagnetic wave effects since the circuit operation is not based upon these effects. In fact, electromagnetic phenomena for such circuits are considered to be parasites.

However, the distinction between the design and packaging of these two classes of high-speed electronic circuits is blurring. Microwave circuits are becoming very complex systems composed of densely spaced elements, discontinuity structures, and passive and active devices. On the digital-circuit side, the rise of everyday clock speeds to 3 GHz implies lowpass signal bandwidths up to about 15 GHz, well into the microwave range. Electromagnetic wave effects that until now were in the domain of the microwave engineer are becoming a limiting factor to digital computer performance. At the chip level or the level of a device comprised of several interconnected chips, these effects are very complex, usually not well understood, and not trivial to analyze either theoretically, numerically, or experimentally. Hard-won experience has shown that the following electromagnetic-wave-related problems exist for high-speed digital circuits:

- Coupling between vias can distort signals, and mismatches between vias and signal lines can lead to ground bounce.
- Holes and other discontinuities in ground planes can increase coupling between the circuit-board layers.
- Metal traces (with or without bends) are likely to have reactive impedance components that can degrade system performance at high clock speeds.
- Signals can couple (crosstalk) from one parallel trace to another.
- Manufacturing tolerances can cause a range of coupling, crosstalk, and impedance parameters.
- Electromagnetic interference and compatibility problems can arise relative to other circuits and systems.

A key bottom line here is the *signal integrity*, a measure of how good the signal quality is at some digital circuit output relative to the specific design goals. Common signal-integrity specifications include rise time, threshold, and ringing.

Overall, the design of complex high-speed circuits, especially those having nonlinear active devices, encounters substantial problems in dealing with electromagnetic coupling and radiation effects between different circuit elements. Until recently, the focus of electromagnetic simulators has been the analysis of passive structures [1, 2]. Design tools have been based on an approach wherein circuits are first divided into small elements, and then the characteristics of each element are cascaded to obtain the overall system performance [3, 4]. Software packages such as Touchstone and MDS use equivalent-circuit models to represent the passive part of the circuits, and then the complete circuit, including the lumped elements, is analyzed using a circuit simulator. The *High-Frequency Structure Simulator* (HFSS) provides a complete electromagnetic solution based on the finite-element method [5].

Recent progress in the FDTD method, including the development of the PML ABC and information-extraction techniques, has made it possible to efficiently analyze realistic high-speed circuits in their complete form using full-vector-field computations [6]. FDTD algorithms now allow the accurate calculation of S parameters, radiation patterns, and coupling coefficients over a very broad frequency band with one run. In addition, visualization of the time-domain data can provide practical insights into the operation of a particular structure.

The goal of this chapter is to communicate this progress, specifically in the FDTD modeling of the electromagnetic wave properties of high-speed circuits that are constructed using primarily stripline and microstrip technology. We consider both passive interconnects and circuit devices. We show how to extract lumped-circuit equivalences and S parameters, and how to embed linear and nonlinear lumped circuit elements within the FDTD space lattice. In the latter area, we also show how Norton's and Thevenin's equivalent circuits can be derived for the FDTD grid, and used to connect the distributed, full-wave physics of Maxwell's equations to the circuit behavior of potentially complex electronic networks. Sample computed results are summarized that show very good accuracy for this extended FDTD algorithm, and the ability to model cases where available circuit simulators cannot be applied. Here, examples include nonlinear cases involving the generation of harmonics and intermodulation products in large-signal microwave circuits. Finally, as a potential means to mitigate the problems of using microstrip / stripline technology in future digital systems clocked faster than10 GHz, we discuss the possibility of utilizing *wireless* microwave / millimeter-wave electromagnetic bandgap structures as interconnects.

15.2 BASIC CIRCUIT PARAMETERS FOR TEM STRIPLINES AND MICROSTRIPS

FDTD modeling provides full-vector $\boldsymbol{E}$ and $\boldsymbol{H}$ distributions in time and space. For metallic stripline and microstrip systems with signal propagation known to be primarily in the *transverse electromagnetic* (TEM) mode, the following fundamental integral expressions can be used to connect the $\boldsymbol{E}$ and $\boldsymbol{H}$ field distributions to the usual circuit quantities of voltage and current:

$$V(t, x_i) = \int_{C_V} \boldsymbol{E}(t, x_i) \cdot d\boldsymbol{l} \quad ; \qquad I(t, x_i) = \oint_{C_I} \boldsymbol{H}(t, x_i) \cdot d\boldsymbol{l} \tag{15.1a, b}$$

Here, C_V is a contour extending from a defined voltage reference point (usually a ground plane) to the metal strip conductor at the adjacent location x_i. $V(t, x_i)$ is independent of the choice of C_V if this contour is confined to the transverse plane. Usually, C_V is conveniently chosen to extend in a perpendicular manner from the ground plane to the adjacent surface of the strip conductor. To provide the local current, contour C_I is chosen to wrap completely around the strip conductor in the transverse plane, as close as possible to the surface of the conductor. Care must be exercised in applying (15.1) if there is evidence that the circuit trace of interest is propagating non-TEM modes. In this case, the accuracy of this (V, I) formulation can degrade rapidly.

15.2.1 Transmission Line Parameters

Wideband frequency-domain transmission line parameters can be found by applying the Fourier transform to the voltage and current responses of (15.1) for an impulsive excitation of the line [7]. For example, the line characteristic impedance as a function of frequency is obtained from

$$Z_0(\omega, x_i) = \mathcal{F}[V_{inc}(t, x_i)] \,/\, \mathcal{F}[I_{inc}(t, x_i)] \tag{15.2}$$

where $\mathcal{F}[\ ]$ is the Fourier transform operator, and V_{inc} (I_{inc}) is the incident voltage (current) observed at x_i. Further, if $g(t, x_i)$ denotes an incident voltage or current waveform at $x = x_i$, and $g(t, x_j)$ is the corresponding waveform at $x = x_j$, then the propagation constant γ is obtained from

$$\mathcal{F}[g(t, x_j)] = \mathcal{F}[g(t, x_i)]\, e^{-\gamma(\omega) d} \tag{15.3}$$

where $d = x_j - x_i$. Rearranging this expression provides γ as a function of frequency:

$$\gamma(\omega) = \frac{1}{d} \ln\left\{ \frac{\mathcal{F}[g(t, x_i)]}{\mathcal{F}[g(t, x_j)]} \right\} \tag{15.4}$$

Defining $\gamma(\omega) = \alpha(\omega) + j\beta(\omega)$, the group velocity $v_g(\omega)$ is calculated as

$$v_g(\omega) = \left[\frac{\partial \beta(\omega)}{\partial \omega} \right]^{-1} \tag{15.5}$$

15.2.2 Impedance

For an impulsive excitation, the wideband reflection coefficient from a given transmission line load is calculated using

$$\Gamma(\omega, x_i) = \frac{\mathcal{F}\left[V_{\text{ref}}(t, x_i)\right]}{\mathcal{F}\left[V_{\text{inc}}(t, x_i)\right]} \tag{15.6}$$

where V_{ref} is the reflected voltage, and V_{inc} is the incident voltage observed at x_i. This reflection coefficient is subsequently transformed to the plane of the load via

$$\Gamma_{\text{L}}(\omega) = \Gamma(\omega, x_i)\, e^{2\gamma(\omega)\ell} \tag{15.7}$$

where ℓ is the distance from x_i to the load. The effective load impedance is then

$$Z_{\text{L}}(\omega) = Z_0(\omega, x_i)\left[\frac{1 + \Gamma_{\text{L}}(\omega)}{1 - \Gamma_{\text{L}}(\omega)}\right] \tag{15.8}$$

where $Z_{\text{L}}(\omega)$, $Z_0(\omega, x_i)$, and $\Gamma_{\text{L}}(\omega)$ are all complex values.

The above derivation does not take into account the fact that the V and I values derived from corresponding E- and H-fields in the FDTD space lattice are offset from each other by one-half space cell and one-half time-step. Ignoring these offsets can lead to errors for certain geometries. Reference [8] reported a simple interpolation that permits the V and I data used in the impedance calculation to be at the same space-time point. For a voltage $V|_i$ and adjacent currents $I|_{i+1/2}$ and $I|_{i-1/2}$, the desired spatially interpolated current is given by

$$I_{\text{interpolated}}\big|_i = 0.5\left(I|_{i+1/2} + I|_{i-1/2}\right) \tag{15.9}$$

Further, the one-half time-step offset can be accounted for by multiplying $V|_i$ by a factor of $e^{+j\omega\Delta t/2}$. For example, (15.2) for the characteristic impedance now becomes

$$Z_0(\omega, x_i) = \mathcal{F}\left[V_{\text{inc}}(t, x_i)\right] e^{+j\omega\Delta t/2} \Big/ \mathcal{F}\left\{0.5\left[I_{\text{inc}}(t, x_{i+1/2}) + I_{\text{inc}}(t, x_{i-1/2})\right]\right\} \tag{15.10}$$

15.2.3 *S*-Parameters

Given a multiport network, wideband, complex-valued scattering parameters S_{mn} can be obtained for an impulsive excitation as follows [7, 9]:

$$S_{mn}(\omega, x_m, x_n) = \frac{\tilde{V}_m(\omega, x_m)}{\tilde{V}_n(\omega, x_n)} \sqrt{\frac{Z_{0,n}(\omega)}{Z_{0,m}(\omega)}} \tag{15.11}$$

where $\tilde{V}_m(\omega, x_m)$ is the phasor voltage at port m at observation plane x_m, $\tilde{V}_n(\omega, x_n)$ is the phasor voltage at port n at observation plane x_n, and $Z_{0,m}$ and $Z_{0,n}$ are the characteristic impedances of the lines connected to these ports. For example, to obtain S_{11}, the incident and reflected pulses at port 1 must be known. To obtain S_{21}, we must observe the transmitted pulse emerging at port 2, corresponding to the known incident pulse at port 1.

In most cases, the magnitudes of the S-parameters are the primary data used by engineers to characterize, for example, the filtering properties of a network. The magnitude data are independent of the observation positions on the transmission lines feeding the corresponding ports, assuming that the feeding lines are either infinitely long or matched at their far ends. However, the phases of the S-parameters are clearly a function of the positions of the observation planes. As discussed later, phase data can be important when extracting a lumped-circuit equivalent network from the observed S-parameter variation with frequency [10].

15.2.4 Differential Capacitance

The capacitance per unit length of a TEM stripline or microstrip can be calculated indirectly using (15.2) and (15.5) as

$$C(\omega) = 1/v_g(\omega) Z_0(\omega) \tag{15.12}$$

Of more interest is the possibility that this capacitance can be calculated directly from the E-field that FDTD simulations provide. Given a designated incremental $\Delta x\,\Delta y$ surface patch ΔA on the bottom of a circuit trace (facing the ground plane), Gauss's law can be used to calculate the electric charge ΔQ on this patch from the normal component of $\boldsymbol{E}$ originating from the patch:

$$\Delta Q = \oiint_{S_E} \boldsymbol{D} \cdot d\hat{\boldsymbol{S}}_E \tag{15.13}$$

Here, $\boldsymbol{D} = \varepsilon \boldsymbol{E}$, and S_E is a virtual surface positioned parallel to the trace, between ΔA and the ground plane, that captures all of the electric flux emanating from ΔA. Combined with the voltage expression of (15.1a), the incremental trace capacitance (in farads/meter) is given by the ratio of the charge on ΔA to the voltage:

$$\frac{\Delta C}{\Delta x} = \frac{\Delta Q}{V} \tag{15.14}$$

This approaches the true differential capacitance per unit length in the limit that $\Delta x << \lambda$ for the smallest relevant wavelength λ present on the trace. In this quasistatic limit, $\Delta C/\Delta x$ is independent of the precise nature of ΔA, S_E, and the excitation.

Note that this procedure lends itself to calculating the capacitance between two surface regions of any shape. Patch ΔA could be defined to wrap around the entire circuit trace. The conductor at the other end of the integral path used to define the voltage (i.e., the other node of the capacitor) could be an adjacent trace rather than the ground plane. Therefore, this electric flux collection method can directly provide the mutual capacitance for subsequent SPICE calculations of crosstalk between traces.

15.2.5 Differential Inductance

In a manner analogous to that discussed above, the differential inductance of a TEM circuit trace geometry can be calculated directly from the vector magnetic field that FDTD simulations provide. It is well known that the self and/or mutual inductance of any physical structure is entirely determined by the geometric relationship between the electric circuit that transports the current I, and the surface S_M through which the magnetic flux generated by I penetrates. In fact, inductance L is defined as the ratio of the magnetic flux Φ penetrating S_M to its generating I:

$$L = \frac{\Phi}{I} = \frac{\iint_{S_M} \boldsymbol{B} \cdot d\hat{\boldsymbol{S}}_M}{I} \tag{15.15}$$

To obtain inductance from the full-wave vector electromagnetic field, properly defining the magnetic-flux-collection virtual surface S_M is crucial. As shown in Fig. 15.1(a), the perimeter of S_M for circuit trace 1 is bounded on the top by trace 1 carrying the signal current I_1, and on the bottom by the ground plane carrying the return (image) current i_1. The incremental self-inductance of trace 1 (in henrys/meter) is given by

$$\frac{\Delta L_{11}}{\Delta x} = \frac{\Delta \Phi_{11}}{I_1} = \frac{\iint_{S_M} \boldsymbol{B}_1 \cdot d\hat{\boldsymbol{S}}_M}{I_1} \tag{15.16}$$

where $\boldsymbol{B}_1$ is the magnetic flux density vector generated by I_1. The incremental mutual inductance of trace 1 relative to the adjacent trace 2 is found by obtaining the magnetic flux penetrating S_M due to current I_2 flowing on trace 2:

$$\frac{\Delta L_{21}}{\Delta x} = \frac{\Delta \Phi_{21}}{I_2} = \frac{\iint_{S_M} \boldsymbol{B}_2 \cdot d\hat{\boldsymbol{S}}_M}{I_2} \tag{15.17}$$

where $\boldsymbol{B}_2$ is the magnetic flux density vector generated by I_2. The inductance values given in (15.16) and (15.17) approach the true differential inductance per unit length in the limit that $\Delta x << \lambda$ for the smallest relevant wavelength λ present on the trace. In this quasistatic limit, $\Delta L/\Delta x$ is independent of the precise nature of S_M and the excitation of the trace.

15.3 LUMPED INDUCTANCE DUE TO A DISCONTINUITY

A key element of current engineering practice in the design of high-speed circuits is the development of lumped-circuit equivalences (especially inductances) for discontinuities in the signal and ground return paths. The equivalent circuits for the discontinuities are then substituted into SPICE or SPICE-like circuit modeling software to obtain the desired overall circuit response. Using the Fourier-transformed FDTD fields, the desired equivalent circuits may be derived in two ways [10]:

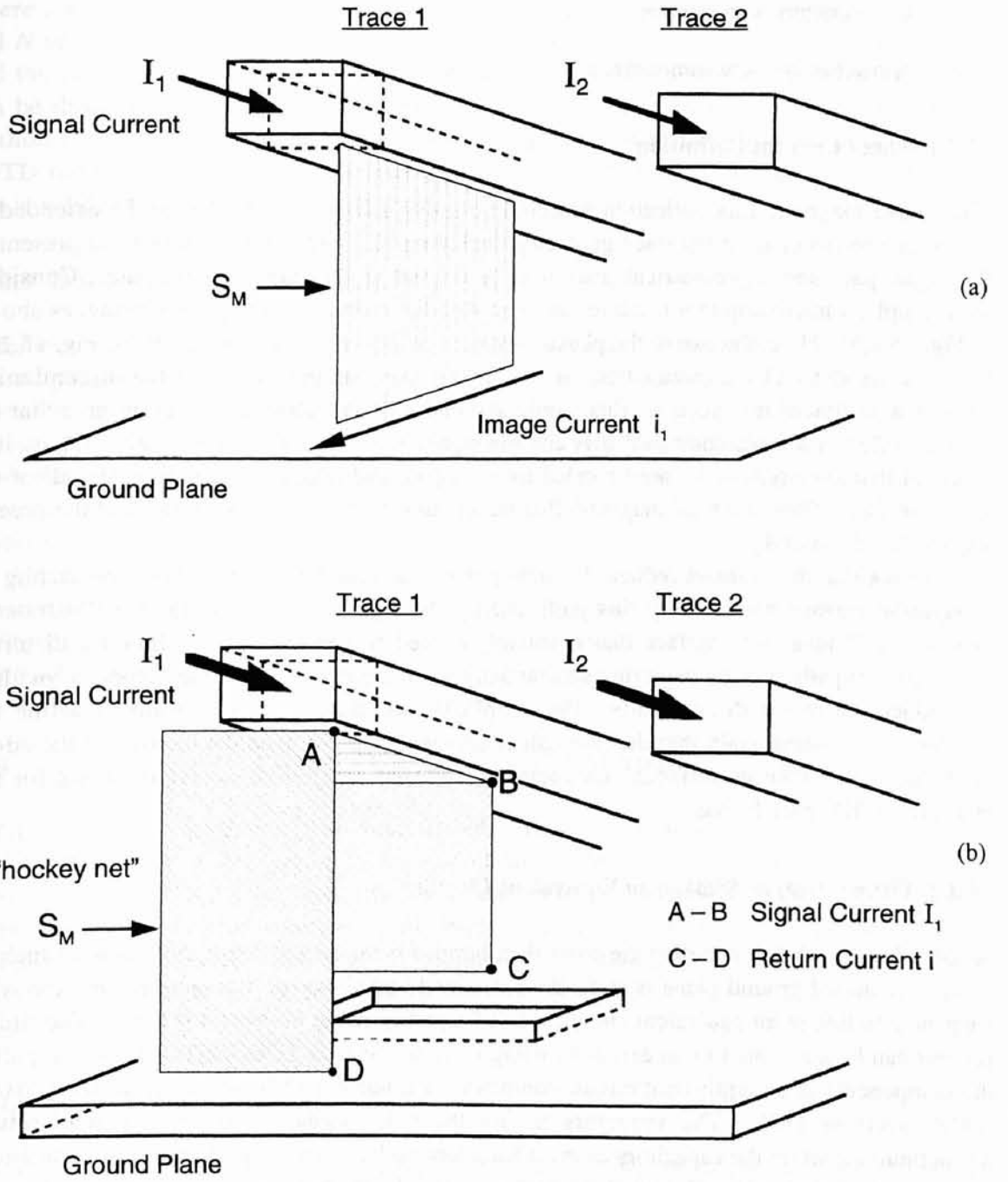


Fig. 15.1 Magnetic flux integration surfaces suitable for inductance calculations for a circuit trace over a ground plane: (a) "window pane" surface suitable for a continuous ground plane; (b) "hockey net" surface suitable for a slotted ground plane.

- By using the basic flux / current definition of inductance, as discussed in Section 15.2.5;
- By fitting an equivalent circuit to the calculated variation of impedance or S-parameters with frequency.

These approaches are now summarized.

15.3.1 Flux / Current Definition

The simple magnetic flux collection procedure discussed in Section 15.2.5 can be extended to any three-dimensional circuit trace geometry, including the interesting cases of a via present in the signal path and a geometrical discontinuity present in the current return path. Consider, for example, a microstrip that traverses an open slot discontinuity in the ground plane, as shown in Fig. 15.1(b). Here, the use of the planar "window-pane" integration surface S_M of Fig. 15.1(a) between the slot and the circuit trace would underreport the inductance of the discontinuity, because a portion of the magnetic flux would slip under its collection area. But we note that the magnetic flux is always contained between the signal and the return current paths. Thus, it is essential that the *edges* of S_M are bounded by the signal and return currents to collect all of the magnetic flux. Then, the total magnetic flux penetrating S_M would be independent of the precise surface definition of S_M.

We see that the problem reduces to finding the return-current path and then constructing an integration surface bounded by this path and by the signal trace. Fig. 15.1(b) illustrates a "hockey net" integration surface that is suitably curved around the slot to follow the distorted return current path, thereby capturing substantially all of the magnetic flux and properly yielding the inductance of the discontinuity. For simplicity, we assign Cartesian walls to define the "hockey net," noting again that the flux calculation is dependent upon the location of the edges of the net rather than its surfaces. Of course, a Cartesian integration surface is natural for the standard FDTD space lattice.

15.3.2 Fitting $Z(\omega)$ or $S(\omega)$ to an Equivalent Circuit

A second approach for obtaining the equivalent lumped inductance due to a discontinuity such as a via or a slotted ground plane is to fit the calculated impedance or S-parameter variation with frequency to that of an equivalent circuit over a frequency range of interest [11–13]. The fitting process can be automated to an extent by using software such as TOUCHSTONE to manipulate the components in an equivalent circuit composed of a number of *inductor-capacitor-resistor* (LCR) sections [13]. The inductors model the vias, signal traces, and ground return discontinuities, while the capacitors account for coupling between the package planes. Resistors are associated with internal impedances of sources, lossy materials, and leakage radiation. The distributed nature of these processes is modeled by using several LCR sections.

For situations where an equivalent circuit is needed over only a limited range of frequencies, [13] reported an alternative to the use of a sophisticated software. This approach is suitable for deriving a simple equivalent circuit by means of a rational function representation of the input impedance:

$$Z_{in}(s) = \frac{V(s)}{I(s)} = R_{in}(s) + jX_{in}(s) \approx \sum_{i=0}^{M} a_i s^i \Big/ \sum_{k=0}^{N} b_k s^k \tag{15.18}$$

where $s = j2\pi f$ (for f spanning the desired range of frequencies), and the summation orders M and N are selected to achieve a balance between the accuracy of the impedance-fitting process and the complexity of the resulting equivalent circuit. The unknown real coefficients a_i and b_k can be determined by using either a least-squares minimization procedure or an interpolation method to achieve a best-fit to the variation of impedance with frequency obtained from the FDTD modeling. Subsequently, these coefficients are used to synthesize an equivalent circuit that is valid in the desired range of frequencies.

15.3.3 Discussion: Choice of Methods

Reference [10] provided data indicating that certain important distributed components, such as ground pads connected to the ground plane using vias, cannot be easily modeled using the fitting procedure of Section 15.3.2. The key data provided (see Fig. 8 of [10]) indicate that the phase of S_{11} (equivalently, the phase of Z_{in}) can have a dramatically different variation with frequency, depending upon the choice of the reference plane. This sensitivity to the observation point makes it very hard to find an equivalent inductance that is independent of the location of the observation plane *and* the frequency.

On the other hand, [10] reported that inductance calculated using the flux / current method is accurate and essentially independent of frequency. The reader is therefore cautioned that the fitting procedure of Section 15.3.2 may not be as robust as the more basic flux / current definition of inductance of Section 15.3.1.

15.4 INDUCTANCE OF COMPLEX POWER-DISTRIBUTION SYSTEMS

Determining the equivalent-circuit lumped 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. This section discusses FDTD modeling of the lumped inductance of a complex printed-circuit-board power-distribution structure. The example shown demonstrates the versatility of the method for arbitrary structures of this type.

15.4.1 Method Description

The previous two sections discuss calculating the signal-path inductance directly from FDTD field data based upon the fundamental flux / current definition. This technique works well for signal paths where the return current flows along a reasonably linear route adjacent to the signal trace. However, by design, 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 not amenable to the flux / current definition.

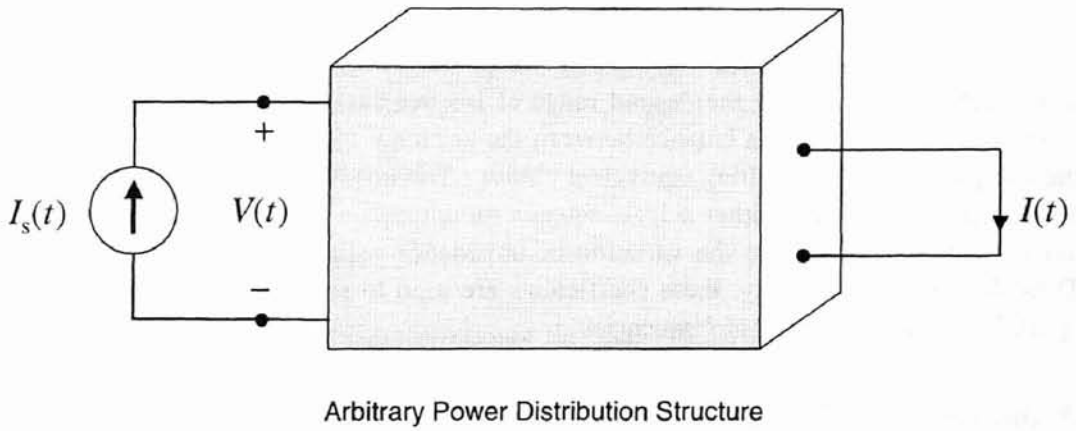


Fig. 15.2 Black-box model of FDTD inductance simulation using the method of current-pulse injection.

Referring to Fig. 15.2, consider an alternate approach [14] for modeling the inductance of an arbitrary power-distribution system. Here, we inject a current pulse $I_s(t)$ at the input of the power-distribution structure. Short-circuiting the output of this structure simulates the operation of a load that is drawing significant current (e.g., a CMOS integrated circuit experiencing a current spike). In this state, the conduction current causes magnetic flux to develop throughout the power-distribution structure, thereby emphasizing its inductive characteristic. A voltage $V(t)$ develops across the input of this structure in response to the current $I(t)$ flowing through the effective lumped inductance of the structure, according to the fundamental relation $V = L\mathrm{d}I/\mathrm{d}t$. Data for both $V(t)$ and $\mathrm{d}I/\mathrm{d}t$ are derived from the full-wave FDTD-calculated fields, as described in Section 15.2. For many power-distribution structures, the temporal waveforms of these quantities are sufficiently similar that they can be related to each other by a constant ratio, which is, in fact, the equivalent inductance $L = V/(\mathrm{d}I/\mathrm{d}t)$.

This approach permits the calculation of inductance to be free from any special considerations of the internal geometry of the power-distribution structure. In fact, the full-wave solution of Maxwell's equations by FDTD already accounts for the physics of the structure's internal and external geometry to the resolution limit of the space lattice. Thus, this approach can be used systematically to calculate the inductance of arbitrarily shaped, three-dimensional structures. In addition, the full-wave FDTD solver frees the inductance calculation from any potentially limiting TEM assumptions or approximations. This provides an inherently three-dimensional environment for the required simulation and parameter extraction.

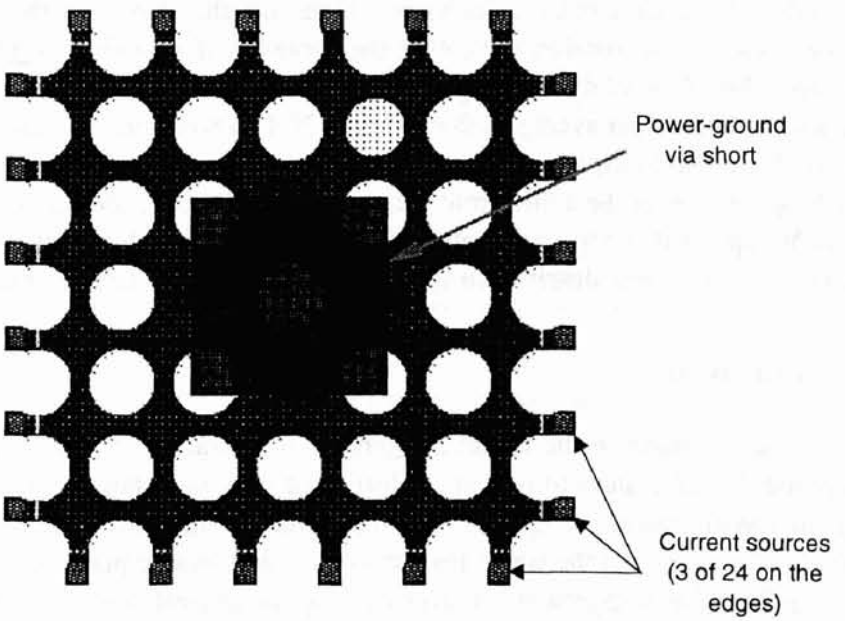
Of key importance to the correct application of this method is the choice of the injected current waveform. There are two directly competing conditions for this waveform:

- The leading edge should rise *quickly* enough to yield a voltage response that corresponds mainly to the inductive component of the structure.
- The leading edge should rise *slowly* enough so that, upon propagation, this edge extends spatially over the entire structure to be characterized for inductance. This provides a comparable dynamic condition over each part of the structure, even though these parts may be distributed in space, and permits the structure to be characterized as a single lumped element.

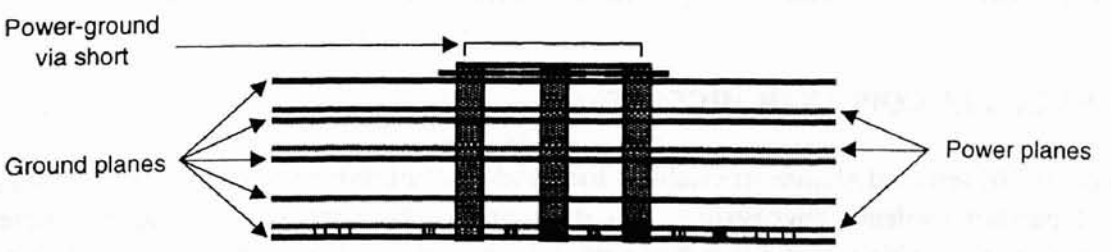
A Gaussian exciting current pulse can be used to meet these conditions. Validation studies using FDTD modeling of canonical microstrip and coaxial geometries [14] indicate that Gaussian leading-edge rise times that are 5 to 10 times longer than the propagation delay across the structure yield reliable inductance data over a 20% to 80% span of the rising edge of the FDTD-calculated response. We also find it useful to average the observed inductance over a time-windowed portion of the response. This averaging window can extend up to the 20% to 80% span. These findings are further supported by analogous circuit-based studies using SPICE to model the current-pulse excitation of an inductor [14].

15.4.2 Example: Multiplane Meshed Printed-Circuit Board

A significant problem in modeling and designing high-speed digital circuits has been understanding how to extract system-level parameters from complex power-distribution systems. Fig. 15.3 shows one such structure, a 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.



(a) Top view.



(b) Side view.

Fig. 15.3 Example of a complex power-distribution system modeled using FDTD.

This model is based on a portion of an actual PCB power-distribution structure that lies under a 2,499-pin, 30-mm *multichip module* (MCM) [15]. The MCM has a pad pitch of 0.5 mm and is built with copper / polyimide technology. The modeled portion of the PCB 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. In applying the inductance-modeling technique discussed in Section 15.4.1, the voltage response is measured between a power plane and a ground plane at the edge of the PCB structure. A via short to a ground plane on top of the structure (simulating a real load) provides a measurement of the input current that causes a voltage to develop across the structure.

In the past, inductance calculations for this type of structure were done in a piecewise manner; that is, by examining the effects of the power / ground planes on the vias, and vice versa. However, by applying the method of Section 15.4.1 (based strictly on the structure's stimulus and response), it is possible to represent the entire power-distribution structure as a lumped inductor with one straightforward calculation.

Fig. 15.4 depicts the FDTD-computed V and I data for the structure of Fig. 15.3, along with the results of the inductance calculation. Note that the shapes of the dI/dt stimulus and the voltage response are similar, indicating the presence of an inductive element linking the two variables. The final plot in the figure is the structure's inductance as given by $L = V/(dI/dt)$. This shows an 83.8 pH average value over the 20% to 80% time window. This period, in which L shows little variation, corresponds to the state where the leading edge of the exciting pulse extends spatially over the entire structure, as discussed earlier. Because the complete PCB under the MCM repeats this 3×3 mm subsection eight times in both lateral directions, the composite inductance of the power-distribution system feeding the MCM is approximately 10 pH.

15.4.3 Discussion

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. An extension of this approach involves developing and validating a dual method for lumped-capacitance modeling and optimization. Overall, this method could be integral in the development of design guidelines for digital structures similar to the one in Fig. 15.3.

15.5 PARALLEL COPLANAR MICROSTRIPS

Reference [16] reported studies to establish the validity of FDTD modeling for the important case of parallel coplanar microstrips. Initial studies set up models for single x-directed microstrips of negligible metallization thickness and a variety of widths over dielectric substrates about 1 mil (25.4 μm) thick. For these cases, $\Delta \cong 0.1$ mil (2.54 μm) and $\Delta t \cong 4.2$ fs. In all cases, the conductors were "on grid"; that is, located at loci of tangential $\boldsymbol{E}$ components in the lattice set to zero for all time-steps.

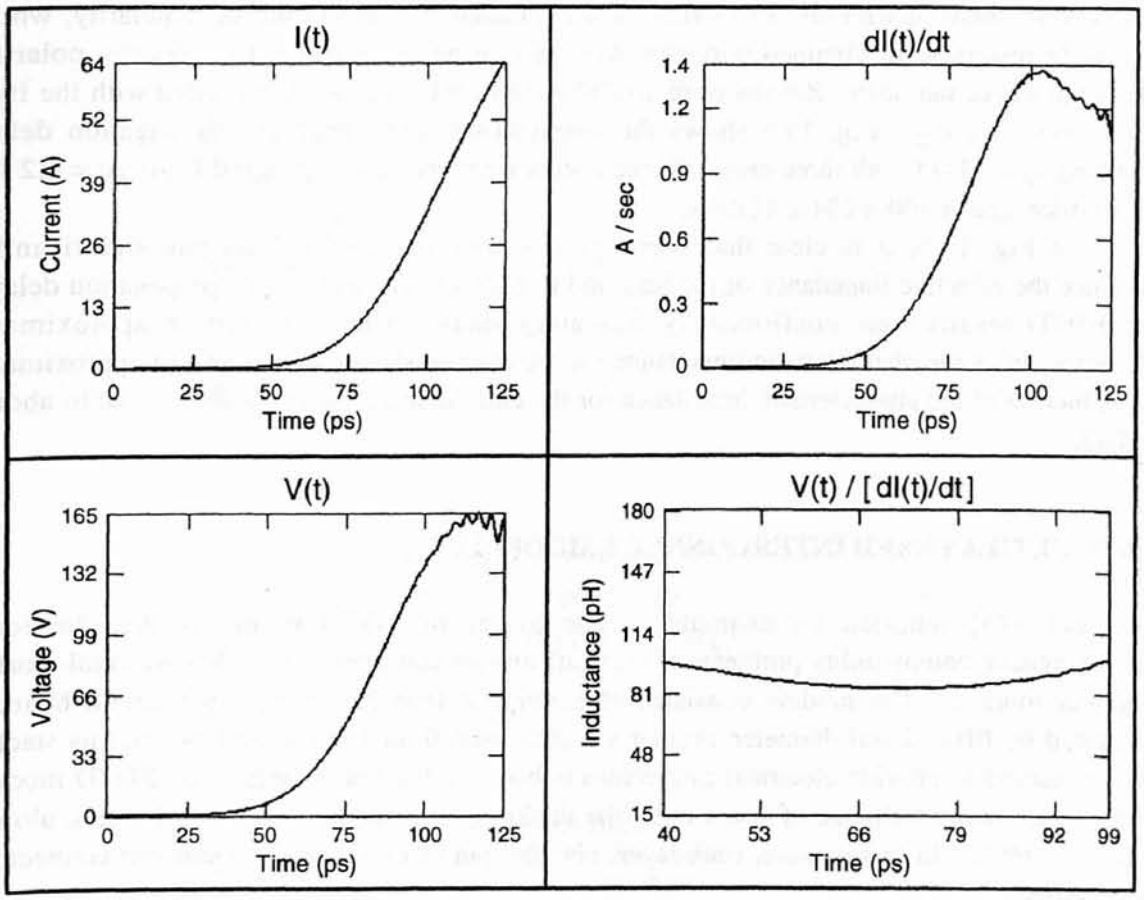


Fig. 15.4 FDTD data used to calculate the inductance of the multiplane power-distribution system of Fig. 15.3.

Excitation of a microstrip was provided by specifying a Gaussian pulse waveform for a group of collinear $\boldsymbol{E}$ components bridging the gap between the ground plane and the strip conductor at the desired source location. Application of (15.2) showed the FDTD-computed values of Z_0 to be virtually independent of frequency up to 1.0 GHz, and on the order of 1% agreement with textbook values given in [17].

Next considered were single microstrips with finite metallization thickness, possibly fully embedded within a dielectric layer. Here, FDTD predictions were compared to measurements. In one example, a 1.1-mil-thick, 1.4-mil-wide metal strip was assumed to be suspended 1.1 mils above a large ground plane within a 3.1-mil-thick dielectric layer having $\varepsilon_r = 3.2$. The FDTD simulation predicted a flat characteristic impedance of $48\,\Omega$ up to 1.0 GHz, agreeing with the experimental results to within the measurement uncertainty (approximately $0.2\,\Omega$). The computed variation of Z_0 above 1 GHz was found to be $\pm 2\,\Omega$. Similar agreement was found for the propagation delay, where the experimental value of 150.5 ± 1.5 ps/in compared very well with the 149.5 ps/in delay calculated using the FDTD model.

Reference [16] also reported a study of the impedance and propagation delay of three parallel, coplanar microstrips of the type discussed above, separated by 3.6 mils. Even-mode results were obtained with all three strips excited simultaneously with the same polarity, while odd-mode results were obtained with the two outer strips excited with the opposite polarity relative to the center strip. Results were also observed for the center strip excited with the two outer strips floating. Fig. 15.5 shows the characteristic impedance and propagation delay predicted by FDTD for all three cases. These results were obtained with $\Delta = 0.1$ mil, $\Delta t = 4.2$ fs, and a lattice size of $200 \times 134 \times 42$ cells.

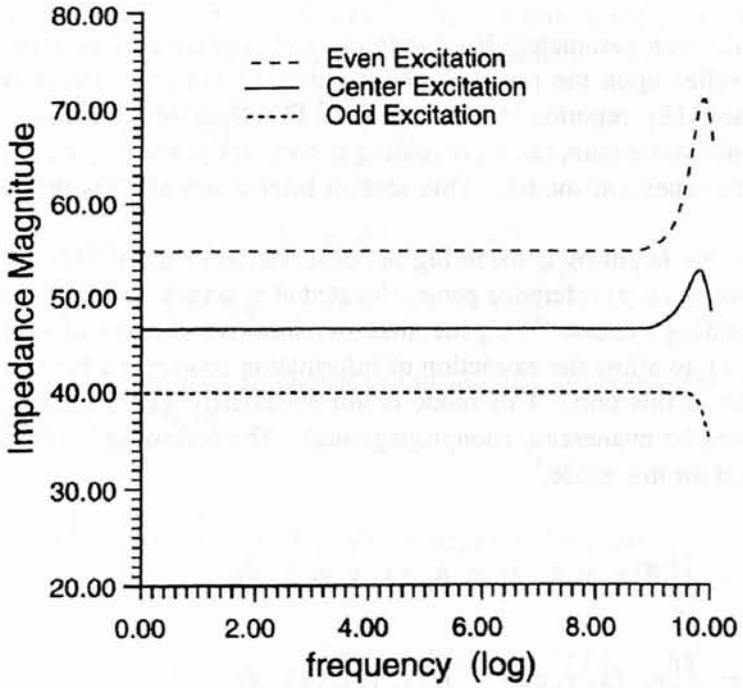
From Fig. 15.5, it is clear that signal propagation on adjacent lines can significantly influence the effective impedance of the line, and to a lesser degree affect the propagation delay. The FDTD results were confirmed by laboratory studies, which showed an approximate $7\,\Omega$ elevation of the characteristic impedance for the even-mode excitation, and an approximate $7\,\Omega$ reduction of the characteristic impedance for the odd-mode excitation (both from dc to about 1 GHz).

15.6 MULTILAYERED INTERCONNECT MODELING

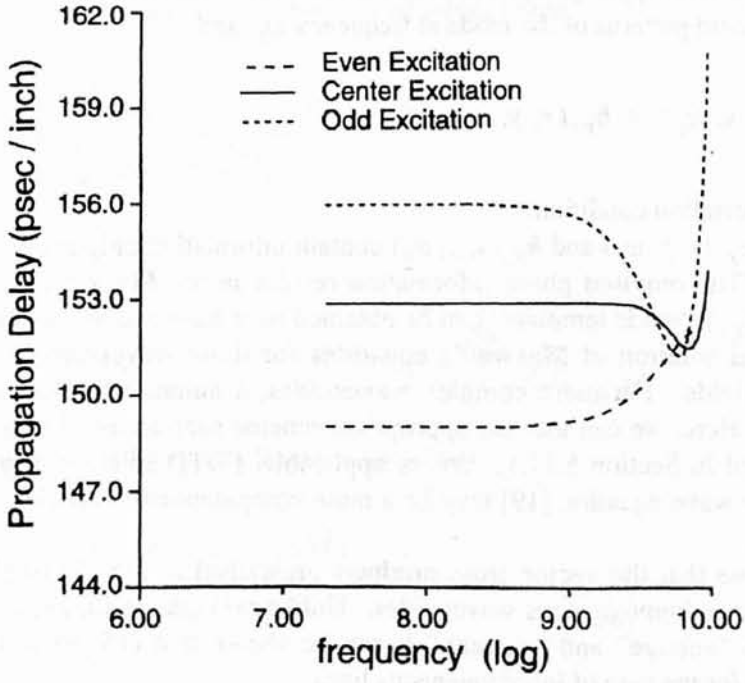
Reference [16] reported an example of the power of FDTD to model the classical electromagnetic compatibility problems of crosstalk and ground-loop current flow in a real-world computer module. The module consisted of a stack of four 22-layer printed circuit boards penetrated by fifty 12-mil-diameter circular via pins located on 100-mil centers. In this stack, the vias served to provide electrical interconnects between the four boards. The FDTD model used a space lattice cell size of $\Delta = 4$ mils (the thickness of a single circuit-board layer), along with $\Delta t = 169$ fs. In this manner, each layer, via, and pin of every circuit board and connector was modeled.

The study first focused on the early-time response of the top circuit board to a 90-ps Gaussian pulse propagating down a *single* vertical via pin. Excitation was provided by pulsing a single $\boldsymbol{E}$ component in the FDTD space lattice, just above the via pin and below a simulated ground strap connected to the outer via pins designated as ground returns. The short duration of the pulse was selected primarily to permit time resolution of layer-to-layer pulse reflection effects within the circuit board. A video of the dynamics of the pulse propagation along the excited via pin showed repeated bursts of outward-propagating waves linking all points within transverse cross sections of the board as the leading edge of the pulse passed vertically through the multiple layers of the board. The resulting early-time crosstalk was calculated to be of sufficient amplitude to upset the operation of the digital circuits using the pins immediately adjacent to the excited pin.

The study then focused on the late-time response of the complete module to the same via-pin excitation. Keeping the same space resolution as for the single-board model, the FDTD lattice was enlarged to $300 \times 92 \times 340$ cells. Currents were derived in a postprocessing step by numerically evaluating the curl of the H-field obtained from the FDTD model. In this manner, it was determined that signal current had proceeded down the excited via through all four boards and all three connectors of the module. However, return current was observed to flow upward on the adjacent signal vias. This represents undesired ground-loop coupling to the digital circuits using these vias at a level predicted to be of sufficient amplitude to cause circuit upset.



(a) Impedance versus frequency.



(a) Propagation delay versus frequency.

Fig. 15.5 FDTD-calculated transmission line characteristics of three parallel coplanar microstrips. *Source:* Piket-May et al., *IEEE Trans. Microwave Theory and Techniques*, 1994, pp. 1514–1523, © 1994 IEEE.

15.7 *S*-PARAMETER EXTRACTION FOR GENERAL WAVEGUIDES

Extraction of transmission parameters from stripline and microstrip structures, discussed in the previous sections, relies upon the primarily TEM nature of the propagating fields along these structures. Reference [18] reported means to process FDTD modeling data to extract wideband *S* parameters for much more complex waveguiding geometries potentially having a superposition of propagating and evanescent modes. This section briefly summarizes the foundation of this technique.

Following [18], we begin by considering the observation of the FDTD-calculated $\boldsymbol{E}$ and $\boldsymbol{H}$ fields along a transverse (x, y) reference plane s located at z_p across a particular geometric port of a multiport waveguiding system. We generalize our intuitive concept of voltage and current, as presented in (15.1), to allow the extraction of information concerning the i'th electromagnetic mode that may exist at this port. This mode is *not* necessarily TEM. It may be one of many modes, and may even be evanescent (nonpropagating). The following "voltage" and "current" quantities are defined for this mode:[1]

$$V_i(z_p, t) = \iint_s \boldsymbol{E}(x, y, z_p, t) \times \boldsymbol{h}_{T,i}(x, y, \omega_T) \cdot \boldsymbol{ds} \tag{15.19}$$

$$I_i(z_p, t) = \iint_s \boldsymbol{e}_{T,i}(x, y, \omega_T) \times \boldsymbol{H}(x, y, z_p, t) \cdot \boldsymbol{ds} \tag{15.20}$$

where $\boldsymbol{e}_{T,i}(x, y, \omega_T)$ and $\boldsymbol{h}_{T,i}(x, y, \omega_T)$ are real-valued vector functions representing the cross-sectional $\boldsymbol{E}$ and $\boldsymbol{H}$ field patterns of the mode at frequency ω_T, and

$$\iint_s \boldsymbol{e}_{T,i}(x, y, \omega_T) \times \boldsymbol{h}_{T,i}(x, y, \omega_T) \cdot \boldsymbol{ds} = 1 \tag{15.21}$$

provides the normalization condition.

We note that $\boldsymbol{e}_{T,i}(x, y, \omega_T)$ and $\boldsymbol{h}_{T,i}(x, y, \omega_T)$ contain information only about field amplitudes, and not phases. [The required phase information resides in the $\boldsymbol{E}(x, y, z_p, t)$ and $\boldsymbol{H}(x, y, z_p, t)$ data.] The $\{\boldsymbol{e}_{T,i}, \boldsymbol{h}_{T,i}\}$ "mode template" can be obtained in at least two ways. The first approach is via an analytical solution of Maxwell's equations for those waveguide geometries having solvable internal fields. For more complex waveguides, a numerical solution of the internal fields is required. Here, we can use any appropriate scheme such as the "bootstrapping" FDTD procedure discussed in Section 5.11.1. Where applicable, FDTD solution of a simplified two-dimensional vector wave equation [19] may be a more computationally efficient means to obtain the mode template.

We further note that the vector cross products prescribed in (15.19) to (15.21) present a robust procedure for inhomogeneous waveguides. Unlike previous techniques employing scalar products to define "voltage" and "current," it can be shown that (15.19) to (15.21) properly separate the modes for the case of inhomogeneous lines.

---

[1] It is only by analogy to TEM transmission lines that V and I of (15.19) and (15.20) are called "voltage" and "current." In actual fact, these quantities may be quite far from such physical interpretations.

The next step in calculating the *S*-parameters is to obtain the Fourier transforms of the modal "voltage" and "current," taking care to account for the space-time offsets inherent in the FDTD algorithm. Linear interpolation is used as needed, as implemented in [8]:

$$\breve{V}_i(z_p, \omega) = \mathcal{F}\left[V_i(z_p, t)\right] \tag{15.22}$$

$$\breve{I}_i(z_p, \omega) = \mathcal{F}\left\{0.5\begin{bmatrix} I_i(z_p + \Delta z/2,\ t + \Delta t/2) + \\ I_i(z_p - \Delta z/2,\ t + \Delta t/2) \end{bmatrix}\right\} e^{-j\omega\Delta t/2} \tag{15.23}$$

In addition, following [18], we require knowledge of the Fourier transforms of the z-derivatives of the modal "voltage" and "current":

$$\breve{V}_i'(z_p, \omega) = \mathcal{F}\left[\frac{\partial V_i(z, t)}{\partial z}\right]_{z=z_p} = \mathcal{F}\left\{\left[V_i(z_p + \Delta z, t) - V_i(z_p - \Delta z, t)\right]/2\Delta z\right\} \tag{15.24}$$

$$\breve{I}_i'(z_p, \omega) = \mathcal{F}\left[\frac{\partial I_i(z, t)}{\partial z}\right]_{z=z_p} = \mathcal{F}\left\{\begin{bmatrix} I_i(z_p + \Delta z/2,\ t + \Delta t/2) - \\ I_i(z_p - \Delta z/2,\ t + \Delta t/2) \end{bmatrix}\Big/\Delta z\right\} e^{-j\omega\Delta t/2} \tag{15.25}$$

Then, from [18], we have the modal impedance at reference plane z_p:

$$Z_i(z_p, \omega) = \sqrt{\frac{\breve{V}_i(z_p, \omega)\,\breve{V}_i'(z_p, \omega)}{\breve{I}_i(z_p, \omega)\,\breve{I}_i'(z_p, \omega)}} \tag{15.26}$$

and the modal propagation constant:

$$\gamma_i(z_p, \omega) = \alpha_i(z_p, \omega) + j\beta_i(z_p, \omega) = \sqrt{\frac{\breve{I}_i'(z_p, \omega)\,\breve{V}_i'(z_p, \omega)}{\breve{I}_i(z_p, \omega)\,\breve{V}_i(z_p, \omega)}} \tag{15.27}$$

and the incident and reflected wave coefficients, respectively:

$$\breve{a}_i(z_p, \omega) = \frac{\breve{V}_i(z_p, \omega) + Z_i(z_p, \omega)\breve{I}_i(z_p, \omega)}{2\sqrt{Z_i(z_p, \omega)}} \tag{15.28}$$

$$\breve{b}_i(z_p, \omega) = \frac{\breve{V}_i(z_p, \omega) - Z_i(z_p, \omega)\breve{I}_i(z_p, \omega)}{2\sqrt{Z_i(z_p, \omega)}} \tag{15.29}$$

The incident and reflected wave coefficients of (15.28) and (15.29) can be inserted directly into (15.11), thereby yielding the *S*-parameters for the mode of interest.

The above procedure allows wideband extraction of the S-matrix with compensation for imperfect matching of the loads. Frequency dependence of the reference (modal) impedances is automatically taken into account. Moreover, knowledge of the propagation constant γ provided by (15.27) also can be used for a virtual shift of the port reference planes at the postprocessing stage.

Furthermore, knowledge of γ allows us to refine the averaged and differenced quantities calculated in (15.23) to (15.25) for improved accuracy [8, 18]. For propagating modes, we first implement (15.22) to (15.25), calculate γ based upon these results, and then recalculate (15.23) to (15.25) as follows:

$$\breve{I}_i(z_p, \omega) = \mathcal{F}\left\{ \frac{0.5}{\cos(\beta_i \Delta z/2)} \cdot \begin{bmatrix} I_i(z_p + \Delta z/2,\ t + \Delta t/2) + \\ I_i(z_p - \Delta z/2,\ t + \Delta t/2) \end{bmatrix} \right\} e^{-j\omega \Delta t/2} \tag{15.30}$$

$$\breve{V}_i'(z_p, \omega) = \mathcal{F}\left\{ \frac{V_i(z_p + \Delta z, t) - V_i(z_p - \Delta z, t)}{2\Delta z} \cdot \frac{\beta_i \Delta z}{\sin(\beta_i \Delta z)} \right\} \tag{15.31}$$

$$\breve{I}_i'(z_p, \omega) = \mathcal{F}\left\{ \frac{\begin{bmatrix} I_i(z_p + \Delta z/2,\ t + \Delta t/2) - \\ I_i(z_p - \Delta z/2,\ t + \Delta t/2) \end{bmatrix}}{\Delta z} \cdot \frac{\beta_i \Delta z/2}{\sin(\beta_i \Delta z/2)} \right\} e^{-j\omega \Delta t/2} \tag{15.32}$$

These corrections enforce a sinusoidal field variation in the z-direction between the FDTD field evaluation points. Without these corrections, direct use of (15.23) to (15.25) would amount to the assumption of a piecewise-linear field variation between these points.[2] For evanescent modes, a similar correction procedure can be followed, but with the trigonometric functions in (15.30) to (15.32) replaced by their respective hyperbolic versions, and α_i used in place of β_i.

15.8 DIGITAL SIGNAL PROCESSING AND SPECTRUM ESTIMATION

Typical FDTD models of high-speed circuit structures use a lattice resolution Δ that is dictated by the dimensions of the circuit board layers and vias. This resolution is almost always *much finer* than needed to resolve the smallest spectral wavelength propagating in the circuit. As a result, with the time-step Δt bound to Δ by numerical stability considerations, it may be necessary to run FDTD simulations for tens of thousands of time-steps in order to fully evolve the impulse responses needed for calculating impedances, S-parameters, and resonant frequencies. This is particularly probable if the structure being modeled has a high *quality* (Q) factor, with little energy dissipated relative to the energy stored.

---

[2] While the correction approach described here is formally iterative, for all practical purposes it converges after just one iteration.

A brute-force way to mitigate this burden is to truncate a run before the impulse response fully evolves. However, this has the effect of viewing the true time-domain response through a rectangular window of duration $T = N_{\max}\Delta t$. In the frequency domain, this windowing is translated into the convolution of the true spectrum with the function $\sin(f)/f$. This convolution widens the peaks in the spectral response, causes other distortions, and can mask weak spectral signatures. Distortion can be reduced and the spectral resolution improved by lengthening the window. However, this may cause the FDTD computer simulation times for several important classes of problems (particularly resonators) to be so lengthy as to be virtually prohibitive.

A fruitful approach in addressing this problem is to apply analysis techniques from the discipline of digital signal processing and spectrum estimation. The strategy is to extrapolate the electromagnetic field time waveform by 10:1 or more beyond the actual FDTD time-stepping window, allowing a good estimate of the complete system response with 90% or greater reduction in the FDTD computation time. In effect, a low-computational-burden extrapolation process at a *limited number* of reflection or transmission observation points in the lattice replaces the computationally intensive FDTD process that is necessarily applied at *every point* in the lattice. Then, the reflected and transmitted field spectra and the associated impedances and S-parameters at the observation points can be efficiently obtained by FFT of the extrapolated waveforms or, in certain cases, from the coefficients of the extrapolation process itself.

This section discusses a number of digital signal processing and spectrum estimation techniques that have appeared in the FDTD literature. While good results have been obtained with each of these methods, there are tradeoffs. A goal of this section is to illuminate the tradeoffs as currently understood.

15.8.1 Prony's Method

Prony's method [20] has been used to extrapolate *transmission-line matrix* (TLM) and FDTD-computed waveforms for microwave circuits by a number of investigators [21–23]. Because it is a technique for modeling sampled data as a linear combination of complex exponentials, it is particularly suitable for calculating the resonant frequency and Q of a resonating structure, since the impulse response of such a structure is characterized by a superposition of decaying exponentials.

To begin our discussion, we assume the existence of N equally spaced time samples of the FDTD-computed impulse response of field component F at observation point (i, j, k) in a high-speed circuit:

$$\left\{ F\big|_{i,j,k}^{M \to M+N-1} \right\} \equiv F\big|_{i,j,k}^{M},\ F\big|_{i,j,k}^{M+1},\ F\big|_{i,j,k}^{M+2},\ \ldots,\ F\big|_{i,j,k}^{M+N-1} \tag{15.33}$$

These samples constitute an observation "window" that begins at time-step M and ends at $M+N-1$. Because the FDTD data are very oversampled relative to what is needed for Prony's method (and the other signal-processing methods discussed in this section), $\{F\}$ is obtained by *decimating* the actual FDTD data record by a factor of 10:1 or greater. Now, let each time sample in $\{F\}$ be approximated by a sum of p exponentials:

$$F\big|_{i,j,k}^{n} = \sum_{\ell=1}^{p} C_\ell\, e^{(\alpha_\ell + j2\pi f_\ell)(n-M)\Delta t} \equiv \sum_{\ell=1}^{p} C_\ell \left(\mu_\ell\right)^{n-M} \tag{15.34}$$

where $0 \le n - M \le N - 1$, C_ℓ is the complex-valued amplitude of the ℓ'th mode, $\mu_\ell = \exp(\alpha_\ell + j2\pi f_\ell)\Delta t$, α_ℓ is the damping factor of the ℓ'th mode, f_ℓ is the frequency of the ℓ'th mode, and p is the order of the model.

The direct solution of (15.34) is a difficult nonlinear least-squares problem. An alternative solution is based on Prony's method. This is a two-step procedure [20] which solves two sequential sets of linear equations with an intermediate polynomial rooting step that concentrates the nonlinearity of the problem. Following [24], the first step is to set up the following $N - p$ equations:

$$\begin{aligned}
&F|_{i,j,k}^{M+p} + A_1 F|_{i,j,k}^{M+p-1} + A_2 F|_{i,j,k}^{M+p-2} + \ldots + A_p F|_{i,j,k}^{M} = 0 \\
&F|_{i,j,k}^{M+p+1} + A_1 F|_{i,j,k}^{M+p} + A_2 F|_{i,j,k}^{M+p-1} + \ldots + A_p F|_{i,j,k}^{M+1} = 0 \\
&\quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\quad \ddots \qquad \vdots \qquad\quad \vdots \\
&F|_{i,j,k}^{M+N-1} + A_1 F|_{i,j,k}^{M+N-2} + A_2 F|_{i,j,k}^{M+N-3} + \ldots + A_p F|_{i,j,k}^{M+N-1-p} = 0
\end{aligned} \tag{15.35}$$

This overdetermined system is solved using a least-squares algorithm for the $\{A_\ell\}$. This permits the $\{\mu_\ell\}$ to be found as the roots of the polynomial

$$\mu^p + A_1 \mu^{p-1} + A_2 \mu^{p-2} + \ldots + A_{p-1}\mu + A_p = 0 \tag{15.36}$$

The second step is to find the $\{C_\ell\}$ in (15.34). This is accomplished by writing out (15.34) for each of the time samples, thereby obtaining a set of N equations in the p values of the C_ℓ:

$$\begin{aligned}
&C_1 + C_2 + \ldots + C_p = F|_{i,j,k}^{M} \\
&\mu_1 C_1 + \mu_2 C_2 + \ldots + \mu_p C_p = F|_{i,j,k}^{M+1} \\
&(\mu_1)^2 C_1 + (\mu_2)^2 C_2 + \ldots + (\mu_p)^2 C_p = F|_{i,j,k}^{M+2} \\
&\quad\vdots \qquad\quad \vdots \qquad\quad \ddots \qquad \vdots \qquad\quad \vdots \\
&(\mu_1)^{N-1} C_1 + (\mu_2)^{N-1} C_2 + \ldots + (\mu_p)^{N-1} C_p = F|_{i,j,k}^{M+N-1}
\end{aligned} \tag{15.37}$$

This overdetermined system is solved using a least-squares algorithm for the $\{C_\ell\}$. Now all of the parameters in (15.34) are known, and it becomes possible to:

1. Write down by inspection the set of resonant frequencies and Q-factors $Q_\ell = |\pi f_\ell / \alpha_\ell|$ directly from the Prony parameter sets just obtained;
2. Calculate the complete estimated impulse response at times $n >> M + N - 1$ until it is observed to decay to zero, at which point an FFT provides the magnitude and phase of the frequency-domain transfer function with no windowing artifacts.

Fig. 15.6 shows results using Prony's method reported in [23]. This figure graphs the resonant frequencies and Q-factors of the three lowest TE_{0n} modes of a cylindrical dielectric resonator of permittivity $\varepsilon_r = 38$, radius $a = 5.25$ mm, and length $l = 4.6$ mm. Here, a cylindrical FDTD grid was used, with parameters $\Delta r = 0.29167$ mm, $\Delta z = 0.2875$ mm, and $\Delta t = 0.4795$ ps. It can be seen that a time-stepping window extending over about 2^{17} (131,072) iterations was needed to obtain converged results for the Q of the TE_{02} mode when using an FFT applied to the windowed FDTD data. However, using Prony's method of order $p = 10$ and a decimation factor of 50, only 3,000 iterations were required, a reduction of about 98%. In fact, reductions can be even greater in the case of structures having even higher Q or with very closely spaced resonant frequencies. The accuracy of the results was verified in [23] by comparing the data to those available in the literature, and the agreement was found to be good.

Prony's method shows good resolution with relatively short data sequences and presents no windowing problem. The main difficulty with this approach lies in its determination of p, the user-selected order of the model. If the value of p is less than the number of actual modes excited in the structure, the spectral resolution is poor.

However, if p is selected to be too high, spurious (nonphysical) modes appear. The following are two guidelines given in [23] to deal with these problems:

- Before Prony's method is applied, digitally lowpass filter the time-domain response to limit the number of resonant modes and therefore the number of parameters to calculate.
- If spurious modes are suspected, apply Prony's method with the time sequence of the sample points in reverse order. Here, the real modes appear with positive damping factors $\alpha_\ell > 0$, but the spurious modes have negative damping factors $\alpha_\ell < 0$ [25].

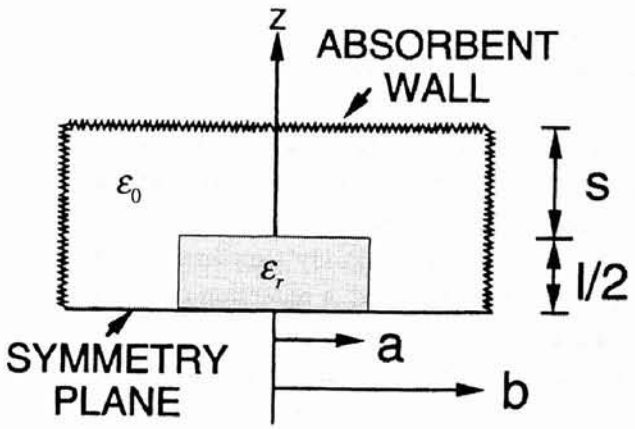
15.8.2 Autoregressive Models

The most popular time series modeling approach used in modern spectral estimation is the class of linear predictors or *autoregressive* (AR) models. This is because an accurate estimate of the AR parameters can be derived by solving a set of linear equations. In contrast to Prony's method, which uses a sum of deterministic exponential functions to fit the data, the AR approach constructs a random model to fit a statistical database to the second-order.

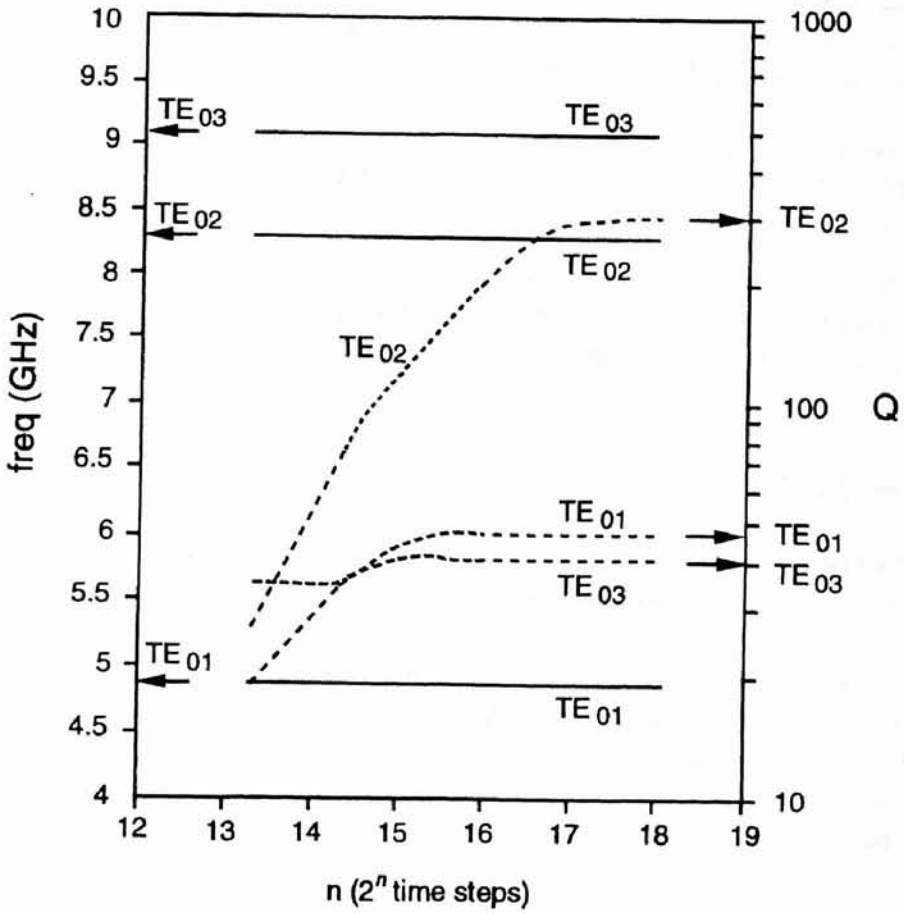
We again assume the FDTD-computed impulse response (15.33) (i.e., N equally spaced time samples beginning at time-step M). This time series is said to represent the realization of an AR process of order p if it satisfies the following relation:

$$f|_{i,j,k}^{n} = -a_1 f|_{i,j,k}^{n-1} - a_2 f|_{i,j,k}^{n-2} - \ldots - a_p f|_{i,j,k}^{n-p} + q(n) \tag{15.38}$$

where the constants $a_1, a_2, \ldots, a_p$ are the AR parameters to be determined from $\{f\}$, and $q(n)$ is a white-noise process whose variance also has to be found to carry out the extrapolation of $\{f\}$. It is clear that (15.38) permits the prediction of a new value of the time series from known previous values, once the AR parameters are in hand.



(a) Geometry.



(b) Legend: —— resonant frequencies, - - - Q-factors from FDTD / FFT after 2^n time-steps; left and right arrows: FDTD / Prony results using only a 3,000 time-step window.

Fig. 15.6 Convergence of the resonant frequencies and Q-factors of the three lowest TE_{0n} modes of an isolated dielectric resonator as a function of the number of FDTD time-steps. *Source:* Pereda et al., *IEEE Microwave and Guided Wave Letters*, 1992, pp. 431–433, © 1992 IEEE.

Many published approaches to evaluating the a_i exist [26]. In fact, a primary categorization of AR techniques is by the means used to calculate these parameters. In this section, we consider three specific applications of AR to the extrapolation of FDTD electromagnetic field records, categorized according to this criterion: (1) the covariance method, (2) the forward-backward method, and (3) the nonlinear predictor. The choice of approach is important, since these methods have substantially different operational characteristics.

Covariance Method

Reference [27] reported the use of the covariance method for determining the AR parameters. This method involves setting up and solving the following $p \times p$ linear system of equations:

$$\begin{bmatrix} c_{ff}(1,1) & c_{ff}(1,2) & \cdots & c_{ff}(1,p) \\ c_{ff}(2,1) & c_{ff}(2,2) & \cdots & c_{ff}(2,p) \\ \vdots & \vdots & \ddots & \vdots \\ c_{ff}(p,1) & c_{ff}(p,2) & \cdots & c_{ff}(p,p) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} c_{ff}(1,0) \\ c_{ff}(2,0) \\ \vdots \\ c_{ff}(p,0) \end{bmatrix} \tag{15.39}$$

where c_{ff} are covariances defined by

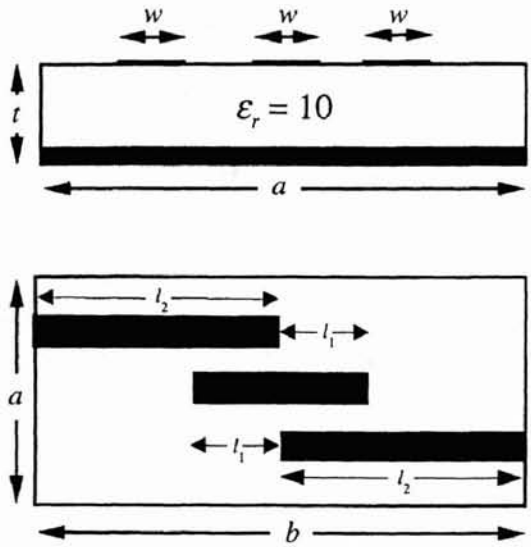
$$c_{ff}(\alpha, \beta) = \frac{1}{N-p} \sum_{n=p}^{N-1} \left(f\big|_{i,j,k}^{M+n-\alpha} \, f\big|_{i,j,k}^{M+n-\beta} \right) \tag{15.40}$$

The matrix of (15.39) is Hermitian and positive semidefinite, and can be solved by Cholesky decomposition.

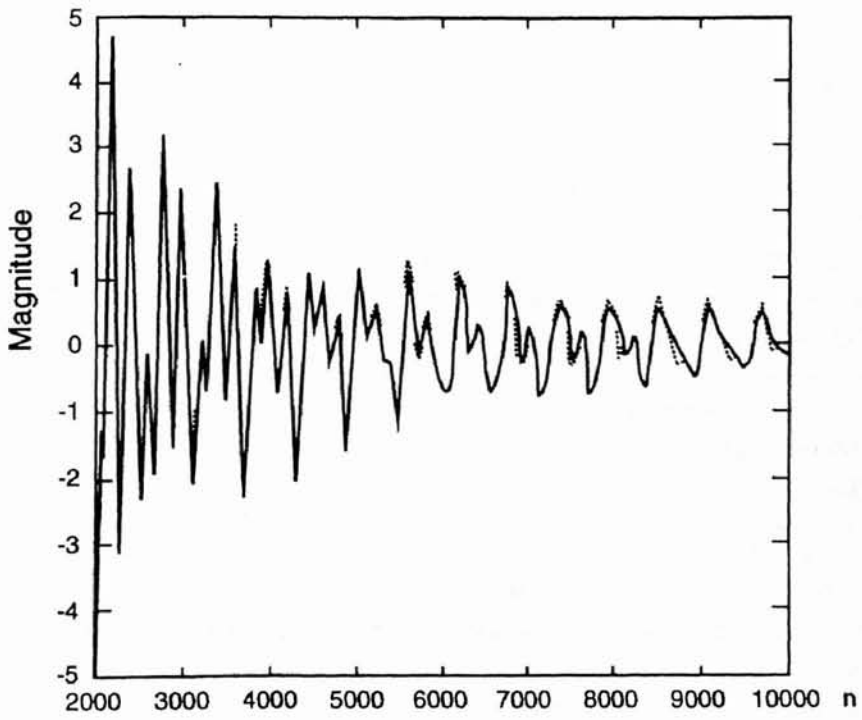
As stated in [27], a key issue that arises when using the covariance-based AR model is choosing the model order p. The use of a low-order model can cause the extrapolated waveform to attenuate nonphysically. However, a high-order model can have divergence problems due to statistical instabilities. These competing effects cause the results to be quite sensitive to p. Two common ways to estimate p are the final-prediction-error technique and the Akaike Information Criterion [28]. Both are based upon the estimated predictor error power, and are regarded as general guides for AR model selection.

Using the covariance-based AR method of (15.39) and (15.40), [27] reported the successful extrapolation to over 30,000 time-steps of FDTD field records spanning only 2,000 to 3,500 time-steps. A decimation factor of 10 was used, and the order of the AR model was 50. As shown in Fig. 15.7, very good accuracy was observed for the extrapolated time waveform for an edge-coupled microstrip bandpass filter.

The final example discussed in [27] involved modeling a microstrip double-stub structure. Here, lengthy wave propagation delays along the stubs required extrapolating the FDTD data to 100,000 time-steps to achieve convergence of the transmission coefficient. This extrapolation was accomplished using a 91st-order AR model working from an FDTD database covering iterations 3,100 to 8,900, and decimated by a factor of 25. Although very good agreement was observed with benchmark frequency-domain data, the results were found to be more sensitive to the model order than the 50th-order model used for the previous two examples.



(a) Geometry, showing the cross section and plan views. Key dimensions: $l_1 = 6.36$ mm; $l_2 = b/2 = 8.48$ mm; $a = 11.62$ mm; $w = t = 1.27$ mm; $\Delta x = \Delta y = \Delta z = w/4$.



(b) Comparison of AR-extrapolated field record with an extended FDTD data set.

Fig. 15.7 Application of a 50th-order covariance-based AR model to extrapolate a windowed FDTD data segment (2,500 time-steps) to 10,000 time-steps. *Source:* Chen et al., *IEEE Trans. Microwave Theory and Techniques*, 1994, pp. 1992–1997, © 1994 IEEE.

Forward-Backward Method

The results of a covariance-based AR model can be sensitive to the order of the model because its a_i parameters often have low accuracy. The low accuracy results, in turn, because the ensemble calculation that is rigorously required to find the covariance is instead substituted in these methods by using the law of large numbers, and by approximating the covariance with inexact functions of the known time signal. A useful alternative approach based on an unconstrained least-squares estimation of the AR parameters was proposed independently in [29, 30] and calculated more efficiently in [31]. This approach, called the forward-backward prediction method [30], avoids the problems of the covariance-based AR models by working directly with the time-domain data, rather than calculating the covariance functions of the data. Following [31], the method involves setting up and solving the following $(p+1)\times(p+1)$ system of linear equations:

$$\begin{bmatrix} r(0,0) & \cdots & r(0,p) \\ \vdots & \ddots & \vdots \\ r(p,0) & \cdots & r(p,p) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} e_p \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{15.41}$$

where, for $0 \le \alpha$ and $\beta \le p$, we have

$$r(\alpha,\beta) = \sum_{\ell=1}^{N-p} \left(f\big|_{i,j,k}^{M+p+\ell-\beta} f\big|_{i,j,k}^{M+p+\ell-\alpha} + f\big|_{i,j,k}^{M+\ell+\beta} f\big|_{i,j,k}^{M+\ell+\alpha} \right) \tag{15.42a}$$

$$e_p = \sum_{\ell=0}^{p} a_\ell \, r(0,\ell) \tag{15.42b}$$

Reference [32] was the first to report the application of the forward-backward method to AR extrapolation of FDTD data records and corresponding spectral estimation. Results were given for a test problem involving a rectangular PEC plate embedded in a high-permittivity half-space and excited by a sinusoidal point source located in the free-space region. Both an autocorrelation-based AR model and the forward-backward AR model were "trained" on the same FDTD field record, a set of datapoints from 201 to 800 time-steps that were "appropriately decimated and lowpass filtered." Both models were then used to generate an extrapolated sequence extending over 5,000 time-steps. The forward-backward method was found to provide more accurate spectra and achieve better computational efficiency, since its order was much lower ($p = 5$ versus $p = 50$).

Nonlinear Predictor

In addition to the use of covariance-based AR linear prediction to extrapolate FDTD waveform records, [27] reported the application of a nonlinear predictor for the same purpose. The form of this predictor is the *exponential autoregressive* (EXPAR) model of order p given by [33, 34]:

$$f\big|_{i,j,k}^{n} = \sum_{\ell=1}^{P}\left\{a_\ell + b_\ell \exp\left[-\delta\left(f\big|_{i,j,k}^{n-\ell}\right)^2\right]\right\} f\big|_{i,j,k}^{n-\ell} + \text{error term} \tag{15.43}$$

where $\{a\}$, $\{b\}$, and δ are the modeling parameters that are "trained" from the windowed and decimated FDTD field-versus-time record.

The EXPAR model was applied in [27] to the same large double-stub microstrip problem discussed earlier, and similarly used to extrapolate an 8,900-step FDTD data record to 100,000 time-steps. As shown in Fig. 15.8, the nonlinear EXPAR model provided excellent accuracy for S_{21} over a wide dynamic range upon Fourier transformation of the 100,000-step record. Moreover, the EXPAR results were much less sensitive to the model order than were the results of the 91st-order covariance-based linear AR model previously discussed in [27]. However, the reported order of the EXPAR model in this paper exceeded 80, which is still very high. A necessary line of inquiry involves applying the forward-backward linear AR model to problems of this size and complexity, to establish whether its superior characteristics can markedly reduce the required model order.

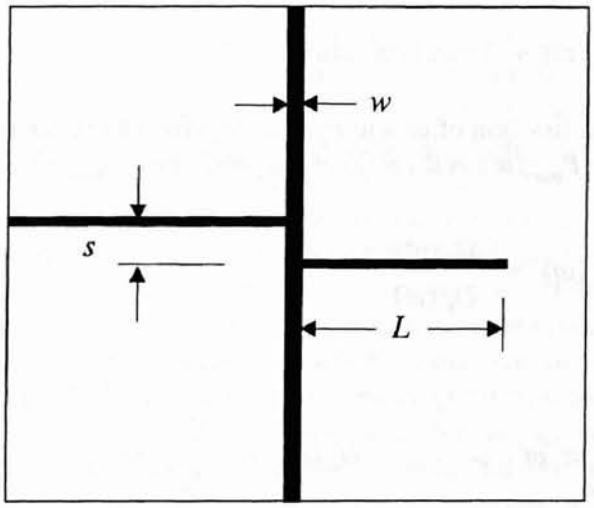
15.8.3 Padé Approximation

In [35], Dey and Mittra reported an efficient, robust alternative to Prony's method (Section 15.8.1) and the generalized pencil-of-function technique [36] for calculating the resonant frequencies and Q factors of cavities from a small time window. Unlike Prony's method and the generalized pencil-of-function technique, both of which represent the FDTD-computed field time response as a sum of exponentials, the Dey-Mittra approach is based upon the use of the Padé approximation [37] in conjunction with the FFT. Because its accuracy appears to be less sensitive to the field-sampling conditions, the Padé approximation technique may be more suited for use in a general-purpose FDTD simulation package.

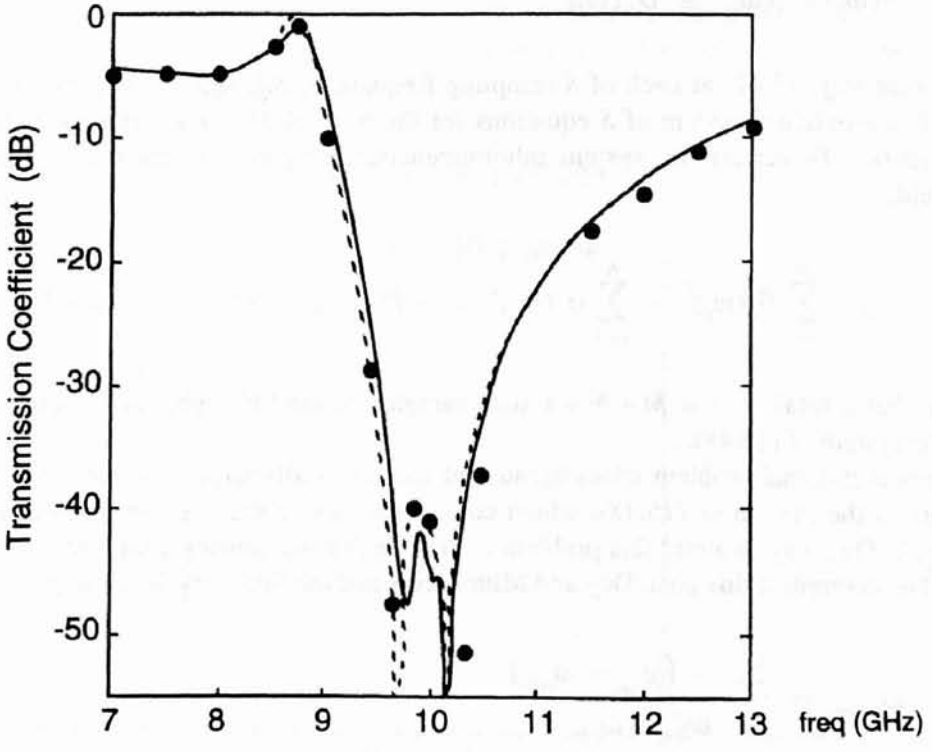
Formulation

To obtain cavity resonant frequencies and Q's using FDTD, we must transform the cavity's field-versus-time response obtained from the FDTD simulation to the frequency domain, for example by using the FFT. An inherent limitation of this approach is its inadequacy in frequency resolution, which is reciprocal to the product of the total number of time-steps and the time-step size. Thus, to achieve some desired frequency resolution, it is necessary to run the FDTD simulation for a sufficiently long time.

Dey and Mittra employ a two-step procedure to overcome the above limitation. First, they apply the FFT to the FDTD-calculated field-versus-time data of one of the six electromagnetic field components to obtain the cavity's spectral response. Then, they process the spectral data using the Padé approximation to improve the accuracy of the frequency response.



(a) Plan view of the double-stub microstrip structure. Key dimensions: $w = 0.122$ mm, $L = 2.92$ mm, $s = 0.76$ mm, $h = 12.7$ mm, $\Delta x = h/4$, $\Delta y = \Delta z = w/4$, $\varepsilon_r = 9.9$.



(b) Comparison of EXPAR and benchmark data for S_{21}.

Fig. 15.8 Use of covariance-based EXPAR model of order greater than 80 to model S_{21} of a double-stub device by extrapolating the FDTD record from 8,900 to 100,000 time-steps. *Source:* Chen et. al., *IEEE Trans. Microwave Theory and Techniques*, 1994, pp. 1992–1997, © 1994 IEEE.

The frequency response can be represented as

$$P(\omega) = P_{\text{poles}}(\omega) + \text{remainder term} \tag{15.44}$$

where $P(\omega)$ is a complex function of ω, and $P_{\text{poles}}(\omega)$ contains all of the poles of $P(\omega)$. The Padé approximation expresses $P_{\text{poles}}(\omega)$ as the ratio of the polynomials $Q_N(\omega)$ and $D_M(\omega)$, such that

$$P(\omega) \cong P_{\text{poles}}(\omega) = \frac{Q_N(\omega)}{D_M(\omega)} \tag{15.45}$$

where

$$Q_N(\omega) = \sum_{i=0}^{N} \alpha_i \omega^i \quad ; \qquad D_M(\omega) = \sum_{i=0}^{M} \beta_i \omega^i \tag{15.46a, b}$$

The unknown coefficients $\{\alpha_i\}$ and $\{\beta_i\}$ in (15.46) can now be obtained. Let $P(\omega_j)$ denote the value of the FFT of the FDTD field-versus-time response as sampled at the frequency ω_j. Then, we rewrite (15.45) as

$$P(\omega_j)\, D_M(\omega_j) \cong Q_N(\omega_j) \tag{15.47}$$

Upon enforcing (15.47) at each of S sampling frequencies $\omega_0, \omega_1, \ldots, \omega_{S-1}$ within the range of the FFT, we obtain a system of S equations for the total of $N + M + 2$ unknown $\{\alpha_i\}$ and $\{\beta_i\}$ coefficients. To render the system inhomogeneous, Dey and Mittra set β_0 equal to unity. This yields

$$P(\omega_j) \sum_{i=1}^{M} \beta_i (\omega_j)^i - \sum_{i=0}^{N} \alpha_i (\omega_j)^i = -P(\omega_j), \qquad j = 0, 1, \ldots, S-1 \tag{15.48}$$

We see that a total of $S \geq M + N + 1$ data samples of the FFT spectral response is needed to solve the system of (15.48).

A computational problem arises because of the potentially large dynamic range of the matrix elements in the system of (15.48), which contain powers of the angular frequency up to $(\omega_j)^M$ and $(\omega_j)^N$. One way to avoid this problem is to scale the frequencies used, such that each is near unity. To accomplish this goal, Dey and Mittra proposed the following frequency scaling:

$$\omega_{\text{scaled}} = \frac{2\omega - (\omega_{\max} - \omega_{\min})}{\omega_{\max} + \omega_{\min}} \tag{15.49}$$

where $\omega_{\max}$ and $\omega_{\min}$ are the maximum and minimum angular frequencies of the samples used.

Upon solution of (15.48), the coefficients $\{\alpha_i\}$ and $\{\beta_i\}$ are known. Then, it is straightforward to interpolate the sampled data for $P(\omega)$ to obtain the desired resolution. In the application reported in [35], N is chosen to equal M. This is called the diagonal Padé approximation, and requires samples of the FFT response at $2N + 1$ frequencies.

Numerical Results

We now review results obtained using the Padé approximation technique, as reported by Dey and Mittra in [35]. These involve calculation of the resonant frequencies and Q's of a $2.286 \times 1.016 \times 2.286$-mm rectangular cavity with lossless walls, where the cavity is filled with a slightly lossy medium having the conductivity 5.162×10^{-4} S/m. The corresponding spatial discretizations in the FDTD lattice are 0.057, 0.058, and 0.057 mm.

In the first numerical experiment, the resonant frequencies and Q's are determined using: (1) the FFT approach alone, and (2) the combined FFT/Padé method working on only 13 samples of the FFT data set, yielding 2,700 interpolated outputs. For total FDTD time-step counts between 4,096 and 32,768, the error in the resonant frequency for the TE_{101} mode obtained using the FFT/Padé method is less than 0.01%, nearly ten times less error than that of the regular FFT approach. While the error in the cavity Q for this mode obtained using the FFT/Padé method is only 0.3%, the frequency resolution for the regular FFT technique is so coarse that the Q cannot be meaningfully obtained even after 32,768 time steps.

Fig. 15.9 vividly illustrates the ability of the FFT/Padé technique to provide high resolution of cavity resonant frequencies and Q factors, using far fewer time-steps than the conventional FFT approach. This figure shows the normalized cavity frequency response calculated after only 1,024 FDTD time-steps. (Note that the normalization in this figure is carried out independently for the two curves to enhance the clarity of the graph.) It is evident that the conventional FFT method is unable to resolve any of the resonances, whereas this is easily accomplished using the FFT/Padé technique. Comparable frequency resolution is obtained using the conventional FFT method only when running the FDTD simulation for *several orders of magnitude* more time-steps than used in the FFT/Padé technique.

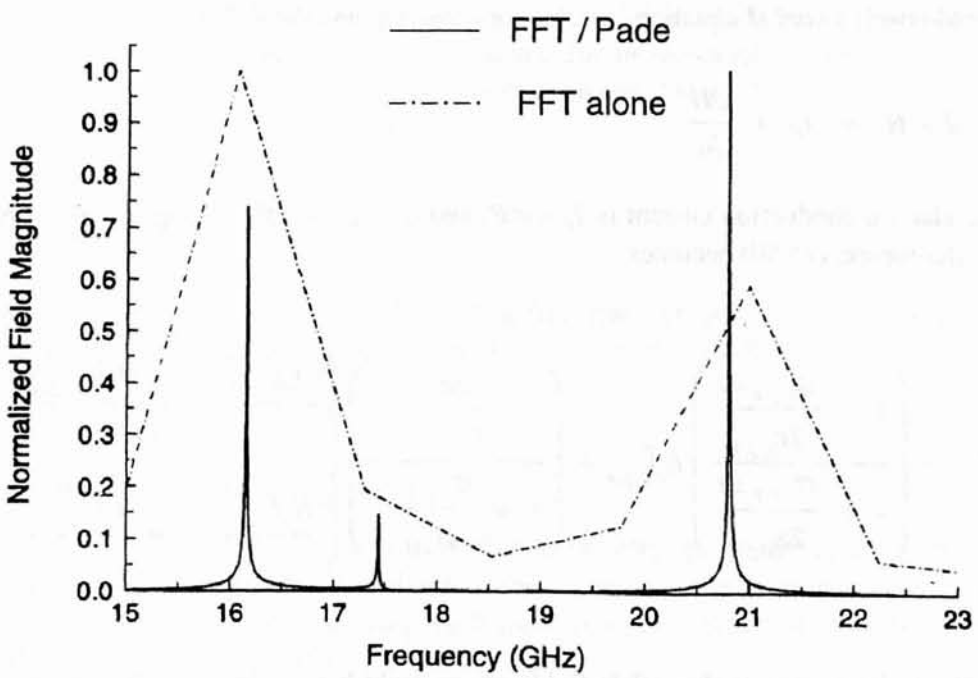


Fig. 15.9 Comparison of FFT and FFT/Padé results for the normalized frequency response of the rectangular cavity for 1,024 time-steps. *Source:* Dey and Mittra, *IEEE Microwave and Guided Wave Lett.*, 1998, pp. 415–417, © 1998 IEEE.

The second numerical experiment in [35] examines the effect of varying the number of input data samples used in the FFT/Padé method for a fixed number (4,096) of FDTD time-steps. It is found that using only 7 frequency samples is sufficient to resolve both the resonant frequency and Q. Using 9, 11, 13, 15, ... frequency samples yields neither improvement nor degradation of the accuracy. Dey and Mittra therefore conclude in [35] that the FFT/Padé technique is quite insensitive to the choice of the input parameters, in contrast to the behavior of Prony's method and the generalized pencil-of-function approach. Their approach appears to be robust.

15.9 MODELING OF LUMPED CIRCUIT ELEMENTS

In this section, three-dimensional FDTD modeling of the connection of linear and nonlinear lumped circuit elements to a microstrip trace is discussed. For convenience in the discussion, the microstrip is assumed to be oriented in the x-direction, and the lumped load is assumed to be oriented in the z-direction. The extension of the theory to other Cartesian orientations in the FDTD space lattice is straightforward.

A resistive source of the type discussed in Section 15.9.3 is used in all of the FDTD simulations of this section. The source resistance is matched to the line impedance so that retroreflections at the source are minimized.

15.9.1 FDTD Formulation Extended to Circuit Elements

Assumptions and Notation

Consider Maxwell's ***curl H*** equation, suitable for time-stepping the E-field:

$$\nabla \times \boldsymbol{H} = \boldsymbol{J}_C + \frac{\partial \boldsymbol{D}}{\partial t} \tag{15.50}$$

Here, the electric conduction current is $\boldsymbol{J}_C = \sigma \boldsymbol{E}$, and the electric flux density is $\boldsymbol{D} = \varepsilon \boldsymbol{E}$. Using central differencing, (15.50) becomes

$$E_z\Big|_{i,j,k}^{n+1} = \left(\frac{1 - \dfrac{\sigma_{i,j,k}\,\Delta t}{2\varepsilon_{i,j,k}}}{1 + \dfrac{\sigma_{i,j,k}\,\Delta t}{2\varepsilon_{i,j,k}}} \right) E_z\Big|_{i,j,k}^{n} + \left(\frac{\dfrac{\Delta t}{\varepsilon_{i,j,k}}}{1 + \dfrac{\sigma_{i,j,k}\,\Delta t}{2\varepsilon_{i,j,k}}} \right) \left(\frac{H_y\Big|_{i+1/2,j,k}^{n+1/2} - H_y\Big|_{i-1/2,j,k}^{n+1/2}}{\Delta x} + \frac{H_x\Big|_{i,j-1/2,k}^{n+1/2} - H_x\Big|_{i,j+1/2,k}^{n+1/2}}{\Delta y} \right) \tag{15.51}$$

An important observation is that all H-fields on the right-hand side of (15.51) are at time-step $n + 1/2$, centered in time relative to the stored E-field $E_z\Big|_{i,j,k}^{n}$ and the newly updated E-field $E_z\Big|_{i,j,k}^{n+1}$. Further, the bracketed coefficients are derived assuming that $\boldsymbol{J}_C$ is also evaluated at time-step $n + 1/2$, taking

$$J_C\Big|_{i,j,k}^{n+1/2} = \sigma_{i,j,k} E_z\Big|_{i,j,k}^{n+1/2} = \frac{\sigma_{i,j,k}}{2}\left(E_z\Big|_{i,j,k}^{n} + E_z\Big|_{i,j,k}^{n+1}\right) \tag{15.52}$$

We call this the semi-implicit formulation for the conduction current, since this current relies in part upon the updated E-field to be determined as a result of the time-stepping, and yet does not result in a system of simultaneous equations. This results in a numerically stable algorithm for arbitrary positive values of σ.

For convenience in this chapter, we shall assume that all circuit components are located in a free-space region ($\varepsilon = \varepsilon_0$, $\sigma = 0$, $J_C = 0$). Additional notational convenience is provided by denoting the Yee finite-difference analog to the curl of $\boldsymbol{H}$ observed at $E_z|_{i,j,k}$ as

$$(\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2} = \frac{H_y\Big|_{i+1/2,j,k}^{n+1/2} - H_y\Big|_{i-1/2,j,k}^{n+1/2}}{\Delta x} + \frac{H_x\Big|_{i,j-1/2,k}^{n+1/2} - H_x\Big|_{i,j+1/2,k}^{n+1/2}}{\Delta y} \tag{15.53}$$

Then, (15.51) can be rewritten more simply as

$$E_z\Big|_{i,j,k}^{n+1} = E_z\Big|_{i,j,k}^{n} + \frac{\Delta t}{\varepsilon_0}(\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2} \tag{15.54}$$

Extended FDTD Formulation

Reference [16] reported the modification of the three-dimensional E-field time-stepping algorithm of (15.54) to allow for the addition of lumped linear and nonlinear circuit elements. The basis of this formulation, reported originally by Sui et al. [38] for two-dimensional problems, is that circuit elements can be accounted for in Maxwell's equations by adding a lumped electric current density term $\boldsymbol{J}_L$ to the conduction and displacement currents on the right-hand side of (15.50), which now becomes

$$\nabla \times \boldsymbol{H} = \boldsymbol{J}_C + \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J}_L \tag{15.55}$$

We assume that the lumped element is located in free space at field component $E_z|_{i,j,k}$, and is z-oriented in the grid. Further, we assume that this element provides a local current density related to the total element current I_L by

$$J_L = I_L / \Delta x \, \Delta y \tag{15.56}$$

Here, I_L can be a scalar multiple, time derivative, time integral, or nonlinear function of the electric potential $V = E_z|_{i,j,k} \Delta z$ developed across the element. (Note that the assumed positive direction of I_L is $+z$.) Then, the following modified version of (15.54) specifies the presence of the lumped circuit element in the FDTD lattice:

$$E_z\Big|_{i,j,k}^{n+1} = E_z\Big|_{i,j,k}^{n} + \frac{\Delta t}{\varepsilon_0}(\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2} - \frac{\Delta t}{\varepsilon_0 \, \Delta x \, \Delta y} I_L^{n+1/2} \tag{15.57}$$

Note that the lumped current in (15.57) is evaluated at time $n + 1/2$, as done in (15.52) for the conduction current. Since the lumped current is a function of the E-field at the circuit element, this requires averaging $E_z|_{i,j,k}^{n+1}$ and $E_z|_{i,j,k}^{n}$, and yields a stable semi-implicit time-stepping algorithm. Generalization to x- and y-orientations of a lumped element is straightforward by permuting the coordinate subscripts of the fields.

15.9.2 The Resistor

Assuming a z-directed resistor R located in free space at the field component $E_z|_{i,j,k}$, the voltage-current characteristic that describes its behavior in a semi-implicit manner appropriate for stable operation of the FDTD field solver is [16]

$$I_z|_{i,j,k}^{n+1/2} = \frac{\Delta z}{2R}\left(E_z|_{i,j,k}^{n+1} + E_z|_{i,j,k}^{n}\right) ; \qquad J_L = \frac{I_z|_{i,j,k}^{n+1/2}}{\Delta x\,\Delta y} \tag{15.58a, b}$$

From (15.57), the corresponding time-stepping relation is [16]

$$E_z|_{i,j,k}^{n+1} = \left(\frac{1 - \dfrac{\Delta t\,\Delta z}{2R\varepsilon_0\,\Delta x\,\Delta y}}{1 + \dfrac{\Delta t\,\Delta z}{2R\varepsilon_0\,\Delta x\,\Delta y}}\right) E_z|_{i,j,k}^{n} + \left(\frac{\dfrac{\Delta t}{\varepsilon_0}}{1 + \dfrac{\Delta t\,\Delta z}{2R\varepsilon_0\,\Delta x\,\Delta y}}\right) (\nabla \times \boldsymbol{H})_z|_{i,j,k}^{n+1/2} \tag{15.59}$$

Implementation of (15.59) in the FDTD algorithm requires only redefining the C_a and C_b updating coefficients for the E-field component at the location of the resistor, as defined in Section 3.6.

Reference [16] reported a comparison of the impedance match provided by the numerical resistor of (15.59) and a benchmark physical resistor model composed of a one-cell-thick resistive slab in the FDTD lattice. Using one or the other to terminate a 50Ω transmission line excited by a 90-ps Gaussian pulse (spectral width of 20 GHz), it was found that both resistor models provide reflection coefficients of less than 1% up to 1 GHz. Above 1 GHz, the two resistor models continue to provide very close agreement in terms of reflection coefficient, although, as expected, the match to the line degrades because of parasitic capacitance or inductance present in each model. It is important to note that implementing (15.59) at two adjacent E-field components between the microstrip signal line and ground plane simulates the presence of two parallel resistors terminating the line, halving the effective load resistance.

15.9.3 The Resistive Voltage Source

With the ability to model lumped elements in the context of FDTD, it is a simple matter to model a nonreflecting (matched) source as a resistive voltage source. Again assuming a z-directed lumped element, the voltage-current characteristic that describes the behavior of a resistive voltage source in a semi-implicit manner is [16]

$$I_z\Big|_{i,j,k}^{n+1/2} = \frac{\Delta z}{2R_S}\left(E_z\Big|_{i,j,k}^{n+1} + E_z\Big|_{i,j,k}^{n}\right) + \frac{V_S^{n+1/2}}{R_S} \quad ; \quad J_L = \frac{I_z\Big|_{i,j,k}^{n+1/2}}{\Delta x\,\Delta y} \tag{15.60a, b}$$

where $V_S^{n+1/2}$ is the source voltage and R_S is the internal source resistance. The corresponding time-stepping relation is [16]

$$E_z\Big|_{i,j,k}^{n+1} = \left(\frac{1 - \dfrac{\Delta t\,\Delta z}{2R_S\,\varepsilon_0\,\Delta x\,\Delta y}}{1 + \dfrac{\Delta t\,\Delta z}{2R_S\,\varepsilon_0\,\Delta x\,\Delta y}}\right) E_z\Big|_{i,j,k}^{n} + \left(\frac{\dfrac{\Delta t}{\varepsilon_0}}{1 + \dfrac{\Delta t\,\Delta z}{2R_S\,\varepsilon_0\,\Delta x\,\Delta y}}\right) (\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2}$$

$$+ \left(\frac{\dfrac{\Delta t}{R_S\,\varepsilon_0\,\Delta x\,\Delta y}}{1 + \dfrac{\Delta t\,\Delta z}{2R_S\,\varepsilon_0\,\Delta x\,\Delta y}}\right) V_S^{n+1/2} \tag{15.61}$$

15.9.4 The Capacitor

We next consider the insertion of a numerical lumped capacitor C into the FDTD space lattice. Again assuming a z-directed lumped element located in free space at $E_z|_{i,j,k}$, the voltage-current characteristic that describes the capacitor's behavior in a semi-implicit manner is [16]

$$I_z\Big|_{i,j,k}^{n+1/2} = \frac{C\,\Delta z}{\Delta t}\left(E_z\Big|_{i,j,k}^{n+1} - E_z\Big|_{i,j,k}^{n}\right) \quad ; \quad J_L = \frac{I_z\Big|_{i,j,k}^{n+1/2}}{\Delta x\,\Delta y} \tag{15.62a, b}$$

This formulation differs from that of [38] in that the E-field samples are separated here by one time-step rather than by two. In this manner, we are consistent with the sampling used for the numerical resistor described above. The corresponding time-stepping relation is [16]

$$E_z\Big|_{i,j,k}^{n+1} = E_z\Big|_{i,j,k}^{n} + \left(\frac{\dfrac{\Delta t}{\varepsilon_0}}{1 + \dfrac{C\,\Delta z}{\varepsilon_0\,\Delta x\,\Delta y}}\right) (\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2} \tag{15.63}$$

In (15.63), we assume that the value of C being modeled includes the additive contribution of the intrinsic lattice capacitance defined in Chapter 5, Section 5.4.4.

For the parallel combination of a capacitor C and resistor R located at $E_z|_{i,j,k}$, the results of (15.58), (15.59), (15.62), and (15.63) can be readily combined to yield the following time-stepping relation [16]:

$$E_z\Big|_{i,j,k}^{n+1} = \left(\frac{1 - \dfrac{\Delta t\,\Delta z}{2R\varepsilon_0\,\Delta x\,\Delta y} + \dfrac{C\,\Delta z}{\varepsilon_0\,\Delta x\,\Delta y}}{1 + \dfrac{\Delta t\,\Delta z}{2R\varepsilon_0\,\Delta x\,\Delta y} + \dfrac{C\,\Delta z}{\varepsilon_0\,\Delta x\,\Delta y}} \right) E_z\Big|_{i,j,k}^{n}$$

$$+ \left(\frac{\dfrac{\Delta t}{\varepsilon_0}}{1 + \dfrac{\Delta t\,\Delta z}{2R\varepsilon_0\,\Delta x\,\Delta y} + \dfrac{C\,\Delta z}{\varepsilon_0\,\Delta x\,\Delta y}} \right) (\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2} \tag{15.64}$$

Implementation of (15.64) in the FDTD algorithm is again simple, requiring only an appropriate redefinition of the C_a and C_b updating coefficients for the E-field component at the location of the parallel resistor / capacitor.

Reference [16] reported tests of this FDTD model wherein a variety of numerical capacitive loads were modeled at the end of a long 50 Ω microstrip line (approximately 1 mil scale in the transverse plane) that was excited by a rectangular pulse 1,000 time-steps long. The FDTD-computed voltage response versus time across each capacitor was then compared to the exact theoretical response. Results are shown in Fig. 15.10 for microstrips terminated with 4- and 20-nF capacitors. The theoretical and FDTD curves are indistinguishable.

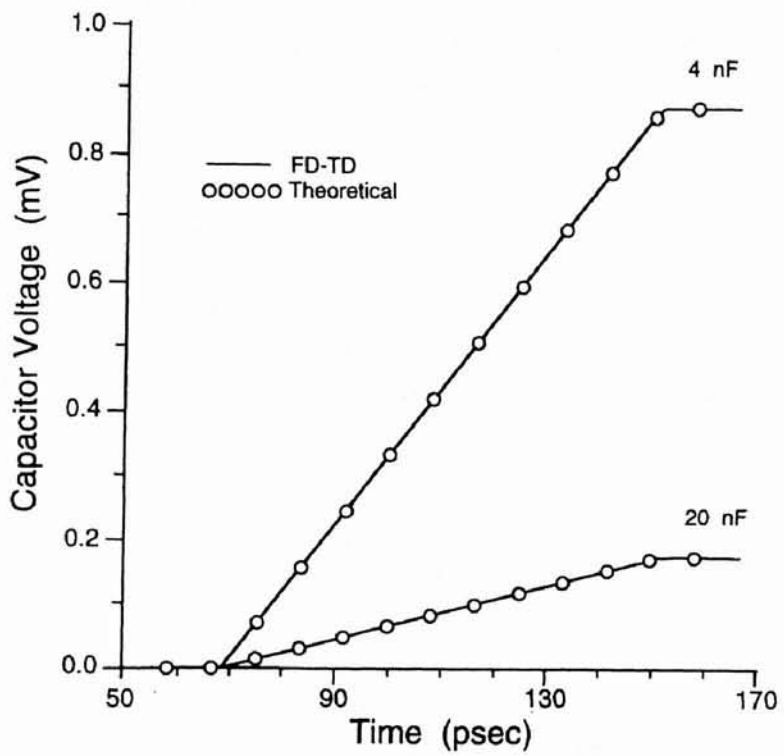


Fig. 15.10 Agreement of FDTD and exact solutions for the voltage across the numerical capacitor terminating a microstrip for two values of capacitance (step incident pulse). *Source:* Piket-May et al., *IEEE Trans. Microwave Theory and Techniques*, 1994, pp. 1514–1523, © 1994 IEEE.

15.9.5 The Inductor

We next consider the insertion of a numerical lumped inductor L into the FDTD space lattice. Again assuming a z-directed lumped element located in free space at $E_z|_{i,j,k}$, the voltage-current characteristic that describes the inductor's behavior in a manner appropriate for stable operation of the FDTD field solver is [16]

$$I_z\Big|_{i,j,k}^{n+1/2} = \frac{\Delta z\,\Delta t}{L}\sum_{m=1}^{n} E_z\Big|_{i,j,k}^{m} \;; \qquad J_L = \frac{I_z\Big|_{i,j,k}^{n+1/2}}{\Delta x\,\Delta y} \tag{15.65a, b}$$

This formulation differs from that of [38] in that the E-field samples are summed only through time-step n, which is consistent with the observation of the lumped current at time-step $n + 1/2$ that is employed throughout this development. The corresponding time-stepping relation is [16]

$$E_z\Big|_{i,j,k}^{n+1} = E_z\Big|_{i,j,k}^{n} + \left(\frac{\Delta t}{\varepsilon_0}\right)(\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2} - \frac{\Delta z\,(\Delta t)^2}{\varepsilon_0\, L\,\Delta x\,\Delta y}\sum_{m=1}^{n} E_z\Big|_{i,j,k}^{m} \tag{15.66}$$

15.9.6 The Arbitrary Two-Terminal Linear Lumped Network

Reference [39] reported an efficient technique to incorporate an arbitrary two-terminal linear lumped network into a single cell of a three-dimensional FDTD space lattice. This approach has been demonstrated to have good accuracy and numerical stability.

Following [39], we first express the impedance of the two-terminal linear lumped network of interest in the form of a rational function as

$$Z(s) = \frac{V(s)}{I(s)} = \sum_{r=0}^{R} a_r (1/s)^r \Bigg/ \sum_{m=0}^{M} b_m (1/s)^m \tag{15.67}$$

where $s = j\omega$, and it is assumed that the order of the numerator and the denominator of the rational function is R and M, respectively. The sets of coefficients $\{a_r\}$ and $\{b_m\}$ are determined strictly by the values of the resistors, inductors, and capacitors comprising the network. From $\{a_r\}$ and $\{b_m\}$, the following composite coefficients are calculated:

$$p_1 = \sum_{m=0}^{M} b_m (\Delta t)^m \Bigg/ \sum_{r=0}^{R} a_r (\Delta t)^r \;; \qquad p_2 = 1 \Bigg/ \sum_{r=0}^{R} a_r (\Delta t)^r \tag{15.68a, b}$$

Now, we can begin the actual time-stepping procedure.

Step 1: Update Auxiliary Variables $\tilde{V}_{(m)}$ Via a Recursion Relation

$$\tilde{V}_{(m=0)}^{n+1} = 0 \tag{15.69a}$$

$$\tilde{V}_{(m)}^{n+1} = \tilde{V}_{(m)}^{n} + \tilde{V}_{(m-1)}^{n+1}\Delta t + E_z\Big|_{i,j,k}^{n}\Delta z\,(\Delta t)^m \;, \quad m = 1, 2, \ldots, M \tag{15.69b}$$

Step 2: Update Auxiliary Variables $\tilde{I}_{(r)}$ Via a Recursion Relation

$$\tilde{I}_{(r=0)}^{n+1/2} = 0 \tag{15.70a}$$

$$\tilde{I}_{(r)}^{n+1/2} = \tilde{I}_{(r)}^{n-1/2} + \tilde{I}_{(r-1)}^{n+1/2}\Delta t + I^{n-1/2}(\Delta t)^r \quad , \quad m = 1, 2, \ldots, M \tag{15.70b}$$

Step 3: Update the Composite Auxiliary Variable Q

$$Q^{n+1} = \sum_{m=1}^{M} b_m \tilde{V}_{(m)}^{n+1} - \sum_{r=1}^{R} a_r \tilde{I}_{(r)}^{n+1/2} \tag{15.71}$$

Step 4: Update E_z at the FDTD Lattice Cell Containing the Lumped Network

$$E_z\Big|_{i,j,k}^{n+1} = \frac{E_z\Big|_{i,j,k}^{n} + \dfrac{\Delta t}{\varepsilon_0}(\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2} - \dfrac{p_2 \Delta t}{\varepsilon_0 \Delta x \Delta y} Q^{n+1}}{1 + \dfrac{p_1 \Delta t \Delta z}{\varepsilon_0 \Delta x \Delta y}} \tag{15.72}$$

Step 5: Update the Auxiliary Current Variable I

$$I^{n+1/2} = p_1 E_z\Big|_{i,j,k}^{n+1} \Delta z + p_2 Q^{n+1} \tag{15.73}$$

This completes the updating sequence at the FDTD space cell containing the lumped network of interest. Subsequently, the Step 1 to Step 5 sequence is repeated during each cycle of E-field updates in the space lattice. A total of $M + R + 2$ auxiliary variables is needed to implement the algorithm at the cell where the lumped network is embedded.

Reference [39] reported a test of this algorithm for a parallel-strip transmission line of length $l = 240$ mil, width $w = 8$ mil, substrate thickness $h = 16$ mil, and substrate permittivity $\varepsilon_r = 2.2$. The line was assumed to be driven by pulsed 50Ω voltage source having a bandwidth of 100 GHz, and terminated by a seven-element lumped network having the configuration of Fig. 15.11:

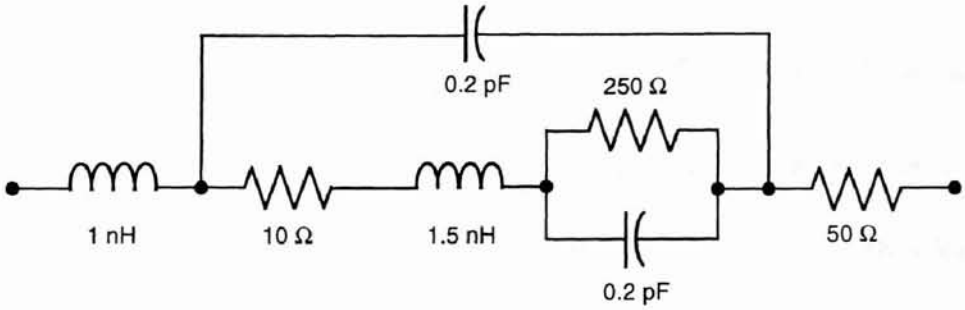
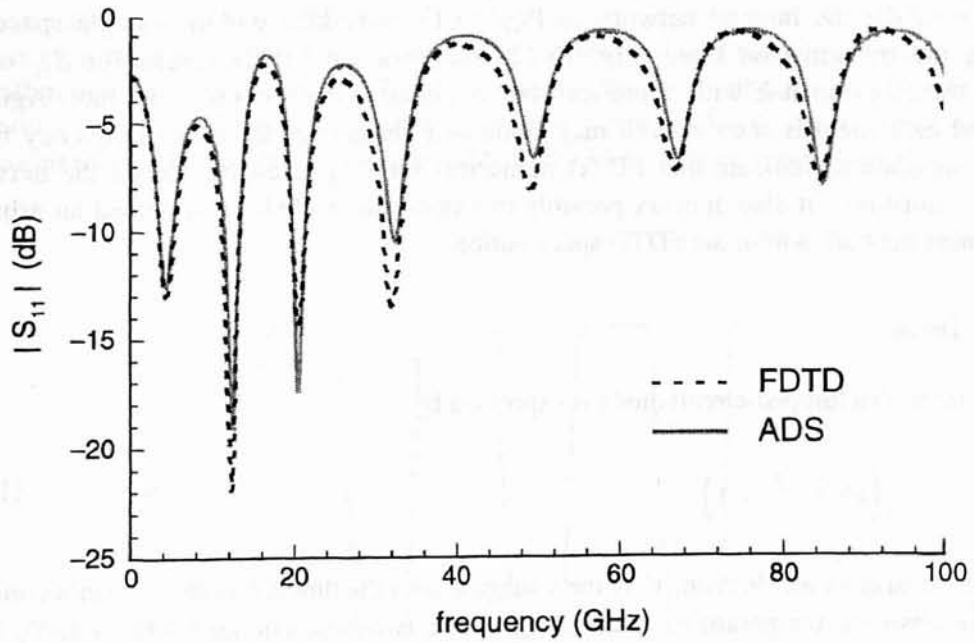
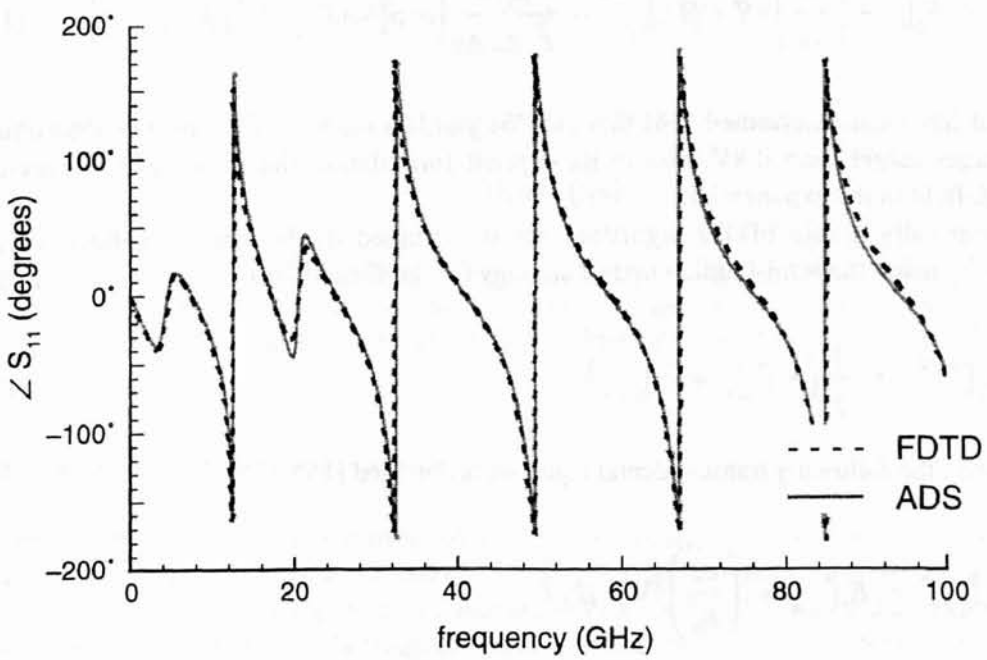


Fig. 15.11 Test circuit used to validate modeling of an arbitrary two-terminal linear lumped network embedded within a single FDTD space-lattice cell. *Adapted from:* Wu et al., *IEEE Microwave and Wireless Components Lett.*, 2004, pp. 74–76, © 2004 IEEE.



(a) Magnitude.



(b) Phase.

Fig. 15.12 Agreement of FDTD and Agilent ADS results for S_{11} of a transmission line terminated by the seven-element network of Fig. 15.11. *Adapted from:* Wu et al., *IEEE Microwave and Wireless Components Lett.*, 2004, pp. 74–76, © 2004 IEEE.

The FDTD model of the parallel-strip transmission line was implemented using a cubic space lattice of cell size $\Delta = 4$ mil with a time-step $\Delta t = 0.25$ ps. The algorithm of this section was used to model the lumped network of Fig. 15.11 embedded within a single space cell terminating the transmission line. Fig. 15.12 compares the FDTD results for S_{11} of the terminated transmission line with numerical data obtained using a commercial tool, Agilent's ADS. Good agreement is seen in both magnitude and phase over the entire frequency range. Additional simulations indicate that FDTD numerical stability is unaffected by the network-embedding algorithm. It also appears possible to extend this technique to embed an arbitrary multiport linear network within the FDTD space lattice.

15.9.7 The Diode

The current through a lumped-circuit diode is expressed by

$$I_d = I_0\left(e^{qV_d/kT} - 1\right) \tag{15.74}$$

where q is the charge of an electron, V_d is the voltage across the diode, k is Boltzmann's constant, and T is the absolute temperature. According to the two-dimensional study of [38], for a z-directed diode located in free-space at $E_z|_{i,j}$, the E-field time-stepping relation is given by

$$E_z\Big|_{i,j}^{n+1} = E_z\Big|_{i,j}^{n} + \left(\frac{\Delta t}{\varepsilon_0}\right)(\nabla \times \boldsymbol{H})_z\Big|_{i,j}^{n+1/2} - \frac{I_0\,\Delta t}{\varepsilon_0\,\Delta x\,\Delta y}\left[\exp\left(-q\,E_z\Big|_{i,j}^{n}\,\Delta z \,\Big/\, kT\right) - 1\right] \tag{15.75}$$

However, it has been determined [16] that (15.75) yields a numerically unstable algorithm for diode voltages larger than 0.8V, due to its explicit formulation that employs the previously computed E-field in the exponential.

A numerically stable FDTD algorithm for the lumped diode can be realized in three dimensions by using the semi-implicit update strategy for the E-field [16]

$$E_z\Big|_{i,j,k}^{n+1/2} = \frac{1}{2}\left(E_z\Big|_{i,j,k}^{n+1} + E_z\Big|_{i,j,k}^{n}\right) \tag{15.76}$$

In this manner, the following transcendental equation is obtained [16]:

$$E_z\Big|_{i,j,k}^{n+1} = E_z\Big|_{i,j,k}^{n} + \left(\frac{\Delta t}{\varepsilon_0}\right)(\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2} - \frac{I_0\,\Delta t}{\varepsilon_0\,\Delta x\,\Delta y}\left\{\exp\left[-q\left(E_z\Big|_{i,j,k}^{n+1} + E_z\Big|_{i,j,k}^{n}\right)\Delta z \,\Big/\, 2kT\right] - 1\right\} \tag{15.77}$$

Upon solving (15.77) for the updated E-field using Newton's method, the new model is found to be numerically stable over a diode voltage range up to 15V.

Reference [16] reported numerical results for a diode with a saturation current of 1×10^{-14} A located at the end of a 50 Ω microstrip line (approximately 1 mil scale in the transverse plane). The excitation was a matched resistive voltage source providing a 10V, 200-MHz sinusoid. This high-level source was selected both to challenge the numerical stability of the diode algorithm and to test whether FDTD can properly simulate driving a diode into hard clipping. As shown in Fig. 15.13, there is excellent agreement of the diode voltage response versus time as calculated by FDTD and SPICE. No instability of the FDTD solver is observed.

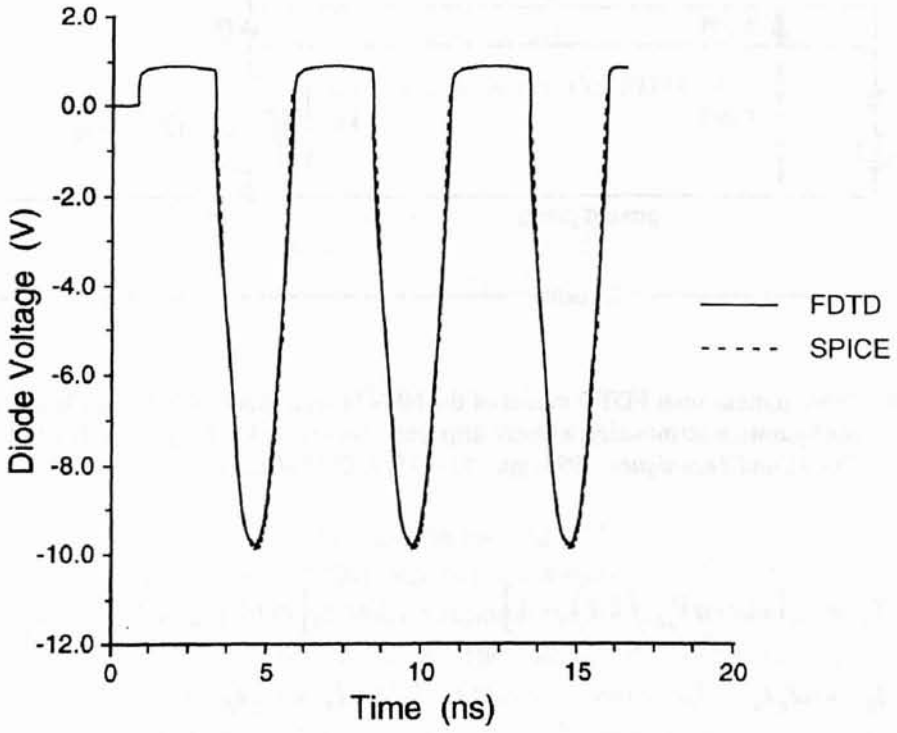


Fig. 15.13 Agreement of FDTD and SPICE calculations for the voltage across a diode terminating a microstrip line and driven into hard clipping by a sinusoidal RF source. *Source:* Piket-May et al., *IEEE Trans. Microwave Theory and Techniques*, 1994, pp. 1514–1523, © 1994 IEEE.

15.9.8 The Bipolar Junction Transistor

Reference [38] presented a two-dimensional FDTD model for the linearized, small-signal behavior of the bipolar junction transistor. While good results were obtained using this model for selected parameters, some problems remained: (1) numerical instability for very high or low values of base resistance r_b, (2) fringing fields between the base and collector terminals, and (3) lack of generality for large-signal and digital switching applications.

Referring to Fig. 15.14, we now consider a numerically stable FDTD algorithm for the lumped NPN bipolar junction transistor in three dimensions [16]. This algorithm permits study of the large-signal behavior, including switching. In the example discussed here, the transistor is assumed to terminate a microstrip in a grounded-emitter configuration. The simulation is based upon a simple Ebers-Moll transistor model described by the following circuit equations [40]:

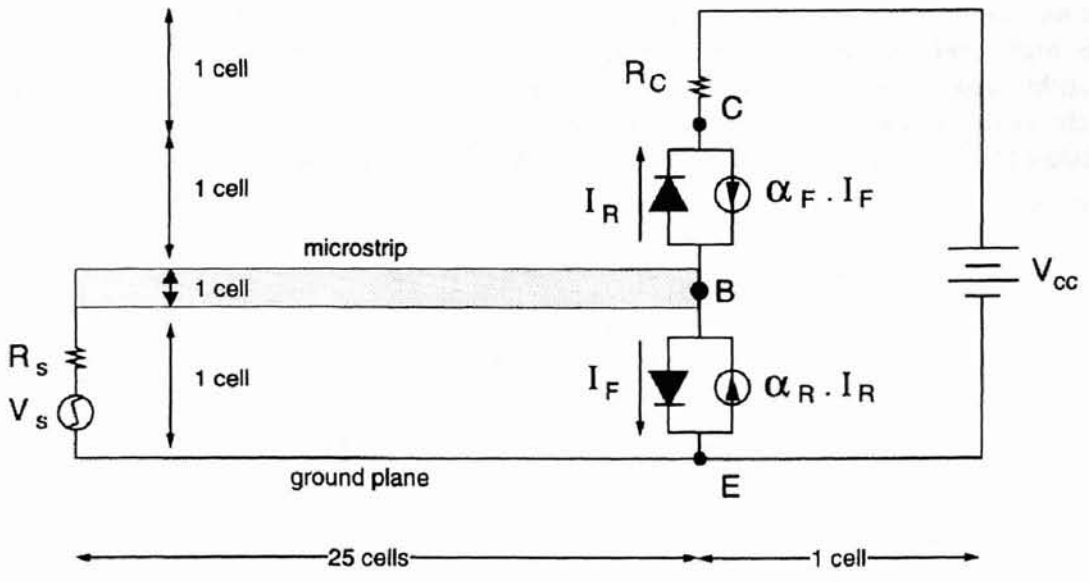


Fig. 15.14 Three-dimensional FDTD model of the NPN bipolar junction transistor in the common-emitter configuration terminating a microstrip line. *Source:* Piket-May et al., *IEEE Trans. Microwave Theory and Techniques*, 1994, pp. 1514–1523, © 1994 IEEE.

$$I_F = I_0\left[\exp\left(q V_{BE} / kT\right) - 1\right] \; ; \qquad I_R = I_0\left[\exp\left(q V_{BC} / kT\right) - 1\right] \tag{15.78a, b}$$

$$I_E = \alpha_R I_R - I_F \; ; \qquad I_C = I_R - \alpha_F I_F \tag{15.79a, b}$$

Now, assuming a transistor that is located in free space and oriented in the z-direction in the FDTD space lattice, we can use a semi-implicit strategy to express the base-emitter voltage V_{BE} in terms of $E_z|_{EB}$, which is the FDTD-computed E-field in the one-cell gap between the ground plane and the end of the microstrip [16]:

$$V_{BE}^{n+1/2} = -\frac{\Delta z}{2}\left(E_z\big|_{EB}^{n+1} + E_z\big|_{EB}^{n}\right) \tag{15.80}$$

A similar semi-implicit strategy is used to express the base-collector voltage V_{BC} in terms of $E_z|_{BC}$, which is the FDTD-computed E-field in the one-cell gap between the end of the microstrip and the collector load:

$$V_{BC}^{n+1/2} = \frac{\Delta z}{2}\left(E_z\big|_{BC}^{n+1} + E_z\big|_{BC}^{n}\right) \tag{15.81}$$

Substituting (15.80) and (15.81) into (15.78) and (15.79), we obtain:

$$I_{\mathrm{E}}^{n+1/2} = \alpha_R I_0 \left\{ \exp\left[q\left(E_z\big|_{\mathrm{BC}}^{n+1} + E_z\big|_{\mathrm{BC}}^{n} \right) \Big/ 2kT \right] - 1 \right\} - I_0 \left\{ \exp\left[-q\left(E_z\big|_{\mathrm{EB}}^{n+1} + E_z\big|_{\mathrm{EB}}^{n} \right) \Big/ 2kT \right] - 1 \right\} \tag{15.82}$$

$$I_{\mathrm{C}}^{n+1/2} = I_0 \left\{ \exp\left[q\left(E_z\big|_{\mathrm{BC}}^{n+1} + E_z\big|_{\mathrm{BC}}^{n} \right) \Big/ 2kT \right] - 1 \right\} - \alpha_F I_0 \left\{ \exp\left[-q\left(E_z\big|_{\mathrm{EB}}^{n+1} + E_z\big|_{\mathrm{EB}}^{n} \right) \Big/ 2kT \right] - 1 \right\} \tag{15.83}$$

Now we obtain two coupled transcendental equations for the FDTD E-field updates at the transistor:

$$E_z\big|_{\mathrm{EB}}^{n+1} = E_z\big|_{\mathrm{EB}}^{n} + \left(\frac{\Delta t}{\varepsilon_0} \right) (\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2} + \frac{\Delta t}{\varepsilon_0 \, \Delta x \, \Delta y} I_{\mathrm{E}}^{n+1/2} \tag{15.84}$$

$$E_z\big|_{\mathrm{BC}}^{n+1} = E_z\big|_{\mathrm{BC}}^{n} + \left(\frac{\Delta t}{\varepsilon_0} \right) (\nabla \times \boldsymbol{H})_z\Big|_{i,j,k}^{n+1/2} + \frac{\Delta t}{\varepsilon_0 \, \Delta x \, \Delta y} I_{\mathrm{C}}^{n+1/2} \tag{15.85}$$

This system can be solved using the Newton-Raphson method [16].

Reference [16] reported FDTD modeling results for a transistor at $T = 300°\mathrm{K}$ ($I_0 = 10^{-16}$ A, $\alpha_R = 0.5$, and $\alpha_F = 0.9901$) located at the end of a 50Ω microstrip line (approximately 1 mil scale in the transverse plane) in the manner of Fig. 15.14. The collector dc supply was included in the electromagnetic simulation. Both the active ($R_{\mathrm{C}} = 50\Omega$) and saturated ($R_{\mathrm{C}} = 10\Omega$) regions of operation were observed for a step-function excitation of the microstrip. A typical result is shown in Fig. 15.15, where the FDTD-computed base-to-emitter voltage waveform is compared with that obtained by a SPICE model. Excellent agreement is observed.

15.10 DIRECT LINKING OF FDTD AND SPICE

The computational technology discussed so far in this chapter demonstrates the feasibility of combining full-wave FDTD Maxwell's equations' models of complex three-dimensional multilayer circuit boards and modules with generic linear and nonlinear loads. At this point, one key question remains to be answered:

> Must we "reinvent the wheel" for the lumped-circuit side of the model? That is, for each electronic device of interest, must we develop from scratch a specialized time-marching circuit model appropriate for insertion into the FDTD space lattice and algorithm? Or, can we somehow build upon the substantial existing knowledge base inherent in popular circuit-modeling tools such as SPICE?

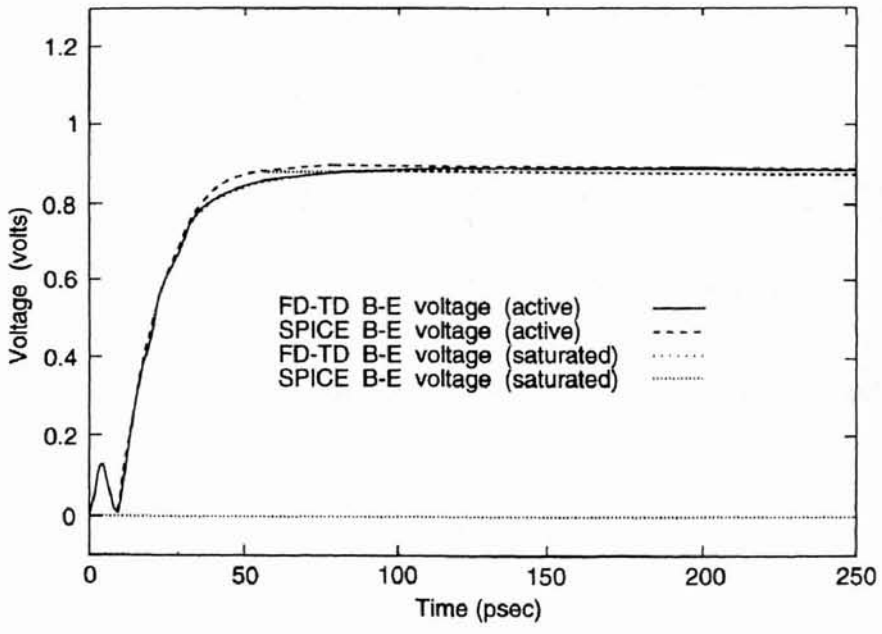


Fig. 15.15 Agreement of FDTD and SPICE calculations for the base-to-emitter voltage waveforms of the transistor of Fig. 15.14 for a step incident pulse. *Source:* Piket-May et al., *IEEE Trans. Microwave Theory and Techniques*, 1994, pp. 1514–1523, © 1994 IEEE.

Reference [41] demonstrated that it is possible to link FDTD and SPICE. This forms an interface between time-stepping the set $\{\boldsymbol{E}, \boldsymbol{H}\}$ of vector electromagnetic fields in a volume of space and time-stepping the set $\{V, I\}$ of scalar voltages and currents characterizing an electronic circuit embedded within this volume. Effectively, the SPICE link permits the systematic construction of FDTD subgrid models of entire circuits that incorporate all important aspects of the circuit physics including gains, nonlinearities, device parasitics, and time delays.

15.10.1 Basic Idea

The basic idea underlying the FDTD–SPICE link can be expressed simply. Consider a multiport lumped circuit embedded within the FDTD space lattice. We want to conveniently and self-consistently interface the physics of this circuit at each of its ports with the distributed electromagnetic field physics of the entire space lattice. Therefore, we ask the question:

> Can we define Norton or Thevenin equivalent circuits "looking into" the FDTD space lattice from the local perspective of each port of the embedded circuit?

If this is possible, the Norton or Thevenin equivalent circuits for the local physics of the FDTD space lattice could be combined with the lumped circuit of interest at its corresponding ports, and the combined circuit time-stepped using SPICE. The resulting updated port voltage or current would then be interpreted as an equivalent updated port E- or H-field, which would then be fed back to the FDTD space lattice. Fields elsewhere in the space lattice would be time-stepped in the usual manner of the Yee algorithm. Repetition of these steps would synchronously time-march the lattice $\{\boldsymbol{E}, \boldsymbol{H}\}$ and the circuit $\{V, I\}$ until evolution to the desired endpoint.

15.10.2 Norton Equivalent Circuit "Looking Into" the FDTD Space Lattice

Fig. 15.16 illustrates how we can define a Norton equivalent circuit "looking into" the FDTD space lattice from the two-terminal port of an embedded circuit. We assume that the port is located in free space at the site of E_w, a particular E_x, E_y, or E_z component in a cubic-cell lattice.

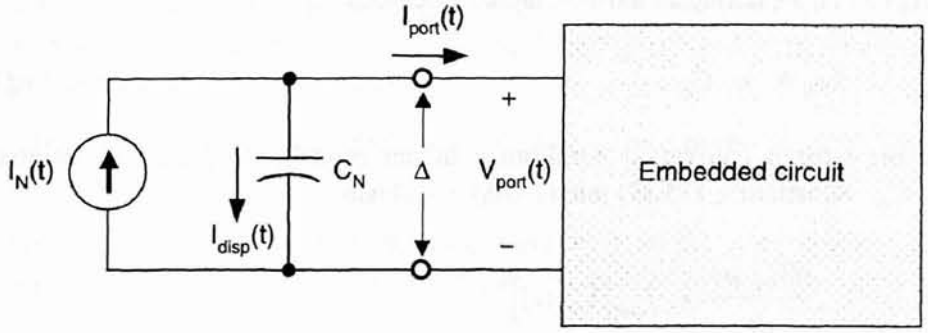


Fig. 15.16 Schematic diagram of a two-terminal port of a circuit embedded in a single cell of the FDTD space lattice, and the Norton equivalent circuit "looking into" the FDTD lattice from this port.

From the Norton equivalent circuit standpoint, the excitation provided by the complete FDTD lattice to this port can be distilled down to the action of the four immediately adjacent $\boldsymbol{H}$-components that loop around E_w. Applying Ampere's law, the total Norton current $I_N(t)$ is obtained from these looping $\boldsymbol{H}$-components by implementing the closed contour integral

$$I_N(t) = \oint_\Gamma \boldsymbol{H}(t) \cdot d\boldsymbol{\Gamma} \tag{15.86}$$

where the contour Γ bounds the space-cell cross section A where the circuit port is embedded. For example, the following expressions implement (15.86) for the three possible Cartesian orientations of $I_N(t)$ within a cubic space cell:

$$I_x\Big|_{i,j,k}^{n+1/2} = \left(H_y\Big|_{i,j,k-1/2}^{n+1/2} - H_y\Big|_{i,j,k+1/2}^{n+1/2} + H_z\Big|_{i,j+1/2,k}^{n+1/2} - H_z\Big|_{i,j-1/2,k}^{n+1/2}\right)\Delta \tag{15.87a}$$

$$I_y\Big|_{i,j,k}^{n+1/2} = \left(H_x\Big|_{i,j,k+1/2}^{n+1/2} - H_x\Big|_{i,j,k-1/2}^{n+1/2} + H_z\Big|_{i-1/2,j,k}^{n+1/2} - H_z\Big|_{i+1/2,j,k}^{n+1/2}\right)\Delta \tag{15.87b}$$

$$I_z\Big|_{i,j,k}^{n+1/2} = \left(H_x\Big|_{i,j-1/2,k}^{n+1/2} - H_x\Big|_{i,j+1/2,k}^{n+1/2} + H_y\Big|_{i+1/2,j,k}^{n+1/2} - H_y\Big|_{i-1/2,j,k}^{n+1/2}\right)\Delta \tag{15.87c}$$

From Fig. 15.16, we note that $I_N(t)$ sources two currents flowing in parallel through the space cell, the displacement current $I_{disp}(t)$ and the port current $I_{port}(t)$:

$$I_N(t) = I_{disp}(t) + I_{port}(t) \tag{15.88}$$

From Chapter 5, Section 5.4.4, $I_{\text{disp}}(t)$ can be related to the voltage across the space cell [which equals the voltage $V_{\text{port}}(t)$ across the embedded circuit port] as

$$I_{\text{disp}}(t) = C_{\text{lattice}} \frac{dV_{\text{port}}(t)}{dt} \tag{15.89a}$$

where, adapting (5.21), we define an intrinsic lattice space-cell capacitance

$$C_{\text{lattice}} = 3\varepsilon_0 \Delta \equiv C_{\text{N}} \tag{15.89b}$$

that provides the Norton equivalent admittance in our model. In fact, we designate this capacitance as C_{N}. Substituting (15.89) into (15.88), we obtain

$$I_{\text{N}}(t) = C_{\text{N}} \frac{dV_{\text{port}}(t)}{dt} + I_{\text{port}}\left[V_{\text{port}}(t)\right] \tag{15.90}$$

where $I_{\text{port}}(V_{\text{port}})$ is the known (possibly nonlinear) current-voltage characteristic of the circuit port.

We see that (15.90) is merely a statement of Kirchhoff's current law applied to the circuit of Fig. 15.16. If this equation could be integrated in time to evolve $V_{\text{port}}(t)$, then it would be a simple matter to obtain the corresponding E-field across the cell containing the circuit port, since

$$E_w(t) = -V_{\text{port}}(t)/\Delta \tag{15.91}$$

This would complete the updating of all of the E-fields in the space lattice, since for all cells where circuit ports are not present, the E-fields are updated by the conventional Yee algorithm. Subsequently, all of the H-fields in the lattice could be updated according to the standard Yee algorithm, and the complete cycle would begin anew.

One possibility is to perform a direct time integration of (15.90) using a standard technique such as the Runge-Kutta method. Specifically, we could start at time t_0 and integrate to time $t_0 + \Delta t$, assuming knowledge of $V_{\text{port}}(t_0)$, $I_{\text{N}}(t)$, and $I_{\text{port}}(V_{\text{port}})$:

$$V_{\text{port}}(t_0 + \Delta t) = V_{\text{port}}(t_0) + \frac{1}{C_{\text{N}}} \int_{t_0}^{t_0+\Delta t} \left\{ I_{\text{N}}(t) - I_{\text{port}}\left[V_{\text{port}}(t)\right] \right\} dt \tag{15.92}$$

However, as stated, this requires knowledge of $I_{\text{port}}(V_{\text{port}})$, a possibly complicated nonlinear function that may not be immediately available.

A potentially more fruitful approach [41] is to simply apply SPICE to model the *complete* circuit of Fig. 15.16, *including* the contents of the embedded circuit. This strategy permits taking advantage of SPICE's built-in algorithms for the integration of the large systems of ordinary differential equations that arise in circuit analysis, as well as SPICE's versatile device library. In this case, there would be no need to have knowledge of $I_{\text{port}}(V_{\text{port}})$, since this function would arise naturally from the interconnection of the linear and nonlinear elements placed into the SPICE model.

For convenience, we now summarize the entire procedure, assuming a SPICE interface. Starting at time $t_0 + \Delta t/2$ with an assumed known H-field distribution everywhere in the lattice:

1. I_N is computed at time $t_0 + \Delta t/2$ using the H-fields circulating around the cell containing the circuit port. Depending upon the orientation of the port, either (15.87a), (15.87b), or (15.87c) is used for this step.

2. SPICE is used to model the circuit of Fig. 15.16, assuming as an excitation $I_N(t_0 + \Delta t/2)$ obtained in step 1. This updates $V_{port}(t_0)$ to $V_{port}(t_0 + \Delta t)$, thereby providing $E_w(t_0 + \Delta t)$ from (15.91).

3. Using the normal Yee algorithm, the E-fields everywhere else in the space lattice are updated to time $t_0 + \Delta t$.

4. Using the normal Yee algorithm, the H-fields everywhere in the space lattice are updated to time $t_0 + 3\Delta t/2$, and the computation cycle of steps 1 to 3 is repeated.

The above technique is readily extended to model a multiport lumped-element network embedded in the FDTD space lattice. Here, an equivalent circuit as shown in Fig. 15.16 is set up at each of the network ports, wherein the Norton excitation $I_N(t)$ at each port is provided by the local FDTD-calculated H-field components in the manner of (15.86). SPICE is then applied to the entirety of the multiport network and its associated Norton equivalent driving circuits. This yields updates for each port's associated E_w component, which are fed back to the FDTD solver to permit subsequent updating of the H-fields in the usual Yee manner.

15.10.3 Thevenin Equivalent Circuit "Looking Into" the FDTD Space Lattice

Fig. 15.17 illustrates how we can define a Thevenin equivalent circuit "looking into" the FDTD space lattice from the two-terminal port of an embedded circuit. We again assume that the port is located in free space at the site of E_w, a particular E_x, E_y, or E_z component in a cubic-cell lattice.

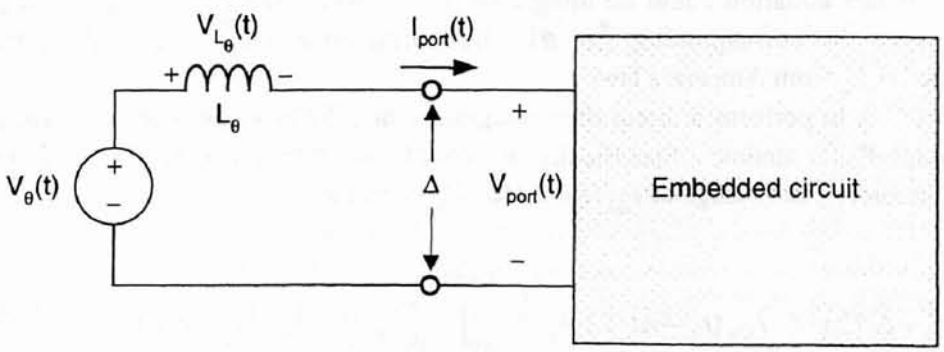


Fig. 15.17 Schematic diagram of a two-terminal port of a circuit embedded in a single cell of the FDTD space lattice, and the Thevenin equivalent circuit "looking into" the FDTD lattice from this port.

From the Thevenin equivalent circuit standpoint, the excitation provided by the complete FDTD lattice to this port can be distilled down to the action of the four immediately adjacent single-cell E-field loops that have E_w along their common edge. The total Thevenin voltage $V_\theta(t)$ is the sum of the loop-voltage contributions:

$$V_\theta(t) = \frac{1}{4}\sum_{i=1}^{4} V_{\text{loop},i}(t) \tag{15.93}$$

where

$$V_{\text{loop},i}(t) = -\oint_{\Gamma_i} \boldsymbol{E}\cdot d\boldsymbol{\Gamma}_i \qquad i = 1, 2, 3, 4 \tag{15.94}$$

Here, Γ_i is the i'th cell-face E-field contour having E_w along one of its edges. There are four such contours. Further, with $L_i = \mu_0\Delta$ being the inductance associated with the i'th contour, we identify from (5.23) an intrinsic lattice space-cell inductance

$$L_{\text{lattice}} = \mu_0\,\Delta/4 \equiv L_\theta \tag{15.95}$$

that provides the Thevenin equivalent source impedance in our model. In fact, we designate this inductance as L_θ.

Having defined the Thevenin equivalent voltage source and source inductance of the FDTD space lattice "seen" by the embedded circuit port, the remainder of the formulation is most easily understood as a duality transformation of (15.90):

$$V_\theta(t) = L_\theta\frac{dI_{\text{port}}(t)}{dt} + V_{\text{port}}\left[I_{\text{port}}(t)\right] \tag{15.96}$$

where $V_{\text{port}}(I_{\text{port}})$ is the known (possibly nonlinear) voltage-current characteristic of the circuit port. We see that (15.96) is merely a statement of Kirchhoff's voltage law applied to the circuit of Fig. 15.17. If this equation could be integrated in time to evolve $I_{\text{port}}(t)$, then it would be possible to evolve the corresponding $\int \boldsymbol{H}\cdot d\boldsymbol{\Gamma}$ circulation around the circuit port, thereby permitting updating E_w from Ampere's law.

One approach is to perform a direct time integration of (15.96) using a standard technique such as the Runge-Kutta method. Specifically, we could start at time $t_0 - \Delta t/2$ and integrate to time $t_0 + \Delta t/2$, assuming knowledge of $I_{\text{port}}(t_0 - \Delta t/2)$, $V_\theta(t)$, and $V_{\text{port}}(I_{\text{port}})$:

$$I_{\text{port}}(t_0 + \Delta t/2) = I_{\text{port}}(t_0 - \Delta t/2) + \frac{1}{L_\theta}\int_{t_0-\Delta t/2}^{t_0+\Delta t/2}\left\{V_\theta(t) - V_{\text{port}}\left[I_{\text{port}}(t)\right]\right\}dt \tag{15.97}$$

However, just as for the case of the Norton equivalent circuit, a potentially more fruitful approach is to apply SPICE to model the complete circuit of Fig. 15.17, *including* the contents of the embedded circuit.

For convenience, we now summarize the entire procedure, assuming a SPICE interface. Starting at time t_0 with an assumed known E-field distribution everywhere in the space lattice:

1. V_θ is computed at time t_0 using (15.93) applied to the E-fields circulating around the four cell-face loops having the circuit port as a common edge.

2. SPICE is used to model the circuit of Fig. 15.17 assuming as an excitation $V_\theta(t_0)$ obtained in step 1. This updates $I_{\text{port}}(t_0 - \Delta t/2)$ to $I_{\text{port}}(t_0 + \Delta t/2)$; and hence from Ampere's law, the $\int \boldsymbol{H} \cdot d\boldsymbol{\Gamma}$ circulation at time $t_0 + \Delta t/2$ around the circuit port. The H-fields everywhere else in the lattice are updated to time $t_0 + \Delta t/2$ using the normal Yee algorithm.

3. E_w can now be updated from t_0 to $t_0 + \Delta t$ using the $\int \boldsymbol{H} \cdot d\boldsymbol{\Gamma}$ circulation around the circuit port at time $t_0 + \Delta t/2$ obtained in step 2.

4. Using the normal Yee algorithm, the E-fields everywhere else in the lattice are updated to time $t_0 + \Delta t$, and the computation cycle of steps 1 to 3 is repeated.

The above technique is readily extended to model a multiport lumped-element network embedded in the FDTD space lattice. Here, an equivalent circuit as shown in Fig. 15.17 is set up at each of the network ports, wherein the Thevenin excitation $V_\theta(t)$ at each port is provided by the local FDTD-calculated E-field components in the manner of (15.93). SPICE is then applied to the entirety of the multiport network and its associated Thevenin equivalent driving circuits. This yields updates for each port's associated E_w component, which are fed back to the FDTD solver to permit subsequent updating of the H-fields in the usual Yee manner.

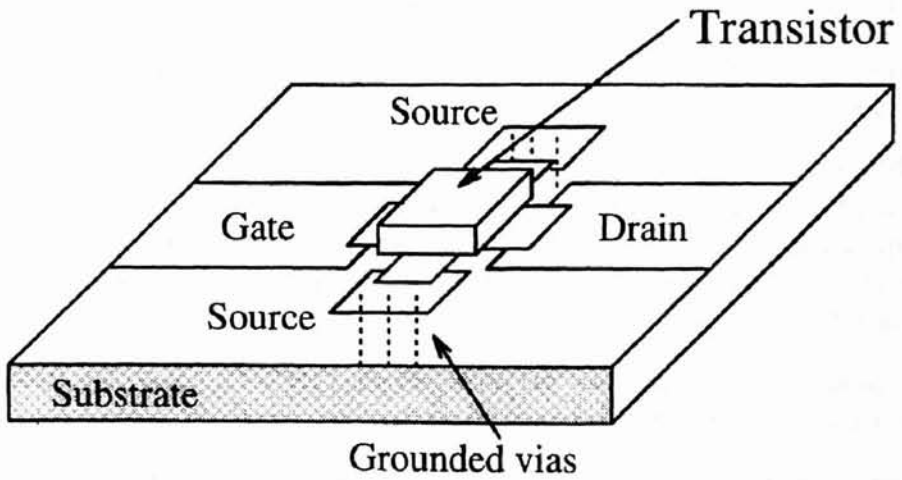
15.11 CASE STUDY: A 6-GHz MESFET AMPLIFIER MODEL

This section discusses the application of the hybrid FDTD / lumped-circuit analysis of Section 15.10 to model a *metal-semiconductor field-effect transistor* (MESFET) used in a two-port, common-source, 6-GHz amplifier. Both linear and nonlinear operation of the MESFET are considered. Fig. 15.18(a) shows the mounting of the MESFET in a microstrip circuit. Here, the MESFET is connected to the ground plane through vias at its source terminal.

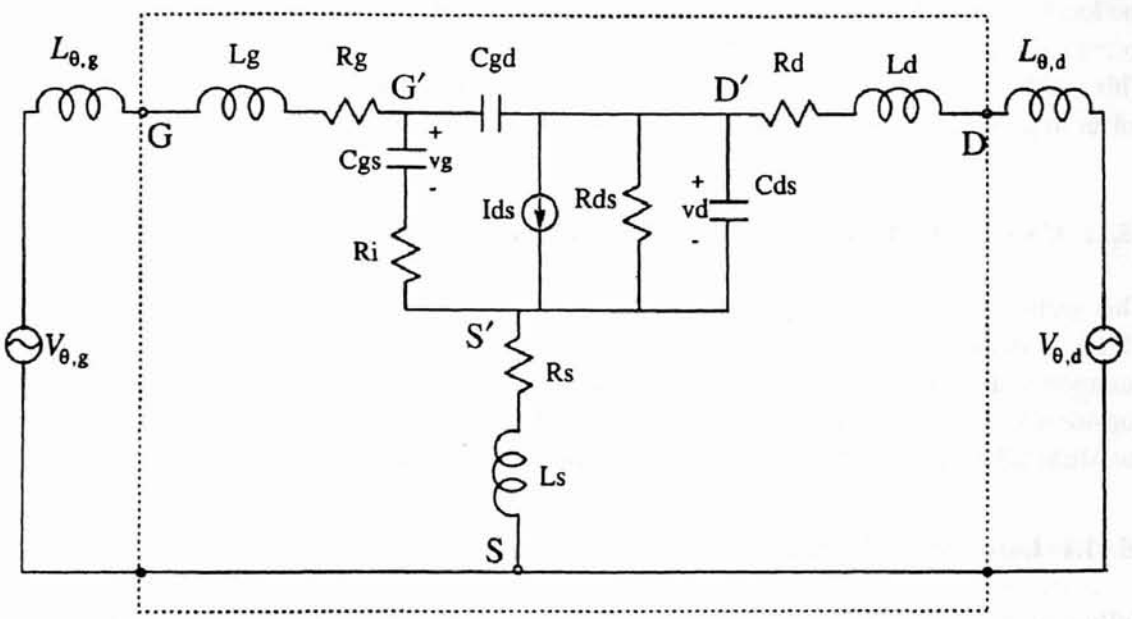
15.11.1 Large-Signal Nonlinear Model

Following [42], Fig. 15.18(b) illustrates the large-signal model of the MESFET used in the hybrid FDTD / lumped-circuit analysis. In this figure, the MESFET circuit model is enclosed by a dashed box, while the Thevenin equivalent circuits outside of this box represent the action of the FDTD space lattice located at the gate and drain terminals of the transistor. Internal nodes G', D', and S' represent the intrinsic part of the MESFET. This model contains two nonlinear elements: the gate-source capacitor C_{gs} and the drain current source I_{ds}. Governed by the PN-junction capacitance model, the gate-source capacitor is expressed as

$$C_{gs}(v_g) = C_{gso} \Big/ \sqrt{1 - v_g/\phi_{bi}} \tag{15.98}$$



(a) Mounting of the packaged MESFET in a microstrip circuit. The transistor is connected to the ground plane through vias at its source terminal.



Rg = 0.5Ω	Rs = 0.7Ω	Rd = 0.5Ω	Ri = 1.0Ω	Cdg = 0.2 pF
Cds = 0.6 pF	Lg = 0.05 nH	Ld = 0.05 nH	Ls = 0.1 nH	

(b) Large-signal model of the MESFET, combined with Thevenin equivalent circuits for the FDTD space lattice at the gate and drain terminals.

Fig. 15.18 MESFET transistor used in the FDTD model of a common-source amplifier. *Source:* Kuo et al., *IEEE Trans. Microwave Theory and Techniques*, 1997, pp. 819–826, © 1997 IEEE.

In (15.98), $C_{gso} = 3$ pf and $\phi_{bi} = 0.7$V for the transistor of interest. The transistor drain current source I_{ds}, describing its dc characteristics, relates to v_g and v_d as

$$I_{ds}(v_g, v_d) = \left[A_0 + A_1 v_{G'S'} + A_2 (v_{G'S'})^2 + A_3 (v_{G'S'})^3\right] \tanh(\alpha v_d) \tag{15.99}$$

where $A_0 = 0.5304$, $A_1 = 0.2595$, $A_2 = -0.0542$, $A_3 = -0.0305$, and $\alpha = 1.0$. Fig. 15.19 plots the transistor's dc characteristics.

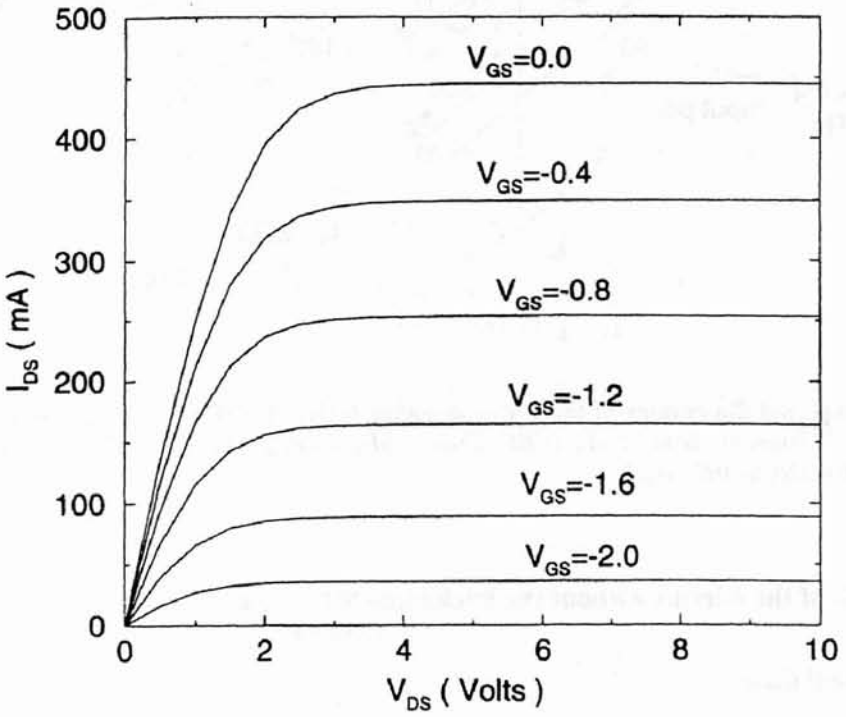


Fig. 15.19 DC characteristics of the MESFET transistor used in the 6-GHz amplifier case study. *Source:* Kuo et al., *IEEE Trans. Microwave Theory and Techniques*, 1997, pp. 819–826, © 1997 IEEE.

15.11.2 Amplifier Configuration

Fig. 15.20 shows the structure and dimensions of the 6-GHz MESFET amplifier used in our case study. This amplifier contains the MESFET having the electrical characteristics described in Section 15.11.1. We note that the physical size of the MESFET (approximately 2.0 mm) is much smaller than the guided wavelength at 6 GHz. A PEC box is used to model the packaging structure. There are two rectangular holes at the input / output ports for feeding dc power from the sources V_{gg} and V_{dd}, each with a source impedance of $50\,\Omega$. The biasing conditions chosen are $V_{GS} = -0.81$V and $V_{DS} = 6.4$V.

The FDTD computation domain is a uniform space lattice of dimensions $74 \times 40 \times 128$ cells with $\Delta x = \Delta z = 0.254$ mm (10 mil) and $\Delta y = 0.197$ mm (7.75 mil). Higdon's second-order ABC is applied at the lattice truncation to absorb outgoing waves.

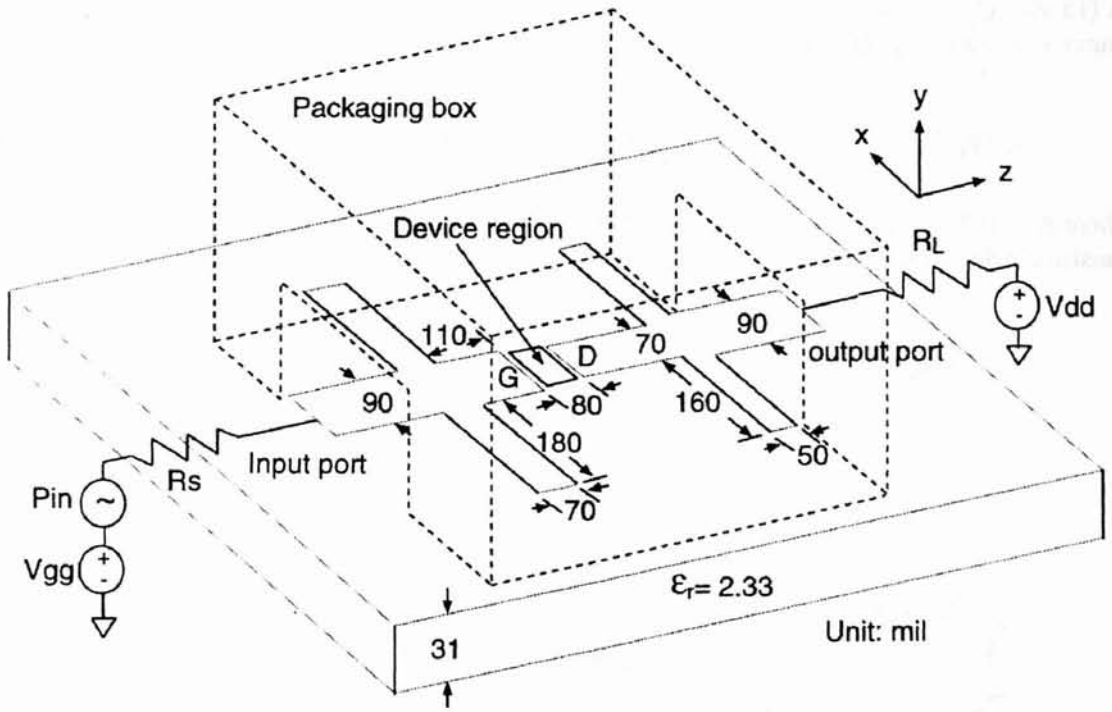


Fig. 15.20 Structure and dimensions of the common-source 6-GHz MESFET amplifier used in the case study. *Source:* Kuo et al., *IEEE Trans. Microwave Theory and Techniques*, 1997, pp. 819–826, © 1997 IEEE.

15.11.3 Analysis of the Circuit without the Packaging Structure

Linear Small-Signal Case

The FDTD model is first applied to obtain the small-signal response of the amplifier of Fig. 15.20 without the packaging box [42]. Here, a Gaussian pulse modulating a 6-GHz carrier is used as the input to the FDTD MESFET circuit model shown in Fig. 15.18(b). The signal amplitude is small enough to allow the circuit to operate in the linear region.

Fig. 15.21 compares the FDTD-computed S parameters with results obtained using the Hewlett-Packard MDS circuit simulator. In this figure, the dip of S_{11} (the impedance-matched condition) in the FDTD simulation is at 5.58 GHz, 7% below that of the HP MDS simulation. This difference may result from a different modeling of the vias at the source port in the two simulations. The frequency of the S_{11} dip is sensitive to the effective inductance of the vias, which causes a series-feedback effect. In the HP MDS simulation, the frequency of the S_{11} dip decreases as the inductance of the vias decreases, approaching the FDTD value.

Nonlinear Large-Signal Case, Harmonic Generation

Fig. 15.22 compares the FDTD- and MDS-calculated output power at the 6-GHz fundamental frequency, the 12-GHz second-harmonic, and the 18-GHz third-harmonic as a function of the input power for a single-tone excitation [42]. Very good agreement is observed.

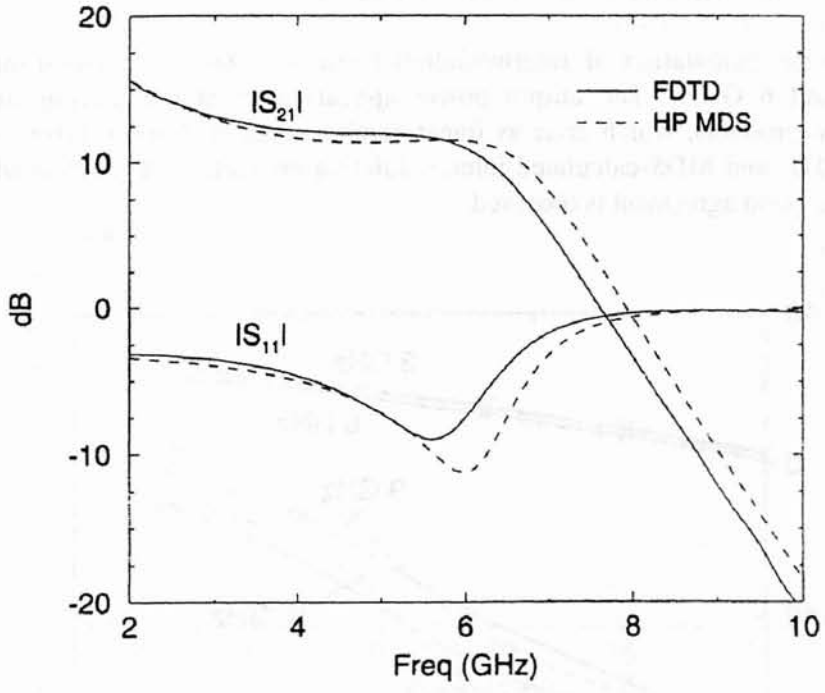


Fig. 15.21 Small-signal response of the common-source 6-GHz MESFET amplifier of Fig. 15.20 without the packaging box. *Source:* Kuo et al., *IEEE Trans. Microwave Theory and Techniques*, 1997, pp. 819–826, © 1997 IEEE.

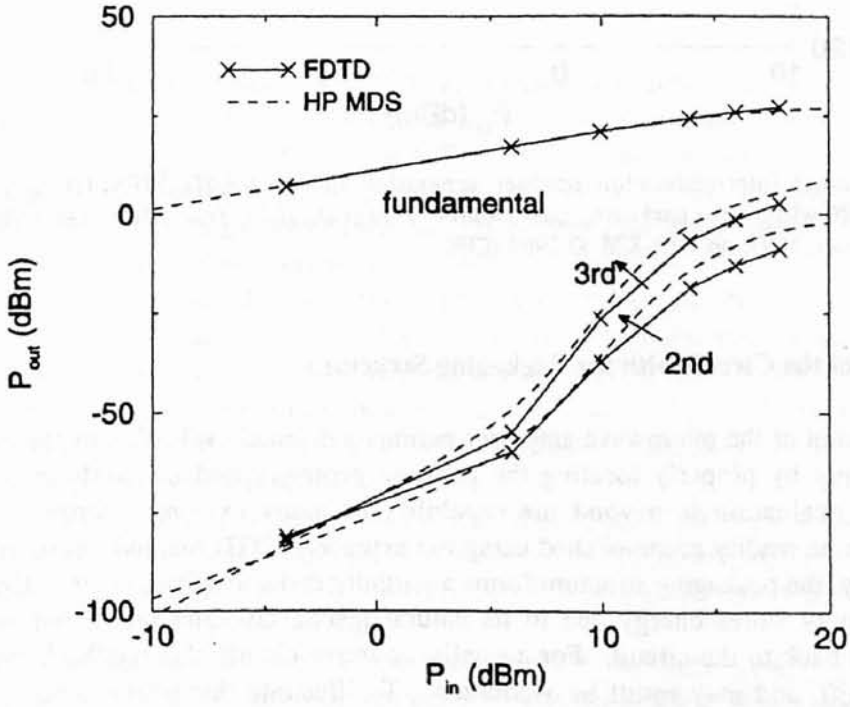


Fig. 15.22 Large-signal harmonic generation of the 6-GHz MESFET amplifier of Fig. 15.20 without the packaging box. *Source:* Kuo et al., *IEEE Trans. Microwave Theory and Techniques*, 1997, pp. 819–826, © 1997 IEEE.

Nonlinear Large-Signal Case, Intermodulation Product Generation

Next, we review the calculation of intermodulation due to a two-tone, equal-input power excitation at 3 and 6 GHz. The output power appears only at the mixing frequencies, or intermodulation products, which arise as linear combinations of 3 and 6 GHz. Fig. 15.23 compares the FDTD- and MDS-calculated intermodulation products as a function of the input power [42]. Again, good agreement is observed.

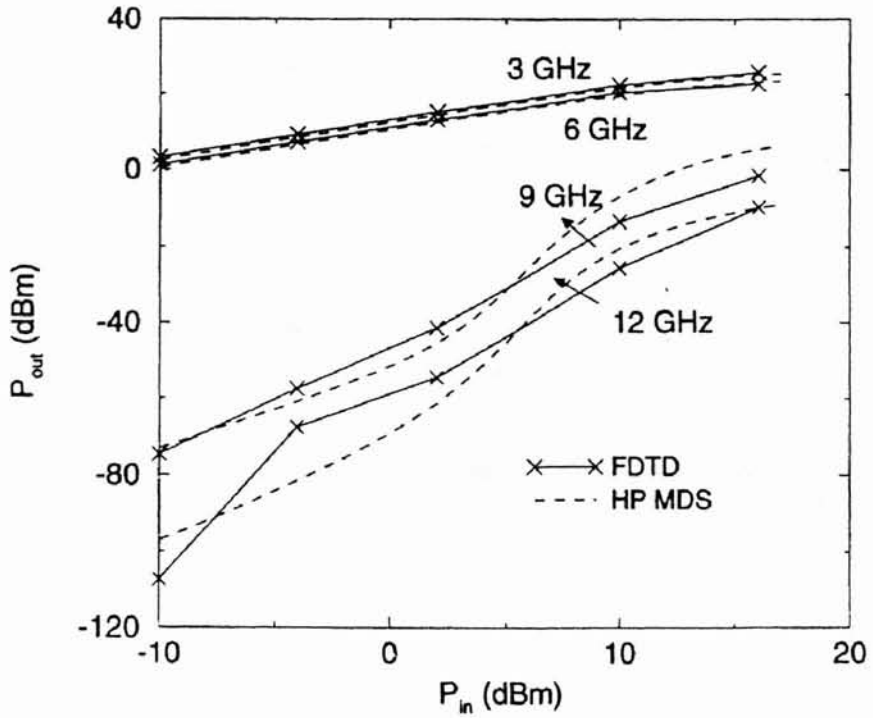


Fig. 15.23 Large-signal intermodulation product generation of the 6-GHz MESFET amplifier of Fig. 15.20 without the packaging box. *Source:* Kuo et al., *IEEE Trans. Microwave Theory and Techniques*, 1997, pp. 819–826, © 1997 IEEE.

15.11.4 Analysis of the Circuit with the Packaging Structure

The FDTD simulation of the microwave amplifier permits a detailed evaluation of the effects of its packaging simply by properly locating the package geometry and materials in the space lattice. Such an evaluation is beyond the capability of many existing microwave circuit simulators, but can be readily accomplished using the extended FDTD method discussed in this chapter. Physically, the packaging structure forms a partially dielectric-filled cavity. Excited by the circuit, the cavity stores energy due to its natural resonances, and the stored energy is inevitably coupled back to the circuit. For a nonlinear active circuit, this feedback makes the stability circles shift, and may result in oscillation. To illustrate this phenomenon, we first consider the example of our MESFET circuit contained within a PEC box of dimensions $39.6 \times 4.7 \times 31.8$ mm ($1{,}560 \times 186 \times 1{,}250$ mil). The first resonant frequency of this box is found to be at 5.72 GHz by an FDTD presimulation of a packaged uniform microstrip line.

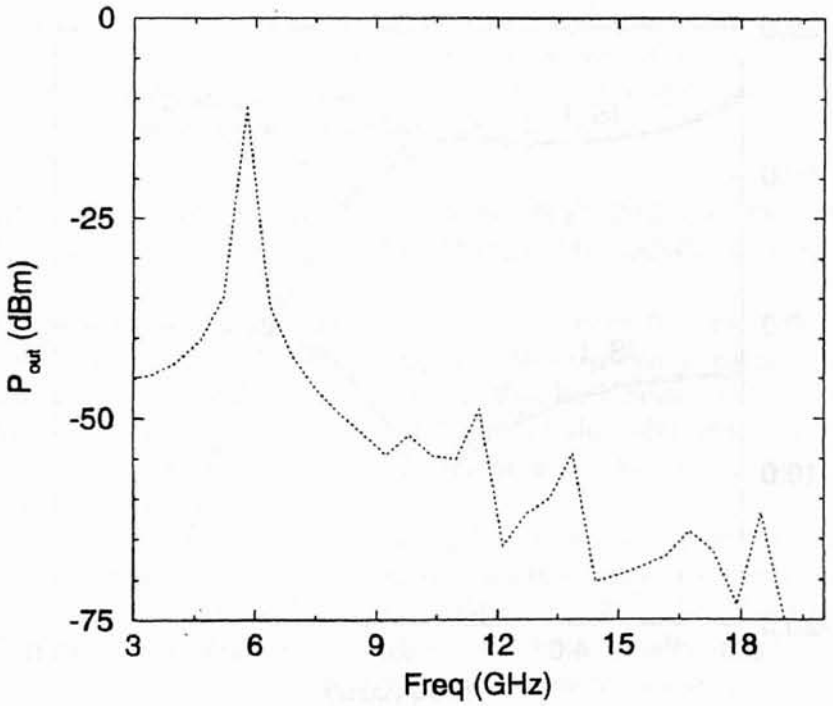


Fig. 15.24 Spectrum of the output power for a small-signal exciting the initial packaged MESFET amplifier, showing the circuit oscillating after being placed in the packaging structure. *Source:* Kuo et al., *IEEE Trans. Microwave Theory and Techniques*, 1997, pp. 819–826, © 1997 IEEE.

Fig. 15.24 shows the results of FDTD calculations indicating that this packaging interacts strongly with the circuit and causes oscillation at the resonant frequency [42]. Here, a very-low-level input Gaussian pulse modulating a 6-GHz carrier (close to the center frequency of the amplifier) generates a much larger output power than warranted by the nominal amplifier gain. In fact, this indicates the presence of essentially unlimited gain at the amplifier center frequency, and thus a condition of potential instability leading to self-oscillation.

In order to avoid instability, the dimensions of the packaging structure are usually chosen such that its resonant frequency is raised well above the frequency range of interest. To explore this possibility, we now consider a second packaging example for our MESFET circuit, where the dimensions of the enclosing PEC box are reduced to $16.3 \times 4.7 \times 17.5$ mm ($640 \times 186 \times 690$ mil) [42]. This elevates the FDTD-calculated first resonant frequency to 11.79 GHz, rendering the packaged circuit electrically stable.

Fig. 15.25 shows the effect of the modified packaging upon the small-signal S-parameters of our MESFET amplifier. In this case, the frequency of the FDTD-calculated match point remains at 5.58 GHz, but the magnitude of S_{11} drops from -8.98 to -10.07 dB, while the amplifier gain S_{21} at 6 GHz rises from 10.95 to 11.76 dB.

Fig. 15.26 illustrates the effect of the packaging on the output power at the fundamental, second-harmonic, and third-harmonic frequencies for a large-signal, single-tone input at 6 GHz. The resonant frequency of the package is close to that of the second harmonic at 12 GHz. Therefore, package energy is coupled strongly to the second harmonic, and its output-power curve varies more significantly than the others.

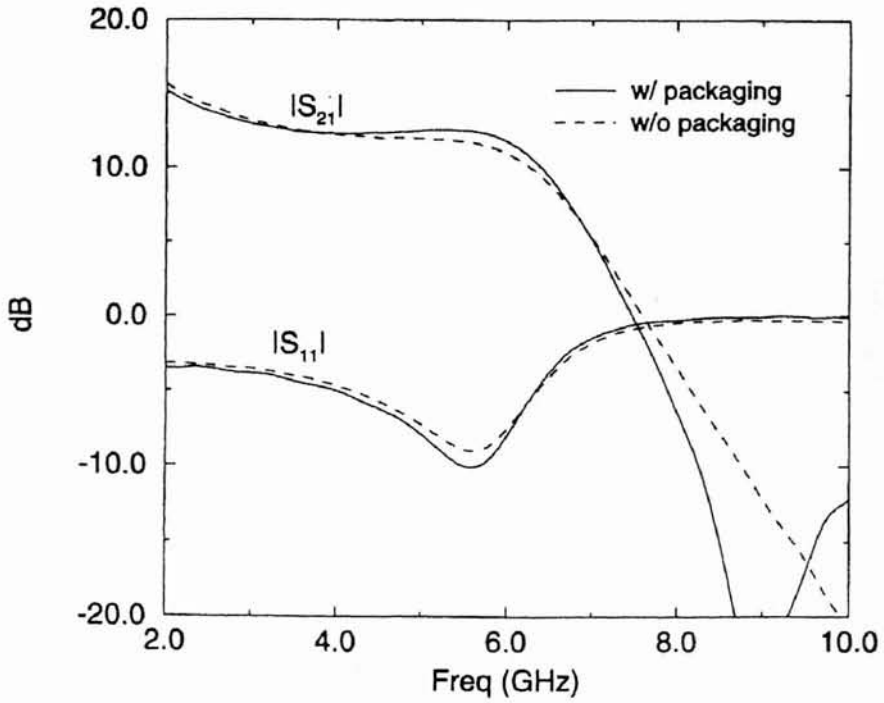


Fig. 15.25 Effect of modified packaging upon the *S*-parameters of the MESFET amplifier. *Source:* Kuo et al., *IEEE Trans. Microwave Theory and Techniques*, 1997, pp. 819–826, © 1997 IEEE.

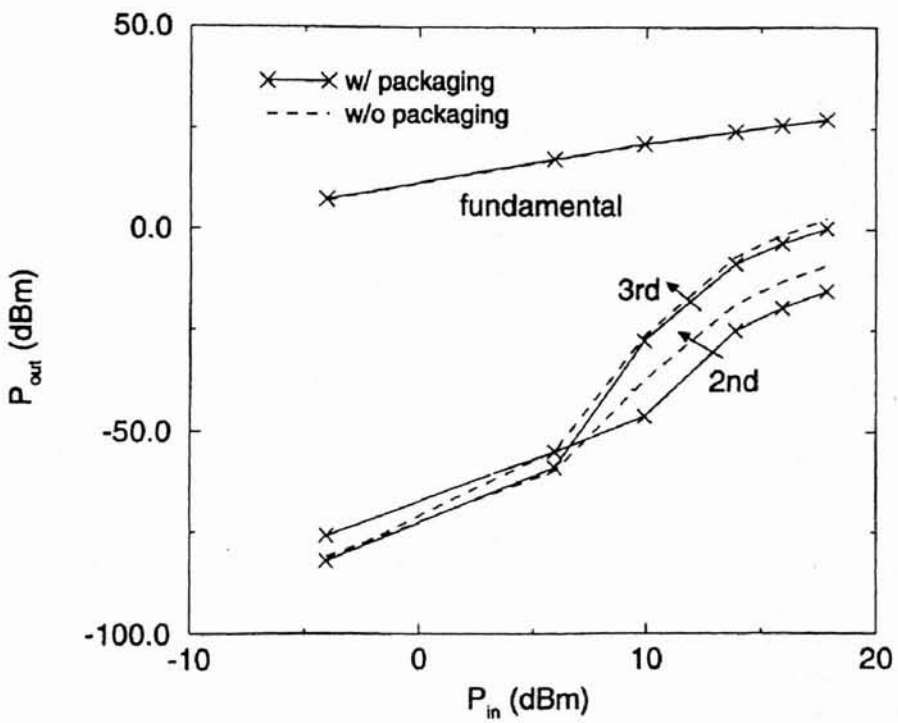


Fig. 15.26 Effect of the MESFET amplifier packaging on the output power at the fundamental and harmonic frequencies for a large-signal, single-tone input at 6 GHz. *Source:* Kuo et al., *IEEE Trans. Microwave Theory and Techniques*, 1997, pp. 819–826, © 1997 IEEE.

Overall, this case study demonstrates that the extended FDTD / electronic-circuit modeling technique discussed in this chapter is capable of providing accurate and useful information regarding the operation of active, nonlinear microwave circuits of realistic complexity. While the example discussed in this section focuses on an analog circuit, it should be clear that digital circuits can be simulated as well by this approach.

15.12 EMERGING TOPIC: WIRELESS HIGH-SPEED DIGITAL INTERCONNECTS USING DEFECT-MODE ELECTROMAGNETIC BANDGAP WAVEGUIDES

As computer clock rates continue to rise in the microwave frequency range above 3 GHz, problems with signal integrity, cross-coupling, and radiation will eventually render useless the baseband metal-strip circuit-board interconnects that have been employed since the 1940s. While replacing metal strips with optical fibers would solve the problem, the required incorporation of optoelectronics would represent a revolution in both chip-making and interconnect technologies.

This section summarizes work reported in [43, 44], which proposes an alternate solution to the digital-interconnect problem: *bandpass wireless interconnects* implemented using "defect-mode" *electromagnetic bandgap* (EBG) waveguides [45–48].[3] (See also Chapter 16, Section 16.14.) The EBG structures studied in [43, 44] are dimensionally scaled to operate at center frequencies of 10 and 50 GHz. These structures are simply square arrays of copper via pins embedded either in free space or in circuit-board dielectric material. One or more rows of pins are removed to create a linear waveguide. Operation at higher center frequencies well above 100 GHz is conceptually feasible because of the recent development of silicon transistors having gain-bandwidths above 1 THz [49]. Relative to metal strips or optical fibers, such millimeter-wave EBG waveguides would have the following advantages when used for board-level digital interconnects:

- Sufficient high-quality bandwidth (i.e., with flat transmission magnitude, linear phase shift, and broadband impedance matching to available loads) to support computer processors clocked up to 30 GHz;
- Construction using evolutionary extensions of existing circuit-board and connector technologies;
- Low copper loss;
- Little signal distortion, coupling, and radiation, even at right-angle bends;
- Nearly speed-of-light signal transmission via the usage of low-permittivity dielectric media.

---

[3] In the 1990s, the photonics community was the first to develop engineering applications of EBG structures. Hence, the terms "photonic bandgap" (PBG) and "photonic crystal" (PC) came into wide use. Photonics applications of EBG structures having point or linear defects include low-loss dielectric or semiconductor waveguides and cavities for optical interconnects, couplers, filters, and lasers. At present, interest in applying EBG concepts is rapidly increasing in the microwave and millimeter-wave community. In fact, the growth in radio-frequency applications has caused the term "PBG" to be generalized to "EBG" in the electrical engineering literature. Here, EBG concepts are being adapted to impose frequency-selective behavior upon familiar baseband metal signal paths such as microstrip lines and coplanar striplines.

15.12.1 Stopband of the Defect-Free Two-Dimensional EBG Structure

Following [43], we first review the application of FDTD modeling to calculate the stopband of defect-free two-dimensional EBG structures employing uniform square arrays of copper pins in free space. Here, referring to Fig. 15.27, numerical experiments involved simulations of the electromagnetic field transmission through these structures for an impulsive, normally incident, E-polarized plane wave (electric field parallel to the copper pins). Many combinations of pin radius r and center-to-center pin spacing a were tested to obtain the maximum possible spectral width of the EBG stopband, in combination with the maximum possible transmission bandwidth of a linear defect-mode waveguide having a desired center frequency of 10 GHz.

In all of the simulations, a standard two-dimensional TM_z FDTD code was implemented on a uniform Cartesian grid using square space cells of size $\Delta = 0.48$ mm. This cell size was found to yield numerically converged results relative to rendering the geometry of the array of copper pins comprising the EBG structure.

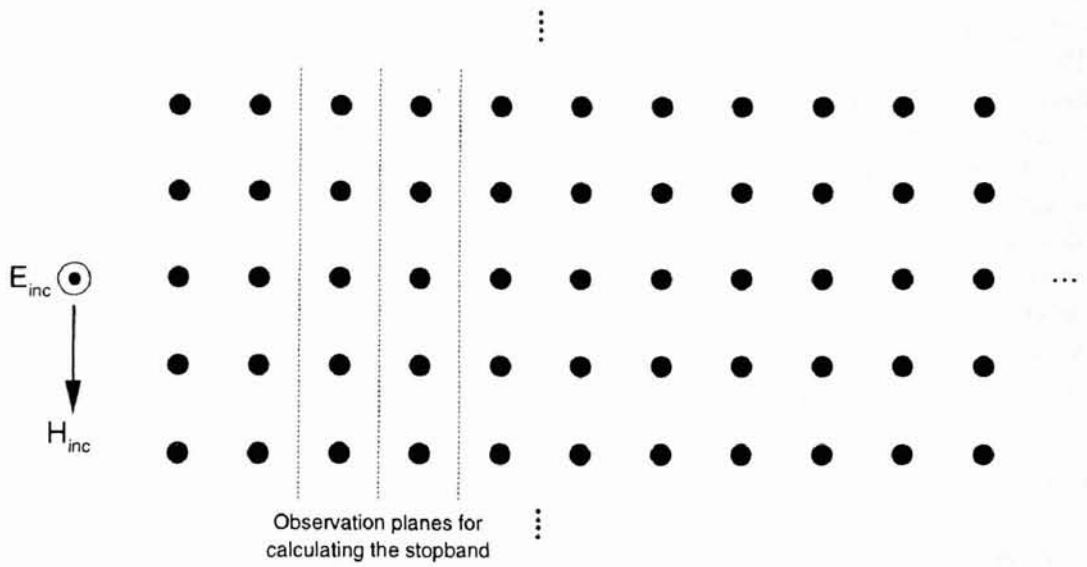


Fig. 15.27 Geometry of the two-dimensional defect-free EBG structure modeled using FDTD.

Fig. 15.28 shows the FDTD-calculated stopband observed two, three, and four rows deep within the EBG structure for the nearly optimum case $r = 1.7$ mm and $a = 9.7$ mm. We see that the stopband extends from dc to approximately 14.1 GHz. In our subsequent discussion, we shall take 14.1 GHz as the upper limit of the permissible passband of the EBG structure with a waveguiding defect, since above this frequency the field confinement is lost.

15.12.2 Passband of the Two-Dimensional EBG Structure with Waveguiding Defect

Reference [43] then reported FDTD modeling results for the transmission characteristics of waveguiding defects created in the EBG geometry of Fig. 15.27. Either one row of pins was deleted from an approximately optimum EBG structure in air having initially five rows in the lateral direction, or two rows of pins were deleted from an EBG structure having initially six rows in the lateral direction.

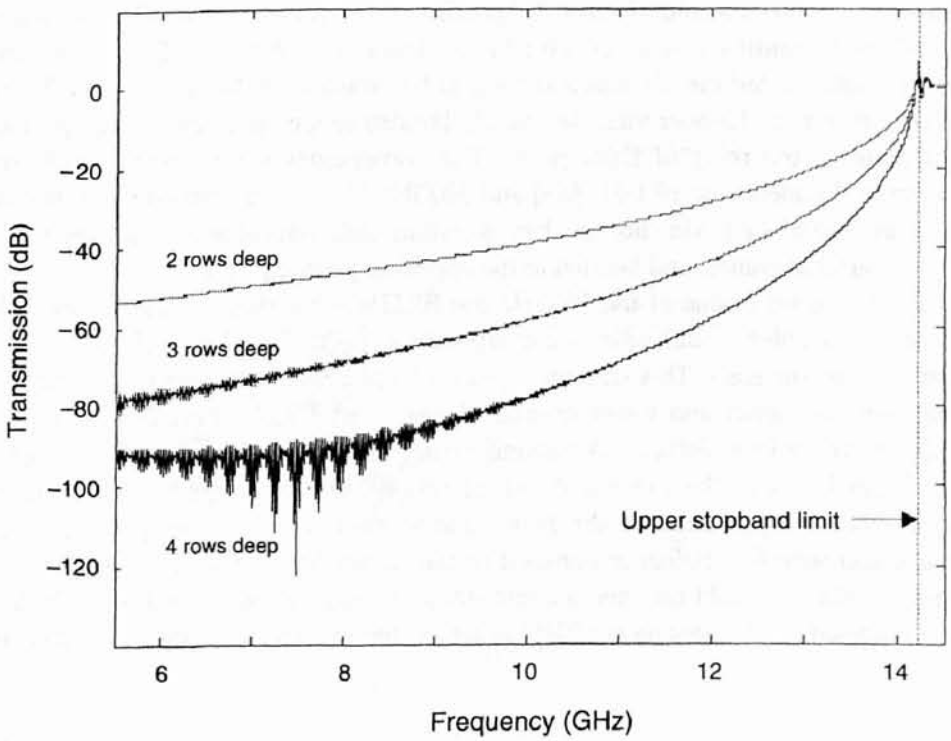


Fig. 15.28 FDTD-calculated magnitude (*E*-polarization case) of the electric-field transmission transfer function through the first two, three, and four rows of metal pins of the EBG structure illustrated in Fig. 15.27 for the nearly optimum parameters pin radius $r = 1.7$ mm and center-to-center spacing $a = 9.7$ mm. *Source:* Simpson et al., *IEEE Microwave and Wireless Components Lett.*, 2004, pp. 343–345, © 2004 IEEE.

For each waveguiding defect, FDTD was applied to calculate the magnitude and phase of the electric field transmission characteristic over a 45-cm longitudinal span along the center axis of the waveguide. Defining the usable transmission bandwidth as the frequency span over which the transmission magnitude variations are less than ±1 dB, the transmission bandwidth for the single-row waveguiding defect was calculated to be 55% (6.1 GHz) centered about 10 GHz, and 85% (8.7 GHz) for the double-row waveguiding defect. Further analysis indicated that the phase of the electric field transmission characteristic corresponding to the single-row waveguiding defect is linear to within ±1.8°/GHz over the full passband, and to within ±1.1°/GHz for the double-row waveguiding defect. The latter is therefore preferable from the standpoint of bandwidth and phase linearity.

To examine the possibility of multimoding in the double-row waveguiding defect [48], [43] reported additional numerical experiments that transversely offset the source distribution within the defect. These experiments indicated little (if any) evidence of odd-mode generation.

15.12.3 Laboratory Experiments and Supporting FDTD Modeling

Based upon the optimizations determined with the two-dimensional FDTD modeling, [43, 44] reported laboratory measurements and supporting three-dimensional FDTD modeling of prototype EBG waveguiding structures with linear double-row defects. These structures were realized using double-sided circuit board having either standard FR4 or low-loss Rogers 5880 as the dielectric material. Copper vias electrically bonded to the upper and lower ground planes served to implement the rows of EBG pins. The waveguides were dimensionally scaled to operate at a center frequency of 10 GHz [43] and 50 GHz [44]. These structures were coupled to input and output coaxial lines via short probes extending transversely across the gap between the upper and lower ground planes, and bonded to the opposing ground plane.

Fig. 15.29 is a photo of one of the 50 GHz test EBG waveguiding structures [44]. (See Fig. 1.6 of Chapter 1 for color visualizations corresponding to the FDTD modeling and laboratory measurements of this device.) This structure spans 8.6 cm between the input and output probes, 0.76 mm between the upper and lower ground planes, and 3.9 mm between the rows of vias bordering the waveguiding defect. A second structure identical to that of Fig. 15.29, but spanning 12.7 cm between the input and output probes, was also constructed to isolate and measure the dielectric loss. Because the radius and spacing of the copper pins scale inversely with frequency and with the dielectric constant of the circuit-board material, the pins of the test structure have a radius $r = 0.23$ mm and a center-to-center separation $a = 1.3$ mm. Note that the waveguide is bounded on all sides by the EBG structure, thereby representing a closed cavity.

Fig. 15.29 Photograph of the FDTD-designed laboratory test structure for the wireless digital interconnect using EBG technology, achieving a measured 42-GHz bandwidth (28 to 70 GHz). Coaxial connectors are probe-coupled to a linear, double-row-defect, EBG waveguide. *Source:* Simpson et al., *Proc. IEEE Antennas and Propagation Society Intl. Symp.*, © 2005 IEEE.

Fig. 15.30 shows the measured results for the insertion loss of the waveguiding structure of Fig. 15.29. A standard Agilent microwave network analyzer having 50Ω nominal source and load impedances was used in these measurements. We see that there exists a sharp transition from a stopband of at least −65 dB to the passband above approximately 28 GHz. This passband extends to approximately 70 GHz with a gain flatness of ±1.5 dB. This result supports the basic concept that an approximate 85% fractional bandwidth is feasible using double-row defect EBG structures, and also that such structures can be scaled in a straightforward manner to higher center frequencies.

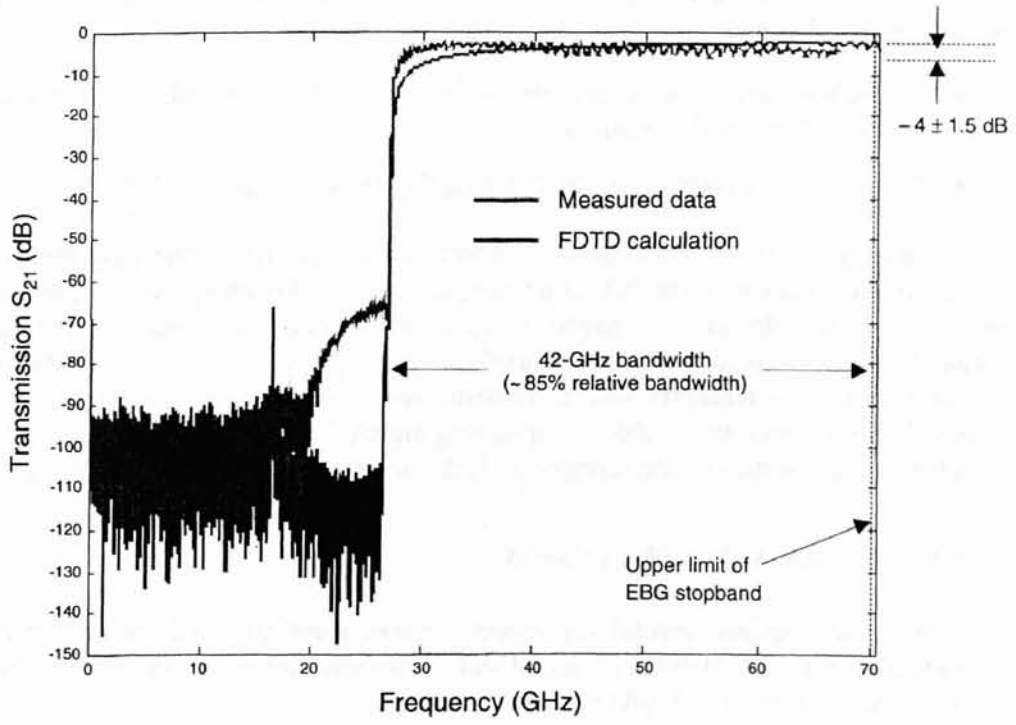


Fig. 15.30 Comparison of FDTD-calculated and measured results for the transmission of the prototype double-row EBG waveguide shown in Fig. 15.29. *Source:* Simpson et al., *Proc. IEEE Antennas and Propagation Society Intl. Symp.*, © 2005 IEEE.

Fig. 15.30 shows that the total insertion loss at midband (50 GHz) of the 8.6-cm-long waveguiding defect is 4 dB. Comparison of this result with that of an identical 12.7-cm-long waveguiding defect indicates that approximately 2 dB of the 4-dB total loss is caused by the propagation attenuation due to the Rogers 5880 dielectric. Subtracting this dielectric loss from the total insertion loss yields a coupling loss of only approximately 1 dB at each coaxial transition. This indicates the possibility of excellent broadband matching into the waveguiding defect using simple thin metal probes.

Overall, the results summarized in this section demonstrate the utility of FDTD modeling in designing a novel wireless digital interconnect technology which employs linear defects in electromagnetic bandgap structures as ultrawideband waveguides. This technology has the following desirable features:

- Ultrawideband (greater than 80%) relative bandwidth;
- Compatibility with existing circuit-board technology;
- Excellent stopband, insertion-loss, and impedance-matching characteristics;
- Negligible crosstalk and radiation.

If a low-permittivity, low-loss dielectric medium such as an aerogel can be used for the insulating layers within a circuit board comprising the EBG structure, the following additional advantages would accrue:

- High-characteristic-impedance operation, thereby reducing copper losses relative to conventional 50Ω striplines;
- Signal velocities potentially approaching the free-space speed of light.

Laboratory measurements conducted at both 10- and 50-GHz center frequencies have shown very good agreement with the FDTD design predictions. Assuming the availability of suitable low-loss dielectrics to serve as insulating layers within the circuit boards, this technology will ultimately be scalable to millimeter-wave center frequencies well above 100 GHz, thereby leveraging emerging terahertz silicon transistor technology. Then, the wireless interconnects discussed herein would be capable of supporting digital data rates in the hundreds of gigabits per second, adequate for the elevated computer clock rates expected over the next decade.

15.13 SUMMARY AND CONCLUSIONS

This chapter summarized several key aspects of the formulation and application of FDTD for modeling contemporary high-speed digital and microwave circuits, including those having active and nonlinear components. Topics reviewed included:

- Processing FDTD-computed E- and H-fields in a TEM transmission-line system to obtain the line's characteristic impedance, propagation constant, load impedance, S-parameters, differential capacitance, and differential inductance;
- Processing FDTD-computed E- and H-fields in a TEM transmission-line system to obtain the equivalent series lumped inductance due to a discontinuity in the signal or ground-return paths;
- Processing FDTD-computed E- and H-fields to obtain the equivalent series lumped inductance of a complex power-distribution system in a high-speed digital-circuit module;
- FDTD modeling of parallel coplanar microstrips;
- FDTD modeling of early-time and late-time coupling modes in a realistic multilayered interconnect module in a high-speed digital circuit;
- S-parameter extraction from FDTD models of general waveguides that support a superposition of propagating and evanescent modes;
- Digital signal processing and spectrum-estimation techniques for extrapolating time-windowed FDTD impulse responses;
- Modeling of linear and nonlinear lumped circuit elements embedded within a single unit cell of the FDTD space lattice, including: resistors, resistive voltage sources, capacitors, inductors, general two-terminal linear lumped networks (containing arbitrary combinations of resistors, capacitors, and inductors), diodes, and bipolar junction transistors;

- Direct linking of FDTD and SPICE via Norton and Thevenin equivalent circuits "looking into" the FDTD space lattice, thereby enabling the coupling of FDTD electromagnetic wave simulations with SPICE circuit models of arbitrary linear and nonlinear components and subsystems;
- A case study involving the hybrid FDTD / circuit modeling of a 6-GHz microstrip MESFET amplifier for small-signal *S*-parameters, large-signal harmonic and intermodulation generation, and the potentially destabilizing effects of the amplifier packaging;
- An emerging wireless high-speed digital-interconnect technology involving defect-mode electromagnetic bandgap waveguides, wherein FDTD modeling has been crucial from the initial concept to the design of laboratory prototypes.

We conclude that recent progress in the FDTD method has made it possible to analyze many high-speed digital and microwave circuits of engineering interest in their complete form using full-wave electromagnetic field computations, and accounting for relevant circuit nonlinearities. A self-consistent interface with SPICE and related circuit solvers is feasible. In the future, this interfacing of FDTD and SPICE may help to unify the electromagnetics and circuit-analysis disciplines within electrical engineering. Overall, FDTD can provide ultrawideband modeling of linear and nonlinear devices and circuits with a single simulation run. FDTD is straightforward, robust, and physically appealing. Given these attributes, FDTD modeling of microwave and digital circuits will likely become widely used in industry as the capabilities of workstations and personal computers improve from year to year.

ACKNOWLEDGMENTS

Professor Piket-May is pleased to recognize the contributions of Andy Byers of the University of Colorado and Roger Gravrok of Sequent Corporation for their work on characterization of the inductance of distributed power systems, and Cray Research for its computing support. Professor Gwarek is pleased to acknowledge the collaboration and coauthorship of Dr. Malgorzata Celuch-Marcysiak of the Warsaw University of Technology. Dr. Houshmand and Professor Itoh are pleased to acknowledge the original contributions made by Brent Toland, Tian Huang, Chien-nan Kou, Min Chen, and K. P. Ma in the extended FDTD algorithm. Ms. Simpson is pleased to acknowledge the collaboration, support, and coauthorship of Jason Mix and Howard Heck of Intel Corporation.

REFERENCES

[1] Maas, S. A., *Microwave Mixers*, 2nd ed., Norwood, MA: Artech House, 1993.

[2] Pozar, D. M., *Microwave Engineering*, Reading, MA: Addison-Wesley, 1990.

[3] Cendes, Z. C., "EM simulators = CAE tools," *IEEE Spectrum*, Vol. 27, 1990, pp. 73–79.

[4] Horng, T. S., "Extending the three-dimensional spectral-domain approach to hybrid microwave integrated circuits with passive and active lumped elements," *1994 IEEE MTT-S Intl. Microwave Symp. Dig.*, San Diego, CA, 1994, pp. 709–712.

[5] Guillouard, K., M. F. Wong , V. F. Hanna, and J. Citerne, "A new global finite-element analysis of microwave circuits including lumped elements," *1996 IEEE MTT-S Intl. Microwave Symp. Dig.*, San Francisco, CA, 1996, pp. 355–358.

[6] Itoh, T., and B. Houshmand, (eds.), *Time-Domain Methods for Microwave Circuits*, New York: IEEE Press, 1998.

[7] Zhang, X., and K. K. Mei, "Time-domain finite-difference approach for the calculation of the frequency-dependent characteristics of microstrip discontinuities," *IEEE Trans. Microwave Theory Tech.*, Vol. 36, 1988, pp. 1775–1787.

[8] Gwarek, W. K., and M. Celuch-Marcysiak, "A differential method of reflection coefficient extraction from FDTD simulations," *IEEE Microwave Guided Wave Lett.*, Vol. 6, 1996, pp. 215–217.

[9] Sheen, D. M., S. M. Ali, M. D. Abouzahra, and J. A. Kong, "Application of the three-dimensional finite-difference time-domain method to the analysis of planar microstrip circuits," *IEEE Trans. Microwave Theory Tech.*, Vol. 38, 1990, pp. 849–857.

[10] Yook, J.-G., N. I. Dib, and L. P. B. Katehi, "Characterization of high-frequency interconnects using finite difference time domain and finite element methods," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 1994, pp. 1727–1736.

[11] Becker, W. D., P. H. Harms, and R. Mittra, "Time-domain electromagnetic analysis of interconnects in a computer chip package," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, 1992, pp. 2155–2163.

[12] Harms, P., J. Lee, and R. Mittra, "Characterizing the cylindrical via discontinuity," *IEEE Trans. Microwave Theory Tech.*, Vol. 41, 1993, pp. 153–156.

[13] Mittra, R., S. Chebolu, and W. D. Becker, "Efficient modeling of power planes in computer packages using the finite difference time domain method," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 1994, pp. 1791–1795.

[14] Gravrok, R., A. Byers, and M. Piket-May, "Numerical modeling of inductance for a distributed system," *Proc. IEEE 6th Conf. Electronics Packaging*, 1997, pp. 83–88.

[15] Gravrok, R., D. Scheid, J. Ficke, K. VanGoor, and V. Rao, "Advanced MCM-Ds break the I/O bottleneck with embedded area-array integrated circuits," *Proc. ISHM MCM Conf.*, Denver, CO, 1996.

[16] Piket-May, M. J., A. Taflove, and J. Baron, "FDTD modeling of digital signal propagation in 3D circuits with passive and active loads," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 1994, pp. 1514–1523.

[17] Ramo, S., J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communications Electronics*, 2nd ed., New York: Wiley, 1984.

[18] Gwarek, W. K., and M. Celuch-Marcysiak, "Wide-band *S*-parameter extraction from FDTD simulations for propagating and evanescent modes in inhomogeneous guides," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, 2003, pp. 1920–1928.

[19] Gwarek, W. K., T. Morawski, and C. Mroczkowski, "Application of the FDTD method to the analysis of circuits described by the two-dimensional vector wave equation," *IEEE Trans. Microwave Theory Tech.*, Vol. 41, 1993, pp. 311–317.

[20] Hildebrand, F. B., *Introduction to Numerical Analysis*, New York: Dover, 1974, pp. 457–462.

[21] Dubard, J. L., D. Pompei, J. Le Roux, and A. Papiernik, "Characterization of microstrip antennas using the TLM simulation associated with a Prony-Pisrenko method," *Intl. J. Numerical Modeling*, Vol. 3, 1990, pp. 269–285.

[22] Ko, W. L., and R. Mittra, "A combination of FDTD and Prony's methods for analyzing microwave integrated circuits," *IEEE Trans. Microwave Theory Tech.*, Vol. 39, 1991, pp. 2176–2181.

[23] Pereda, J. A., L. A. Vielva, and A. Prieto, "Computation of resonant frequencies and quality factors of open dielectric resonators by a combination of the finite-difference time-domain (FDTD) and Prony's methods," *IEEE Microwave Guided Wave Lett.*, Vol. 2, 1992, pp. 431–433.

[24] Maloney, J. G., and G. S. Smith, "The use of surface impedance concepts in the finite-difference time-domain method," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 38–48.

[25] Kumaresan, R., and D. W. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise," *IEEE Trans. Acoustics, Speech, Signal Proc.*, Vol. 30, 1982, pp. 833–840.

[26] Kay, S. M., and S. L. Marple, "Spectrum analysis—A modern perspective," *Proc. IEEE*, Vol. 69, 1981, pp. 1380–1419.

[27] Chen, J., C. Wu, T. K. Y. Lo, K.-L. Wu, and J. Litva, "Using linear and nonlinear predictors to improve the computational efficiency of the FDTD algorithm," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 1994, pp. 1992–1997.

[28] Kay, S. M., *Modern Spectral Estimation*, Englewood Cliffs, NJ: Prentice-Hall, 1988.

[29] Ulrych, T. J., and R. W. Clayton, "Time series modeling and maximum entropy," *Phys. Earth and Plan. Int.*, Vol. 12, 1976, pp. 188–200.

[30] Nuttal, A. H., "Spectral analysis of a univariate process with bad data points via maximum entropy and linear predictive techniques," Naval Underwater Systems Center, New London, CT, Tech. Rept. 5303, 1976.

[31] Marple, L., "A new autoregressive spectrum analysis algorithm," *IEEE Trans. Acoustics, Speech, Signal Proc.*, Vol. 28, 1980, pp. 441–454.

[32] Jandhyala, V., E. Michielssen, and R. Mittra, "FDTD signal extrapolation using the forward-backward autoregressive (AR) model," *IEEE Microwave Guided Wave Lett.*, Vol. 4, 1994, pp. 163–165.

[33] Haggan, V., and T. Ozaki, "Modeling nonlinear random vibrations using an amplitude-dependent autoregressive time series model," *Biometrika*, Vol. 68, 1981, pp. 189–196.

[34] Tong, H., *Non-Linear Time Series*, Cambridge, U.K.: Oxford University Press, 1990.

[35] Dey, S., and R. Mittra, "Efficient computation of resonant frequencies and quality factors of cavities via a combination of the finite-difference time-domain technique and the Padé approximation," *IEEE Microwave Guided Wave Lett.*, Vol. 8, 1998, pp. 415–417.

[36] Monorchio, A., and R. Mittra, "A hybrid finite-element finite-difference time-domain (FE/FDTD) technique for solving complex electromagnetic problems," *IEEE Microwave Guided Wave Lett.*, Vol. 8, 1998, pp. 93–95.

[37] Pozzi, A., *Application of Padé Approximation Theory in Fluid Dynamics*, World Science, 1994.

[38] 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 Tech.*, Vol. 40, 1992, pp. 724–730.

[39] Wu, T.-L., S.-T. Chen, and Y.-S. Huang, "A novel approach for the incorporation of arbitrary linear lumped network into FDTD method," *IEEE Microwave Wireless Components Lett.*, Vol. 14, 2004, pp. 74–76.

[40] Sze, S. M., *Physics of Semiconductor Devices*, New York: Wiley, 1981.

[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 Guided Wave Lett.*, Vol. 4, 1994, pp. 141–143.

[42] Kuo, C.-N., B. Houshmand, and T. Itoh, "Full-wave analysis of packaged microwave circuits with active and nonlinear devices: An FDTD approach," *IEEE Trans. Microwave Theory Tech.*, Vol. 45, 1997, pp. 819–826.

[43] Simpson, J. J., A. Taflove, J. A. Mix, and H. Heck, "Computational and experimental study of a microwave electromagnetic bandgap structure with waveguiding defect for potential use as a bandpass wireless interconnect," *IEEE Microwave Wireless Components Lett.*, Vol. 14, 2004, pp. 343–345.

[44] Simpson, J. J., A. Taflove, J. A. Mix, and H. Heck, "Advances in hyperspeed digital interconnects using electromagnetic bandgap technology: Measured low-loss 42-GHz passband centered at 50 GHz," *Proc. IEEE Antennas Propagat. Soc. Intl. Symp.*, Washington, D.C., 2005.

[45] Yablonovitch, E., "Photonic band-gap structures." *J. Optical Soc. America B*, Vol. 10, 1993, pp. 283–295.

[46] Joannopoulos, J. D., R. D. Mead, and J. N. Winn, *Bandgap Structures: Molding the Flow of Light.* Princeton, N.J.: Princeton Univ. Press, 1995.

[47] Thevenot, M., A. Reineix, and B. Jecko, "FDTD approach for modeling PBG structures," *J. Optical Soc. America A*, 1999, pp. 495–500.

[48] Qiu, M., and S. He, "Guided mode in a two-dimensional metallic bandgap structure waveguide," *Phys. Lett. A.*, 2000, pp. 425–429.

[49] Service, R. F., "World's smallest transistor," *Science*, Vol. 293, Aug. 3, 2001, p. 786.

SELECTED BIBLIOGRAPHY

Choi, D. H., and W. J. R. Hoefer, "The finite-difference time-domain method and its application to eigenvalue problems," *IEEE Trans. Microwave Theory Tech.*, Vol. 34, 1986, pp. 1464–1470.

Gwarek, W. K., "Analysis of an arbitrarily-shaped planar circuit: A time-domain approach," *IEEE Trans. Microwave Theory Tech.*, Vol. 33, 1985, pp. 1067–1072.

Gwarek, W. K., "Computer-aided analysis of arbitrarily-shaped coaxial discontinuities," *IEEE Trans. Microwave Theory Tech.*, Vol. 36, 1988, pp. 337–342.

Gwarek, W. K., "Analysis of arbitrarily-shaped two-dimensional microwave circuits by finite-difference time-domain method," *IEEE Trans. Microwave Theory Tech.*, Vol. 36, 1988, pp. 738–744.

Kuo, C.-N., V. A. Thomas, S. T. Chew, B. Houshmand, and T. Itoh, "Small-signal analysis of active circuits using FDTD algorithm," *IEEE Microwave Guided Wave Lett.*, Vol. 5, 1995, pp. 216–218.

Kuo, C.-N., R.-B Wu, B. Houshmand, and T. Itoh, "Modeling of microwave active devices using the FDTD analysis based on the voltage-source approach," *IEEE Microwave Guided Wave Lett.*, Vol. 6, 1996, pp. 199–201.

Liang, G.- C., Y.-W. Liu, and K. K. Mei, "Full-wave analysis of coplanar waveguide and slotline using the time-domain finite-difference method," *IEEE Trans. Microwave Theory Tech.*, Vol. 37, 1989, pp. 1949–1957.

Maeda, S., T. Kashiwa, and I. Fukai, "Full wave analysis of propagation characteristics of a through hole using the finite-difference time-domain method," *IEEE Trans. Microwave Theory Tech.*, Vol. 39, 1991, pp. 2154–2159.

Mezzanotte, P., M. Mongiardo, L. Roselli, R. Sorrentino, and W. Heinrich, "Analysis of packaged microwave integrated circuits by FDTD," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 1994, pp. 1796–1801.

Sano, E., and T. Shibata, "Full wave analysis of picosecond photoconductive switches," *IEEE J. Quantum Electronics,* Vol. 26, 1990, pp. 372–379.

Shibata, T., and E. Sano, "Characterization of MIS structure coplanar transmission lines for investigation of signal propagation in integrated circuits," *IEEE Trans. Microwave Theory Tech.,* Vol. 38, 1990, pp. 881–890.

Toland, B., B. Houshmand, and T. Itoh, "Modeling of nonlinear active regions with the FDTD method," *IEEE Microwave Guided Wave Lett.,* Vol. 3, 1993, pp. 333–335.

Toland, B., J. Lin, B. Houshmand, and T. Itoh, "FDTD analysis of an active antenna," *IEEE Microwave Guided Wave Lett.,* Vol. 3, 1993, pp. 423–425.

Tsuei, Y.-S., A. C. Cangellaris, and J. L. Prince, "Rigorous electromagnetic modeling of chip-to-package (first-level) interconnections," *IEEE Trans. Components, Hybrids, Manufacturing Tech.,* Vol. 16, 1993, pp. 876–883.

Zhang, X., J. Fang, K. Mei, and Y. Liu, "Calculation of the dispersive characteristics of microstrips by the time-domain finite difference method," *IEEE Trans. Microwave Theory Tech.,* Vol. 36, 1988, pp. 263–267.

PROJECTS

P15.1 Construct a three-dimensional Cartesian FDTD code with a simple first-order Mur ABC. This will be used as a tool in a number of subsequent exercises. Model a 50Ω microstrip transmission line, making sure that there is ample clearance between the microstrip and the edges of the PEC ground plane, and similar clearance between the edges of the ground plane and the ABC. Excite the line by pulsing collinear E-field components located between the ground plane and the metal trace. Set up a subroutine to implement DFTs of the incident and reflected pulses in the line. Check to see that the model is working by using a single-space-cell resistive block to terminate the trace to ground with 50Ω, 100Ω, and 25Ω. For these cases, observe the time-domain reflected pulses and compare to the simple analytical theory. Calculate the frequency-domain reflection coefficient and compare with the theory.

P15.2 Using the method of Section 15.2.1, obtain $Z_0(\omega)$ and $\gamma(\omega)$ for the microstrip of P15.1 and compare with the analytical theory.

P15.3 Using the method of Section 15.2.2, obtain $Z_L(\omega)$ when the microstrip is terminated in the three loads tested in P15.1. Locate the observation point 20 space lattice cells from the load in each case.

P15.4 Using the methods of Sections 15.2.4 and 15.2.5, calculate the differential capacitance and differential inductance of the microstrip. Compare with the analytical theory.

P15.5 Introduce a thin slot (air gap) in the ground plane under the microstrip. Use the flux / current method of Section 15.3.1 to calculate the equivalent lumped series inductance due to the slot, and compare with the differential inductance obtained in P15.4.

P15.6 Repeat P15.5, but use the impedance-fitting method of Section 15.3.2. Examine the sensitivity of the answer to the location of the observation plane.

P15.7 Construct a one-dimensional FDTD code that models the impulsive illumination of a finite-thickness dielectric slab of high permittivity. Choose the slab thickness and dielectric constant so that more than 10,000 time-steps are required for the transmitted impulse response to decay essentially to zero. Use DFTs to calculate and graph $S_{21}(\omega)$.

P15.8 Repeat P15.7, but apply Prony's method of Section 15.8.1 to extrapolate the transmitted waveform. Window the FDTD data for only about 1,000 time-steps, beginning after the transmitted wave unambiguously begins to decay. Compare the extrapolated waveform and the corresponding $S_{21}(\omega)$ with that obtained with the full-term FDTD computation of P15.7.

P15.9 Repeat P15.1, but use the resistive load and resistive voltage source models of Sections 15.9.2 and 15.9.3 to excite and terminate the microstrip.

P15.10 Repeat P15.9, but use the numerical capacitor load of Section 15.9.4 to terminate the microstrip. Replicate the results of Fig. 15.10.

P15.11 Repeat P15.9, but use the numerical diode load of Section 15.9.7 to terminate the microstrip. Replicate the results of Fig. 15.13.

Chapter 16

Photonics

Geoffrey Burr, Susan Hagness, and Allen Taflove

16.1 INTRODUCTION

Recently, there has been a pronounced expansion in the application of FDTD and related *pseudospectral time-domain* (PSTD) computational techniques to simulate man-made and natural structures at optical wavelengths. This expansion is being driven by the worldwide emergence of interest in lightwave communications and computing, nanometer-scale technology, and biomedical applications of light. While one word, *photonics*, serves to describe the totality of this field, we realize that the various component technologies are simultaneously disparate and yet linked strongly together via the elegant unification posed by Maxwell's equations. FDTD and PSTD can put Maxwell's equations to work across a wide range of photonics technologies.

This chapter reviews applications of FDTD and PSTD techniques to photonics in six distinct areas: (1) index-contrast guided-wave structures, especially waveguides and microcavity ring, racetrack, and disk resonators; (2) distributed Bragg reflector devices; (3) photonic crystals; (4) frequency conversion in second-order nonlinear materials; (5) nanoplasmonic devices; and (6) biophotonics. Almost 300 references are provided to assist the reader in exploring further how FDTD and PSTD computational models are assisting in the development of these areas.

16.2 INTRODUCTION TO INDEX-CONTRAST GUIDED-WAVE STRUCTURES

Advances in nanofabrication techniques permit optical structures with physical dimensions on the order of the optical wavelength to be used as components in high-density photonic integrated circuits. Such components include waveguides, couplers, filters, multiplexers, switches, and lasers [1–4]. Micro-optical structures can be categorized into the following two groups, based on the physical mechanism by which they confine light [5]:

- Group 1: Confinement by refractive-index contrast, leading to total internal reflection as in *whispering-gallery-mode* (WGM) disk cavities;
- Group 2: Confinement by multiple wave reflections within periodic structures, as in *vertical-cavity surface-emitting lasers* (VCSELs) and photonic crystals.

The next four sections of this chapter review FDTD applications to index-contrast guided-wave structures, especially optical waveguides and microcavity ring, racetrack, and disk resonators.

16.3 FDTD MODELING ISSUES

The FDTD method permits accurate analysis of the complex electromagnetic wave phenomena inherent in the operation of index-contrast guided-wave structures. This section introduces basic issues related to FDTD modeling of such structures.

16.3.1 Optical Waveguides

The submicron width and large refractive index contrast of waveguides commonly found in micron-scale integrated optical devices leads to significant waveguide dispersion. The relatively strong dependence of the waveguide propagation constant upon frequency impacts several key design parameters for microcavity resonators composed of such waveguides, including their Q and the *free spectral range* (FSR), defined as the spacing in frequency or wavelength between adjacent resonances. Accurate FDTD modeling of highly dispersive waveguides requires careful consideration of two important issues:

- How to terminate waveguides that extend beyond the lattice boundaries;
- How to source specific waveguide modes for the case of pulsed excitation.

The accuracy of the FDTD method for problems in linear optics was first demonstrated for the directional coupler [6]. Since then, Berenger's PML ABC [7] has provided the increased computational dynamic range required for sensitive numerical simulations, such as the analysis of antireflection coatings [8]. Of particular importance here is the demonstration in [9] that multimodal and dispersive waveguiding structures extending beyond the edges of the FDTD space lattice can be terminated accurately by the PML ABC, with reflection coefficients below –75 dB for a wide range of group velocities.

Consider a dielectric slab waveguide and its TE optical modes. Sourcing the numerical analog of a propagating TE mode is accomplished by specifying an electric field distribution along a one-dimensional transverse cross section of the waveguide. For the case of sinusoidal excitation, the transverse field distribution of the desired guided mode is obtained by numerically solving a determinant equation for the transverse propagation constant and decay constants. For example, given the refractive indexes of the core and cladding layers, the dispersion relations for the various waveguide modes can be computed [10]. With these data, the longitudinal propagation constant β corresponding to the desired frequency of excitation and mode number is selected, and used to calculate the transverse propagation constant and decay constants.

The case of a pulsed source, such as a Gaussian pulse modulating a carrier wave, poses more of a challenge. Even if the frequency spectrum of the pulse is narrow enough and the carrier frequency is low enough to allow only the fundamental mode to propagate, there still exists a spread of β values along the fundamental mode dispersion curve corresponding to the spread of frequencies of interest, as illustrated in Fig. 16.1.

In this case, the transverse mode profile for the pulse can be approximated using the value of β at the carrier frequency. The fact that the transverse distribution of the source is not exact for all frequencies in the pulse spectrum means that a small amount of energy is shed from the waveguide early in the FDTD simulation. The accuracy of simulations in which the waveguide is immediately surrounded by the PML ABC is not affected by this shed energy. However, shed energy from the waveguide could interact with other structures in the grid, thereby introducing unwanted noise into the computation. This should be carefully examined.

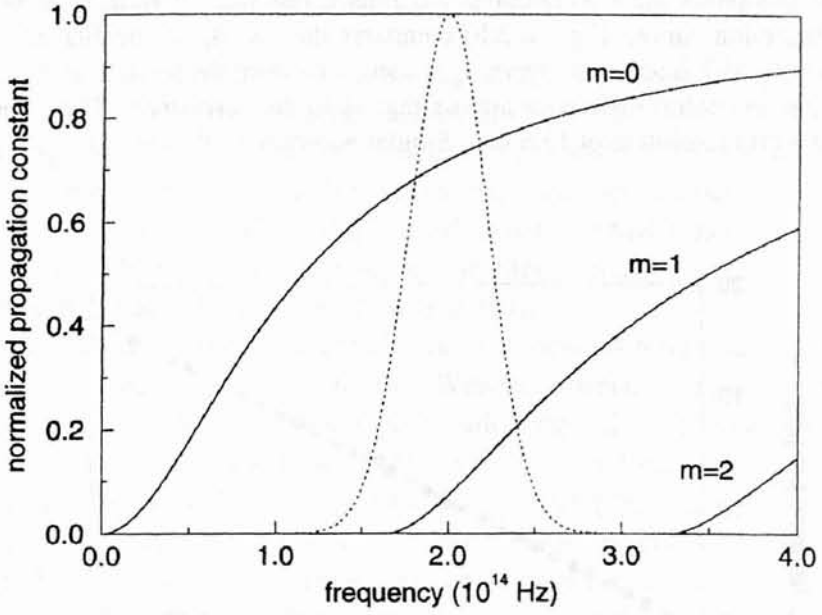


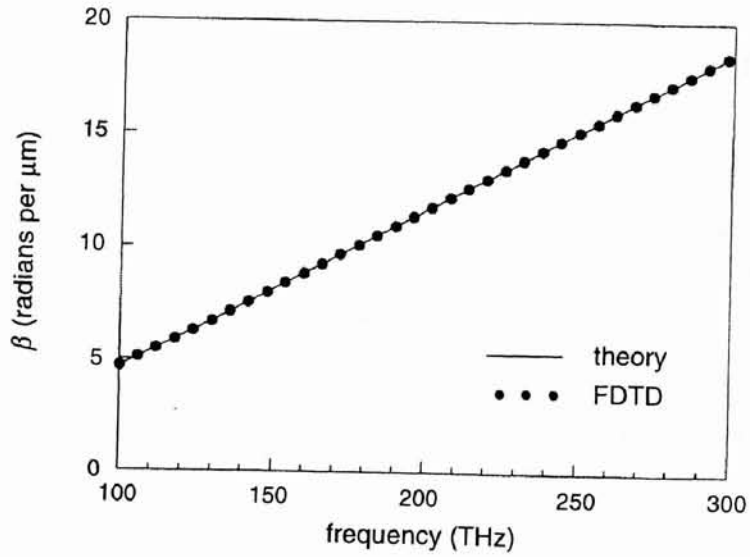
Fig. 16.1 Excitation spectrum (dotted line) superimposed on the waveguide dispersion relations (solid lines) for the first three TE modes of a symmetric, strongly guiding dielectric slab waveguide. The index of refraction n for the 0.3-μm-wide core is 3.2; for the cladding layers, $n = 1.0$.

An effective solution to this problem is to use "bootstrapping"; that is, to perform a preliminary FDTD simulation of a sufficiently long version of the waveguide, and store the time history of the transverse E-field distribution as the pulse passes an observation plane at the far end of the lattice. The stored data are used as the waveguide source excitation in all subsequent simulations. The level of numerical noise introduced into the grid from this bootstrapped source is very low, comparable to that arising from reflections at the PML ABC.

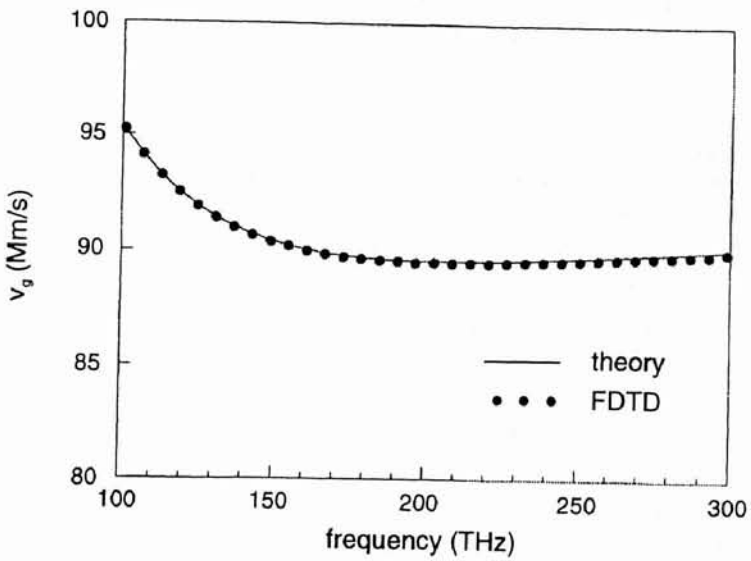
The accuracy of the FDTD method for modeling propagating waveguide modes is now demonstrated in a sample computation of the longitudinal propagation constant β and the group velocity v_g of a strongly guiding planar waveguide. Here, the waveguide consists of a 0.3-μm-wide core of high refractive index ($n = 3.2$) having low-refractive-index ($n = 1.0$) cladding on both sides. The PML ABC is used to terminate the waveguide in the lattice. Since the only structure in this simulation is the straight waveguide, a bootstrapped source is not required, and the inexact mode profile for the pulsed source of the waveguide is used. (Any energy shed from the waveguide is absorbed by the PML ABC.) The waveguide must simply be made long enough to ensure that the mode has settled down before recording any data.

A single 20-fs Gaussian pulse modulating a carrier of frequency $f_0 = 2.0 \times 10^{14}$ Hz is launched at one end of the grid, using the fundamental ($m = 0$) mode profile computed for f_0. Fig. 16.1 shows the source spectrum superimposed on the normalized waveguide dispersion curves for the first three TE modes. At f_0, this waveguide can support only the $m = 0$ mode, which has spatially even symmetry about the longitudinal axis of the waveguide, and the $m = 1$ mode, which has odd symmetry. Because the numerical excitation has even symmetry, it generates only the $m = 0$ mode. At the far end of the grid, two fixed observation points located along the longitudinal axis of the core record E-field data as the pulse passes. By taking the ratio of the DFTs of the two time histories, β is computed over the full bandwidth of the pulse.

Fig. 16.2(a) compares the FDTD-computed β as a function of frequency with the exact denormalized dispersion curve. Fig. 16.2(b) compares the corresponding FDTD-computed and exact variations of v_g with frequency. Here, v_g is computed from the numerical data for β using a second-order-accurate central-difference approximation to the derivative. The error in β at f_0 is less than 0.2% at a grid resolution of 13.6 nm. Similar accuracy is obtained for v_g.



(a) Propagation constant β.



(b) Group velocity v_g.

Fig. 16.2 Comparison of FDTD-computed and theoretical values for properties of the fundamental TE mode propagating in the strongly guiding dielectric-slab waveguide characterized in Fig. 16.1.

16.3.2 Material Dispersion and Nonlinearities

To fully treat the broadband behavior of integrated optical devices, the effect of material dispersion should be included in the FDTD calculations. Over the past several years, a number of different methods have been proposed to account for material dispersion with absorption. These include *recursive-convolution* (RC) methods that implement a discrete convolution of the dispersion relation [11, 12], *auxiliary differential equation* (ADE) methods that discretize a differential equation obtained from the relevant constitutive relation [13, 14], and Z-transform methods based on digital signal processing techniques [15].

Additional extensions to the ADE and Z-transform methods have enabled the modeling of $\chi^{(2)}$ and $\chi^{(3)}$ nonlinear optical materials [16–19]. With this advent, the range of FDTD modeling applications has been substantially expanded to include temporal and spatial soliton propagation [16, 18–21], self-focusing of optical beams [17], scattering from linear-nonlinear interfaces [22], pulse propagation through nonlinear corrugated waveguides [23], pulse-selective behavior in nonlinear Fabry-Perot cavities [24], and nonlinear second-harmonic generation [18].

FDTD algorithms for modeling optical gain media on a macroscopic level have also been reported [25, 26]. Advances in modeling the interaction of light with optical materials on a microscopic level have enabled the modeling of ultrafast pulse interactions with two-level atoms [27] and lasing by four-level atoms [28, 29]. See Chapter 9 for details of FDTD algorithms used to model a broad range of nonlinear, dispersive, and gain materials.

16.4 LATERALLY COUPLED MICROCAVITY RING RESONATORS

This section reviews the use of the FDTD method for modeling laterally coupled microcavity ring resonators. Such devices are useful components for wavelength filtering, routing, switching, modulation, and multiplexing/demultiplexing applications [30]. The quality of a resonator for optical communications is measured by its: (1) FSR, or wavelength spacing between adjacent resonances; (2) finesse, or ratio of the FSR to the width of the resonance; and (3) 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 an FSR greater than 30 nm is desirable for accommodating the WDM channels within the erbium-amplifier communications window.

To achieve a large FSR, the diameter of the ring (or disk, as discussed in Section 16.5) must be just a few microns. Semiconductor ring and disk resonators based on high-index-contrast, strongly guiding waveguides can be designed with diameters as small as 1 to 2 μm 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 [1, 31]. For example, Fig. 16.3 shows scanning electron microscope images of 10.5-μm-diameter AlGaAs/GaAs microcavity ring and disk resonators laterally coupled to 0.5-μm-wide waveguides across 100-nm air gaps.

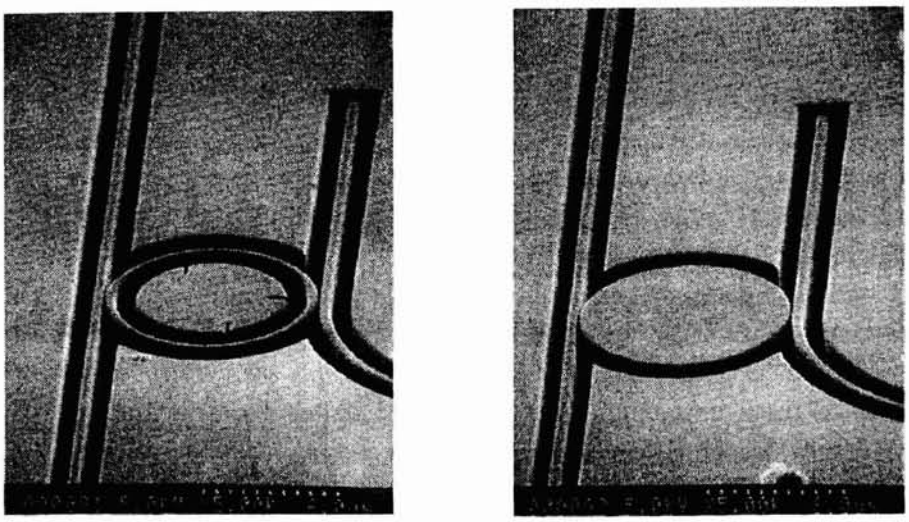


Fig. 16.3 Scanning electron microscope images of 10.5-μm-diameter AlGaAs / GaAs ring and disk resonators laterally coupled to 0.5-μm-wide waveguides across 100-nm air gaps. *Source:* Rafizadeh et al., *Optics Letters*, 1997, pp. 1244–1246.

16.4.1 Modeling Considerations: Two-Dimensional FDTD Simulations

The resonator geometries of Fig. 16.3 are clearly three dimensional. In the vertical direction, each semiconductor waveguide consists of a 0.45-μm-thick GaAs guiding layer with AlGaAs cladding layers above and below. Strong lateral confinement is achieved by etching air-filled trenches down through the lower cladding layer. Confinement in the vertical direction is deliberately weak to allow optical coupling from an external fiber or planar-emitting device. To conserve computer resources, two-dimensional FDTD models can be applied. In these models, the vertical waveguide thickness and inhomogeneous material composition is accounted for by computing the effective refractive index n_{eff} for the fundamental mode at $\lambda = 1.5$ μm, and then using n_{eff} as the bulk material index of the core in the two-dimensional simulations. In the cases considered here, each waveguide core ($n = 3.2$) is bordered on each side by air ($n = 1$).

Fig. 16.4 is a schematic diagram of the two-dimensional FDTD model of a microcavity ring or disk resonator of the type shown in Fig. 16.3. Waveguides WG1 and WG2 serve as evanescent wave input and output couplers. Using the bootstrapping technique discussed in Section 16.3.1, a fundamental mode is sourced at the left end of WG1. In the simulations discussed here, the excitation is a 20-fs Gaussian pulse modulating a 200-THz carrier.

The standard FDTD method is used in two dimensions for either the TM or TE mode. Transverse field components in the pulsed waveguide mode are recorded as the mode propagates past cross sections P1 and P2 in WG1, and P3 in the microring or microdisk. By taking DFTs of the time waveforms of the fields, calculating the Poynting power densities along the cross sections, and then summing, the longitudinal power flux passing through each cross section is obtained as a function of frequency. The total power computed in this manner is found to be conserved to within parts per 10,000 for cases where the optical radiation is negligible, making possible a reliable spectral characterization of the coupling efficiency (flux through P3 normalized by the incident flux through P1), and transmittance (normalized flux through P2).

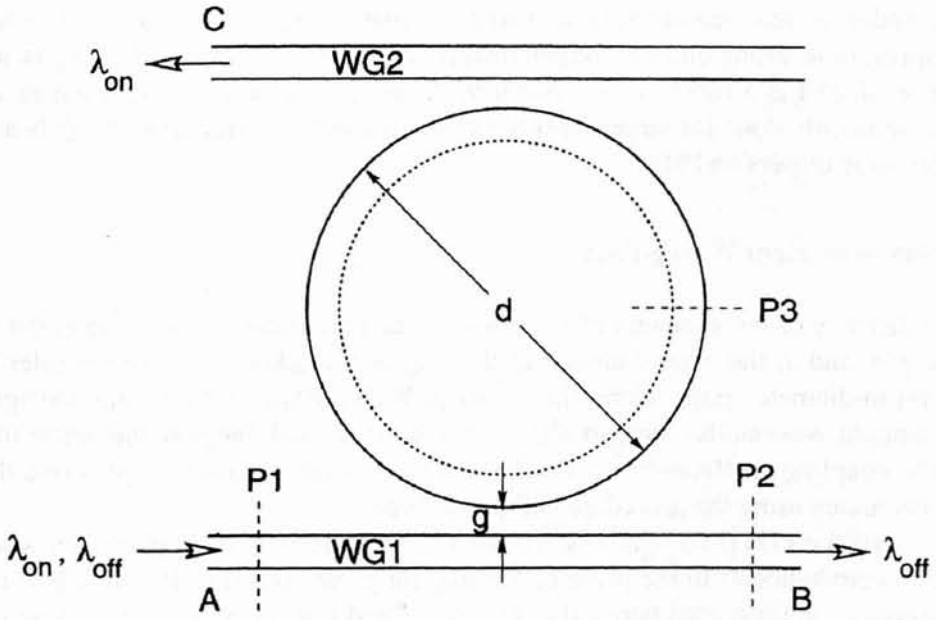


Fig. 16.4 Schematic diagram of a microcavity ring or disk resonator laterally coupled to two straight waveguides. An off-resonance signal entering port A remains in waveguide WG1 and exits at port B. An on-resonance signal is switched to waveguide WG2 and exits at port C. *Source:* Hagness et al., *J. Lightwave Technology*, 1997, pp. 2154–2165, © 1997 IEEE.

We note that the etching process that yields the physical structures shown in Fig. 16.3 results in a 10 to 20 nm roughness of the waveguide sidewalls. This fabrication artifact scatters light propagating within the waveguide, and causes back-reflection into the counterpropagating mode [32]. Since FDTD grid-cell sizes in the 10 to 20 nm range are computationally feasible for two-dimensional problems spanning 10 μm or more, we can directly and simply model this surface roughness via the usual staircasing inherent in Cartesian FDTD grids. In fact, several variations of a staircase fit can be implemented to examine the optical effects of randomness in the fabrication process. The fine 10 to 20 nm grid-cell size also permits detailed modeling of optical coupling across the approximately 100-nm waveguide-resonator air gaps.

Optical scattering due to waveguide surface roughness also causes radiation. This results in a partial loss of the stored energy and a reduction of the resonator quality factor, Q [33]. We note that such radiation losses can also be caused by waveguide bends. However, because of the strong mode confinement, bending losses are negligible in these resonators. For the passive devices considered here, the intrinsic absorption loss due to the materials is also negligible.

The importance of accurate modeling of semiconductor waveguides having highly dispersive wave-propagation characteristics is illustrated with the following equation for the FSR between longitudinal modes m and $m + 1$:

$$\Delta f_{\text{FSR}} = \frac{c}{\pi}\left[\frac{m+1}{d_{\text{eff}}(f_{m+1})\, n_{\text{eff}}(f_{m+1})} - \frac{m}{d_{\text{eff}}(f_m)\, n_{\text{eff}}(f_m)}\right] \tag{16.1}$$

where d_{eff} is the effective diameter of the ring or disk. The quantity πd_{eff} corresponds to the circumferential path traveled by the peak of the guided mode.

For semiconductor microcavity ring and disk resonators, $n_{\text{eff}}(f_{m+1}) \neq n_{\text{eff}}(f_m)$, since the waveguide dispersion is strong and the longitudinal mode spacing is wide. Also, d_{eff} is smaller than the diameter d, and is a function of frequency. Thus, (16.1) cannot be reduced to $\Delta f_{\text{FSR}} = c/\pi d n_{\text{eff}}$, as is commonly done for larger weakly guiding resonators. Similarly, the Q is affected by the relatively large dispersion [34].

16.4.2 Coupling to Straight Waveguides

The coupling efficiency of the geometry of Fig. 16.4 can be investigated by varying g, the size of the coupler air gap, and d, the outer diameter of the ring or disk [31]. We now consider such a study for a 5.0-μm-diameter ring. Here, the width of WG1, WG2, and the ring waveguide is 0.3 μm. The straight waveguides support only one symmetric and one antisymmetric mode at $\lambda = 1.5$ μm. The coupling coefficient κ, defined as the percentage of power coupled into the ring from WG1, is computed using the procedure outlined above.

Fig. 16.5 shows the FDTD-computed κ for the vertical-polarization (V-pol) case, where the optical E-field is perpendicular to the plane of the ring for g varying from 191 to 259 nm. Two coupling phenomena are illustrated here [31]: (1) for a fixed frequency, κ decreases as the gap widens; and (2) for a fixed gap, κ decreases at higher frequencies. At $\lambda = 1.5$ μm ($f = 200$ THz), κ ranges between 0.7% and 2.8% for a change of only 68 nm in g. This compares well with the 0.5% to 3% κ range needed to achieve good transmission characteristics and large extinction ratios. Convenient rules of thumb are that κ should be greater than the round-trip cavity loss, and should be symmetric for both WG1 and WG2 relative to the ring or disk. Since the FDTD simulations indicate that κ is sensitive to g, and since it is difficult to fabricate identical gaps to very high precision, the coupling from the ring to WG1 and WG2 may be slightly asymmetric.

Fig. 16.6 compares the FDTD-computed V-pol coupling coefficient curve of Fig. 16.5 for $g = 232$ nm with FDTD results for the same gap size computed for the horizontal-polarization (H-pol) case, where the E-field is in the plane of the ring [31]. The coupling coefficient for the H-pol case at 1.5 μm is 1.9%, compared with 1.2% for the V-pol case.

The results of Fig. 16.6 show that the H-pol coupling exceeds the V-pol coupling for most of the spectrum of interest. In order to avoid high lateral coupling of light to the media surrounding the waveguides and the ring or disk, relatively wide boundary trenches must be etched, as seen in the images of Fig. 16.3. While the remainder of this section focuses on the V-pol case, note that with a sufficiently wide trench, the H-pol case is qualitatively the same.

The V-pol FDTD results for a 10-μm-diameter ring have also been obtained for the same range of gap sizes. The 10-μm ring has about twice the κ of the 5.0-μm ring for a given gap size and frequency, ranging from 1.3% to 5.5% at $\lambda = 1.5$ μm. The interaction length for the larger ring is still short enough to avoid back-coupling from the ring to the waveguide.

16.4.3 Coupling to Curved Waveguides

In addition to varying the gap size and ring diameter, the waveguide-to-ring coupling can be manipulated by adding curvature to the coupling waveguides. Detailed FDTD modeling studies [31] have shown that the coupling *drops* and becomes less sensitive to frequency as we bend the input waveguide from the straight tangent path shown in Fig. 16.4 toward the ring. Although curving the waveguide toward the ring increases their interaction length, it introduces a path-length difference due to their differing bending radii. This causes a phase mismatch between the waveguide and ring signals, reduces the net coupling, and smooths out frequency variations.

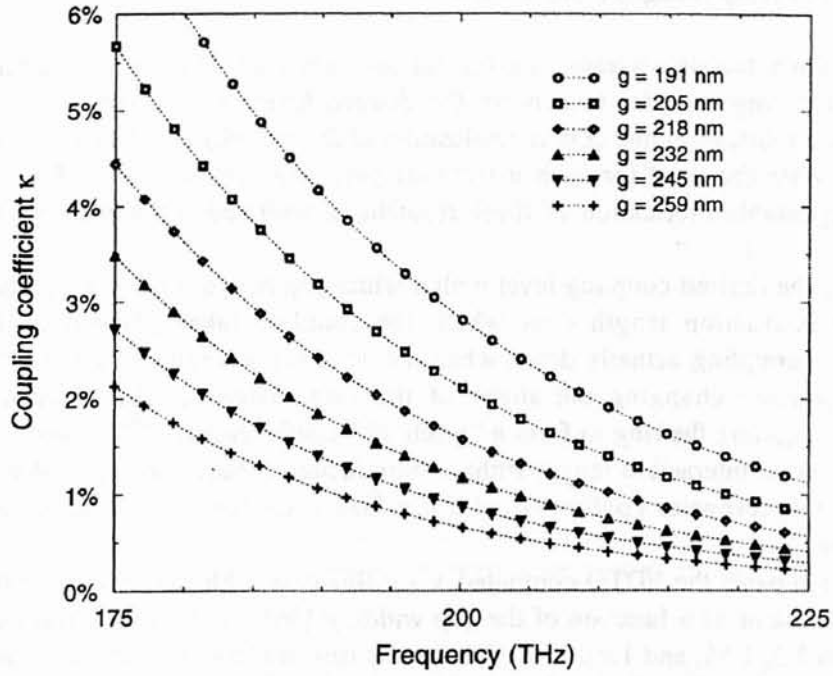


Fig. 16.5 FDTD-computed coupling coefficient κ as a function of frequency and gap size g for the geometry of Fig. 16.4, $d = 5.0\ \mu$m ring resonator case. The E-field is vertically polarized; that is, perpendicular to the plane of Fig. 16.4. *Source:* Hagness et al., *J. Lightwave Technology*, 1997, pp. 2154–2165, © 1997 IEEE.

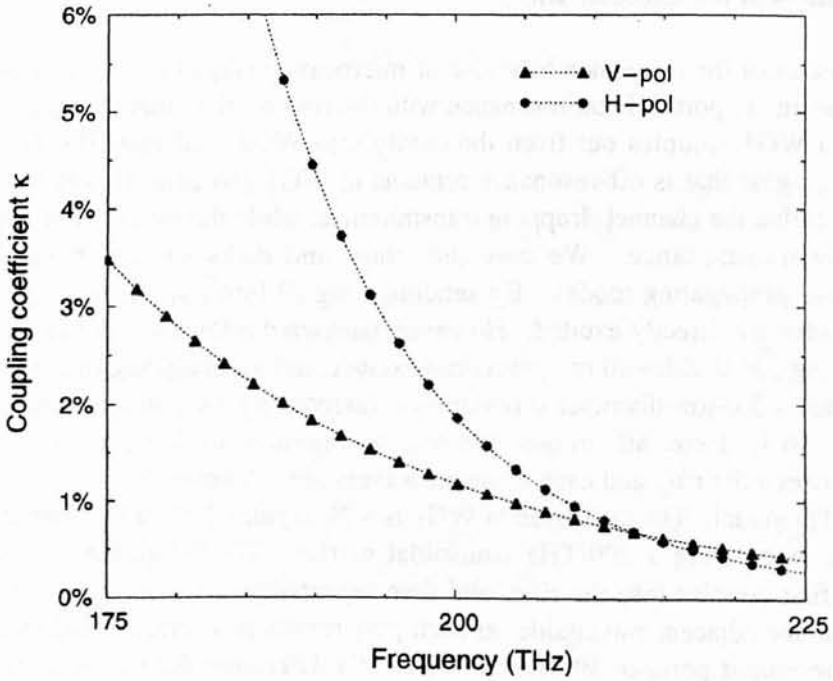


Fig. 16.6 Comparison of the FDTD-computed V-pol coupling-coefficient curve of Fig. 16.5 for $g = 232$ nm with H-pol FDTD results for the same gap size. *Source:* Hagness et al., *J. Lightwave Technology*, 1997, pp. 2154–2165, © 1997 IEEE.

16.4.4 Elongated Ring Designs ("Racetracks")

As discussed above, the strong transverse optical confinement of the waveguide requires it to be very close to the ring in order to achieve the desired level of coupling. An important step towards practical fabrication and commercialization of these devices is the exploration of designs that might alleviate the need for such narrow air gaps. In particular, a wider gap may allow reliable and repeatable fabrication of these resonators, with reduced sensitivity to fabrication variations.

To achieve the desired coupling level with a wider gap between the waveguide and the ring resonator, the interaction length over which the coupling takes place must be increased. As noted above, coupling actually drops when the adjacent waveguide curves toward the ring. Here, we investigate changing the shape of the ring instead of the adjacent waveguide. For example, elongating the ring to form a "racetrack" configuration [35], as shown in Fig. 16.7, leads to an increased interaction length without introducing a phase mismatch. We shall consider as an example this resonator configured with $d = 5.0$ μm semicircular sections and $L = 1.5$ μm straight sections.

Fig. 16.8 compares the FDTD-computed κ for this case with the previous 5.0-μm-diameter circular ring resonator as a function of the gap width, g [36]. The dotted curves represent κ at the wavelengths 1.5, 1.55, and 1.6 μm for the circular ring, and the solid curves represent κ at the same three wavelengths for the racetrack geometry. At a given wavelength, a specific coupling level is achieved at a larger gap width with the racetrack resonator, in comparison with the circular ring resonator. These results illustrate the potential advantage of the racetrack structure in allowing wider air gaps while providing the desired coupling levels.

16.4.5 Resonances of the Circular Ring

This section discusses the resonance behavior of microcavity rings [31]. Referring to Fig. 16.4, if the signal that enters port A is on-resonance with the ring or disk, then that signal couples into the cavity from WG1, couples out from the cavity into WG2, and exits the device at port C. Alternatively, a signal that is off-resonance remains in WG1 and exits at port B. The output at port B is often called the channel-dropping transmittance, while the output at port C is called the channel-passing transmittance. We note that rings and disks support both clockwise and counterclockwise propagating modes. By sending a signal into port A, only counterclockwise propagating modes are directly excited. However, backward reflections in the coupling regions and backscattering due to sidewall roughness can excite counterpropagating modes.

We consider a 5.0-μm-diameter circular ring coupled by two straight waveguides in the manner of Fig. 16.4. Here, all straight and ring waveguides are 0.3 μm wide, and there is a 232-nm gap between the ring and each adjacent waveguide. A uniform 13.6-nm grid-cell size is used in the FDTD model. The excitation in WG1 is a 20-fs *full-width at half-maximum* (FWHM) Gaussian pulse modulating a 200-THz sinusoidal carrier. Time-stepping is continued as the exciting pulse first couples into the ring, and then repeatedly circumnavigates it. Coupling of optical energy to the adjacent waveguides at each pass results in a series of pulses of diminishing amplitude at the output ports of WG1 and WG2. To determine the spectral properties of the output signals, DFTs are performed on these signals concurrently with the FDTD time-stepping. High Q-factors require either extended time-stepping windows or extrapolation of unconverged windowed responses using, for example, Prony's method (see Chapter 15, Section 15.8.1).

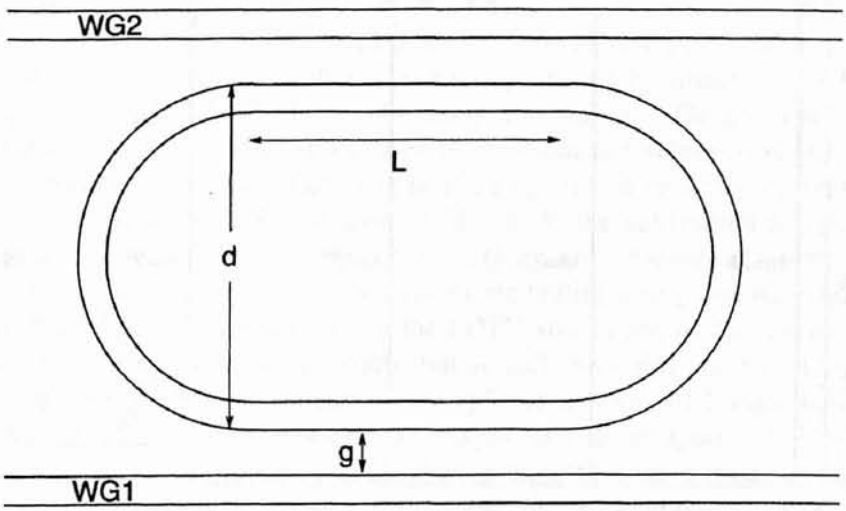


Fig. 16.7 Schematic diagram of an elongated-ring "racetrack" microcavity resonator laterally coupled to straight waveguides.

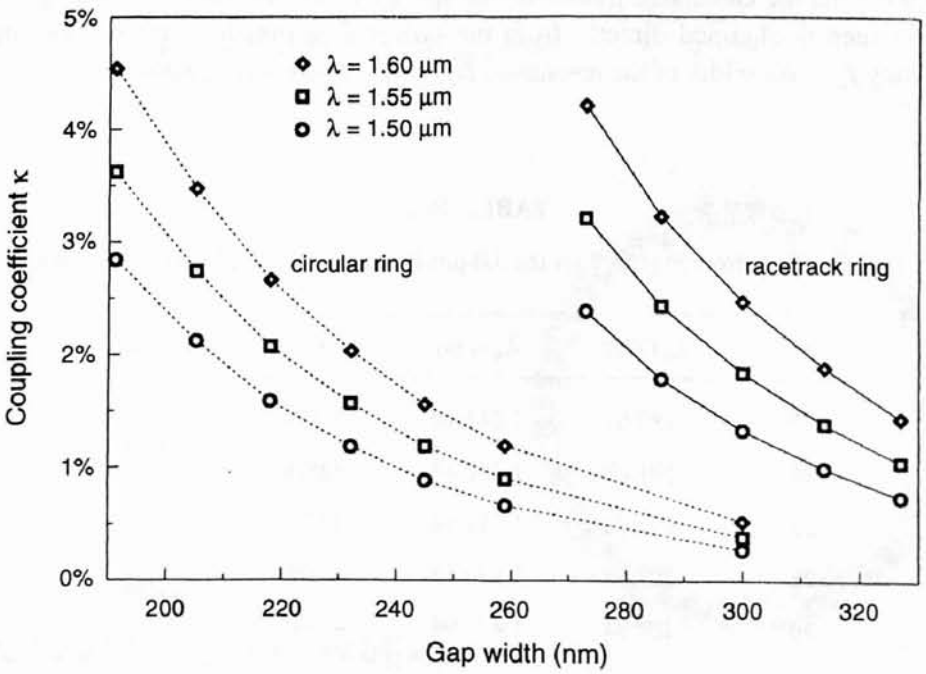


Fig. 16.8 FDTD-computed coupling coefficient κ as a function of the wavelength and the gap width g for a 5.0-μm-diameter circular ring resonator (dotted lines), and a racetrack ring resonator having 5.0-μm-diameter semicircular sections and 1.5-μm-long straight sections (solid lines). The E-field is vertically polarized for this case. *Source:* Hagness et al., *Integrated Photonics Research*, OSA Technical Digest Series.

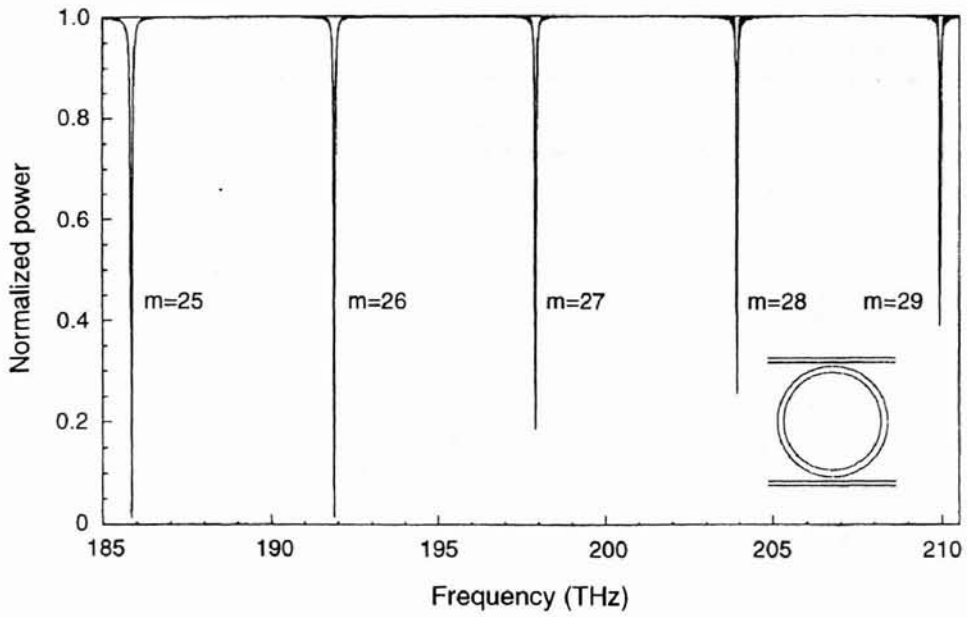


Fig. 16.9 FDTD-computed transmittance spectrum of a 5.0-μm-diameter circular ring resonator.

Fig. 16.9 graphs the calculated transmittance spectrum for this microcavity ring. The Q of the m'th resonance is obtained directly from the power spectrum by forming the ratio of the center frequency f_m to the width of the resonance Δf_m at the half-power points.

TABLE 16.1

Resonance Data from Fig. 16.9 for the 5.0-μm-Diameter Circular Microcavity Ring

m	f_m (THz)	λ_m (nm)	Q	FSR (nm)
25	185.85	1,613.10	3,700	
				50.66
26	191.88	1,562.44	5,000	
				47.56
27	197.90	1,514.88	4,000	
				44.73
28	203.92	1,470.15	4,000	
				42.17
29	209.94	1,427.98	4,000	

Table 16.1 lists the resonant frequencies, Q factors, and FSR values calculated from the data in Fig. 16.9. We note that the FSR decreases as the frequency increases. This is a consequence of the relatively large waveguide dispersion, as seen in (16.1). The finesse, calculated from the FSR values and the widths of the resonances, falls in the range between 115 and 160.

Impact of Sidewall Roughness

Based only on coupling, which decreases at higher frequencies as seen in Fig. 16.5, the Q should increase with frequency. This is observed for the $m = 25, 26$ resonances in Table 16.1, but not for the higher-frequency resonances due to scattering caused by imperfections in the etched sidewalls (approximated in the FDTD model using staircasing). On an expanded frequency scale, each of the $m = 28, 29$ resonances is seen to represent two subresonances. This is caused by sidewall roughness, which backscatters optical energy into a counterpropagating mode that splits the resonance peak [32]. In the case of Fig. 16.9, the subresonances are not split far enough to be individually discernible, so the apparent Q of each resonance is lowered.

The drop in the apparent Q has been verified by simulating a ring that has rougher sidewalls by a factor of two. This involves modifying the FDTD staircasing of the sidewalls to mimic a twice-coarser grid in the vicinity of the walls, but actually retaining the fine 13.6-nm grid-cell size everywhere. Here, all five resonances are split as a result of higher scattering losses. Further, the two highest-frequency resonances are split far enough apart to individually resolve each subresonance. This study reveals that the Q does increase with increasing mode number for each subresonance, as expected based on the strong dependence of coupling on frequency.

Extinction Ratio

Fig. 16.10 illustrates the potential for high extinction ratios. First, the sinusoidal steady-state E-field is computed for excitation at the nonresonant frequency 193.4 THz (1.55 μm), as shown in Fig. 16.10(a). In this case, nearly 100% of the signal remains in WG1. Second, the sinusoidal steady-state E-field is computed for excitation at the $m = 26$ resonant frequency of 191.9 THz. As shown in Fig. 16.10(b), nearly 100% of the signal in this case is switched to WG2.

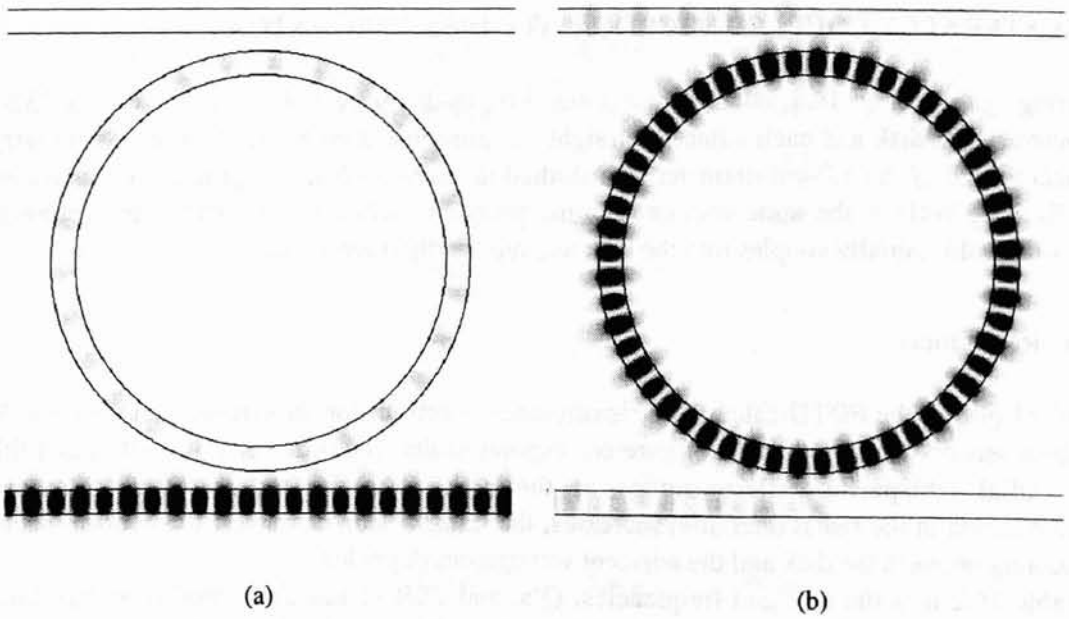


Fig. 16.10 Grayscale visualizations of the FDTD-computed sinusoidal steady-state E-field in the 5.0-μm diameter ring. Here, a single-frequency (sinusoidal) excitation is applied at port A of WG1: (a) off-resonance signal at 193.4 THz; (b) $m = 26$ on-resonance signal at 191.9 THz.

Rejection Ratio

The on/off ratio, or rejection ratio, is defined as the ratio of power transmitted at a resonance frequency to the power not transmitted at that frequency. Consider the 30-nm communications window of erbium amplifiers centered around 1.55 μm. Within this window, that is, around the $m = 26$ resonance (191.9 THz) in Fig. 16.9, the rejection ratio is approximately 72:1. By slightly decreasing the gap, the rejection ratio is increased to more than 500:1. In general, this ratio can be optimized over any desired band spanning a few terahertz by adjusting the gap size. This changes the coupling with respect to the scattering losses in that wavelength range.

Effective Diameter

Since n_{eff} is a known function of frequency for straight waveguides, we can estimate d_{eff} for the ring resonator using $mc/\pi f_m n_{\text{eff}}$. For the resonances in Fig. 16.9, d_{eff} increases with frequency from 4.717 to 4.720 μm. If the circumferential path of the mode were along the center of the waveguide, d_{eff} would be 4.7 μm, halfway between the outer-rim diameter of 5.0 μm and the inner-rim diameter of 4.4 μm. The estimate of d_{eff} is slightly larger, agreeing with the expectation that the peak of the mode profile shifts toward the outer edge of the curved waveguide.

Overall Assessment

The 5.0-μm-diameter circular ring resonator demonstrates narrow channel wavelength selectivity (high finesse) and constant transmittance over FSR values as wide as 6.026 THz (50.66 nm). This structure exhibits excellent rejection and extinction ratios. Its Q is limited primarily by the degree of sidewall roughness, which is a function of the fabrication process.

16.5 LATERALLY COUPLED MICROCAVITY DISK RESONATORS

Referring again to Fig. 16.4, consider a circular 5.0-μm-diameter disk resonator with a 232-nm gap between the disk and each adjacent straight 0.3-μm-wide waveguide. The disk geometry is identical to that of the 5.0-μm-diameter ring studied in Section 16.4, except that there is no inner rim. Exciting WG1 in the same manner as in the previous studies, the simulation is continued in time as the pulse initially couples into the disk and repeatedly travels around it.

16.5.1 Resonances

Fig. 16.11 graphs the FDTD-calculated transmittance spectrum for the circular microcavity disk. The three sets of resonances in this figure correspond to the first-order, second-order, and third-order radial whispering-gallery modes of the disk. As the radial order (number of E-field maxima in the radial direction) increases, the transmission characteristics worsen because the coupling between the disk and the adjacent waveguides degrades.

Table 16.2 lists the resonant frequencies, Q's, and FSR values calculated from the data in Fig. 16.11. In this table, q indicates the radial order of the mode and "NA" indicates that the resonance is too weak to obtain accurate data. Comparing with the data for the ring in Table 16.1, the disk's first-order radial-mode resonances exhibit several interesting properties:

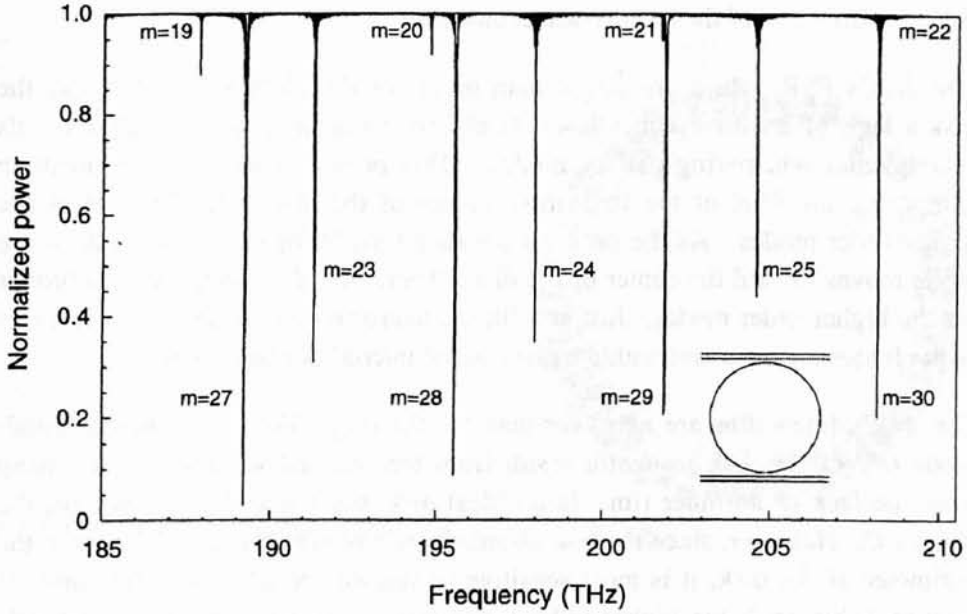


Fig. 16.11 FDTD-computed transmittance spectrum of a 5.0-μm-diameter circular disk resonator.

TABLE 16.2

Resonance Data from Fig. 16.11 for the 5.0-μm-Diameter Circular Microcavity Disk

q	m	f_m (THz)	λ_m (nm)	Q	FSR (nm)
1	27	189.20	1,584.51	7,600	
					51.32
1	28	195.54	1,533.19	7,500	
					47.98
1	29	201.85	1,485.21	7,000	
					45.03
1	30	208.16	1,440.18	9,100	
2	23	191.29	1,567.24	8,700	
					52.69
2	24	197.94	1,514.55	9,900	
					49.12
2	25	204.58	1,465.43	11,000	
3	19	187.83	1,596.10	9,900	
					57.09
3	20	194.80	1,539.01	9,700	
					52.96
3	21	201.74	1,486.05	8,800	
					NA
3	22	~208.65	NA	NA	

1. The disk's resonances are at lower frequencies than those of the ring because the effective index of the fundamental whispering-gallery modes in the solid disk is different from that of the narrow-waveguide ring.

2. The disk's FSR values are larger than those for the circular ring because the disk's lack of an inner rim allows its effective diameter to be smaller for its fundamental whispering-gallery modes. This phenomenon is also evident in comparing the FSR of the first-order modes of the disk with the FSR of the higher-order modes. As the order of the radial mode increases, the peak of the mode moves toward the center of the disk. Therefore, d_{eff} decreases even further for the higher-order modes. Just as with the microcavity ring, the FSR decreases as the frequency increases within a given set of microdisk resonances.

3. The disk's linewidths are narrower than for the ring. The higher fundamental-mode Q's of the disk resonator result from the reduced scattering loss arising from the lack of an inner rim. In an ideal disk, the low-order modes have the highest Q. However, since the lowest-order mode is confined most closely to the perimeter of the disk, it is most sensitive to sidewall roughness. Therefore, as seen in Table 16.2, the higher-order mode resonances have slightly higher Q's because of lower scattering loss.

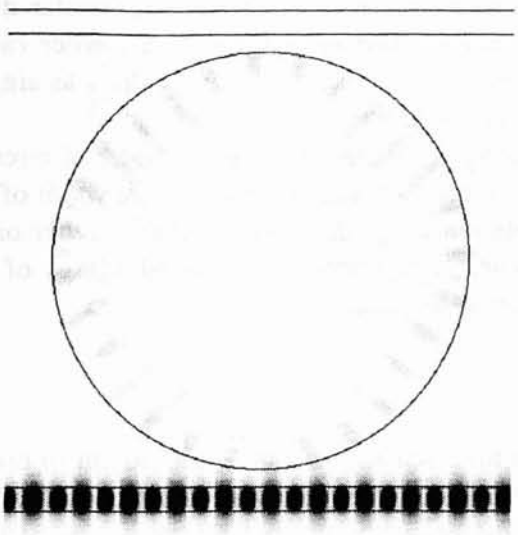
To observe the whispering-gallery modes of the microdisk resonator, single-frequency FDTD simulations are performed at the resonance frequencies pinpointed by Fig. 16.11. Fig. 16.12(a) visualizes the sinusoidal steady-state E_z distribution for a continuous excitation at the nonresonant frequency, $f = 193.4$ THz ($\lambda = 1.55$ μm). Calculations made from the FDTD-computed field data for this case show that 99.98% of the power in the incident signal at 193.4 THz remains in WG1.

Fig. 16.12(b) visualizes the E_z distribution within the microdisk at the $m = 27$ resonance of the first-order radial whispering-gallery mode. We note the resonant field enhancement inside the disk. While 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. 16.12(a), since the grayscale is normalized to the peak field amplitude. For this on-resonance case, 99.79% of the incident power in WG1 switches to WG2.

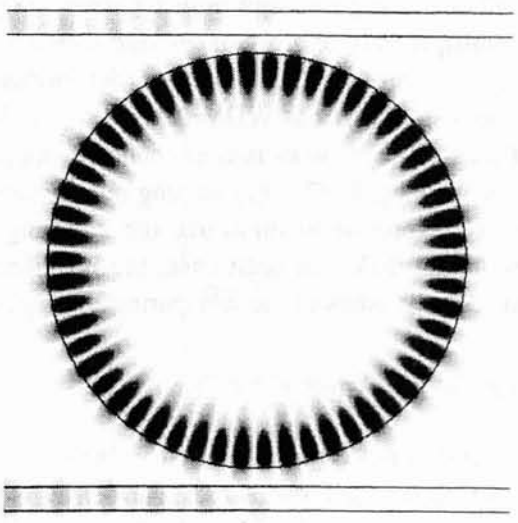
Fig. 16.12(c) visualizes the E_z distribution within the microdisk at the $m = 23$ resonance of the second-order radial whispering-gallery mode. Here, the field distribution has two local maxima in the radial direction, with the higher-amplitude field maximum shifted radially toward the center of the disk, away from the disk's edge.

Finally, Fig. 16.12(d) visualizes the E_z distribution within the microdisk at the $m = 19$ resonance of the third-order radial whispering-gallery mode. Here, the field distribution has three local maxima in the radial direction, with the highest-amplitude field maximum shifted radially even further toward the center of the disk than in Fig. 16.12(c).

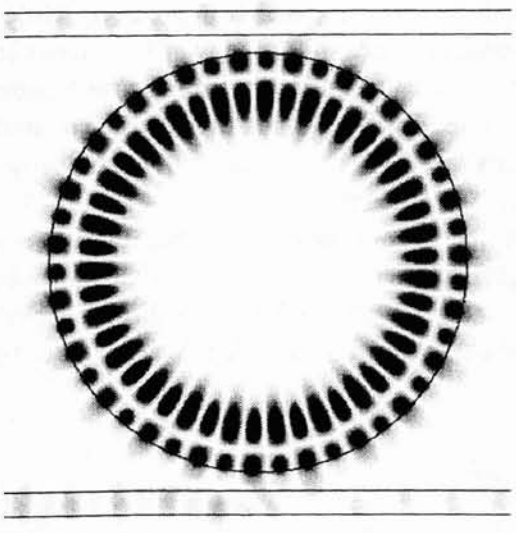
The visualizations of Fig. 16.12 show how each resonant whispering-gallery mode is confined to an annulus around the perimeter of the disk, wherein the width of the annulus increases as the radial order increases. A comparison between Figs. 16.12(a and b) indicates the potential for high extinction ratios and low crosstalk between channels in these devices if the higher order modes of the microdisk can be suppressed.



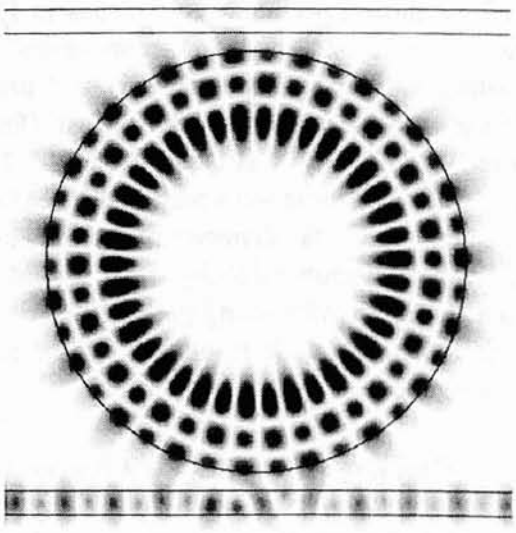
(a) Off-resonance: 193.4 THz.



(b) 1st-order radial WGM resonance: 189.2 THz.



(c) 2nd-order radial WGM resonance: 191.3 THz.



(d) 3rd-order radial WGM resonance: 187.8 THz.

Fig. 16.12 Grayscale visualizations of the FDTD-computed sinusoidal steady-state E_z distributions in the 5.0-μm-diameter circular disk resonator for monochromatic excitations at four different frequencies. *Source:* Hagness et al., *J. Lightwave Technology*, 1997, pp. 2154–2165, © 1997 IEEE.

16.5.2 Suppression of Higher-Order Radial Whispering-Gallery Modes

As demonstrated above, circular disk resonators supporting whispering-gallery modes have less sidewall scattering loss and higher Q than single-mode circular rings. However, circular disks have multiple sets of resonances due to the presence of fundamental and higher-order radial modes. Suppression of the higher-order modes is necessary to permit using the disks as single-mode laser sources, or as WDM devices having low crosstalk across a wide spectrum.

This section reviews two methods for suppressing the higher-order radial modes of circular disk resonators [31, 37]: (1) etching out the center of the disk, and (2) choosing the width of the adjacent waveguide to minimize the coupling between its guided mode and the higher-order modes of the disk. In each case, the resonances are manipulated by taking advantage of the spatial characteristics of the whispering-gallery modes of the disk.

Etching Out the Center of the Disk

A concentric circular hole located in the center of the circular disk can be etched out to create larger losses for the higher-order whispering-gallery modes that propagate in that region, thereby suppressing these modes. For the 5-μm circular disk resonator, the two limiting cases of hole diameter are $d_{\text{hole}} = 0$ (yielding the solid disk considered above), and $d_{\text{hole}} = 4.4$ μm (yielding the narrow circular waveguide ring considered in Section 16.4). Note that only the microcavity geometry is perturbed; the air gap between the laterally coupling waveguides and the microcavity remains constant.

FDTD studies have shown that, as d_{hole} increases from zero to 3.5 μm, the weak $m = 19$, 20, 21, and 22 third-order-mode resonances in Fig. 16.11 are completely suppressed, while the stronger $m = 23$, 24, and 25 second-order resonances are shifted upward in frequency. Proceeding further, as d_{hole} increases to 4 μm, there is substantial weakening and frequency upshifting of the second-order resonances. Here, the fundamental-mode resonances are slightly upshifted and their Q's slightly lowered. Finally, as d_{hole} increases to 4.4 μm, only the fundamental-mode resonances previously seen in Fig. 16.9 remain.

In summary, as the diameter of the etched-out hole within the circular disk increases, each set of resonances monotonically shifts toward higher frequencies. A smooth transition occurs between the multi-whispering-gallery-mode resonances of the solid 5.0-μm-diameter circular disk and the single-mode resonances of the 5.0-μm-diameter circular ring formed by 0.3-μm-wide waveguide.

Optimally Choosing the Width of the Adjacent Waveguide

We can minimize coupling to the higher-order whispering-gallery modes of the circular disk by properly choosing the width of the adjacent waveguide. Consider the FDTD-computed transmittance spectra for the 5.0-μm-diameter circular disk resonator coupled laterally to two straight adjacent waveguides across air gaps of 200 nm. As the width of the adjacent waveguides increases from 0.2 μm, the radial whispering-gallery-mode resonances of the third order and higher weaken and disappear. At the optimum waveguide width of approximately 0.38 μm, the second-order resonance verges on extinction, as seen in Fig. 16.13. Now, the air gap can be adjusted to give the best fundamental-mode characteristics, namely, low on-resonance transmission and high Q.

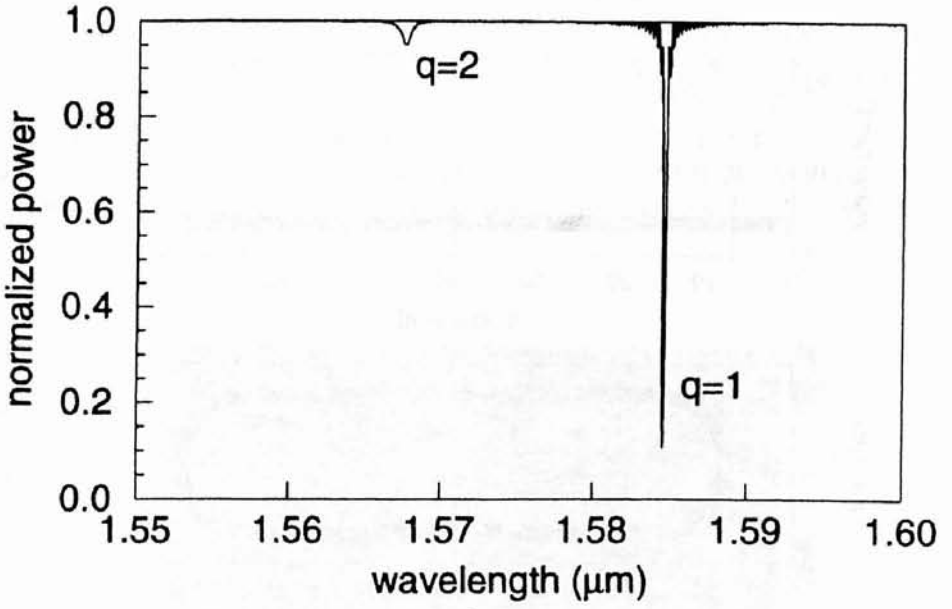


Fig. 16.13 FDTD-computed transmittance of a 5.0-μm-diameter circular solid disk laterally coupled to an adjacent 0.38-μm-wide waveguide across a 200-nm air gap. Nearly single-mode operation is observed. *Adapted from:* Hagness et al., *Proc. IEEE LEOS Annual Meeting*, 1997, pp. 160–161, © 1997 IEEE.

16.6 VERTICALLY COUPLED RACETRACK

The previous two sections discussed two-dimensional FDTD modeling studies of micron-scale photonic resonators having the shape of a circular ring, an elongated racetrack, and a solid circular disk. However, the very nature of two-dimensional modeling prevents simulation of the important class of guided-wave photonic structures wherein the waveguide-to-resonator coupling occurs in the vertical direction, rather than sideways in the plane of the resonator.

Reference [38] reported the formulation and results of what may be the first fully three-dimensional FDTD model of vertical waveguide coupling and subsequent optical beam confinement in a photonic racetrack. Referring to Figs. 16.14 and 16.15, the geometry considered in this paper consists of a straight bus waveguide located 1.343 μm below and parallel to one of the straight sections of a photonic racetrack that spans 64.0 × 17.6 μm.

Fig. 16.14 shows the x–y cross section planes containing the bus waveguide and the racetrack, and Fig. 16.15 shows the y–z cross section plane at P_2, the center of the waveguide coupling region. The bus waveguide consists of a 2.015 × 0.775-μm rectangular cross section channel of refractive index $n = 3.33$ embedded within a uniform substrate of index $n = 3.17$. The racetrack consists of a 2.015 × 0.775-μm rectangular cross section channel of refractive index $n = 3.33$ sandwiched above and below by rectangular cross section channels of refractive index $n = 3.17$, all embedded within a uniform superstrate of index $n = 1.5$.

By use of the bootstrapping technique, the straight bus waveguide is excited in its fundamental TE mode with a numerical Gaussian pulsed source having a center free-space wavelength $\lambda_0 = 2.32$ μm. Because the bootstrapping runs indicate 3- and 1.5-μm λ_0 cutoffs for the fundamental and the first higher-order modes, respectively, the source spectrum is adjusted to span 2.17 μm < λ_0 < 2.47 μm (FWHM) to avoid multimoding in the straight waveguide.

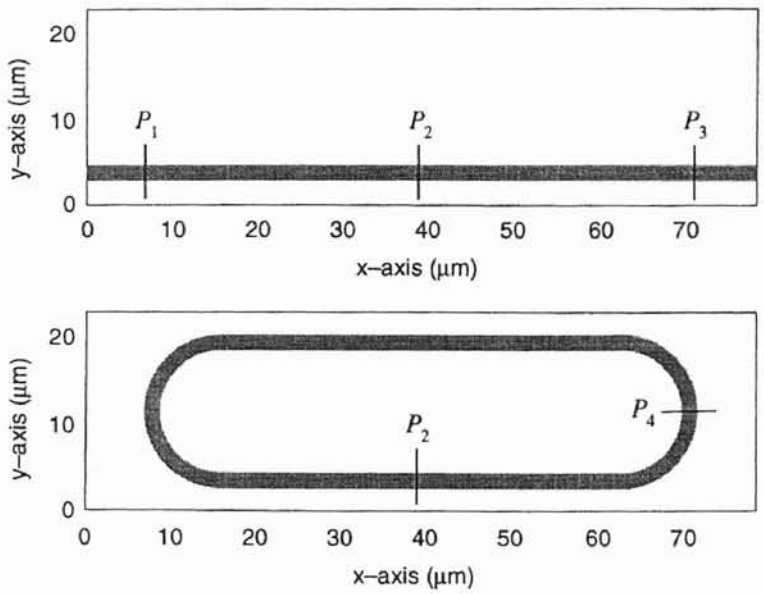


Fig. 16.14 Horizontal (x–y) cuts through the three-dimensional FDTD modeling geometry: (a) through the bus waveguide beneath the racetrack; (b) through the racetrack. Vertical transverse-field observation cut-planes P_1, P_2, P_3, and P_4 are also shown. *Source:* Greene and Taflove, *Optics Letters*, 2003, pp. 1733–1735.

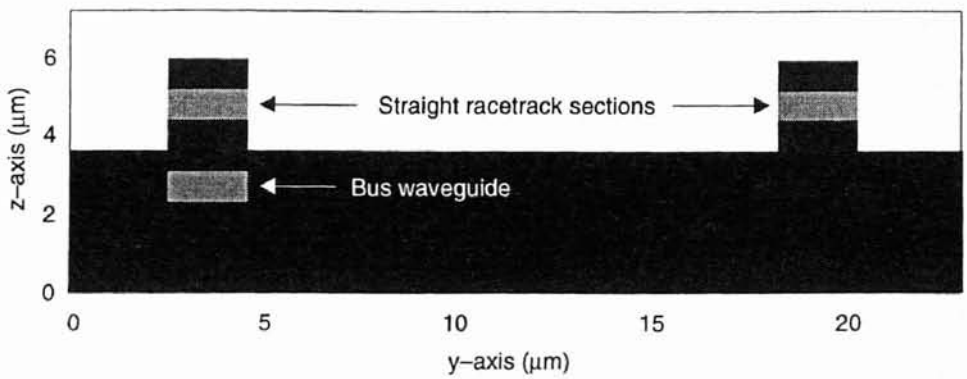


Fig. 16.15 Vertical (y–z) cross section through the three-dimensional FDTD modeling geometry at P_2 in the center of the waveguide coupling region. *Source:* Greene and Taflove, *Optics Letters*, 2003, pp. 1733–1735.

The FDTD model is implemented on a uniform $1500 \times 440 \times 138$-cell Cartesian cubic space lattice of resolution 51.7 nm ($\lambda_d/13.5$ within the waveguide at the center of the excitation spectrum). A 12-cell uniaxial perfectly matched layer absorbing boundary condition is used to terminate the mesh with wave reflections below 0.01%.

Figs. 16.16 and 16.17 provide grayscale visualizations of FDTD-calculated electric fields and Poynting vectors which illustrate key aspects of the operation of the racetrack. In Fig. 16.16, the absolute value of E_y within the racetrack is shown at three key points during the propagation of the optical pulse coupled from the bus waveguide: (1) to the right of P_2 immediately after the initial coupling, showing the establishment of the fundamental propagating mode; (2) approaching P_4 in the right curved section, showing the onset of multimoding; and (3) after exiting the right curved section, showing residual modal-distortion effects and pulse dispersion.

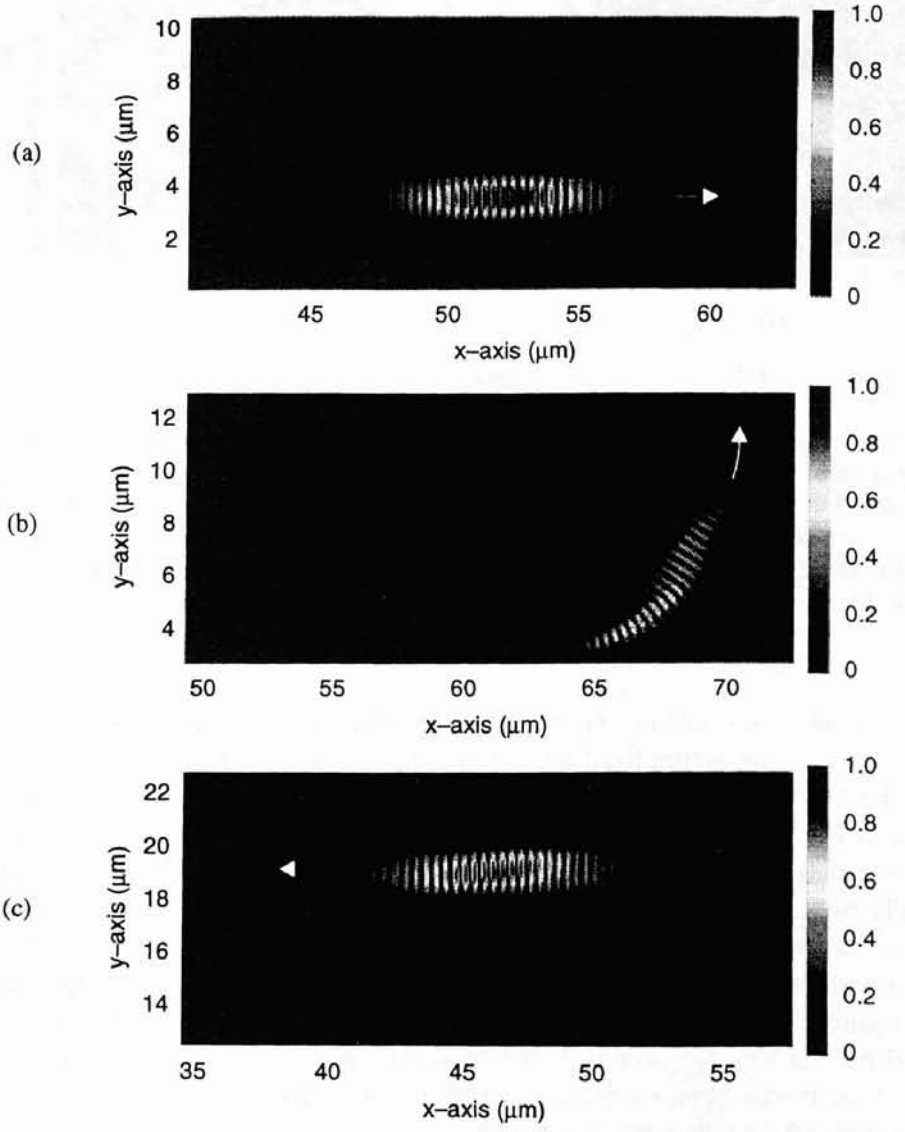


Fig. 16.16 Grayscale visualizations of the FDTD-calculated absolute value of E_y of the optical pulse within the racetrack: (a) $+x$-directed propagation to the right of P_2 of Fig. 16.14 after initial coupling from the bus waveguide; (b) in the right curved section approaching P_4 of Fig. 16.14, showing modal distortion; (c) $-x$-directed propagation after exiting the right curved section, showing residual modal-distortion effects. *Source:* Greene and Taflove, *Optics Letters*, 2003, pp. 1733-1735.

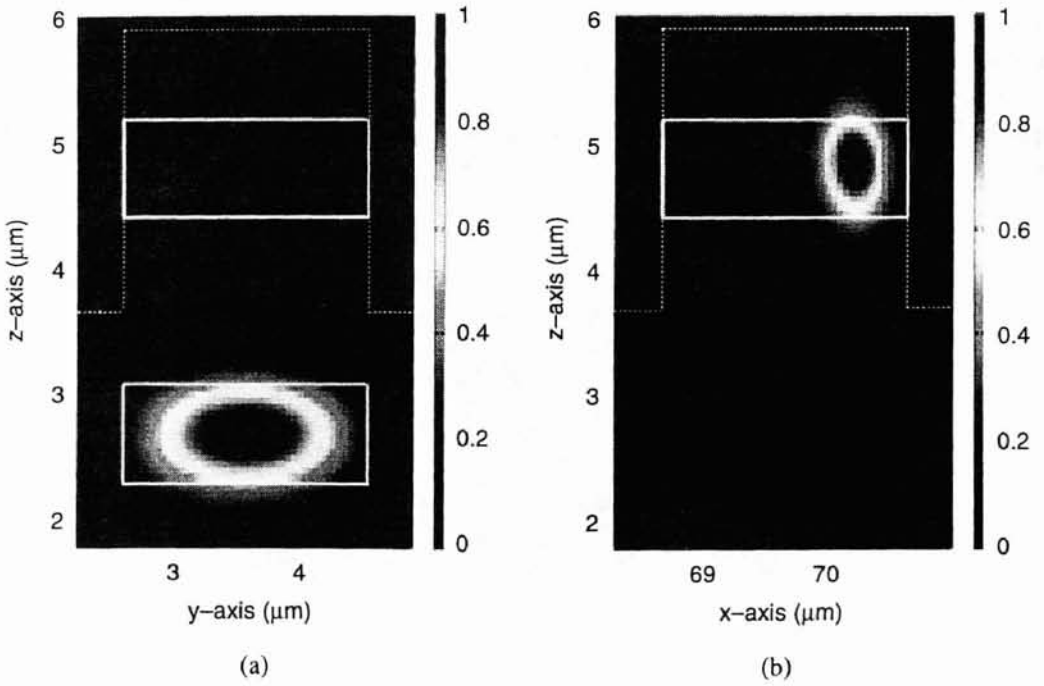


Fig. 16.17 Grayscale visualizations of the relative FDTD-calculated Poynting vector in the direction of propagation: (a) $+x$-component of the Poynting vector along the y–z cross section at P_2 of Fig. 16.14 at the instant of peak excitation within the bus waveguide; (b) $+y$-component of the Poynting vector along the x–z cross section at P_4 of Fig. 16.14 at the instant of peak excitation within the racetrack waveguide. *Source:* Greene and Taflove, *Optics Letters*, 2003, pp. 1733–1735.

Fig. 16.17 visualizes the relative Poynting vector in the direction of propagation. In this figure, the magnitudes are normalized by the maximum magnitude in each plot, and the boundaries of the bus and racetrack waveguides are shown as white rectangles. From Fig. 16.17(b) [see also Fig. 16.16(b)], we see that transient multimoding arises within the curved section of the racetrack. This multimoding spatially distorts the propagating optical pulse both by longitudinally broadening it and by shifting its local power flow toward the outer curved boundary of the racetrack waveguide. The latter effect increases radiation from the curved sections, since a portion of the near-surface power flow is launched as freely propagating energy by the surface roughness inherent in the fabrication of the racetrack waveguide. Interestingly, as seen from Fig. 16.16(c), the modal distribution within the racetrack returns to a closer approximation of the fundamental mode seen in Fig. 16.16(a), once it leaves the curved section, albeit with some residual distortion and broadening.

One of the advantages of vertical coupling to rings, racetracks, or disks is that this type of coupling exhibits reduced sensitivity to dimensional tolerances in the fabrication process relative to the sensitivity of the lateral coupling considered previously (illustrated by Fig. 16.8, which showed lateral-coupling variations over an approximate 2:1 range for gap-width changes of only 40 nm). In turn, increased interest in vertically coupled photonic structures is causing fully three-dimensional FDTD modeling to be more seriously considered for detailed simulations.

16.7 INTRODUCTION TO DISTRIBUTED BRAGG REFLECTOR DEVICES

Distributed Bragg reflector (DBR) mirrors are comprised of essentially one-dimensional stacks of pairs of dielectric layers having contrasting (high-low) refractive indices n_H and n_L. By choosing the layer thicknesses, the values of n_H and n_L, and the number of high-low-index pairs, we can adjust the reflection characteristics of the composite DBR structure. Qualitatively, light is reflected by a DBR mirror due to multiple internal wave reflections within its periodic structure. Positioning two DBR mirrors in close proximity yields a simple Fabry-Perot cavity.

An important application of DBR mirrors involves the fabrication of *vertical-cavity surface-emitting lasers* (VCSELs). These have attracted much attention because of their desirable lasing characteristics: low power consumption, high output power, and single-mode operation. The low lasing threshold of these devices can be further improved by using a periodic-gain configuration within the cavity. By placing thin gain segments along the E-field standing-wave maxima, the longitudinal confinement factor is maximized, thereby reducing the material gain threshold [39].

16.8 APPLICATION TO VERTICAL-CAVITY SURFACE-EMITTING LASERS

This section reviews the application of the classical (nonquantum) FDTD gain model described in Chapter 9, Section 9.7 to model VCSELs having *uniform-gain* and *periodic-gain structures* (UGS and PGS) [26]. Fig. 16.18 illustrates the microcavity geometries of interest.

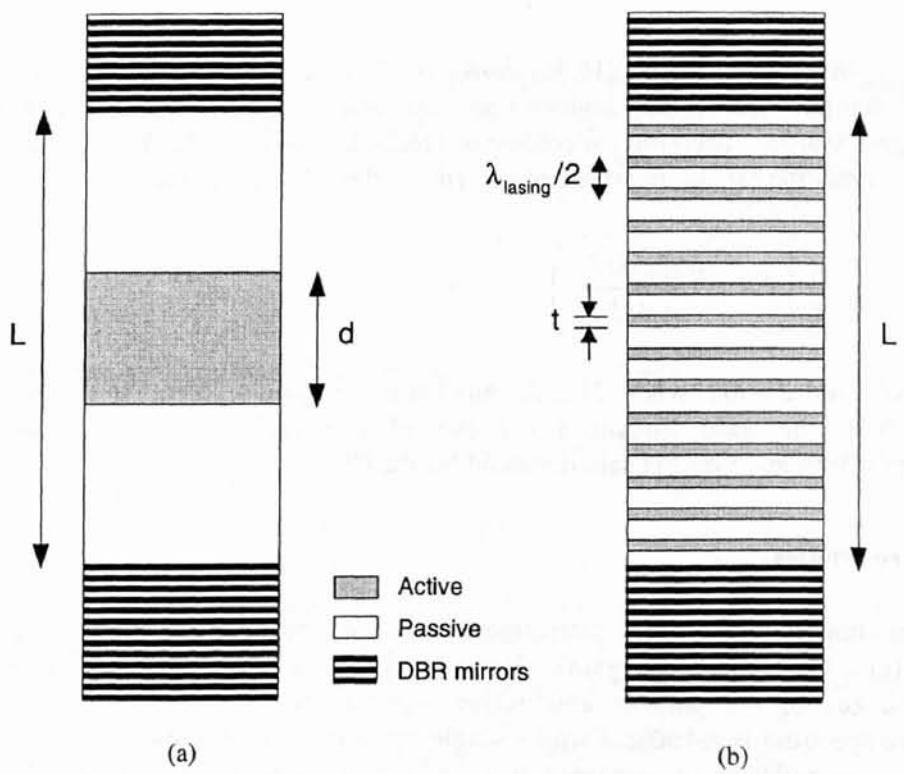


Fig. 16.18 Schematic diagrams of VCSELs: (a) uniform gain within the cavity; (b) periodic gain within the cavity. *Source:* Hagness et al., *Radio Science*, 1996, pp. 931–941.

The VCSELs of Fig. 16.18 can be modeled as one-dimensional structures since their lateral dimensions are large relative to λ [39]. Each is assumed to be grown on a GaAs substrate ($n = 3.59$), with DBR mirrors comprised of pairs of alternating layers of AlAs ($n_L = 2.971$) and $Al_{0.2}Ga_{0.8}As$ ($n_H = 3.452$) having a Bragg wavelength $\lambda_B = 0.87\ \mu m$. The top DBR mirror has 9 pairs, while the bottom DBR mirror has 12.5 pairs. Both DBR mirrors are designed to have a peak reflectivity of 92.8%. The optical path length of each cavity is assumed to be $27\lambda_B/2$.

For the uniform-gain case, shown in Fig. 16.18(a), the cavity consists of a solid, active GaAs region ($n = 3.59$) of length d surrounded on each side by a passive GaAs region (also $n = 3.59$). The fill factor d/L is chosen to be 0.5. For the periodic-gain case, shown in Fig. 16.18(b), the cavity consists of thin active GaAs segments, each with a thickness t separated by passive GaAs segments. The spacing between the gain segments is equal to one-half the lasing wavelength λ_{lasing}, and the total thickness of the gain segments is d. These design parameters permit comparison of the FDTD results with the analysis presented in [39].

Following [39], for a cavity with no loss, the material gain required at threshold is given by

$$\alpha_{th} = \ln R/2\Gamma L \tag{16.2}$$

where R is the geometric mean mirror reflectivity and Γ is the longitudinal confinement factor. For the UGS, Γ is the fill factor: $\Gamma_{UGS} = d/L$. Assuming a uniform standing-wave pattern and ideal mirror reflectivities, the longitudinal confinement factor for the PGS is approximated as

$$\Gamma_{PGS} \cong \frac{d}{L}\left[1 + \frac{\sin(2\pi t/\lambda)}{2\pi t/\lambda}\right] \tag{16.3a}$$

where $\lambda = \lambda_{lasing}/n$. When $t = \lambda/2$, (16.3a) shows that Γ for the PGS reduces to that of the UGS. When $t \to 0$, that is, when the gain segments are extremely thin, Γ is maximized at twice that of the uniform-gain system. Therefore, according to (16.2), the gain threshold can be reduced by as much as 50%. Equation (16.3a) is expressed in terms of the fill factor as follows:

$$\Gamma_{PGS} \cong \frac{d}{L}\left[1 + \frac{\sin(\pi d/L)}{\pi d/L}\right] \tag{16.3b}$$

since $L = M\lambda/2$ and $d = Mt$, where M is the number of thin gain segments in the PGS. For this case ($d/L = 0.5$), the gain threshold for the PGS according to (16.2) and (16.3b) is approximately 39% less than the gain threshold for the UGS.

16.8.1 Passive Studies

The FDTD method is first used to determine the passive "cold cavity" characteristics of the microcavity (i.e., for zero optical gain). Here, the UGS and PGS are identical because the refractive indexes of the passive and active regions are the same. For each mirror, the reflectivity spectrum is calculated with a single run using a short-pulse excitation and DFTs of the incident and reflected time-domain data. At a grid resolution of $\Delta x = \lambda_B/124n_H \cong 2$ nm, errors of less than three parts per 10,000 are obtained in the squared-amplitudes of the reflection coefficients at the reflectivity peaks of the bottom and top mirrors.

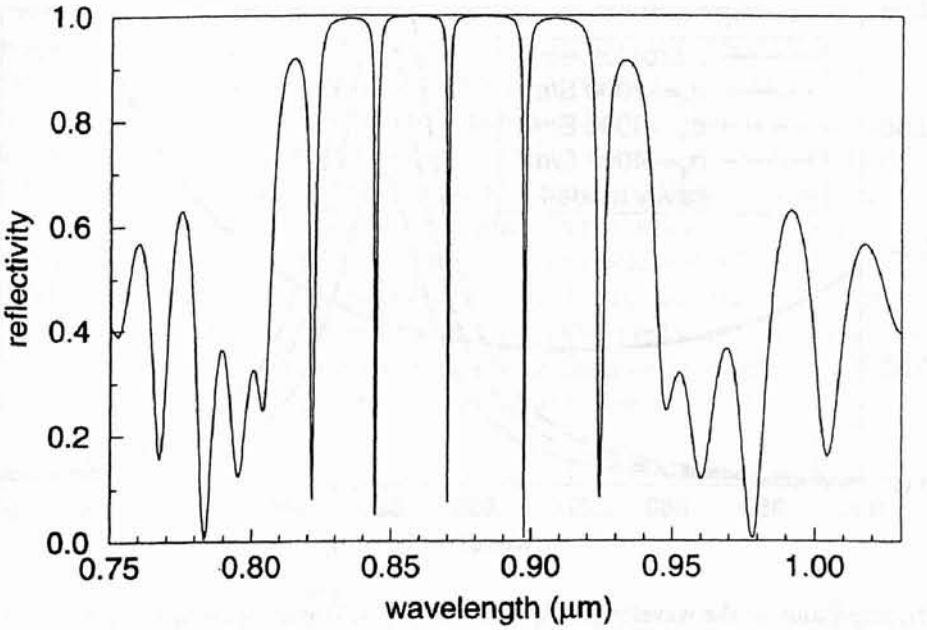


Fig. 16.19 FDTD-computed reflectivity spectrum of the each passive vertical microcavity illustrated in Fig. 16.18. *Source:* Hagness et al., *Radio Science*, 1996, pp. 931–941.

Fig. 16.19 graphs the FDTD-computed reflectivity spectrum of the entire microcavity. Five cold-cavity resonances at 821, 843, 870, 897, and 923 nm fall within the broad stopband. Knowledge of the wavelength spacing between these resonances, here approximately 27 nm, permits determination of the effective cavity length. This is greater than the physical length of the cavity because of field penetration into the DBR structures. In this case, the effective cavity length is approximately $34\lambda_B/2$, compared to the physical length of $27\lambda_B/2$.

16.8.2 Active Studies: Application of the Classical Gain Model

Fig. 16.20 graphs the superposition of the wavelength dependencies of: (1) the optical gain of GaAs for three different values of the baseline negative conductance σ_0 [as defined in the classical Lorentzian gain model of (9.112), with sample results shown in Fig. 9.8]; (2) the losses of the DBR mirrors of Fig. 16.18; and (3) the cold-cavity resonances obtained from Fig. 16.19 [40]. We see that the gain spectrum is centered at 890 nm, while the mirror losses are minimum at the Bragg wavelength of 870 nm. At the 870-nm cavity resonance, the mirror losses exceed the optical gain for each value of σ_0 shown, even though the mirror losses are at a minimum here. However, at the 897-nm cavity resonance, the optical gain exceeds the mirror losses for values of σ_0 more negative than approximately –3,000 S/m. Therefore, since the 897-nm longitudinal mode experiences the greatest net gain, lasing should occur at this wavelength.

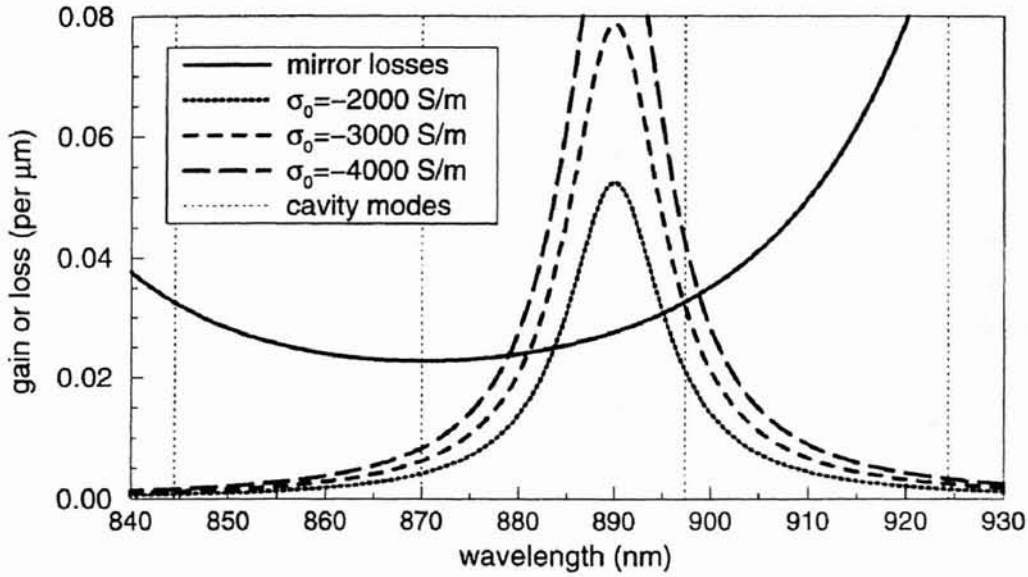


Fig. 16.20 Superposition of the wavelength dependencies of: (1) optical gain of GaAs for three different values of baseline negative conductance σ_0; (2) losses of the DBR mirrors of Fig. 16.18; and (3) cold-cavity resonances of Fig. 16.19 (shown as vertical dotted lines). *Source:* Hagness et al., *Computational Electromagnetics and Its Applications*, 1997, pp. 229–251.

The classical-gain FDTD algorithm is now used to determine the lasing wavelength and gain threshold for the VCSELs of Fig. 16.18, using a simulated noise current as an initial condition. Fig. 16.21 plots the FDTD-computed envelope of the *E*-field within the uniform-gain VCSEL at the sinusoidal steady state.

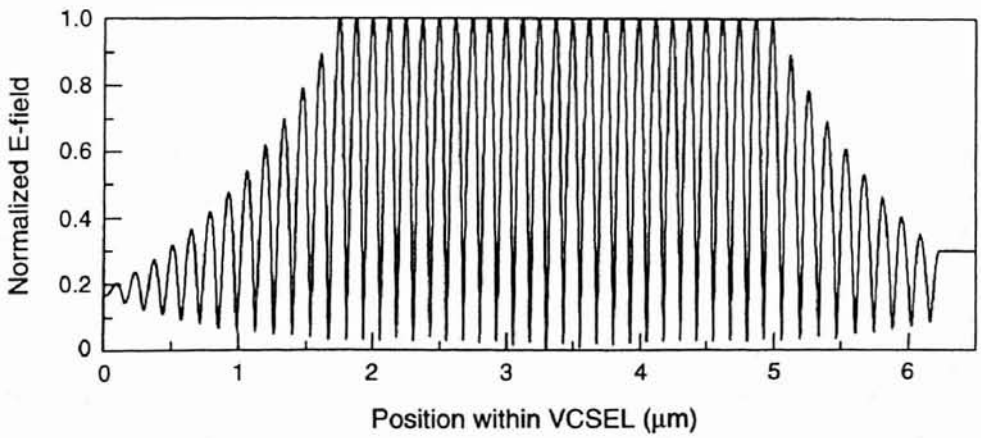


Fig. 16.21 FDTD-computed normalized *E*-field envelope within the uniform-gain-structure VCSEL of Fig. 16.18 under lasing conditions at the sinusoidal steady state. *Adapted from:* Hagness et al., *Radio Science*, 1996, pp. 931–941.

In Fig. 16.21, the decaying fields on either side of the cavity illustrate the penetration of the mode into the passive DBR mirrors. This simulation yields a lasing wavelength of just under 900 nm, as expected from the discussion of Fig. 16.20. The standing-wave maxima within the cavity are spaced periodically by an optical path length of $\lambda_{lasing}/2$. These maxima determine the optimum locations of the thin gain segments within the periodic-gain VCSEL.

For each lasing structure, the simulation is performed for several different values of σ_0 and the output intensity is recorded, as shown in Fig. 16.22. A comparison of the FDTD-computed UGS and PGS data sets reveals the following advantages of the PGS: (1) a substantial reduction in the gain threshold for lasing, and (2) a higher output intensity for a given gain level. Specifically, extrapolating the data of Fig. 16.22 to zero output yields a lasing gain threshold $\sigma_{0,th} \cong -3{,}260$ S/m for the UGS (as expected from Fig. 16.20), but only $-1{,}790$ S/m for the PGS, approximately 45% less. This compares to a 38.5% threshold reduction predicted by the analysis of (16.2) and (16.3). The FDTD calculation is a more accurate measure of the lasing threshold reduction, since it accounts for the fact that the DBR mirrors are not perfect reflectors.

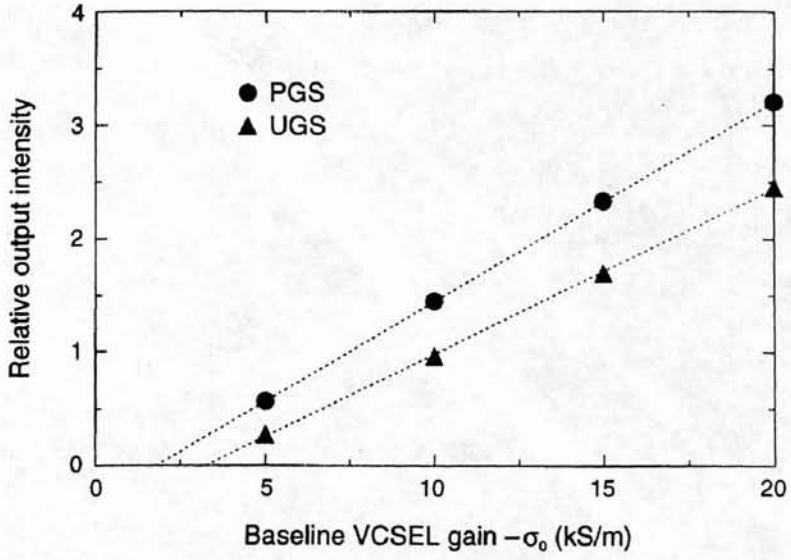


Fig. 16.22 FDTD-computed sinusoidal steady-state output intensity versus gain for the uniform-gain-structure and periodic-gain-structure VCSELs of Fig. 16.18. *Source:* Hagness et al., *Radio Science*, 1996, pp. 931–941.

16.8.3 Application of a New Semiclassical Gain Model

The reader is referred to Chapter 9, Section 9.8 for the FDTD algorithm and sample results for a new semiclassical formulation [28, 29] for modeling lasing dynamics, which treats the atom quantum mechanically and the electromagnetic wave classically. This algorithm models four quantized electron energy levels for each of two interacting electrons. Electron transitions between energy levels are governed by four coupled differential (rate) equations and the Pauli Exclusion Principle. In addition, there are two oscillator-type differential equations for temporal evolution of the induced polarization. This provides a robust, self-consistent treatment of the dynamics of the four-level atomic system and the instantaneous ambient optical electromagnetic field, regardless of the complexity of the material geometry that confines this field.

16.9 QUASI-ONE-DIMENSIONAL DBR STRUCTURES

An important class of photonic structures is based on an index-contrast semiconductor waveguide that confines light transversely, and a quasi-one-dimensional DBR structure that confines light longitudinally. Using FDTD, numerical investigations have been carried out for a variety of such structures, including waveguides with holes or notches and multilayer dielectric stacks [26, 40, 41–45]. These simulations provide transmission and reflection characteristics, radiation losses, field patterns, and cavity Q.

In this section, we discuss FDTD modeling of one such structure, the "air-bridge" [41] or "photonic-wire" [42] microcavity shown in the scanning electron microscope image of Fig. 16.23. Here, two linear arrays of etched air holes along an air-suspended 0.45×0.79-μm InGaAsP/InGaAs waveguide bound the microcavity.

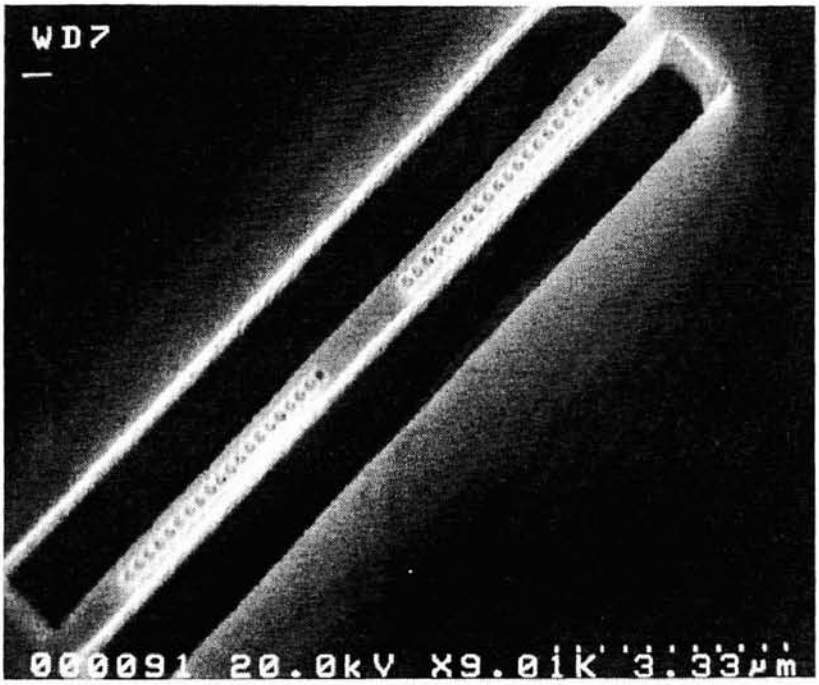


Fig. 16.23 Scanning electron microscope image of an InGaAsP/InGaAs suspended photonic wire microcavity. Air holes etched in the waveguide form quasi-one-dimensional DBR structures. *Source:* Zhang et al., *IEEE Photonics Technology Letters*, 1996, pp. 491–493, © 1996 IEEE.

Fig. 16.24 is a schematic diagram of a quasi-one-dimensional DBR reflector formed by a six-hole array, and a microcavity formed by two six-hole arrays at either end of a photonic wire. The specific geometry considered here is a 0.3×0.79-μm semiconductor waveguide ($n = 3.3$) surrounded by air ($n = 1$). To permit fine resolution (12.5 nm), simulations are performed in two dimensions (the horizontal plane of the structures) for the field components E_z, H_x, and H_y.

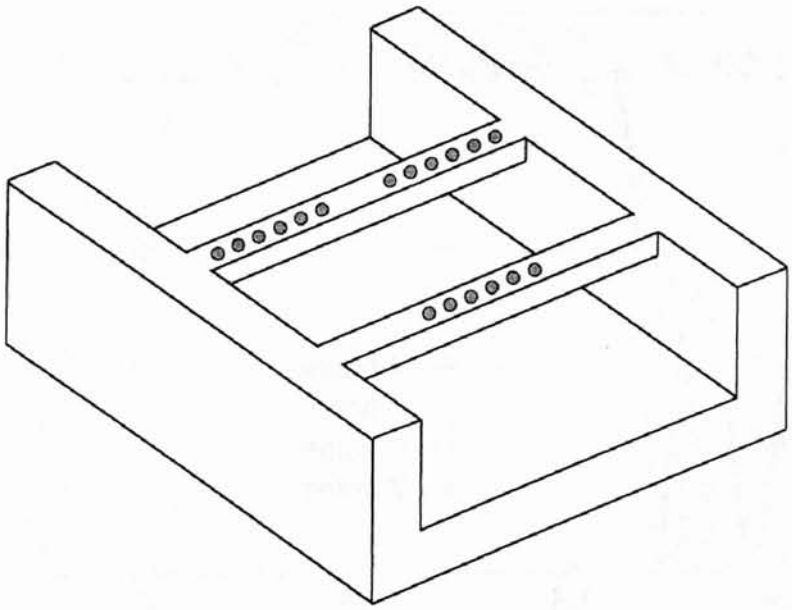


Fig. 16.24 Schematic diagram of a photonic wire mirror and the corresponding microcavity. *Adapted from*: Zhang et al., *IEEE Photonics Technology Letters*, 1996, pp. 491–493, © 1996 IEEE.

In the two-dimensional FDTD model, the finite extent of each optical waveguide of Fig. 16.24 in the vertical direction is accounted for by using the theoretical effective index of propagation in the actual three-dimensional waveguide as the bulk material index in the model. All waveguide and grid outer boundaries are terminated using the PML ABC. To obtain broadband spectral information from a single computer run, a 30-fs Gaussian pulse modulating a $\lambda = 1.5$ μm carrier is excited at one end of the waveguide using a fundamental-mode transverse source profile. Reflection and transmission spectra for the quasi-one-dimensional DBR structures are obtained from the FDTD data by taking the appropriate DFTs.

For such DBR structures, the center-frequency and width of the reflectivity spectrum are governed by the size, shape, number, and spacing of the etched holes. Fig. 16.25 graphs as a function of wavelength the FDTD-computed reflectivity of DBRs of the type shown in Fig. 16.24 formed by arrays of four, five, six, and seven rectangular air holes (each 0.25×0.1 μm). Here, the center-to-center spatial periodicity of the holes along the waveguide is 0.325 μm. As the number of holes increases, we see that the edges of the stopband sharpen, and the peak reflectivity increases. With seven holes, the DBR provides a reflectivity bandwidth of approximately 500 nm (about 33% of the carrier wavelength), and a peak reflectivity that approaches 99%.

Additional FDTD studies show that the shape of the holes impacts the reflectivity spectrum. Specifically, the bandwidth and peak reflectivity of a quasi-one-dimensional DBR of the type shown in Fig. 16.24 that uses elliptical holes is reduced relative to DBRs using rectangular holes. The bandwidth of the circular-hole DBR structure is similar to that of the rectangular-hole DBR, but is shifted towards shorter wavelengths.

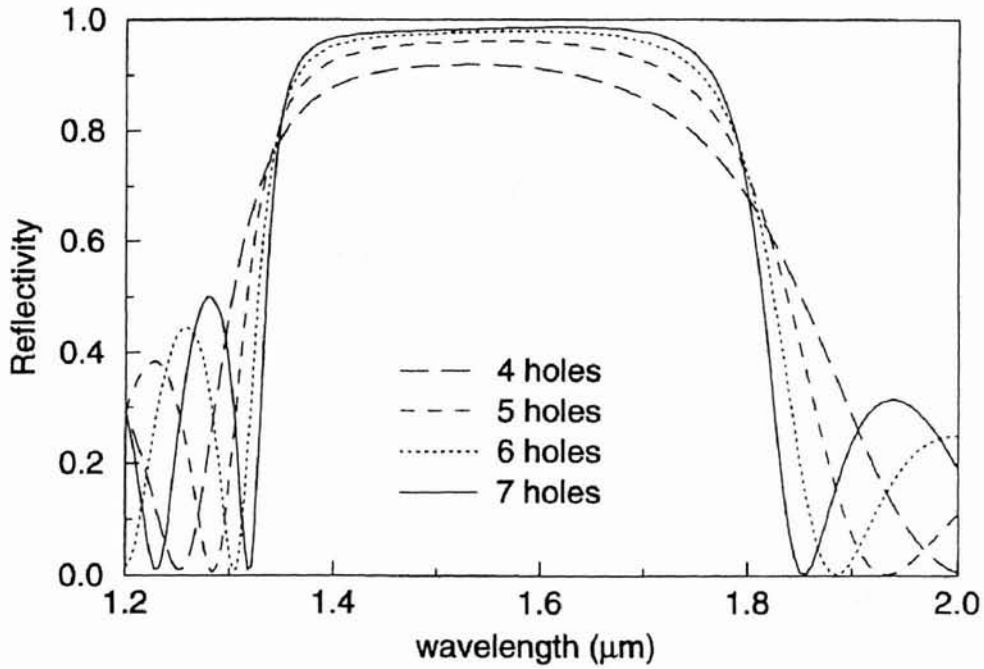


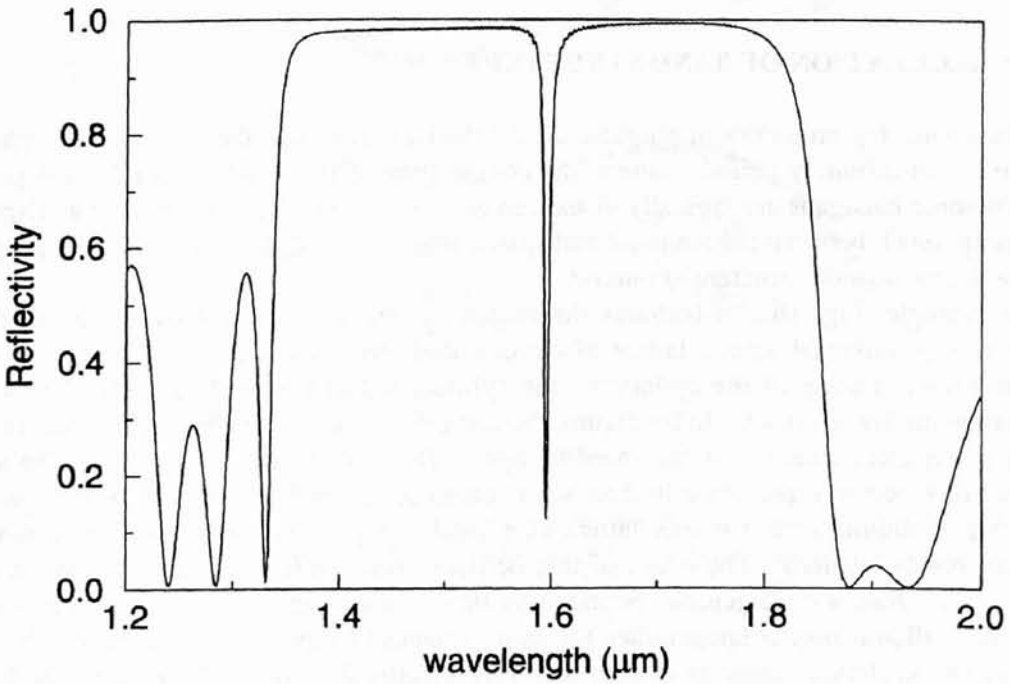
Fig. 16.25 FDTD-computed reflectivity spectra for quasi-one-dimensional DBR structures formed by arrays of rectangular air holes etched along a strongly guiding semiconductor waveguide.

Finally, we consider a microcavity of the type shown in Fig. 16.24 that is formed by an array of 10 rectangular holes with a 0.6625-μm-long defect introduced in the center. Fig. 16.26 plots the FDTD-computed reflectivity spectrum of this microcavity obtained from the time-domain data using a high-resolution DFT. A single resonant mode is seen within the stopband at $\lambda = 1.595$ μm. From the width and location of the resonance, the cavity Q is estimated to be about 300 [26]. By changing the length of the defect, the number and location of resonant modes within the stopband can be customized.

16.10 INTRODUCTION TO PHOTONIC CRYSTALS

In much the same way that the atomic lattice of a semiconductor establishes an electronic bandgap between conduction and valence bands, a periodic distribution of materials of disparate optical properties can create a *photonic bandgap* [46–49]. Within such a material, termed a *photonic crystal*, electromagnetic waves with frequencies within the bandgap cannot propagate. By itself, this behavior is extremely interesting, offering the potential for high reflectivity over a wide range of incident angles and wavelengths, or for useful dispersion properties at the edges of the bandgaps. However, it is when defects are introduced into an otherwise perfect lattice that suddenly the field of photonic crystals offers a host of new opportunities for manipulating light in ways never before considered. To date, researchers have used photonic crystals to bend, combine, filter, modulate, route, self-collimate, slow, split, switch, and trap light; to enhance nonlinearities; and to manipulate spontaneous emission [50]. A very incomplete list of applications includes waveguides, filters, lasers, modulators, and sensors.

(a) Geometry (approximately to scale).



(b) Calculated reflectivity spectrum.

Fig. 16.26 FDTD simulation of a photonic microcavity formed by introducing a 0.6625-μm-long defect at the center of a quasi-one-dimensional DBR structure formed by an array of 10 rectangular air holes etched along a strongly guiding semiconductor waveguide.

In many ways, photonic crystals and the FDTD method seem uniquely suited for each other. For instance, a photonic bandgap of noticeable extent requires a large index contrast, which is typically implemented by combining two or more widely different materials (e.g., air and a semiconductor such as silicon). However, such sharp discontinuities in the index of refraction would seem to mandate a modeling technique based directly on Maxwell's equations, rather than on the wave equation, since in general the latter is accurate only when the index varies slowly along the propagation direction. In turn, given the appropriate boundary conditions, FDTD can model either the properties of an underlying photonic crystal lattice, the performance of a device engineered by appropriate defect placement, or even the effects of fabrication imperfections. In addition, the current emphasis inherent in photonic crystal structures on enabling functionality in compact areas or volumes implies that many important problems can be addressed without having to resort to inordinately large FDTD simulations.

In the next six sections, we shall review the use of FDTD for modeling photonic crystal structures and devices. While the emphasis is on optical wavelengths, the physical principles involved work equally well at all wavelengths. However, the size of the periodic lattice that must be engineered tends to scale with the wavelength. This implies that most near-term applications will be for visible light, the near-infrared, millimeter waves, and microwaves, since outside these spectral bands, the required geometrical features are either too small to accurately fabricate or too large to be practical.

16.11 CALCULATION OF BAND STRUCTURE

Since the interesting properties of photonic crystal devices arise from the existence of a photonic bandgap for an infinitely periodic lattice, the computation of this bandgap is a logical place to start. Photonic bandgaps are typically visualized and investigated by computing the dispersion relationship, $\omega(k)$, between the temporal and spatial frequencies of the modes that can propagate in the particular periodic structure of interest.

For example, Fig. 16.27 illustrates the dispersion relationship or *band diagram* for an infinite two-dimensional square lattice of circular dielectric cylinders in air. Here, a is the center-to-center spacing of the cylinders; the cylinder radius equals $0.2a$; and the cylinder relative permittivity, ε_r, is 8.9. In the figure, the vertical axis represents the normalized temporal frequency $\omega a/2\pi c$, where c is the speed of light. The horizontal axis represents the spatial frequency or k-vector, represented in three segments to correspond to the edges of the triangular irreducible Brillouin zone for this lattice (the shaded region in the figure inset, with the symmetry points labeled). The edges of this Brillouin zone suffice because the modes in the interior of this zone are, in general, bounded by those on the periphery, and because once the irreducible Brillouin zone is known, then the entire extent of reciprocal space is known by either symmetry or translation. Only modes for which the electric field is parallel to the cylinders are shown (e.g., this is a TM_z band diagram)[1]. Note the bandgap that exists for waves with frequencies between $\omega a/2\pi c = 0.322$ and 0.442. To place this in context, in order to center this bandgap at $\lambda = 1.5$ μm, the cylinders would need to be 229.6 nm in diameter on 574-nm centers, and the bandgap would then run from 1,295 to 1,780 nm.

Fig. 16.27 actually shows the dispersion diagram computed in two different ways: the solid curves show the results of the plane-wave expansion method using the MIT Photonic Bands package [51, 52], while the black dots show the band diagram as computed by a two-dimensional FDTD model. To learn more about the plane-wave method and photonic crystal band diagrams in general, the reader is directed to [48, 49, 51]. In particular, the example shown here is described in detail on page 56 of [48]. We turn now to discussing how to use FDTD to compute photonic crystal band diagrams.

---

[1]We follow the FDTD and photonic crystal convention that the transverse nature of a two-dimensional simulation is dictated by the plane of the simulation; so having H_x and H_y in the plane and E_z out of the plane as in Fig. 16.27 is designated as a simulation that is transverse magnetic relative to the z-direction (TM_z). However, in the *same* two-dimensional geometry, any resulting guided-wave modes with these *same* field components would be designated as TE modes in the lexicon of standard waveguide theory, which takes the modal propagation direction as the reference. Therefore, the designations TE and TM must be used with considerable care.

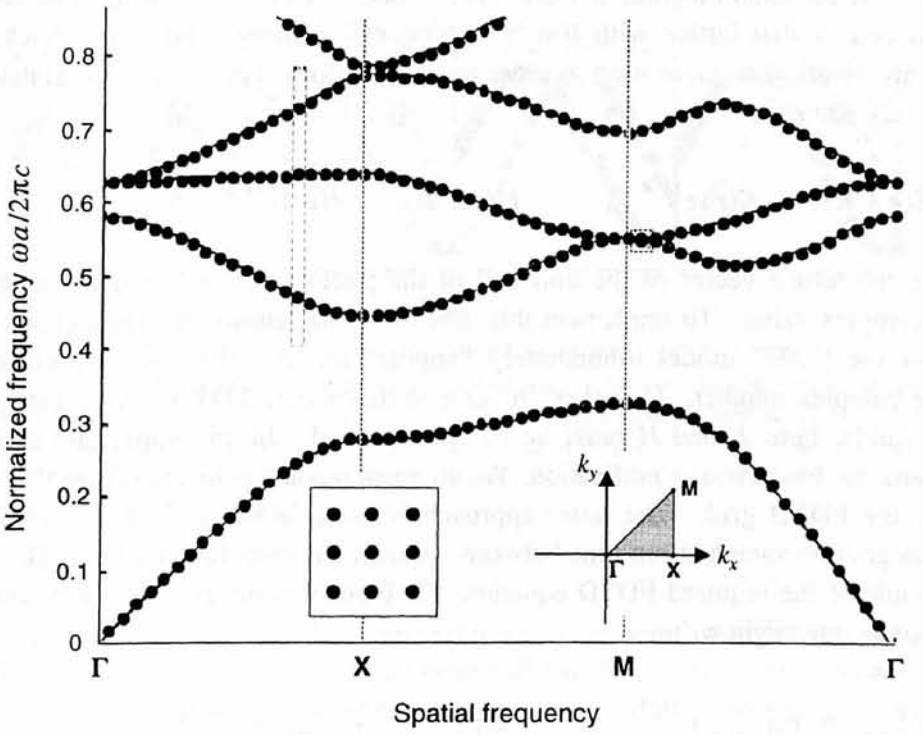


Fig. 16.27 Dispersion diagram for an infinite two-dimensional square lattice of high-index dielectric cylinders in air for TM_z polarization. Insets show the real-space and the reciprocal-space lattices. The normalized frequency is in units of $2\pi c/a$. The spatial frequency is normalized to the symmetry points $\mathbf{\Gamma}$ ($k_x = k_y = 0$), $\mathbf{X}$ ($k_x = \pi/a$, $k_y = 0$), and $\mathbf{M}$ ($k_x = k_y = \pi/a$). Two dotted boxes show regions of the dispersion diagram that are discussed in detail later: one is near the $\mathbf{X}$ point (Fig. 16.29), and the other is near the $\mathbf{M}$ point (Fig. 16.30).

16.11.1 The "Order-N" Method

The FDTD technique described here was introduced in [53] as the "Order-N" method. It was motivated by the fact that the computational burden of plane-wave expansion techniques in use at the time scaled as $O(N^3)$, where N denotes the number of grid cells in the simulation. Since an FDTD simulation of N grid cells scales as $O(N)$, at least in memory size, the advantages of FDTD could be significant. Subsequently, the Order-N method was extended to Green functions [54], nonorthogonal grids [54, 55], and metallic materials [56–58]. A parallel can also be drawn between the Order-N method and earlier work using FDTD to solve eigenvalue problems in microwave engineering [59]. A review of some of the features of the Order-N method and a brief discussion of its application to photonic crystals can be found in [60].

We note that, rather than the $O(N^3)$ referred to in earlier papers, the block-iterative plane-wave expansion techniques described in [51], and widely available through download [52], scale computationally as $O[pN\log(N)] + O(p^2N)$, where p denotes the number of bands to be evaluated. While this affects the tradeoff between using FDTD and plane-wave methods in terms of speed, particularly for two-dimensional simulations, there remain numerous useful and unique features in using FDTD to compute photonic crystal band diagrams.

To compute the band diagram of a photonic crystal lattice, we use (nearly) standard FDTD on the unit cell of that lattice with Bloch or Floquet boundary conditions. Such boundary conditions are identical to those used in order to source infinite plane waves at arbitrary oblique angles [61, 62], namely

$$\breve{\boldsymbol{E}}(\boldsymbol{r}+\boldsymbol{R}) = \breve{\boldsymbol{E}}(\boldsymbol{r})e^{j\boldsymbol{k}\cdot\boldsymbol{R}} \quad ; \qquad \breve{\boldsymbol{H}}(\boldsymbol{r}+\boldsymbol{R}) = \breve{\boldsymbol{H}}(\boldsymbol{r})e^{j\boldsymbol{k}\cdot\boldsymbol{R}} \tag{16.4a, b}$$

where $\boldsymbol{R}$ is the lattice vector of the unit cell of the photonic crystal, and the saucer symbol denotes a complex value. To implement this, one need only ensure that the fields that "leave" one side of the FDTD model immediately "appear" on the other side, multiplied by the appropriate complex number. However, in contrast to standard FDTD, which uses purely real E- and H-fields, here E and H must be complex-valued. In an appropriate programming language, this can be a trivial modification. An alternate method is to add yet another dimension or layer to the FDTD grid. The latter approach is more flexible, since it allows the same compiled program to switch at run time between standard and complex-valued FDTD, as needed. As an example of the required FDTD equations for Floquet boundary conditions along x in a TM_z simulation, one might write

$$\breve{E}_z\big|_{i=0} = \breve{E}_z\big|_{i=i_{\max}} e^{jk_x L_x} \quad ; \qquad \breve{E}_z\big|_{i=i_{\max}+1} = \breve{E}_z\big|_{i=1} e^{-jk_x L_x} \tag{16.5a, b}$$

Here, for simplicity, we assume that the boundary field components are redundantly stored (perhaps for ease of parallelization), so that the photonic crystal unit cell of interest spans from FDTD grid cell 1 to $i_{\max}$, with cell 0 serving as a phase-adjusted copy of cell $i_{\max}$, and cell $i_{\max+1}$ serving in the same role for cell 1. Similar equations follow for the two H-components in the TM_z simulation (or all six field components in a three-dimensional simulation).

Note that these equations are the only point at which the "layer-1" components ($\breve{E}_z$, $\breve{H}_x$, $\breve{H}_y$) interact with the "layer-0" field components (E_z, H_x, H_y), and thus the only point at which complex numbers are considered. For the E-components, (16.5) is implemented after the normal E-updates from the H-components and the subsequent data exchange for parallelization, but before any H-updates. The H-components are phase-adjusted at a commensurate spot with respect to their normal updates.

With these Floquet or Bloch boundary conditions, one can consider the FDTD algorithm to be modeling wave propagation through an infinite expanse, which just happens to be tiled into identical squares or cubes, as shown schematically in Fig. 16.28. In this figure, we consider the simple case of a plane wave propagating from left to right through various "copies" of the unit cell. Two snapshots of the variation in the amplitude of one field component, say E_z, along the horizontal direction x are shown for two points slightly separated in time. Since the wavefront is constant along y, we are using the vertical direction to plot both the y-dimension of the cell and its monitor, as well as the amplitude of the wavefront snapshots. Here, the k value established by the $\pm x$ boundaries corresponds to a wavelength slightly longer than the unit cell width L_x. This wave undergoes slightly less than 2π of phase shift within the unit cell, but the boundary conditions add just the right amount of phase to get exactly to 2π. As a result, the wavefront appears identical in each unit cell, jumping in phase at each unit-cell boundary.

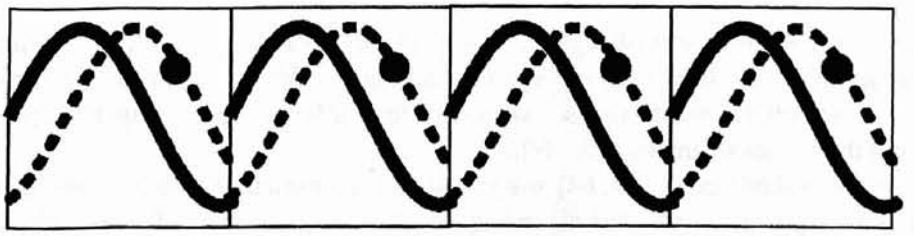


Fig. 16.28 Four "copies" of a square unit cell are shown, each containing its own copy of a monitor point (dot). A wavefront with wavevector k_1 is illustrated at two instants in time, showing that when the phase jump imparted by the Floquet boundary conditions completes a 2π phase excursion, every copy of the unit cell has the same snippet of waveform. Under these conditions, the field amplitude in every copy constructively interferes, resulting in a strong monitor signal at the associated temporal frequency ω_1.

If one considers a monitor point at some arbitrary point within each cell (represented by a small dot in Fig. 16.28), the amplitude at that point will oscillate in time with a frequency ω_1 and some arbitrary phase offset. In fact, every copy of this monitor point will see an identical oscillation with the same phase offset. However, for some other wavefront of different k_x value (and its own frequency ω_2), the phase excursion for each unit cell is not 2π, and a snapshot at any point in time would find a different snippet of the wavefront in each unit cell. As a result, while each monitor point indeed sees an amplitude variation at some frequency ω_2, the phase in each unit cell "copy" differs.

However, since there is only one unit cell in the actual FDTD simulation, there is really only one monitor. The field measured at this point is a superposition of all of the unit-cell "copies" shown in Fig. 16.28. The phase differences between "copies" causes ω_2 to die out through destructive interference, leaving only an oscillation at ω_1 (and at the ω values corresponding to all other modes for which the transit across the unit cell plus boundary conditions represents a multiple of 2π phase excursion). Note that this does not depend on any relationship between the values of ω_1 and ω_2. In addition, even without the spatial-superposition effect, the signal contribution at ω_2 at any one monitor "copy" at some time t_1 is still out-of-phase with the contribution at some later time t_2 because of the non-2π phase excursion. Thus, the Fourier transform of the time evolution of the fields at the monitor point combines both effects, resulting in a strong response or resonance at ω_1 and no response at ω_2. Hence, all of the modes (at various values of ω) that share the same underlying k vector have their frequencies represented in the Fourier spectrum, and can be computed with the same FDTD run.

There are, of course, several ways for this to fail to find all of the correct modes. If the monitor point we choose is located at or near a node where the fields associated with a mode go to zero, then the contribution at ω_1 can be so low as to be accidentally discarded. However, it is fairly straightforward to record the time evolution at several appropriately selected monitor points during the FDTD simulation, and then combine the spectra afterwards. This greatly reduces the probability of missing a mode. Unfortunately, such redundancy does not help if the original excitation of the FDTD unit cell fails to inject any energy into one of the modes we want to identify. The original Order-N paper suggested a linear combination of a small number of plane waves **G** as an initial condition, or:

$$\breve{\boldsymbol{H}}(\boldsymbol{r}) = \sum_{\mathbf{G}} \left[\hat{\boldsymbol{v}} \times (\boldsymbol{k} + \mathbf{G})\right] e^{j(\boldsymbol{k}+\mathbf{G})\cdot\boldsymbol{r}} \tag{16.6}$$

where $\hat{v}$ is a unit vector controlling the resulting polarization [53]. This choice of initial condition ensures that $\nabla \cdot \boldsymbol{H} = 0$ throughout the simulation. However, the user must still choose how many and which $\mathbf{G}$ values to use, which in turn affects the overlap between the initial conditions and the modes of interest [53, 60].

Alternatively, Sakoda et al. [63, 64] suggested that if a radiating dipole is introduced into the FDTD grid at a frequency ω near the eigenfrequency ω_d (equal to ω_1) of a resonant mode of the structure being simulated, then the field $\boldsymbol{E}(\boldsymbol{r})$ emitted by the dipole (as well as the radiated energy) should depend strongly on the resonant eigenmode $\boldsymbol{E}_d(\boldsymbol{r})$, expressed as

$$\boldsymbol{E} \simeq \frac{2\pi\omega_d \left[\boldsymbol{\gamma} \cdot \boldsymbol{E}_d^*(\boldsymbol{r}_0)\right] \boldsymbol{E}_d(\boldsymbol{r}) e^{-j\omega t}}{V(\omega - \omega_d + j\Gamma)} \tag{16.7}$$

where Γ is the decay rate of the mode (here effectively zero, since the medium is not lossy and there are no radiation losses in an infinite crystal); V is the volume of the unit cell; and $\boldsymbol{\gamma}$ is the dipole moment of the excitation located at position $\boldsymbol{r}_0$ [56, 58, 63, 64]. To map out the dispersion diagram with this, one then would scan along ω looking for the strongest response.

In contrast, we have found it simpler to combine both of these ideas by just "pinging" the FDTD model with a single delta function in both space and time. This has the advantage of exciting all temporal and spatial frequencies, so that the only way to fail is to be so unlucky as to "ping" at or near a node point. In the context of (16.7), we are attempting to excite all possible ω simultaneously without having $\boldsymbol{\gamma} \cdot \boldsymbol{E}_d^*(\boldsymbol{r}_0) \sim 0$ for any mode of interest to us. In practice, we can avoid this condition by ensuring that the source delta function does not sit at any symmetry point. However, prudence would suggest that several such source points be tried, especially when computing a new band structure for the first time, in order to ensure that the resulting band structure is indeed independent of the source placement.

Through the denominator of (16.7), we can be confident that only the portion of injected energy that is very close in frequency to each ω_d contributes significantly to the measured spectral peaks. In the context of the two-layer complex-valued FDTD formulation described above, the delta function appears on only one of the two layers, with energy eventually making its way to the other layer through the boundary conditions (except when $k = 0$). The field component chosen for the "ping" dictates the polarization of the modes that are computed, as dictated by the dot product between $\boldsymbol{\gamma}$ and $\boldsymbol{E}_d^*(\boldsymbol{r}_0)$ in (16.7). For instance, setting just E_z to 1 for one time-step should excite the TM modes, while H_z excites the TE modes. In three-dimensional simulations, all modes wherein the chosen field component is nonzero at the "ping" position are excited.

16.11.2 Frequency Resolution

There remains an unanswered question: How many FDTD time-steps are necessary in order to unambiguously identify the spectral peaks ω_1? Through the usual properties of the Fourier transform, the width of each spectral bin in the calculated frequency spectrum is inversely proportional to the total duration of the monitored field component's temporal response. This seems to imply that accurate determination of the resonance frequencies could require inordinately long FDTD simulations.

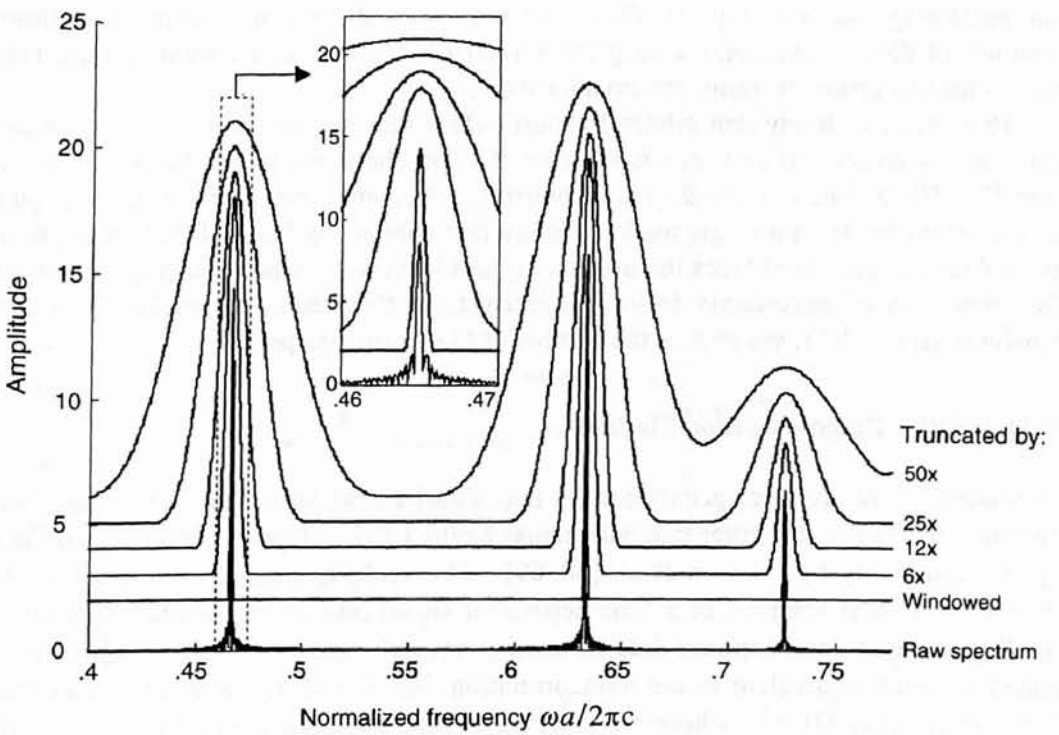


Fig. 16.29 Spectral peaks at $k_x = 0.8\pi$ for the infinite square lattice of dielectric cylinders, corresponding to the left-hand dotted box in Fig. 16.27. Successive curves show the effect of windowing the raw temporal data with a Blackman window before performing the Fourier transform, and the effects of various amounts of truncation (as if the FDTD simulations had been stopped earlier). Curves are vertically offset for visualization purposes. The inset displays only a small portion of the spectral range, showing the strong ripple in the raw, unwindowed spectrum.

As an additional complication, due to the inherently abrupt truncation of the ongoing temporal response, the raw spectrum tends to have a significant amount of ripple, as shown by the lowest trace in Fig. 16.29. Here, the spectral peaks corresponding to the particular k value of $k_x = 0.8\pi$ are shown, corresponding to the dotted box in Fig. 16.27 located to the left of the **X** point. This undesirable ripple tends to complicate the identification of spectral peaks, since the derivative of the spectrum goes through zero in many places other than at a peak.

However, it is important to keep in mind that the main interest here is the presence and position of the spectral peaks, not their width. Thus, multiplying the original temporal response data with a windowing function such as the Blackman window

$$f(t) = 0.42 - 0.5\cos\left(2\pi t/t_{\max}\right) + 0.08\cos\left(4\pi t/t_{\max}\right) \tag{16.8}$$

can reduce the ripples in the spectrum without significantly affecting the spectral peak position, as shown by the second trace in Fig. 16.29, labeled as "Windowed." In essence, the underlying delta functions we are trying to locate (representing a lossless mode of essentially infinite Q) are simply being convolved by a symmetric function. As a smaller and smaller portion of the temporal response captured during the FDTD run is used when performing the Fourier transform (as if we had truncated the run earlier), the resulting spectral peaks get wider and wider

(the remaining traces in Fig. 16.29, shown only with Blackman windowing). However, the centroids of these peaks, representing the normalized frequencies we want to extract to place as dots on the dispersion diagram, remain almost completely unchanged.

This seems to imply that relatively short FDTD runs are sufficient to construct an accurate dispersion diagram. In fact, this is the case if all of the peaks are isolated. However, we see from Fig. 16.29 that, at some degree of blurring, it becomes impossible to distinguish a closely set pair of modes from a single mode. We see that conducting longer FDTD runs provides finer spectral resolution and reduces the uncertainty in identifying all possible pairs of spectral peaks. This reduction of uncertainty follows directly from the frequency resolution of the Fourier transform, $\Delta f = 1/N\Delta t$, where N is the number of FDTD time-steps of size Δt.

16.11.3 Filter Diagonalization Method

Fortunately, there are more accurate ways than Fourier analysis to extract resonant frequencies. One such technique, the filter-diagonalization method [65–67] was introduced to the photonic crystal community by Johnson et al. [68, 69]. This technique recasts the problem of spectral analysis of a short segment of a time-dependent signal into an eigenvalue problem. A set of FDTD-calculated time-response data containing several resonances, $C(t) = \sum_k d_k \exp(-j\omega_k t)$, is treated as being equivalent to the autocorrelation function of an unknown dynamical system, $C(t) = \langle \Phi_0, \exp(-jt\hat{\Omega})\Phi_0 \rangle$, where $\langle \cdot, \cdot \rangle$ represents the complex symmetric inner product [66]. By choosing Φ_0 appropriately, the unknown operator $\hat{\Omega}$ can be completely described in terms of the original data $C(t)$, and thus can be diagonalized [66]. The advantage of this technique over Fourier analysis is that, rather than simply measuring the correlation of the signal against multiple sinusoids $\exp(-j\omega_k t)$ in multiple fixed-width spectral bins, a smaller number of spectral bins are precisely tuned in size and position to identify the complex frequencies where there is maximal correlation (e.g., the desired resonances). This allows the filter-diagonalization method to distinguish pairs of spectral peaks using much shorter-duration time signals than possible with Fourier analysis. Relative to older harmonic-inversion techniques, such as those based on Prony's method, filter diagonalization remains stable and efficient while requiring small matrices, even when an enormous number of resonances may exist outside of the spectral range of interest [65, 66].

Implementing the filter-diagonalization method in a Fourier basis involves following the checklist given on page 6761 of [66], with one minor addition described below. We first choose a matrix size J based on the desired frequency range $(\omega_{\min}, \omega_{\max})$, the signal extent N, and the time-step Δt. We then fill the basis vectors $\omega_{\min}\Delta t \le \phi_j \le \omega_{\max}\Delta t$ and the matrices $[\mathrm{U}]^{(p)}$ for $p = 0, 1, 2$; solve the generalized eigenvalue problem $[\mathrm{U}]^{(1)}\{\mathrm{B}_k\} = u_k[\mathrm{U}]^{(0)}\{\mathrm{B}_k\}$ for eigenvalues $u_k = \exp(-j\omega_k\Delta t)$ and eigenvectors $\{\mathrm{B}_k\}$; and then discard the spurious eigenvalues by checking them against $[\mathrm{U}]^{(2)}$. The minor addition is that the matrices $[\mathrm{U}]^{(p)}$ can be assembled much more efficiently by using (25) of [67], so that the various inner products between the raw time data $C(t)$ and the vectors derived from ϕ_j are performed only once, and are stored temporarily as G_j^p so they can be reused appropriately. (Note that while (29) of [67] describes how to iteratively derive $[\mathrm{U}]^{(1)}$ using G_j^0, it seems to be more efficient to iteratively compute G_j^1 from G_j^0 instead, since that can then lead to $[\mathrm{U}]^{(2)}$ as well.) Of course, an even easier approach to implementing the filter-diagonalization method is to simply download Johnson's "harminv" program from the Web [69]. In addition to the above checklist, the harminv program refines the eigenvalues by repeated iteration, converging smoothly to an accepted set of resonances.

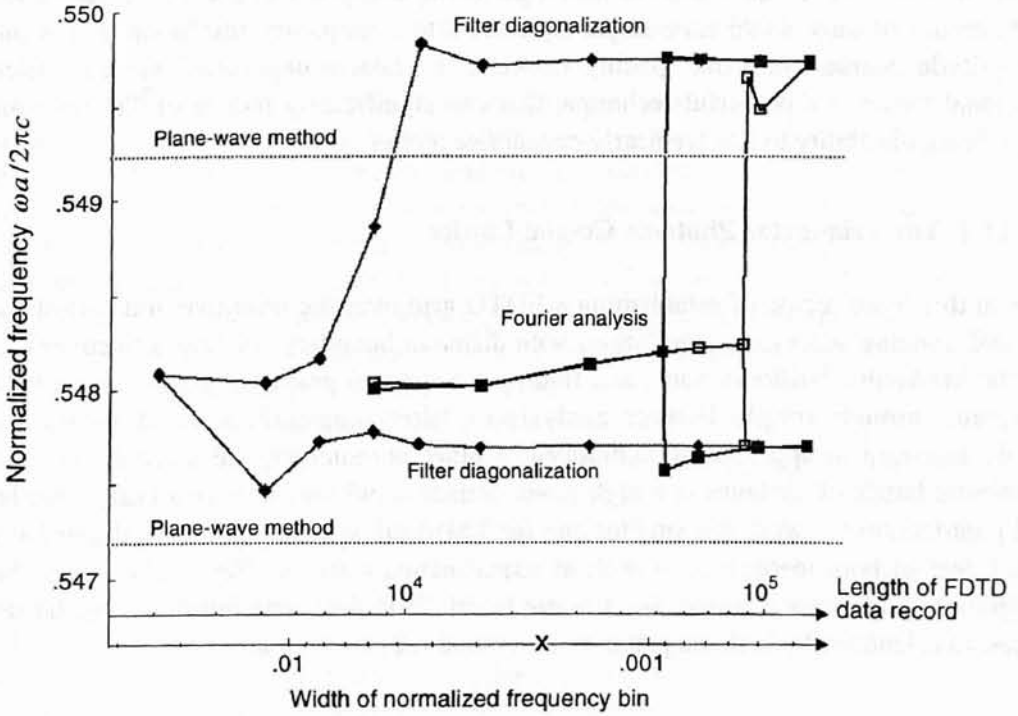


Fig. 16.30 Advantages of the filter-diagonalization method relative to Fourier analysis in using short FDTD datasets to identify the two nearly degenerate modes at $k_x = k_y = 0.95\pi$ for the infinite square lattice of dielectric cylinders, corresponding to the right-hand dotted box in Fig. 16.27. Filter-diagonalization method data: diamonds; Fourier data: open and filled squares; dotted horizontal lines: predictions of the plane-wave method.

Fig. 16.30 demonstrates the advantages of the filter-diagonalization method relative to Fourier analysis. Here, we plot as a function of the length of the FDTD data record the calculated resonant frequencies of the two nearly degenerate modes at $k_x = k_y = 0.95\pi$ for the infinite square lattice of dielectric cylinders (corresponding to the right-hand dotted box in Fig. 16.27, just to the right of the **M** point). A second horizontal axis is also shown in this figure to indicate the width of the normalized frequency bin corresponding to the FDTD dataset length. The diamonds show the results of the filter-diagonalization method, and the squares show the results of Fourier analysis (open squares—direct transformation of the FDTD data record; filled square—transformation after zero-padding the FDTD data to a uniform length of 131,072 time-steps). The two dotted horizontal lines show the resonant frequencies calculated using the plane-wave method [52] (128 plane waves with the meshing value set to estimate the local dielectric constant over 7×7 subsamples).

Fig. 16.30 shows that, by applying the filter-diagonalization method, we can separate the resonances of the two nearly degenerate modes using FDTD data records extending over 11,000 time-steps, only about one-fifth the record length required using Fourier analysis. This corresponds to a spectral resolution that is more than twice as coarse as the separation between the peaks of the degenerate modes, indicated by an x on the lower horizontal axis. A relative frequency difference of only 0.1% is noted in comparison with the plane-wave method.

In addition, the filter-diagonalization method continues to identify the presence of a double peak (although the frequencies are no longer being computed as accurately) down to an FDTD data record of only 4,096 time-steps, equivalent to a frequency resolution that is one order of magnitude coarser than the spacing between the nearly degenerate modes. Clearly, filter diagonalization is a powerful technique that can significantly reduce FDTD run times without sacrificing the ability to resolve nearly degenerate modes.

16.11.4 The Triangular Photonic Crystal Lattice

Given this basic recipe of establishing a FDTD grid over the primitive unit cell of the photonic crystal, running successive simulations with different boundary conditions to cover the periphery of the irreducible Brillouin zone, and finding the spectral peaks for placement on the dispersion diagram through simple Fourier analysis or filter diagonalization, it would seem rather straightforward to apply these techniques to other photonic crystal lattices. For instance, the triangular lattice of air holes in a high-index dielectric, which can have a fairly wide bandgap for TE polarization (as well as a smaller one for TM) [70], is a photonic crystal that has received a great deal of both theoretical as well as experimental interest. Fig. 16.31 shows the TE band diagram, the real-space lattice, and the irreducible Brillouin zone for the triangular lattice of air holes, as calculated by both the plane-wave method [52] and by FDTD.

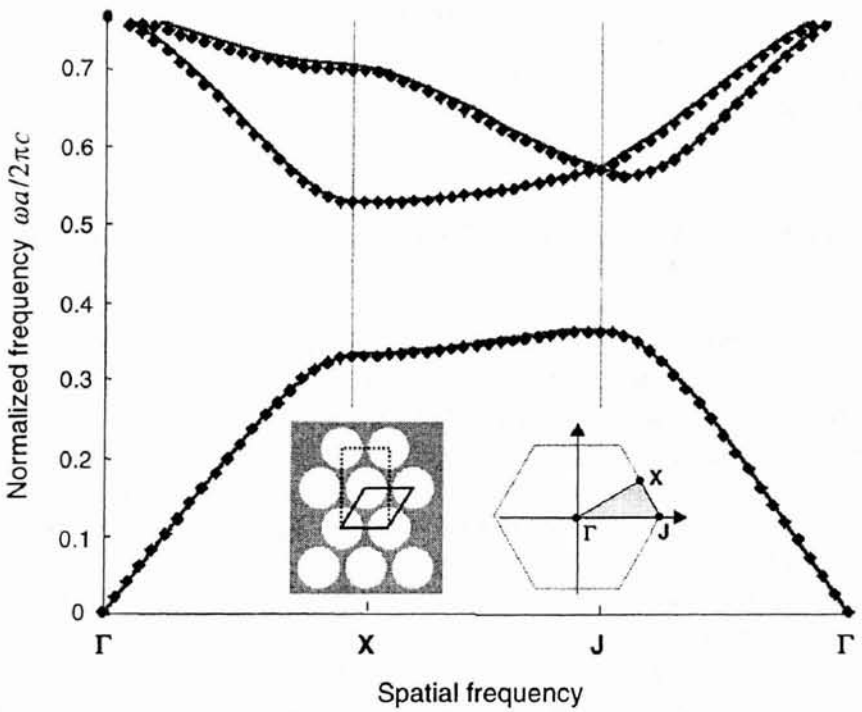


Fig. 16.31 TE dispersion diagram of an infinite two-dimensional triangular lattice of air holes ($r/a = 0.48$) in a dielectric ($\varepsilon_r = 13$), along with the real-space and reciprocal-space lattices. The unit cell used here is the true primitive unit cell (the rhombus shown in the inset), rather than the larger rectangle that would lead to a folded band structure.

Fig. 16.31 seems unremarkable until one realizes that the primitive unit cell for this photonic crystal lattice is the rhombus marked in the left-hand inset, rather than a rectangle. If we had used the rectangle marked with dotted lines in this inset as the unit cell, we actually would have been simulating a "folded" version of the real band structure [71] — one having extra bands to be identified and unfolded. As a result, some researchers have concluded that such non-Cartesian-friendly photonic crystal lattices are simply inaccessible to standard FDTD [60], and require nonorthogonal FDTD [72]. However, one way to overcome this unit-cell problem and yet still use standard FDTD is to place a dipole source in the same spot of every copy of the primitive cell represented within the larger rectangular grid being simulated. (In the unit cell shown in the inset, there are two such copies.) Then, we would adjust the phase of the waveform applied to each to satisfy each particular k-vector [71]. This can be adapted to the simple one-time-step "ping" discussed earlier by having the second impulse (at point $\boldsymbol{r}'$ relative to the first) appear on both layers of the FDTD grid, weighted by $\cos(\boldsymbol{k} \cdot \boldsymbol{r}')$ and $\sin(\boldsymbol{k} \cdot \boldsymbol{r}')$, respectively.

An easier approach [73] is to take the smallest rectangle in which we can inscribe the correct rhomboidal primitive unit cell, perform the standard FDTD algorithm over this entire grid, and then modify the boundary conditions to periodically wrap the sides of the rhombus rather than the sides of the rectangle. (A similar approach was likely taken in [74].) The unit cell is shown in Fig. 16.32 for $r/a = 0.48$ and 30 cells/a, along with some circular arcs added to help visualize the cylindrical air holes. While the additional Yee time-marching steps performed on the cells outside the rhombus are wasted, the advantage is that these customized boundary conditions are a relatively simple computational step that can be added to the usual FDTD algorithm. In fact, in this implementation, this extra computational step is added by a dynamic-link library, meaning that the underlying FDTD program does not even need to be recompiled.

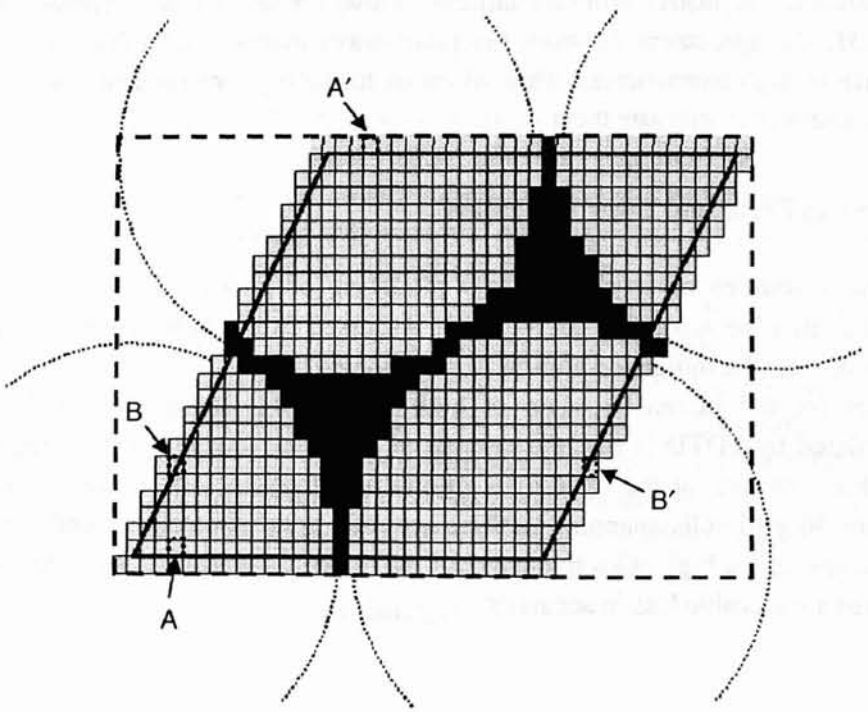


Fig. 16.32 FDTD grid used to compute the dispersion diagram for a photonic crystal consisting of a triangular lattice of air holes in a dielectric. A rectangular grid of FDTD cells is used (dashed box), but the boundary conditions are performed across the rhombus-shaped primitive cell.

The triangular photonic crystal lattice can be thought of as holes arranged in equilateral triangles of side length a, or equivalently as two interlaced comb functions $\mathrm{comb}[x/a, y/2a'] + \mathrm{comb}[(x-a/2)/a, (y-a')/2a']$, where $a' = a\sqrt{3}/2$ is the spacing between rows of holes. Thus, the real-space lattice vectors are $(a/2)(\hat{x} \pm \sqrt{3}\hat{y})$, and the reciprocal lattice vectors are $(2\pi/a)(\hat{u} \pm \hat{v}/\sqrt{3})$.[2] Since symmetry point **X** is located halfway between the center **Γ** point and the next reciprocal lattice point, the k-vector for point **X** is $k_x = \pi/a$, $k_y = \pi/a\sqrt{3}$, and it follows from trigonometry that point **J** is $k_x = 4\pi/3a$, $k_y = 0$.

In Fig. 16.32, Floquet conditions are used which correspond to translating the bottom horizontal segment of the rhombus by $a/2$ in x and $a\sqrt{3}/2$ in y (point A maps onto point A′), and translating the slanted left side of the rhombus by a in x (point B maps onto point B′). The top and right segments of the rhombus are similarly mapped in the opposite direction. Note that using square FDTD grid cells ($\Delta x = \Delta y$) prevents straightforward enforcement of these conditions in both x and y. However, it is simple to choose $\Delta y = a'/\lceil a'/\Delta x \rceil$, where $\lceil \cdot \rceil$ is the "ceil" function. Thus, while Δy may end up slightly smaller than Δx, the unit cell still tiles the plane with an integer number of FDTD grid cells in both dimensions.

As an example of the custom boundary conditions, points A and B shown in Fig. 16.32 serve in the role of the $i = 1$ cells from (16.5), with point A′ receiving the complex field from A multiplied by $\exp[-ja(k_x + k_y\sqrt{3})/2]$, and point B′ similarly from point B multiplied by $\exp(-jak_x)$. As in the simple rectangular grid, these boundary conditions replace the field components in the outermost rows and columns of active grid cells by phase-adjusting the components from the cells that are exactly one row or column "in" on the opposite side.

As can be seen in Fig. 16.31, this approach allows FDTD to accurately compute the photonic crystal band structure of nearly arbitrary lattices. However, as can be observed in both Figs. 16.27 and 16.31, the agreement between the plane-wave method and FDTD is better at low frequencies than at high frequencies. This brings us to the topic of the sources of error in the FDTD models, and how to mitigate them.

16.11.5 Sources of Error and Their Mitigation

The two primary sources of error in FDTD modeling of photonic crystals are numerical dispersion and staircasing artifacts. As these are discussed extensively elsewhere in this book, we only briefly discuss the mitigation approaches considered here.

The reason for the increasing error at high normalized frequencies in the dispersion diagrams calculated by FDTD is that the number of cells per wavelength is reduced for these frequencies. For instance, at the uppermost normalized frequency of $\omega a/2\pi c = a/\lambda_0 = 0.8$ in Fig. 16.27, with 30 grid cells spanning a, there are only approximately 13 cells per dielectric wavelength because of the high refractive index of the material (nearly 3). As a result, numerical dispersion causes a noticeable loss in accuracy.

---

[2]This follows from the Fourier transform of the two comb functions, $\mathrm{comb}(au, 2a'v)\{1 + \exp[-j2\pi(au/2 + a'v)]\}$, where destructive interference in the term with the exponential reduces the full comb function back to a triangular lattice rotated by 90°. Or see Appendix B of [48], noting there the rotated coordinate systems and an unfortunate typo in the reciprocal lattice vectors of the triangular lattice.

To help compensate for numerical dispersion, we can apply the simple correction suggested in Chapter 4, Section 4.9.1, adjusting the permittivity and permeability of free space to force a particular frequency and angle to propagate at exactly the speed of light within the FDTD grid. As suggested in [75], we apply a different adjustment to each axis to allow for nonsquare or noncubic grid cells. However, in contrast to the procedure detailed in [75], we choose to solve an index-adjusted version of (4.12):

$$\left[\frac{n}{c\Delta t}\sin\left(\frac{\pi c\Delta t}{\lambda_0}\right)\right]^2 = \left[\frac{1}{\Delta x}\sin\left(\frac{\pi n\alpha_x \Delta x}{\lambda_0}\right)\right]^2 + \left[\frac{1}{\Delta y}\sin\left(\frac{\pi n\alpha_y \Delta y}{\lambda_0}\right)\right]^2 + \left[\frac{1}{\Delta z}\sin\left(\frac{\pi n\alpha_z \Delta z}{\lambda_0}\right)\right]^2 \tag{16.9}$$

for the values of $\alpha_{x,y,z}$ in a two-step process. Here for readability, we have written out ω in terms of the vacuum wavelength λ_0. First, we remove the effects of the different grid-cell sizes by solving (16.9) for the three Cartesian directions, resulting in

$$\alpha_x = \frac{\lambda_0}{\pi n\Delta x}\sin^{-1}\left[\frac{n\Delta x}{c\Delta t}\sin\left(\frac{\pi c\Delta t}{\lambda_0}\right)\right] \tag{16.10}$$

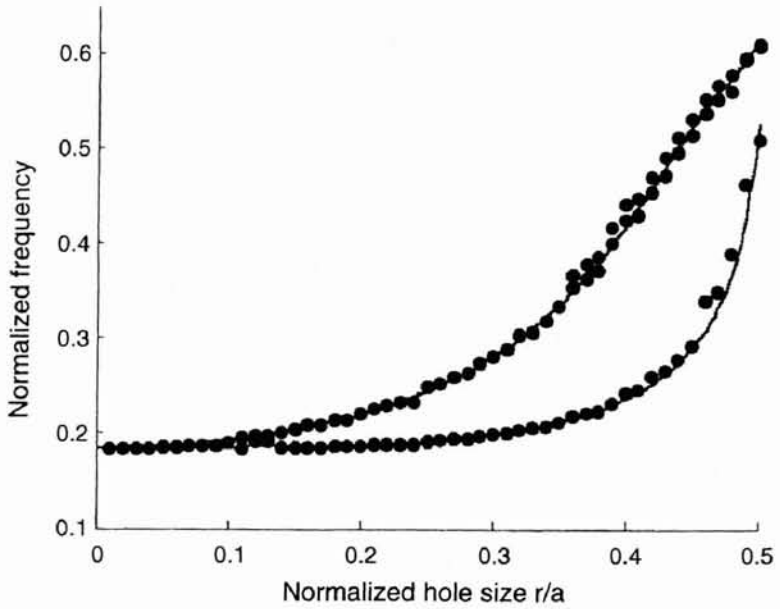
and the equivalent for y and z. At this point, the numerical dispersion is corrected such that the speed of light in the FDTD grid is identically c along any of the three coordinate axes, and is faster at all intermediate angles. It is then straightforward to apply the iterative Newton's method of (4.16) to compute the effective phase velocity at any of these intermediate angles, and then to use this incorrect speed of light to further adjust $\alpha_{x,y,z}$ as in (4.69). Thus, the numerical dispersion can be exactly compensated at this particular angle and wavelength. Alternatively, an angle can be picked which minimizes the numerical dispersion over a particular range of nearby angles, or which minimizes the worst-case numerical dispersion.

The simple mitigation technique for numerical dispersion discussed above provides relatively small improvements in the accuracy of photonic crystal band diagrams [73][3]. This is inherent in the variation of the speed of light within the classic Yee grid at frequencies and angles that are not close to the set-point chosen for compensation. One advantage of even this simple form of numerical-dispersion compensation, which we only briefly mention here, is that it can help to reduce the nonphysical reflections associated with the changes in mesh-cell size in a nonuniform grid [76]. Note that in generating Figs. 16.27 and 16.31, dispersion compensation was implemented at the angle corresponding to the particular k-vector being simulated, and at a wavelength near the center of the photonic bandgap.

[3]The reader is referred to Chapters 4 and 17 for more sophisticated approaches to reduce numerical dispersion.

In contrast with the subtle improvements provided by compensating numerical dispersion, mitigating staircasing errors provides significantly improved accuracy [73]. This is illustrated in Fig. 16.33, which plots as a function of the size of the cylindrical holes the first two FDTD-calculated bands of the triangular-lattice photonic-crystal dispersion diagram at the **J** point (spatial frequency $k_x = 4\pi/3a$, $k_y = 0$) for the transverse electric (TE_z) case. Fig. 16.33(a) shows results using the classic staircased FDTD algorithm for a moderate spatial resolution, 20 cells spanning a. Fig. 16.33(b) shows results for the same grid resolution, but using a subcell method to compute an effective dielectric constant for every E-component near a material interface.

(a)



(b)

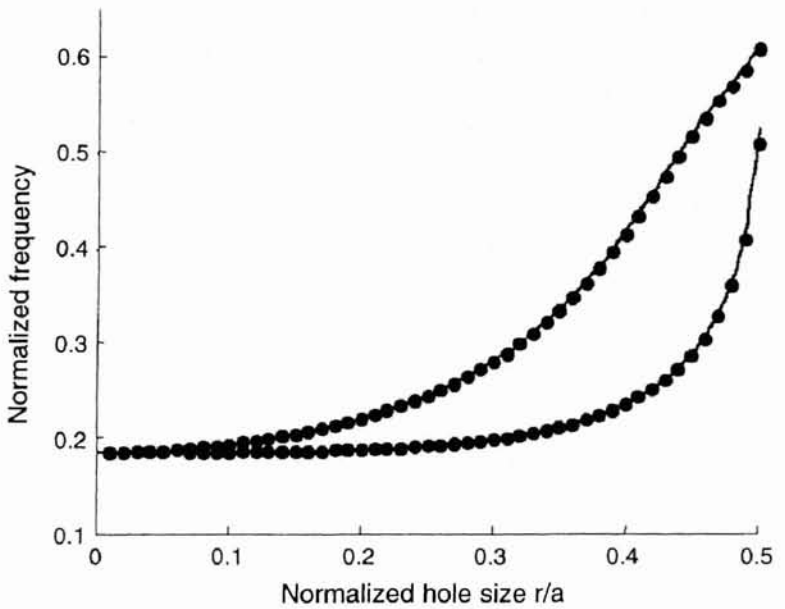


Fig. 16.33 First two FDTD-calculated bands of the triangular-lattice photonic-crystal dispersion diagram at the **J** point (spatial frequency $k_x = 4\pi/3a$, $k_y = 0$) as a function of the size of the cylindrical holes for the TE_z case: (a) staircased hole surfaces; (b) subcell model for hole surfaces.

The particular subcell method used here assigns an effective dielectric constant to each E-component, based on a volume integration of the dielectric-constant distribution within the FDTD grid cell containing this component [77]. For a general Cartesian component E_u that is oriented perpendicular to the v-w grid plane and is located at grid point (u_0, v_0, w_0), this scheme is given by

$$\varepsilon_{u,\text{eff}} = \left\{ \frac{1}{\Delta u} \int_{u_0-\Delta u/2}^{u_0+\Delta u/2} \left[\frac{\Delta v\, \Delta w}{\int_{v_0-\Delta v/2}^{v_0+\Delta v/2} \int_{w_0-\Delta w/2}^{w_0+\Delta w/2} \varepsilon(u,v,w)\, dv\, dw} \right] du \right\}^{-1} \tag{16.11}$$

Here, the direction, u, along which the outer integration is performed, corresponds to the direction of E_u. This allows satisfying the required continuity of the tangential $\boldsymbol{E}$ and normal $\boldsymbol{D}$ fields at the dielectric interface. For E_u parallel to a dielectric interface, the outer integration is straightforward, while the inner double-integral results in a simple averaging of the dielectric constant over the v-w cross section of the grid cell. For E_u in a plane that is perpendicular to a dielectric interface, the inner double-integral is straightforward, while the outer integral leads to an averaging of the inverses of the dielectric constants, just like capacitors in series. Note that this technique allows modeling of an arbitrarily tilted interface in three dimensions. Furthermore, it has no impact on the stability of the underlying FDTD algorithm, and only a minor impact on the memory requirements and execution speed.

The literature contains a variety of other subcell approaches for FDTD modeling of tilted and curved dielectric interfaces [60, 78–82]. An analysis of the impact of these various subcell techniques on the accuracy of FDTD as applied to photonic crystals is currently under way [73].

In summary, relative to staircasing, subcell methods for modeling tilted and curved dielectric interfaces have two important benefits for FDTD modeling of photonic crystals. First, subcell techniques permit improved accuracy in calculating resonant frequencies by properly modeling the amount of high- and low-index material. This minimizes oscillations in the error signal as a function of the grid's spatial resolution. Second, even when there is residual error in computing a resonant frequency, only a few coarse-resolution FDTD simulations may be needed to accurately measure trends or derivatives of interest that involve a change of material boundaries. Of course, these considerations may not be very significant in the analysis and design of two-dimensional photonic crystals, where each FDTD run might take at most a minute, and where even faster alternatives are readily available [52]. However, for large three-dimensional FDTD simulations, the use of subcell methods can significantly reduce the amount of required memory, the execution time, and the number of simulation runs, thereby enabling studies that would otherwise be intractable.

16.12 CALCULATION OF MODE PATTERNS

In addition to the frequencies (eigenvalues) of the photonic crystal modes, it is often important to know the mode patterns (eigenfunctions) associated with each mode. In essence, this should require nothing beyond the analysis described above, with monitor locations at every gridpoint in the cell. The relative amplitude and phase of the response at ω_1 at each point in the unit cell describes the mode patterns:

$$\boldsymbol{E}_{x,y,z}(\boldsymbol{r},t) = \boldsymbol{e}_{x,y,z}(\boldsymbol{r})\, e^{j(\omega t - \boldsymbol{k}\cdot\boldsymbol{r})} = \tilde{\boldsymbol{e}}_{x,y,z}(\boldsymbol{r})\, e^{j\omega t} \tag{16.12}$$

$$\boldsymbol{H}_{x,y,z}(\boldsymbol{r},t) = \boldsymbol{h}_{x,y,z}(\boldsymbol{r})\, e^{j(\omega t - \boldsymbol{k}\cdot\boldsymbol{r})} = \tilde{\boldsymbol{h}}_{x,y,z}(\boldsymbol{r})\, e^{j\omega t} \tag{16.13}$$

However, it is often intractable to retain the full time evolution of all field components at all gridpoints. For this reason, it is more convenient to perform two FDTD runs. The first, as described in the previous sections, uses only a few monitor points to obtain the eigenfrequencies of all the modes. The second then performs a running Fourier transform at every gridpoint for each frequency of interest. Here, one requires two extra variables per cell per field component per frequency of interest, one to retain the cosine (real) part and the other for the sine (imaginary) part, as

$$U_{l,\cos} = U_{l,\cos} + a(t)\left[U_l \cos(\omega_l t) - \hat{U}_l \sin(\omega_l t) \right] \tag{16.14}$$

$$U_{l,\sin} = U_{l,\sin} + a(t)\left[U_l \sin(\omega_l t) + \hat{U}_l \cos(\omega_l t) \right] \tag{16.15}$$

where U stands for any of the three (or six) $\boldsymbol{E}$ and $\boldsymbol{H}$ components, $\hat{U}$ identifies the corresponding field from the second (imaginary) layer of the FDTD grid, and ω_l is the angular frequency of interest. Often, it is sufficient to extract the mode patterns in a single plane, thus reducing the portion of three-dimensional simulations for which this additional storage / computation is performed. With fully complex-valued fields, one could simply multiply by $a(t)\exp(j\omega_l t)$. The envelope function $a(t)$ allows us to apply a Blackman window or other kinds of spectral filtering during the FDTD run. This is important, since it may not be straightforward to use filter diagonalization here to help separate near-degenerate modes from each other.

However, assuming that the initial run is performed with sufficient frequency resolution, we have a priori knowledge of the location of the modes that might interfere, which we can use to our advantage [83]. For instance, if we are interested in finding the eigenmode at ω_2 while avoiding any contribution from a nearby mode at ω_1, we can choose an envelope function whose spectrum has a zero at ω_1.

Fig. 16.34 shows the E_z mode patterns for the two nearly degenerate modes of the square photonic crystal lattice near the **M** point of the dispersion diagram of Fig. 16.27, as also discussed in the context of Fig. 16.30. The top row of Fig. 16.34 shows the lower-frequency mode pattern, while the bottom row shows the pattern for the higher-frequency mode. The higher-frequency mode has a weaker response in these simulations due to poorer coupling from the particular initial excitation "ping." Figs. 16.34(a, d) represent the modes calculated by a long FDTD run (120,300 time-steps), so that the two frequencies are clearly resolved by the end of the running Fourier transform. Figs. 16.34(b, e) represent the modes obtained when the running Fourier transform is accumulated over only 12,000 time-steps. Since the lower frequency mode is stronger, it dominates the truncated computation of both mode patterns. These are obtained by Fourier transforming at the exact frequencies indicated by filter diagonalization on a previous run, without any Blackman window. (With a Blackman window applied, the higher-frequency mode pattern is almost identical to the lower frequency pattern.)

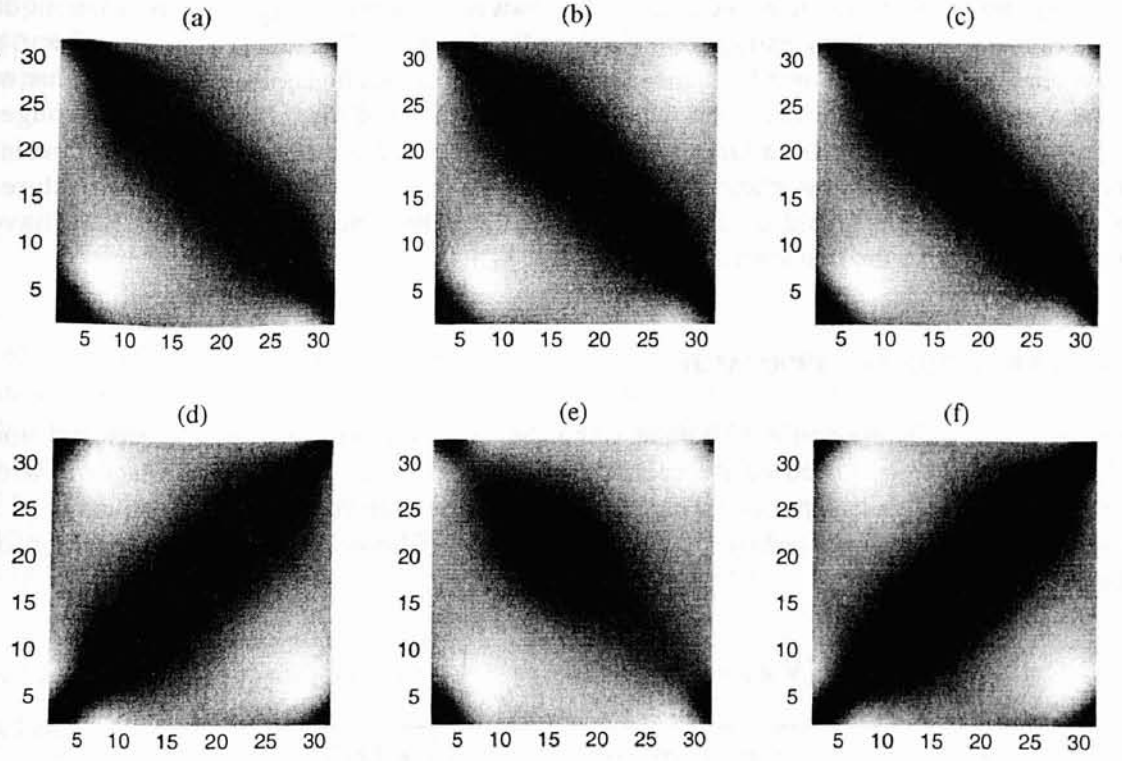


Fig. 16.34 Modal patterns ($|E_z|$) for the two nearly degenerate modes of the square photonic crystal lattice near the **M** point of the dispersion diagram of Fig. 16.27. One unit cell of the photonic crystal is shown, with the dielectric cylinders (not shown) located at the four corners. Parts (a), (b), and (c) correspond to different calculations of the lower-frequency ($f \sim 0.5477$) mode; while parts (d), (e), and (f) correspond to the simultaneous computation of the higher-frequency ($f \sim 0.5497$) mode. Parts (a) and (d) result from long FDTD runs (120,300 time-steps); while parts (b), (c), (e), and (f) result from much shorter simulations (12,000 time-steps). Although the modes are nearly degenerate, the response to the lower-frequency mode is stronger in these simulations due to the spatial placement of the initial excitation. As a result, for coarse frequency resolution, the running Fourier transform of the higher-frequency mode (e) is swamped by crosstalk from the nearly degenerate mode (b). As described in the text, a simple modification allows the properties of the Fourier transform to null out the crosstalk from the strong mode, revealing the weak mode (f).

Figs. 16.34(c, f), in which the two modes are correctly resolved, are obtained with exactly the same reduced number of time-steps. The only difference is that Δt is adjusted so that the normalized frequency f_1 of the strong mode lies exactly at the center of a frequency bin, as

$$(\Delta t)' = \frac{\lfloor f_1 \, c \, \Delta t \, N \rfloor}{f_1 \, c \, N} \tag{16.16}$$

where N is the number of time-steps. Without any spectral windowing, the effective spectral response of each bin is a sinc function [e.g., $\sin(\pi x)/(\pi x)$] whose nulls land at the center of each

neighboring spectral bin. This means that a running Fourier transform at the frequency corresponding to the center of the next higher bin does not accumulate any signal from the mode at f_1, even as it accumulates energy from f_2 (since this bin's sinc function is nonzero at $f_2 \neq f_1$). This occurs even when this next frequency bin-center is somewhat higher than the actual value of f_2. As a result, the mode patterns for the weaker mode can be distinguished from the stronger modes without having to use a large number of time-steps. Such techniques for the Fourier analysis of nearly degenerate photonic crystal modes are possible because the spectral features we are trying to either avoid or detect are essentially delta functions, and because we have a priori knowledge of the frequencies involved.

16.13 VARIATIONAL APPROACH

References [84, 85] reported a technique using the mode patterns of a photonic crystal and variational expressions to reduce the computational burden of computing band diagrams with FDTD. The variational expressions can be derived by reinserting the periodic solution of a photonic crystal mode described by (16.12) and (16.13) into Maxwell's equations and solving for ω [48], producing

$$\left(\frac{\omega}{c}\right)^2 = \frac{\int_{V_{\text{unit}}} \frac{1}{\mu_r} |\nabla \times \boldsymbol{e} - j\boldsymbol{k} \times \boldsymbol{e}|^2 dv}{\int_{V_{\text{unit}}} \varepsilon_r |\boldsymbol{e}|^2 dv} = \frac{\int_{V_{\text{unit}}} \frac{1}{\mu_r} |\nabla \times \breve{\boldsymbol{e}}|^2 dv}{\int_{V_{\text{unit}}} \varepsilon_r |\breve{\boldsymbol{e}}|^2 dv} \tag{16.17}$$

$$\left(\frac{\omega}{c}\right)^2 = \frac{\int_{V_{\text{unit}}} \frac{1}{\varepsilon_r} |\nabla \times \boldsymbol{h} - j\boldsymbol{k} \times \boldsymbol{h}|^2 dv}{\int_{V_{\text{unit}}} \mu_r |\boldsymbol{h}|^2 dv} = \frac{\int_{V_{\text{unit}}} \frac{1}{\varepsilon_r} |\nabla \times \breve{\boldsymbol{h}}|^2 dv}{\int_{V_{\text{unit}}} \mu_r |\breve{\boldsymbol{h}}|^2 dv} \tag{16.18}$$

In turn, these expressions can be used to compute the eigenfrequency ω of a mode from its eigenfunctions $\breve{\boldsymbol{e}}$ or $\breve{\boldsymbol{h}}$. Since $\breve{\boldsymbol{e}}$ and $\breve{\boldsymbol{h}}$ are the same functions computed by the running Fourier transform described in the previous section, these mode patterns can be directly inserted into the right-hand sides of either (16.17) or (16.18) to compute ω.

However, this functionality would seem to be somewhat superfluous, since by the time we can compute the eigenfunction with FDTD, we must have already computed the eigenfrequency with an earlier FDTD run. However, [84, 85] described two significant advantages to using these variational expressions. First, by interpolating the eigenfunction computed by coarsely gridded FDTD onto a finer-resolution grid before inserting it into the variational expression, the accuracy of the computed frequencies can be greatly improved. This is another means to reduce the frequency errors introduced by staircasing and numerical dispersion, which can be applied in addition to the error-reduction techniques discussed previously.

The second advantage is that variational expressions can provide information about the band structure *around* the k value used in the FDTD run, rather than just at that single spatial frequency. One way to do this is to approximate the modal pattern for some k value of interest by interpolating between two FDTD-computed modal patterns (of the same mode) at two

bracketing k values. This linear superposition can then be inserted into the variational expression to compute the eigenfrequency [84], making it possible to fill in the complete band structure using only a few FDTD runs. Finally, from even a single modal pattern, the variational expressions can be manipulated to allow the computation of the group velocity [84] as

$$\nabla_k \omega = \frac{2\,\mathrm{Re} \int_{V_{\mathrm{unit}}} \boldsymbol{e} \times \boldsymbol{h}^* \, dv}{\int_{V_{\mathrm{unit}}} \varepsilon_0 \varepsilon_r |\boldsymbol{e}|^2 \, dv + \int_{V_{\mathrm{unit}}} \mu_0 \mu_r |\boldsymbol{h}|^2 \, dv} \tag{16.19}$$

For dispersive media, we note that the denominator of (16.19) must be modified to obtain the correct energy density, as described in Section 6.8 of [86]. We also note that, to ensure accuracy, the input data ($\boldsymbol{e}$, $\boldsymbol{h}$, and ε_r) must be adjusted to account for the offset spatial positions inherent in the underlying Yee lattice. This is similar to the treatment required for accurate computation of the Poynting vector [87]. In particular, when interpolating E-fields, it is important to consider the impact of any nearby material boundaries [88]. In [85], this latter point was avoided by using only TM polarization wherein the E-field is tangent to all interfaces.

16.14 MODELING OF DEFECT-MODE PHOTONIC CRYSTAL WAVEGUIDES

The intentional introduction of defects into an otherwise perfect photonic crystal opens up opportunities to engineer novel photonic devices. For example, since we know that light with a frequency within the photonic bandgap is not allowed to enter a region of photonic crystal, we can create a waveguide along which light *must* travel, simply by hemming it on all other sides with photonic crystal. An example would be a row of defects created by either removing or adding material to a photonic crystal. Given such a waveguide, we can fabricate: (1) sharp bends by connecting the ends of row and column defects, (2) couplers and splitters by bringing two or more waveguides close together, and (3) filters by modifying the photonic crystal in the vicinity of such a coupler.

While there has been progress in fabricating fully three-dimensional photonic crystals, it is challenging to realize such wavelength-scale structures with the required precision. Conventional semiconductor lithographic techniques are most efficient when fabricating structures having a small number of layers, and when the need for precise layer-to-layer registration can be avoided or finessed. In contrast, assembly techniques for three-dimensional photonic crystals are often subject to both long- and short-range disorder. These techniques face the challenge of having to suppress unwanted defects, and yet permit the introduction of desired defects at arbitrary positions. As a result, many researchers in the photonic crystal community have chosen to work with relatively simple planar or slab geometries.

Fig. 16.35 illustrates an example of a planar photonic crystal. Here, instead of the infinitely long air holes considered previously, a triangular lattice of air holes is drilled into a finite-thickness slab. A row of holes remains filled to serve as the defect-mode waveguide, which is connected at both ends to a conventional ridge waveguide. An air region exists both above and below the slab, at least down to an underlying substrate. Such an "air-bridge" structure [see also Figs. 16.23 and 16.24] can be fabricated using conventional semiconductor lithography, with one of the final steps being to etch away the originally underlying oxide layer of a silicon-on-insulator (SOI) wafer in the vicinity of the photonic crystal [89, 90].

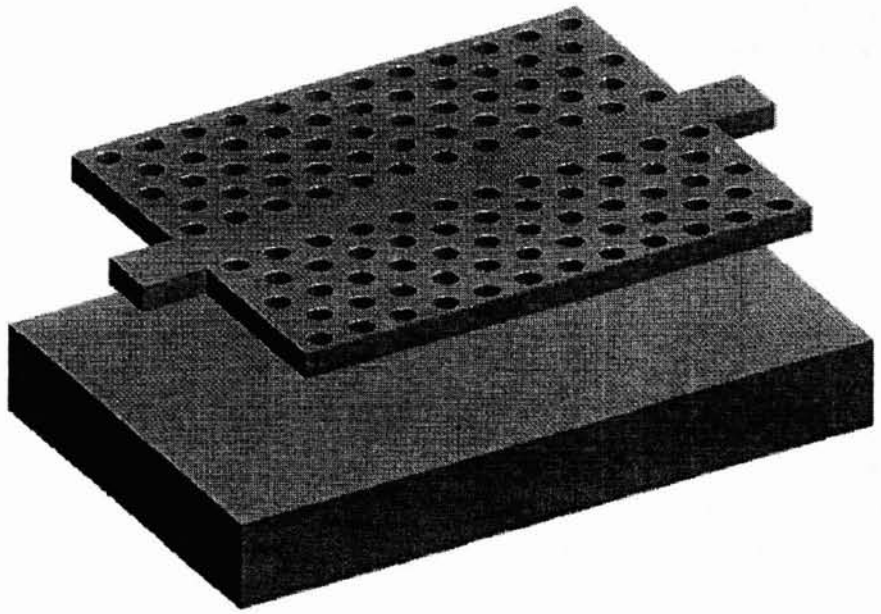


Fig. 16.35 Example of a planar photonic crystal consisting of a triangular lattice of air holes (lattice spacing a) etched into a thin high-index slab. A defect-mode waveguide is formed by not etching one row of holes. Light entering this waveguide at either of its ends from one of the ridge waveguides is confined vertically within the slab by total internal reflection (index guiding), and laterally within the defect row by the photonic bandgap effect. Here, an air-bridge structure is shown that offers improved index guiding, since the underlying substrate is located across an air gap some distance below the slab.

In Fig. 16.35, the optical confinement within the photonic crystal waveguide takes two forms. Within the plane, the presence of the photonic bandgap in the neighboring regions of the slab confines light of relevant frequencies to the defect row. Out of the plane, light is confined by the total internal reflection established by the index contrast between the high-index semiconductor slab and the surrounding air regions. This same index confinement keeps the light in the conventional ridge waveguide as well, but only if the waveguide is not bent too sharply. The presence of the photonic crystal not only offers the potential for sharp bends [91–96], but also for wavelength filtering [97–104], slow group velocity [105–108], and enhancement of nonlinear effects [105–108]. It is even possible to have the photonic bandgap force the light to travel in a lower-index region surrounded by higher index, which has no analogy in conventional index-guiding [109–110].

In this section, we discuss how to extend the FDTD techniques reviewed in the previous sections to the study of photonic crystal waveguides. We do not attempt a full exposition on this subject, but instead refer the reader to numerous excellent articles wherein FDTD modeling, other numerical techniques, and experiments have been used to investigate the properties of these fascinating devices.

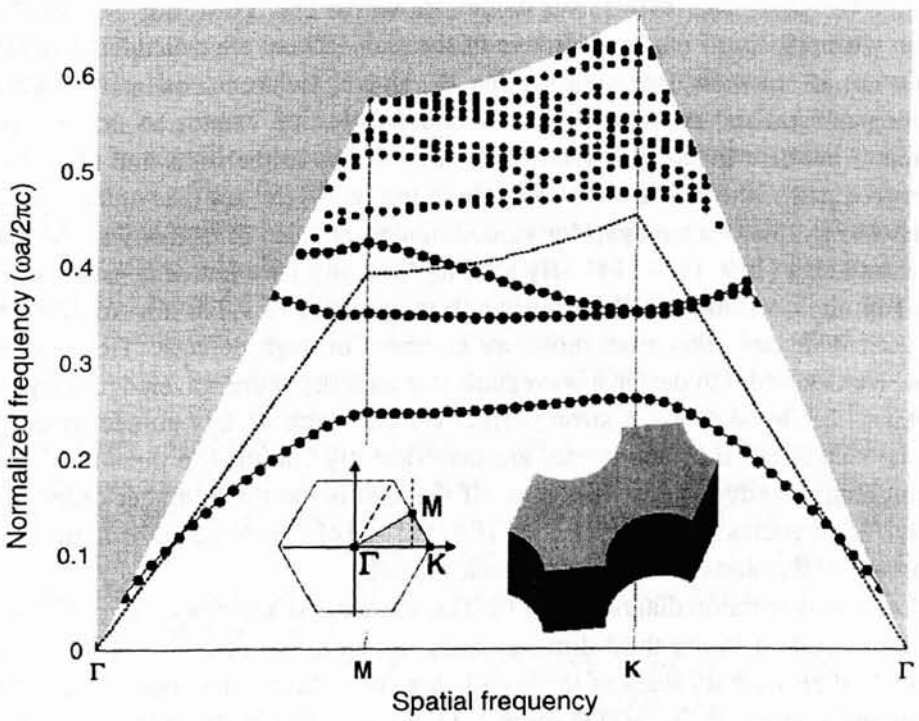


Fig. 16.36 Dispersion diagram for a planar photonic crystal triangular lattice of air holes (hole radius $r = 0.3a$, slab thickness $h = 0.6a$) in a high-permittivity dielectric ($\varepsilon_r = 11.56$), along with the reciprocal lattice. The unit cell used here is the same as the primitive two-dimensional unit cell in x and y (the rhombus shown in the inset), extruded in z through surrounding air regions and the underlying substrate of identical high-index material (located $3a$ below the slab). Light lines are shown for both air ($n = 1.0$, gray-shaded region), as well as for a higher-index cladding ($n = 1.47$, dotted line). Note that the naming convention for the symmetry points of this triangular lattice, and the different convention shown in Fig. 16.31, are both in active use within the photonic-crystal waveguide and resonator communities, respectively.

16.14.1 Band Diagram of a Photonic Crystal Slab

When the third dimension of a photonic crystal structure is made finite, the temporal frequencies of the modes shift and a new feature, the "light line," is added to the dispersion diagram, as shown in Fig. 16.36. Here the spatial frequency on the horizontal axis represents the in-plane component of the k-vector. This diagram is sometimes referred to as a "projected" dispersion diagram (projected to $k_z = 0$). The shift in temporal frequency is due to the fact that some part of each mode now travels in the lower-index material outside the slab [111]. The light line, corresponding to the lower boundary of the gray shaded region in the figure, can be thought of in several ways. In one sense, it marks the critical angle for the total internal reflection that confines modes to the slab. In another, it marks the lower boundary for radiation modes that are not confined to the slab [109, 112]. Thus, any photonic crystal slab mode that lies in the shaded region above the light line can freely couple energy to one of these radiation modes, and in general tends to be lossy. Since the light line has the same slope in all directions, it is sometimes referred to as the "light cone." The curved shape between the **M** and **K** symmetry points can then be understood as the projected shadow of this cone down onto the **M**-**K** leg of the triangular irreducible Brillouin zone (shown shaded in the figure inset).

At least for the first three bands, the modes shown in Fig. 16.36 are "TE-like" that are vertically even with respect to the center plane of the slab. These are called TE-like since they have most, but not all, of their energy in the H_z, E_x, and E_y field components. While a gap is evident between normalized frequencies 0.261 and 0.346, this cannot be considered a true photonic bandgap because of the presence of the modes above the light line [109, 111–113]. (At any given frequency, there is always some direction that light can take and be permitted to propagate.) However, this is a bandgap for guided modes, and can offer many of the features of a full photonic bandgap [109, 111–114]. By spacing the holes by a lattice constant $a = 455$ nm to center this gap on $\lambda = 1{,}500$ nm, the wavelength range 1,315 to 1,740 nm would be available without any guided modes other than those we engineer through defects. However, it is not always feasible (or desirable) to design a waveguide that uses the entire wavelength range.

The thickness of the slab has a strong effect on the width of this guided-mode photonic bandgap. If the slab is too thin, the modes are insufficiently confined to the slab, reducing the bandgap by hugging the edge of the light line. If the slab is too thick, higher-order modes can exist at similar frequencies [112]. See [50, 109, 111–114] for more information about the dispersion diagram of this and other planar photonic crystals.

To calculate this dispersion diagram with FDTD, we take the same two-dimensional unit cell used before and extrude it in the third dimension, as shown in the inset to Fig. 16.36, adding a vertical sandwich of air on both sides of the high-index slab. Absorbing boundary conditions are used at the top and bottom of the FDTD mesh [111, 114]. The CPML [115] is recommended. CPML is an improved version of uniaxial PML [116], and has a straightforward implementation for widely different types of materials (homogenous, lossy, inhomogeneous, and dispersive). See Chapter 7, Sections 7.7 and 7.9 to 7.11 for a thorough discussion of CPML theory and numerical implementation.

In the case of an isolated slab, the surrounding air regions should be at least one to two lattice constants a thick. If a low-index cladding is used on one side, then more space may be needed between the high-index slab and PML, since the modes then penetrate farther into the cladding. Before performing many potentially inaccurate three-dimensional simulations, it is prudent to pick a few k points and try different amounts of cladding to assess the impact on the frequencies of interest. For instance, in the next section, we discuss convergence tests supporting the use of nonuniform gridding along the z-direction of the unit cell. While 20 grid-cells are used to span a in x and z (with the y-gridding closely related, as described in the previous section), portions of the mesh that are farther than a from the slab have only 12 grid-cells spanning a. We compensate as described before for numerical dispersion in the fine- and coarse-gridding regions, and in the intersection of these two regions (since the z-distance between H-components here is the average of $a/12$ and $a/20$). This suppresses most of the nonphysical reflections from the interface between the two grid regions so long as the change in grid-cell size is moderate (less than a factor of three to five).

If the structure and its excitation is symmetric with respect to the center plane, it is possible to halve (and with additional symmetry, sometimes even quarter) the FDTD simulation space by using symmetry conditions. The easiest way to implement symmetry conditions is with a PMC plane for even symmetry, and a PEC plane for odd symmetry. This is easy to remember, since the tangential E-field goes to zero at a PEC plane, and thus the E-fields in the mirrored structure are "odd" with respect to this fold plane. See Chapter 14, Section 14.2.3 for a discussion of the use of PEC and PMC planes in the FDTD mesh to efficiently model the symmetry of a structure and its excitation.

By definition, symmetric and antisymmetric field boundary conditions at a mirror plane in the center of the photonic crystal slab are consistent only with the vertically even and odd modes, respectively [111]. For the case of Fig. 16.36, neither boundary condition could be used because of the asymmetric location of the substrate. In this case, odd modes would always be generated despite attempts to provide a precisely even field excitation. However, we could identify and remove the portions of the first few odd modes that would have otherwise appeared on the dispersion diagram simply by using five to seven monitors arranged at a common x, y point but with different z values. By looking at the shape and strength of the H_z response along these monitors, we could label the modes sufficiently well to visually pick out the even modes from the full dispersion diagram. However, this classification could not be perfect even for the first few bands. For higher bands, the assignment to odd and even is not at all straightforward [112]. For Fig. 16.36, such errors could be corrected manually, and the first few odd bands could be removed for clarity. Had this been an issue, it would have been straightforward to repeat the simulation with running Fourier transforms at the frequencies of interest to obtain modal patterns suitable for unambiguous assignment of frequencies to modes of various symmetries [114]. As before, the spatial and temporal extent of the initial excitation could have been modified to further control which modes were excited. However, our interest here is the simple delineation of the region of the guided mode gap.

The light line is not produced by FDTD, but needs to be drawn on the plot afterwards. At a constant frequency, the k-vectors in the air above the waveguide ($n = 1$) fall on a k-sphere of diameter $2\pi n/\lambda_0$, where λ_0 is the vacuum wavelength. If the wavevector of light in the slab is larger than this radius, the light cannot couple to a mode in the air. The lower boundary of the light line corresponds to the condition in which the transverse k-vector in the waveguide shrinks to the point that matches the radius of the k-sphere in air (e.g., a plane wave of near-grazing incidence to the surface). Since the normalized frequency $\omega a/2\pi c$ can also be written as a/λ_0, the equation of the light line (normalized temporal frequency as a function of k-vector) is simply $a/\lambda_0 = ka/2\pi n$. For the **K** point where $k = k_x = 4\pi/3a$, then for air the light line passes through a normalized frequency of 2/3. Also drawn as a dotted line in Fig. 16.36 is the light line for a surrounding index of $n = 1.47$, corresponding to the case where the oxide underlying the slab is not removed. (Note that only the light line is shown for this; the slab bands themselves would also shift slightly for such a structure.) From the point of view of the light line, whether the higher-index cladding (and its associated radiation modes at lower frequencies) is present on one or both sides of the slab is often not considered important. As we will see below, the light line has a significant impact on the loss of photonic crystal waveguides. In the case of an inhomogeneous cladding (such as in many AlGaAs–GaAs structures where the air holes pierce both the slab and the cladding), the effective index of the overall cladding layer should be used. This is equivalent to the lowest band of the two-dimensional photonic structure [112], which corresponds intuitively in the discussion above to the wavevector of light at near-grazing incidence within the punctured cladding layer.

16.14.2 Band Diagram of a Photonic Crystal Waveguide

Now that we have identified the guided-mode gap described by Fig. 16.36, we can engineer a waveguide by creating a row of defects in the photonic crystal. As described in [112], the best way to do this is to introduce defects along the nearest-neighbor direction (i.e., the Γ–**K** direction, as shown in Fig. 16.35). This produces a waveguide structure that is periodic in x with periodicity at the lattice constant a.

As a result, we need a new dispersion diagram that shows the modes from Fig. 16.36, but only as a function of k_x. To create this from Fig. 16.36, we fold its triangular Brillouin zone at the point $k_x = \pi/a$, marked in its inset as a vertical dashed line running through the **M** point, and plot every mode from the figure as a function of the k_x value (after folding). Thus, the mode at the **K** point appears at the spatial frequency $k_x = 2\pi/3a$. It is important to consider the presence of the other copies of the irreducible Brillouin zone, such as the one shown outlined by the gray dotted line. For example, because of this copy, the mode at the **M** point also appears mapped to the point $k_x = 0$ (the "**M2**" point) as well as at $k_x = \pi/a$.

Fig. 16.37 shows the folded slab modes from Fig. 16.36 as the boundaries of two light gray regions of slab modes, bracketing the mode-gap region between normalized frequencies of 0.261 and 0.346. These gray regions represent the presence of modes from throughout the entire Brillouin zone, as bounded by the modes shown in Fig. 16.36 from the periphery of the Brillouin zone. The light line (for air cladding) has also been placed on the dispersion diagram, running from $(k_x = 0,\ \omega a/2\pi c = 0)$ to $(k_x = \pi/a,\ \omega a/2\pi c = 0.5)$. It is clear that the regions of primary interest are those left unshaded on this figure. If we create a guided mode through defect placement that is in the light-gray region, then we expect it to be lossy by coupling from the defect mode to the many slab modes of the photonic crystal that are free to guide light away from the waveguide. If a guided mode runs above the light line, then we expect it to be lossy by coupling to the many radiation modes that are free to radiate light vertically away from the slab.

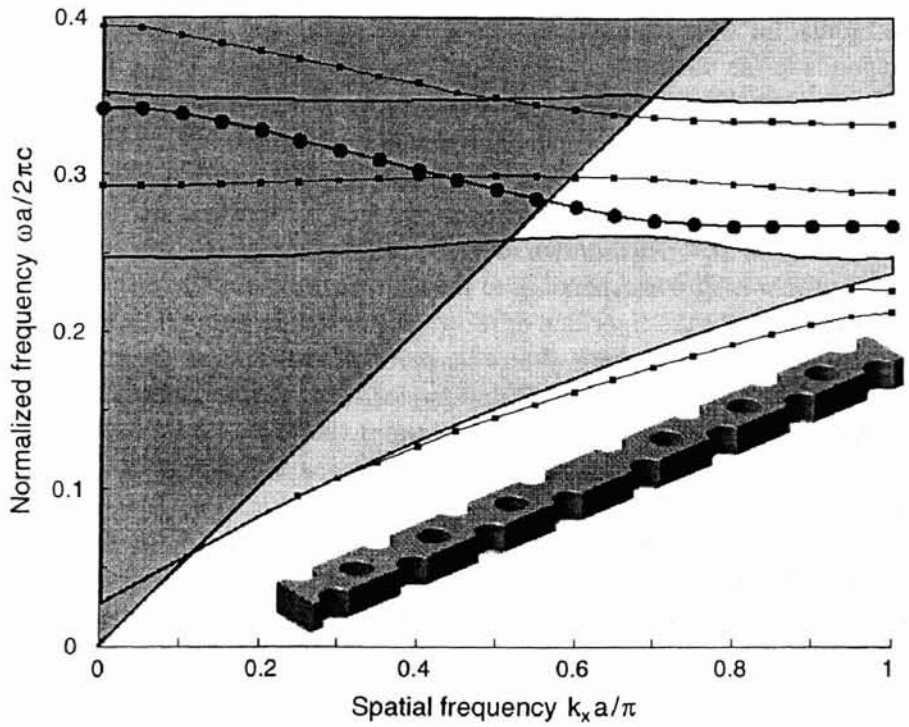


Fig. 16.37 Dispersion diagram for the planar photonic-crystal waveguide formed by removing a row of holes from the triangular lattice along the **Γ–K** direction. The unit cell shown in the inset spans a in the x-direction and $14a$ in the y-direction, with a vertical geometry identical to that of Fig. 16.36. PML absorbing boundary conditions are placed on the $\pm y$ and $\pm z$ outer mesh boundaries, with Bloch/Floquet conditions on the $\pm x$ mesh boundaries. When the lattice of holes continues into the PML, the slab modes are greatly suppressed, but also require a reduced time-step for values of k_x greater than approximately $0.8\pi/a$ to avoid late-time instability.

To compute the defect modes of the photonic crystal waveguide with FDTD, we use a three-dimensional unit cell of the waveguide, as shown in the inset of Fig. 16.37. Here, we take a slice out of the waveguide of length a in the guiding direction x, and which spans a number of photonic crystal rows in y [111–117]. As before, there is air above and below the waveguide, and a substrate is located $3a$ below the slab (not shown in the inset). PML terminates the mesh in both the $\pm z$ and $\pm y$ directions, and Bloch/Floquet conditions are used in the $\pm x$ direction [111, 113, 118]. (While it is possible to use periodic boundary conditions in $\pm y$, the modes of the resulting supercell would also be calculated.) The same FDTD models as before are run, with each determining the eigenfrequencies of the modes at the k-vector established by the Bloch/Floquet boundary conditions. The difference here is that we need only to vary k_x from 0 to π/a.

When the FDTD mesh is terminated in the $\pm y$ direction with a contiguous PML (e.g., the lattice of air holes is continued up to and into the PML region), there is the possibility of a late-time numerical instability, especially for k_x values greater than approximately $0.8\pi/a$ [118]. This problem can be avoided by reducing the time-step to 60% of the nominal three-dimensional Courant stability limit. Other fixes include imposing a vertical notch in the slab outside of the last hole, or continuing the slab without holes for a distance of approximately a before introducing the PML. With these fixes, slab modes are also produced because light propagating along y is retroreflected by the termination of the photonic crystal lattice. However, most of these modes can be readily identified since they fall within the light-gray slab-mode region.

The guided modes introduced by the defect are shown in Fig. 16.37 as small black dots (connecting lines are present only to guide the eye). The defect mode that is horizontally even, and which has received considerable attention in both theoretical and experimental work, is shown as large black dots. Typically, interest in this particular mode is driven by a need to obtain strong coupling to and from the photonic crystal waveguide, since a conventional ridge waveguide has a similarly shaped fundamental mode. The other two modes in the guided-mode gap have different symmetries [109, 111]. While it is clear that part of this mode exists above the light line and part exists below the light line, the dispersion diagram does not provide any quantitative information on the losses. We will take up this topic in the next section.

One of the advantages of using FDTD for computing these guided modes over an alternate technique such as the plane-wave method is that such defect modes can be readily followed into the region above the light line or out of the guided-mode gap. This can be important, since a full understanding of the waveguiding often depends on knowing what other modes are lying above the light line at the same frequency. Of course, if the mode enters the shaded region that denotes slab modes, it is critical to be able to differentiate slab modes from defect modes using PML or other absorbing boundary conditions in the $\pm y$ direction.

Fig. 16.38 is a convergence study that illustrates the effects of coarsening the FDTD grid resolution (from $a/20$ used in Fig. 16.37) and decreasing the number of surrounding photonic crystal rows in the FDTD model (from seven used in Fig. 16.37). Three $k_x a/\pi$ values (0.25, 0.60, and 0.85) are considered. These results demonstrate that the choices used to obtain Fig. 16.37 were conservative. While it would be prudent to check these choices in combination, it appears that a grid resolution as coarse as $a/12$ and only three rows modeled on each side would be sufficient to provide an acceptable error level. We note that this convergence is assisted by the use of the Kaneda (or another) subcell technique, as described earlier. The relative frequency error is less than 0.1% when modeling three or more surrounding rows, and less than 0.3% for grid resolutions finer than $a/12$. As described in previous sections, it is likely that the number of time-steps used to compute the eigenfrequencies could be significantly reduced as well. Here, the number of time-steps is sufficient to resolve two frequencies near the center of the bandgap that differ by 1% (i.e., to place them two frequency bins apart in the resulting FFT).

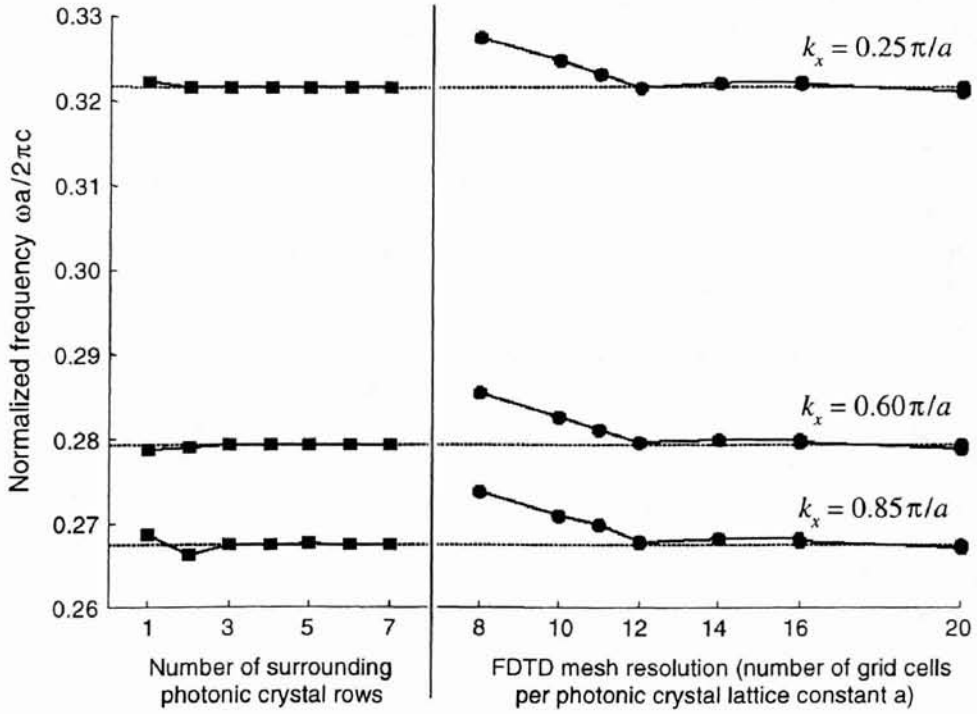


Fig. 16.38 Convergence study of the normalized frequency at three k_x values. Left side: as a function of the number of surrounding rows modeled in the photonic crystal for a fixed $a/20$ mesh resolution. Right side: as a function of the FDTD mesh resolution for a fixed number (7) of surrounding photonic crystal rows. The FDTD mesh cells are scaled so that their x and z spans are identical within the slab; the y-span of the mesh cell is slightly smaller to allow an integer number of time-steps per $a' = a\sqrt{3}/2$.

It is common to read in the photonic crystal literature that FDTD modeling is excessively demanding in computational resources. While this perspective has some merit, it is often based upon overly conservative specifications for the required mesh size, mesh resolution, and number of time-steps. As Fig. 16.38 shows, relaxing these specifications can lead to significant reductions of the computational burden without a significant impact on accuracy. Again, the presence of the subcell methods described earlier are invaluable in enabling such reductions.

16.14.3 Intrinsic Loss in Photonic Crystal Waveguides

There are many types of loss in a planar photonic crystal waveguide. Since the number of rows of photonic crystal in the slab on each side of the defect row is by definition finite, some amount of light can escape this way. Above the light line, light can couple from the waveguide into radiation modes; and below the light line, light can couple from the waveguide into the adjacent high-index substrate. Furthermore, the semiconductor material itself can absorb light in the wavelength range of interest. Finally, modes with k-values very close to the Brillouin zone edges can be coupled into backward-traveling modes through distributed Bragg reflection [119, 120]. These factors are all considered intrinsic losses, since they are present even in a perfectly fabricated device.

In addition, there are extrinsic losses caused by random errors in fabrication. Since semiconductor fabrication procedures usually begin with a wafer whose vertical layer structure can be controlled very precisely, fabrication errors and roughness are often associated with the definition of the holes themselves. Thus, typical errors are variations in the placement of holes relative to the lattice [121, 122], errors in their radius or degree of circularity [123, 124], and any undesired cone angle [125, 126] or roughness [124, 127, 128] on the sidewalls of the cylindrical holes.

Finally, many experiments and even some numerical simulations are affected by coupling issues, both from conventional ridge waveguides to the photonic crystal waveguide, as well as (in experiments) the coupling of light onto and off the device under test. Recently, the use of calibrated tapers at the photonic crystal and polymer-encapsulated mode converters at the edge of the chip have led to significant improvements in the ability to experimentally isolate the loss of the photonic crystal devices [89].

Many numerical and theoretical techniques have been used to estimate the intrinsic and extrinsic losses in photonic crystals, including FDTD [74, 126, 129–132], transmission-line matrix theory [133–136], and coupled-mode theory or modal expansions [121, 128, 137–141]. Besides the FDTD methods of [74] described in this section, and of [124] described in a later section, one quite flexible non-FDTD method was introduced in [123, 142, 143]. In this method, the coupling between the guided and radiation modes is calculated by a time-dependent perturbation theory. While it is not clear how this method would handle the presence of numerous guided modes, it has been successfully used to compute a wide range of both intrinsic and extrinsic losses.

One popular way to measure both loss and waveguide coupling with FDTD has been to simulate a section of photonic crystal waveguide and to then measure the transmission of a pulse through it. However, the measurement of very small levels of intrinsic loss intuitively requires long propagation lengths. Such large simulations are often intractable for three-dimensional FDTD modeling without significant code parallelization. (We will discuss other issues inherent to such transmission measurements in the next section). In fact, motivated by exactly these problems, a phenomenological method was introduced to allow two-dimensional FDTD models to simulate the out-of-plane scattering of the real three-dimensional planar photonic crystal waveguide. This approach encapsulates the out-of-plane losses in an effective dielectric loss coefficient, which is added as an imaginary term to the dielectric constant of the air holes [144]. While this approach has given researchers a useful tool that can be used to fit experimental results and to qualitatively motivate new designs [125, 130, 145, 146], it is clear that a true three-dimensional simulation capable of rapid quantitative prediction would be preferable.

Other researchers have used the temporal decay of the defect-mode resonance to compute the associated losses. As mentioned in a previous section, the eigenfunctions of an infinite two-dimensional lattice, or of a vertically isolated and laterally infinite slab photonic crystal below the light line, should be lossless, and thus have an infinitely high quality factor, Q. However, the defect modes in a finite structure are not lossless, and thus the energy in the mode $U(t)$ decays as [114, 147, 148]

$$U(t) = U(t_0)\, e^{-\omega_0 (t - t_0)/Q} \tag{16.20}$$

By rearranging this equation, or by taking the Fourier transform of the underlying field evolution [147, 149], we can obtain two useful alternate definitions of Q:

$$Q = \frac{-\omega_0 E}{dE/dt} \quad ; \qquad Q = \frac{\omega_0}{\Delta\omega} \tag{16.21}$$

where $\Delta\omega$ is the FWHM of the Lorentzian energy spectrum centered at ω_0, and

$$U(\omega) = \frac{\Delta\omega}{(\omega - \omega_0)^2 + (\Delta\omega/2)^2} \tag{16.22}$$

As can be seen from (16.20), the quality factor describes the loss per unit time. In order to convert this to loss per distance, we can use the group velocity v_g of the mode to derive an equation for the loss coefficient α:

$$U(L) = U(0)e^{-\alpha L} = U(0)e^{-\omega_0 L/Q v_g} \tag{16.23}$$

that can be represented in conventional units such as loss per centimeter. Since this is the loss of the mode propagating in an infinitely periodic structure, this approach is typically used to compute intrinsic loss. However, it could be used for extrinsic losses if they are due to an error that is periodic with respect to the waveguide unit cell (for example, all of the holes in the first row next to the defect having the same slightly elliptical shape).

This approach works well if the Q is relatively low and the group velocity is not too small. Often, Q can be calculated from the noticeable decay of the field envelope, which should fall as $\exp(-\omega_0 t/2Q)$. However, sometimes a lower-loss mode at a different frequency masks the decay of the mode of interest. In this case, the Q can still be computed from the width and position of the spectral line in the FFT. However, when the Q gets very high, it becomes difficult to estimate the decay from either the envelope time response or the spectrum without very long FDTD run times. Fortunately, the filter-diagonalization method described in Section 16.11.3 already provides the decay constants and Q as the imaginary part of the eigenfrequency it computes. This allows one to accurately estimate Q values that would be too high to measure reliably with the FFT spectrum, and to measure more than just the mode with the lowest loss. However, while FDM can extend the effective frequency resolution, we saw in Fig. 16.30 that the range of this resolution extension might be at most four to eight times. Thus, it is perhaps unreasonable to expect FDM to accurately estimate Q values that are 100 or 1,000 times higher than the inverse of the frequency resolution. In addition, if the group velocity is being calculated by taking the difference between two eigenfrequencies, then the impact of frequency errors can be greatly compounded when the group velocity is small. Although correlated errors in frequency (for instance, perhaps all the frequencies are slightly too high) subtract away, even small amounts of uncorrelated frequency error can produce a large relative error when the group velocity is low.

As a result of these considerations, the loss as calculated by (16.23) is often not sufficiently reliable for high-Q and low-group-velocity modes. Fortunately, an alternate technique for computing the intrinsic loss was recently introduced in [74]. This technique is similar to the procedure we used earlier to compute modal patterns, in that once the eigenfrequency of a defect mode of interest is known, we run another nearly identical simulation to focus in on this eigenfrequency. However, rather than "ping" all temporal frequencies and take a running Fourier transform, instead we preferentially excite just the mode of interest by a continuous-

wave excitation of our single spatial dipole as before. By driving the real part of the complex E-fields (the top FDTD "layer") with a sinusoidal carrier, and the imaginary part or bottom layer with a cosine, we can produce an excitation that travels in only one direction across the periodic boundaries [62].

Fig. 16.39 illustrates many of these concepts. This figure plots the out-of-plane loss for the horizontally even mode featured in Fig. 16.37 of the air-bridge, triangular lattice planar photonic crystal shown in Fig. 16.35. This design is identical to the first of the five cases studied in [74], and thus can be directly compared with the lowest curve from Fig. 2(a) of that paper.

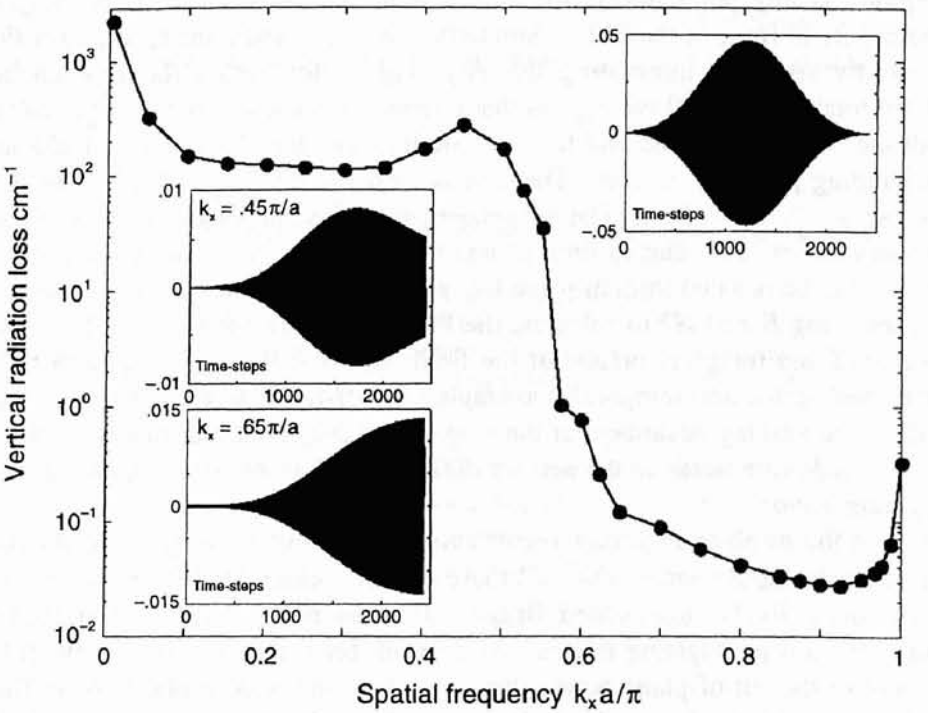


Fig. 16.39 Intrinsic loss due to out-of-plane radiation as a function of spatial frequency for the photonic-crystal slab waveguide described in Figs. 16.35 to 16.37. This curve can be directly compared with the lowest curve from Fig. 2(a) of [74], corresponding to the "deep undercut" design. Insets show the Blackman-windowed incident field (upper right); the resulting field evolution for a lossy spatial frequency above the light line (center left); and the field evolution for a low-loss point below the light line (lower left).

We first consider the insets in Fig. 16.39 as a means to understand its main plot. In order to reduce the excitation of modes at other frequencies, we apodize the temporal waveform with a Blackman window [74]. The envelope of the resulting temporal waveform is shown as the upper right inset in Fig. 16.39. (Here, the carrier oscillates too rapidly to be seen clearly, with one sinusoidal period spanned by about 140 time-steps.) At the end of this excitation waveform, a large amount of energy has been injected into the mode of interest, specified in spatial frequency k by the boundary conditions and in temporal frequency ω by the incident source. If this mode is lossy, this energy continuously leaks out of the waveguide; if it is not, the injected energy would continue to propagate along the guide.

These two cases are shown in the left-hand insets in Fig. 16.39. The upper-left inset shows the field evolution at a k-value located above the light line. For this case, we see that energy leaks out of the mode even before the excitation pulse is finished. In contrast, the lower-left inset shows the field evolution at a k-value located below the light line. For this case, the field amplitude continues to build throughout the excitation. These effects combine to yield the overall loss characteristic plotted in Fig. 16.39, wherein losses are high for the portions of the dispersion curve located above the light line, and are low below the light line.

For the low-loss case, we could choose to prolong the FDTD time-stepping and attempt to measure the slow decay of the field. However, there is an alternative. Since all lost energy exits the simulation space via the PML absorbing boundary, we can measure the total power flow out of the waveguide at any point during the steady-state condition by spatially integrating the Poynting vector just in front of the PML. Similarly, we can measure the total power flow along the waveguide by spatially integrating the Poynting vector across the Floquet boundary. The ratio of the total power outflow, P_{out}, to that remaining in the waveguide, P_0, can be used to measure both the out-of-plane loss and the intrinsic loss into the slab due to a finite number of rows of surrounding photonic crystal. The loss is then $-\log[1-(P_{out}/P_0)]/a$. In the present example, we use $a = 0.42\ \mu m$ [74]. The advantages of this technique are twofold. First, we can measure very low losses (an accuracy limit of less than 0.01 cm^{-1} is estimated in [74]). Second, out-of-plane loss can be isolated from in-plane loss and from effects such as mode coupling.

Before combining E and H^* to calculate the Poynting vector for a power flow, we account for the Yee spatial and temporal offsets of the fields by spatially averaging both fields to the centers of the mesh cells, and temporally averaging the H-fields to coincide with the E-fields [87]. Overall, we are taking advantage of the single-frequency nature of this simulation and the presence of the quadrature fields in the second FDTD "layer" to produce a monitor of the time-averaged Poynting vector.

We note that the in-plane Poynting vector evolves differently during these simulations for k-values very close to the Brillouin zone. At these k-values, either bidirectionality is imperfectly suppressed or, more likely, distributed Bragg reflection redistributes power into both the forward- and backward-propagating modes. As a result, for k-vectors close to either 0 or π/a, we take the ratio of the out-of-plane power flow relative to the peak in-plane power flow, rather than the final in-plane power flow. While it is possible to change the transition point between these two operations without introducing a nonphysical jump in the calculated losses, the values of loss calculated with this algorithm for k-values within $0.05\pi/a$ of the Brillouin zone edges may not be as reliable as at other spatial frequencies. This issue was apparently finessed in [74] by simply avoiding such k-values.

Fig. 16.40 provides the results of a convergence study of the FDTD-calculated vertical radiation loss with respect to the number of surrounding rows of the photonic crystal in the model, the mesh resolution, and the number of time-steps. We see that as few as four surrounding rows, 12 mesh cells spanning a, or 6,000 time-steps could be used without significant deviation of the results. (Of course, the impact of making all three of these choices together would need to be verified.)

Interestingly, initially seeding these tests with the accurate eigenfrequency obtained with the conservative choices of seven surrounding rows and 20 cells spanning a, yields somewhat larger errors at the left edge of each subplot in Fig. 16.40 than those shown here. This can be traced to the fact that very little energy is injected into the mode of interest when we combine a slight frequency error and a long time window. Simply using the same conditions (coarse gridding or reduced number of surrounding photonic crystal rows) for both the frequency and loss calculations avoids this problem.

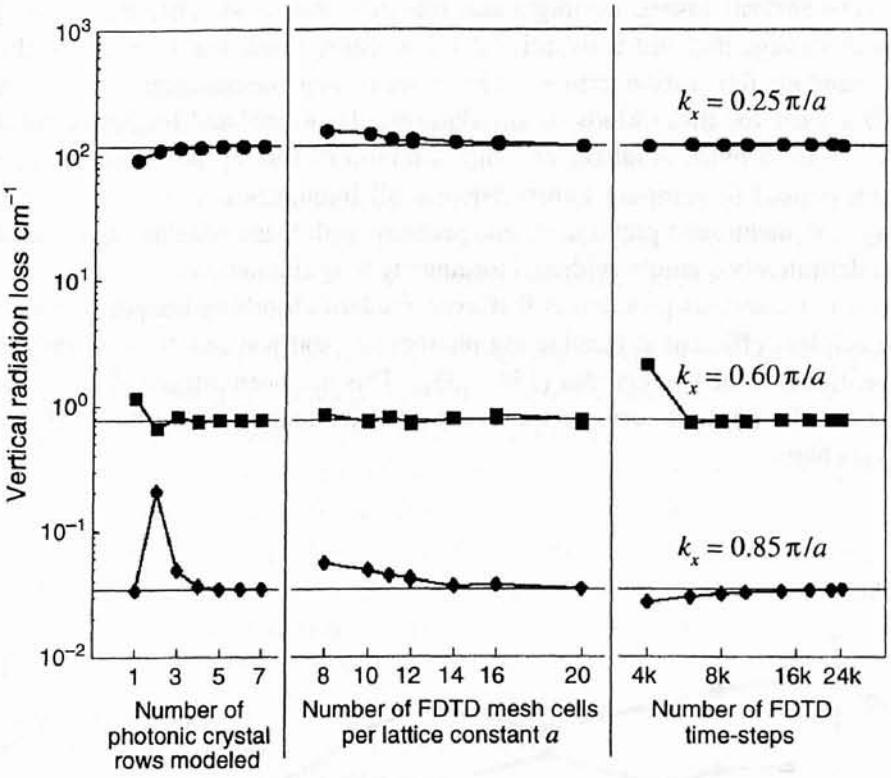


Fig. 16.40 Convergence study of the FDTD-calculated vertical radiation loss at three k_x values from Fig. 16.39 with respect to the number of surrounding rows of the photonic crystal in the model, the mesh resolution, and the number of time-steps. The length of the Blackman window is chosen to be five optical cycles (approximately 700 time-steps) shorter than the overall length of the simulation.

Similarly, the use of a shorter time window on the excitation allows any nearby frequency to produce the correct excitation, although this should be avoided if a second defect mode were located at a nearby frequency. Still, we can reduce the computational load in terms of the number of time-steps so long as no competing defect (or slab) mode is nearby.

16.14.4 Transmission in Photonic Crystal Waveguides

The method of [74] described in the previous section can calculate extremely low losses, and can be used to isolate the intrinsic losses of individual modes. However, it requires a pair of simulations per k-point, and thus does not exploit the ability of FDTD to compute broadband frequency response with a single simulation. One approach that does was introduced in [131]. Here, a larger simulation space of 10 lattice constants, surrounded on all sides by absorbing boundary conditions, is excited with a pulsed input from a dipole. The waveform is captured along the waveguide, Fourier transformed to isolate each ω in the excitation bandwidth, and then filtered spatially by correlation against previously captured mode profiles. This was used to study the loss above the light line, but was shown to have difficulties measuring loss when the group velocity is low [131].

In addition to intrinsic losses, we might also like to be able to simulate the coupling between photonic crystal waveguides and conventional ridge waveguides, and to estimate the extrinsic losses due to random fabrication errors. The transmission measurements briefly mentioned before are often used for these kinds of simulations. Here, isolated boundary conditions are imposed on all sides, a pulse is introduced, and the ratio of flux at the far end to flux near the input excitation is used to compute transmission at all frequencies within the pulse bandwidth simultaneously. As mentioned previously, one problem with these simulations is that low losses are difficult to definitively quantify without inordinately long simulations.

However, a more serious problem is that even modern absorbing boundary conditions such as PML are much less efficient at terminating photonic crystal waveguides than they are for free space or conventional ridge waveguides [150–153]. This has been attributed to imperfect mode matching between the photonic crystal waveguide and its analog within the PML. Fig. 16.41 illustrates this problem.

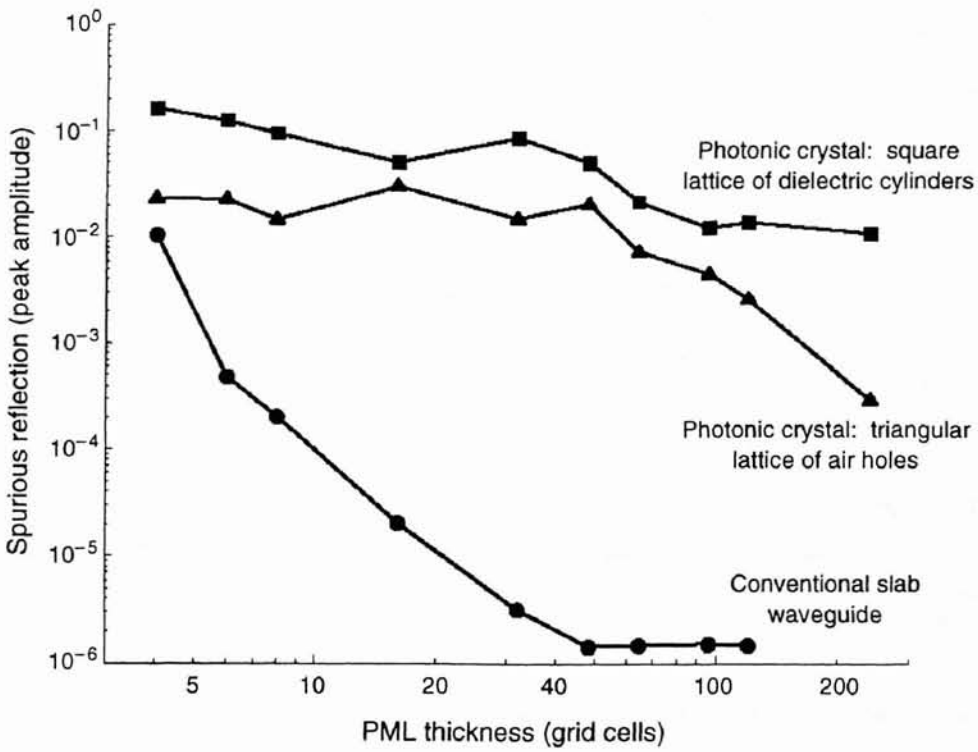


Fig. 16.41 Spurious reflections as a function of PML thickness when terminating a conventional high-index slab waveguide and two different photonic crystal waveguides. All FDTD simulations employ two-dimensional TM_z grids. The slab waveguide (350-nm thick, $n = 3.4$) propagates a TE waveguide mode centered at $\lambda \cong 1{,}550$ nm within a square-cell grid of resolution 25 nm. The square-lattice photonic crystal waveguide has a single missing row of dielectric cylinders ($r/a = 0.2$, $n = 3.4$), and is excited at a center frequency $\omega a/2\pi c = 0.37$ in a grid of resolution $a/18$. The triangular-lattice photonic crystal waveguide has a single missing row of air holes ($r/a = 0.4$) in a high-index background ($n = 3.4$), and is excited at a center frequency $\omega a/2\pi c = 0.37$ in a grid having $\Delta x = a/18$ (with Δy related as described earlier).

Fig. 16.41 compares the spurious reflections as a function of PML thickness when terminating a conventional high-index slab waveguide and two different photonic crystal waveguides. (Note that, for the latter two cases, the air-dielectric variation of the photonic crystal is continued throughout the entire extent of the PML.) From this figure, we see that PML is much more effective for the conventional slab waveguide than for the photonic crystal waveguides, even for dramatically increased PML thickness. Similarly (but not shown in this figure), little improvement is obtained in terminating the photonic crystal waveguides upon varying the usual PML parameters of σ_{max} and polynomial-grading order m, as defined in Chapter 7, Section 7.6.2.

The low efficiency of PML terminations for photonic crystal waveguides negatively impacts FDTD simulations. There is the obvious error introduced by the PML reflection artifact, which causes the calculated transmission spectra to have oscillations that are dependent upon the length of the waveguide. Another uncertainty is how to measure flux: whether with one monitor, a few scattered monitors, all of the flux within a few a of the waveguide, or the entire lateral extent of the simulation. Yet another problem is that it is more difficult to develop an accurate "bootstrapped" wave source; that is, a record of a pulsed waveform supported by the waveguide for use in sourcing subsequent simulations (see Chapter 5, Section 5.11). Data records of this type can also be contaminated by spurious reflections.

One strategy to attack the reflective-termination problem involves modeling the coupling of the input and output ports of the photonic crystal waveguide of interest to a waveguide that *can* be efficiently terminated with PML. For example, [150] reported for this purpose the design of a DBR waveguide (a one- or two-dimensional Bragg grating orthogonal to the guide direction). DBR waveguides can be terminated with PML because they possess continuous rather than translational symmetry. However, for the smoothest coupling, such a guide should be optimally designed for each photonic crystal waveguide of interest, and even then, may have difficulty in achieving spurious reflections below a few percent at the input and output ports.

In the same genre, we can model the coupling of ridge waveguides to the input and output ports of the photonic crystal waveguide. This has the advantage of potentially corresponding to the actual physical measurements of the device. (For example, see the moderate-Q photonic crystal resonator studied in Section 16.15 [154].) However, due to limitations on the size of the simulation space, the input and output ridge waveguides likely are much closer together in the FDTD model than in the physical experiment. As a result, such simulations could include artifacts caused by interactions of the input and output couplings, especially potential Fabry-Perot oscillations.

Of course, the brute-force solution to the problem of terminating photonic crystal waveguides is to simply use enormous simulation spaces. This allows temporal isolation of the reflected and transmitted pulses of interest (e.g., those due to a bend in the waveguide [90, 155]), from the spurious pulses due to the outer boundaries of the simulation space. In such simulations, the pulse bandwidth is chosen carefully: too small, and the pulse wavepacket is so long that it overlaps at the monitor despite the long propagation lengths; too large, and the group-velocity dispersion inherent to many photonic crystal waveguides creates a broad wavepacket anyway [90]. Unfortunately, this approach also means that such simulations, especially for three-dimensional FDTD, mandate parallelization across a large number of processors.

16.14.5 Aperiodic Photonic Crystal Waveguides

Recently, [156–158] reported a photonic crystal waveguide design that is intentionally aperiodic. Here, the row of holes closest to the defect row is made periodic along the waveguide with period $b \neq a$. For this structure, the simulations described earlier, based on the unit cell along the waveguide, are no longer appropriate. In an alternate approach [156, 157], FDTD is used to model the injection of energy from a ridge waveguide into the photonic crystal waveguide. The k-vector of the light propagating in the photonic crystal waveguide region is then analyzed using two Fourier transforms: one transform in time at each grid cell to isolate particular temporal frequencies ω; and then a transform in space to identify the particular k-vector. Comparison of the results of this *spatial Fourier-transform* (SFT) technique to the conventional Order-N method described previously indicates good agreement [156, 157].

An advantage of the SFT technique is that it can potentially obtain the entire dispersion of defect modes, ω versus k, with one simulation. A disadvantage is the size of the simulation relative to the single unit cells, especially given that the shorter the waveguide, the fewer oscillations along x that are available for identification of k. It might be expected that such simulations could also be adversely affected by imperfect termination of the photonic crystal waveguide, although [156, 157] do not indicate the thickness of the PML used to terminate the waveguide. Fortunately, any power reflected from the waveguide port at temporal frequency ω and wavevector k is more likely to establish a noise signal at $2k$, corresponding to a standing wave. Another interesting consideration for the SFT method is how well two or more modes with common symmetry can be distinguished if they share the same temporal frequency.

16.14.6 Photonic Crystal Waveguide Extrinsic Scattering Loss from the Green Function

As mentioned above, the method for loss calculation introduced in [74] is primarily intended for intrinsic loss (as well as extrinsic loss due to periodic errors). However, the ability to focus on a single k-vector at a time improves the reliability of the results and eases subsequent interpretation of the underlying physics. Recently, [124] introduced a method of computing the extrinsic scattering loss using FDTD. Here, FDTD is used as described above to compute the defect modes and their mode profiles. Then, pairs of dipoles and monitors are successively applied to map out the Green function tensor at the positions within the photonic crystal lattice that are expected to be the primary contributors to scattering loss (for instance, the edges of the air holes) [159].

Using this Green function and a suitable model for the error in the permittivity, $\Delta\varepsilon$, between the desired photonic crystal structure and the structure that is actually fabricated, the power loss due to backscatter and the overall extrinsic power loss can be computed as [124]:

$$\alpha_{\text{back}}(\omega) = \frac{a^2\omega^2}{4|v_g|^2}\int\int dr\,dr'\,\Delta\varepsilon(r)\Delta\varepsilon(r')\left[E_k^*(r)\cdot E_k^*(r)\right]\left[E_k(r')\cdot E_k(r')\right] \tag{16.24}$$

$$\alpha_{\text{total}}(\omega) = \frac{a\omega}{|v_g|}\int\int dr\,dr'\,\Delta\varepsilon(r)\Delta\varepsilon(r')\ \text{Im}\left[E_k^*(r)\cdot\ddot{G}^B(r,r';\omega)\cdot E_k(r')\right] \tag{16.25}$$

using the mode patterns $E_k(r) = e_k \exp(jkx)$ and the assembled Green function tensor $\ddot{G}^B$.

Even before computing any elements of the Green function, it is already apparent from (16.24) and (16.25) that slow group velocity is associated with large extrinsic losses, and in particular with very large backscatter. (Note that second-order scattering from the backward mode into the forward mode and similar higher-order terms are not represented in these equations. See citation #17 in [124].) From (16.24) and (16.25), it is also apparent that any particular error in shape or roughness (for example, where there should be dielectric, there is air instead) results in a larger value of $\Delta\varepsilon$ (and thus more extrinsic scattering) for a high-index-contrast slab than for structures where the materials involved are closer in index. By choosing an appropriate value of root-mean-square surface roughness and correlation length, it has proven possible to closely match experimentally measured loss spectra of both conventional W1 as well as W0.7 photonic-crystal waveguides [160]. (In a W0.7 waveguide, the photonic crystal sections on either side of the defect mode are moved 30% closer together, improving the group-velocity dispersion and pushing competing modes out of the bandgap [106].) Thus, this Green function method offers the opportunity to compute extrinsic loss.

In a certain sense, this method has many similarities with the methods of [123, 142, 143]. Both depend on enumerating all the guided modes, and both then quantify the coupling between these modes and the radiation modes.

16.15 MODELING OF PHOTONIC CRYSTAL RESONATORS

Just as a row of defects creates a photonic crystal waveguide, an isolated defect in an otherwise perfect photonic crystal can create a photonic crystal resonator. Light is trapped at such a defect because it is hemmed in on all sides by regions of photonic crystal through which it cannot propagate. Such resonators can be used as out-of-plane couplers [100, 161–164], wavelength filters [97–99, 165], lasers[4] [166–172], cavities designed for *quantum electrodynamics* (QED) experiments [173–175], sensors [154, 176–179], tunable filters [180–186], and switches [187, 188]. Planar as well as fully three-dimensional photonic crystal structures such as the "woodpile" have been studied [152, 189–193].

For most of these applications, the parameter of primary interest is the quality factor Q of the photonic crystal resonator. A high-Q cavity has a narrow spectral response and a long lifetime for its resonant photons; equivalently, the resonator loss is low. This is certainly desirable for filters and lasers. Additional desirable attributes of a photonic crystal resonator vary according to the application. For example, a key goal in the design of photonic crystal lasers is to produce an electrically pumped device. This implies that the maximum optical field within the photonic crystal cavity should be in a solid region where one can position gain material and electrical contacts [168, 172]. In contrast, in QED experiments, wherein the goal is to manipulate the quantum state of a neutral gas-phase atom, the photonic crystal cavity should have its maximum optical field at a hole [173]. Here, in addition, to obtain strong coupling between the atom and the cavity mode, not only should the Q be large, but the effective volume of the cavity mode, V_{eff}, should be small [173].

---

[4]In addition to defect-based resonators, a "band-edge" laser can be engineered by using the distributed Bragg reflection inherent at symmetry points of the band diagram where the bands become flat [194–197]. Here, the center of the resonator is dictated not by the engineered defect, but by the optical or electrical pumping.

The suite of FDTD tools currently used to study photonic crystal resonators has essentially been covered in previous sections. Reference [166] first laid out this set of tools, starting with two-dimensional band diagrams of a perfect photonic crystal (Section 16.11.1); extending this to a three-dimensional band diagram delineated by the light line (Section 16.14.1); and finally three-dimensional simulations to explore the impact of the defect. The latter parallel the defect-mode and intrinsic-loss simulations discussed in Section 16.14, but with absorbing boundary conditions on all six boundaries. Once the band diagram of the planar photonic crystal is known, the resonator is probed with delta functions in both time and space to identify the modes of the resonator [166]. Alternatively, the temporal pulse used for excitation can be tailored to reduce its spectral content. Symmetries of the resonator structure can be used to both reduce the size of the simulation space and isolate those modes sharing the given symmetry [166]. Alternatively, multiple dipoles can be arranged in symmetric positions to have the same effect [198], or the results of multiple FDTD runs can be combined to project onto the proper irreducible representation according to group theory [189]. Since, in general, there are no longer any Bloch/Floquet boundaries, there is no need to scan over spatial frequency.

Once the frequencies of the defect modes of interest are known, Q can be obtained by adapting the intrinsic-loss simulations described in Section 16.14.3.[5] The resonator is excited by a pulsed excitation centered on the frequency of the mode of interest, and the Q is calculated using (16.20) from either the energy decay or the relative amount of power leaving the simulation space. For the latter, $Q = \omega_0 U(t)/P(t)$, where $U(t)$ is the energy in the mode and $P(t)$ is the amount of power leaving the simulation space. One advantage of measuring the outgoing power over the energy decay is that we avoid the necessity of accurately measuring what could be a very slow decay of the energy. However, both methods give similar results [166, 199]. A more important advantage of computing Q by the power shed from the resonator is that different contributions to the total quality factor, Q_T, can be studied separately—specifically, the portion due to vertical or out-of-plane losses ($Q_\perp$), and the portion due to power loss within the plane of the slab ($Q_\parallel$), where $Q_T^{-1} = Q_\perp^{-1} + Q_\parallel^{-1}$.

However, while this method is fairly straightforward, there are still cautionary points. First, it is important to be sure that the outgoing flux at the boundaries reaches a steady-state condition before using it to calculate Q, especially when Q is large. Second, if one wants to concentrate on $Q_\perp$ or $Q_\parallel$ individually, the division between which portions of the outer boundaries correspond to $P_\perp(t)$ and which to $P_\parallel(t)$ must be chosen with care. For instance, portions of the sidewalls far from the slab almost certainly intercept some of the outgoing vertical radiation, yet the in-plane loss cannot be considered to be completely confined to the high-index slab. One common choice is to consider the slab plus one-half wavelength of cladding above and below as the in-plane component, with everything else assigned to $P_\perp(t)$ [173–175].

In addition to the choice of flux planes, the excitation bandwidth must be chosen carefully. As described in previous sections, it is often important when investigating a particular mode to avoid injecting any energy into other modes. However, if the bandwidth of the excitation is made too narrow, the high Q value obtained may be the spectral width of the excitation rather than that of the underlying resonator [199].

---

[5] An alternative approach is to improve the estimates of resonant frequency and Q by iteratively inserting the modal patterns into a variational expression [200].

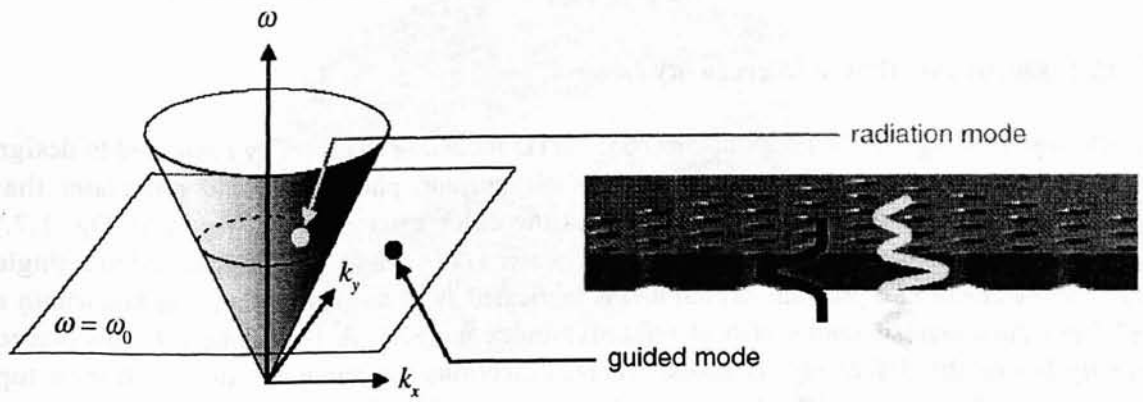


Fig. 16.42 Modes in a planar photonic crystal resonator with components inside the light cone tend to be lossy due to coupling to radiation modes. A *k*-space design thus concentrates on placing the spatial-frequency content of modes outside the light cone. *Adapted from:* Srinivasan and Painter, *Optics Express*, 2002, pp. 670–684; and *Optics Express*, 2003, pp. 579–593.

Just as with the planar photonic crystal waveguide, the light line represents an avenue for light to escape a planar photonic crystal resonator, and thereby reduce its Q. Fig. 16.42, adapted from [169, 170], elegantly shows the relationship between spatial frequency, temporal frequency, the light line, and the relative positioning of guided and radiation modes. Many researchers have used FDTD to help redesign their cavities to suppress this loss mechanism and thus increase Q [167, 169–171, 174, 199]. Reference [174] pointed out that by applying the paraxial approximation to the analysis of the far-field radiation from a planar photonic crystal resonator, the vertically radiated power at some vacuum wavelength λ_0 could be written in terms of the in-plane field components as

$$P = \frac{1}{16\pi^2}\int\int\left\{\eta_0\left[\left|\mathcal{F}(H_x)\right|^2+\left|\mathcal{F}(H_y)\right|^2\right]+\frac{1}{\eta_0}\left[\left|\mathcal{F}(E_x)\right|^2+\left|\mathcal{F}(E_y)\right|^2\right]\right\}\boldsymbol{dk} \qquad (16.26)$$

where $\eta_0 = \sqrt{\mu_0/\varepsilon_0}$; $\mathcal{F}$ represents the two-dimensional spatial Fourier transform; and the integration is performed over all spatial frequencies inside the light cone (e.g., with magnitude $|k| \le 2\pi/\lambda_0$). Thus, the modal patterns of the resonator (Section 16.12) become an avenue to reducing its out-of-plane losses by redesigning the cavity mode to decrease the low-spatial-frequency content. This redesign can be driven by symmetry [169, 170], by solving the inverse problem [201], by destructively interfering lower-order multipole radiation components [68], or by creating a dislocation [174] or a shift [202] of the holes. It is conventional to use the field patterns at the center plane of the resonator for this *k*-space design. It is important to suppress the effects of the field pattern that lie outside the real-space boundaries of the resonator, but without shifting or altering the *k*-space pattern during such windowing [199]. FDTD can then be used to investigate the sensitivity of the Q to slight variations in the device parameters, as a way to develop resonator designs that are more tolerant to potential fabrication errors [203, 204].

16.16 MODELING EXAMPLES OF PHOTONIC CRYSTAL RESONATORS

16.16.1 Electrically Driven Microcavity Laser

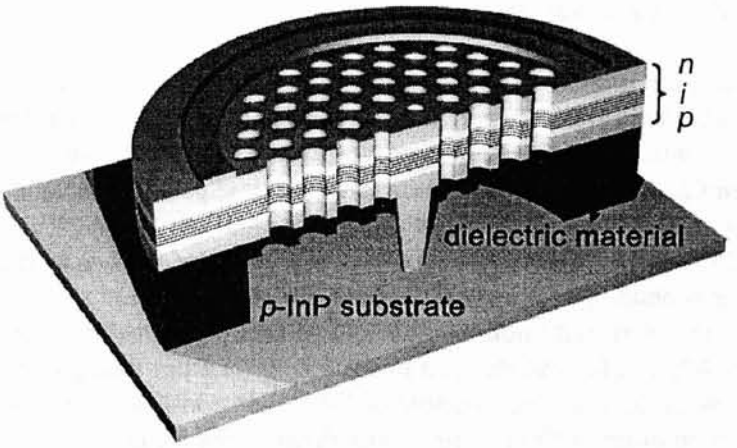
As illustrated in Fig. 16.43, three-dimensional FDTD modeling has recently been used to design an electrically driven, single-mode, low-threshold-current, photonic crystal microlaser that operates at room temperature [172]. (See also the color version of this figure in Fig. 1.7.) Fig. 16.43(a) is a schematic diagram of this microlaser [172]. Here, light is confined to a single defect at the center of a photonic crystal that is fabricated as an array of air holes etched within a 282.5-nm-thick semiconductor slab of refractive index $n = 3.4$. A 1-μm-long InP post placed directly below the defect injects holes, whereas electrons are supplied laterally from a top circumferential electrode. The holes and electrons recombine in six InGaAsP quantum wells within the photonic crystal slab. These quantum wells are designed to have an electroluminescence peak near the communications wavelength of 1,500 nm.

Fig. 16.43(b) is a scanning electron microscope image of the top view of the fabricated photonic crystal slab [172]. The defect-mode cavity is surrounded by five photonic-crystal regions I, II, III, IV, and V having the same lattice constant $a \cong 510$ nm, but different air-hole radii: 0.28a, 0.35a, 0.385a, 0.4a, and 0.41a, respectively. Optical electromagnetic fields are confined within the cavity in the vertical direction by total internal reflection at the slab-air interface, and laterally by the action of the photonic crystal bandgap. Therefore, light can escape from this cavity only by either tunneling laterally through the photonic crystal, or by exiting vertically by impinging on the slab-air interface at a sufficiently large angle.

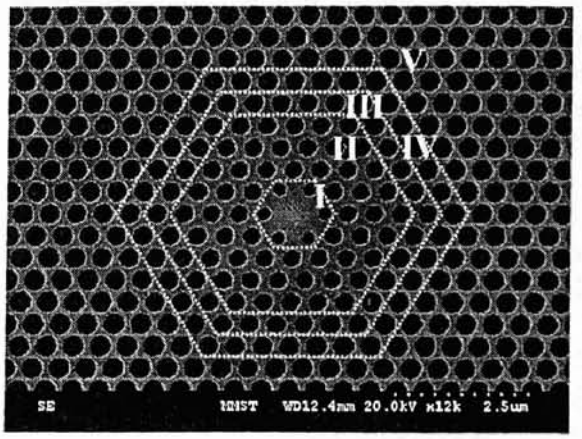
Fig. 16.43(c) visualizes the FDTD-calculated E-field intensity profile (log scale) along a planar central cut through the photonic crystal slab [172]. Note that, in creating the FDTD model, actual structural data for the fabricated microlaser was transferred directly from the scanning electron microscope image to the FDTD geometry dataset. This allowed the model to incorporate actual imperfections of the fabrication process. As a result, the FDTD model predicted the actual observed monopole-mode operation of the laser, and reasonably reproduced the field asymmetry resulting from the laser's imperfect fabrication. Cavity Q factors for this study were in the order of 3,000 for a lasing wavelength of 1,519.7 nm and a modal volume of 0.0587 μm^3. The latter corresponds to 0.684 $(\lambda/n)^3$, which approaches the smallest theoretical value.

The microcavity laser of [172] discussed here has a flexible geometry, which allows fine tuning of its radiation pattern and emission wavelength. The compact size and high spontaneous-emission coupling factor of its defect microcavity also make this laser interesting as a low-noise light source, potentially approaching thresholdless operation or a single-photon source. In addition, such microcavity lasers 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.

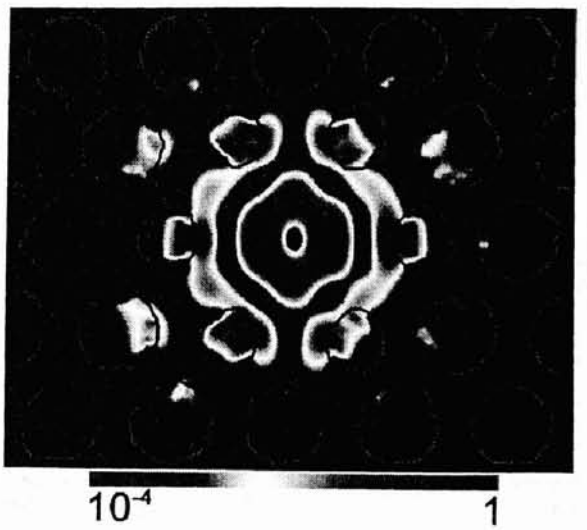
Finally, we note that FDTD can be used to predict the far-field pattern of photonic-crystal lasers for comparison with measured near- and far-field patterns [199, 205, 206]. This can be a useful way to verify that the device operation is well understood. Here, it is important to avoid evanescent contributions to the simulated far-field by taking the flux at a plane some distance above the resonator surface. It is also important to match the finite collection angle inherent in any experimentally measured data [199, 206].



(a)



(b)



(c)

Fig. 16.43 Electrically driven, single-mode, low-threshold-current photonic crystal microlaser operating at room temperature: (a) schematic diagram; (b) top view of the fabricated sample; (c) FDTD-calculated *E*-field intensity profile of the monopole mode (log scale). *Source:* Park et al., *Science*, Sept. 3, 2004, pp. 1444–1447. See Fig. 1.7 of Chapter 1 for the color visualizations.

16.16.2 Photonic Crystal Cross-Waveguide Switch

FDTD has been used to study micron-scale all-optical switching in photonic crystals [187, 188]. Here, photonic crystal microcavities confine photons to submicron dimensions, leading to very high peak power levels which enhance nonlinear processes by many orders of magnitude. This enhancement in nonlinearity can be used to construct ultralow-power all-optical switches.

Fig. 16.44 shows grayscale visualizations of the operation of an all-optical photonic crystal cross-waveguide switch that can operate at 10 Gbits/s with pulse energy levels as small as 100 fJ/pulse. (The corresponding color visualizations are shown in Fig. 1.8.) The switching action employs only the intrinsic nonresonant Kerr nonlinearities that are present in semiconductors such as AlGaAs for wavelengths in the order of 1.5 μm (compatible with modern lightwave communication systems). The structure of Fig. 16.44 consists of two photonic crystal waveguides that cross each other at right angles. The microcavity located at the intersection of the waveguides possesses two orthogonal dipole-like cavity modes. Each cavity mode couples to only one of the waveguides, and isolates the signals in the two waveguides. As proven by the FDTD simulations, this accomplishes both spatial and spectral isolation between the signal and the control inputs, even in the nonlinear regime.

Fig. 16.44 illustrates the two states of the device as simulated with FDTD. Initially, the microcavity at the intersection is designed to be out of resonance with the signal input in the absence of the control input. This causes a very low signal transmission, as shown in Fig. 16.44(a). When a control pulse is launched, as shown in Fig. 16.44(b), it causes a strong electric field buildup in the microcavity, which shifts its resonant frequency due to the material nonlinearity. The overall result is a bistable transition of the cavity resonance, which switches the signal transmission from low to high. Since the switching process occurs via bistable transitions, the device is digital; that is, only low and high transmission states are allowed. As discussed in [187, 188], the same device can function either as a switch or as a memory element, depending upon the design parameters.

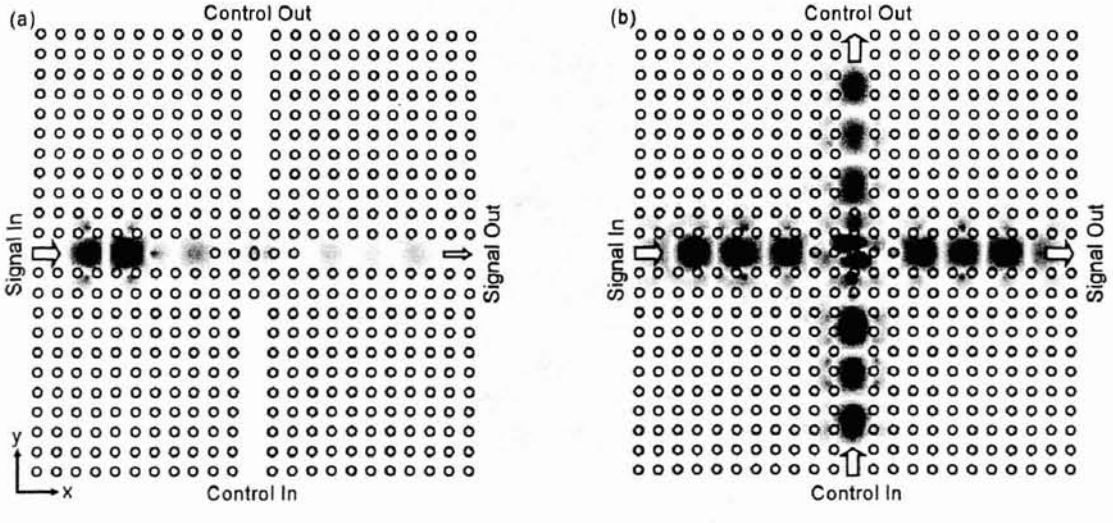


Fig. 16.44 FDTD-calculated electric field distributions in a photonic crystal cross-waveguide switch: (a) control input is absent, and the signal output is low; (b) control input is present, and the signal output is high. Black circles indicate the positions of the dielectric rods in the photonic crystal. *Source:* Yanik et al., *Optics Letters*, 2003, pp. 2506–2508.

16.17 INTRODUCTION TO FREQUENCY CONVERSION IN SECOND-ORDER NONLINEAR OPTICAL MATERIALS

Frequency-conversion processes in second-order ($\chi^{(2)}$) nonlinear materials such as $LiNbO_3$ have been explored for a variety of optical communications, signal-processing, data-storage, and sensing applications that rely upon components such as wavelength routers, add/drop multiplexers, tunable coherent light sources, and all-optical switches. Frequency conversion is extremely sensitive to the phase velocities of interacting optical waves. Therefore, *quasi-phase-matching* (QPM) techniques are needed to compensate for material dispersion that would otherwise degrade the frequency conversion efficiency. First-principles modeling tools based on the full-wave vector Maxwell's equations are useful for the analysis and design of $\chi^{(2)}$ nonlinear optical devices, including QPM structures.

Since the early 1990s, the FDTD method has been refined and extended to treat a wide range of problems involving the interaction of electromagnetic waves with nonlinear materials (see examples in Chapter 9). However, the numerical dispersion characteristics associated with FDTD can be problematic in modeling frequency-conversion processes. Since interacting optical waves with different phase velocities fall out-of-step as they propagate through a nonlinear material, a coarsely resolved FDTD simulation of a homogeneous frequency-independent second-order nonlinear material erroneously predicts a finite coherence length. A computationally inefficient solution to this problem is to use an extremely fine grid resolution to reduce the effects of numerical dispersion. For example, [207] reported that at least 80 points per wavelength at the second harmonic are required to achieve reasonably accurate results in modeling second-harmonic generation.

Reference [208] reported an alternative solution to the problem of modeling phase-sensitive nonlinear optical phenomena. This solution is based on the PSTD method [209, 210], adapted for modeling optical wave propagation in $\chi^{(2)}$ materials. (For detailed discussions of PSTD techniques, see Chapter 4, Section 4.9.4, and Chapter 17.) The spatial derivatives in Maxwell's equations are calculated using the differentiation theorem for Fourier transforms, thereby providing spectral accuracy. Hence, the time-stepping is the only source of numerical dispersion errors. The dispersion errors of the standard second-order-accurate-in-time PSTD algorithm, denoted here as PSTD-2, can be reduced by employing a PSTD scheme with fourth-order time-stepping, denoted here as PSTD-4.

16.18 PSTD-4 ALGORITHM

For simplicity, consider an electromagnetic plane wave propagating in a linear, lossless, nondispersive, and isotropic medium with no electric or magnetic current sources. For this one-dimensional case with E_z polarization and x-directed propagation, the PSTD-4 algorithm is given by the following update equations [208]:

$$H_y\big|_i^{n+1/2} = H_y\big|_i^{n-1/2} + \frac{\Delta t}{\mu_i}\left\{\mathcal{F}_x^{-1}\left[-jk_x\mathcal{F}_x\left(E_z\right)\right]\Big|_i^n + \frac{v_i^{\,2}\Delta t^2}{24}\mathcal{F}_x^{-1}\left[jk_x^{\,3}\mathcal{F}_x\left(E_z\right)\right]\Big|_i^n\right\} \tag{16.27}$$

$$E_z\big|_i^{n+1} = E_z\big|_i^{n} + \frac{\Delta t}{\varepsilon_i}\left\{\mathcal{F}_x^{-1}\left[-jk_x\mathcal{F}_x\left(H_y\right)\right]\Big|_i^{n+1/2} + \frac{v_i^{\,2}\Delta t^2}{24}\mathcal{F}_x^{-1}\left[jk_x^{\,3}\mathcal{F}_x\left(H_y\right)\right]\Big|_i^{n+1/2}\right\} \tag{16.28}$$

where k_x is the Fourier-transform variable representing the x-component of the numerical wave vector, and $\mathcal{F}_x$ and $\mathcal{F}_x^{-1}$ denote, respectively, the forward and inverse DFTs along the x-direction. The complete derivation of this algorithm is given in [208]. Without the last term on the right-hand side of (16.27) and (16.28), the update equations reduce to the PSTD-2 algorithm for one dimension.

An FFT algorithm is used to implement the DFTs in (16.27) and (16.28). Note that one forward and two inverse FFTs are required in PSTD-4 to update either field, whereas only one forward and one inverse FFT is required for PSTD-2. Hence, in one dimension, PSTD-4 requires a total of six FFT subroutine calls per grid cell per time-step, whereas PSTD-2 requires a total of only four. However, PSTD-2 requires a much smaller time-step to achieve the same level of accuracy offered by PSTD-4. Thus, PSTD-4 offers improved computational efficiency relative to PSTD-2, as long as the computational overhead is compensated by the reduction in the total number of time-steps required by PSTD-4 relative to PSTD-2. The break-even point in the computational burden of PSTD-4 versus PSTD-2 occurs when the accuracy requirements of the problem force PSTD-2 to have a time-step that is 1.5, 2.25, and 3.5 times smaller than that used by PSTD-4 in one, two, and three dimensions, respectively.

16.19 EXTENSION TO SECOND-ORDER NONLINEAR MEDIA

For the linear case, one time-step in the PSTD numerical solution consists of first updating H from the previously computed E using (16.27), and then updating E from the previously computed H using (16.28). In contrast, the PSTD scheme for the nonlinear case involves three stages to complete one time-step: (1) updating H from E, (2) updating the electric flux density D from H, and (3) calculating the updated E from D using the nonlinear constitutive relation

$$D = \varepsilon_0 \varepsilon_r E + \varepsilon_0 \chi^{(2)} E^2 \tag{16.29}$$

This equation can be solved directly as follows:

$$E = \frac{-\varepsilon_0 \varepsilon_r + \sqrt{(\varepsilon_0 \varepsilon_r)^2 + 4\varepsilon_0 \chi^{(2)} D}}{2\varepsilon_0 \chi^{(2)}} \tag{16.30}$$

Note that this equation is replaced by the linear constitutive relation $E = D/\varepsilon_0 \varepsilon_r$ in any region of the grid where $\chi^{(2)} = 0$.

16.20 APPLICATION TO A NONLINEAR WAVEGUIDE WITH A QPM GRATING

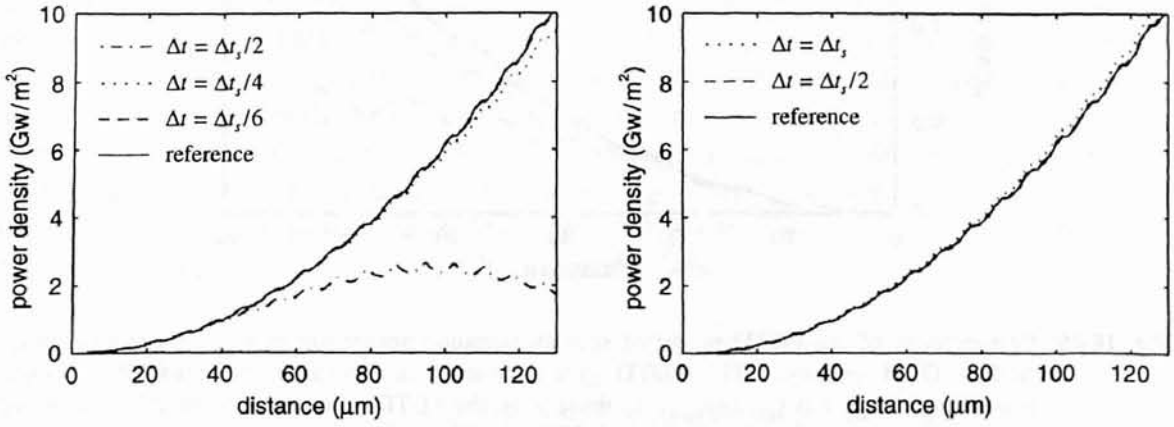
Reference [208] reported the results of numerical experiments designed to test the performance of the PSTD algorithms for modeling *second-harmonic generation* (SHG). Here, a symmetric dielectric slab waveguide with a nonlinear core guiding layer was modeled in a two-dimensional TE PSTD computational domain.

In [208], the dielectric constants of the waveguide core and cladding layers were assumed to be 4.84 and 4.0, respectively. The core was assumed to be 0.682 μm thick, with the nonlinear susceptibility $\chi^{(2)} = 44$ pm/V (corresponding to $LiNbO_3$). A QPM grating (a periodic reversal of the nonlinearity) was modeled in the core. This grating was designed to restore the proper phase relationship between the fundamental wave and the second harmonic in the presence of the waveguide dispersion, thereby improving the efficiency of SHG. The sign (polarization) of $\chi^{(2)}$ was assumed to be constant over the coherence length for the waveguide (7.885 μm, assuming a fundamental wavelength of 1.5 μm).

Reference [208] first compared the numerical convergence of PSTD-2 and PSTD-4 for modeling SHG in a 130-μm length of the nonlinear waveguide, using a grid sampling density of approximately ten points per fundamental wavelength in the core region. The maximum allowed time-step was the two-dimensional stability limit for the linear case [209, 210]:

$$\Delta t_s = \frac{2\sqrt{\varepsilon_r}}{\pi c \sqrt{\Delta x^{-2} + \Delta y^{-2}}} \tag{16.31}$$

where ε_r denotes the dielectric constant of the cladding region. In each simulation, a DFT at the second-harmonic frequency was applied to the fields along the center axis of the waveguide, and the power density was computed. Starting with Δt_s, the PSTD-2 and PSTD-4 simulations were repeated with smaller time-steps until numerical convergence in the power density was reached. Fig. 16.45 shows the results of these convergence tests.



(a) PSTD-2 model.

(b) PSTD-4 model.

Fig. 16.45 Convergence of the PSTD-2 and PSTD-4 models of SHG in a nonlinear waveguide with a QPM grating. *Source:* Lee and Hagness, *J. Opt. Soc. America B*, 2004, pp. 330–342.

Fig. 16.45(a) shows that the power density of the second harmonic was grossly underestimated when PSTD-2 was used with a large time-step, but converged to the reference data ($\Delta t = \Delta t_s/30$) as the time-step decreased. Fig. 16.45(b) shows that PSTD-4 slightly overestimated the power density of the second harmonic when the maximum possible time-step was used, but very quickly converged to the reference data as the time-step decreased.

By way of a quantitative comparison, the time-step required to achieve an error of no more than approximately 2% in the computed second-harmonic power density at a propagation distance of 130 μm was determined to be $\Delta t_s/5$ for PSTD-2 and $\Delta t_s/1.5$ for PSTD-4. With this selection of time-steps, the PSTD-4 model ran about 1.5 times faster than the PSTD-2 model, despite the extra FFTs needed to implement PSTD-4.

Reference [208] then evaluated the numerical convergence of FDTD with respect to grid resolution for modeling second-harmonic generation in the nonlinear waveguide with a QPM grating. Since the numerical dispersion error of FDTD is directly controlled by the grid resolution, the grid-cell size in the FDTD model was gradually decreased until the FDTD results were as accurate as the PSTD-4 results. Fig. 16.46 shows the results of the FDTD convergence tests for a 60-μm propagation distance. The FDTD error drops to the 2% range when its grid-cell size is decreased by a factor of 10:1 relative to that used for PSTD-4. With this choice of grid resolution, the FDTD model using a time-step of Δt_s ran about 51 times slower than the PSTD-4 model using a time-step of $\Delta t_s/1.5$.

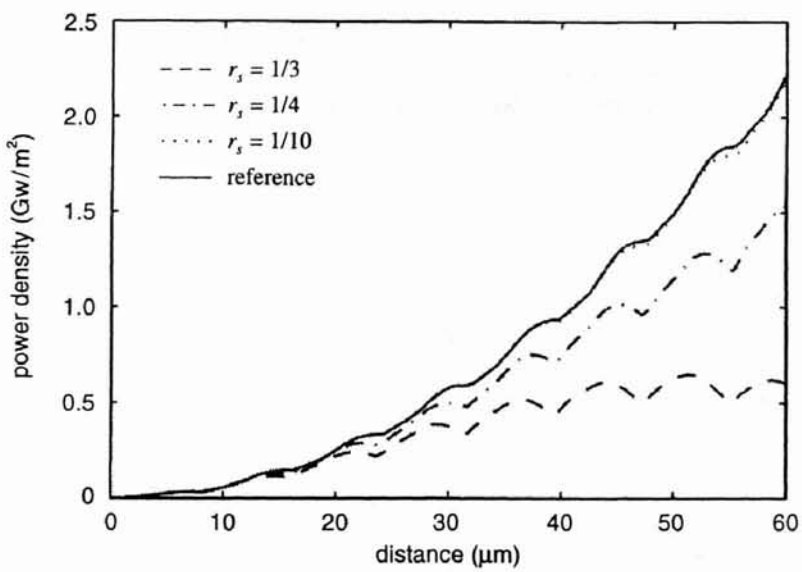


Fig. 16.46 Convergence of the FDTD model of second-harmonic generation in a nonlinear waveguide with a QPM grating. The FDTD grid resolution is expressed in terms of the ratio $r_s \equiv \Delta x_{\text{FDTD}}/\Delta x_{\text{PSTD}} = \Delta y_{\text{FDTD}}/\Delta y_{\text{PSTD}}$. In these tests, the FDTD time-step was set at Δt_s. *Source:* Lee and Hagness, *J. Opt. Soc. America B*, 2004, pp. 330–342.

These benchmark simulations demonstrate that low-dispersion schemes such as PSTD-2 and PSTD-4 offer significant computational advantages over the standard FDTD scheme for modeling phase-sensitive frequency-conversion processes in second-order nonlinear optical materials. Since errors due to numerical dispersion are cumulative, the computational savings of PSTD relative to FDTD increases with the distance of wave propagation. Hence, PSTD techniques (especially PSTD-4) enable much more efficient modeling of frequency conversion in larger-scale nonlinear optical devices, such as the nonlinear photonic crystals to be discussed next.

16.21 APPLICATION TO NONLINEAR PHOTONIC CRYSTALS

A *nonlinear photonic crystal* (NPC) is an artificially engineered material comprised of a periodic spatial variation in a nonlinear susceptibility tensor [211]. An NPC created via two-dimensional periodic poling of the second-order susceptibility has a unique spatio-spectral feature that permits quasi-phase-matching in two dimensions. Thus, in contrast to one-dimensional QPM, the harmonic wave in a two-dimensional NPC can propagate in a different direction than the fundamental wave. This noncollinear propagation characteristic has been demonstrated experimentally for SHG in $LiNbO_3$ NPCs with hexagonal poling patterns [212], and has been exploited for simultaneous wavelength interchange of optical signals [213].

In this section, we review the application of the PSTD-4 technique (introduced in [208] and discussed in the previous section) to model SHG processes in $LiNbO_3$ NPCs, as reported in [214]. The material properties of $LiNbO_3$ were incorporated using an auxiliary differential equation technique [215] adapted for nonlinear PSTD simulations [216]. A two-pole Lorentz dispersion model was developed to account for the linear frequency-dependent susceptibility [216]. The instantaneous nonlinear susceptibility was described by $\chi^{(2)} = +44$ pm/V for the inverted poling region, and $\chi^{(2)} = -44$ pm/V for the surrounding areas.

Fig. 16.47 shows the computational domain assumed in [214] for the case of a two-dimensional NPC lattice structure designed to generate the second harmonic at an angle of 6.3° with respect to the direction of propagation of the fundamental. In this example, a hexagonal poling pattern was assumed at each lattice point of the NPC. The size of the computational domain was 24.2 μm in the y-direction and 270 μm in the x-direction (between the source plane and the far right boundary). The fundamental wave ($\lambda_0 = 1.5$ μm) propagating in the $+x$-direction was launched as a uniform plane wave from the left side of the domain, using the source condition of [217]. UPML absorbing boundary conditions were used to terminate the left and right sides of the computational domain, while periodic boundary conditions were applied at the top and bottom sides of the domain. Once the sinusoidal steady state was reached in the time-domain simulation, discrete Fourier transforms were applied to the PSTD-computed E- and H-fields at the fundamental and second-harmonic frequencies. Then, the spatial variation of the power density and propagation direction of the second-harmonic wave was determined via Poynting vector calculations performed throughout the domain.

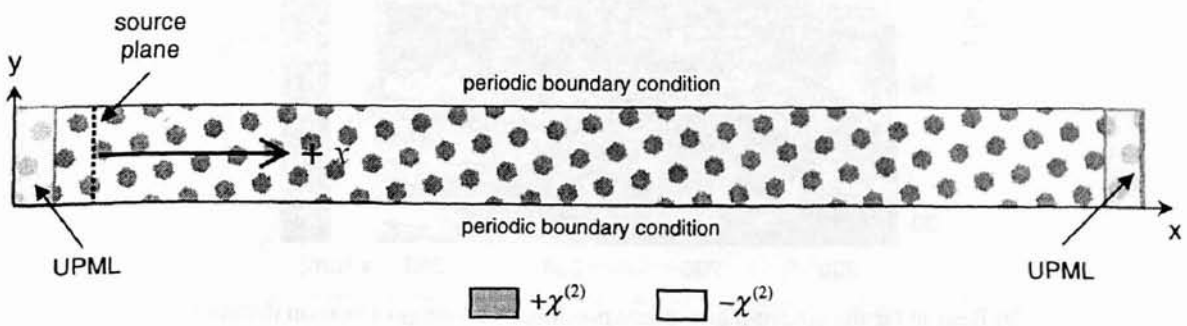
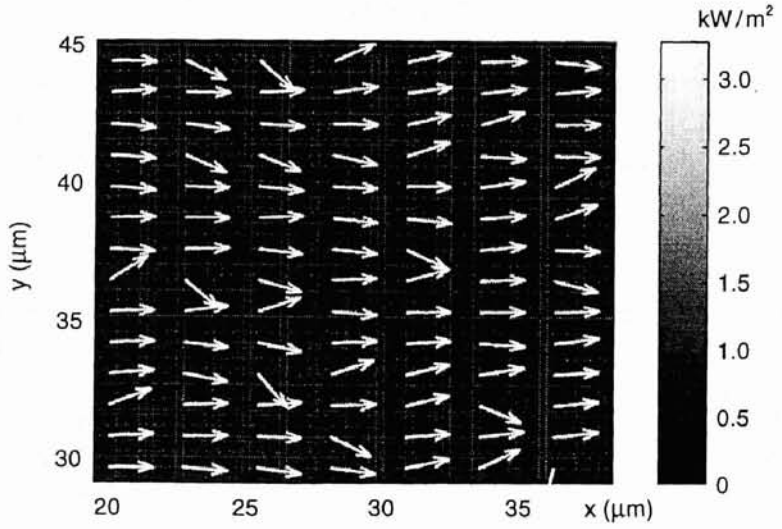


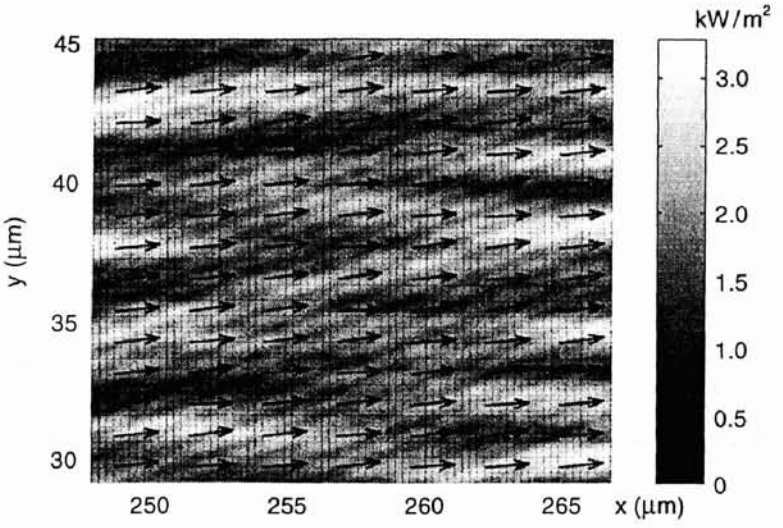
Fig. 16.47 PSTD model of a two-dimensional nonlinear photonic crystal with a hexagonal poling pattern. The thick arrow indicates the direction of incident wave propagation. *Adapted from:* Lee and Hagness, *Proc. IEEE LEOS Annual Meeting*, 2003, Vol. 1, pp. 194–195.

Three different shapes of poling patterns were investigated in [214]: hexagonal, circular, and elliptical. The circular and hexagonal poling patterns were parameterized in terms of the radius r of the circle and the distance h between the center and the side of the hexagon. The elliptical pattern was parameterized in terms of the major axis a and the minor axis b, where the major axis was aligned with the intended direction of propagation for the second harmonic.

Fig. 16.48 shows the PSTD simulation results of [214] for second-harmonic generation in a nonlinear photonic crystal having a circular poling pattern ($r = 2.6$ μm). Here, the magnitude (background grayscale) and the direction (arrows) of the local Poynting vector for the second-harmonic wave are displayed for two regions of the computational domain.



(a) Region near the source plane, corresponding to a short propagation distance.



(b) Region far the source plane, corresponding to a long propagation distance.

Fig. 16.48 Power density (grayscale) and propagation direction (arrows) of the second-harmonic wave computed in two 15-μm-wide regions of the PSTD nonlinear photonic crystal model. A circular poling pattern was assumed. *Adapted from:* Lee and Hagness, *Proc. IEEE LEOS Annual Meeting*, 2003, Vol. 1, pp. 194–195.

Fig. 16.48(a) shows that, close to the source, the second-harmonic wave is weak and quite randomly directed. However, Fig. 16.48(b) shows a much stronger second-harmonic wave far from the source, with Poynting vectors uniformly oriented at the design angle of approximately 6° with respect to the $+x$-axis. These PSTD results illustrate that a finite propagation distance is required for the second-harmonic wave to emerge with a well-defined direction of propagation.

Fig. 16.49 shows the results of [214] for the power densities and average propagation angles of second-harmonic waves generated in nonlinear photonic crystals having the three different poling patterns. Here, the dimensions for each pattern had been optimized for maximum frequency-conversion efficiency. The results shown in Fig. 16.49(a) reveal that the elliptical poling pattern yields the highest frequency-conversion efficiency among these poling patterns. Fig. 16.49(b) shows how the propagation direction of the second-harmonic wave approaches the design angle of 6.3° as a function of distance of propagation along the $+x$-axis.

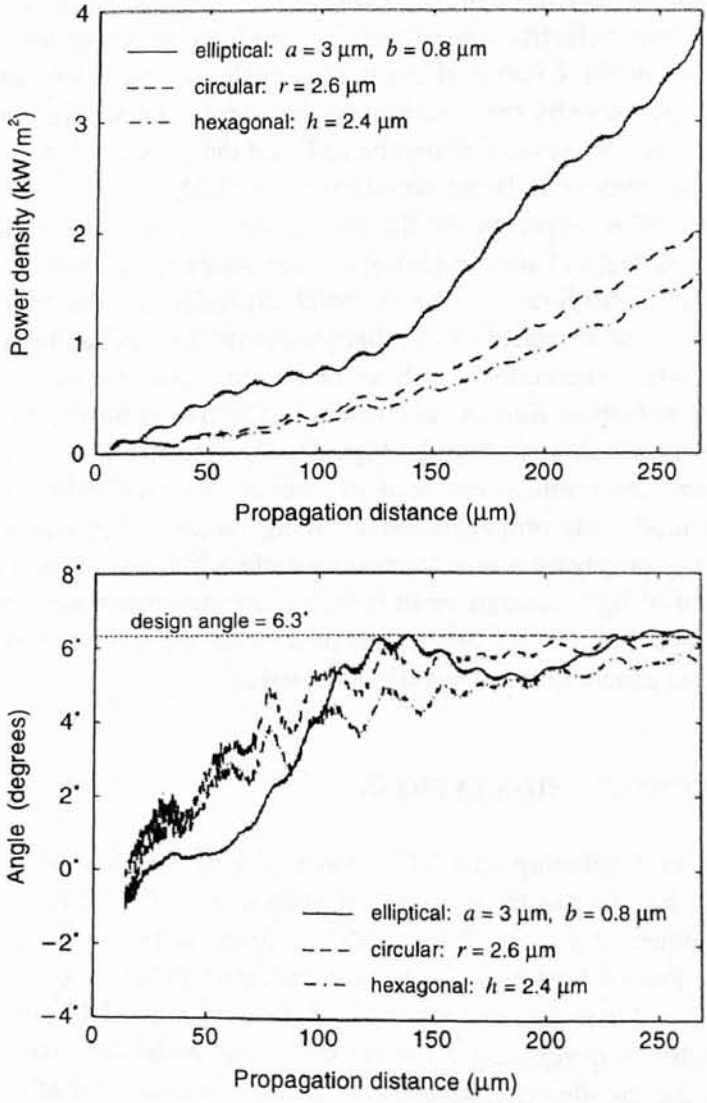


Fig. 16.49 PSTD modeling results for three different optimized poling patterns: (a) power density, and (b) propagation direction of the second-harmonic wave as a function of propagation distance. *Source:* Lee and Hagness, *Proc. IEEE LEOS Annual Meeting*, 2003, Vol. 1, pp. 194–195.

16.22 INTRODUCTION TO NANOPLASMONIC DEVICES

In addition to nanophotonic devices based on dielectric materials, there has been much recent interest in nanoscale devices that exploit the unique features of surface plasmons. Surface plasmons are the joint interaction between the collective oscillation of electrons and the associated electromagnetic wave bound to the surface between two different materials [218–220]. The requirement here is that the two materials have opposite signs in the real part of the dielectric constant, so that the field strength decays away from the interface in both directions [218–220]. Typically, this implies a metal such as gold or silver that interfaces either air, glass, or another dielectric material.

The classical experiment for surface plasmons is excitation via *attenuated total reflection* (ATR). Here, an optical beam is incident on the inside face of a silver-coated glass prism at an incident angle past the critical angle. In one configuration [218–220], evanescent waves from the reflected beam couple energy through the silver film into surface plasmons on the interface between the silver and the dielectric beyond it. By carefully selecting the incidence angle, the transverse component of the k-vector of the light is made to match the spatial frequency of the plasmon. Further, by an optimum selection of the gap thickness, the coupling from the evanescent wave to the surface plasmon is maximized, and the peak field strength at the silver-dielectric interface can be many times larger than the incident field.

Even though such ATR experiments do not necessarily involve nanoscale features, they demonstrate the advantages of surface plasmons: the ability to concentrate and enhance the optical electric field in configurations that depend critically on the local surroundings. The potentially large field enhancement means that plasmonic devices can be used to maximize the strength of light-matter interactions such as two-photon fluorescence, second-harmonic generation, and *surface-enhanced Raman spectroscopy* (SERS). In turn, the tight confinement means one can implement nanolithography and high-density optical storage. Further, the strong dependence on the local environment can lead to sensors of either index or film thickness. Since surface plasmon modes can propagate energy along surfaces, they can be used as optical waveguides. Such energy propagation or collection by surface plasmons may also have a role in the enhanced throughput of light through small holes in nanostructured metallic films. Finally, periodic structures of metallic material can be designed to show photonic bandgaps, to display negative refraction, and to generally act as novel metamaterials.

16.23 FDTD MODELING CONSIDERATIONS

The first consideration in developing an FDTD model of a nanoplasmonic device is properly simulating the dielectric behavior of the materials of interest at optical wavelengths. In general, we can apply the techniques of Chapter 9 for modeling linear dispersive media, especially the Drude dispersion algorithm of Section 9.4.3 for the dielectric behavior of metals. Reference [221] provides an example of matching real experimental data for n and k for this purpose.

A complicating factor in developing a correct dielectric model for metal nanoparticles at optical wavelengths is that the dielectric constant of a metal in a nanoparticle can be perturbed from the baseline (macroscopic-scale) value by the particle size and shape. Such perturbations are caused by changes in the effective mean free path of the electrons in the metal that result from scattering from the nanoparticle surface [222, 223].

The second fundamental consideration in developing an FDTD model of a nanoplasmonic device is properly choosing the spatial resolution. This issue arises because the plasmon wavelength may be much shorter than the free-space optical wavelength. Experience indicates that very fine grid cells of approximately 0.5 nm are needed to accurately simulate the exact angle of plasmon coupling [224], absent any corrections for numerical dispersion [225] and staircasing. Hence, to reduce the computational demands of simulating nanoplasmonic devices with FDTD, it is appropriate to employ subcell techniques, or perhaps more robustly, the nonuniform gridding or subgridding techniques described in Chapter 11, Sections 11.2 and 11.8.

16.24 FDTD MODELING APPLICATIONS

Oscillating charges in a metallic nanoparticle have a resonant interaction with incident light at a frequency that depends on the incident polarization, the particle shape, and the properties of the metal and the surrounding dielectric. Such nanoparticles are of interest because the field enhancement associated with them can be used to enhance interactions between light and nearby molecules [226–228]. Here, FDTD has been used to both understand experiments as well as to engineer the field enhancement via the morphology of the nanoparticle [222, 229–234]. These local field enhancements also can be utilized in lithography, where FDTD has been used to study the near-field intensity of nanoparticles resting on photoresist substrates [235, 236].

Another approach to enhance the local field either for Raman spectroscopy or other nonlinear effects is to use a sharp metal tip, such as on a scanning probe microscope. FDTD has been used to try to understand and optimize the field enhancement with such near-field optical probes [237–243].

Because of the tight localization of light, it is also possible to create a surface plasmon waveguide. This can be accomplished by either coupling the individual plasmon resonances of a line of closely spaced nanoparticles [244, 245], or by coupling into a very narrow metal stripe [246, 247]. Both nanoparticle chain devices [248–250] and metallic-stripe waveguides [251–253] have been studied with FDTD.

A number of studies have used FDTD to help understand and optimize the images generated by near-field microscopes, including the important effects of surface plasmons generated at the sharp metallized tip [254–262]. An important application outside of microscopy for such near-field tips is optical data storage, where FDTD has been used to study punctured tips [263, 264], nanoparticles in an aperture for field localization [265–268], and field enhancement provided by either a triangular metal aperture [269] or a triangular metal plate on a slider [270]. In addition, FDTD has been used to model the readout from super-resolution near-field structures [271–274], wherein silver nanoparticles embedded in one or more of the disk layers act as a virtual near-field probe or aperture.

A topic that has received much attention is the transmission of light through thin metal films perforated with single or multiple holes [275, 276], and containing nanostructure on one or both sides [277]. Essentially, the optical throughput is much higher than would be expected according to the original theory of diffraction through small holes proposed by Bethe [278]. It is believed that the nanostructure on the incident side of the film collects energy from portions of the incident beam outside the holes, either by surface plasmons [279] or other evanescent waves [280]. This energy makes its way through the hole(s) to the shadow side, where it can be reradiated into a beam whose diffraction properties can be affected by the backside nanostructure [281–283]. Many researchers in this area have used FDTD to study these effects [284–291].

16.25 INTRODUCTION TO BIOPHOTONICS

Biophotonics is an exciting emerging discipline that involves theoretical studies and clinical applications of optical interactions with living tissues. Currently, primary applications of biophotonics include the accelerated diagnosis and treatment of human disease.

An important emerging research area in biophotonics with potentially near-term clinical applications is early-stage cancer detection. This involves investigation of possible correlations of the linear ("elastic") light-scattering properties of tissues with precancerous alterations in their cellular composition and nanostructure. Until the introduction of FDTD and PSTD techniques to this field, exploring these correlations was impeded by an inability to robustly and accurately model the electrodynamics of optical interactions with living tissues. In the following two sections, we review recent progress in this area. Topics include: (1) FDTD modeling results showing how optical interactions can be sensitive to submicron, even nanometer-scale features embedded within micron-scale models of living cells; and (2) PSTD modeling results showing promise for studies of optical interactions with random arrangements of hundreds, even thousands of living cells, spanning in aggregate macroscopic tissue regions. FDTD and PSTD techniques are providing means to strengthen the science base for cellular-level and tissue-level biophotonics, and to accelerate the development of corresponding novel clinical technologies.

16.26 FDTD MODELING APPLICATIONS

16.26.1 Vertebrate Retinal Rod

Arguably the first application of FDTD to cellular-level biophotonics was reported in [292], wherein visible light interactions with a retinal photoreceptor were modeled for the two-dimensional TM_z and TE_z polarization cases. The working hypothesis was that the detailed physical structure of a photoreceptor impacts the physics of its optical absorption and thereby, vision. One such photoreceptor was studied: the vertebrate retinal rod. The bulk structure of the retinal rod exhibits the physics of an optical waveguide, while the periodic internal disk-stack structure adds the physics of an optical interferometer. These effects combine to generate a complex optical standing wave within the rod, thereby creating a pattern of local intensifications of the optical field.

The FDTD model of the rod reported in [292] had the cross section dimensions of $2 \times 20\ \mu m$, corresponding to $(3.8\lambda_d - 5.7\lambda_d) \times (38\lambda_d - 57\lambda_d)$ over the range of wavelengths considered, where λ_d denotes the optical wavelength within the rod's dielectric media. A uniform Cartesian space grid having 5.0-nm square unit cells was utilized. This permitted resolution of the 15-nm-thick outer wall membrane of the rod and the 15-nm-thick internal disk membranes. There was a total of 799 disks distributed uniformly along the length of the rod, separated from each other by 10 nm of fluid, and separated from the outer wall membrane by 5 nm of fluid. The index of refraction of the membrane was chosen to be 1.43, and the index of refraction of the fluid was chosen to be 1.36, in accordance with generally accepted physiological data. These parameters implied a resolution within the dielectric media of $\lambda_d/70$ to $\lambda_d/105$, depending on the incident wavelength.

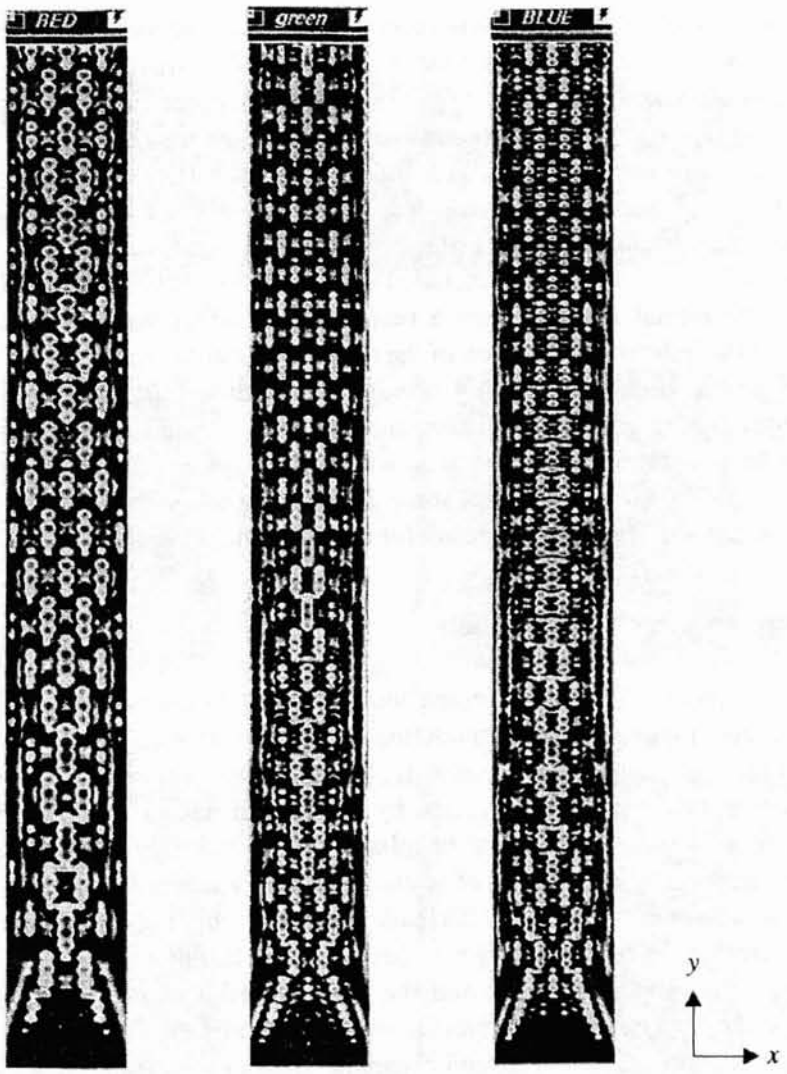


Fig. 16.50 Grayscale visualizations of the FDTD-computed optical E-field standing wave within the retinal rod model for TM_z illumination at free-space wavelengths $\lambda_0 = 714$, 505, and 475 nm. *Source:* Piket-May et al., *Optics Letters,* 1993, pp. 568–570.

As reported in [292], Fig. 16.50 provides grayscale visualizations of the FDTD-computed magnitude of the normalized electric field values of the optical standing wave within the retinal rod model for TM_z illumination at the free-space wavelengths $\lambda_0 = 714$, 505, and 475 nm. Similar visualizations were obtained for the TE_z illumination case.

To assist in understanding the physics of the retinal rod as an optical structure, the standing-wave magnitude data at each λ_0 were reduced as follows. First, at each transverse plane located at a given y_0 in the rod, the electric field values, $E(x, y_0)$, of the optical standing wave were integrated over the x-coordinate to obtain a single number, $E_{int}(y_0)$. Second, a discrete spatial Fourier transform of the set of $E_{int}(y_0)$ values was performed over the y-coordinate. With the exception of isolated peaks unique to each λ_0, the spatial-frequency spectra for each polarization were found to be essentially independent of the illumination wavelength. It was concluded that the retinal rod exhibits a type of frequency-independent electrodynamic behavior.

The agreement of the spatial-frequency spectra for the three incident wavelengths for each polarization was so remarkable that the overall procedure was tested for computational artifacts. The test involved perturbing the indices of refraction of the membrane and fluid from those of the vertebrate rod to those of glass and air, while leaving the geometry unchanged. It was found that the glass-air spectrum exhibited little correlation (i.e., numerous sharp high-amplitude oscillations) over the entire spatial-frequency range considered. On the other hand, the normalized membrane-fluid spectrum varied in a tight range near unity through spatial frequencies of 3.6 μm^{-1}. It was concluded that the agreement of the spatial-frequency spectra for the vertebrate retinal rod indicates a real physical effect that is dependent upon the proper definition of the indices of refraction of the components of the rod structure.

From an electrical engineering standpoint, frequency-independent structures have found major usages in broadband transmission and reception of radio frequency and microwave signals. There is a limited set of such structures, and it is always exciting to find a new one. Reference [292] concluded by speculating that some engineering usage of wavelength-independent retinal-rod-like structures may eventually result for optical signal processing.

16.26.2 Precancerous Cervical Cells

In a series of papers ([293, 294] being most relevant to the present discussion), the Richards-Kortum group pioneered FDTD modeling of light scattering from cervical cells during their earliest stages of cancer development. This group investigated how the light-scattering properties of cervical cells are affected by changes in nuclear morphology, DNA content, and chromatin texture that occur during neoplastic progression. FDTD was applied to calculate the magnitude and angular distribution of scattered light as a function of pathologic grade.

We now consider work by the Richards-Kortum group on two-dimensional FDTD models of cellular scattering, as illustrated in Fig. 16.51 [293]. In this example, the cell cytoplasm, when present, had a diameter of 8 μm, and the nucleus had a diameter of 4 μm. Refractive index values for the cytoplasm and the nucleus were 1.37 and 1.40, respectively. Organelle refractive indices ranged from 1.38 to 1.42, and organelle sizes ranged from 0.1 to 1 μm. Approximately 25% of the available space within the cell (i.e., space not already occupied by the nucleus) was filled with organelles. Wavelengths spanned from 600 to 1,000 nm in 5-nm increments.

From Fig. 16.51, we note how the introduction of heterogeneities in the form of small organelles impacts scattering. Closely following the discussion of [293], the addition of cytoplasmic organelles begins to obscure the interference peaks visible in the simulations using homogeneous geometries. The effects of the heterogeneities are most noticeable at angles over 90°, partially because the scattered intensity values in this region are five to six orders-of-magnitude smaller than the scattered intensity values at low angles.

Reference [293] then proceeded to consider more complicated two-dimensional descriptions of cellular morphology. In the example illustrated in Fig. 16.52, two cells containing multiple sizes and shapes of organelles and heterogeneous nuclei were considered. In the first cell, the morphology was defined using histological features of normal cervical cells. In the second cell, the morphology was defined based on the features of cervical cells staged as high-grade dysplasia. In order to emphasize differences due to the internal contents, both cells were assumed to be circular with 9-μm diameters. The most significant differences between the dysplastic cell relative to the normal cell included increased nuclear size and nuclear-to-cytoplasmic ratio (normal 0.2, dysplastic 0.67), asymmetric nuclear shape, increased DNA content, and hyperchromatic nucleus with areas of coarse chromatin clumping and clearing.

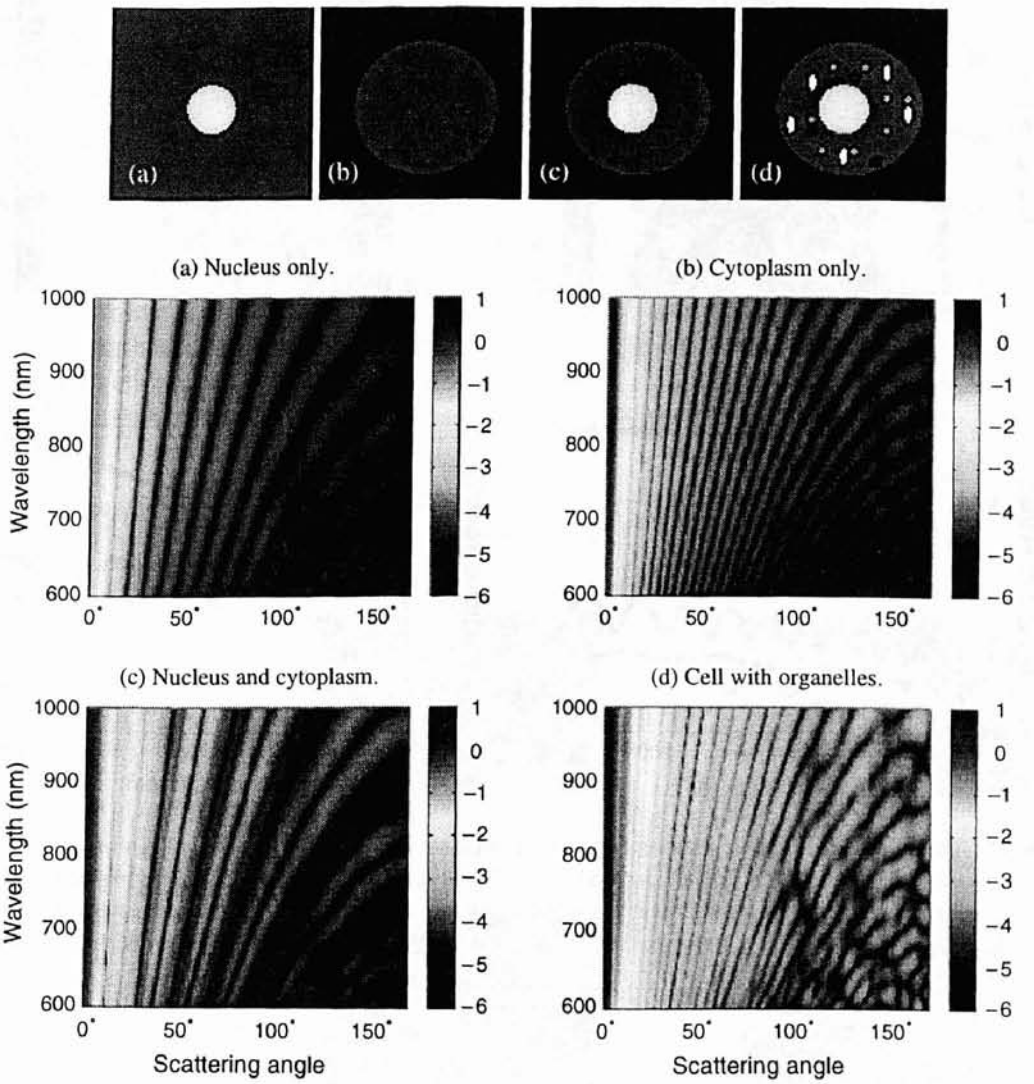


Fig. 16.51 Grayscale visualizations of the FDTD-computed optical scattering of four models of a cell: (a) nucleus only; (b) cytoplasm only; (c) nucleus and cytoplasm; and (d) nucleus and cytoplasm containing organelles. The grayscale corresponds to the log of the scattered intensity. *Source:* Drezek et al., *Optics Express*, March 27, 2000, pp. 147–157.

For the normal cell considered in Fig. 16.52, nuclear refractive-index variations were assumed to be uniformly distributed in the range $n = 1.40 \pm 0.02$ at spatial frequencies ranging from 10 to 30 μm^{-1}, thereby simulating a fine, heterogeneous chromatin structure [293]. In the dysplastic cell, nuclear refractive-index variations were distributed in the range $n = 1.42 \pm 0.04$ at spatial frequencies ranging from 3 to 30 μm^{-1}, thereby simulating a coarser, more heterogeneous chromatin structure. Both normal and dysplastic cells contained several hundred organelles (radii from 50 to 500 nm; $n = 1.38$ to 1.40) randomly distributed throughout the cytoplasm.

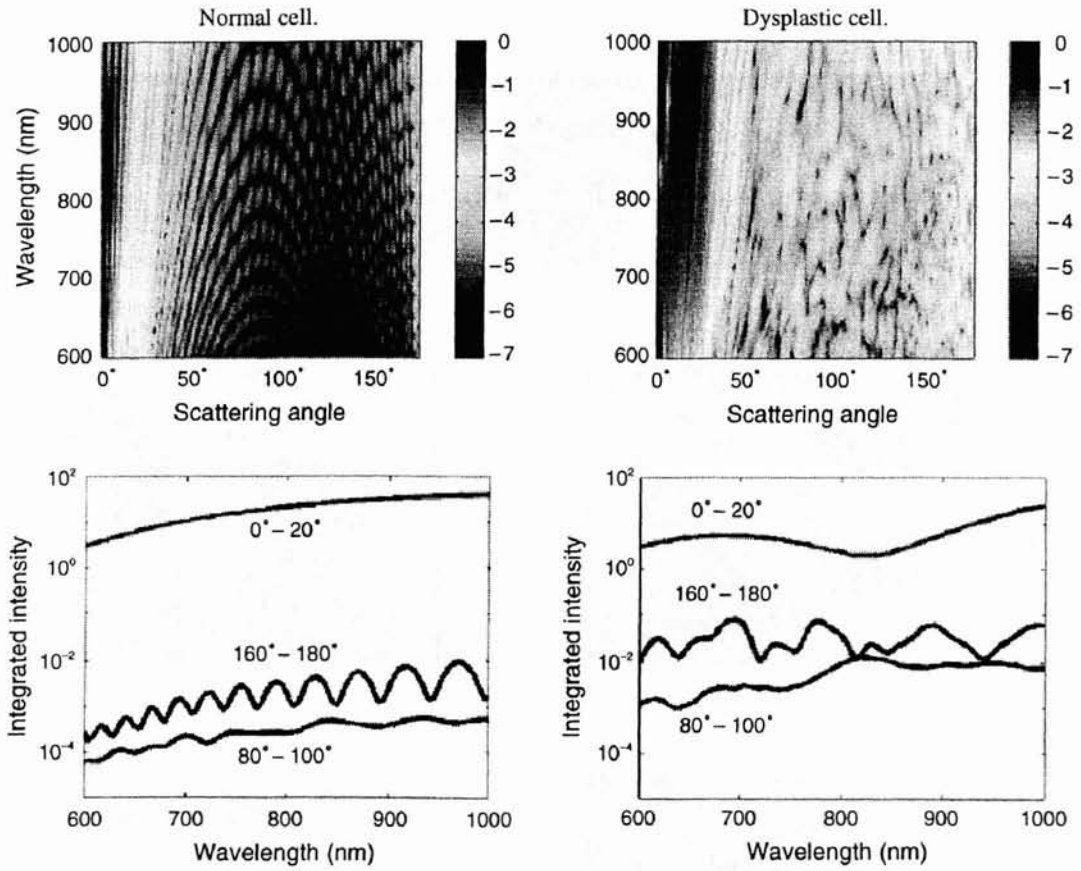


Fig. 16.52 Top: Grayscale visualizations of the FDTD-computed optical scattering from models of normal (left) and dysplastic (right) cervical cells. The grayscale corresponds to the log of the scattered intensity. Bottom: Integrated scattered intensities over three angular ranges for normal (left) and dysplastic (right) cervical cells. *Source:* Drezek et al., *Optics Express*, March 27, 2000, pp. 147–157.

Fig. 16.52 shows the results of this FDTD modeling study. Closely following the discussion of [293], the dysplastic cell exhibits elevated scattering, with increased scattering at small angles due to the larger nucleus, and with increased scattering at larger angles due to alterations in the chromatin structure, which results in increased heterogeneity of the refractive index. Since the dysplastic cell contains a large heterogeneous nucleus that is comprised of an assortment of scatterer sizes and refractive indexes, distinct interference peaks are not present. Although heterogeneities are present in the structure of a normal cell, they are not significant enough to disrupt the peaks resulting from the cytoplasm and nuclear boundaries.

The bottom half of Fig. 16.52 displays the integrated scattered intensity as a function of wavelength for three angular ranges: 0° to 20°, 80° to 100°, and 160° to 180°. Closely following the discussion of [293], these results show that the integrated intensity is a function of both angle and cellular structure. Here, changes in the wavelength dependence of the scattering between the normal and dysplastic cells are especially evident at large angles. To develop optimized optical probes and measurement techniques that can discriminate between normal and dysplastic tissue based on differences in the wavelength dependence of cellular scattering, it is important to be aware of which angular regions offer the greatest potential for differential diagnosis.

16.26.3 Sensitivity of Backscattering Signatures to Nanometer-Scale Cellular Changes

Recent experimental evidence indicates that light-scattering signals can provide means for ultra-early-stage detection of colon cancer [295] before any other biomarker that is currently known. In combination with the findings reported by the Richards-Kortum group, it is now quite clear that light scattering is very sensitive to minute differences in tissue and cellular structures. An important question then arises for researchers investigating optical tissue diagnostic techniques: Which light-scattering parameters provide the best sensitivity to detect cellular changes that are at the nanometer scale (i.e., those that may indicate cancer)?

Fig. 16.53 illustrates the application of FDTD to evaluate the sensitivity of optical backscattering and forward-scattering signatures to refractive index fluctuations spanning nanometer length scales [296]. Here, the spectral and angular distributions of scattered light from inhomogeneous dielectric particles with identical sizes and volume-averaged refractive indices are compared with corresponding data calculated for their homogeneous counterparts.

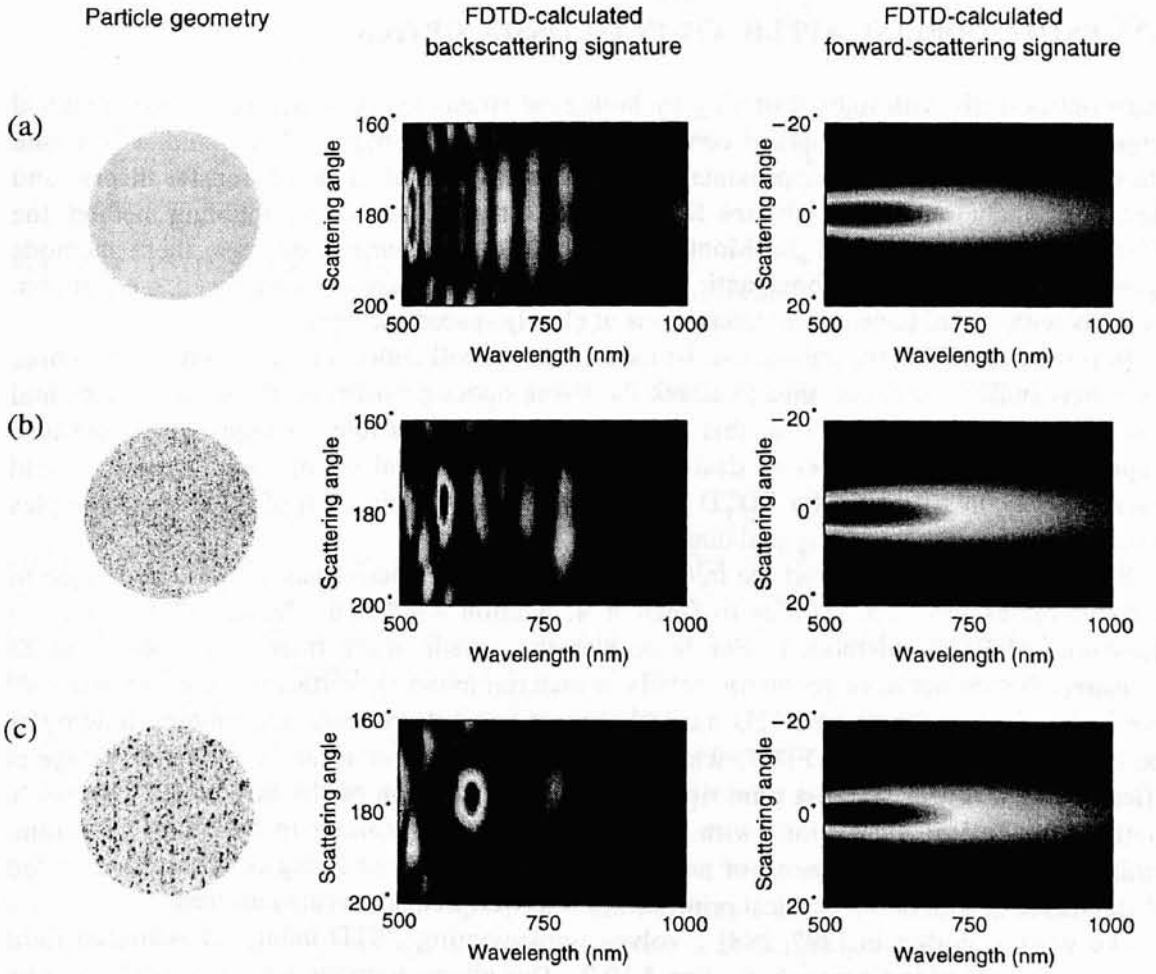


Fig. 16.53 Grayscale visualizations of the FDTD-computed optical scattering signatures of a 4-μm-diameter particle with a volume-averaged refractive index $n_{avg} = 1.1$. (a) homogeneous particle; (b) inhomogeneous particle with refractive index fluctuations $\Delta n = \pm 0.03$ spanning distance scales of approximately 50 nm; and (c) inhomogeneous particle with refractive index fluctuations $\Delta n = \pm 0.03$ spanning distance scales of approximately 100 nm.

The optical backscattering signatures (shown in the center panels of Fig. 16.53) are of particular interest. These are grayscale visualizations of the FDTD-calculated backscattering intensity distributions as functions of wavelength and scattering angle within a $\pm 20^{\circ}$ range of direct backscatter. Relative to the homogeneous case of Fig. 16.53(a), we observe distinctive features of the backscattering signatures for the randomly inhomogeneous cases of Figs. 16.53(b, c). This is despite the fact that the inhomogeneities for these cases have characteristic sizes of only 50 and 100 nm, respectively, which are much smaller than the illumination wavelength of 750 nm. In contrast, the forward-scattering signatures shown in the right-hand panels exhibit no distinctive features.

These FDTD calculations strongly support the hypothesis that there exists signatures in backscattered light that are sufficiently sensitive to detect alterations in the cellular architecture at the nanometer scale. Importantly, this sensitivity is *not bound by the diffraction limit.* Potentially, backscattering signatures can serve as biomarkers to detect and characterize slight alterations in tissue structure which may be precursors of cancer [295].

16.27 PSTD MODELING APPLICATION TO TISSUE OPTICS

Tissue optics deals with light scattering by biological structures, on which noninvasive optical imaging techniques such as optical coherence tomography are based. Most studies of tissue optics have utilized heuristic approximations in the categories of radiative transfer theory and Mie theory, including Beer's law, the Kubelka-Monk theory, the adding-doubling method, the diffusion approximation, and the Monte Carlo method. To various degrees, these methods neglect the full-vector electromagnetic wave nature of light based on Maxwell's equations, especially with regard to near-field interactions of closely spaced particles.

In principle, FDTD techniques can be used to model cell collections spanning macroscopic dimensions (millimeters) and thus to attack the tissue-optics problem on the most fundamental basis. However, using FDTD for this purpose may not be feasible for many years, because computers lack the capabilities to deal with the enormous database of electromagnetic field vector components mandated for FDTD by its mesh-density requirement of 20 or more samples per optical wavelength in each spatial dimension.

References [297, 298] report the initial application of the Fourier-based PSTD technique to the tissue-optics problem. (Refer to Chapter 4, Section 4.9.4 and Chapter 17 for detailed discussions of PSTD methods.) For large electromagnetic wave interaction models in D dimensions that do not have geometric details or material inhomogeneities smaller than one-half wavelength, the Fourier-based PSTD method reduces computer storage and running time by at least 8^D:1 relative to standard FDTD while it achieves comparable accuracy. This advantage is sufficient to permit for the first time rigorous numerical solution of the full-vector Maxwell's equations for optical interactions with dielectric structures spanning in the order of 1 mm. At this scale, random arrangements of hundreds, even thousands, of living cells can be modeled with assurance that all of the physical principles of Maxwell's equations are enforced.

The work reported in [297, 298] involves implementing PSTD using the scattered-field formulation reviewed in Chapter 5, Section 5.10.2. This allows sourcing a plane wave having an arbitrary angle of incidence, polarization, and time waveform. The total E- and H-fields can be obtained in a postprocessing step by adding the known (analytical) incident fields to the PSTD-computed scattered fields. In implementing these studies, it was determined that temporal convergence of the PSTD calculations requires that the time-step must be less than 1/60th of the sinusoidal period at the maximum frequency of interest [298].

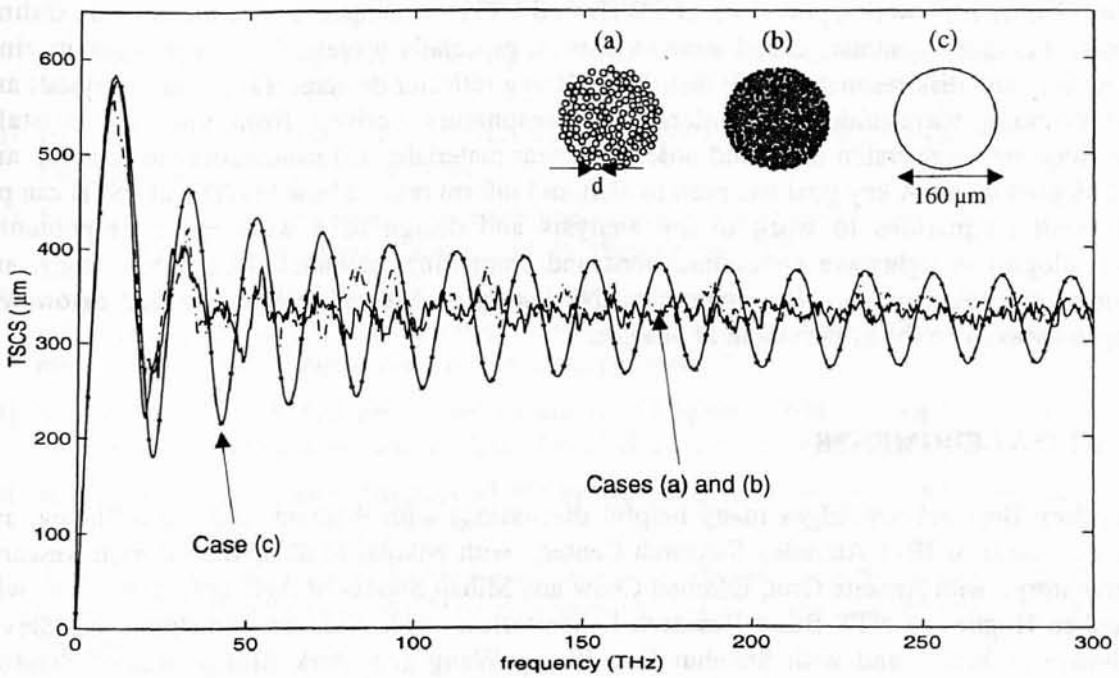


Fig. 16.54 PSTD-computed total scattering cross section of: (a) 160-μm overall-diameter cylindrical bundle of 120 randomly positioned, noncontacting dielectric cylinders of individual diameter $d = 10$ μm and refractive index $n = 1.2$; (b) as in (a), but for 480 cylinders of individual diameter $d = 5$ μm; and (c) single cylinder of $n = 1.0938$, the volume-average refractive index of the random bundles of (a) and (b). *Source:* Tseng et al., *Optics Letters*, 2005, pp. 56–57.

Fig. 16.54 illustrates a principal finding of [297, 298], namely that when the average dielectric coverage of a bundle of cylindrical scatterers has increased beyond a certain threshold, the *total scattering cross section* (TSCS) of this bundle becomes independent of its internal geometric details such as the size, position, and number of its constituent cylinders. In this regime, the frequency dependence (spectrum) of the TSCS of the cylinder bundle represents essentially the average behavior of the TSCS spectrum of the volume-averaged homogeneous cylinder of the same diameter. The primary difference is that the homogeneous cylinder exhibits ripples of its TSCS spectrum as a result of coherent internal wave-interference effects that are suppressed by scattering events within the random clusters.

The results reported in [297, 298] point toward the emerging feasibility of direct, exact Maxwell's equations modeling of light propagation through, and scattering by, millimeters of biological tissues. More generally, these results have a wider implication. Namely, the study of electromagnetic wave propagation within random media is moving toward exact rather than approximate solutions of Maxwell's equations. The driving force behind this fundamental advance is the advent of robust PSTD algorithms, which permit using space grids having coarse resolutions approaching one-half wavelength. PSTD achieves spectral accuracy, thereby yielding numerical dispersion errors approaching theoretically minimum values.

16.28 SUMMARY

This chapter reviewed applications of FDTD and PSTD techniques to photonics in six distinct areas: (1) index-contrast guided-wave structures, especially waveguides and microcavity ring, racetrack, and disk resonators; (2) distributed Bragg reflector devices; (3) photonic crystals and defect-mode waveguides and microcavity resonators derived from photonic crystals; (4) frequency conversion in second-order nonlinear materials; (5) nanoplasmonic devices; and (6) biophotonics. A key goal has been to alert and inform readers how FDTD and PSTD can put Maxwell's equations to work in the analysis and design of a wide range of photonics technologies in lightwave communications and computing, nanometer-scale technology, and biomedical applications. The comprehensive listing of 298 references provided below will hopefully assist in the achievement of this goal.

ACKNOWLEDGMENTS

Geoffrey Burr acknowledges many helpful discussions with William Risk, Bob Shelby, and Bülent Kurdi of IBM Almaden Research Center; with Nikolaj Moll of IBM Zurich Research Laboratory; with Annette Grot, Edmond Chow and Mihail Sigalas of Agilent Laboratories; with Stephen Hughes of NTT Basic Research Laboratories; with Ardavan Farjadpour and Steven Johnson of MIT; and with Shanhui Fan, Zheng Wang and Mark Brongersma of Stanford University. Susan Hagness acknowledges contributions from Tae-Woo Lee of Argonne National Laboratory and A. Chowdhury and L. McCaughan of the University of Wisconsin-Madison. Allen Taflove acknowledges key collaborations with Vadim Backman, Jay Walsh, Xu Li, Snow Tseng, and Zhigang Chen of Northwestern University.

REFERENCES

[1] Rafizadeh, D., J. P. Zhang, S. C. Hagness, A. Taflove, K. A. Stair, S. T. Ho, and R. C. Tiberio, "Waveguide-coupled AlGaAs / GaAs microcavity ring and disk resonators with high finesse and 21.6-nm free spectral range," *Optics Lett.*, Vol. 22, 1997, pp. 1244–1246.

[2] Blom, F. C., D. R. van Dijk, H. J. W. M. Hoekstra, A. Driessen, and Th. J. A. Popma, "Experimental study of integrated-optics microcavity resonators: Toward an all-optical switching device," *Applied Physics Lett.*, Vol. 71, 1997, pp. 747–749.

[3] Yamamoto, Y., and R. Slusher, "Optical processes in microcavities," *Physics Today*, Vol. 46, 1993, pp. 66–73.

[4] Fan, S., P. R. Villeneuve, J. D. Joannopoulos, and E. F. Schubert, "High extraction efficiency of spontaneous emission from slabs of photonic crystals," *Phys. Rev. Lett.*, Vol. 78, 1997, pp. 3294–3297.

[5] Baba, T., "Photonic crystals and microdisk cavities based on GaInAsP-InP systems," *IEEE J. Selected Topics Quantum Electronics*, Vol. 3, 1997, pp. 808–830.

[6] Chu, S.-T., and S. K. Chaudhuri, "A finite-difference time-domain method for the design and analysis of guided-wave optical structures," *J. Lightwave Tech.*, Vol. 7, 1989, pp. 2033–2038.

[7] Berenger, J.-P., "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comp. Phys.*, Vol. 114, 1994, pp. 185–200.

[8] Yamauchi, J., M. Mita, S. Aoki, and H. Nakano, "Analysis of antireflection coatings using the FDTD method with the PML ABC," *IEEE Photonics Tech. Lett.*, Vol. 8, 1996, pp. 239–241.

[9] 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 Guided Wave Lett.*, Vol. 4, 1994, pp. 344–346.

[10] Kogelnik, H., and V. Ramaswamy, "Scaling rules for thin-film optical waveguides," *Applied Optics*, Vol. 13, 1974, pp. 1857–1862.

[11] Luebbers, R. J., and F. Hunsberger, "FDTD for Nth-order dispersive media," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 1297–1301.

[12] Kelley, D. F., and R. J. Luebbers, "Piecewise linear recursive convolution for dispersive media using FDTD," *IEEE Trans. Antennas Propagat.*, Vol. 44, 1996, pp. 792–797.

[13] 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 Lett.*, Vol. 16, 1991, pp. 1412–1414.

[14] Korner, T. O., and W. Fichtner, "Auxiliary differential equation: Efficient implementation in the finite-difference time-domain method," *Optics Lett.*, Vol. 22, 1997, pp. 1586–1588.

[15] Sullivan, D. M., "Frequency-dependent FDTD methods using Z transforms," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 1223–1230.

[16] Goorjian, P. M., A. Taflove, R. M. Joseph, and S. C. Hagness, "Computational modeling of femtosecond optical solitons from Maxwell's equations," *IEEE J. Quantum Electronics*, Vol. 28, 1992, pp. 2416–2422.

[17] Ziolkowski, R. W., and J. B. Judkins, "Full-wave vector Maxwell's equations modeling of the self-focusing of ultrashort pulses in a nonlinear Kerr medium exhibiting a finite response time," *J. Opt. Soc. America B*, Vol. 10, 1993, pp. 186–198.

[18] Joseph, R. M., and A. Taflove, "FDTD Maxwell's equations modeling for nonlinear electrodynamics and optics," *IEEE Trans. Antennas Propagat.*, Vol. 45, 1997, pp. 364–374.

[19] Sullivan, D. M., "Nonlinear FDTD formulations using Z transforms," *IEEE Trans. Microwave Theory Tech.*, Vol. 43, 1995, pp. 676–692.

[20] Joseph, R. M., P. M. Goorjian, and A. Taflove, "Direct time integration of Maxwell's equations in two-dimensional dielectric waveguides for propagation and scattering of femtosecond electromagnetic solitons," *Optics Lett.*, Vol. 18, 1993, pp. 491–493.

[21] Joseph, R. M., and A. Taflove, "Spatial soliton deflection mechanism indicated by FDTD Maxwell's equations modeling," *IEEE Photonic Tech. Lett.*, Vol. 6, 1994, pp. 1251–1254.

[22] Ziolkowski, R. W., and J. B. Judkins, "Applications of the nonlinear finite difference time domain (NL-FDTD) method to pulse propagation in nonlinear media: Self-focusing and linear-nonlinear interfaces," *Radio Science*, Vol. 28, 1993, pp. 901–911.

[23] Ziolkowski, R. W., and J. B. Judkins, "Nonlinear finite-difference time-domain modeling of linear and nonlinear corrugated waveguides," *J. Opt. Soc. America B*, Vol. 11, 1994, pp. 1565–1575.

[24] Basinger, S. A., and D. J. Brady, "Finite-difference time-domain modeling of dispersive nonlinear Fabry-Perot cavities," *J. Opt. Soc. America B*, Vol. 11, 1994, pp. 1504–1511.

[25] Hawkins, R. J., and J. S. Kallman, "Lasing in tilted-waveguide semiconductor laser amplifiers," *Optical Quantum Electronics*, Vol. 26, 1994, pp. S207–S217.

[26] Hagness, S. C., R. M. Joseph, and A. Taflove, "Subpicosecond electrodynamics of distributed Bragg reflector microlasers: Results from finite difference time domain simulations," *Radio Science*, Vol. 31, 1996, pp. 931–941.

[27] Ziolkowski, R. W., J. M. Arnold, and D. M. Gogny, "Ultrafast pulse interactions with two-level atoms," *Phys. Rev. A*, Vol. 52, 1995, pp. 3082–3094.

[28] Huang, Y., *Simulation of Semiconductor Materials Using FDTD Method*, M.S. thesis, Northwestern University, Evanston, IL, May 2002.

[29] Chang, S.-H., and A. Taflove, "Finite-difference time-domain model of lasing action in a four-level two-electron atomic system," *Optics Express*, Vol. 12, 2004, pp. 3827–3833.

[30] Oda, K., N. Takato, and H. Toba, "A wide-FSR waveguide double-ring resonator for optical FDM transmission systems," *J. Lightwave Tech.*, Vol. 9, 1991, pp. 728–736.

[31] Hagness, S. C., D. Rafizadeh, S. T. Ho, and A. Taflove, "FDTD microcavity simulations: Design and experimental realization of waveguide-coupled single-mode ring and whispering-gallery-mode disk resonators," *J. Lightwave Tech.*, Vol. 15, 1997, pp. 2154–2165.

[32] Little, B. E., J.-P. Laine, and S. T. Chu, "Surface-roughness-induced contradirectional coupling in ring and disk resonators," *Optics Lett.*, Vol. 22, 1997, pp. 4–6.

[33] Little, B. E., and S. T. Chu, "Estimating surface-roughness loss and output coupling in microdisk resonators," *Optics Lett.*, Vol. 21, 1996, pp. 1390–1392.

[34] Slusher, R. E., A. F. J. Levi, U. Mohideen, S. L. McCall, S. J. Pearton, and R. A. Logan, "Threshold characteristics of semiconductor microdisk lasers," *Applied Physics Lett.*, Vol. 63, 1993, pp. 1310–1312.

[35] Griffel, G., and S. Arnold, "Synthesis of variable optical filters using meso-optical ring arrays," *IEEE LEOS Annual Meeting*, San Francisco, CA, Nov. 1997, paper WT4.

[36] Hagness, S. C., D. Rafizadeh, S. T. Ho, and A. Taflove, "FDTD analysis and comparison of circular and elongated ring designs for waveguide-coupled microcavity ring resonators," in *Integrated Photonics Research*, Opt. Soc. America Technical Digest Series, 1998.

[37] Hagness, S. C., D. Rafizadeh, S. T. Ho, and A. Taflove, "Suppression of higher-order radial whispering gallery modes in waveguide-coupled microcavity disk resonators," in *Proc. IEEE LEOS Annual Meeting*, San Francisco, CA, Nov. 1997, pp. 160–161.

[38] Greene, J. H., and A. Taflove, "Initial three-dimensional FDTD phenomenology study of the transient response of a large vertically coupled photonic racetrack," *Optics Lett.*, Vol. 28, 2003, pp. 1733-1735.

[39] Corzine, S. W., R. S. Geels, J. W. Scott, R. H. Yan, and L. A. Coldren, "Design of Fabry-Perot surface-emitting lasers with a periodic gain structure," *IEEE J. Quantum Electronics*, Vol. 25, 1989, pp. 1513–1524.

[40] Hagness, S. C., S. T. Ho, and A. Taflove, "Finite-difference time-domain (FDTD) computational electrodynamics simulations of microlaser cavities in one and two spatial dimensions," in *Computational Electromagnetics and Its Applications*, T. G. Campbell, R. A. Nicolaides, and M. D. Salas, (eds.), Dordrecht, Netherlands: Kluwer, 1997, pp. 229–251.

[41] Villeneuve, P. R., S. Fan, J. D. Joannopoulos, K. Y. Lim, G. S. Petrich, L. A. Kolodziejski, and R. Reif, "Air-bridge microcavities," *Applied Physics Lett.*, Vol. 67, 1995, pp. 167–169.

[42] Zhang, J. P., D. Y. Chu, S. L. Wu, W. G. Bi, R. C. Tiberio, R. M. Joseph, A. Taflove, C. W. Tu, and S. T. Ho, "Nanofabrication of 1D photonic bandgap structures along a photonic wire," *IEEE Photonic Tech. Lett.*, Vol. 8, 1996, pp. 491–493.

[43] Ziolkowski, R. W., and T. Liang, "Design and characterization of a grating-assisted coupler enhanced by a photonic bandgap structure for effective wavelength-division demultiplexing," *Optics Lett.*, Vol. 22, 1997, pp. 1033–1035.

[44] Villeneuve, P. R., D. S. Abrams, S. Fan, and J. D. Joannopoulos, "Single-mode waveguide microcavity for fast optical switching," *Optics Lett.*, Vol. 21, 1996, pp. 2017–2019.

[45] Chen, J. C., H. A. Haus, S. Fan, P. R. Villeneuve, and J. D. Joannopoulos, "Optical filters from photonic band gap air bridges," *J. Lightwave Tech.*, Vol. 14, 1996, pp. 2575–2580.

[46] John, S., "Strong localization of photons in certain disordered dielectric superlattices," *Phys. Rev. Lett.*, Vol. 58, 1987, pp. 2486–2489.

[47] Yablonovitch, E., "Inhibited spontaneous emission in solid-state physics and electronics," *Phys. Rev. Lett.*, Vol. 58, 1987, pp. 2059–2062.

[48] Joannopoulos, J. D., R. D. Meade, and J. N. Winn, *Photonic Crystals: Molding the Flow of Light*, Princeton, N.J.: Princeton University Press, 1995.

[49] Johnson, S. G., and J. D. Joannopoulos, *Photonic Crystals: The Road from Theory to Practice*, Norwell, MA: Kluwer, 2002.

[50] [online]: www.pbglink.com

[51] Johnson, S. G., and J. D. Joannopoulos, "Block-iterative frequency-domain methods for Maxwell's equations in a plane wave basis," *Optics Express*, Vol. 8, 2001, pp. 173–190.

[52] [online]: http://ab-initio.mit.edu/mpb/

[53] Chan, C. T., Q. L. Yu, and K. M. Ho, "Order-N spectral method for electromagnetic waves," *Phys. Rev. B*, Vol. 51, 1995, pp. 16635–16642.

[54] Ward, A. J., and J. B. Pendry, "Calculating photonic Green functions using a nonorthogonal finite-difference time-domain method," *Phys. Rev. B*, Vol. 58, 1998, pp. 7252–7259.

[55] Wu, L., and S. He, "Revised finite difference time domain algorithm in a nonorthogonal coordinate system and its application to the computation of the band structure of a photonic crystal," *J. Applied Phys.*, Vol. 91, 2002, pp. 6499–6506.

[56] Sakoda, K., and J. Kawamata, "Novel approach to photonic bands with frequency-dependent dielectric constants," *Optics Express*, Vol. 3, 1998, pp. 12–18.

[57] Arriaga, J., A. J. Ward, and J. B. Pendry, "Order-N photonic band structures for metals and other dispersive materials," *Phys. Rev. B.*, Vol. 59, 1999, pp. 1874–1877.

[58] Sakoda, K., N. Kawai, T. Ito, A. Chutinan, S. Noda, T. Mitsuyu, and K. Hirao, "Photonic bands of metallic systems: 1. Principle of calculation and accuracy," *Phys. Rev. B.*, Vol. 64, 2001, 045116.

[59] Choi, D. H., and W. J. R. Hoefer, "The finite-difference time-domain method and its application to eigenvalue problems," *IEEE Trans. Microwave Theory Tech.*, Vol. 34, 1986, pp. 1464–1470.

[60] Hess, O., C. Hermann, and A. Klaedtke, "Finite-difference time-domain simulations of photonic crystal defect structures," *Physica Status Solidi A*, Vol. 197, 2003, pp. 605–619.

[61] Prescott, D. T., and N. V. Shuley, "Extensions to the FDTD method for the analysis of infinitely periodic arrays," *IEEE Microwave Guided Wave Lett.*, Vol. 4, 1994, pp. 352–354.

[62] Celuch-Marcysiak, M., and W. K. Gwarek, "Spatially looped algorithms for time-domain analysis of periodic structures," *IEEE Trans. Microwave Theory Tech.*, Vol. 43, 1995, pp. 860–865.

[63] Sakoda, K., and H. Shiroma, "Numerical method for localized defect modes in photonic lattices," *Phys. Rev. B*, Vol. 56, 1997, pp. 4830–4835.

[64] Sakoda, K., T. Ueta, and K. Ohtaka, "Numerical analysis of eigenmodes localized at line defects in photonic lattices," *Phys. Rev. B*, Vol. 56, 1997, pp. 14905–14908.

[65] Wall, M. R., and D. Neuhauser, "Extraction, through filter-diagonalization of general quantum eigenvalues or classical normal-mode frequencies from a small number of residues or a short-time segment of a signal: 1. Theory and application to a quantum-dynamics model," *J. Chem. Phys.*, Vol. 102, 1995, pp. 8011–8022.

[66] Mandelshtam, V. A., and H. S. Taylor, "Harmonic inversion of time signals and its applications," *J. Chem. Phys.*, Vol. 107, 1997, pp. 6756–6769.

[67] Chen, R. Q., and H. Guo, "Efficient calculation of matrix elements in low storage filter diagonalization," *J. Chem. Phys.* Vol. 111, 1999, pp. 464–471.

[68] Johnson, S. G., S. Fan, A. Mekis, and J. D. Joannopoulos, "Multiple-cancellation mechanism for high-Q cavities in the absence of a complete photonic band gap," *Applied Phys. Lett.*, Vol. 78, 2001, pp. 3388–3390.

[69] [online]: http://ab-initio.mit.edu/harminv/

[70] Meade, R. D., K. D. Brommer, A. M. Rappe, and J. D. Joannopoulos, "Existence of a photonic band-gap in two dimensions," *Applied Phys. Lett.*, Vol. 61, 1992, pp. 495–497.

[71] Fan, S. H., P. R. Villeneuve, and J. D. Joannopoulos, "Large omnidirectional bandgaps in metallodielectric photonic crystals," *Phys. Rev. B*, Vol. 54, 1996, pp. 11245–11251.

[72] Qiu, M., and S. L. He, "A nonorthogonal finite-difference time-domain method for computing the band structure of a two-dimensional photonic crystal with dielectric and metallic inclusions," *J. Appl. Phys.*, Vol. 87, 2000, pp. 8268–8275.

[73] Farjadpour, A., and G. W. Burr, "Accurate FDTD simulation of dielectric nanophotonics structures despite coarse gridding," in preparation.

[74] Kuang, W., C. Kim, A. Stapleton, W. J. Kim, and J. D. O'Brien, "Calculated out-of-plane transmission loss for photonic-crystal slab waveguides," *Optics Lett.*, Vol. 28, 2003, pp. 1781–1783.

[75] Juntunen, J. S., and T. D. Tsiboukis, "Reduction of numerical dispersion in FDTD method through artificial anisotropy," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 582–588.

[76] Yu, W. H., and R. Mittra, "A technique for improving the accuracy of the nonuniform finite-difference time-domain algorithm," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, 1999, pp. 353–356.

[77] Kaneda, N., B. Houshmand, and T. Itoh, "FDTD analysis of dielectric resonators with curved surfaces," *IEEE Trans. Microwave Theory Tech.*, Vol. 45, 1997, pp. 1645–1649.

[78] Hirono, T., Y. Shibata, W. W. Lui, S. Seki, and Y. Yoshikuni, "The second-order condition for the dielectric interface orthogonal to the Yee-lattice axis in the FDTD scheme," *IEEE Microwave Guided Wave Lett.*, Vol. 10, 2000, pp. 359–361.

[79] Hwang, K. P., and A. C. Cangellaris, "Effective permittivities for second-order accurate FDTD equations at dielectric interfaces," *IEEE Microwave Wireless Components Lett.*, Vol. 11, 2001, pp. 158–160.

[80] Dey, S., and R. Mittra, "A conformal finite-difference time-domain technique for modeling cylindrical dielectric resonators," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, 1999, pp. 1737–1739.

[81] Yu, W. H., and R. Mittra, "A conformal finite-difference time-domain technique for modeling curved dielectric surfaces," *IEEE Microwave Wireless Components Lett.*, Vol. 11, 2001, pp. 25–27.

[82] Nadobny, J., D. Sullivan, W. Wlodarczyk, P. Deuflhard, and P. Wust, "A 3D tensor FDTD formulation for treatment of sloped interfaces in electrically inhomogeneous media," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 1760–1770.

[83] Burr, G. W., and A. Farjadpour, "Accurate FDTD simulation of dielectric nanophotonics structures despite truncated computations," in preparation.

[84] Ohtera, Y., K. Kurokawa, and S. Kawakami, "Variational expression for the analysis of photonic crystal devices," *IEEE J. Quantum Electronics*, Vol. 38, 2002, pp. 919–926.

[85] Ohtera, Y., Y. Sasaki, and S. Kawakami, "Postprocessing of FDTD solutions for precise calculations of eigenfrequencies of photonic periodic structures utilizing the variational expression," *J. Lightwave Tech.*, Vol. 22, 2004, pp. 1628–1636.

[86] Jackson, J. D., *Classical Electrodynamics, 3rd ed.*, New York: Wiley, 1999.

[87] De Moerloose, J., and D. DeZutter, "Poynting's theorem for the finite-difference time-domain method," *Microwave Optical Tech. Lett.*, Vol. 8, 1995, pp. 257–260.

[88] Nadobny, J., D. Sullivan, P. Wust, M. Seebass, P. Deuflhard, and R. Felix, "A high-resolution interpolation at arbitrary interfaces for the FDTD method," *IEEE Trans. Microwave Theory Tech.*, Vol. 46, 1998, pp. 1759–1766.

[89] McNab, S. J., N. Moll, and Y. A. Vlasov, "Ultralow loss photonic integrated circuit with membrane-type photonic crystal waveguides," *Optics Express*, Vol. 11, 2003, pp. 2927–2939.

[90] Sugimoto, Y., Y. Tanaka, N. Ikeda, K. Kanamoto, Y. Nakamura, S. Ohkouchi, H. Nakamura, K. Inoue, H. Sasaki, Y. Watanabe, K. Ishida, H. Ishikawa, and K. Asakawa, "Two dimensional semiconductor-based photonic crystal slab waveguides for ultra-fast optical signal processing devices," *IEICE Trans. Electronics*, Vol. E87C, 2004, pp. 316–327.

[91] Mekis, A., J. C. Chen, I. Kurland, S. H. Fan, P. R. Villeneuve, and J. D. Joannopoulos, "High transmission through sharp bends in photonic crystal waveguides," *Phys. Rev. Lett.*, Vol. 77, 1996, pp. 3787–3790.

[92] Lin, S. Y., E. Chow, V. Hietala, P. R. Villeneuve, and J. D. Joannopoulos, "Experimental demonstration of guiding and bending of electromagnetic waves in a photonic crystal," *Science*, Vol. 282, 1998, pp. 274–276.

[93] Mekis, A., S. H. Fan, and J. D. Joannopoulos, "Bound states in photonic crystal waveguides and waveguide bends," *Phys. Rev. B.*, Vol. 58, 1998, pp. 4809–4817.

[94] Tokushima, M., H. Kosaka, A. Tomita, and H. Yamada, "Lightwave propagation through a 120° sharply bent single-line-defect photonic crystal waveguide," *Appl. Phys. Lett.*, Vol. 76, 2000, pp. 952–954.

[95] Chow, E., S. Y. Lin, J. R. Wendt, S. G. Johnson, and J. D. Joannopoulos, "Quantitative analysis of bending efficiency in photonic-crystal waveguide bends at $\lambda = 1.55$ μm wavelengths," *Optics Lett.*, Vol. 26, 2001, pp. 286–288.

[96] Espinola, R. L., R. U. Ahmad, F. Pizzuto, M. J. Steel, and R. M. Osgood, "A study of high-index-contrast 90° waveguide bend structures," *Optics Express*, Vol. 8, 2001, pp. 517–528.

[97] Fan, S. H., P. R. Villeneuve, and J. D. Joannopoulos, "Channel drop tunneling through localized states," *Phys. Rev. Lett.*, Vol. 80, 1998, pp. 960–963.

[98] Fan, S. H., P. R. Villeneuve, J. D. Joannopoulos, and H. A. Haus, "Channel drop filters in photonic crystals," *Optics Express*, Vol. 3, 1998, pp. 4–11.

[99] Fan, S. H., P. R. Villeneuve, J. D. Joannopoulos, M. J. Khan, C. Manolatou, and H. A. Haus, "Theoretical analysis of channel drop tunneling processes," *Phys. Rev. B.*, Vol. 59, 1999, pp. 15882–15892.

[100] Noda, S., A. Chutinan, and M. Imada, "Trapping and emission of photons by a single defect in a photonic bandgap structure," *Nature*, Vol. 407, 2000, pp. 608–610.

[101] Koshiba, M., "Wavelength division multiplexing and demultiplexing with photonic crystal waveguide couplers," *J. Lightwave Tech.*, Vol. 19, 2001, pp. 1970–1975.

[102] Sharkawy, A., S. Y. Shi, and D. W. Prather, "Multichannel wavelength division multiplexing with photonic crystals," *Appl. Optics*, Vol. 40, 2001, pp. 2247–2252.

[103] Takano, K., and K. Nakagawa, "Frequency analysis of wavelength demultiplexers and optical filters with finite 2D photonic crystals," *IEICE Trans. Electronics*, Vol. E84C, 2001, pp. 669–677.

[104] Bayindir, M., and E. Ozbay, "Band-dropping via coupled photonic crystal waveguides," *Optics Express*, Vol. 10, 2002, pp. 1279–1284.

[105] Yariv, A., Y. Xu, R. K. Lee, and A. Scherer, "Coupled-resonator optical waveguide: A proposal and analysis," *Optics Lett.*, Vol. 24, 1999, pp. 711–713.

[106] Notomi, M., K. Yamada, A. Shinya, J. Takahashi, C. Takahashi, and I. Yokohama, "Extremely large group-velocity dispersion of line-defect waveguides in photonic crystal slabs," *Phys. Rev. Lett.*, Vol. 87, 2001, 253902.

[107] Soljacic, M., S. G. Johnson, S. H. Fan., M. Ibanescu, E. Ippen, and J. D. Joannopoulos, "Photonic-crystal slow-light enhancement of nonlinear phase sensitivity," *J. Opt. Soc. America B*, Vol. 19, 2002, pp. 2052– 2059.

[108] Sugimoto, Y., S. Lan, S. Nishikawa, N. Ikeda, H. Ishikawa, and K. Asakawa, "Design and fabrication of impurity band-based photonic crystal waveguides for optical delay lines," *Appl. Phys. Lett.*, Vol. 81, 2002, pp. 1946–1948.

[109] Johnson, S. G., P. R. Villeneuve, S. H. Fan, and J. D. Joannopoulos, "Linear waveguides in photonic-crystal slabs," *Phys. Rev. B*, Vol. 62, 2000, pp. 8212–8222.

[110] Assefa, S., P. T. Rakich, P. Bienstman, S. G. Johnson, G. S. Petrich, J. D. Joannopoulos, L. A. Kolodziejski, E. P. Ippen, and H. I. Smith, "Guiding 1.5 μm light in photonic crystals based on dielectric rods," *Appl. Phys. Lett.*, Vol. 85, 2004, pp. 6110–6112.

[111] Loncar, M., T. Doll, J. Vuckovic, and A. Scherer, "Design and fabrication of silicon photonic crystal optical waveguides," *J. Lightwave Tech.*, Vol. 18, 2000, pp. 1402–1411.

[112] Johnson, S. G., S. H. Fan, P. R. Villeneuve, J. D. Joannopoulos, and L. A. Kolodziejski, "Guided modes in photonic crystal slabs," *Phys. Rev. B*, Vol. 60, 1999, pp. 5751–5758.

[113] Loncar, M., J. Vuckovic, and A. Scherer, "Methods for controlling positions of guided modes of photonic-crystal waveguides," *J. Opt. Soc. America B*, Vol. 18, 2001, pp. 1362–1368.

[114] Ochiai, T., and K. Sakoda, "Dispersion relation and optical transmittance of a hexagonal photonic crystal slab," *Phys. Rev. B*, Vol. 63, 2001, 125107.

[115] Roden, J. A., and S. D. Gedney, "Convolutional PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media," *Microwave Optical Tech. Lett.*, Vol. 27, 2000, pp. 334–339.

[116] Gedney, S. D., "An anisotropic perfectly matched layer absorbing medium for the truncation of FDTD lattices," *IEEE Trans. Antennas Propagat.*, Vol. 44, 1996, pp. 1630–1639.

[117] D'Urso, B., O. Painter, J. O'Brien, T. Tombrello, A. Yariv, and A. Scherer, "Modal reflectivity in finite-depth two-dimensional photonic-crystal microcavities," *J. Opt. Soc. America B*, Vol. 15, 1998, pp. 1155–1159.

[118] Baba, T., A. Motegi, T. Iwai, N. Fukaya, Y. Watanabe, and A. Sakai, "Light propagation characteristics of straight single-line-defect waveguides in photonic crystal slabs fabricated into a silicon-on-insulator substrate," *IEEE J. Quantum Electronics*, Vol. 38, 2002, pp. 743–752.

[119] Adibi, A., Y. Xu, R. K. Lee, M. Loncar, A. Yariv, and A. Scherer, "Role of distributed Bragg reflection in photonic-crystal optical waveguides," *Phys. Rev. B*, Vol. 64, 2001, 041102.

[120] Adibi, A., Y. Xu, R. K. Lee, A. Yariv, and A. Scherer, "Guiding mechanisms in dielectric-core photonic-crystal optical waveguides," *Phys. Rev. B*, Vol. 64, 2001, 033308.

[121] Langtry, T. N., A. A. Asatryan, L. C. Botten, C. M. de Sterke, R. C. Mcphedran, and P. A. Robinson, "Effects of disorder in two-dimensional photonic crystal waveguides," *Phys. Rev. E*, Vol. 68, 2003, 026611.

[122] Frei, W. R., and H. T. Johnson, "Finite-element analysis of disorder effects in photonic crystals," *Phys. Rev. B*, Vol. 70, 2004, 165116.

[123] Gerace, D., and L. C. Andreani, "Disorder-induced losses in photonic crystal waveguides with line defects," *Optics Lett.*, Vol. 29, 2004, pp. 1897–1899.

[124] Hughes, S., L. Ramunno, J. F. Young, and J. E. Sipe, "Extrinsic optical scattering loss in photonic crystal waveguides: Role of fabrication disorder and photon group velocity," *Phys. Rev. Lett.*, Vol. 94, 2005, 033903.

[125] Ferrini, R., B. Lombardet, B. Wild, R. Houdre, and G. H. Duan, "Hole depth- and shape-induced radiation losses in two-dimensional photonic crystals," *Appl. Phys. Lett.*, Vol. 82, 2003, pp. 1009–1011.

[126] Tanaka, Y., T. Asano, Y. Akahane, B. S. Song, and S. Noda, "Theoretical investigation of a two-dimensional photonic crystal slab with truncated cone air holes," *Appl. Phys. Lett.*, Vol. 82, 2003, pp. 1661–1663.

[127] Bogaerts, W., P. Bienstman, and R. Baets, "Scattering at sidewall roughness in photonic crystal slabs," *Optics Lett.*, Vol. 28, 2003, pp. 689–691.

[128] Povinelli, M. L., S. G. Johnson, E. Lidorikis, J. D. Joannopoulos, and M. Soljacic, "Effect of a photonic band gap on scattering from waveguide disorder," *Appl. Phys. Lett.*, Vol. 84, 2004, pp. 3639–3641.

[129] Hadley, G. R., "Out-of-plane losses of line-defect photonic crystal waveguides," *IEEE Photonics Tech. Lett.*, Vol. 14, 2002, pp. 642–644.

[130] Ferrini, R., R. Houdre, H. Benisty, M. Qiu, and J. Moosburger, "Radiation losses in planar photonic crystals: Two-dimensional representation of hole depth and shape by an imaginary dielectric constant," *J. Opt. Soc. America B*, Vol. 20, 2003, pp. 469–478.

[131] Witzens, J., T. Baehr-Jones, M. Hochberg, M. Loncar, and A. Scherer, "Photonic crystal waveguide-mode orthogonality conditions and computation of intrinsic waveguide losses," *J. Opt. Soc. America A*, Vol. 20, 2003, pp. 1963–1968.

[132] Lavrinenko, A., P. I. Borel, L. H. Frandsen, M. Thorhauge, A. Harpoth, M. Kristensen, T. Niemi, and H. M. H. Chong, "Comprehensive FDTD modeling of photonic crystal waveguide components," *Optics Express*, Vol. 12, 2004, pp. 234–248.

[133] Desieres, Y., T. Benyattou, R. Orobtchouk, A. Morand, P. Benech, C. Grillet, C. Seassal, X. Letartre, P. Rojo-Romeo, and P. Viktorovitch, "Propagation losses of the fundamental mode in a single line-defect photonic crystal waveguide on an InP membrane," *J. Appl. Phys.*, Vol. 92, 2002, pp. 2227–2234.

[134] Li, Z. Y., and K. M. Ho, "Light propagation in semi-infinite photonic crystals and related waveguide structures," *Phys. Rev. B*, Vol. 68, 2003, 155101.

[135] Morand, A., C. Robinson, Y. Desieres, T. Benyattou, P. Benech, O. Jacquin, and M. Le Vassor d'yerville, "3D numerical modeling of propagation losses of a single line-defect photonic crystal," *Optics Com.*, Vol. 221, 2003, pp. 353–357.

[136] Li, Z. Y., and K. M. Ho, "Anomalous propagation loss in photonic crystal waveguides," *Phys. Rev. Lett.*, Vol. 92, 2004, 063904.

[137] Bogaerts, W., P. Bienstman, D. Taillaert, R. Baets, and D. De Zutter, "Out-of-plane scattering in photonic crystal slabs," *IEEE Photonics Tech. Lett.*, Vol. 13, 2001, pp. 565–567.

[138] Lalanne, P., "Electromagnetic analysis of photonic crystal waveguides operating above the light cone," *IEEE J. Quantum Electron.*, Vol. 38, 2002, pp. 800–804.

[139] Olivier, S., H. Benisty, C. Weisbuch, C. J. M. Smith, T. F. Krauss, and R. Houdre, "Coupled-mode theory and propagation losses in photonic crystal waveguides," *Optics Express*, Vol. 11, 2003, pp. 1490–1496.

[140] Sauvan, C., P. Lalanne, J. C. Rodier, J. P. Hugonin, and A. Talneau, "Accurate modeling of line defect photonic crystal waveguides," *IEEE Photonics Tech. Lett.*, Vol. 15, 2003, pp. 1243–1245.

[141] Choi, J. M., W. Liang, Y. Xu, and A. Yariv, "Loss optimization of transverse Bragg resonance waveguides," *J. Opt. Soc. America A*, Vol. 21, 2004, pp. 426–429.

[142] Andreani, L. C., and M. Agio, "Photonic bands and gap maps in a photonic crystal slab," *IEEE J. Quantum Electron.*, Vol. 38, 2002, pp. 891–898.

[143] Andreani, L. C., "Photonic bands and radiation losses in photonic crystal waveguides," *Physica Status Solidi B*, Vol. 234, 2002, pp. 139–146.

[144] Benisty, H., D. Labilloy, C. Weisbuch, C. J. M. Smith, T. F. Krauss, D. Cassagne, A. Beraud, and C. Jouanin, "Radiation losses of waveguide-based two-dimensional photonic crystals: Positive role of the substrate," *Appl. Phys. Lett.*, Vol. 76, 2000, pp. 532–534.

[145] Benisty, H., P. Lalanne, S. Olivier, M. Rattier, C. Weisbuch, C. J. M. Smith, T. F. Krauss, C. Jouanin, and D. Cassagne, "Finite-depth and intrinsic losses in vertically etched two-dimensional photonic crystals," *Optical Quantum Electron.*, Vol. 34, 2002, pp. 205–215.

[146] Ferrini, R., A. Berrier, L. A. Dunbar, R. Houdre, M. Mulot, S. Anand, S. de Rossi, and A. Talneau, "Minimization of out-of-plane losses in planar photonic crystals by optimizing the vertical waveguide," *Appl. Phys. Lett.*, Vol. 85, 2004, pp. 3998–4000.

[147] Villeneuve, P. R., S. H. Fan, and J. D. Joannopoulos, "Microcavities in photonic crystals: Mode symmetry, tunability, and coupling efficiency," *Phys. Rev. B.*, Vol. 54, 1996, pp. 7837–7842.

[148] Moll, N., and G. L. Bona, "Comparison of three-dimensional photonic crystal slab waveguides with two-dimensional photonic crystal waveguides: Efficient butt coupling into these photonic crystal waveguides," *J. Appl. Phys.*, Vol. 93, 2003, pp. 4986–4991.

[149] Yariv, A., *Optical Electronics, 4th ed.*, Saunders, 1991.

[150] Mekis, A., S. H. Fan, and J. D. Joannopoulos, "Absorbing boundary condition for FDTD simulations of photonic crystal waveguides," *IEEE Microwave Guided Wave Lett.*, Vol. 9, 1999, pp. 502–504.

[151] Koshiba, M., Y. Tsuji, and S. Sasaki, "High-performance absorbing boundary conditions for photonic crystal waveguide simulations," *IEEE Microwave Wireless Components Lett.*, Vol. 11, 2001, pp. 152–154.

[152] Okano, M., S. Kako, and S. Noda, "Coupling between a point defect cavity and a line defect waveguide in three dimensional photonic crystal," *Phys. Rev. B*, Vol. 68, 2003, 235110.

[153] Weily, A. R., L. Horvath, K. P. Esselle, and B. C. Sanders, "Performance of PML absorbing boundary conditions in 3D photonic crystal waveguides," *Microwave Optical Tech. Lett.*, Vol. 40, 2004, pp. 1–3.

[154] Chow, E., A. Grot, L. W. Mirkarimi, M. Sigalas, and G. Girolami, "Ultracompact biochemical sensor built with two-dimensional photonic crystal microcavity," *Optics Lett.*, Vol. 29, 2004, pp. 1093–1095.

[155] Lidorikis, E., M. L. Povinelli, S. G. Johnson, and J. D. Joannopoulos, "Polarization independent linear waveguides in 3D photonic crystals," *Phys. Rev. Lett.*, Vol. 91, 2003, 023902.

[156] Jafarpour, A., A. Adibi, Y. Xu, and R. K. Lee, "Mode dispersion in biperiodic photonic crystal waveguides," *Phys. Rev. B.*, Vol. 68, 2003, 233102.

[157] Jafarpour, A., C. M. Reinke, A. Adibi, Y. Xu, and R. K. Lee, "A new method for the calculation of the dispersion of nonperiodic photonic crystal waveguides," *IEEE J. Quantum Electron.*, Vol. 40, 2004, pp. 1060–1067.

[158] Jafarpour, A., E. Chow, C. M. Reinke, J. Huang, A. Adibi, A. Grot, L. W. Mirkarimi, G. Girolami, R. K. Lee, and Y. Xu, "Large-bandwidth ultra-low-loss guiding in bi-periodic photonic crystal waveguides," *Appl. Phys. B*, Vol. 79, 2004, pp. 409–414.

[159] Hughes, S., personal communication to G. Burr.

[160] Hughes, S., L. Ramunno, E. Kuramochi, T. Watanabe, A. Shinya, and M. Notomi, "Disorder-induced optical scattering loss in planar photonic crystal waveguides: Theory and experiment," *CLEO 2005*, submitted.

[161] Akahane, Y., M. Mochizuki, T. Asano, Y. Tanaka, and S. Noda, "Design of a channel drop filter by using a donor-type cavity with high quality factor in a two-dimensional photonic crystal slab," *Appl. Phys. Lett.*, Vol. 82, 2003, pp. 1341–1343.

[162] Akahane, Y., T. Asano, and S. Noda, "Investigation of high-Q channel drop filters using donor type defects in two dimensional photonic crystal slabs," *Appl. Phys. Lett.*, Vol. 83, 2003, pp. 1512–1514.

[163] Asano, T., M. Mochizuki, S. Noda, M. Okano, and M. Imada, "A channel drop filter using a single defect in a 2-D photonic crystal slab—Defect engineering with respect to polarization mode and ratio of emissions from upper and lower sides," *J. Lightwave Tech.*, Vol. 21, 2003, pp. 1370–1376.

[164] Asano, T., B. S. Song, Y. Tanaka, and S. Noda, "Investigation of a channel add/drop filtering device using acceptor type point defects in a two dimensional photonic crystal slab," *Appl. Phys. Lett.*, Vol. 83, 2003, pp. 407–409.

[165] Takano, H., Y. Akahane, T. Asano, and S. Noda, "In plane type channel drop filter in a two-dimensional photonic crystal slab," *Appl. Phys. Lett.*, Vol. 84, 2004, pp. 2226–2228.

[166] Painter, O., J. Vuckovic, and A. Scherer, "Defect modes of a two-dimensional photonic crystal in an optically thin dielectric slab," *J. Opt. Soc. America B*, Vol. 16, 1999, pp. 275–285.

[167] Painter, O., K. Srinivasan, J. D. O'Brien, A. Scherer, and P. D. Dapkus, "Tailoring of the resonant mode properties of optical nanocavities in two-dimensional photonic crystal slab waveguides," *J. Optics A—Pure Appl. Optics*, Vol. 3, 2001, pp. S161–S170.

[168] Ryu, H. Y., H. G. Park, and Y. H. Lee, "Two-dimensional photonic crystal semiconductor lasers: Computational design, fabrication, and characterization," *IEEE J. Selected Topics Quantum Electron.*, Vol. 8, 2002, pp. 891–908.

[169] Srinivasan, K., and O. Painter, "Momentum space design of high-Q photonic crystal optical cavities," *Optics Express*, Vol. 10, 2002, pp. 670–684.

[170] Srinivasan, K., and O. Painter, "Fourier space design of high-Q cavities in standard and compressed hexagonal lattice photonic crystals," *Optics Express*, Vol. 11, 2003, pp. 579–593.

[171] Painter, O., and K. Srinivasan, "Localized defect states in two-dimensional photonic crystal slab waveguides: A simple model based upon symmetry analysis," *Phys. Rev. B*, Vol. 68, 2003, 035110.

[172] Park, H.-G., S.-H. Kim, S.-H. Kwon, Y.-G. Ju, J.-K. Yang, J.-H. Baek, S.-B. Kim, and Y.-H. Lee, "Electrically driven single-cell photonic crystal laser," *Science*, Vol. 305, 2004, pp. 1444–1447.

[173] Vuckovic, J., M. Loncar, H. Mabuchi, and A. Scherer, "Design of photonic crystal microcavities for cavity QED," *Phys. Rev. E*, Vol. 65, 2001, 016608.

[174] Vuckovic, J., M. Loncar, H. Mabuchi, and A. Scherer, "Optimization of the Q factor in photonic crystal microcavities," *IEEE J. Quantum Electron.*, Vol. 38, 2002, pp. 850–856.

[175] Vuckovic, J., and Y. Yamamoto, "Photonic crystal microcavities for cavity quantum electrodynamics with a single quantum dot," *Appl. Phys. Lett.*, Vol. 82, 2003, pp. 2374–2376.

[176] Loncar, M., A. Scherer, and Q. Yueming, "Photonic crystal laser sources for chemical detection," *Appl. Phys. Lett.*, Vol. 82, 2003, pp. 4648–4650.

[177] Suh, W., M. F. Yanik, O. Solgaard, and S. H. Fan, "Displacement-sensitive photonic crystal structures based on guided resonance in photonic crystal slabs," *Appl. Phys. Lett.*, Vol. 82, 2003, pp. 1999–2001.

[178] Loncar, M., T. Yoshie, K. Okamoto, Y. M. Qiu, J. Vuckovic, and A. Scherer, "Planar photonic crystal nanolasers (i): Porous cavity lasers," *IEICE Trans. Electron.*, Vol. E87C, 2004, pp. 291–299.

[179] Chow, E., A. Grot, L. W. Mirkarimi, M. Sigalas, and G. W. Burr, "Photonic crystal microcavities as ultracompact film-thickness monitors for biosensing," in preparation.

[180] Lam, C. F., R. B. Vrijen, P. P. L. Chang-Chien, D. F. Sievenpiper, and E. Yablonovitch, "A tunable wavelength demultiplexer using logarithmic filter chains," *J. Lightwave Tech.*, Vol. 16, 1998, pp. 1657–1662.

[181] Leonard, S. W., J. P. Mondia, H. M. van Driel, O. Toader, S. John, K. Busch, A. Birner, U. Gosele, and V. Lehmann, "Tunable two-dimensional photonic crystals using liquid-crystal infiltration," *Phys. Rev. B*, Vol. 61, 2000, pp. R2389–R2392.

[182] Hillmer, H., J. Daleiden, C. Prott, F. Romer, S. Irmer, V. Rangelov, A. Tarraf, S. Schuler, and M. Strassner, "Potential for micromachined actuation of ultra-wide continuously tunable optoelectronic devices," *Appl. Phys. B*, Vol. 75, 2002, pp. 3–13.

[183] Pustai, D. M., A. Sharkawy, S. Y. Shi, and D. W. Prather, "Tunable photonic crystal microcavities," *Appl. Optics*, Vol. 41, 2002, pp. 5574–5579

[184] Yi, Y., P. Bermel, K. Wada, X. M. Duan, J. D. Joannopoulos, and L. C. Kimerling, "Tunable multichannel optical filter based on silicon photonic band gap materials actuation," *Appl. Phys. Lett.*, Vol. 81, 2002, pp. 4112–4114.

[185] Nakagawa, W., and Y. Fainman, "Tunable optical nanocavity based on modulation of near-field coupling between subwavelength periodic nanostructures," *IEEE J. Selected Topics Quantum Electron.*, Vol. 10, 2004, pp. 478–483.

[186] Panoiu, N. C., M. Bahl, and R. M. Osgood, "All-optical tunability of a nonlinear photonic crystal channel drop filter," *Optics Express*, Vol. 12, 2004, pp. 1606–1610.

[187] Yanik, M. F., S. H. Fan, and M. Soljacic, "High contrast all optical bistable switching in photonic crystal microcavities," *Appl. Phys. Lett.*, Vol. 83, 2003, pp. 2739–2741.

[188] Yanik, M. F., S. H. Fan, M. Soljacic, and J. D. Joannopoulos, "All-optical transistor action with bistable switching in a photonic crystal cross-waveguide geometry," *Optics Lett.*, Vol. 28, 2003, pp. 2506–2508.

[189] Okano, M., and S. Noda, "Analysis of multimode point-defect cavities in three-dimensional photonic crystals using group theory in frequency and time domains," *Phys. Rev. B*, Vol. 70, 2004, 125105.

[190] Chutinan, A., and S. Noda, "Highly confined waveguides and waveguide bends in three-dimensional photonic crystal," *Appl. Phys. Lett.*, Vol. 75, 1999, pp. 3739–3741.

[191] Noda, S., K. Tomoda, N. Yamamoto, and A. Chutinan, "Full three-dimensional photonic bandgap crystals at near-infrared wavelengths," *Science*, Vol. 289, 2000, pp. 604–606.

[192] Okano, M., A. Chutinan, and S. Noda, "Analysis and design of single-defect cavities in a three-dimensional photonic crystal," *Phys. Rev. B*, Vol. 66, 2002, 165211.

[193] Ogawa, S., M. Imada, S. Yoshimoto, M. Okano, and S. Noda, "Control of light emission by 3D photonic crystals," *Science*, Vol. 305, 2004, pp. 227–229.

[194] Meier, M., A. Mekis, A. Dodabalapur, A. Timko, R. E. Slusher, J. D. Joannopoulos, and O. Nalamasu, "Laser action from two-dimensional distributed feedback in photonic crystals," *Appl. Phys. Lett.*, Vol. 74, 1999, pp. 7–9.

[195] Imada, M., S. Noda, A. Chutinan, T. Tokuda, M. Murata, and G. Sasaki, "Coherent two-dimensional lasing action in surface-emitting laser with triangular-lattice photonic crystal structure," *Appl. Phys. Lett.*, Vol. 75, 1999, pp. 316–318.

[196] Kwon, S. H., H. Y. Ryu, G. H. Kim, Y. H. Lee, and S. B. Kim, "Photonic band edge lasers in two-dimensional square-lattice photonic crystal slabs," *Appl. Phys. Lett.*, Vol. 83, 2003, pp. 3870–3872.

[197] Ryu, H. Y., M. Notomi, and Y. H. Lee, "Finite difference time-domain investigation of band edge resonant modes in finite size two dimensional photonic crystal slab," *Phys. Rev. B*, Vol. 68, 2003, 45209.

[198] Guo, S. P., and S. Albin, "Numerical techniques for excitation and analysis of defect modes in photonic crystals," *Optics Express*, Vol. 11, 2003, pp. 1080–1089.

[199] Yokoyama, M., and S. Noda, "Finite-difference time-domain simulation of two-dimensional photonic crystal surface-emitting laser having a square-lattice slab structure," *IEICE Trans. Electron.*, Vol. E87C, 2004, pp. 386–392.

[200] Xu, Y., R. K. Lee, and A. Yariv, "Propagation and second-harmonic generation of electromagnetic waves in a coupled-resonator optical waveguide," *J. Opt. Soc. America B*, Vol. 17, 2000, pp. 387–400.

[201] Geremia, J. M., J. Williams, and H. Mabuchi, "Inverse-problem approach to designing photonic crystals for cavity QED experiments," *Phys. Rev. E*, Vol. 66, 2002, 066606.

[202] Akahane, Y., T. Asano, B. S. Song, and S. Noda, "High-Q photonic nanocavity in a two-dimensional photonic crystal," *Nature*, Vol. 425, 2003, pp. 944–947.

[203] Srinivasan, K., P. E. Barclay, O. Painter, J. X. Chen, A. Y. Cho, and C. Gmachl, "Experimental demonstration of a high quality factor photonic crystal microcavity," *Appl. Phys. Lett.*, Vol. 83, 2003, pp. 1915–1917.

[204] Srinivasan, K., P. E. Barclay, and O. Painter, "Fabrication-tolerant high quality factor photonic crystal microcavities," *Optics Express*, Vol. 12, 2004, pp. 1458–1463.

[205] Painter, O., and K. Srinivasan, "Polarization properties of dipole-like defect modes in photonic crystal nanocavities," *Optics Lett.*, Vol. 27, 2002, pp. 339–341.

[206] Srinivasan, K., O. Painter, R. Colombelli, C. Gmachl, D. M. Tennant, A. M. Sergent, D. L. Sivco, A. Y. Cho, M. Troccoli, and F. Capasso, "Lasing mode pattern of a quantum cascade photonic crystal surface-emitting microcavity laser," *Appl. Phys. Lett.*, Vol. 84, 2004, pp. 4164–4166.

[207] Bourgeade, A., and E. Freysz, "Computational modeling of second-harmonic generation by solution of full-wave vector Maxwell equations," *J. Optical Soc. America B*, Vol. 17, 2000, pp. 226–234.

[208] Lee, T. W., and S. C. Hagness, "Pseudospectral time-domain methods for modeling optical wave propagation in second-order nonlinear materials," *J. Optical Soc. America B*, Vol. 21, 2004, pp. 330–342.

[209] Liu, Q. H., *The PSTD Algorithm: A Time-Domain Method Requiring Only Two Grids Per Wavelength*, New Mexico State Univ., Las Cruces, NM, Tech. Rept. NMSU-ECE96-013, Dec. 1996.

[210] Liu, Q. H., "The pseudospectral time-domain (PSTD) method: A new algorithm for solutions of Maxwell's equations," *Proc. 1997 IEEE Antennas Propagat. Soc. Intl. Symp.*, Vol. 1, pp. 122–125.

[211] Berger, V., "Nonlinear photonic crystals," *Phys. Rev. Lett.*, Vol. 81, 1998, pp. 4136–4139.

[212] Broderick, N. G. R., G. W. Ross, H. L. Offerhaus, D. J. Richardson, and D. C. Hanna, "Hexagonally poled lithium niobate: A two-dimensional nonlinear photonic crystal," *Phys. Rev. Lett.*, Vol. 84, 2000, pp. 4345–4348.

[213] Chowdhury, A., S. C. Hagness, and L. McCaughan, "Simultaneous optical wavelength interchange with a two-dimensional second-order nonlinear photonic crystal," *Optics Lett.*, Vol. 25, 2000, pp. 832–834.

[214] Lee, T. W., and S. C. Hagness, "Pseudospectral time-domain analysis of second-harmonic generation in 2D nonlinear photonic crystals," *Proc. IEEE LEOS Ann. Mtng.*, Vol. 1, 2003, pp. 194–195.

[215] Korner, T. O., and W. Fichnter, "Auxiliary differential equation: Efficient implementation in the finite-difference time-domain method," *Optics Lett.*, Vol. 22, 1997, pp. 1586–1588.

[216] Lee, T. W., *Development and Application of Full-Wave Time-Domain Numerical Modeling Techniques for the Analysis of Linear and Nonlinear Photonic Micro / Nanostructures*, Ph.D. diss., Univ. Wisconsin-Madison, 2003.

[217] Lee, T. W., and S. C. Hagness, "A compact wave source condition for the pseudospectral time-domain method," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 3, 2004, pp. 253–256.

[218] Raether, H., *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, Vol. 111, *Springer Tracts in Modern Physics*, Berlin: Springer-Verlag, 1988.

[219] Sambles, J. R., G. W. Bradbery, and F. Z. Yang, "Optical excitation of surface plasmons—An introduction," *Contemporary Physics*, Vol. 32, 1991, pp. 173–183.

[220] Welford, K., "Surface-plasmon polaritons and their uses," *Optical Quantum Electron.*, Vol. 23, 1991, pp. 1–27.

[221] Rakic, A. D., A. B. Djurisic, J. M. Elazar, and M. L. Majewski, "Optical properties of metallic films for vertical-cavity optoelectronic devices," *Applied Optics*, Vol. 37, 1998, pp. 5271–5283.

[222] Oubre, C., and P. Nordlander, "Optical properties of metallodielectric nanostructures calculated using the finite difference time domain method," *J. Phys. Chem. B*, Vol. 108, 2004, pp. 17740–17747.

[223] Coronado, E. A., and G. C. Schatz, "Surface plasmon broadening for arbitrary shape nanoparticles: A geometrical probability approach," *J. Chem. Phys.*, Vol. 119, 2003, pp. 3926–3934.

[224] Burr, G. W., "FDTD as a nanophotonics design optimization tool," *Proc. Intl. Symp. Photonic Electromagn. Crystals*, 2004.

[225] Young, J. L., and R. O. Nelson, "A summary and systematic analysis of FDTD algorithms for linearly dispersive media," *IEEE Antennas Propagat. Magazine*, Vol. 43(1), 2001, pp. 61–77.

[226] Kneipp, K., H. Kneipp, I. Itzkan, R. R. Dasari, and M. S. Feld, "Ultrasensitive chemical analysis by Raman spectroscopy," *Chemical Rev.*, Vol. 99, 1999, p. 2957.

[227] Kelly, K. L., E. Coronado, L. L. Zhao, and G. C. Schatz, "The optical properties of metal nanoparticles: The influence of size, shape, and dielectric environment," *J. Phys. Chem. B*, Vol. 107, 2003, pp. 668–677.

[228] Jackson, J. B., and N. J. Halas, "Surface-enhanced Raman scattering on tunable plasmonic nanoparticle substrates," *Proc. National Acad. Sciences*, Vol. 101, 2004, pp. 17930–17935.

[229] Charnay, C., A. Lee, S. Q. Man, C. E. Moran, C. Radloff, R. K. Bradley, and N. J. Halas, "Reduced symmetry metallodielectric nanoparticles: Chemical synthesis and plasmonic properties," *J. Phys. Chem. B*, Vol. 107, 2003, pp. 7327–7333.

[230] Tamaru, H., H. Kuwata, H. T. Miyazaki, and K. Miyano, "Resonant light scattering from individual Ag nanoparticles and particle pairs," *Appl. Phys. Lett.*, Vol. 80, 2002, pp. 1826–1828.

[231] Crozier, K. B., A. Sundaramurthy, G. S. Kino, and C. F. Quate, "Optical antennas: Resonators for local field enhancement," *J. Appl. Phys.*, Vol. 94, 2003, pp. 4632–4642.

[232] Gray, S. K., and T. Kupka, "Propagation of light in metallic nanowire arrays: Finite-difference time-domain studies of silver cylinders," *Phys. Rev. B*, Vol. 68, 2003, 045415.

[233] Oliva, J. M., and S. K. Gray, "Theoretical study of dielectrically coated metallic nanowires," *Chem. Phys. Lett.*, Vol. 379, 2003, pp. 325–331.

[234] Fromm, D. P., A. Sundaramurthy, P. J. Schuck, G. Kino, and W. E. Moerner, "Gap-dependent optical coupling of single 'bowtie' nanoantennas resonant in the visible," *Nano Lett.*, Vol. 4, 2004, pp. 957–961.

[235] Kik, P. G., S. A. Maier, and H. A. Atwater, "Plasmon printing — A new approach to near-field lithography," *Proc. Materials Rsrch. Soc.*, Vol. 705, 2002, p. Y3.6.

[236] Kik, P. G., A. L. Martin, S. A. Maier, and H. A. Atwater, "Metal nanoparticle arrays for near field optical lithography," *Proc. SPIE*, 2002.

[237] Furukawa, H., and S. Kawata, "Local field enhancement with an apertureless near-field microscope probe," *Optics Comm.*, Vol. 148, 1998, pp. 221–224.

[238] Baida, F. I., D. Van Labeke, and J. M. Vigoureux, "Theoretical study of near-field surface plasmon excitation, propagation, and diffraction," *Optics Comm.*, Vol. 171, 1999, pp. 317–331.

[239] Hayazawa, N., Y. Inouye, Z. Sekkat, and S. Kawata, "Metallized tip amplification of near-field Raman scattering," *Optics Comm.*, Vol. 183, 2000, pp. 333–336.

[240] Milner, R. G., and D. Richards, "The role of tip plasmons in near-field Raman microscopy," *J. Microscopy–Oxford*, Vol. 202, 2001, pp. 66–71.

[241] Akhremitchev, B. B., S. Pollack, and G. C. Walker, "Apertureless scanning near-field infrared microscopy of a rough polymeric surface," *Langmuir*, Vol. 17, 2001, pp. 2774–2781.

[242] Krug, J. T., E. J. Sanchez, and X. S. Xie, "Design of near-field optical probes with optimal field enhancement by finite difference time domain electromagnetic simulation," *J. Chem. Phys.* Vol. 116, 2002, pp. 10895–10901.

[243] Chau, Y. F., T. J. Yang, and D. P. Tsai, "Imaging properties of three dimensional aperture near-field scanning optical microscopy and optimized near-field fiber probe designs," *Japanese J. Appl. Phys. Part 1*, Vol. 43, 2004, pp. 8115–8125.

[244] Quinten, M., A. Leitner, J. R. Krenn, and F. R. Aussenegg, "Electromagnetic energy transport via linear chains of silver nanoparticles," *Optics Lett.*, Vol. 23, 1998, pp. 1331–1333.

[245] Brongersma, M. L., J. W. Hartman, and H. A. Atwater, "Electromagnetic energy transfer and switching in nanoparticle chain arrays below the diffraction limit," *Phys. Rev. B*, Vol. 62, 2000, pp. R16356–R16359.

[246] Berini, P., "Plasmon-polariton modes guided by a metal film of finite width," *Optics Lett.*, Vol. 24, 1999, pp. 1011–1013.

[247] Weeber, J. C., A. Dereux, C. Girard, J. R. Krenn, and J. P. Goudonnet, "Plasmon polaritons of metallic nanowires for controlling submicron propagation of light," *Phys. Rev. B*, Vol. 60, 1999, pp. 9061–9068.

[248] Maier, S. A., P. G. Kik, and H. A. Atwater, "Observation of coupled plasmon-polariton modes in Au nanoparticle chain waveguides of different lengths: Estimation of waveguide loss," *Appl. Phys. Lett.*, Vol. 81, 2002, pp. 1714–1716.

[249] Maier, S. A., P. E. Barclay, T. J. Johnson, M. D. Friedman, and O. Painter, "Low-loss fiber accessible plasmon waveguide for planar energy guiding and sensing," *Appl. Phys. Lett.*, Vol. 84, 2004, pp. 3990–3992.

[250] Panoiu, N. C., and R. M. Osgood, "Subwavelength nonlinear plasmonic nanowire," *Nano Lett.*, Vol. 4, 2004, pp. 2427–2430.

[251] Hochberg, M., T. Baehr-Jones, C. Walker, and A. Scherer, "Integrated plasmon and dielectric waveguides," *Optics Express*, Vol. 12, 2004, p. 5481.

[252] Wang, B., and G. P. Wang, "Surface plasmon polariton propagation in nanoscale metal gap waveguides," *Optics Lett.*, Vol. 29, 2004, pp. 1992–1994.

[253] Wang, B., and G. P. Wang, "Metal heterowaveguides for nanometric focusing of light," *Appl. Phys. Lett.*, Vol. 85, 2004, pp. 3599–3601.

[254] Christensen, D. A., "Analysis of near-field tip patterns including object interaction using finite-difference time-domain calculations," *Ultramicroscopy*, Vol. 57, 1995, pp. 189–195.

[255] Furukawa, H., and S. Kawata, "Analysis of image formation in a near-field scanning optical microscope: Effects of multiple scattering," *Optics Comm.*, Vol. 132, 1996, pp. 170–178.

[256] Vasilyeva, E., and A. Taflove, "Three-dimensional modeling of amplitude-object imaging in scanning near-field optical microscopy," *Optics Lett.*, Vol. 23, 1998, pp. 1155–1157.

[257] Minh, P. N., T. Ono, S. Tanaka, and M. Esashi, "Near-field optical apertured tip and modified structures for local field enhancement," *Applied Optics*, Vol. 40, 2001, pp. 2479–2484.

[258] Nakamura, H., T. Sato, H. Kambe, K. Sawada, and T. Saiki, "Design and optimization of tapered structure of near-field fibre probe based on finite-difference time-domain simulation," *J. Microscopy-Oxford*, Vol. 202, 2001, pp. 50–52.

[259] Parent, G., D. Van Labeke, and F. I. Baida, "Theoretical study of transient phenomena in near-field optics," *J. Microscopy-Oxford*, Vol. 202, 2001, pp. 296–306.

[260] Yatsui, T., K. Itsumi, M. Kourogi, and M. Ohtsu, "Metallized pyramidal silicon probe with extremely high throughput and resolution capability for optical near-field technology," *Appl. Phys. Lett.*, Vol. 80, 2002, pp. 2257–2259.

[261] Baida, F. I., D. Van Labeke, and Y. Pagani, "Body-of-revolution FDTD simulations of improved tip performance for scanning near-field optical microscopes," *Optics Comm.*, Vol. 225, 2003, pp. 241–252.

[262] Kim, J., J. Kim, K. B. Song, S. Q. Lee, E. K. Kim, and K. H. Park, "Scanning near-field optical microscope study of Ag nanoprotrusions fabricated by nano-oxidation with atomic force microscope," *Japanese J. Appl. Phys. Part 1*, Vol. 42, 2003, pp. 7635–7639.

[263] Goto, K., Y. J. Kim, T. Kirigaya, and Y. Masuda, "Near-field evanescent wave enhancement with nanometer-sized metal grating and microlens array in parallel optical recording head," *Japanese J. Appl. Phys. Part 1*, Vol. 43, 2004, pp. 5814–5818.

[264] Kato, K., S. Ichihara, M. Oumi, H. Maeda, T. Niwa, Y. Mitsuoka, K. Nakajima, T. Ohkubo, and K. Itao, "Signal readout using small near-field optical head with horizontal light introduction through optical fiber," *Japanese J. Appl. Phys. Part 1*, Vol. 42, 2003, pp. 5102–5106.

[265] Tanaka, K., T. Ohkubo, M. Oumi, Y. Mitsuoka, K. Nakajima, H. Hosaka, and K. Itao, "Numerical simulation of readout characteristics of the planar aperture-mounted head with a minute scatterer," *Japanese J. Appl. Phys. Part 1*, Vol. 40, 2001, pp. 1542–1547.

[266] Tanaka, K., H. Hosaka, K. Itao, M. Oumi, T. Niwa, T. Miyatani, Y. Mitsuoka, K. Nakajima, and T. Ohkubo, "Improvements in near-field optical performance using localized surface plasmon excitation by a scatterer-formed aperture," *Appl. Phys. Lett.*, Vol. 83, 2003, pp. 1083–1085.

[267] Hashizume, J., and F. Koyama, "Plasmon enhanced optical near-field probing of metal nanoaperture surface emitting laser," *Optics Express*, Vol. 12, 2004, pp. 6391–6396.

[268] Hashizume, J., and F. Koyama, "Plasmon enhancement of optical near-field of metal nanoaperture surface-emitting laser," *Appl. Phys. Lett.*, Vol. 84, 2004, pp. 3226–3228.

[269] Tanaka, K., M. Oumi, T. Niwa, S. Ichihara, Y. Mitsuoka, Nakajima, T. Ohkubo, H. Hosaka, and K. Itao, "High spatial resolution and throughput potential of an optical head with a triangular aperture for near-field optical data storage," *Japanese J. Appl. Phys. Part 1*, Vol. 42, 2003, pp. 1113–1117.

[270] Matsumoto, T., T. Shimano, H. Saga, H. Sukeda, and M. Kiguchi, "Highly efficient probe with a wedge-shaped metallic plate for high density near-field optical recording," *J. Appl. Phys.*, Vol. 95, 2004, pp. 3901–3906.

[271] Tominaga, J., H. Johoo Kim Fuji, D. Buchel, T. Kikuawa, H. Liqiu Men Fukuda, A. Sato, T. Nakano, A. Tachibana, Y. Yamakawa, M. Kumagai, T. Fukaya, and N. Atoda, "Super-resolution near-field structure and signal enhancement by surface plasmons," *Japanese J. Appl. Phys. Part 1*, Vol. 40, 2001, pp. 1831–1834.

[272] Liu, W. C., and D. P. Tsar, "Nonlinear near-field optical effects of the agox-type super-resolution near-field structure," *Japanese J. Appl. Phys. Part 1*, Vol. 42, 2003, pp. 1031–1032.

[273] Kataja, K., J. Olkkonen, J. Aikio, and D. Howe, "Readout modeling of super resolution disks," *Japanese J. Appl. Phys. Part 1*, Vol. 43, 2004, pp. 4718–4723.

[274] Liu, W. C., M. Y. Ng, and D. P. Tsai, "Surface plasmon effects on the far-field signals of agox-type super-resolution near-field structure," *Japanese J. Appl. Phys. Part 1*, Vol. 43, 2004, pp. 4713–4717.

[275] Thio, T., K. M. Pellerin, R. A. Linke, H. J. Lezec, and T. W. Ebbesen, "Enhanced light transmission through a single subwavelength aperture," *Optics Lett.*, Vol. 26, 2001, pp. 1972–1974.

[276] Ebbesen, T. W., H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, "Extraordinary optical transmission through subwavelength hole arrays," *Nature*, Vol. 391, 1998, pp. 667–669.

[277] Barnes, W. L., A. Dereux, and T. W. Ebbesen, "Surface plasmon subwavelength optics," *Nature*, Vol. 424, 2003, pp. 824–830.

[278] Bethe, H. A., "Theory of diffraction by small holes," *Phys. Rev.*, Vol. 66, 1944, pp. 163–182.

[279] Barnes, W. L., W. A. Murray, J. Dintinger, E. Devaux, and T. W. Ebbesen, "Surface plasmon polaritons and their role in the enhanced transmission of light through periodic arrays of subwavelength holes in a metal film," *Phys. Rev. Lett.*, Vol. 92, 2004, 107401.

[280] Lezec, H. J., and T. Thio, "Diffracted evanescent wave model for enhanced and suppressed optical transmission through subwavelength hole arrays," *Optics Express*, Vol. 12, 2004, pp. 3629–3651.

[281] Lezec, H. J., A. Degiron, E. Devaux, R. A. Linke, L. Martin-Moreno, F. J. Garcia-Vidal, and T. W. Ebbesen, "Beaming light from a subwavelength aperture," *Science*, Vol. 297, 2002, pp. 820–822.

[282] Garcia-Vidal, F. J., L. Martin-Moreno, H. J. Lezec, and T. W. Ebbesen, "Focusing light with a single subwavelength aperture flanked by surface corrugations," *Appl. Phys. Lett.*, Vol. 83, 2003, pp. 4500–4502.

[283] Martin-Moreno, L., F. J. Garcia-Vidal, H. J. Lezec, A. Degiron, and T. W. Ebbesen, "Theory of highly directional emission from a single subwavelength aperture surrounded by surface corrugations," *Phys. Rev. Lett.*, Vol. 90, 2003, 167401.

[284] Baida, F. I., and D. Van Labeke, "Light transmission by subwavelength annular aperture arrays in metallic films," *Optics Comm.*, Vol. 209, 2002, pp. 17–22.

[285] Hohng, S. C., Y. C. Yoon, D. S. Kim, V. Malyarchuk, R. Muller, C. Lienau, J. W. Park, K. H. Yoo, J. Kim, H. Y. Ryu, and Q. H. Park, "Light emission from the shadows: Surface plasmon nano-optics at near and far fields," *Appl. Phys. Lett.*, Vol. 81, 2002, pp. 3239–3241.

[286] Egorov, D., B. S. Dennis, G. Blumberg, and M. I. Haftel, "Two-dimensional control of surface plasmons and directional beaming from arrays of subwavelength apertures," *Phys. Rev. B*, Vol. 70, 2004, 033404.

[287] Muller, R., C. Ropers, and C. Lienau, "Femtosecond light pulse propagation through metallic nanohole arrays: The role of the dielectric substrate," *Optics Express*, Vol. 12, 2004, pp. 5067–5081.

[288] Yin, L., V. K. Vlasko-Vlasov, A. Rydh, J. Pearson, U. Welp, S. H. Chang, S. K. Gray, G. C. Schatz, D. B. Brown, and C. W. Kimball, "Surface plasmons at single nanoholes in Au films," *Appl. Phys. Lett.*, Vol. 85, 2004, pp. 467–469.

[289] Zakharian, A. R., M. Mansuripur, and J. V. Moloney, "Transmission of light through small elliptical apertures," *Optics Express*, Vol. 12, 2004, pp. 2631–2648.

[290] Shi, X. L., and L. Hesselink, "Design of a C aperture to achieve $\lambda/10$ resolution and resonant transmission," *J. Opt. Soc. America B*, Vol. 21, 2004, pp. 1305–1317.

[291] Matteo, J. A., and L. Hesselink, "Fractal extensions of near-field aperture shapes for enhanced transmission and resolution," *Optics Express*, Vol. 13, 2005, pp. 636–647.

[292] Piket-May, M. J., A. Taflove, and J. B. Troy, "Electrodynamics of visible light interactions with the vertebrate retinal rod," *Optics Lett.*, Vol. 18, 1993, pp. 568–570.

[293] Drezek, R., A. Dunn, and R. Richards-Kortum, "A pulsed finite-difference time-domain (FDTD) method for calculating light scattering from biological cells over broad wavelength ranges," *Optics Express*, Vol. 6, 2000, pp. 147–157.

[294] Drezek, R., M. Guillaud, T. Collier, I. Boiko, A. Malpica, C. Macaulay, M. Follen, and R. Richards-Kortum, "Light scattering from cervical cells throughout neoplastic progression: Influence of nuclear morphology, DNA content, and chromatin texture," *J. Biomedical Optics*, Vol. 8, 2003, pp. 7–16.

[295] Roy, H. K., Y. Liu, R. K. Wali, Y. L. Kim, A. K. Kromin, M. J. Goldberg, and V. Backman, "Four-dimensional elastic light-scattering fingerprints as preneoplastic markers in the rat model of colon carcinogenesis," *Gastroenterology*, Vol. 126, 2004, pp. 1071–1081.

[296] Li, X., A. Taflove, and V. Backman, "Recent progress in exact and reduced-order modeling of light-scattering properties of complex structures," *IEEE J. Selected Topics Quantum Electron.*, submitted.

[297] Tseng, S. H., J. H. Greene, A. Taflove, D. Maitland, V. Backman, and J. T. Walsh, Jr., "Exact solution of Maxwell's equations for optical interactions with a macroscopic random medium," *Optics Lett.*, Vol. 29, 2004, pp. 1393–1395.

[298] Tseng, S. H., J. H. Greene, A. Taflove, D. Maitland, V. Backman, and J. T. Walsh, Jr., "Addendum to 'Exact solution of Maxwell's equations for optical interactions with a macroscopic random medium,' " *Optics Lett.*, Vol. 30, 2005, pp. 56–57.

Chapter 17

Advances in PSTD Techniques

Qing Liu and Gang Zhao

17.1 INTRODUCTION

As demonstrated in the previous chapters, the FDTD method is a simple, robust, and powerful technique to simulate transient electromagnetic phenomena. However, numerous examples have indicated that a spatial sampling density of at least 10 to 20 cells per minimum wavelength is necessary to ensure that the FDTD method produces acceptable results even for an electrically small problem. To maintain acceptable accuracy for an electrically large problem, it may be necessary to increase the spatial sampling rate beyond this range to reduce the cumulative numerical dispersion error, which is linearly proportional to the length of wave propagation. In other words, given a prescribed accuracy, the spatial sampling rate is not just determined by the minimum wavelength, but also by the electrical size of the modeled structure. This makes FDTD modeling of electrically large problems very challenging.

In order to efficiently solve time-domain electromagnetic problems, various techniques have been proposed to improve the FDTD method. This chapter focuses on recent advances in *pseudospectral time-domain* (PSTD) techniques. PSTD techniques use either trigonometric functions [1–3] or Chebyshev polynomials [4–11] to approximate spatial derivatives in order to greatly reduce the numerical dispersion error. When applied to single domains having smooth internal media, PSTD techniques based upon these functions have *spectral accuracy*, meaning that the numerical dispersion error decreases exponentially with the sampling density. Spectral accuracy also can be achieved for problems with multiple regions of inhomogeneities when PSTD techniques are coupled with appropriate boundary-patching conditions.

We will first summarize the basic finite-difference, Fourier pseudospectral, and Chebyshev pseudospectral approximation methods used for the derivative of a function. Next, we will present a single-domain Fourier PSTD method that uses the fast Fourier transform algorithm. Then, a single-domain Chebyshev PSTD method will be presented where the fields are represented by Chebyshev polynomials. Finally, several multidomain PSTD methods are described for modeling structures with discontinuous material compositions.

17.2 APPROXIMATION OF DERIVATIVES

In this section, we summarize the basic finite-difference, Fourier pseudospectral, and Chebyshev pseudospectral approximation methods used for the derivative of a function. A convenient compact matrix notation is used.

17.2.1 Derivative Matrix for the Second-Order Finite-Difference Method

For a uniformly sampled function, the central-differencing scheme

$$u_m = \frac{df(x_m)}{dx} \approx \frac{f(x_{m+1}) - f(x_{m-1})}{2\Delta x} \tag{17.1}$$

has a second-order accuracy [i.e., the error is $O(\Delta x)^2$], as can be verified by a Taylor expansion. Assuming a set of periodic data $\{f_m, m = 1, 2, \ldots, N\}$, where $f_{m+N} = f_m$ for all integers m, we can write this derivative in terms of a derivative matrix $[D]$ such that

$$\{u\} = [D]\{f\} \tag{17.2}$$

where

$$\{u\} = \{u_1, u_2, \ldots, u_N\}^{\mathrm{T}}, \quad \{f\} = \{f_1, f_2, \ldots, f_N\}^{\mathrm{T}} \tag{17.3}$$

$$[D] = \frac{1}{2\Delta x}\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & -1 \\ -1 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & -1 & 0 \end{pmatrix} \tag{17.4}$$

This can be also viewed from the interpolation approach of a second-order polynomial $p^{(2)}(x)$:

$$\begin{aligned} f(x) &\approx f_{m-1}\phi_{-1}^{(2)}(x) + f_m\phi_0^{(2)}(x) + f_{m+1}\phi_1^{(2)}(x) \\ &= \sum_{\ell=-1}^{1} f_{m+\ell}\phi_\ell^{(2)}(x) \qquad x_{m-1} \le x < x_{m+1} \end{aligned} \tag{17.5}$$

where $\phi_m^{(2)}(x)$ are the Lagrange interpolation polynomials

$$\phi_{-1}^{(2)}(x) = \frac{(x - x_m)(x - x_{m+1})}{2(\Delta x)^2} \tag{17.6a}$$

$$\phi_0^{(2)}(x) = -\frac{(x - x_{m-1})(x - x_{m+1})}{(\Delta x)^2} \tag{17.6b}$$

$$\phi_1^{(2)}(x) = \frac{(x - x_{m-1})(x - x_m)}{2(\Delta x)^2} \tag{17.6c}$$

Hence, the derivative is given by

$$\{u_m\} = \left\{\frac{df(x_m)}{dx}\right\} = \sum_{\ell=-1}^{1} f_{m+\ell}\frac{d\phi_\ell^{(2)}(x_m)}{dx} \equiv [D_{mn}]\{f_n\} \tag{17.7}$$

where $[D]$ is a Toeplitz matrix given by

$$D_{mn} = \frac{1}{2\Delta x}\left(\delta_{n,m+1} - \delta_{n,m-1}\right) \equiv \frac{1}{2\Delta x}\left(\delta_{m-n+1,0} - \delta_{m-n-1,0}\right) \equiv a_{m-n}$$

$$= \begin{cases} 1/(2\Delta x) & n = m+1 \\ -1/(2\Delta x) & n = m-1 \\ 0 & \text{otherwise} \end{cases} \tag{17.8a}$$

$$a_k = \frac{1}{2\Delta x}\left(\delta_{k-1,0} - \delta_{k+1,0}\right) \tag{17.8b}$$

17.2.2 Derivative Matrices for Fourth-Order and *N*'th-Order Finite-Difference Methods

For the fourth-order finite-difference scheme, we have

$$f(x) \approx \sum_{\ell=-2}^{2} f_{m+\ell}\,\phi_\ell^{(4)}(x) \qquad x_{m-2} \le x < x_{m+2} \tag{17.9}$$

The derivative matrix is given by

$$D_{mn} \equiv a_{m-n} = \frac{1}{12\Delta x}\left(8\delta_{m-n+1,0} - 8\delta_{m-n-1,0} - \delta_{m-n+2,0} + \delta_{m-n-2,0}\right)$$

$$= \begin{cases} 1/(12\Delta x) & n = m-2 \\ -2/(3\Delta x) & n = m-1 \\ 2/(3\Delta x) & n = m+1 \\ -1/(12\Delta x) & n = m+2 \\ 0 & \text{otherwise} \end{cases} \tag{17.10}$$

Similarly, the finite-difference scheme can be increased to N'th order with all N points, yielding

$$D_{mn} = \frac{d\phi_{n-m}^{(N)}(x_m)}{dx} \tag{17.11}$$

where $\phi_{n-m}^{(N)}(x)$ are the N'th-order Lagrange polynomials. The required $N+1$ data points are provided by the N points, plus the additional point from the periodic boundary condition.

17.2.3 Trigonometric Interpolation and FFT Method

If the period of the computational domain is $T = x_{\max} - x_{\min}$, with spatial sampling points at $x_m = x_{\min} + (m-1)\Delta x$ (where $m = 0, 1, \ldots, N-1$ and $\Delta x = T/N$), we can write the periodic function as the truncated Fourier series

$$f(x) \approx \frac{1}{N} \sum_{n=-N/2}^{N/2-1} \hat{f}_n \, e^{j2\pi n(x-x_0)/N\Delta x} \tag{17.12}$$

where the Fourier series coefficients are

$$\hat{f}_n = \frac{N}{T} \int_{x_{\min}}^{x_{\max}} f(x) e^{-j2\pi n(x-x_0)/N\Delta x} dx \approx \sum_{m=0}^{N-1} f(x_m) e^{-j2\pi mn/N} \equiv \mathrm{DFT}(\{f\})_n \tag{17.13}$$

Thus, from (17.12), we have

$$\begin{aligned} \frac{df(x_m)}{dx} &\approx \frac{1}{N^2} \sum_{n=-N/2}^{N/2-1} \frac{j2\pi n}{\Delta x} \hat{f}_n \, e^{j2\pi mn/N} \equiv \frac{2\pi}{N\Delta x} \mathrm{DFT}^{-1}\left(jn\{\hat{f}\}_n\right)_m \\ &= \frac{2\pi}{N\Delta x} \mathrm{DFT}^{-1}\left[jn\{\mathrm{DFT}(\{f\})_n\right]_m \end{aligned} \tag{17.14}$$

where DFT and DFT^{-1} denote the forward and inverse discrete Fourier transforms, respectively.

When substituting the Fourier series coefficients $\hat{f}_n$ into (17.12), one can obtain the explicit derivative matrix

$$\frac{df(x_m)}{dx} \approx \frac{2\pi}{N^2\Delta x} \sum_{p=0}^{N-1} f(x_p) \sum_{n=-N/2}^{N/2-1} jn e^{j2\pi(m-p)n/N} = \sum_{p=0}^{N-1} f(x_p) D_{mp} \tag{17.15}$$

where

$$\begin{aligned} D_{mp} &= \frac{2\pi}{N^2\Delta x} \sum_{n=-N/2}^{N/2-1} jn e^{j2\pi(m-p)n/N} = \frac{\pi}{N\Delta x}(-1)^{m-p}\left(1-\delta_{m-p,\,0}\right) \cot\left[\frac{(m-p)\pi}{N}\right] \\ &\equiv a_{m-p} \end{aligned} \tag{17.16}$$

Therefore, the derivative matrix is again Toeplitz. Hence, the derivative vector is given by

$$\frac{d\{f\}}{dx} = [D]\{f\} = \mathrm{DFT}^{-1}[\mathrm{DFT}(\{f\}) \cdot \mathrm{DFT}(\{a\})] \tag{17.17}$$

This derivative costs $O(N \log N)$ operations. For an analytic function, the accuracy of this algorithm is exponential; that is, the error decreases as $O(\alpha^N)$ where $0 < \alpha < 1$.

17.2.4 Nonperiodic Functions and Chebyshev Method

In the preceding discussion, we assume that the domain is periodic. However, many problems in practice are not periodic. For example, a cavity problem is not periodic, and thus the trigonometric representation of the derivatives cannot be used. If one attempts to apply such a representation, the discontinuity at the endpoints generates the Gibbs' phenomenon. Furthermore, the wave field "wraps around" because of the periodicity, thereby corrupting the fields within the computational domain.

Gibbs' Phenomenon and Wraparound Effect

When trigonometric interpolation is used to approximate a discontinuous wave function, a significant overshoot and ringing error is introduced adjacent to each discontinuity. This error is called the Gibbs' phenomenon. Furthermore, when a nonperiodic wave function is interpolated using trigonometric functions, a wraparound effect causes the wave field from other periods to spuriously propagate into the domain of interest.

Runge Phenomenon

One may be tempted to choose a uniform grid and the Lagrange interpolation method to interpolate a nonperiodic function. This is fine if the order of the Lagrange interpolation polynomial is low. However, as one increases the order of the Lagrange polynomials, the numerical error near the edges grows exponentially. This is called the Runge phenomenon for a uniform grid. For example, if

$$f(x) \approx \sum_{i=1}^{N+1} f_i \phi^{(N)}_{i-[N/2-1]}(x) \qquad |x| \le 1 \tag{17.18}$$

the Runge phenomenon causes the error to increase exponentially with N near $x = \pm 1$. Because of the Runge phenomenon, one must choose a nonuniform grid where the gridpoints are clustered near the edge.

Chebyshev Interpolation

The basis of Chebyshev interpolation is that the gridding density per unit length should change with N, so that the density is proportional to $N/(\pi\sqrt{1-\xi^2})$, where $\xi \in [-1, 1]$. An example of such gridding is the set of *Gauss-Chebyshev-Lobatto* (GCL) points given by

$$\xi_m = -\cos(m\pi/N) \qquad m = 0, 1, \cdots, N \tag{17.19}$$

If $x_{\min} \le x \le x_{\max}$, we can first transform x into ξ by

$$x = J_x \xi + (x_{\min} + x_{\max})/2 \tag{17.20}$$

where $J_x = (x_{\max} - x_{\min})/2$ is the Jacobian of the transformation.

Given $\{f_m = f(x_m) = f(\xi_m)\}$ $(m = 0, 1, \ldots, N)$, the function can be interpolated by Lagrange polynomials as

$$\phi_m^{(N)}(x) = \prod_{\substack{n=0 \\ n \neq m}}^{N} \frac{x - x_n}{x_m - x_n} = \prod_{\substack{n=0 \\ n \neq m}}^{N} \frac{\xi - \xi_n}{\xi_m - \xi_n} = \phi_m^{(N)}(\xi) \qquad m = 0, 1, \ldots, N \tag{17.21}$$

where the interpolation polynomial can be written in closed form as

$$\phi_m^{(N)}(\xi) = \frac{(1 - \xi^2)\, T_N'(\xi)\, (-1)^{m+1+N}}{c_m N^2 (\xi - \xi_m)} \tag{17.22}$$

where $c_m = 1 + \delta_{m,0} + \delta_{m,N}$ and T is the Chebyshev polynomial given by $T_n(\xi) = \cos[n \cos^{-1}(\xi)]$.

For example, if $N = 1$, then $\xi_0 = -1$, $\xi_1 = 1$, and

$$f(\xi) \approx 0.5(1 - \xi) f_0 + 0.5(1 + \xi) f_1 \tag{17.23a}$$

$$\frac{df}{dx} = \frac{1}{J_x}(-0.5 f_0 + 0.5 f_1) \tag{17.23b}$$

The derivatives at the gridpoints are

$$\begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = [D]^{(1)} \begin{pmatrix} f_0 \\ f_1 \end{pmatrix} \tag{17.24}$$

where

$$[D]^{(1)} = \frac{1}{2J_x} \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \tag{17.25}$$

is the derivative matrix. Similarly, if $N = 2$, then $\xi_0 = -1$, $\xi_1 = 0$, $\xi_2 = 1$, and we have

$$f(\xi) = 0.5\xi(\xi - 1) f_0 + (1 - \xi^2) f_1 + 0.5\xi(1 + \xi) f_2 \tag{17.26a}$$

$$\frac{df}{dx} = \frac{1}{J_x}\left[(\xi - 0.5) f_0 - 2\xi f_1 + (\xi + 0.5) f_2\right] \tag{17.26b}$$

The derivative matrix is now

$$[D]^{(2)} = \frac{1}{2J_x} \begin{pmatrix} -3 & 4 & -1 \\ -1 & 0 & 1 \\ 1 & -4 & 3 \end{pmatrix} \tag{17.27}$$

For an arbitrary positive integer N, one can obtain the general formula

$$D_{mn}^{(N)} = \frac{1}{J_x} \cdot \begin{cases} \dfrac{c_m}{c_n} \dfrac{(-1)^{m+n}}{\xi_m - \xi_n} & m \neq n \\ -\dfrac{\xi_n}{2(1-\xi_n^2)} & 1 \leq m = n \leq N-1 \\ -\dfrac{2N^2+1}{6} & m = n = 0 \\ \dfrac{2N^2+1}{6} & m = n = N \end{cases} \tag{17.28}$$

It seems that, in general, to find the derivative, one would need to multiple this dense matrix $[D]^{(N)}$ with the array $\{f\}$, thus requiring $O(N^2)$ operations. However, this can be circumvented by the fast cosine transform if one views this in an alternative way.

Since the Lagrange polynomials used above are of order N, then from (17.22) we can expand the function $f(x)$ by Chebyshev polynomials up to order N:

$$f(\xi) = \sum_{n=0}^{N} a_n T_n(\xi) \tag{17.29}$$

where a_n are the expansion coefficients. Examples of Chebyshev polynomials are

$$T_0(\xi) = 1 \; ; \quad T_1(\xi) = \xi \; ; \quad T_2(\xi) = 2\xi^2 - 1 \; ; \quad T_3(\xi) = 4\xi^3 - 3\xi \tag{17.30a, b, c, d}$$

Some useful recursion relations for Chebyshev polynomials are

$$T_{n+1}(\xi) = 2\xi T_n(\xi) - T_{n-1}(\xi) \tag{17.31a}$$

$$\frac{T'_{n+1}(\xi)}{n+1} - \frac{T'_{n-1}(\xi)}{n-1} = 2T_n(\xi) \tag{17.31b}$$

$$\left(1-\xi^2\right) T'_n(\xi) = -n\xi T_n(\xi) + n T_{n-1}(\xi) \tag{17.31c}$$

$$2T_m(\xi)\, T_n(\xi) = T_{n+m}(\xi) + T_{|n-m|}(\xi) \tag{17.31d}$$

Therefore, from $T_0(\xi)$ and $T_1(\xi)$, one can obtain all higher-order Chebyshev polynomials.

Now, using (17.29), we can obtain the derivative of $f(\xi)$ as

$$\frac{df(\xi)}{d\xi} = \sum_{n=0}^{N} a_n \frac{dT_n(\xi)}{d\xi} = \sum_{n=0}^{N} b_n T_n(\xi) \tag{17.32}$$

The coefficients $\{b_n\}$ can be derived through the recursion relations of the Chebyshev polynomials. Comparing (17.32) and the derivative of (17.29), we then obtain

$$\begin{aligned} b_N &= 0 \\ b_{N-1} &= 2Na_N \\ b_{N-2} &= 2(N-1)a_{N-1} \\ &\vdots \\ b_{n-1} &= b_{n+1} + 2na_n \qquad n = N-2,\ N-3,\ \cdots,\ 2 \\ &\vdots \\ b_0 &= a_1 + 0.5b_2 \end{aligned} \tag{17.33}$$

With the choice of the gridpoints in (17.19), we can obtain the coefficients $\{a_n\}$ and $\{b_n\}$ using the *fast cosine transform* (FCT) algorithm. First, $\{a_n\}$ can be obtained by the inverse FCT, since

$$\begin{aligned} f(\xi_m) &= \sum_{n=0}^{N} a_n T_n(\xi_m) = \sum_{n=0}^{N} a_n \cos\left\{n\cos^{-1}\left[-\cos(m\pi/N)\right]\right\} \\ &= \sum_{n=0}^{N} a_n \cos\left\{n\left[\pi - (m\pi/N)\right]\right\} \end{aligned} \tag{17.34}$$

In other words, coefficients $\{a_n\}$ are just the inverse cosine transform of $\{f_m\}$. Then, $\{b_n\}$ can be obtained from $\{a_n\}$ using (17.33). Finally, the derivative at the gridpoints can be obtained via (17.32) through the forward cosine transform:

$$\frac{df(\xi_m)}{d\xi} = \sum_{n=0}^{N} b_n \cos\left\{n\left[\pi - (m\pi/N)\right]\right\} \tag{17.35}$$

Thus, the cost of finding the derivative is $O(N\log N)$ using the GCL points.

Legendre Interpolation

A similar approach can be developed for the Legendre interpolation method. However, in this case, the cost is, in general, $O(N^2)$, since one cannot use the FCT algorithm to speed up the derivative computation. Nevertheless, this is not a problem if $N \leq 16$, since it is usually faster to do the direct matrix-vector multiplication than the FCT for smaller N values.

17.3 SINGLE-DOMAIN FOURIER PSTD METHOD

In contrast to the FDTD method, which uses a staggered grid, the Fourier PSTD method uses a grid wherein all vector field components are located at the cell centers. This centered grid provides an important advantage over FDTD in specifying material properties, especially for anisotropic media [3], and removes the field-singularity problem in cylindrical coordinates [12].

The single-domain Fourier PSTD mesh includes the modeling volume and the PML in a single, uniform space lattice. If the dimensions of the problem are $L_x \times L_y \times L_z$, and $N_x \times N_y \times N_z$ cells are used to discretize the problem, then we have $\Delta w = L_w / N_w$, where $w = x, y, z$. Then, using classic Yee leapfrog time-stepping, the spatially collocated field components are given by

$$\boldsymbol{E}^n(i, j, k) = \boldsymbol{E}\big|^{n\Delta t}_{(i+1/2)\Delta x,\,(j+1/2)\Delta y,\,(k+1/2)\Delta z} \tag{17.36a}$$

$$\boldsymbol{H}^{n+1/2}(i, j, k) = \boldsymbol{H}\big|^{(n+1/2)\Delta t}_{(i+1/2)\Delta x,\,(j+1/2)\Delta y,\,(k+1/2)\Delta z} \tag{17.36b}$$

A higher-order time-integration scheme can be used to improve the accuracy of the time-stepping. Such a scheme based on the fourth-order, five-stage Runge-Kutta method will be presented later in the context of the Chebyshev PSTD technique. If this is used for the Fourier PSTD method, the components of $\boldsymbol{E}$ and $\boldsymbol{H}$ must be collocated in time instead of being staggered by one-half time-step, as for Yee updates.

17.3.1 Approximation of Spatial Derivatives

In the Fourier PSTD method, computation of the field-component spatial derivatives is accomplished by the FFT algorithm through (17.14). For example:

$$\frac{\partial\left\{E_y^n(:, j_0, k_0)\right\}}{\partial x} = \frac{2\pi}{N_x \Delta x} \mathrm{FFT}_x^{-1}\left[j n_x \mathrm{FFT}_x\left(\left\{E_y^n(:, j_0, k_0)\right\}\right)_{n_x} \right] \tag{17.37}$$

where n_x is the index of the FFT, and the : symbol denotes all i-coordinates along the straight-line cut through the space lattice at (j_0, k_0). Note that this formulation involves only the one-dimensional FFT and the one-dimensional inverse FFT, each of which is calculated efficiently with only $O(N_x \log N_x)$ operations.

Equation (17.37) yields an approximation to the set of spatial derivatives that has *spectral accuracy* (i.e., the error decreases exponentially as N_x increases), given that the fields are analytic (i.e., the medium is a very smooth function of space). For such media, a discretization of only two grid cells per minimum wavelength (i.e., the Nyquist sampling limit) is required.

We note that, for piecewise-smooth media, there is some degradation in accuracy because of the Gibbs' phenomenon. However, because Maxwell's equations guarantee that tangential field components are continuous across material interfaces, the resulting degree of smoothness permits the spectral calculation of derivatives to be very accurate, even for material interfaces exhibiting reasonably large contrasts in properties across the interface. For the extreme case where the contrast is infinity (e.g., a perfect conductor), however, one of these field components (e.g., the tangential component of the magnetic field) is no longer continuous, and a significant Gibbs' phenomenon arises. In practice, for very high material contrasts, a higher grid density is required in order to adequately model the higher spatial frequency components. Alternatively, the *nonuniform fast Fourier transform* (NUFFT) algorithm [13, 14] can be used, which samples more densely near high contrasts. The latter approach represents a better use of the field-sampling locations. This improvement is implemented in [15]. Alternatively, one can use the nonuniform cosine transform algorithm in a Chebyshev PSTD method [16].

17.3.2 Numerical Stability and Dispersion

References [1, 2] reported numerical stability and dispersion analyses for the Fourier PSTD algorithm when modeling a homogeneous, lossless medium. The numerical stability condition can be written compactly as

$$\frac{c\Delta t}{\Delta x} \leq \frac{2}{\pi\sqrt{D}} \tag{17.38}$$

for a problem of dimensionality D. Note that, for a given Δx, Δt must be reduced by a factor of $\pi/2 \approx 1.57{:}1$ relative to FDTD. However, in practice, for electrically large problems not requiring modeling of geometrical features smaller than one-quarter wavelength, the choice of Δt in the Fourier PSTD method is usually dictated by considerations of accuracy rather than numerical stability, as discussed in Chapter 4, Section 4.9.4, and reviewed below.

For the Fourier PSTD method, the following expressions, (4.96) and (4.97), repeated for convenience from Section 4.9.4, are obtained for the numerical phase-velocity figures of merit:

$$\Delta\tilde{v}_{\text{physical}}\Big|_{\text{PSTD}} = \left[\frac{\pi / N_T}{\sin\left(\pi / N_T\right)} - 1\right] \times 100\%$$

$$\Delta\tilde{v}_{\text{aniso}}\Big|_{\text{PSTD}} = 0 \qquad N_\lambda \geq 2$$

where $N_T = T/\Delta t$ is the temporal sampling density in time samples per wave-oscillation period.

However, the remarkable nulling of the numerical phase-velocity anisotropy implied by (4.97) for spatial sampling densities exceeding the Nyquist rate does *not* mean that the Fourier PSTD technique yields perfect results analogous to the magic-time-step case of the one-dimensional scalar wave equation. In fact, (4.96) shows that there remains a numerical phase-velocity error relative to c. This residual velocity error is not a function of the wave-propagation direction ϕ, and is therefore isotropic within the space grid. The residual velocity error arises from the Yee-type leapfrog time-stepping used in the algorithm, and is a function only of N_T. As noted in Section 4.9.4, the density of the PSTD time sampling must therefore increase with the electrical size of the modeling problem if we require a fixed upper bound on the maximum total phase error of propagating waves within the mesh. Recent numerical experiments have shown that, when modeling large, low-contrast dielectric structures ($\varepsilon_r \sim 1.44$), values of N_T of at least 60 are required to properly resolve the shortest-duration sinusoidal period of interest [17].

Despite the potential need for a small Δt, the Fourier PSTD technique can provide a *very large* reduction in computer resources relative to the classic Yee algorithm for electrically large problems not having spatial details or material inhomogeneities smaller than $\lambda_{\min}/2$. References [1–3] report that, for problem sizes up to 64 wavelengths, the use of Fourier PSTD permits approximately an $8^D{:}1$ reduction in computer storage and running-time relative to the Yee algorithm to produce results with comparable accuracy, where D is the problem dimensionality. Even greater reductions in computer resources are obtainable for larger problems. Additional results using the Fourier PSTD method for large-scale problems have been reported in [12] for cylindrical coordinates, in [18] for large-scale scattering objects, in [19–21] for dispersive media, and in [3] for anisotropic media. The reader is referred to these papers for more details.

17.4 SINGLE-DOMAIN CHEBYSHEV PSTD METHOD

17.4.1 Spatial and Temporal Grids

Similar to the single-domain Fourier PSTD method, the single-domain Chebyshev PSTD method adopts a collocated spatial grid, but with nonuniform gridpoints to treat a nonperiodic problem. The temporal grid, however, is a collocated grid instead of a staggered grid, to accommodate higher-order time-integration schemes.

The single-domain Chebyshev PSTD mesh can be used for bounded or unbounded nonperiodic problems with a smooth material distribution. For an unbounded domain, it includes the modeling volume and the PML in a single, nonuniform space lattice with the outer lattice boundary a PEC or PMC surface.

A simple cubic domain of dimension $L_x \times L_y \times L_z$ and $(N_x+1)\times(N_y+1)\times(N_z+1)$ gridpoints in Cartesian coordinates can be transformed into a standard cube in coordinates (ξ, η, ζ) through the linear coordinate transformation:

$$x_i = J_x \xi_i + 0.5(x_{\min} + x_{\max}) \tag{17.39a}$$

$$y_j = J_y \eta_j + 0.5(y_{\min} + y_{\max}) \tag{17.39b}$$

$$z_k = J_z \zeta_k + 0.5(z_{\min} + z_{\max}) \tag{17.39c}$$

where $(\xi, \eta, \zeta) \in [-1, 1]$, and the Jacobians are $J_\alpha = 0.5(\alpha_{\max} - \alpha_{\min})$.

Furthermore, a more general curved hexahedral domain can also be transformed into (ξ, η, ζ) coordinates by a curvilinear transformation, if the curved hexahedral domain is described by a number of anchor points on the surface and inside the domain. In general, we can approximate the curved domain by a high-order polynomial, for example, $O(P, Q, R)$ in the (ξ, η, ζ) directions. If we denote these anchor points as $(x_{pqr}, y_{pqr}, z_{pqr})$ for $(0 \le p \le P, 0 \le q \le Q, 0 \le r \le R)$, then

$$x = \sum_{p=0}^{P}\sum_{q=0}^{Q}\sum_{r=0}^{R} x_{pqr}\,\phi_p^{(P)}(\xi)\,\phi_q^{(Q)}(\eta)\,\phi_r^{(R)}(\zeta) \tag{17.40}$$

$$y = \sum_{p=0}^{P}\sum_{q=0}^{Q}\sum_{r=0}^{R} y_{pqr}\,\phi_p^{(P)}(\xi)\,\phi_q^{(Q)}(\eta)\,\phi_r^{(R)}(\zeta) \tag{17.41}$$

$$z = \sum_{p=0}^{P}\sum_{q=0}^{Q}\sum_{r=0}^{R} z_{pqr}\,\phi_p^{(P)}(\xi)\,\phi_q^{(Q)}(\eta)\,\phi_r^{(R)}(\zeta) \tag{17.42}$$

Note that the simple cubic domain is just a special case of this with $P = Q = R = 1$. Inside the domain, the inverse transform of these equations yields

$$\xi = \xi(x,y,z), \qquad \eta = \eta(x,y,z), \qquad \zeta = \zeta(x,y,z) \tag{17.43a, b, c}$$

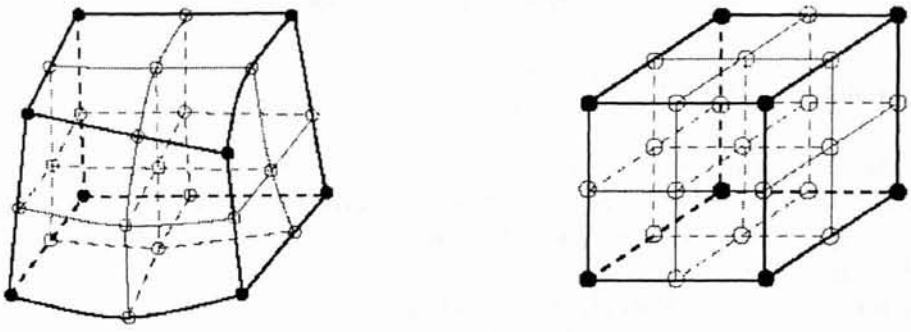


Fig. 17.1 Representative quadratic element in the PSTD mesh with three anchor points on each edge. (a) Curved element in the physical space, and (b) standard cubic element in the transformed coordinates.

Fig. 17.1 illustrates a representative element in the PSTD mesh with curved geometry. Each element is completely described by some anchor points, and can be made conformal to the actual problem geometry with high-order Lagrange polynomials. In this figure, for simplicity a quadratic element is shown with three anchor points on each edge. For the more general high-order elements, the number of anchor points is determined by the polynomial approximation of the geometry.

In the transformed coordinates, the gridpoints can be chosen as the tensor product of GCL points in the ξ, η, and ζ directions:

$$\xi_i = -\cos\left(i\pi / N_x\right) \qquad i = 0, 1, \cdots, N_x \tag{17.44a}$$

$$\eta_j = -\cos\left(j\pi / N_y\right) \qquad j = 0, 1, \cdots, N_y \tag{17.44b}$$

$$\zeta_k = -\cos\left(k\pi / N_z\right) \qquad k = 0, 1, \cdots, N_z \tag{17.44c}$$

The collocated field components are defined at

$$\boldsymbol{E}^n(i, j, k) = \boldsymbol{E}(\xi_i, \eta_j, \zeta_k, n\Delta t) \; ; \quad \boldsymbol{H}^n(i, j, k) = \boldsymbol{H}(\xi_i, \eta_j, \zeta_k, n\Delta t) \tag{17.45a, b}$$

where the temporal grid is also collocated. Note that these gridpoints for the fields can be in general different from the anchor points for the geometry transformation.

17.4.2 Maxwell's Equations in Curvilinear Coordinates

Maxwell's equations can be equivalently written into the conservation form

$$\frac{\partial\{q\}}{\partial t} + \nabla \cdot \boldsymbol{F} = -[D]\{q\} \tag{17.46}$$

where $\{q\} = \{H_x, H_y, H_z, E_x, E_y, E_z\}$, $\{F\} = \{F_x, F_y, F_z\} = \{[A]\{q\}, [B]\{q\}, [C]\{q\}\}$, and

$$[A] = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\mu} \\ 0 & 0 & 0 & 0 & \frac{1}{\mu} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\varepsilon} & 0 & 0 & 0 \\ 0 & -\frac{1}{\varepsilon} & 0 & 0 & 0 & 0 \end{pmatrix} \qquad [B] = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \frac{1}{\mu} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\mu} & 0 & 0 \\ 0 & 0 & -\frac{1}{\varepsilon} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\varepsilon} & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{17.47a, b}$$

$$[C] = \begin{pmatrix} 0 & 0 & 0 & 0 & -\frac{1}{\mu} & 0 \\ 0 & 0 & 0 & \frac{1}{\mu} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\varepsilon} & 0 & 0 & 0 & 0 \\ -\frac{1}{\varepsilon} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad [D] = \frac{\sigma}{\varepsilon} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{17.48a, b}$$

The corresponding flux defined in the curvilinear coordinates $\{\hat{F}\} = \{\hat{F}_\xi, \hat{F}_\eta, \hat{F}_\zeta\}$ must satisfy the contravariant mapping [22] and maintain a divergence-conforming relationship with the original flux in the physical coordinates (x, y, z):

$$\begin{bmatrix} \hat{F}_\xi(i) \\ \hat{F}_\eta(i) \\ \hat{F}_\zeta(i) \end{bmatrix} = [J]^{-1} \left|[J]\right| \begin{bmatrix} F_x(i) \\ F_y(i) \\ F_z(i) \end{bmatrix} \qquad i = 1, 2, \cdots, 6 \tag{17.49}$$

where $\left|[\cdot]\right|$ denotes the determinant of a matrix, and the Jacobian matrix is defined by

$$[J] = \frac{\partial(x, y, z)}{\partial(\xi, \eta, \zeta)} = \begin{bmatrix} x_\xi & x_\eta & x_\zeta \\ y_\xi & y_\eta & y_\zeta \\ z_\xi & z_\eta & z_\zeta \end{bmatrix} \tag{17.50}$$

Defining $\{\hat{q}\} = \left|[J]\right| \{q\}$ for convenience, Maxwell's equations in the transformed curvilinear coordinates become

$$\frac{\partial\{\hat{q}\}}{\partial t} + \nabla_\xi \cdot \hat{\boldsymbol{F}} = -[D]\{\hat{q}\} \tag{17.51}$$

with the divergence of the flux functions related by

$$|[J]|\,\nabla \cdot \boldsymbol{F} = \nabla_\xi \cdot \hat{\boldsymbol{F}} = \frac{\partial[\hat{A}]\{\hat{q}\}}{\partial\xi} + \frac{\partial[\hat{B}]\{\hat{q}\}}{\partial\eta} + \frac{\partial[\hat{C}]\{\hat{q}\}}{\partial\zeta} \tag{17.52}$$

where

$$[\hat{A}] = \frac{\left|\frac{\partial(y,z)}{\partial(\eta,\zeta)}\right|[A] + \left|\frac{\partial(z,x)}{\partial(\eta,\zeta)}\right|[B] + \left|\frac{\partial(x,y)}{\partial(\eta,\zeta)}\right|[C]}{|[J]|} = \xi_x[A] + \xi_y[B] + \xi_z[C] \tag{17.53a}$$

$$[\hat{B}] = \frac{\left|\frac{\partial(y,z)}{\partial(\zeta,\xi)}\right|[A] + \left|\frac{\partial(z,x)}{\partial(\zeta,\xi)}\right|[B] + \left|\frac{\partial(x,y)}{\partial(\zeta,\xi)}\right|[C]}{|[J]|} = \eta_x[A] + \eta_y[B] + \eta_z[C] \tag{17.53b}$$

$$[\hat{C}] = \frac{\left|\frac{\partial(y,z)}{\partial(\xi,\eta)}\right|[A] + \left|\frac{\partial(z,x)}{\partial(\xi,\eta)}\right|[B] + \left|\frac{\partial(x,y)}{\partial(\xi,\eta)}\right|[C]}{|[J]|} = \zeta_x[A] + \zeta_y[B] + \zeta_z[C] \tag{17.53c}$$

can be derived from (17.49). With this transformation, we can then solve Maxwell's equations (17.51) in the (ξ, η, ζ) coordinates for a general curved hexahedron.

17.4.3 Spatial Derivatives

In multidimensions, the spatial derivatives in the Chebyshev PSTD method can be simply obtained by the tensor product of one-dimensional derivatives. A field function $f(\xi, \eta, \zeta)$ defined in the standard cube is interpolated by the Chebyshev-Lagrange interpolation polynomials $\phi_i^{(N_\xi)}(\xi)$, $\phi_j^{(N_\eta)}(\eta)$, and $\phi_k^{(N_\zeta)}(\zeta)$ of (17.22) as

$$f(\xi,\eta,\zeta) = \sum_{i=0}^{N_\xi}\sum_{j=0}^{N_\eta}\sum_{k=0}^{N_\zeta} f(\xi_i,\eta_j,\zeta_k)\,\phi_i^{(N_\xi)}(\xi)\,\phi_j^{(N_\eta)}(\eta)\,\phi_k^{(N_\zeta)}(\zeta) \tag{17.54}$$

Thus, the partial derivatives of $f(\xi, \eta, \zeta)$ with respect to ξ, η, and ζ are approximated as

$$\frac{\partial f(\xi_i,\eta_j,\zeta_k)}{\partial\xi} = \sum_{m=0}^{N_\xi} D_{im}^{(\xi)} f(\xi_m,\eta_j,\zeta_k) \tag{17.55a}$$

$$\frac{\partial f(\xi_i,\eta_j,\zeta_k)}{\partial\eta} = \sum_{m=0}^{N_\eta} D_{jm}^{(\eta)} f(\xi_i,\eta_m,\zeta_k) \tag{17.55b}$$

$$\frac{\partial f(\xi_i,\eta_j,\zeta_k)}{\partial\zeta} = \sum_{m=0}^{N_\zeta} D_{km}^{(\zeta)} f(\xi_i,\eta_j,\zeta_m) \tag{17.55c}$$

with the derivative matrix $[D]$ given in (17.28). Note that, even though these derivatives are written here as the multiplication of derivative matrices and field arrays, for large N_ξ, N_η, and N_ζ (greater than 16), the fast cosine transform algorithm can be used to speed up the derivative calculation, as shown in (17.35).

17.4.4 Time-Integration Scheme

With the spatial derivatives represented by the Chebyshev pseudospectral method, we are now ready to perform the time integration of Maxwell's equations. Equation (17.46) can be written compactly as

$$\frac{\partial\{q\}}{\partial t} = \{p(t,\{q\})\} \tag{17.56}$$

where $\{q\}$ is an array that includes six field components in regular (non-PML) media, and the required additional auxiliary field variables in a PML medium. Array $\{p\}$ contains the right-hand side of (17.46) including the spatial derivative terms of (17.55).

Both two-stage, second-order and five-stage, fourth-order Runge-Kutta schemes to temporally integrate (17.56) have been reported [23]. Denoting $\{q\}^n$ as the value of array $\{q\}$ at time $n\Delta t$, the $2N$-storage, M-stage, K-th order Runge-Kutta method is given by

$$\left.\begin{aligned} \{u\}_0 &= \{q\}^n \\ \{k\}_j &= a_j\{k\}_{j-1} + \left\{p\left[(n+c_j)\Delta t,\ \{u\}_{j-1}\right]\right\}\Delta t \\ \{u\}_j &= \{u\}_{j-1} + b_j\{k\}_j \end{aligned}\right\} \quad \forall\, j \in [1, M]$$

$$\{q\}^{n+1} = \{u\}_M \tag{17.57}$$

For the two-stage, second-order scheme, $a_1 = c_1 = 0$, $b_1 = c_2 = 0.5$, $a_2 = -0.5$, and $b_2 = 1$. This scheme has the advantage of self-starting and $2N$-storage. Appendix 17A tabulates the values of a_j, b_j, and c_j for the five-stage, fourth-order Runge-Kutta method.

In some applications where Δt is small, the two-stage, second-order Runge-Kutta scheme is advantageous because it is faster than the five-stage scheme, and provides sufficient accuracy. For both schemes, Δt must provide numerical stability, as discussed in [5, 23]. Unfortunately, there is no closed-form expression for the stability condition.

17.5 MULTIDOMAIN CHEBYSHEV PSTD METHOD

The single-domain Fourier and Chebyshev PSTD methods are ideal for modeling spatial regions having no discontinuities in the material properties, and thus the field components. Although these methods can be applied to discontinuous media, the accuracy is no longer spectral. A better way to model sharp material discontinuities is to use the multidomain PSTD method. This provides the additional benefit of removing staircasing errors associated with the tensor product of one-dimensional grids that are inherent to single-domain PSTD methods.

The multidomain Chebyshev PSTD method has been developed for general curvilinear coordinates [4 – 7]. This method has recently also been developed for lossy media [8 – 11] and for elastic waves [9]. In this section, we present the general formulation of this technique for Maxwell's equations for the cases of three-dimensional and 2.5-dimensional simulations.

In general, the steps involved in implementing the multidomain PSTD method are as follows:

1. The computational domain is divided into a set of nonoverlapping curved hexahedral subdomains conforming to the boundary of the material distribution.

2. Each general hexahedral subdomain is mapped onto a cube in (ξ, η, ζ) coordinates by a curvilinear transformation. The fields within each subdomain are represented by Chebyshev-Lagrange interpolation polynomials of orders (N_ξ, N_η, N_ζ) for the (ξ, η, ζ) directions, respectively. Spatial derivatives with respect to ξ, η, and ζ are evaluated by the derivative matrices or through fast cosine transforms.

3. Given the spatial derivatives, Maxwell's equations are updated by a time-integration scheme such as the fourth-order Runge-Kutta method.

4. At each stage of the time integration, the fields between adjacent subdomains must be reconciled in order to arrive at the correct fields in the whole computational domain. This important step is termed "boundary patching," and is discussed in more detail in the following sections.

One advantage of this approach is that the derivative calculations can be done in parallel for all of the subdomains. Communication between any pair of adjacent subdomains is needed only at their mutual interface at each substep of the time integration.

17.5.1 Subdomain Spatial Derivatives and Time Integration

In order to develop a multidomain Chebyshev PSTD method, the computational domain is divided into hexahedral subdomains that naturally conform to the problem geometry. In general, each subdomain has curved boundaries in (x, y, z) coordinates, which can be mapped onto a cube in $(\xi, \eta, \zeta) \in [-1, 1]$ coordinates by the curvilinear transformation given by (17.40) to (17.43). Then, one can calculate all spatial derivatives with respect to (ξ, η, ζ) in Maxwell's equations (17.51) within each subdomain.

Since the spatial derivatives are found, the time integration for each subdomain can be performed in the same way as in the single-domain Chebyshev PSTD method, *with the exception that the boundary conditions between adjacent subdomains must be reconciled at each substep of the time integration.* We can use either the two-stage, second-order or the five-stage, fourth-order Runge-Kutta schemes to temporally integrate (17.51) within each subdomain.

The boundary patching procedure is an important step in the multidomain PSTD method to enforce the boundary conditions between adjacent subdomains. We next discuss two ways to perform this boundary patching: first, through the method of characteristics, and second, through the physical boundary conditions.

17.5.2 Subdomain Patching by Characteristics

At the end of each substep of the time integration, the fields at the interfaces between adjacent subdomains do not naturally satisfy the appropriate boundary conditions between subdomains. The first subdomain patching technique presented here is based on the characteristic conditions of the hyperbolic system given by Maxwell's equations.

Note that (17.51) is a hyperbolic partial differential equation. We can diagonalize $[\hat{A}]$ by orthonormal matrices:

$$[\hat{A}] = [S][\Lambda][S]^{-1} \tag{17.58}$$

where

$$[\Lambda] = v \cdot \text{diag}(0, 0, 1, 1, -1, -1), \qquad v = \sqrt{\left(\xi_x^2 + \xi_y^2 + \xi_z^2\right)/\varepsilon\mu} \tag{17.59}$$

are the eigenvalues, and matrix $[S]$ consists of the corresponding eigenvectors, which represent the characteristic waves propagating along the ξ direction. The six eigenmodes include two nonpropagating characteristic waves (R_1, R_2) with a speed of zero, two $+\xi$-directed waves (R_3, R_4) with a speed of v, and two $-\xi$-directed waves (R_5, R_6) with a speed of $-v$. The corresponding characteristic vectors can be written as

$$\{R\} = [S]^{-1}\{q\} = \left\{R_1, R_2, R_3, R_4, R_5, R_6\right\}^T \tag{17.60}$$

Similar results can be obtained for matrices $[\hat{B}]$ and $[\hat{C}]$ for characteristic waves propagating along η and ζ directions, respectively.

The characteristic variables can be used to match the fields at the interface separating two subdomains of the same material. Suppose the ξ-axis points from subdomain 1 to subdomain 2. In this case, we match the characteristic waves at the interface as follows:

$$R_1^{(1,2)} \Leftarrow 0.5\left(R_1^{(1)} + R_1^{(2)}\right) \tag{17.61a}$$

$$R_2^{(1,2)} \Leftarrow 0.5\left(R_2^{(1)} + R_2^{(2)}\right) \tag{17.61b}$$

$$R_3^{(2)} \Leftarrow R_3^{(1)} \tag{17.61c}$$

$$R_4^{(2)} \Leftarrow R_4^{(1)} \tag{17.61d}$$

$$R_5^{(1)} \Leftarrow R_5^{(2)} \tag{17.61e}$$

$$R_6^{(1)} \Leftarrow R_6^{(2)} \tag{17.61f}$$

where the superscript denotes the two different subdomains. Essentially, these conditions state that for each subdomain, the incoming characteristic waves are determined by the corresponding outgoing waves from the adjacent subdomain, the outgoing waves are left unaltered, and the nonpropagating waves remain continuous.

17.5.3 Subdomain Patching by Physical Conditions

The characteristic conditions discussed above are used for an interface separating two subdomains with the same material. However, when the two adjacent subdomains have different materials, these characteristic conditions cannot be applied directly, since the characteristic waves become discontinuous across the interface. Instead, we can use the physical boundary conditions to perform subdomain patching.

On the interface of two adjacent subdomains, the physical boundary conditions require the tangential components of $\boldsymbol{E}$ and $\boldsymbol{H}$ to be continuous:

$$\boldsymbol{E}_{\tan}^{(1)} = \boldsymbol{E}_{\tan}^{(2)} ; \qquad \boldsymbol{H}_{\tan}^{(1)} = \boldsymbol{H}_{\tan}^{(2)} \tag{17.62a, b}$$

where the superscripts denote the two subdomains with a common interface. For a dielectric interface, this implies continuity of the tangential $\boldsymbol{E}$ and $\boldsymbol{H}$ components across the interface, while for a PEC interface, the tangential $\boldsymbol{E}$ and normal $\boldsymbol{H}$ components vanish. We first extract the tangential and normal field components:

$$E_{\tan 1} = \boldsymbol{E} \cdot \hat{\boldsymbol{t}}_1 , \qquad E_{\tan 2} = \boldsymbol{E} \cdot \hat{\boldsymbol{t}}_2 , \qquad E_{\text{norm}} = \boldsymbol{E} \cdot \hat{\boldsymbol{n}} \tag{17.63a}$$

$$H_{\tan 1} = \boldsymbol{H} \cdot \hat{\boldsymbol{t}}_1 , \qquad H_{\tan 2} = \boldsymbol{H} \cdot \hat{\boldsymbol{t}}_2 , \qquad H_{\text{norm}} = \boldsymbol{H} \cdot \hat{\boldsymbol{n}} \tag{17.63b}$$

where $\hat{\boldsymbol{t}}_1$, $\hat{\boldsymbol{t}}_2$, and $\hat{\boldsymbol{n}}$ are local unit tangential and normal vectors, respectively. We next force the tangential field components to be continuous:

$$\left(E_{\tan 1}^{(1)}, E_{\tan 1}^{(2)}\right) \Leftarrow 0.5\left(E_{\tan 1}^{(1)} + E_{\tan 1}^{(2)}\right) \qquad \left(E_{\tan 2}^{(1)}, E_{\tan 2}^{(2)}\right) \Leftarrow 0.5\left(E_{\tan 2}^{(1)} + E_{\tan 2}^{(2)}\right) \tag{17.64a}$$

$$\left(H_{\tan 1}^{(1)}, H_{\tan 1}^{(2)}\right) \Leftarrow 0.5\left(H_{\tan 1}^{(1)} + H_{\tan 1}^{(2)}\right) \qquad \left(H_{\tan 2}^{(1)}, H_{\tan 2}^{(2)}\right) \Leftarrow 0.5\left(H_{\tan 2}^{(1)} + H_{\tan 2}^{(2)}\right) \tag{17.64b}$$

The normal field components are left unchanged. This subdomain patching is performed after the field components are updated at each Runge-Kutta stage. The process for field updating and boundary patching continues until the desired time window is completed.

Combinations of characteristic and physical conditions have been successfully applied to subdomain patching [5, 6, 8, 10]. However, when applied to patch three-dimensional subdomains of the same material, using characteristics can be cumbersome, because one has to track the conditions that make some of the characteristic variables zero [7]. Here, it is simpler to patch subdomains by using only the physical boundary conditions [9, 11, 24].

17.5.4 Filter Design for Corner Singularities

As discussed in detail in [25], numerical oscillations can appear when pseudospectral methods are applied to hyperbolic problems with discontinuous solutions. These oscillations are directly caused by the solution discontinuities, and have a high-frequency character. For the multi-domain PSTD scheme, such discontinuities can arise from edges and corners at subdomain interfaces, potentially causing late-time instability.

A viable means to surmount the oscillation and to help improve the stability of pseudospectral solutions is to employ a filtering procedure [26–28]. Filtering accelerates convergence by multiplying the Fourier or Chebyshev coefficients with a gradually decreasing function $\sigma(\cdot)$ to reduce the high-frequency components. For example, a 2π-periodic discrete function u_m can be written in its Fourier series:

$$u_m = \frac{1}{N}\sum_{n=-N/2}^{N/2-1} \breve{u}_n\, e^{j\frac{2\pi}{N}mn} \tag{17.65}$$

Then, the filtered solution is represented as

$$u_m = \frac{1}{N}\sum_{n=-N/2}^{N/2-1} \sigma_n\, \breve{u}_n e^{j\frac{2\pi}{N}mn} \tag{17.66}$$

where the Fourier coefficients are

$$\breve{u}_n = \sum_{m=0}^{N-1} u_m\, e^{-j\frac{2\pi}{N}mn} \tag{17.67}$$

and σ_n is a decreasing function with $|n|$.

Similarly, a function $u(\xi)$ defined within $\xi \in [-1, 1]$ also can be represented in its discrete polynomial transformation, for example, by Chebyshev polynomials

$$u_m = u(\xi_m) = \sum_{n=0}^{N} \breve{u}_n T_n(\xi_m) \tag{17.68}$$

where ξ_m are the GCL points $\xi_m = -\cos(m\pi/n)$, $0 \le m \le N$; and $T_n(x)$ is the n'th-order Chebyshev polynomial $T_n(x_m) = -\cos(mn\pi/N)$. The inverse relationship is

$$\breve{u}_n = \frac{1}{\sum_{j=0}^{N} T_n^2(x_j) w_j} \sum_{j=0}^{N} u_m T_n(x_m) w_m = \sum_{m=0}^{N} \frac{2}{N c_m c_n} u_m T_n(x_m) \tag{17.69}$$

where $c_n = 1 + \delta_{n,0} + \delta_{n,N}$, and the weights for the GCL points are given by $w_0 = w_N = \pi/2N$ and $w_m = 2\pi/N$ for $1 \le m \le N-1$. Then, the filtered solution is derived as

$$\begin{aligned} u_m &= \sum_{n=0}^{N} \sigma_n \breve{u}_n T_n(x_m) = \sum_{n=0}^{N} \sigma_n T_n(x_m) \sum_{j=0}^{N} \frac{2}{N c_j c_n} u_j T_n(x_j) \\ &= \frac{2}{N}\sum_{j=0}^{N} \frac{u_j}{c_j} \sum_{n=0}^{N} \frac{\sigma_n}{c_n} T_n(x_m) T_n(x_j) \end{aligned} \tag{17.70}$$

which can be conveniently expressed into a tensor product form.

The choice of the filtering function $\sigma(\cdot)$ is discussed in [25] and remains an active research topic [29]. This function must be unity in the neighborhood of the origin, and gradually decrease to zero at higher frequencies to filter out the higher modes. Some of the functions that have been discussed include the raised cosine function:

$$\sigma_n = 0.5\left[1+\cos\left(n\pi/N\right)\right] \tag{17.71}$$

and the exponential cutoff function

$$\sigma_n = \begin{cases} 1 & 0 \le n \le n_c \\ e^{-\alpha\left(\frac{n-n_c}{N-n_c}\right)^{\beta}} & n_c \le n \le N \end{cases} \tag{17.72}$$

where n_c is the cutoff number, β is the order of the filter, and $\alpha = -\ln(\varepsilon_M)$ is a constant coefficient, with ε_M being the machine precision. Decreasing either the cutoff number n_c or the order β leads to a stronger filtering action. Reference [29] proposed to choose β adaptively as a function of both the order of projection and the distance to the discontinuity. It has been demonstrated that, with an appropriate implementation of a filter, spectral accuracy can be maintained in the regions away from discontinuities.

17.5.5 Multidomain PSTD Results for 2.5-Dimensional Problems

This section presents two illustrative cases of 2.5-dimensional multidomain PSTD modeling of electromagnetic wave interaction problems. We first consider the field radiated by an impulsive line current source located at the center of a 4 m diameter, air-filled, circular cylindrical hole within an infinite, lossy, homogeneous medium ($\varepsilon_r = 9$, $\mu_r = 1$, and $\sigma = 0.0015$ S/m). This source has a sinusoidal spatial distribution along the z-axis given by $\boldsymbol{J} = \hat{z}\cos(k_z z)\delta(x)\delta(y)$ for $k_z = 0.5$ m^{-1}, and a time function that yields a Blackman-Harris spectral window centered at 30 MHz. Note that the current varies with position along the z-axis with a periodicity that is comparable to the free-space wavelength at the center frequency of the excitation.

Fig. 17.2(a) illustrates the mesh of subdomains used for the PSTD model. A grid with 16×16 GCL points is used within each subdomain, along with a time-step $\Delta t = 21.5$ ps. This represents a meshing density of only 5.66 *points per wavelength* (PPW) in the high-permittivity surrounding medium for the highest-frequency spectral component of importance, 90 MHz. Fig. 17.2(b) shows excellent agreement between the PSTD numerical result and the analytical solution for the time waveform of the radiated E_z field at observation location $(x, y, z) = (-2.0135, -2.0135, 0)$.

In the second example, we study the same problem as that above, except that a point electric dipole $\boldsymbol{J} = \hat{z}\,\delta(x)\delta(y)\delta(z)$ is used. The fields in this case can be obtained by performing the inverse cosine/sine transform of the fields obtained in the first example with different values of k_z. In this particular case, only the cosine transform is needed for E_z because $\boldsymbol{J}$ is an even function of z. Fig. 17.3 compares the PSTD numerical result and the analytical solution for the time waveform of the radiated E_z field at observation location $(x, y, z) = (-2.0135, -2.0135, 1)$. The agreement between the two results is excellent despite the low sampling density of 5.66 PPW for the highest-frequency spectral component of importance.

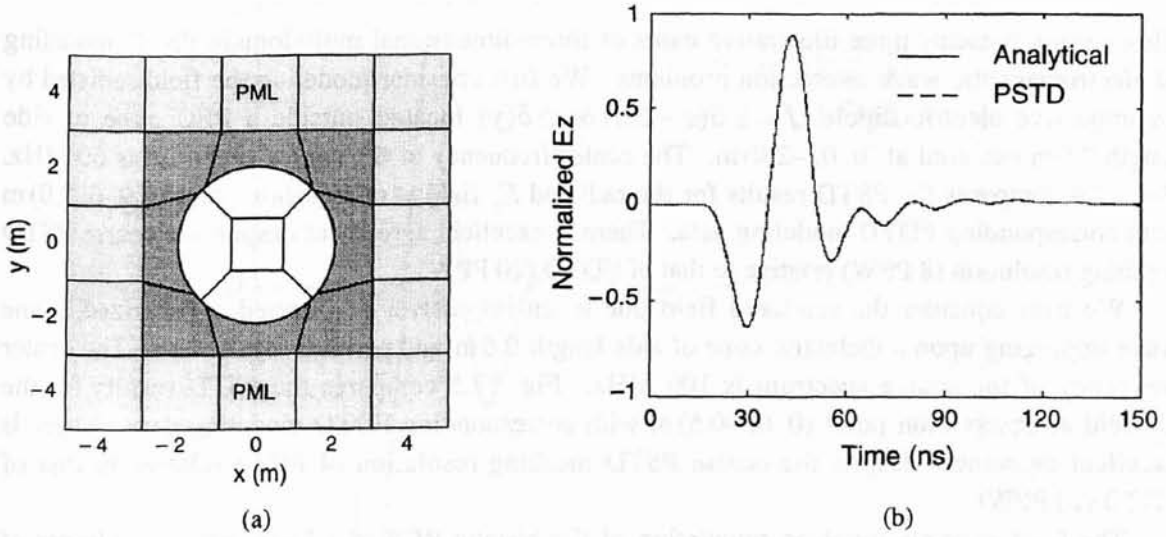


Fig. 17.2 Multidomain PSTD model of a 2.5-dimensional problem involving the field radiated by an impulsive line current source. The current source is located at the center of a 4m diameter, air-filled, circular cylindrical hole within an infinite, lossy, homogeneous medium ($\varepsilon_r = 9$, $\mu_r = 1$, and $\sigma = 0.0015$ S/m). The current source has a sinusoidal spatial distribution along the z-axis given by $\boldsymbol{J} = \hat{z}\cos(k_z z)\,\delta(x)\,\delta(y)$ for $k_z = 0.5$ m^{-1}, and a time function that yields a Blackman-Harris spectral window centered at 30 MHz. (a) Subdomain geometry, and (b) comparison of PSTD numerical result and analytical solution for the time waveform of the radiated E_z field at observation location $(x, y, z) = (-2.0135, -2.0135, 0)$.

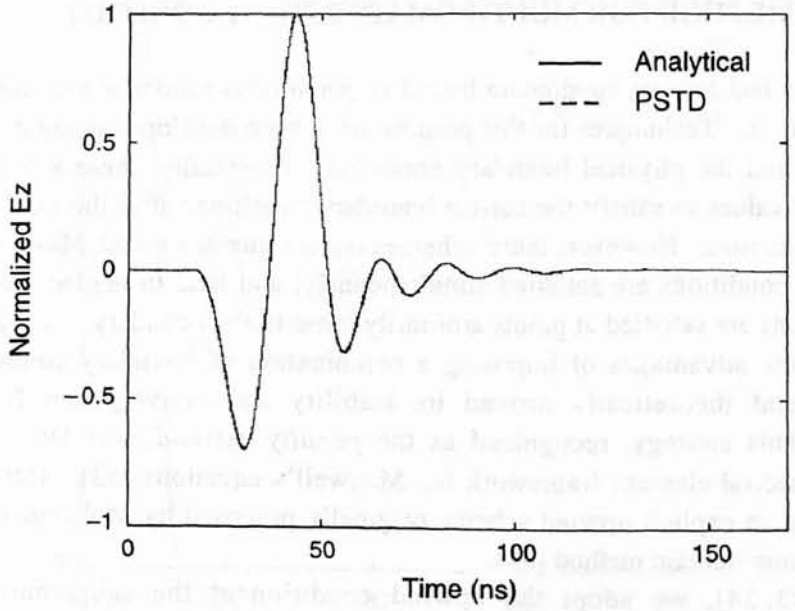


Fig. 17.3 Results of the same multidomain PSTD model as in Fig. 17.2(a), except that a point electric dipole $\boldsymbol{J} = \hat{z}\,\delta(x)\,\delta(y)\,\delta(z)$ is used for the excitation. This graph compares the PSTD numerical result and the analytical solution for the time waveform of the radiated E_z field at observation location $(x, y, z) = (-2.0135, -2.0135, 1)$.

17.5.6 Multidomain PSTD Results for Three-Dimensional Problems

This section presents three illustrative cases of three-dimensional multidomain PSTD modeling of electromagnetic wave interaction problems. We first consider modeling the field radiated by an impulsive electric dipole $\boldsymbol{J} = \hat{z}\,\delta(z-2.0)\,\delta(x)\,\delta(y)$ located outside a PEC cube of side length 2.0 m centered at (0, 0, –2.0) m. The center frequency of the source spectrum is 60 MHz. Fig. 17.4 compares the PSTD results for the radiated E_z field at observation point (2.0, 0, 2.0) m with corresponding FDTD modeling data. There is excellent agreement despite the coarse PSTD meshing resolution (8 PPW) relative to that of FDTD (20 PPW).

We next consider the scattered field due to an impulsive, +z-directed, x-polarized, plane wave impinging upon a dielectric cube of side length 0.6 m and permittivity $\varepsilon_r = 16$. The center frequency of the source spectrum is 100 MHz. Fig. 17.5 compares the PSTD results for the E_x field at observation point (0, 0, –0.6) m with corresponding FDTD modeling data. There is excellent agreement despite the coarse PSTD meshing resolution (4 PPW) relative to that of FDTD (10 PPW).

The final example involves calculation of the bistatic RCS of a lossy circular cylinder of radius and length $a = 2$ m, permittivity $\varepsilon_r = 4$, and conductivity $\sigma = 0.0015$ S/m. Illumination is provided by an impulsive, –z-directed, x-polarized, plane wave with a spectrum centered at 50 MHz. This problem is designed to test the ability of the PSTD algorithm to simulate curved and conducting objects. Fig. 17.6(a) illustrates the geometry of the cylinder and the PSTD mesh in a transverse cross section, showing the conformal subdomain decomposition. Fig. 17.6(b) shows the good agreement of the PSTD RCS calculations with corresponding results obtained using the phasor-domain, stabilized biconjugate-gradient FFT method [30, 31].

17.6 PENALTY METHOD FOR MULTIDOMAIN PSTD ALGORITHM

As discussed in the last section, subdomain boundary patching is needed at the end of each stage of the time integration. Techniques for this purpose have been developed using both the method of characteristics and the physical boundary condition. Essentially, these schemes explicitly reconcile the field values to satisfy the correct boundary conditions after the field is updated by the differential equations. However, these schemes do not guarantee that Maxwell's equations and the boundary conditions are satisfied simultaneously, and tend to neglect the fact that the differential equations are satisfied at points arbitrarily close to the boundary. In [32], Funaro and Gottlieb showed the advantages of imposing a combination of boundary conditions and the equation itself, and theoretically proved its stability and convergence for Chebyshev approximations. This strategy, recognized as the *penalty method*, was later developed by Hesthaven in a spectral-element framework for Maxwell's equations [33]. Here, the penalty term is realized by an explicit upwind scheme originally proposed by Mohammadian et al. for the finite-volume time-domain method [34].

Following [33, 34], we adopt the upwind condition at the subdomain boundaries. The corrected tangential field components at a dielectric interface can be determined by

$$\hat{\boldsymbol{n}} \times \boldsymbol{E}^{*} = \hat{\boldsymbol{n}} \times \frac{(Y\boldsymbol{E} - \hat{\boldsymbol{n}} \times \boldsymbol{H})^{(1)} + (Y\boldsymbol{E} + \hat{\boldsymbol{n}} \times \boldsymbol{H})^{(2)}}{Y^{(1)} + Y^{(2)}} \tag{17.73a}$$

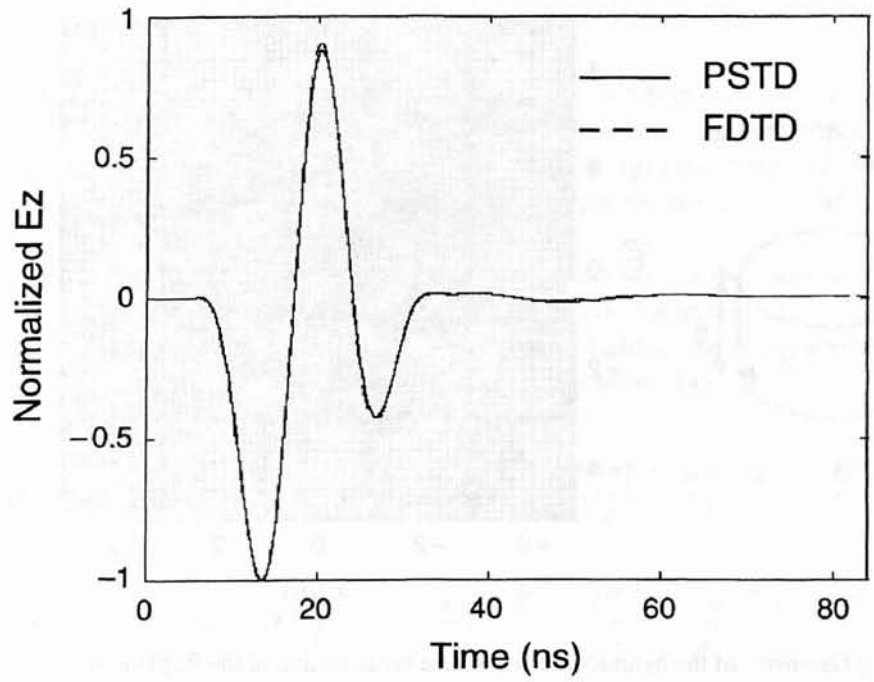


Fig. 17.4 Comparison of PSTD (8 PPW) and FDTD (20 PPW) modeling results for the E_z field radiated by an impulsive current source adjacent to a PEC cube.

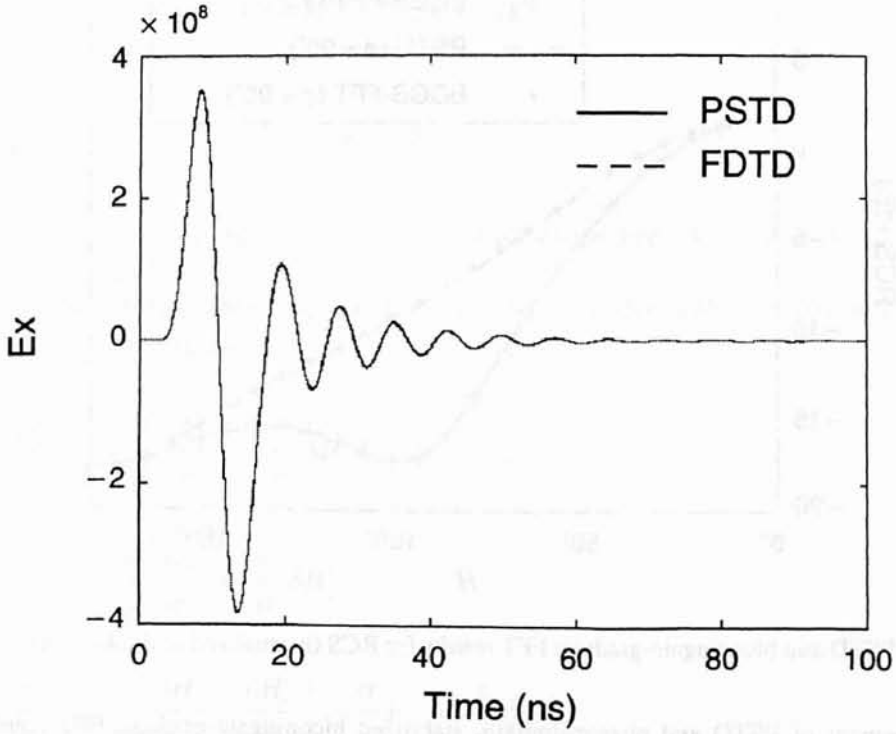
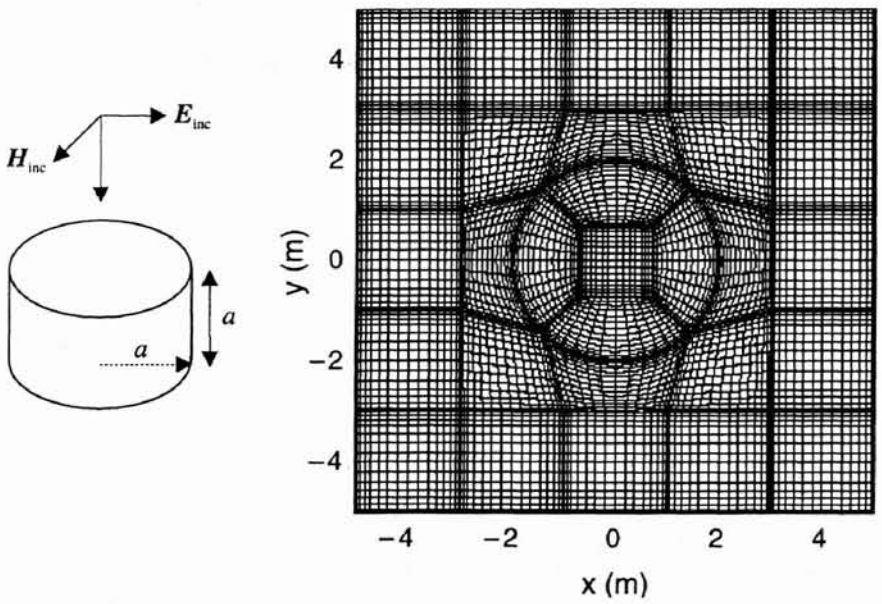
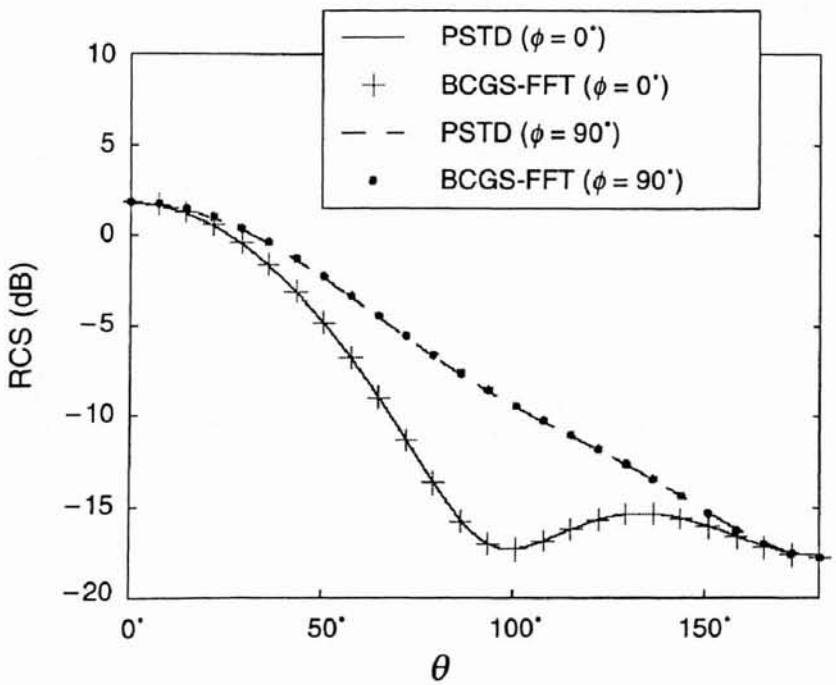


Fig. 17.5 Comparison of PSTD (4 PPW) and FDTD (10 PPW) modeling results for the E_x field scattered by an $\varepsilon_r = 16$ dielectric cube illuminated by an impulsive plane wave.



(a) Geometry of the cylinder and transverse cross section of the PSTD mesh.



(b) PSTD and biconjugate-gradient FFT results for RCS (normalized to λ^2, $\lambda = 6$m).

Fig. 17.6 Agreement of PSTD and phasor-domain, stabilized biconjugate-gradient FFT results for the bistatic RCS of a lossy circular cylinder of radius = length = $a = 2$m, permittivity $\varepsilon_r = 4$, and conductivity $\sigma = 0.0015$ S/m. Illumination is provided by an impulsive, $-z$-directed, x-polarized, plane wave with a spectrum centered at 50 MHz.

$$\hat{\boldsymbol{n}} \times \boldsymbol{H}^{*} = \hat{\boldsymbol{n}} \times \frac{\left(Z\boldsymbol{H} + \hat{\boldsymbol{n}} \times \boldsymbol{E}\right)^{(1)} + \left(Z\boldsymbol{H} - \hat{\boldsymbol{n}} \times \boldsymbol{E}\right)^{(2)}}{Z^{(1)} + Z^{(2)}} \tag{17.73b}$$

where the superscripts (1) and (2) denote two adjacent subdomains with normal direction $\hat{\boldsymbol{n}}$ pointing from subdomain 1 to subdomain 2, and * denotes the corrected field values for both subdomains to replace the current field components on the right-hand side of (17.73). Note that both characteristic and physical boundary conditions are included in (17.73), which is also known as the Riemann solver of Maxwell's equations.

The ideal situation is that the corrected field components equal the current field components, in which case the boundary conditions are already satisfied. Otherwise, it is natural to consider using the difference of the corrected field and the current field as the penalty term for Maxwell's equations. For subdomain 1, this leads to the following equations [33]:

$$\hat{\boldsymbol{n}} \times \Delta\boldsymbol{E}^{(1)} \equiv \hat{\boldsymbol{n}} \times \left(\boldsymbol{E}^{*} - \boldsymbol{E}^{(1)}\right) = \hat{\boldsymbol{n}} \times \frac{Y^{(2)}\left(\boldsymbol{E}^{(2)} - \boldsymbol{E}^{(1)}\right) + \hat{\boldsymbol{n}} \times \left(\boldsymbol{H}^{(2)} - \boldsymbol{H}^{(1)}\right)}{Y^{(1)} + Y^{(2)}} \tag{17.74a}$$

$$\hat{\boldsymbol{n}} \times \Delta\boldsymbol{H}^{(1)} \equiv \hat{\boldsymbol{n}} \times \left(\boldsymbol{H}^{*} - \boldsymbol{H}^{(1)}\right) = \hat{\boldsymbol{n}} \times \frac{Z^{(2)}\left(\boldsymbol{H}^{(2)} - \boldsymbol{H}^{(1)}\right) - \hat{\boldsymbol{n}} \times \left(\boldsymbol{E}^{(2)} - \boldsymbol{E}^{(1)}\right)}{Z^{(1)} + Z^{(2)}} \tag{17.74b}$$

where the jumps in the fields are $\boldsymbol{E}^{(2)} - \boldsymbol{E}^{(1)}$ and $\boldsymbol{H}^{(2)} - \boldsymbol{H}^{(1)}$. Therefore, the penalized Maxwell's equations on the interface for subdomain 1 can be written as

$$\varepsilon \frac{\partial \boldsymbol{E}^{(1)}}{\partial t} = \nabla \times \boldsymbol{H}^{(1)} - \sigma \boldsymbol{E}^{(1)} + \tau\, \hat{\boldsymbol{n}} \times \Delta\boldsymbol{H}^{(1)} \tag{17.75a}$$

$$\mu \frac{\partial \boldsymbol{H}^{(1)}}{\partial t} = -\nabla \times \boldsymbol{E}^{(1)} - \tau\, \hat{\boldsymbol{n}} \times \Delta\boldsymbol{E}^{(1)} \tag{17.75b}$$

where τ is a positive coefficient to adjust the weight of the penalty term. A similar formulation applies to subdomain 2.

For the two-dimensional Maxwell's equations under TM_y polarization, the above formulation can be written as

$$\frac{\partial H_z}{\partial t} = -\frac{1}{\mu}\frac{\partial E_y}{\partial x} + \frac{\tau}{\mu}\,\delta H_z^{(1)} \tag{17.76a}$$

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu}\frac{\partial E_y}{\partial z} + \frac{\tau}{\mu}\,\delta H_x^{(1)} \tag{17.76b}$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}\right) - \frac{\sigma}{\varepsilon}E_y + \frac{\tau}{\varepsilon}\,\delta E_y^{(1)} \tag{17.76c}$$

where the penalty terms δH_x, δH_y, and δH_z are derived from (17.74):

$$\delta H_z^{(1)} = \frac{n_x^2\left(H_z^{(2)} - H_z^{(1)}\right) - n_z n_x\left(H_x^{(2)} - H_x^{(1)}\right) - n_x Y^{(2)}\left(E_y^{(2)} - E_y^{(1)}\right)}{Y^{(1)} + Y^{(2)}} \tag{17.77a}$$

$$\delta H_x^{(1)} = \frac{n_z^2\left(H_x^{(2)} - H_x^{(1)}\right) - n_z n_x\left(H_z^{(2)} - H_z^{(1)}\right) + n_z Y^{(2)}\left(E_y^{(2)} - E_y^{(1)}\right)}{Y^{(1)} + Y^{(2)}} \tag{17.77b}$$

$$\delta E_y^{(1)} = \frac{Z^{(2)}\left[-n_x\left(H_z^{(2)} - H_z^{(1)}\right) + n_z\left(H_x^{(2)} - H_x^{(1)}\right)\right] + \left(E_y^{(2)} - E_y^{(1)}\right)}{Z^{(1)} + Z^{(2)}} \tag{17.77c}$$

This penalty method has been implemented to achieve stable and accurate treatment for the boundary conditions. It replaces the traditional patching schemes discussed in the previous sections. As an example, Fig. 17.7 shows a calculation of the scattering of a $+x$-directed, impulsive plane wave of center frequency 200 MHz by an $\varepsilon_r = 4.0$ dielectric cube of side length 0.6 m. The penalty PSTD method at 4 PPW and the FDTD method at 10 PPW are used to compute the E_z field at the observation point (–0.6, 0, 0) m. There is excellent agreement of the results.

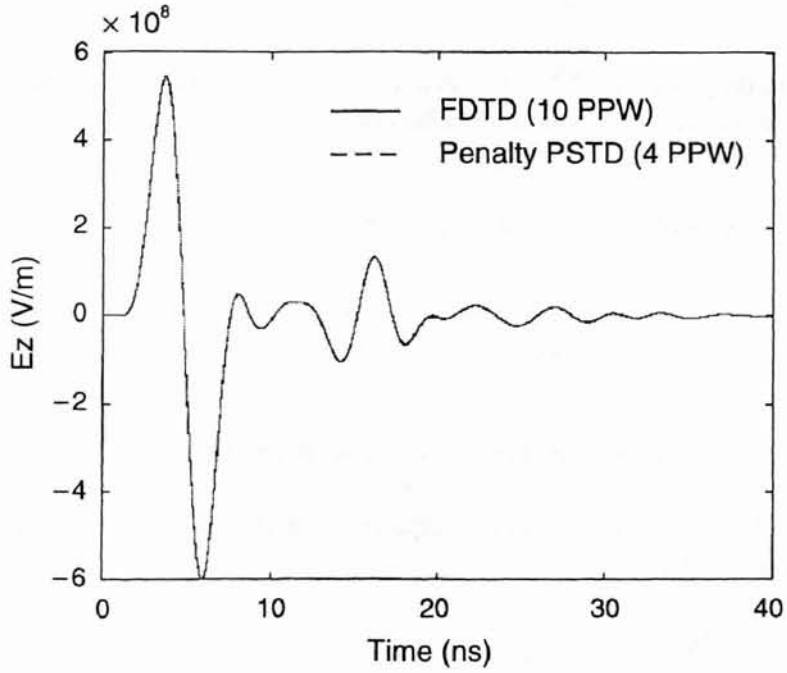


Fig. 17.7 Comparison of results of the penalty PSTD method at 4 PPW and the FDTD method at 10 PPW for the electric field near an $\varepsilon_r = 4.0$ dielectric cube of side length 0.6m illuminated by an impulsive plane wave of center frequency 200 MHz.

The penalty PSTD method achieves stability without the help of filtering [6]. Since implementing an appropriate filter is primarily dependent upon the experience of the modeler, and since an inappropriate filter could degrade the accuracy of the results, we consider the penalty PSTD method superior to the subdomain patching schemes in terms of stability.

17.7 DISCONTINUOUS GALERKIN METHOD FOR PSTD BOUNDARY PATCHING

The *discontinuous Galerkin method* (DGM) has become increasingly popular among the computational fluid dynamics community in the past decade [35]. First introduced by Reed and Hill [36], it has been demonstrated to be a robust, efficient approach for high-order and spectral methods, because it can achieve a uniformly high-order accuracy for complicated geometries.

DGM has been recently applied to solve Maxwell's equations with superior accuracy and efficiency [37, 38]. Here, we apply DGM to the multidomain PSTD algorithm in order to improve the robustness of the interface patching schemes.

17.7.1 Weak Form of Maxwell's Equations

Consider a piecewise homogeneous medium within a subdomain Γ. We can multiply Maxwell's equations in conservation form (17.46) by a testing function $\{p(x, y, z)\}$ and integrate over the subdomain to obtain the weak form of this equation:

$$\left\langle \frac{\partial\{q\}}{\partial t},\ \{p(x,y,z)\} \right\rangle + \left\langle \nabla\cdot \boldsymbol{F},\ \{p(x,y,z)\} \right\rangle = -\left\langle [D]\{q\},\ \{p(x,y,z)\} \right\rangle \tag{17.78}$$

where

$$\langle f, g \rangle = \int_{\Gamma} f(x,y,z)\, g(x,y,z)\, dx\, dy\, dz \tag{17.79}$$

denotes the inner product over the volume Γ. After integration by parts, (17.78) is written as

$$\left\langle \frac{\partial\{q\}}{\partial t}, \{p(x,y,z)\} \right\rangle - \left\langle \boldsymbol{F}, \nabla\{p(x,y,z)\} \right\rangle + \int_{\partial\Gamma} \{p(x,y,z)\}\, \boldsymbol{F}\cdot\hat{\boldsymbol{n}}\, ds = -\left\langle [D]\{q\}, \{p(x,y,z)\} \right\rangle \tag{17.80}$$

17.7.2 Space Discretization and Domain Transformation

Again, a coordinate transformation is performed to transform an arbitrary hexahedron in the physical (x, y, z) domain onto a standard cube $[-1, 1]^3$ in (ξ, η, ζ) coordinates. The spatial discretization of the transformed domain is based on the *Gauss-Legendre-Lobatto* (GLL) points $\{\xi_0 = -1,\ \xi_i\ (1 \le i \le N-1),\ \xi_N = 1\}$, which are zeros of $L'_N(\xi)$, where $L_N(\xi)$ is the N'th order Legendre polynomial. This choice takes advantage of the Gauss quadrature rule

$$\int_{-1}^{1} p(\xi)\, d\xi = \sum_{i=0}^{N} p(\xi_i)\, w_i^{(N)} \tag{17.81}$$

which is exact for polynomial $p(\xi)$ of up to order $(2N - 1)$. Here, the weights for the GLL nodes are given by

$$w_i^{(N)} = 2 \Big/ \left[N(N+1) L_N^2(\xi_i) \right] \qquad 0 \le i \le N \tag{17.82}$$

Therefore, the inner product can be written in the transformed domain (ξ, η, ζ) as

$$\begin{aligned}\langle f, g\rangle &= \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} f\, g\, |[J]|\, d\xi\, d\eta\, d\zeta \\ &= \sum_{i=0}^{I}\sum_{j=0}^{J}\sum_{k=0}^{K} f_{ijk}\, g_{ijk} |[J]|_{ijk}\, w_i^{(I)} w_j^{(J)} w_k^{(K)}\end{aligned} \tag{17.83}$$

where (I, J, K) are the orders of polynomials used in the (ξ, η, ζ) directions, respectively, and $|[J]|_{ijk} = |[J]|(\xi_i, \eta_j, \zeta_k)$.

The testing function is defined in the transformed domain as

$$p_{lmn}(\xi, \eta, \zeta) = \frac{g_l^{(I)}(\xi)\, g_m^{(J)}(\eta)\, g_n^{(K)}(\zeta)}{w_i^{(I)}\, w_j^{(J)}\, w_k^{(K)}} \tag{17.84}$$

where

$$g_l^{(I)}(\xi) = \frac{(1-\xi^2)\, L_I'(\xi)}{I(I+1)(\xi - \xi_l) L_I(\xi_l)} \tag{17.85}$$

is the I'th-order Legendre-Lagrange polynomial having the property $g_l(\xi_i) = \delta_{li}$ for GLL nodes ξ_i $(0 \le i \le I)$.

17.7.3 Mass Matrix and Stiffness Matrix

With the above definitions, and further denoting $|[J]|_{lmn} = |[J]|(\xi_l, \eta_m, \zeta_n)$, the inner product for the time derivative in (17.80) can be written into

$$\begin{aligned}\left\langle \frac{\partial\{q\}}{\partial t}, \{p\} \right\rangle &= \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} \frac{\partial\{q\}}{\partial t} p_{lmn}(\xi, \eta, \zeta)\, |[J]|\, d\xi\, d\eta\, d\zeta \\ &= |[J]|_{lmn} \frac{\partial\{q(\xi_l, \eta_m, \zeta_n)\}}{\partial t} \equiv |[J]|_{lmn} \frac{\partial q_{lmn}}{\partial t}\end{aligned} \tag{17.86}$$

In other words, the mass matrix is diagonal if derived by using the GLL points and the Legendre-Lagrange polynomials as basis and testing functions.

The derivation of the stiffness matrix involves the transformation of the flux from the physical domain to the transformed domain. Given the transform relationship in (17.51), the inner product of the divergence of flux can be directly evaluated in the curvilinear coordinates:

$$\begin{aligned}\left\langle \nabla_\xi \cdot \hat{F}, \{p\} \right\rangle &= \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} \nabla_\xi \cdot \hat{F}\, p_{lmn}\, d\xi\, d\eta\, d\zeta \\ &= \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} \left(\frac{\partial([\hat{A}]\{\hat{q}\})}{\partial \xi} + \frac{\partial([\hat{B}]\{\hat{q}\})}{\partial \eta} + \frac{\partial([\hat{C}]\{\hat{q}\})}{\partial \zeta} \right) p_{lmn}\, d\xi\, d\eta\, d\zeta\end{aligned} \tag{17.87}$$

where $\{\hat{q}\} = |[J]|\{q\}$. For the derivative in the ξ direction alone,

$$\int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} \frac{\partial([\hat{A}]\{\hat{q}\})}{\partial\xi} p_{lmn}\, d\xi\, d\eta\, d\zeta = \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} \frac{\partial([\hat{A}]\{\hat{q}\})}{\partial\xi} \frac{g_l^{(I)}(\xi) g_m^{(J)}(\eta) g_n^{(K)}(\zeta)}{w_l^{(I)} w_m^{(J)} w_n^{(K)}} d\xi\, d\eta\, d\zeta$$

$$= \left([\hat{A}]\,|[J]|\,\{q\}\right)_{lmn} \frac{g_l^{(I)}(\xi)}{w_l^{(I)}}\bigg|_{-1}^{1} - \sum_{i=0}^{I} \frac{w_i^{(I)}}{w_l^{(I)}} \left([\hat{A}]\,|[J]|\,\{q\} \frac{\partial g_l(\xi)}{\partial\xi}\right)_{imn}$$

$$= \left([\hat{A}]\,|[J]|\,\{q\}\right)^{*}_{lmn} \left(\frac{\delta_{l,I} - \delta_{l,0}}{w_l^{(I)}}\right) - \sum_{i=0}^{I} \frac{w_i^{(I)}}{w_l^{(I)}} D_{il}^{(I)} \left([\hat{A}]\,|[J]|\,\{q\}\right)_{imn}$$

$$= \sum_{i=0}^{I} D_{li}^{(I)} \left([\hat{A}]\,|[J]|\,\{q\}\right)_{imn} + \left(\frac{\delta_{l,I} - \delta_{l,0}}{w_l^{(I)}}\right) \left(\left([\hat{A}]\,|[J]|\,\{q\}\right)^{*} - [\hat{A}]\,|[J]|\,\{q\} \right)_{lmn} \tag{17.88}$$

where * denotes the updated flux obtained by the interface conditions, and

$$D_{il}^{(I)} = \frac{\partial g_l(\xi_i)}{\partial\xi} = \begin{cases} \dfrac{L_l(x_i)}{L_l(x_l)(x_i - x_l)} & i \neq l \\ -I(I+1)/4 & i = l = 0 \\ I(I+1)/4 & i = l = I \\ 0 & \text{otherwise} \end{cases} \tag{17.89}$$

represents the derivative matrix for GLL nodes.

With similar manipulations for the η and ζ directions, the weak form of Maxwell's equations in (17.78) can be finally written into the following form:

$$\frac{\partial q_{lmn}}{\partial t} + \frac{1}{|[J]|_{lmn}} \left(\begin{aligned} &\sum_{i=0}^{I} D_{li}^{(I)} \left([\hat{A}]\,|[J]|\{q\}\right)_{imn} + \sum_{j=0}^{J} D_{mj}^{(J)} \left([\hat{B}]\,|[J]|\{q\}\right)_{ljn} \\ &+ \sum_{k=0}^{K} D_{nk}^{(K)} \left([\hat{C}]\,|[J]|\{q\}\right)_{lmk} \end{aligned} \right)$$

$$+ \left(\frac{\delta_{l,I} - \delta_{l,0}}{w_l^{(I)}}\right) \left(\left([\hat{A}]\{q\}\right)^{*} - [\hat{A}]\{q\} \right)_{lmn} + \left(\frac{\delta_{m,J} - \delta_{m,0}}{w_J^{(J)}}\right) \left(\left([\hat{B}]\{q\}\right)^{*} - [\hat{B}]\{q\} \right)_{lmn}$$

$$+ \left(\frac{\delta_{n,K} - \delta_{n,0}}{w_K^{(K)}}\right) \left(\left([\hat{C}]\{q\}\right)^{*} - [\hat{C}]\{q\} \right)_{lmn} = -\left([D]\{q\}\right)_{lmn} \tag{17.90}$$

Note that in the left-hand side of (17.90), the second term represents spectral collocation for the evaluation of spatial derivatives that have been introduced in the Chebyshev PSTD algorithm. The remaining three terms denote the correction of the interface fluxes at the six surfaces of the unit cube in the ξ, η, and ζ directions, respectively. These also can be viewed as penalty terms imposed on the six surfaces. In other words, the discontinuous Galerkin PSTD technique specified in (17.90) is rigorously equivalent to the PSTD method with a penalty term having the fixed coefficient

$$\tau = 1/w_N^{(N)} = N(N+1)/2 \tag{17.91}$$

if one assumes $I = J = K = N$.

17.7.4 Flux on the Boundary

The remaining task in (17.90) is to evaluate the interface fluxes in the ξ, η, and ζ directions, respectively. Note that

$$[\hat{A}] = \frac{\left|\frac{\partial(y,z)}{\partial(\eta,\zeta)}\right|[A] + \left|\frac{\partial(z,x)}{\partial(\eta,\zeta)}\right|[B] + \left|\frac{\partial(x,y)}{\partial(\eta,\zeta)}\right|[C]}{|[J]|} = \xi_x[A] + \xi_y[B] + \xi_z[C] \tag{17.92}$$

Thus, it is natural to rewrite the flux in the physical domain as

$$[\hat{A}]\{q\} = \begin{Bmatrix} \frac{1}{\mu}\xi\times\boldsymbol{E} \\ -\frac{1}{\varepsilon}\xi\times\boldsymbol{H} \end{Bmatrix} = \sqrt{(\xi_x)^2+(\xi_y)^2+(\xi_z)^2}\begin{Bmatrix} \frac{1}{\mu}\hat{\xi}\times\boldsymbol{E} \\ -\frac{1}{\varepsilon}\hat{\xi}\times\boldsymbol{H} \end{Bmatrix} \tag{17.93}$$

Therefore, for the field vectors at the interface of two adjacent subdomains with a normal direction $\hat{\xi}$ pointing from subdomain 1 to subdomain 2, the resolved flux for a dielectric interface or homogeneous medium can be determined by (17.73) [34]. Once the flux at the interface is updated, the discontinuous Galerkin PSTD system (17.90) is ready to be solved by a time-integration scheme. In the results discussed next, we employ a five-stage, fourth-order Runge-Kutta method.

17.7.5 Numerical Results for DG-PSTD Method

Similar to the penalty method, the discontinuous Galerkin PSTD technique is considered to be more robust than the regular PSTD algorithm because of its independence of the filtering procedure. As discussed in earlier sections, filtering degrades the accuracy of the high-frequency components, especially when the PSTD sampling density is close to its limit of π points per wavelength.

To explore the attributes of the new discontinuous Galerkin PSTD technique, we now consider the application of this and other PSTD algorithms to example problems involving electromagnetic wave interaction in two and three dimensions. The first example involves the two-dimensional *electromagnetic bandgap* (EBG) structure shown in Fig. 17.8(a), which is comprised of circular dielectric rods of permittivity $\varepsilon_r = 9.6$. The EBG structure is assumed to extend infinitely in the x-direction, and have a finite thickness of six periods in the y-direction.

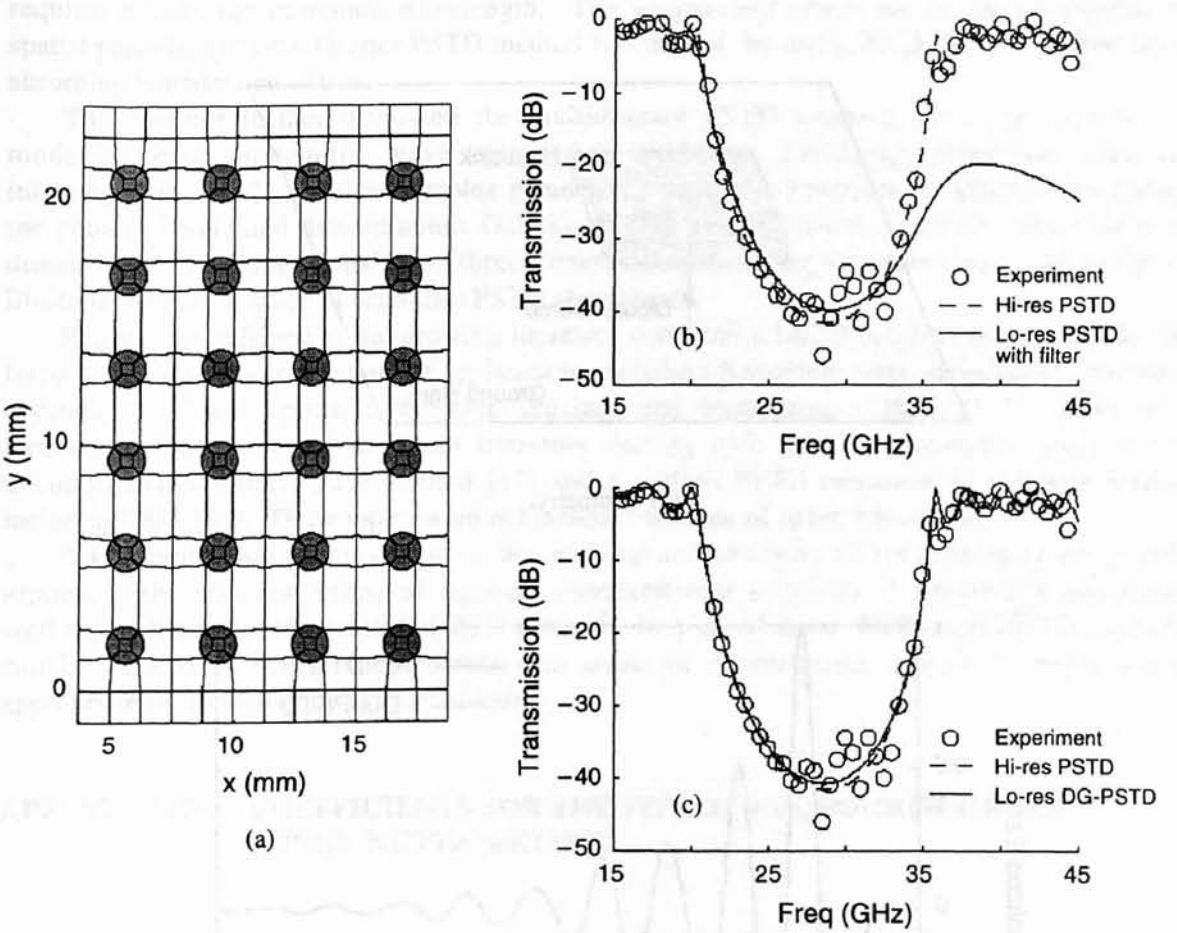
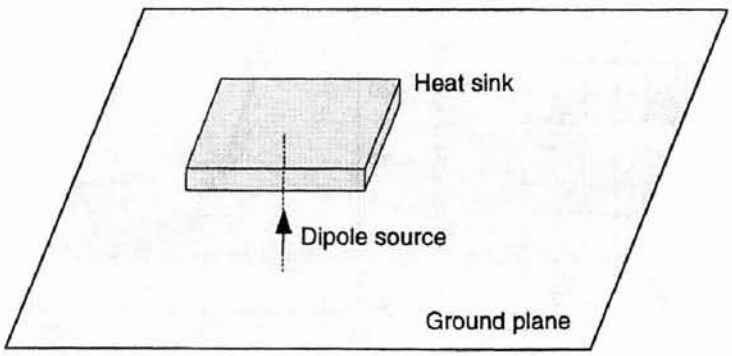


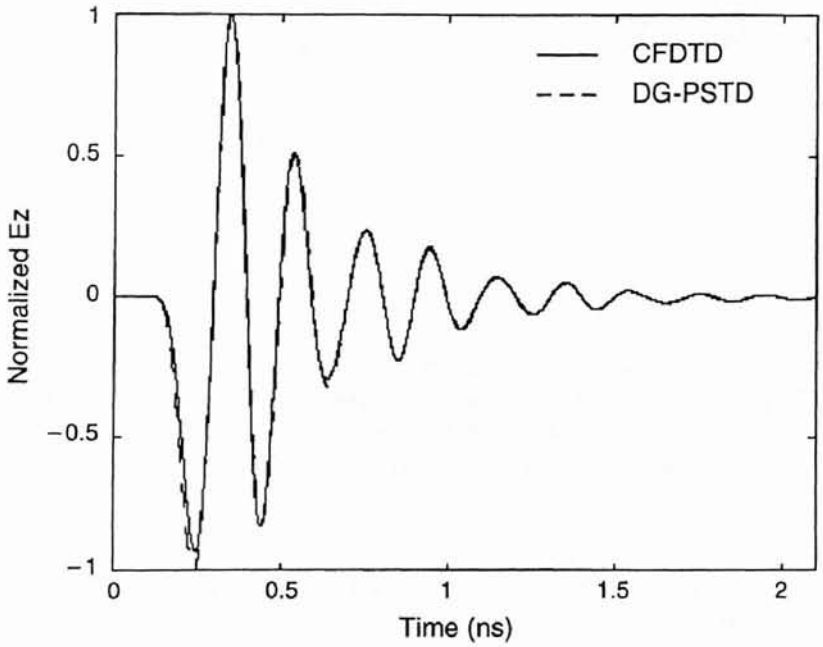
Fig. 17.8 PSTD studies of an electromagnetic bandgap structure having a stopband between approximately 20 and 35 GHz: (a) structure geometry; (b) comparison of benchmark results with a standard low-resolution (2 PPW at 40 GHz) Chebyshev PSTD code with filtering; (c) comparison of benchmark results with a low-resolution DG-PSTD code (2 PPW at 40 GHz).

Fig. 17.8(b) compares experimental results for y-directed electromagnetic wave transmission through the EBG structure [39] with two sets of simulation results: (1) benchmark PSTD data obtained using a high-resolution mesh (4 PPW at 40 GHz), and (2) test PSTD data obtained using a standard low-resolution (2 PPW at 40 GHz) Chebyshev PSTD code with filtering. Here, we see that the PSTD technique with filtering fails to maintain accuracy at high frequencies. Fig. 17.8(c) repeats this study, but now the test PSTD data are obtained using a low-resolution DG-PSTD code (2 PPW at 40 GHz). Here, it is clear that the DG-PSTD method maintains excellent accuracy despite its low spatial resolution.

The second example involves application of the three-dimensional discontinuous Galerkin PSTD method to model the electromagnetic wave interaction of an impulsive dipole source with a 4.8×4.8 cm heat sink slab, as shown in Fig. 17.9(a). The source is assumed to be located between the heat sink and an infinite ground plane, and is assumed to have a spectrum centered at 4 GHz. Fig. 17.9(b) demonstrates excellent agreement of the results of the DG-PSTD technique with results of the enlarged-cell *conformal FDTD* (CFDTD) technique [40] for a nearby radiated electric field component.



(a) Geometry.



(b) Comparison of the DG-PSTD and CFDTD results for a nearby radiated electric field component.

Fig. 17.9 Three-dimensional discontinuous Galerkin PSTD model of the electromagnetic wave interaction of an impulsive dipole source with a 4.8×4.8 cm heat sink slab.

17.8 SUMMARY AND CONCLUSIONS

This chapter has discussed in detail two different types of PSTD techniques for Maxwell's equations: the Fourier and Chebyshev PSTD methods. For a single-domain implementation, spectral accuracy can be achieved with these techniques if the internal medium is continuously inhomogeneous. Under this condition, the Fourier PSTD method can achieve spectral accuracy with a spatial resolution of only two cells per minimum wavelength, while the Chebyshev PSTD requires π cells per minimum wavelength. The wraparound effect due to the assumption of spatial periodicity in the Fourier PSTD method is removed by using the perfectly matched layer absorbing boundary condition.

This chapter further discussed the multidomain PSTD method, which is suitable for modeling electromagnetic wave interaction problems involving piecewise material inhomogeneities and curved or complex geometric structures. Recent developments concerning the penalty PSTD and discontinuous Galerkin PSTD methods were reviewed. Selected two-dimensional, 2.5-dimensional, and three-dimensional modeling examples were discussed to illustrate the performance of candidate PSTD algorithms.

Readers are referred to the growing literature concerning PSTD applications, especially for large-scale problems in a variety of applications including scattering, wave propagation, photonic crystals, nonlinear optics, subsurface sensing, and biophotonics [6, 8, 11, 15, 16, 41–46]. Readers are also referred to recent literature dealing with new developments, such as the unconditionally stable PSTD method [47] and a unified PSTD treatment of complex media, including PML [48]. These topics were not included because of space limitations.

A key conclusion of this chapter is that existing and emerging PSTD techniques are *greatly expanding* the size and scope of rigorous computational solutions of Maxwell's equations, well beyond those of the original FDTD method. In a broad sense, PSTD and FDTD methods can be viewed as being complementary in terms of characteristic electrical length scales appropriate for modeling by each technique.

APPENDIX 17A: COEFFICIENTS FOR THE FIVE-STAGE, FOURTH-ORDER RUNGE-KUTTA METHOD

a_1	0.0
a_2	−567301805773.0 / 1357537059087.0
a_3	−2404267990393.0 / 2016746695238.0
a_4	−3550918686646.0 / 2091501179385.0
a_5	1275806237668.0 / 842570457699.0
b_1	1432997174477.0 / 9575080441755.0
b_2	5161836677717.0 / 13612068292357.0
b_3	1720146321549.0 / 2090206949498.0
b_4	3134564353537.0 / 4481467310338.0
b_5	2277821191437.0 / 14882151754819.0

c_1	0.0
c_2	1432997174477.0 / 9575080441755.0
c_3	2526269341429.0 / 6820363962896.0
c_4	2006345519317.0 / 3224310063776.0
c_5	2802321613138.0 / 2924317926251.0

REFERENCES

[1] Liu, Q. H., "The PSTD algorithm: A time-domain method requiring only two cells per wavelength," *Microwave Optical Tech. Lett.*, Vol. 15, 1997, pp. 158–165.

[2] Liu, Q. H., "Large-scale simulations of electromagnetic and acoustic measurements using the pseudospectral time-domain (PSTD) algorithm," *IEEE Trans. Geoscience Remote Sensing*, Vol. 37, 1999, pp. 917–926.

[3] Liu, Q. H., "PML and PSTD algorithm for arbitrary lossy anisotropic media," *IEEE Microwave Guided Wave Lett.*, Vol. 9, 1999, pp. 48–50.

[4] Yang, B., D. Gottlieb, and J. S. Hesthaven, "Spectral simulation of electromagnetic wave scattering," *J. Comp. Phys.*, Vol. 134, 1997, pp. 216–230.

[5] Hesthaven, J. S., "Spectral collocation time-domain modeling of diffractive optical elements," *J. Comp. Phys.*, Vol. 155, 1999, pp. 287–306.

[6] Yang, B., and J. S. Hesthaven, "A pseudospectral method for time-domain computation of electromagnetic scattering by bodies of revolution," *IEEE Trans. Antennas Propagat.*, Vol. 47, 1999, pp. 132–141.

[7] Yang, B, and J. S. Hesthaven, "Multidomain pseudospectral computation of Maxwell's equations in 3D general curvilinear coordinates," *Appl. Numerical Math.*, Vol. 33, 2000, pp. 281–289.

[8] Fan, G. X., Q. H. Liu, and J. S. Hesthaven, "Multidomain pseudospectral time-domain method for simulation of scattering from buried objects," *IEEE Trans. Geoscience Remote Sensing*, Vol. 40, 2002, pp. 1366–1373.

[9] Zeng, Y., Q. H. Liu, and G. Zhao, "Multidomain pseudospectral time-domain method for acoustic waves in lossy media," *J. Comp. Acoustics*, Vol. 12, 2004, pp. 277–299.

[10] Zhao, G., and Q. H. Liu, "The 2.5D multidomain pseudospectral time-domain algorithm," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 619–627.

[11] Zhao, G., Y. Q. Zeng, and Q. H. Liu, "The 3D multidomain pseudospectral time-domain method for wideband simulation," *IEEE Microwave Wireless Comp. Lett.*, Vol. 13, 2003, pp. 184–186.

[12] Liu, Q. H., and J. Q. He, "An efficient PSTD algorithm for cylindrical coordinates," *IEEE Trans. Antennas Propagat.*, Vol. 49, 2001, pp. 1349–1351.

[13] Liu, Q. H., and N. Nguyen, "An accurate algorithm for nonuniform fast Fourier transforms (NUFFT)," *IEEE Microwave Guided Wave Lett.*, Vol. 8, 1998, pp. 18–20.

[14] Nguyen, N., and Q. H. Liu, "The regular Fourier matrices and nonuniform fast Fourier transforms," *SIAM J. Scientific Computing*, Vol. 21, 1999, pp. 283–293.

[15] Liu, Q. H., X. M. Millard, B. Tian, and Z. Q. Zhang, "Applications of nonuniform fast transform algorithms in numerical solutions of differential and integral equations," *IEEE Trans. Geoscience Remote Sensing*, Vol. 38, 2000, pp. 1551–1560.

[16] Tian, B., and Q. H. Liu, "Nonuniform fast cosine transform and Chebyshev PSTD algorithm," *Progress Electromagn. Rsrch.*, Vol. 28, 2000, pp. 259–279.

[17] Tseng, S. H., J. H. Greene, A. Taflove, D. Maitland, V. Backman, and J. T. Walsh, Jr., "Exact solution of Maxwell's equations for optical interactions with a macroscopic random medium: Addendum," *Optics Lett.*, Vol. 30, 2005, pp. 56–57.

[18] Fan, G.-X., and Q. H. Liu, "Pseudospectral time-domain algorithm applied to electromagnetic scattering from electrically large objects," *Microwave Optical Tech. Lett.*, Vol. 29, 2001, pp. 123–125.

[19] Fan, G.-X., and Q. H. Liu, "FDTD and PSTD simulations for plasma applications," *IEEE Trans. Plasma Science*, Vol. 29, 2001, pp. 341–348.

[20] Liu, Q. H., and G.-X. Fan, "A frequency-dependent PSTD algorithm for general dispersive media," *IEEE Microwave Guided Wave Lett.*, Vol. 9, 1999, pp. 51–53.

[21] Liu, Q. H., and G.-X. Fan, "Simulations of GPR in dispersive media using the PSTD algorithm," *IEEE Trans. Geoscience Remote Sensing*, Vol. 37, 1999, pp. 2317–2324.

[22] Crowley, C. W., *Mixed-Order Covariant Projection Finite Elements for Vector Fields*, Ph.D. diss., McGill University, Montreal, Quebec, 1988.

[23] Carpenter, M. H., and C. A. Kennedy, *Fourth-Order 2N-Storage Runge-Kutta Scheme*, Tech. Rpt. NASA-TM-109112, 1994.

[24] Zhao, G., and Q. H. Liu, "The 3D multidomain pseudospectral time-domain method for inhomogeneous conductive media," *IEEE Trans. Antennas Propagat.*, Vol. 52, 2004, pp. 742–749.

[25] Canuto, C., M. Y. Hussani, A. Quarteroni, and T. Zang, *Spectral Methods in Fluid Dynamics*, New York: Springer-Verlag, 1988.

[26] Majda, A., J. McDonough, and S. Osher, "The Fourier method for nonsmooth initial data," *Math. Comp.*, Vol. 30, 1978, pp. 1041–1081.

[27] Vandeven, H., "Family of spectral filters for discontinuous problems," *J. Sci. Comput.*, Vol. 6, 1991, pp. 159–192.

[28] Mulholland, L. S., and D. M. Sloan, "The effect of filtering on the pseudospectral solution of evolutionary partial differential equations," *J. Comp. Phys.*, Vol. 96, 1991, pp. 369–390.

[29] Tadmor, E., and J. Tanner, "Adaptive mollifiers: High resolution of piecewise smooth data from its spectral information," *Foundations Comp. Math.*, Vol. 2, 2002, pp. 155–189.

[30] Xu, X. M., Q. H. Liu, and Z. Q. Zhang, "The stabilized biconjugate gradient fast Fourier transform method for electromagnetic scattering," *Appl. Comp. Electromagn. Soc. J.*, Vol. 17, 2002, pp. 97–103.

[31] Zhang, Z. Q., Q. H. Liu, and X. M. Xu, "RCS computation of large inhomogeneous objects using a fast integral equation solver," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 613–618.

[32] Funaro, D., and D. Gottlieb, "A new method of imposing boundary conditions in pseudospectral approximations of hyperbolic equations," *Math. Comp.*, Vol. 51, 1988, pp. 599–613.

[33] Hesthaven, J. S., and T. Warbutron, "Nodal high-order methods on unstructured grids," *J. Comp. Phys.*, Vol. 181, 2002, pp. 186–221.

[34] Mohammadian, A., V. Shankar, and W. Hall, "Computation of electromagnetic scattering and radiation using a time-domain finite-volume discretization procedure," *Computer Phys. Commun.*, Vol. 68, 1991, pp. 175–196.

[35] Cockburn, B., G. Karniadakis, and C. W. Shu, *Discontinuous Galerkin Methods: Theory, Computation and Applications*, New York: Springer, 2000.

[36] Reed, W. H., and T. R. Hill, *Triangular Mesh Method for the Neutron Transport Equation*, Tech. Rept. LA-UR-73-479, 1973.

[37] Kopriva, D., S. L. Woodruff, and M.Y. Hussaini, "Discontinuous spectral element approximation of Maxwell's equations," in B. Cockburn, G. Karniadakis, and C. W. Shu, (eds.), *Discontinuous Galerkin Methods: Theory, Computation and Applications*, New York: Springer, 2000, pp. 355–361.

[38] Warburton, T., "Application of the discontinuous Galerkin method to Maxwell's equations using unstructured polymorphic hp-finite elements," in B. Cockburn, G. Karniadakis, and C. W. Shu, (eds.), *Discontinuous Galerkin Methods: Theory, Computation and Applications*, New York: Springer, 2000, pp. 451–458.

[39] Reineix, A., and B. Jecko, "A new photonic bandgap equivalent model using finite difference time domain method," *Annales des Telecommunications*, Vol. 51, 1996, pp. 656–662.

[40] Xiao, T., and Q. H. Liu, "Enlarged cells for the conformal FDTD method to avoid the time step reduction," *IEEE Microwave Wireless Component Lett.*, Vol. 14, 2004, pp. 551–553.

[41] Leung, W. K., and Y. Chen, "Transformed-space nonuniform pseudospectral time-domain algorithm," *Microwave Optical Tech. Lett.*, Vol. 28, 2001, pp. 391–396.

[42] Li, Q., and Y. Chen, "Applications of the PSTD for scattering analysis," *IEEE Trans. Antennas Propagat.*, Vol. 50, 2002, pp. 1317–1319.

[43] Lee, T. W., and S. C. Hagness, "Pseudospectral time-domain methods for modeling optical wave propagation in second-order nonlinear materials," *J. Optical Soc. America B*, Vol. 21, 2004, pp. 330–342.

[44] Li, Q., Y. Chen, and C. K. Li, "Hybrid PSTD-FDTD technique for scattering analysis," *Microwave Optical Tech. Lett.*, Vol. 34, 2002, pp. 19–24.

[45] Leung, W. K., and C. H. Chan, "Combining the FDTD and PSTD methods," *Microwave Optical Tech. Lett.*, Vol. 23, 1999, pp. 249–254.

[46] Tseng, S. H., J. H. Greene, A. Taflove, D. Maitland, V. Backman, and J. T. Walsh, Jr., "Exact solution of Maxwell's equations for optical interactions with a macroscopic random medium," *Optics Lett.*, Vol. 29, 2004, pp. 1393–1395.

[47] Zhao, G., and Q. H. Liu, "The unconditionally stable multidomain pseudospectral time-domain method," *IEEE Microwave Wireless Comp. Lett.*, Vol. 13, 2003, pp. 475–477.

[48] Liu, Q. H., and G. Zhao, "Review of PSTD methods for transient electromagnetics," *Intl. J. Numerical Modeling: Electronic Networks, Devices, Fields*, Vol. 22, 2004, pp. 299–323.

Chapter 18

Advances in Unconditionally Stable Techniques

Hans De Raedt

18.1 INTRODUCTION

This chapter discusses recent progress in the development of unconditionally stable FDTD algorithms to solve Maxwell's equations. We present a general, unified framework that facilitates the construction of FDTD algorithms (including the Yee algorithm [1]) with specific properties. Our approach is *constructive* and *modular*. It gives a recipe for constructing unconditionally stable algorithms that are tuned to particular problems, and that can be combined with other unconditionally stable algorithms to solve more complicated problems.

18.2 GENERAL FRAMEWORK

For simplicity and clarity of the notation, we introduce the basic concepts using Maxwell's equation in one dimension. Specifically, we consider the case of a TEM wave (see Section 3.4). We assume that $M_{\text{source}_y} = 0$, $\sigma = \sigma^* = 0$, and $\varepsilon = \mu = 1$. Equation (3.16) then reduces to

$$\frac{\partial H_y(x,t)}{\partial t} = \frac{\partial E_z(x,t)}{\partial x} \tag{18.1a}$$

$$\frac{\partial E_z(x,t)}{\partial t} = \frac{\partial H_y(x,t)}{\partial x} - J_z(x,t) \tag{18.1b}$$

Using the one-dimensional Yee lattice arrangement of field components illustrated in Fig. 18.1, we represent the spatial derivatives in (18.1) by second-order central-difference approximations, but do not yet discretize the time variable t:

$$\frac{\partial}{\partial t} H_y\Big|_{i+1/2}^{t} = \frac{1}{\Delta x}\left(E_z\Big|_{i+1}^{t} - E_z\Big|_{i}^{t}\right) \tag{18.2a}$$

$$\frac{\partial}{\partial t} E_z\Big|_{i}^{t} = \frac{1}{\Delta x}\left(H_y\Big|_{i+1/2}^{t} - H_y\Big|_{i-1/2}^{t}\right) - J_z\Big|_{i}^{t} \tag{18.2b}$$

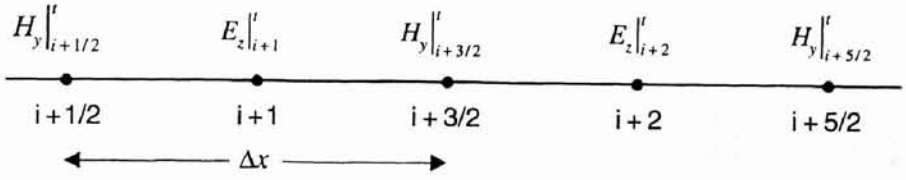


Fig. 18.1 Positions of the two field components in the one-dimensional Yee grid (see also Fig. 3.2).

It is convenient to use a compact notation for one-dimensional arrays of numbers and their associated matrix operators. We introduce the row array

$$\{\Psi(t)\}^{\mathrm{T}} = \left\{ H_y|^t_{1/2} \quad E_z|^t_1 \quad \cdots \quad H_y|^t_{i+1/2} \quad E_z|^t_i \quad \cdots \quad E_z|^t_n \quad H_y|^t_{n+1/2} \right\} \tag{18.3}$$

where the superscript T in (18.3) denotes the transpose of an array (or matrix). Accordingly, $\{\Psi(t)\}$ is a column array containing $2n+1$ field components. The Maxwell equations (18.2) can now be written as

$$\frac{\partial}{\partial t}\{\Psi(t)\} = [L]\{\Psi(t)\} - \{S(t)\} \tag{18.4}$$

where the source term in (18.4) is represented by the array

$$\{S(t)\}^{\mathrm{T}} = \left\{ 0 \quad J_z|^t_1 \quad \cdots \quad 0 \quad J_z|^t_i \quad \cdots \quad J_z|^t_n \quad 0 \right\} \tag{18.5}$$

and the matrix $[L]$ is given by

$$[L] = \frac{1}{\Delta x}\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 0 & 1 \\ 0 & \cdots & 0 & -1 & 0 \end{pmatrix} \tag{18.6}$$

The mathematical structure of (18.4) to (18.6) is much more general than this simple one-dimensional example suggests. We now discuss the general aspects of (18.4) in more detail.

18.3 MATRIX-EXPONENTIAL CONCEPTS

In general, Maxwell's equations [see (3.7) and (3.8) of Chapter 3] can be written in the form (18.4). By direct substitution, it is easy to verify that the solution of (18.4) can be written as

$$\{\Psi(t)\} = e^{t[L]}\{\Psi(0)\} - \int_0^t e^{(t-u)[L]}\{S(u)\}\,du \tag{18.7}$$

where the matrix exponential $e^{t[L]}$ is formally defined by the matrix-valued Taylor series [2–8]

$$e^{t[L]} = \sum_{k=0}^{\infty} \frac{t^k}{k!}[L]^k \tag{18.8}$$

Formal solution (18.7) is a convenient starting point to construct FDTD algorithms. Time-marching methods solve (18.7) by advancing the fields $\{\Psi(t)\}$ in small steps, using a time-step $\tau = \Delta t$. From (18.7) it follows directly that

$$\{\Psi(t+\tau)\} = e^{\tau[L]}\{\Psi(t)\} - \int_0^{\tau} e^{u[L]}\{S(t+\tau-u)\}\,du \tag{18.9}$$

From (18.9), we see that solving Maxwell's equations in the time domain amounts to specifying an algorithm that updates an arbitrary array of field components $\{\Phi(t)\}$ according to

$$\{\Phi(t+\tau)\} = e^{\tau[L]}\{\Phi(t)\} \tag{18.10}$$

The exceptional case for which the matrix exponential $e^{\tau[L]}$ can be calculated explicitly is usually of little relevance for realistic problems. Therefore, we will exclude this case from our discussion. Thus, the general strategy is to develop some approximation scheme that performs the operation (18.10). Later, we show that this strategy can be used to recover algorithms that are obtained in the conventional manner (including the Yee algorithm) by discretizing the time derivatives. The main advantage of approximating $e^{\tau[L]}$ instead of the differential equation is that we can easily construct algorithms with specific properties, such as unconditional stability.

To see why this is a fruitful strategy, we now make a number of general observations. For simplicity, we assume that there are no current sources ($\{S(t)\} = 0$) and exclude the case of lossy materials.[1]

1. Matrix $[L]$ in (18.4) satisfies $[L]^T = -[L]$; that is, $[L]$ is real and skew-symmetric. This is a fundamental, general symmetry property of the pointwise Maxwell equations. Under the conditions stated in Chapter 3, Section 3.2, the integral representation of Maxwell's equations is easily transformed into an equation of the form (18.4). In general, for the two- and three-dimensional Maxwell equations (with spatially varying permittivity and permeability but in the absence of absorption), $[L]$ is skew-symmetric.

2. If $[L]$ is a skew-symmetric matrix, then $(e^{t[L]})^{-1} = e^{-t[L]} = e^{t[L]^T} = (e^{t[L]})^T$. Hence, $e^{t[L]}$ is an orthogonal matrix; that is, it performs a rotation of a vector [9, 10]. We see that the time development in formal solution (18.7) is represented entirely in terms of rotations $e^{t[L]}$.

[1] Appendices 18A.1, 2 and Sections 18.6, 7 discuss the extensions that are necessary to deal with these cases.

3. The energy of the electromagnetic field in the Yee grid is just the inner product $\{\Psi(t)\}^T\{\Psi(t)\}$. Therefore, conservation of electromagnetic energy implies that the length (norm) $\|\{\Psi(t)\}\| = \sqrt{\{\Psi(t)\}^T\{\Psi(t)\}}$ of $\{\Psi(t)\}$ does not change with time. Of course, the time evolution $e^{t[L]}$ exactly conserves the energy (in the absence of external sources) because it only rotates $\{\Psi(t)\}$, and rotation does not change its length.

4. Any approximation to $e^{t[L]}$ that can be written as a product of rotations *exactly* conserves the electromagnetic energy.

5. The general criterion for numerical stability given in Chapter 2, Section 2.7, is formally expressed by $\|\{\Psi(t)\}\| \leq C\|\{\Psi(0)\}\|$ for all t and C being a positive constant [2].

6. If a time-integration algorithm can be written as a product of rotations, then according to observation 3, we have $\|\{\Psi(t)\}\| = \|\{\Psi(0)\}\|$. By observation 4, this algorithm is unconditionally stable by construction. Thus, an algorithm that conserves energy *exactly* is unconditionally stable.

A systematic approach to construct approximations to the matrix exponential $e^{\tau[L]}$ is to invoke the Lie-Trotter-Suzuki formula [6–8]. The basic idea is to split $[L]$ into several parts $[L] = [L_1] + [L_2] + \ldots + [L_p]$, and to approximate $e^{\tau[L]}$ by some product of the matrix exponentials $e^{\tau[L_1]}, e^{\tau[L_2]}, \ldots, e^{\tau[L_p]}$. Thus, we apply the idea of *operator splitting* to the formal solution of the differential equation, rather than to the differential equation itself. Instead of discussing this approach in great detail, we confine ourselves to approximations that are useful for FDTD applications. As a matter of fact, we only need the approximation

$$e^{\tau[L]} \approx e^{\tau[L_1]/2} e^{\tau[L_2]} e^{\tau[L_1/2]} \tag{18.11}$$

Appendix 18A.2 gives a rigorous proof that the error resulting from approximating $e^{\tau[L]}$ by $e^{\tau[L_1/2]} e^{\tau[L_2]} e^{\tau[L_1/2]}$ is less than $C\tau^3$ for any time-step τ, where C is a positive constant.

Without going into the details of a particular problem, there are a number of important consequences of employing (18.11). First, for practical purposes, it is necessary to decompose $[L]$ such that we can easily compute $e^{\tau[L_1]}\{\Phi\}$ and $e^{\tau[L_2]}\{\Phi\}$ for any time step τ and arbitrary $\{\Phi\}$. We have not yet specified how to choose $[L_1]$ or $[L_2]$, except that these should satisfy $[L] = [L_1] + [L_2]$. Therefore, we have much freedom to construct algorithms with specific properties. For example, if $[L_1]$ and $[L_2]$ are skew-symmetric, then by observation 2, $e^{\tau[L_1/2]} e^{\tau[L_2]} e^{\tau[L_1/2]}$ is the product of three orthogonal matrices. Hence, by observation 4, the energy of the fields is conserved, and by observation 6, this choice defines an unconditionally stable algorithm. Second, if necessary, we can split up any of the matrix exponentials in $e^{\tau[L_1/2]} e^{\tau[L_2]} e^{\tau[L_1/2]}$ by reapplying (18.11). The error always remains less than $C\tau^3$, where the actual value of C depends on the details of the decomposition of $[L]$. Finally, if necessary, the accuracy of approximation (18.11) can be increased to $O(\tau^4)$ by simply reusing (18.11); that is, with a minimum of programming effort. A particularly useful fourth-order product formula, also illustrating the modular structure of the approach, is given by [8, 11, 12]

$$e^{\tau[L]} \approx \left(e^{a\tau[L_1]/2}e^{a\tau[L_2]}e^{a\tau[L_1]/2}\right)^2\left(e^{b\tau[L_1]/2}e^{b\tau[L_2]}e^{b\tau[L_1]/2}\right)\left(e^{a\tau[L_1]/2}e^{a\tau[L_2]}e^{a\tau[L_1]/2}\right)^2 \tag{18.12}$$

where $a = 1/(4 - 4^{1/3})$ and $b = 1 - 4a$. The use of (18.12) does not require more storage than its component algorithms.

As an alternative to the product-formula approach, we may also consider computing $e^{\tau[L]}\{\Phi\}$ using a polynomial or rational approximation to $e^{\tau[L]}$. We shall later discuss two cases: the "one-step" algorithm [13, 14] based on an extremely accurate Chebyshev polynomial representation of $e^{\tau[L]}$ [15], and an *alternating-direction implicit* (ADI) algorithm [16] that employs a rational approximation to compute $e^{\tau[L_1]}\{\Phi\}$ and $e^{\tau[L_2]}\{\Phi\}$.

18.4 PRODUCT-FORMULA APPROACH

18.4.1 The Classic Yee Algorithm as a Particular Realization

We first show how the classic Yee algorithm can be viewed as a particular realization of the product-formula approach discussed above [17]. While this discussion uses the example of a TEM wave in a one-dimensional space to illustrate the basic ideas, extension to two- and three-dimensional models does not require new concepts.

In the Yee algorithm, one time-step involves an update of $\boldsymbol{E}$ using the current values of $\boldsymbol{H}$, followed by an update of $\boldsymbol{H}$ using the new values of $\boldsymbol{E}$ (see Chapter 3, Section 3.6.1). Deriving the Yee algorithm via the product-formula approach is straightforward if we rearrange the elements of $\{\Psi(t)\}$. We define a new array of field components by

$$\{Y(t)\}^{\mathrm{T}} = \left\{ H_y\Big|^t_{1/2} \quad \cdots \quad H_y\Big|^t_{n+1/2} \quad E_z\Big|^t_1 \quad \cdots \quad E_z\Big|^t_n \right\} \tag{18.13}$$

This permutation of elements also changes $[L]$ as follows:

$$[L] = \frac{1}{\Delta x}\left(\begin{array}{cccc:ccc} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ \hdashline -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{array}\right) \tag{18.14}$$

where we have taken as an example a one-dimensional Yee lattice of seven gridpoints. From (18.14), we see that $[L]$ remains skew-symmetric, as it should because a permutation of the array elements does not change the symmetry properties of a matrix. The structure of $[L]$ of (18.14) suggests that it is convenient to use a block-matrix notation. Following our general strategy, we can now recover the Yee algorithm if we choose $[L] = [L_1] + [L_2]$, where

$$[L] = \begin{pmatrix} 0 & -[A]^{\mathrm{T}} \\ [A] & 0 \end{pmatrix}, \quad [L_1] = \begin{pmatrix} 0 & -[A]^{\mathrm{T}} \\ 0 & 0 \end{pmatrix}, \quad [L_2] = \begin{pmatrix} 0 & 0 \\ [A] & 0 \end{pmatrix} \tag{18.15}$$

and

$$[A] = \frac{1}{\Delta x}\begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix} \tag{18.16}$$

The decomposition of (18.15) is general; that is, it also works for two- and three-dimensional problems. However, the particular form of $[A]$ given by (18.16) is directly linked to the specific one-dimensional example considered here.

According to our general strategy, the decomposition should be such that we can readily compute $e^{\tau[L_1]}\{\Phi\}$ and $e^{\tau[L_2]}\{\Phi\}$ for arbitrary τ and $\{\Phi\}$. For (18.15), this is straightforward; it is clear that $[L_1][L_1] = [L_2][L_2] = 0$. From Taylor series (18.8), it follows that $e^{\tau[L_1]}\{Y(t)\} = (1 + \tau[L_1])\{Y(t)\}$, which is *exactly* the first part of the Yee algorithm, the $\boldsymbol{H}$ update. Similarly, the second part of the Yee algorithm, the $\boldsymbol{E}$ update, is given by $e^{\tau[L_2]}\{Y(t)\} = (1 + \tau[L_2])\{Y(t)\}$. From the general criterion for numerical stability, it follows that the Yee algorithm is stable if $\|e^{\tau[L_1]}e^{\tau[L_2]}\| = \|1 + \tau[L] - \tau^2[L_1][L_1]^{\mathrm{T}}\| \le 1$. Appendix 18B shows how this condition leads to the well-known result (see Chapter 4, Section 4.7) that the Yee algorithm is numerically stable if $\tau \le \Delta x/\sqrt{d}$, where d denotes the spatial dimension of the system.

We now show that a slightly different approach yields a Yee-like algorithm that uses a collocated time arrangement of $\boldsymbol{E}$ and $\boldsymbol{H}$ [17]. According to our general formalism, we can write

$$\begin{aligned} \{Y(t+2\tau)\} &= (e^{\tau[L_1]/2}e^{\tau[L_2]}e^{\tau[L_1]/2})(e^{\tau[L_1]/2}e^{\tau[L_2]}e^{\tau[L_1]/2})\{Y(t)\} \\ &= e^{-\tau[L_1]/2}(e^{\tau[L_1]}e^{\tau[L_2]})(e^{\tau[L_1]}e^{\tau[L_2]})(e^{\tau[L_1]/2}\{Y(t)\}) \end{aligned} \tag{18.17}$$

In the second line of (18.17), operations are grouped so that the relation to the Yee algorithm is more explicit. Reading from right to the left, the first operation propagates $\boldsymbol{H}$ by $\tau/2$. Then, the first pair of matrix exponentials performs a Yee update of $\boldsymbol{E}$ with a time-step τ, followed by a Yee update of $\boldsymbol{H}$ with a time-step τ. The second pair of matrix exponentials repeats the previous two operations. The final operation restores the $\tau/2$ shift generated by the first matrix exponential by propagating $\boldsymbol{H}$ over $-\tau/2$. In the original Yee algorithm, the last step is missing because, throughout the procedure, $\boldsymbol{H}$ and $\boldsymbol{E}$ are defined for times that differ by $\tau/2$. Overall, (18.17) demonstrates that we can remove the staggered-in-time feature of the conventional Yee algorithm at virtually no extra computational cost, while retaining its other properties.

18.4.2 The ADI Method as a Second Realization

ADI algorithms (see Chapter 4, Section 4.10) are usually derived by applying the idea of operator splitting to the difference equation [16]. In fact, and importantly, we can derive the ADI algorithm starting from formal update rule (18.10).

Obtaining the ADI method from the product-formula approach is instructive for two reasons. First, this derivation shows that ADI algorithms are *approximations* to the second-order product formula relation (18.11). Second, it suggests how to derive ADI algorithms that are unconditionally stable by construction. As an example, we consider the splitting $[L] = [L_1] + [L_2]$ into skew-symmetric matrices $[L_1]$ and $[L_2]$. The second-order product formula reads

$$\begin{aligned} e^{\tau([L_1]+[L_2])} &\approx e^{\tau[L_1]/2} e^{\tau[L_2]} e^{\tau[L_1]/2} = e^{\tau[L_1]/2} e^{\tau[L_2]/2} e^{\tau[L_2]/2} e^{\tau[L_1]/2} \\ &\approx (e^{-\tau[L_2]/2} e^{-\tau[L_1]/2})^{-1} e^{\tau[L_2]/2} e^{\tau[L_1]/2} \end{aligned} \tag{18.18}$$

We now replace each of the matrix exponentials in the right-hand side by its *first*-order accurate Taylor expansion, and obtain a standard ADI scheme [16]

$$\begin{aligned} &(e^{-\tau[L_2]/2} e^{-\tau[L_1]/2})^{-1} e^{\tau[L_2]/2} e^{\tau[L_1]/2} \approx \\ &\qquad (1-\tau[L_1]/2)^{-1}(1+\tau[L_2]/2)(1-\tau[L_2]/2)^{-1}(1+\tau[L_1]/2) \end{aligned} \tag{18.19}$$

where we used the fact that we may interchange the order of the two middle factors in the right-hand side of (18.19). The implicit character of ADI algorithms is reflected by the presence of factors such as $(1 + \tau[L_2]/2)(1 - \tau[L_2]/2)^{-1}$, which is the (1,1) Padé approximation to $e^{\tau[L_2]}$ [5]. Appendix 18C proves that time-stepping based on the ADI algorithm is unconditionally stable if $[L_1]$ and $[L_2]$ are skew-symmetric. This proof is less complicated than the proof of the conditional stability of the Yee algorithm. A discussion of applications of ADI methods to FDTD problems can be found in [18 – 21], and in Chapter 4, Section 4.10.

18.4.3 Unconditionally Stable Algorithms: Real-Space Approach

According to our general theory, all we have to do to construct an unconditionally stable algorithm is to apply the product-formula recipe with a decomposition of $[L]$ into skew-symmetric parts. These parts are chosen such that it easy to compute each matrix exponential that appears in the product. To illustrate the basic idea, we use a simple one-dimensional system with seven gridpoints and decompose $[L] = [L_1] + [L_2]$ as follows [11]:

$$[L_1] = \frac{1}{\Delta x} \left(\begin{array}{ccc:cc:cc:c} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{18.20a}$$

$$[L_2] = \frac{1}{\Delta x}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix} \tag{18.20b}$$

It is clear that both $[L_1]$ and $[L_2]$ are skew-symmetric, and display a block-matrix structure wherein each block matrix has at most a 2×2 dimension. Since the matrix exponential of a block diagonal matrix is also block diagonal (with the same structure as the matrix itself), finding the explicit form of the matrix exponentials of $[L_1]$ and $[L_2]$ requires, at most, the calculation of a matrix exponential of a 2×2 matrix. For this example, the expressions are

$$e^{\tau[L_1]} = \begin{pmatrix} c & s & 0 & 0 & 0 & 0 & 0 \\ -s & c & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c & s & 0 & 0 & 0 \\ 0 & 0 & -s & c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & c & s & 0 \\ 0 & 0 & 0 & 0 & -s & s & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{18.21a}$$

$$e^{\tau[L_2]} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & c & s & 0 & 0 & 0 & 0 \\ 0 & -s & c & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c & s & 0 & 0 \\ 0 & 0 & 0 & -s & c & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & c & s \\ 0 & 0 & 0 & 0 & 0 & -s & c \end{pmatrix} \tag{18.21b}$$

where $c = \cos(\tau/\Delta x)$ and $s = \sin(\tau/\Delta x)$. This completes the construction of an unconditionally stable algorithm that solves the one-dimensional Maxwell's equations.

In practice, implementing the algorithm defined by the decomposition (18.21) is about as simple as writing code to implement the classic Yee algorithm. From (18.21), it is clear that all we need to program is the multiplication of an array of two elements by a 2×2 matrix. This operation needs to be repeated in a loop with stride 2. Table 18.1 lists Fortran code that implements both the real-space algorithm based on (18.21) and the conventional Yee algorithm for the example of Section 18.4.1.

TABLE 18.1

Comparative Fortran Code for the Yee Algorithm and the Real-Space Unconditionally Stable Algorithm Based on (18.21)

Conventional Yee Algorithm	*Real-Space Unconditionally Stable Algorithm*
`r = tau / deltaX`	`r = tau / deltaX`
`if(istart==1) then c=0 else c=psi(1)`	`c = cos(r)`
`do i = istart, L-1, 2`	`s = sin(r)`
`s = psi(i+1)`	`do i = istart, L-1, 2`
`psi(i) = r*(s-c) + psi(i)`	`r = psi(i)`
`c = s`	`psi(i) = c*r + s*psi(i+1)`
`end do`	`psi(i+1) = c*psi(i+1) - s*r`
`if(istart /= 1) psi(L) = psi(L) - r*c`	`end do`

The left side of Table 18.1 lists Fortran code that implements the classic Yee algorithm using the decomposition defined by (18.15) and (18.16). The right side of Table 18.1 lists Fortran code that implements the real-space unconditionally stable algorithm based on the decomposition (18.21). In the latter case, $e^{\tau[L_1]}$ is implemented for `istart = 1` and $e^{\tau[L_2]}$ is implemented for `istart = 2`. Note that, for both algorithms, the arrangement of the fields is given by (18.3) and not by (18.13).

Counting floating-point arithmetic and memory-access operations, we see from Table 18.1 that six operations per iteration are required for the Yee algorithm, and nine operations per iteration are required for the unconditionally stable algorithm. Thus, in terms of operations, the extra cost of using the unconditionally stable algorithm based on (18.21) is not excessive. Furthermore, both algorithms vectorize and parallelize equally well.

The extension of this real-space approach to two- and three-dimensional systems with spatially varying permeability and permittivity and/or using fourth-order-accurate spatial differences is technically more complicated, but does not require new concepts [11, 12]. This class of algorithms is explicit and unconditionally stable by construction.

18.4.4 Unconditionally Stable Algorithms: Fourier-Space Approach

As discussed in Chapter 4, Section 4.9.4, and Chapter 17, evaluation of the spatial derivatives in Maxwell's equations via Fourier transformation [i.e., the *pseudospectral time-domain* (PSTD) technique] greatly reduces numerical dispersion. Here, we show how this idea can be incorporated into unconditionally stable algorithms. The approach we describe may be useful in combination with the decompositions to be discussed in Sections 18.6 and 18.7, in which $[L]$ is split into a free-space propagator and a propagator that accounts for the material properties. The latter is most conveniently treated in a real-space representation. Thus, we only have to find a procedure to perform the free-space time evolution efficiently, with minimal numerical dispersion. As in the Fourier PSTD approach, we work with regular, rectangular spatial lattices.

In three dimensions, Fourier transformation of Maxwell's equations with respect to the spatial coordinates yields

$$\frac{\partial}{\partial t}\begin{pmatrix}\{\boldsymbol{B}\}\\ \{\boldsymbol{E}\}\end{pmatrix} \equiv [L_1]\begin{pmatrix}\{\boldsymbol{B}\}\\ \{\boldsymbol{E}\}\end{pmatrix} = j\begin{pmatrix}0 & [K]\\ [K]^{\mathrm{T}} & 0\end{pmatrix}\begin{pmatrix}\{\boldsymbol{B}\}\\ \{\boldsymbol{E}\}\end{pmatrix} \tag{18.22a}$$

where

$$[K] = \begin{pmatrix}0 & -k_z & k_y\\ k_z & 0 & -k_x\\ -k_y & k_x & 0\end{pmatrix} \tag{18.22b}$$

Straightforward algebra shows that $[L_1]^3 = -(k_x^2 + k_y^2 + k_z^2)[L_1]$. This implies that $[U(t)] = e^{t[L_1]}$ satisfies the differential equation $d^3[U(t)]/dt^3 = -(k_x^2 + k_y^2 + k_z^2)\, d[U(t)]/dt$. Its solution is

$$[U(t)] = [I] + \frac{\sin t\sqrt{k_x^2 + k_y^2 + k_z^2}}{\sqrt{k_x^2 + k_y^2 + k_z^2}}[L_1] + \frac{1 - \cos t\sqrt{k_x^2 + k_y^2 + k_z^2}}{k_x^2 + k_y^2 + k_z^2}[L_1]^2 \tag{18.23}$$

where $[I]$ denotes the unit matrix. Thus, the PSTD unconditionally stable algorithm for the free-space propagator reads

$$\begin{bmatrix}\{\boldsymbol{B}(\boldsymbol{r}, t+\tau)\}\\ \{\boldsymbol{E}(\boldsymbol{r}, t+\tau)\}\end{bmatrix} = \mathcal{F}^{-1}\left\{[U(\tau)]\ \mathcal{F}\begin{bmatrix}\{\boldsymbol{B}(\boldsymbol{r}, t)\}\\ \{\boldsymbol{E}(\boldsymbol{r}, t)\}\end{bmatrix}\right\} \tag{18.24}$$

which is exact for any time-step τ. In practice, we implement (18.24) as follows. First, we Fourier transform the fields, labeling the fields with wavevector $\boldsymbol{k}$ instead of position $\boldsymbol{r}$. Then, for each value of $\boldsymbol{k}$, we multiply the six field components with the 6×6 matrix $[U(\tau)]$. Finally, we apply the inverse Fourier transform to bring the fields back to real space.

The accuracy, unconditional stability, numerical dispersion, and efficiency of the algorithm based on (18.23) and (18.24) do not depend on τ. Therefore, modeling free-space propagation with this algorithm is extremely efficient. Of course, in realistic FDTD applications, some material is present in the simulation space, and it is necessary to use the splitting $[L] = [L_1] + [L_2]$. Then, the efficiency (but not the unconditional stability or the numerical dispersion) depends on the specific material properties, grid size and shape, boundary conditions, and so forth.

18.5 CHEBYSHEV POLYNOMIAL ALGORITHM

This section describes an algorithm that uses Chebyshev polynomials to compute $e^{t[L]}\{\Phi\}$ to *machine accuracy* for arbitrary large times t [15]. We call this a "one-step algorithm," because the solution at time t is obtained directly without making use of time-stepping, while the calculation of $e^{t[L]}\{\Phi\}$ cannot be used to extract information about $e^{t'[L]}\{\Phi\}$ for $t' < t$ [13, 14]. Because of its extreme accuracy, the one-step algorithm can be used for time-stepping (with very large t) without running into numerical instabilities [13, 14, 17].

It is instructive to start from Lagrange's form of the Taylor series:

$$e^{t[L]} = \sum_{k=0}^{K-1}\left(\frac{t^k}{k!}[L]^k\right) + \frac{t^K[L]^K e^{\xi[L]}}{K!}, \qquad 0 \le \xi \le t \tag{18.25}$$

From (18.25), we can estimate the error of truncating the Taylor series after K terms. For a skew-symmetric matrix $[L]$, the error is smaller than some positive number ε if $\|t[L]\| \le (\varepsilon K!)^{1/K}$. For sufficiently large K, we can use Stirling's formula [22] to find that the error vanishes faster than $(e\|t[L]\|/K)^K$. Thus, if $e\|t[L]\| \le K$, the error incurred by retaining the first K terms in the Taylor series vanishes exponentially fast with K. Put differently, for a given (large) time t, we can calculate $e^{t[L]}\{\Phi\}$ to machine precision using $O(K)$ matrix-vector operations $\{\Phi\} \leftarrow [L]\{\Phi\}$ if K is sufficiently large. This would be the most efficient procedure to compute $e^{t[L]}\{\Phi\}$ if summing of the first K terms is a numerically stable procedure, which it is not [4]. In all cases of interest, some of the K terms in (18.25) are so large that it renders the Taylor series useless. However, this numerical instability can be easily removed by economization of the series (18.25) in terms of matrix-valued modified Chebyshev polynomials [22].

We define the matrix $[\hat{L}] = [L]/\|[L]\|$. This ensures that $\|[\hat{L}]\{\Phi\}\| \le \|\{\Phi\}\|$ for all $\{\Phi\}$. We also define new matrices $[\hat{T}_k]$ by the recursion relation

$$[\hat{T}_{k+1}]\{\Phi\} = 2[\hat{L}][\hat{T}_k]\{\Phi\} + [\hat{T}_{k-1}]\{\Phi\}, \quad k \ge 1 \tag{18.26a}$$

where

$$[\hat{T}_0]\{\Phi\} = \{\Phi\}, \qquad [\hat{T}_1]\{\Phi\} = [\hat{L}]\{\Phi\} \tag{18.26b}$$

and express the powers of $[\hat{L}]$ in terms of $[\hat{T}_k]$. After some algebra, or alternatively, using the expansion $e^{jz\cos\theta} = J_0(z) + 2\sum_k j^k J_k(z)\cos(k\theta)$ [22], we find

$$e^{t[L]}\{\Phi\} = J_0\big(t\|[L]\|\big)\{\Phi\} + 2\sum_{k=1}^{\infty} J_k\big(t\|[L]\|\big)[\hat{T}_k]\{\Phi\} \tag{18.27}$$

where $J_k(x)$ is the Bessel function of the first kind of order k. Since $|J_k(x)| \le 1$, each term in (18.27) will be of order one if $\|[\hat{T}_k]\| \le 1$ for all k. The latter is equivalent to the statement that matrix recursion (18.26) is numerically stable. The solutions are given by $[\hat{T}_k]\{\Phi\} = [R]^k\{\Phi\}$, where $[R] = [\hat{L}] \pm (1 + [\hat{L}]^2)^{1/2}$. If $[L]$ is skew-symmetric, the eigenvalues of $[\hat{L}]$ are pure imaginary, and their modulus does not exceed one. Hence, $\|[R]\| \le 1$ and $\|[\hat{T}_k]\| \le 1$. Thus, each term in (18.27) can be calculated in a numerically stable manner.

Retaining the first K terms of series (18.27), we obtain

$$e^{t[L]}\{\Phi\} = J_0\big(t\|[L]\|\big)\{\Phi\} + 2\sum_{k=1}^{K-1} J_k\big(t\|[L]\|\big)[\hat{T}_k]\{\Phi\} \tag{18.28}$$

Using the inequality $|J_k(x)| \le x^k/2^k k!$, we find that the error incurred by retaining the first K terms of series (18.27) vanishes exponentially fast with K if $e\|t[L]\| \le 2K$, and K is large.

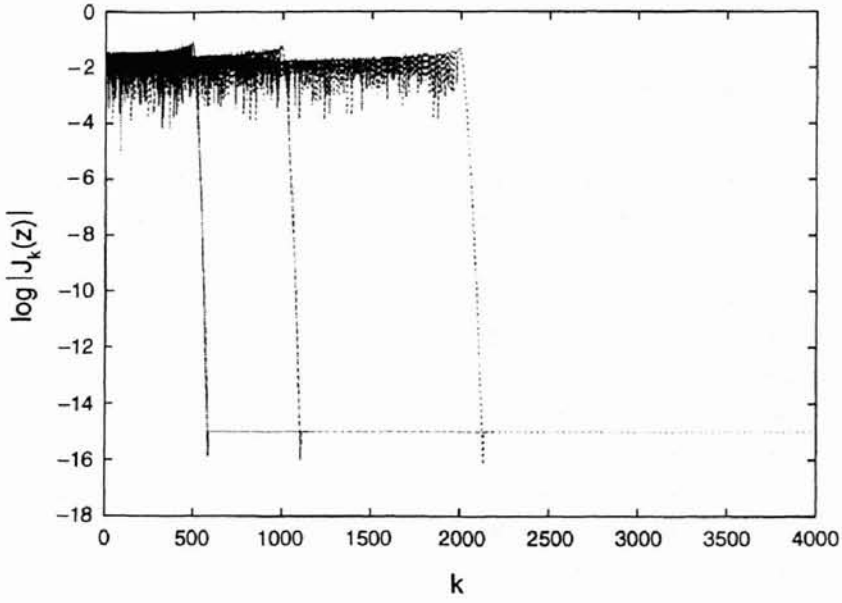


Fig. 18.2 Bessel function $J_k(z)$ as a function of order k. Solid line, $z = 500$; dashed line, $z = 1{,}000$; dotted line, $z = 2{,}000$.

Fig. 18.2 provides numerical results using the downward recursion relation of [16] to calculate Bessel functions. These results confirm that the first K Bessel functions can be calculated to machine precision in $O(K)$ arithmetic operations. We see that, if k is slightly larger than z, the coefficients in expansion (18.28) are zero within machine precision. In practice, calculating the first 20,000 Bessel functions takes less than 1 sec on a Pentium III 600-MHz mobile processor using 14 to 15 digit arithmetic. Hence, this part of the calculation is a negligible fraction of the total computational work for solving Maxwell's equations.

The previous analysis can be repeated to find the Chebyshev polynomial expansion of the source term $\int_0^t f(u)e^{(t-u)[L]}S_0\,du$. For sinusoidal and Gaussian pulsed sources, an efficient computational procedure for the expansion coefficients is described in [13, 14].

Implementing the one-step algorithm is straightforward if we already have a subroutine to perform $\{\Phi'\} = [L]\{\Phi\}$. A few minor modifications (mainly deleting fragments) of an existing Yee code are sufficient to write such a subroutine. For a realistic problem, the calculation of $\|[L]\|$ is very expensive. Therefore, we use the rigorous upper bound $\|[L]\| \le \|[L]\|_1 = \max_j \sum_i |L_{ij}|$ [9, 10], and replace $\|[L]\|$ by $\|[L]\|_1$ in (18.26) and (18.28). Table 18.2 lists the code that implements the one-step algorithm. Here, we assume that a subroutine [Lpsi(X,Y)] to perform $\{\Phi\} \leftarrow [L]\{\Phi\}$ is available, and that the K Bessel functions have been stored in the array cheb(.). This code is not well-optimized.

In summary, since all the terms in (18.28) are of the same order of magnitude, we can calculate $e^{t[L]}\{\Phi\}$ to machine precision with only K matrix-vector operations $\{\Phi\} \leftarrow [L]\{\Phi\}$. This operation can be performed in real space using the Yee grid, or in Fourier space using PSTD. The main limitation of the one-step algorithm is that it can only be applied to problems for which $[L]$ is skew-symmetric [13, 14, 17].

TABLE 18.2

Fortran 90 Code that Implements the One-Step Algorithm

```
real(kind=8)::cheb(0:K-1), Psi(1:L), X(1:L), Y(1:L), Z(1:L), Lnorm1
Lnorm1 = 2/DeltaX                ! ||L||_1
X = Psi                          ! T_0 = 1
call Lpsi(X,Y)                   ! Lpsi(X,Y) ⇔ Y = L X
Y = Y / Lnorm1                   ! cheb(0) = J_0(t||L||_1), cheb(1) = 2*J_1(t||L||_1)
Psi = cheb(0)*X + cheb(1)*Y
do i = 2, K-1
call Lpsi(Y,Z)                   ! T_{i+1} psi = (2/||L||_1)*L*T_{i} psi + T_{i-1} psi
Z = Z*(2/Lnorm1) + X
X = Y                            ! T_{i-1} psi ← T_{i} psi
Y = Z                            ! T_{i} psi ← T_{i+1} psi
Psi = Psi + cheb(i)*Y            ! cheb(i) = 2*J_i(t||L||_1)
enddo
```

18.6 EXTENSION TO LINEAR DISPERSIVE MEDIA

This section shows how the general strategy of constructing unconditionally stable algorithms can be extended to the case of linear dispersive materials. Conceptually, the approach is related to the auxiliary differential equation method discussed in Chapter 9. We consider fully three-dimensional systems, and interpret the symbol $\nabla\times$ as either the differential operator for the continuum case or the corresponding difference operator on the Yee space lattice.

For purposes of illustration, we consider a Lorentz medium characterized by a single pole pair (see Chapter 9, Section 9.2.2). Note that the same technique can be used to treat the Debye and Drude dispersions, as well as generalizations to multipole pairs or magnetic response. The Lorentz model for linear dispersion assumes that: (1) the electrical polarization $\boldsymbol{P} = \boldsymbol{P}(\boldsymbol{r}, t)$ is related to the displacement $\boldsymbol{R} = \boldsymbol{R}(\boldsymbol{r}, t)$ of the local electron cloud in the molecules by $\boldsymbol{P} = eN(\boldsymbol{r})\boldsymbol{R}$, where $N(\boldsymbol{r})$ is the density of molecules at position $\boldsymbol{r}$; and (2) $\boldsymbol{R}$ obeys the equation of motion of a damped oscillator driven by the effective electric field $\boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{P}/3\varepsilon_0$ [23]. Denoting the damping coefficient and bare frequency of the oscillators by $\delta_1 = \delta_1(\boldsymbol{r})$ and $\Omega = \Omega(\boldsymbol{r})$, respectively, the equation of motion for the polarization is given by [23]

$$\frac{\partial^2 \boldsymbol{P}}{\partial t^2} + 2\delta_1 \frac{\partial \boldsymbol{P}}{\partial t} + \Omega^2 \boldsymbol{P} = \frac{e^2 N}{m}\boldsymbol{E} + \frac{e^2 N}{3m\varepsilon_0}\boldsymbol{P} \tag{18.29}$$

where m denotes the mass of the electrons in the molecules. In the notation of Chapter 9, $\omega_1^2 = \Omega^2 - e^2 N/3m\varepsilon_0 \geq 0$, $\Delta\varepsilon_1 = 3(\Omega^2 - \omega_1^2)/\omega_1^2 \geq 0$, and (9.46b) is obtained from (18.29) by differentiating with respect to time. Next, we write Maxwell's curl equations in the matrix form

$$\frac{\partial}{\partial t}\begin{pmatrix}\boldsymbol{E}\\ \boldsymbol{B}\\ \boldsymbol{P}\\ \boldsymbol{Q}\end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{\varepsilon_0\mu_0}\nabla\times & 0 & -\frac{1}{\varepsilon_0}\\ -\nabla\times & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ \omega_1^2\varepsilon_0\Delta\varepsilon_1 & 0 & -\omega_1^2 & -2\delta_1 \end{pmatrix}\begin{pmatrix}\boldsymbol{E}\\ \boldsymbol{B}\\ \boldsymbol{P}\\ \boldsymbol{Q}\end{pmatrix} \tag{18.30}$$

where we have introduced the set of auxiliary variables $\boldsymbol{Q} = \boldsymbol{Q}(\boldsymbol{r})$ to write the system as a set of first-order differential equations. Note that each entry in (18.30) represents a matrix itself.

We can "skew-symmetrize" the matrix in (18.30) by a simple rescaling transformation; that is, we search for a diagonal matrix $[D]$ $(D_{ii} \equiv d_i)$ such that

$$
\begin{aligned}
[L] &= [D]^{-1} \begin{pmatrix} 0 & \dfrac{1}{\varepsilon_0 \mu_0} \nabla \times & 0 & -\dfrac{1}{\varepsilon_0} \\ -\nabla \times & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \omega_1^2 \varepsilon_0 \Delta\varepsilon_1 & 0 & -\omega_1^2 & -2\delta_1 \end{pmatrix} [D] \\
&= \begin{pmatrix} 0 & c\nabla \times & 0 & -\omega_1 \sqrt{\Delta\varepsilon_1} \\ -c\nabla \times & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_1 \\ \omega_1 \sqrt{\Delta\varepsilon_1} & 0 & -\omega_1 & -2\delta_1 \end{pmatrix}
\end{aligned}
\tag{18.31}
$$

After some algebra, we obtain $(d_2/d_1)^2 = \varepsilon_0 \mu_0$, $(d_4/d_1)^2 = \varepsilon_0^2 \Delta\varepsilon_1 \omega_1^2$, and $(d_4/d_3)^2 = \omega_1^2$. Choosing $d_1 = 1$ yields the second line of (18.31). In terms of the rescaled fields, (18.30) reads

$$
\frac{\partial}{\partial t} \begin{pmatrix} \boldsymbol{E}/d_1 \\ \boldsymbol{B}/d_2 \\ \boldsymbol{P}/d_3 \\ \boldsymbol{Q}/d_4 \end{pmatrix} = \begin{pmatrix} 0 & c\nabla \times & 0 & -\omega_1 \sqrt{\Delta\varepsilon_1} \\ -c\nabla \times & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_1 \\ \omega_1 \sqrt{\Delta\varepsilon_1} & 0 & -\omega_1 & -2\delta_1 \end{pmatrix} \begin{pmatrix} \boldsymbol{E}/d_1 \\ \boldsymbol{B}/d_2 \\ \boldsymbol{P}/d_3 \\ \boldsymbol{Q}/d_4 \end{pmatrix}
\tag{18.32}
$$

Note that d_1 and d_2 do not depend on $\boldsymbol{r}$, but d_3 and d_4 do. Disregarding its negative diagonal elements, the matrix of (18.32) is clearly skew-symmetric. According to Appendix 18A.2, $\| e^{t[L]} \| \leq e^{t\rho([L])}$, where $\rho([L])$ is the largest eigenvalue of $([L] + [L]^{\mathrm{T}})/2$. From (18.32), it follows that $\rho([L]) = 0$, implying that the formal solution of (18.32) cannot diverge with time.

We have two main options to construct unconditionally stable algorithms. We can decompose $[L]$ into a skew-symmetric $([L]')$ and nonpositive definite diagonal $([L]'')$ matrix. Appendix 18A.2 shows that employing an unconditionally stable algorithm to compute $e^{t[L]'}\{\Phi\}$ guarantees that $e^{\tau[L]''/2} e^{\tau[L]'} e^{\tau[L]''/2}$ is an unconditionally stable algorithm. Alternatively, we write $[L] = [L_1] + [L_2]$ where

$$
[L_1] = \begin{pmatrix} 0 & c\nabla \times & 0 & 0 \\ -c\nabla \times & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
\tag{18.33}
$$

$$[L_2] = \begin{pmatrix} 0 & 0 & 0 & -\omega_1\sqrt{\Delta\varepsilon_1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_1 \\ \omega_1\sqrt{\Delta\varepsilon_1} & 0 & -\omega_1 & -2\delta_1 \end{pmatrix} \tag{18.34}$$

Clearly, $[L_1]$ describes the propagation of the electromagnetic fields in a vacuum. The calculation of $e^{t[L_2]}\{\Phi\}$ is straightforward, amounting to solving three coupled linear differential equations for each location $\boldsymbol{r}$ at which material is present [otherwise, $\Delta\varepsilon_1 = \Delta\varepsilon_1(\boldsymbol{r}) = 0$]. This can be done analytically. For each position $\boldsymbol{r}$, the matrix exponential of $[L_2]$ reads

$$e^{t[L_2]} = \begin{pmatrix} \dfrac{1+[C(t)+\delta_1 S(t)]\Delta\varepsilon_1}{1+\Delta\varepsilon_1} & 0 & \dfrac{[1-C(t)-\delta_1 S(t)]\sqrt{\Delta\varepsilon_1}}{1+\Delta\varepsilon_1} & -\omega_1\sqrt{\Delta\varepsilon_1}S(t) \\ 0 & 1 & 0 & 0 \\ \dfrac{[1-C(t)-\delta_1 S(t)]\sqrt{\Delta\varepsilon_1}}{1+\Delta\varepsilon_1} & 0 & \dfrac{\Delta\varepsilon_1 + C(t) + \delta_1 S(t)}{1+\Delta\varepsilon_1} & \omega_1 S(t) \\ \omega_1\sqrt{\Delta\varepsilon_1}S(t) & 0 & -\omega_1 S(t) & C(t)-\delta_1 S(t) \end{pmatrix} \tag{18.35}$$

where $C(t) = e^{-\delta_1 t}\cos wt$, $S(t) = w^{-1}e^{-\delta_1 t}\sin wt$, and $w = [\omega_1^2(1+\Delta\varepsilon_1) - \delta_1^2]^{1/2}$. For each position $\boldsymbol{r}$, the eigenvalues of (18.35) are 1, 1, $e^{-\delta_1 t - jwt}$, and $e^{-\delta_1 t + jwt}$. Therefore, the calculation of $e^{t[L_2]}\{\Phi\}$ is not only exact, but also unconditionally stable for any value of the time-step τ and damping $\delta_1 \geq 0$.

The practical implementation of algorithms of this type is much easier than the algebra involved in constructing the algorithms might suggest. We start from the basic product formula

$$\{\Psi(t+\tau)\} = e^{\tau[L_2]/2} e^{\tau[L_1]} e^{\tau[L_2]/2} \{\Psi(t)\} \tag{18.36}$$

where τ denotes the time-step. To compute $e^{\tau[L_1]}\{\Phi\}$, any unconditionally stable algorithm can be used. Given the time-step τ, we compute and store the nine nonzero elements of the matrix of (18.35) for each $\boldsymbol{r}$ at which material is present. This needs to be done once. To compute $e^{\tau[L_2]/2}\{\Phi\}$, we make a loop over all gridpoints at which material is present, and multiply the array $\{\boldsymbol{E}(\boldsymbol{r}, t), \boldsymbol{P}(\boldsymbol{r}, t), \boldsymbol{Q}(\boldsymbol{r}, t)\}$ by the corresponding 3×3 matrix; that is, the nonzero matrix elements of (18.35) for that particular gridpoint $\boldsymbol{r}$. Note that this procedure is fundamentally different from the one described in Chapter 9, Section 9.4: instead of updating the three fields sequentially, we update them simultaneously.

The modularity of the product-formula approach adds extra flexibility to FDTD algorithms. For instance, for some applications we may want to use different algorithms (e.g., the Yee, Fourier-space, or one-step algorithms) to perform the free-space time-stepping. In practice, this means using another code to compute $e^{\tau[L_1]}\{\Phi\}$ without changing the code that simulates the material properties.

18.7 EXTENSION TO PERFECTLY MATCHED LAYER ABSORBING BOUNDARY CONDITIONS

This section discusses the numerical stability of FDTD schemes that implement *uniaxial perfectly matched layer* (UPML) *absorbing boundary conditions* (ABCs) [24, 25]. We explore the conditions that allow constructing unconditionally stable algorithms that use UPML ABCs. We consider only the case of two-dimensional TE modes with a UPML ABC in the x-direction. Other cases are left for future research.

Our starting point is the UPML formulation of [24]. For convenience, we repeat the relevant equations:

$$j\omega s_x^{-1} E_x = \frac{\partial H_z}{\partial y}; \quad j\omega s_x E_y = -\frac{\partial H_z}{\partial x}; \quad j\omega s_x H_z = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \tag{18.37}$$

where $s_x = \kappa_x + \sigma_x / j\omega\varepsilon$. Introducing the field $B_z = s_x H_z$, we have

$$j\omega E_x = \frac{\partial B_z}{\partial y}; \quad j\omega s_x^2 E_y = -\frac{\partial B_z}{\partial x} + B_z \frac{1}{s_x}\frac{\partial s_x}{\partial x}; \quad j\omega B_z = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \tag{18.38}$$

To simplify the notation in the remainder of this section, we put $\varepsilon_0 = \mu_0 = \varepsilon_1 = \mu_1 = 1$, and take $s_x = 1 + \sigma_x / j\omega$. Then, (18.38) can be written as

$$\frac{\partial}{\partial t}\begin{pmatrix} B_z \\ E_x \\ E_y \\ Q \end{pmatrix} \equiv [L]\begin{pmatrix} B_z \\ E_x \\ E_y \\ Q \end{pmatrix} = \begin{pmatrix} 0 & \partial/\partial y & -\partial/\partial x & 0 \\ \partial/\partial y & 0 & 0 & 0 \\ -\partial/\partial x & 0 & -2\sigma_x & -\sigma_x \\ 0 & 0 & \sigma_x & 0 \end{pmatrix}\begin{pmatrix} B_z \\ E_x \\ E_y \\ Q \end{pmatrix} \tag{18.39}$$

where the auxiliary field Q is introduced to write the equations in the standard form (18.4). As in Section 18.6, we make a minor abuse of notation by interpreting the symbols $\partial/\partial x$ and $\partial/\partial y$ as differential operators if we discuss the continuum equations, or as the corresponding difference operators on the Yee grid.

We recognize that (18.33) and (18.39) have the same mathematical properties. The matrix in (18.39) is the sum of a skew-symmetric and a nonpositive diagonal matrix. Thus, we can apply the same strategies to construct unconditionally stable algorithms. For instance, repeating the steps of Section 18.6, we write $[L] = [L_1] + [L_2]$ where

$$[L_1] = \begin{pmatrix} 0 & \partial/\partial y & -\partial/\partial x & 0 \\ \partial/\partial y & 0 & 0 & 0 \\ -\partial/\partial x & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad [L_2] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2\sigma_x & -\sigma_x \\ 0 & 0 & \sigma_x & 0 \end{pmatrix} \tag{18.40}$$

To construct an unconditionally stable algorithm, we use the second-order formula, where τ denotes the time-step. To compute $e^{\tau[L_1]}\{\Phi\}$, any of the unconditionally stable algorithms can be used. The matrix exponential of $[L_2]$ is straightforward to obtain analytically, and is given by

$$e^{\tau[L_2]} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-\tau\sigma_x}(1-\tau\sigma_x) & -\tau\sigma_x e^{-\tau\sigma_x} \\ 0 & 0 & \tau\sigma_x e^{-\tau\sigma_x} & e^{-\tau\sigma_x}(1+\tau\sigma_x) \end{pmatrix} \tag{18.41}$$

The eigenvalues of (18.41) are 1 or $e^{-\tau\sigma_x}$, implying that calculating $e^{\tau[L_2]/2}\{\Phi\}$ is numerically stable if $\tau\sigma_x \geq 0$ (which is the relevant case). Thus, we conclude that we have constructed an unconditionally stable algorithm for the TE mode and the UPML ABC in the x-direction.

Note that, instead of starting from the UPML formulation, we can also construct an unconditionally stable algorithm starting from the stretched-coordinate formulation (Chapter 7, Section 7.4), and repeating the steps leading to (18.39). This yields the same equations, with $D_x = s_x E_x$ replacing E_x.

Finally, as discussed in Chapter 7, in any discrete formation of Maxwell's equations, perfect transmission into the PML is lost and spurious wave reflection occurs. The standard procedure to reduce these artifacts is to introduce a grading of the PML parameters (here, σ_x and κ_x). However, a subtle problem now appears. The construction that leads to (18.39) explicitly uses the fact that $\partial s_x/\partial x = 0$. However, if $\partial s_x/\partial x \neq 0$, we have not yet been able to transform the set of continuum equations (18.38) into a form that explicitly shows that these equations are stable.

In retrospect, this is not surprising. We have repeatedly seen that unconditional stability is intimately related to the behavior of the underlying physical system, requiring energy conservation or energy dissipation. The UPML formulation is based on a physical medium, with constitutive equations of the form $\boldsymbol{D}(\omega, \boldsymbol{k}) = \varepsilon(\omega, \boldsymbol{k})\boldsymbol{E}(\omega, \boldsymbol{k})$ and $\boldsymbol{B}(\omega, \boldsymbol{k}) = \mu(\omega, \boldsymbol{k})\boldsymbol{H}(\omega, \boldsymbol{k})$ [23]. Transformation to the time-space domain yields a convolution over time *and* space. In current implementations of UPML, the latter is approximated by its local value. While very reasonable from a practical point of view, this approximation is not compatible with the derivation of reflectionless transmission at the UPML interface. Finding a UPML medium that yields reflectionless transmission after discretization, and does not violate the constitutive equations, is a challenging theoretical problem that we leave for future research.

18.8 SUMMARY

The main theme of this chapter is that there is a systematic, constructive procedure to generate unconditionally stable FDTD algorithms. Examples of the application of this procedure include:

- Unconditionally stable real-space and Fourier-space algorithms;
- A one-step algorithm that permits the use of very large time-steps;
- Unconditionally stable algorithms for linear dispersive media;
- Unconditionally stable algorithms for UPML in two dimensions.

An important conceptual consequence of the constructive approach is that there is an intimate relation between the unconditional stability of an algorithm and the energy conservation (or dissipation) in the physical model. Further research is necessary to establish whether these recent developments can extend the range of applications of the FDTD method.

APPENDIX 18A: SOME TECHNICAL DETAILS

18A.1 Convolution Equation (18.9)

The convolution of Equation (18.9) is simplified by invoking the n-point Gauss-Legendre quadrature formulas

$$\int_0^\tau e^{u[L]}\{S(t+\tau-u)\}\,du \;=\; \tau\sum_{i=1}^{n} w_i\, e^{(1+x_i)[L]/2}\{S(t+(1-x_i)\tau/2)\} \;+\; O(\tau^{2n+1}) \qquad (18A.1)$$

where x_i is the i'th zero of the Legendre polynomial $P_n(x)$, and the weights are given by $w_i = 1/(1-x_i^2)[P_n'(x_i)]^2$. From (18A.1), we see that an error reduction by a factor τ^2 comes at the cost of an extra calculation of the type $e^{u[L]}\{S(v)\}$. In many applications, the temporal and spatial dependence of the source-term factors is given by $\{S(t)\} = f(t)\{S_0\}$, where $f(t)$ is a scalar function [e.g., $f(t) = \sin\Omega t$]. Then, (18A.1) reduces to [17]

$$\int_0^\tau e^{u[L]}\{S(t+\tau-u)\}\,du \;=\; \tau\sum_{i=1}^{n} f[t+(1-x_i)\tau/2]\; w_i\, e^{(1+x_i)[L]/2}\{S_0\} \qquad (18A.2)$$

where the computational work involves n calculations of the type $e^{u[L]}\{S_0\}$. For $n = 1, 2, 3$, this work is negligible compared to that required to perform the time-stepping itself.

18A.2 Error Bounds on Product Formula Approximations

To give a rigorous proof of the basic results used in this chapter, it is useful to recall several facts from linear algebra [3, 9, 10]. First, we define the norm of a matrix $[A]$ as $\|[A]\| =$ (maximum eigenvalue of $[A]^{\mathrm{T}}[A]$)$^{1/2}$. Second, for any orthogonal matrix $[R]$ (a rotation), we have $[R]^{\mathrm{T}}[R] = 1$, and therefore $\|[R]\| = 1$.

In addition, for two square matrices $[A]$ and $[B]$, we have the identity [6]

$$e^{\tau([A]+[B])} - e^{\tau[A]/2}e^{\tau[B]}e^{\tau[A]/2} = \frac{1}{4}\int_0^\tau d\lambda\int_0^\lambda d\mu\int_0^\mu d\kappa \begin{pmatrix} e^{\lambda[A]/2}e^{\lambda[B]}e^{\kappa[A]/2}\langle[A],\langle[A],[B]\rangle\rangle \\ \cdot\; e^{(\lambda-\kappa)[A]/2}\,e^{(\tau-\lambda)([A]+[B])} \end{pmatrix}$$

$$+\; \frac{1}{2}\int_0^\tau d\lambda\int_0^\lambda d\mu\int_0^\mu d\kappa \begin{pmatrix} e^{\lambda[A]/2}\,e^{(\lambda-\kappa)[B]}\langle[B],\langle[A],[B]\rangle\rangle \\ \cdot\; e^{\kappa[B]}\,e^{\lambda[A]/2}\,e^{(\tau-\lambda)([A]+[B])} \end{pmatrix} \qquad (18A.3)$$

where $\langle [A], [B] \rangle = [A][B] - [B][A]$ is a symbol for the commutator of $[A]$ and $[B]$. From (18A.3), it is clear that the right-hand side vanishes if $[A][B] = [B][A]$. Without loss of generality, we may assume that $\tau \geq 0$. Repeated application of $\|[A][B]\| \leq \|[A]\| \|[B]\|$ to identity (18A.3) yields [7, 26]

$$\left\| e^{\tau([A]+[B])} - e^{\tau[A]/2} e^{\tau[B]} e^{\tau[A]/2} \right\| \leq \frac{\tau^3}{24} \begin{pmatrix} \left\| \langle [A], \langle [A], [B] \rangle \rangle \right\| + \\ 2 \left\| \langle [B], \langle [A], [B] \rangle \rangle \right\| \end{pmatrix} e^{\tau(\|[A]\|+\|[B]\|)} \tag{18A.4}$$

Using the identity $[X]^m - [Y]^m = [X]^{m-1}([X]-[Y]) + [X]^{m-2}([X]-[Y])[Y] + \ldots + ([X]-[Y])\,[Y]^{m-1}$ and the triangle inequality $\|[X] + [Y]\| \leq \|[X]\| + \|[Y]\|$, we can find an upper bound on the error after m time-steps. For general matrices $[A]$ and $[B]$, we have [7]

$$\left\| e^{m\tau([A]+[B])} - (e^{\tau[A]/2} e^{\tau[B]} e^{\tau[A]/2})^m \right\| \leq \frac{m\tau^3}{24} \begin{pmatrix} \left\| \langle [A], \langle [A], [B] \rangle \rangle \right\| + \\ 2 \left\| \langle [B], \langle [A], [B] \rangle \rangle \right\| \end{pmatrix} e^{m\tau(\|[A]\|+\|[B]\|)} \tag{18A.5}$$

Putting $m\tau = t$, we conclude that

$$e^{t([A]+[B])} = \lim_{m \to \infty} \left(e^{t[A]/2m} e^{t[B]/m} e^{t[A]/2m} \right)^m \tag{18A.6}$$

We can use (18A.6) to derive much stronger results than (18A.4) and (18A.5). Substituting $[A] = ([L] - [L]^{\mathrm{T}})/2$ and $[B] = ([L] + [L]^{\mathrm{T}})/2$ in (18A.6), and taking the norms we find

$$\left\| e^{t[L]} \right\| \leq \lim_{m \to \infty} \left(\left\| e^{t([L]-[L]^{\mathrm{T}})/4m} \right\| \left\| e^{t([L]+[L]^{\mathrm{T}})/2m} \right\| \left\| e^{t([L]-[L]^{\mathrm{T}})/4m} \right\| \right)^m = \left\| e^{t([L]+[L]^{\mathrm{T}})/2m} \right\|^m \tag{18A.7}$$

where we used the fact that since $[A] = ([L] - [L]^{\mathrm{T}})/2$ is skew-symmetric, $e^{t([L]-[L]^{\mathrm{T}})/4m}$ is an orthogonal matrix, and therefore $\|e^{t([L]-[L]^{\mathrm{T}})/4m}\| = 1$. Denoting the largest eigenvalue of the symmetric matrix $([L] + [L]^{\mathrm{T}})/2$ by $\rho([L])$, the largest eigenvalue of $e^{t([L]+[L]^{\mathrm{T}})/2m}$ is $e^{t\rho([L])/m}$. We find that

$$\left\| e^{t[L]} \right\| \leq e^{t\rho([L])} \tag{18A.8}$$

If $[L]$ itself is skew-symmetric, then $\rho([L]) = 1$ and (18A.8) implies that $\|e^{t([L])}\| = 1$, as it should be.

It is now straightforward to show that we can improve the bounds (18A.4) and (18A.5) by repeated use of (18A.8). For general matrices $[A]$ and $[B]$, we obtain

$$\left\| e^{\tau([A]+[B])} - e^{\tau[A]/2} e^{\tau[B]} e^{\tau[A]/2} \right\| \leq \frac{\tau^3}{24} \begin{pmatrix} \left\| \langle [A], \langle [A], [B] \rangle \rangle \right\| + \\ 2 \left\| \langle [B], \langle [A], [B] \rangle \rangle \right\| \end{pmatrix} e^{\tau\rho([A])+\tau\rho([B])} \tag{18A.9}$$

and

$$\left\| e^{m\tau([A]+[B])} - (e^{\tau[A]/2}\, e^{\tau[B]}\, e^{\tau[A]/2})^m \right\| \le \frac{m\tau^3}{24}\begin{pmatrix} \left\| \langle [A], \langle [A],[B] \rangle \rangle \right\| + \\ 2\left\| \langle [B], \langle [A],[B] \rangle \rangle \right\| \end{pmatrix} e^{m\tau\rho([A])+m\tau\rho([B])} \tag{18A.10}$$

This shows that the error between the exact time-evolution operator and the product-formula approximation cannot grow faster than the time $t = m\tau$, and vanishes like τ^2. If $[A]$ and $[B]$ are skew-symmetric matrices, the last factor in (18A.9) and (18A.10) is equal to one.

In the presence of dissipation, we have $[L] = [L'] + [L'']$, where $[L']$ is a skew-symmetric matrix, and the absorbing medium is represented by a *symmetric* matrix $[L'']$. As all the eigenvalues of $[L'']$ are nonpositive, we have $\|e^{\tau[L]}\| \le e^{\tau\rho([L''])} \le 1$. Furthermore, $\|e^{\tau[L']/2}\, e^{\tau[L'']}\, e^{\tau[L']/2}\| = \|e^{\tau[L'']}\| \le 1$, implying that the product formula $e^{\tau[L']/2}\, e^{\tau[L'']}\, e^{\tau[L']/2}$ defines an unconditionally stable algorithm.

APPENDIX 18B: STABILITY ANALYSIS OF EQUATION (18.17)

The stability analysis of Chapter 4 assumed a homogenous FDTD modeling space, the existence of sinusoidal waves, and the knowledge of the dispersion relation for waves propagating on the Yee lattice. Here, we derive a general stability criterion for the Yee algorithm by exploiting the underlying product-formula structure. First we note that from

$$\begin{aligned} \left\| \left(e^{\tau[L_1]/2}\, e^{\tau[L_2]}\, e^{\tau[L_1]/2} \right)^m \right\| &= \left\| e^{-\tau[L_1]/2} \left(e^{\tau[L_1]}\, e^{\tau[L_2]} \right)^m e^{\tau[L_1]/2} \right\| \\ &\le \left\| e^{-\tau[L_1]/2} \right\| \left\| e^{\tau[L_1]/2} \right\| \left\| e^{\tau[L_1]}\, e^{\tau[L_2]} \right\|^m \end{aligned} \tag{18B.1}$$

and

$$\begin{aligned} \left\| \left(e^{\tau[L_1]}\, e^{\tau[L_2]} \right)^m \right\| &= \left\| e^{\tau[L_1]/2} \left(e^{\tau[L_1]/2}\, e^{\tau[L_2]}\, e^{\tau[L_1]/2} \right)^m e^{-\tau[L_1]/2} \right\| \\ &\le \left\| e^{\tau[L_1]/2} \right\| \left\| e^{-\tau[L_1]/2} \right\| \left\| e^{\tau[L_1]/2}\, e^{\tau[L_2]}\, e^{\tau[L_1]/2} \right\|^m \end{aligned} \tag{18B.2}$$

it follows immediately that stability of the Yee algorithm implies stability of the equal-time formulation (18.17) and vice versa. Next, we observe that

$$[L]^{\mathrm{T}}[L] = \begin{pmatrix} [A]^{\mathrm{T}}[A] & \\ 0 & [A]\,[A]^{\mathrm{T}} \end{pmatrix} \tag{18B.3}$$

In general [as in example (18.14)], $[A]$ is not a square matrix. Since $([L]\{\Phi\})^{\mathrm{T}}([L]\{\Phi\}) \ge 0$, all eigenvalues of $[L]^{\mathrm{T}}[L]$ are nonnegative. Because of the block-diagonal structure of (18B.3), all eigenvalues of $[A]^{\mathrm{T}}[A]$ and $[A][A]^{\mathrm{T}}$ are also nonnegative.

The eigenvalues of the time-step operator

$$e^{\tau[L_1]}e^{\tau[L_2]} = \begin{pmatrix} [I]-\tau^2[A]^{\mathrm{T}}[A] & -\tau[A]^{\mathrm{T}} \\ \tau[A] & [I] \end{pmatrix} \tag{18B.4}$$

where $[I]$ denotes the unit matrix, are the solution of the secular equation

$$\left(e^{\tau[L_1]}e^{\tau[L_2]} - \lambda[I]\right)\{\Phi\} = \begin{pmatrix} (1-\lambda)[I]-\tau^2[A]^{\mathrm{T}}[A] & -\tau[A]^{\mathrm{T}} \\ \tau[A] & (1-\lambda)[I] \end{pmatrix}\begin{pmatrix} \{\Phi_1\} \\ \{\Phi_2\} \end{pmatrix} \tag{18B.5}$$

Assuming that $\lambda \neq 1$, we eliminate $\{\Phi_2\}$ from (18B.5) and obtain

$$\left((1-\lambda)^2[I]+\lambda\tau^2[A]^{\mathrm{T}}[A]\right)\{\Phi_1\} = 0 \tag{18B.6}$$

The linear set of equations (18B.6) has nonzero solutions $\{\Phi_1\}$ if

$$\det\left((1-\lambda)^2[I] + \lambda\tau^2[A]^{\mathrm{T}}[A]\right) = \prod_{j=1}^{(n+1)/2}\left((1-\lambda)^2[I]+\lambda\tau^2 a_j^2\right) = 0 \tag{18B.7}$$

where the eigenvalues of $[A]^{\mathrm{T}}[A]$ are denoted by a_j^2. The solutions of (18B.7) are given by

$$\lambda = 1 - \frac{\tau^2 a_j^2}{2} \pm \sqrt{\left(1-\frac{\tau^2 a_j^2}{2}\right)^2 - 1} \tag{18B.8}$$

Numerical stability of the algorithm requires that $|\lambda| \leq 1$ for all j. This is equivalent to the condition that $\tau^2 a_j^2 \leq 4$ for all j. From our definition of the matrix norm, it follows that this condition is satisfied if $\tau^2\|[A]\|^2 \leq 4$. In general, finding the largest eigenvalues of (the very large) $[A]^{\mathrm{T}}[A]$ is of comparable difficulty as solving Maxwell's equations in the frequency domain. Fortunately, it is easy to a find good approximation to $\|[A]\|^2$. From linear algebra, we know that $\|[A]\|^2 \leq \|[A]^{\mathrm{T}}\|_1 \|[A]\|_1$, where $\|[A]\|_1 = \max_j(\sum_i |A_{ij}|)$ [9]. Thus, the Yee algorithm is numerically stable if

$$\tau \leq \frac{2}{\|[A]\|} \leq \frac{2}{\sqrt{\|[A]^{\mathrm{T}}\|_1 \|[A]\|_1}} \tag{18B.9}$$

For FDTD problems, $[A]$ is sparse [see example (18.16)], and the right-hand side of (18B.9) is easy to compute. This is also the case when the permittivity and permeability vary in space. For the matrix (18.16) of the one-dimensional example, we have $\|[A]\|_1 = 2/\Delta x$ and $\|[A]^{\mathrm{T}}\|_1 = 2/\Delta x$, and we recover the stability condition of (2.52), $\tau \leq \Delta x$. For a homogeneous system represented on the Yee space lattice (see Fig. 3.1), $\|[A]\|_1$ "counts" the maximum number of $\boldsymbol{H}$-components that appear in the update rule of the $\boldsymbol{E}$-components. Similarly, $\|[A]^{\mathrm{T}}\|_1$ "counts" the maximum number of $\boldsymbol{E}$-components that appear in the update rule of the $\boldsymbol{H}$-components.

For the TM_z mode in a two-dimensional system, inspection of (4.2c) shows that $\|[A]\|_1 = 2/\Delta x + 2/\Delta y$ and $\|[A]^T\|_1 = 2/\min(\Delta x, \Delta y)$. If $\Delta x = \Delta y$, condition (18B.9) reduces to $\tau \le \Delta x/\sqrt{2}$, which agrees with (4.66). For a cubic Yee lattice, we find $\tau \le \Delta x/2$, which is not as sharp as the result $\tau \le \Delta x/\sqrt{3}$ obtained by using the exact value $\|[A]\| = 2\sqrt{3}/\Delta x$ for the homogeneous system.

APPENDIX 18C: STABILITY ANALYSIS OF EQUATION (18.19)

After m applications of the ADI time-step matrix (18.19), the growth of the arrays of field components is bounded by $\| ((1-\tau[L_1]/2)^{-1} (1+\tau[L_2]/2) (1-\tau[L_2]/2)^{-1} (1+\tau[L_1]/2))^m \|$. Regrouping factors and introducing $[X_i] = (1-\tau[L_i]/2)^{-1} (1+\tau[L_i]/2)$ for $i = 1, 2$, we can write this as $\| (1+\tau[L_1]/2)^{-1} ([X_1][X_2])^m (1+\tau[L_1]/2) \|$. It is straightforward to show that $[X_i]$ is an orthogonal matrix. Hence, we have

$$\left\| \left(1 + \tau[L_1]/2\right)^{-1} \left([X_1][X_2]\right)^m \left(1 + \tau[L_1]/2\right) \right\| \le \left\| \left(1 + \tau[L_1]/2\right)^{-1} \right\| \left\| \left(1 + \tau[L_1]/2\right) \right\| \tag{18C.1}$$

Since the eigenvalues of the skew-symmetric matrix $[L_1]$ are pure imaginary, the matrix $(1+\tau[L_1]/2)^{-1}$ is nonsingular [9]. Therefore, for any value of the number of time-steps m, the ADI time-step matrix (18.19) is bounded by a finite number.

REFERENCES

[1] Yee, K. S., "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, Vol. 14, 1966, pp. 302–307.

[2] Richtmyer, R. D., *Difference Methods for Initial-Value Problems*, New York: Interscience, 1957.

[3] Bellman, R., *Introduction to Matrix Analysis,* New York: McGraw-Hill, 1970; Philadelphia: SIAM, 1997.

[4] Moler, C., and C. F. Van Loan, "Nineteen dubious ways to compute the exponential of a matrix," *SIAM Review*, Vol. 20, 1978, pp. 801–836.

[5] Smith, G. D., *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, Oxford: Clarendon Press, 1998.

[6] Suzuki, M., "Decomposition formulas for exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics," *J. Math. Phys.*, Vol. 26, 1985, pp. 601–612.

[7] De Raedt, H., "Product formula algorithms for solving the time dependent Schrödinger equation," *Comp. Phys. Rep.*, Vol. 7, 1987, pp. 1–72.

[8] Suzuki, M., "General theory of fractal path integrals with applications to many-body theories and statistical physics," *J. Math. Phys.*, Vol. 32, 1991, pp. 400–407.

[9] Wilkinson, J. H., *The Algebraic Eigenvalue Problem*, Oxford: Clarendon Press, 1965.

[10] Golub, G. H., and C. F. Van Loan, *Matrix Computations*, Baltimore: John Hopkins University Press, 1996.

[11] Kole, J. S., M. T. Figge, and H. De Raedt, "Unconditionally stable algorithms to solve the time-dependent Maxwell equations," *Phys. Rev. E*, Vol. 64, 2001, 066705.

[12] Kole, J. S., M. T. Figge, and H. De Raedt, "Higher-order unconditionally stable algorithms to solve the time-dependent Maxwell equations," *Phys. Rev. E*, Vol. 65, 2002, 066705.

[13] De Raedt, H., K. Michielsen, J. S. Kole, and M. T. Figge, "Solving the Maxwell equations by the Chebyshev method: A one-step finite difference time-domain algorithm," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 3155–3160.

[14] De Raedt, H., K. Michielsen, J. S. Kole, and M. T. Figge, "One-step finite difference time-domain algorithm to solve the Maxwell equations," *Phys. Rev. E*, Vol. 67, 2003, 056706.

[15] Tal-Ezer, H., "Spectral methods in time for hyperbolic equations," *SIAM J. Numer. Anal.*, Vol. 23, 1986, pp. 11–26.

[16] Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, New York: Cambridge, 1986.

[17] De Raedt, H., J. S. Kole, K. Michielsen, and M. T. Figge, "Unified framework for numerical methods to solve the Maxwell equations," *Comp. Phys. Comm.*, Vol. 156, 2003, pp. 43–61.

[18] 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 Tech.*, Vol. 48, 2000, pp. 1550–1558.

[19] Namiki, T., "3D ADI-FDTD method—Unconditionally stable time-domain algorithm for solving full vector Maxwell's equations," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2000, pp. 1743–1748.

[20] Zheng, F., and Z. Chen, "Numerical dispersion analysis of the unconditionally stable 3D ADI-FDTD method," *IEEE Trans. Microwave Theory Tech.*, Vol. 49, 2001, pp. 1006–1009.

[21] García, S. G., T. W. Lee, and S. C. Hagness, "On the accuracy of the ADI-FDTD method," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 1, 2002, pp. 31–34.

[22] Abramowitz, M., and I. Stegun, *Handbook of Mathematical Functions*, New York: Dover, 1964.

[23] Born, M., and E. Wolf, *Principles of Optics*, Oxford: Pergamon, 1964.

[24] Gedney, S. D, "The perfectly matched layer absorbing medium," Ch. 5 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Taflove, A., (ed.), Norwood, MA: Artech House, 1998.

[25] Bécache, E., P. G. Petropoulos, and S. D. Gedney, "On the long-time behavior of unsplit perfectly matched layers," *IEEE Trans. Antennas Propagat.*, Vol. 52, 2004, pp. 1335–1342.

[26] Hams, A. H., and H. De Raedt, "Fast algorithm for finding the eigenvalue distribution of very large matrices," *Phys. Rev. E*, Vol. 62, 2000, pp. 4365–4377.

PROJECTS

P18.1 Implement the Chebyshev polynomial algorithm for the one-dimensional TM mode problem. Use the Yee grid of Fig. 18.1 to discretize the continuum equations and assume PEC boundary conditions. Compute the time evolution of a Gaussian wave packet and compare with the exact analytical solution for a cavity of length X:

$$E_z(x,t) = \sum_{n=-\infty}^{\infty} \left(e^{-(2nX+x+x_0+ct)^2/\sigma^2} - e^{-(2nX+x-x_0-ct)^2/\sigma^2} \right)$$

P18.2 Write a code to compare the Yee algorithm and the real-space second-order unconditionally stable algorithm with the numerically exact (for a given mesh size δ) solution generated by the Chebyshev polynomial method. Use as an initial state a uniform random distribution of the E and H fields, normalized such that $\|\{\Psi(t=0)\}\| = 1$. Study the behavior of the three algorithms as a function of the Courant number $S = c\tau/\Delta x$. Hint: To facilitate the comparison, consider employing the form of the Yee algorithm that does not use the staggered-in-time formulation.

P18.3 Reuse the codes of P18.2 to implement the fourth-order time-accurate algorithm defined by (18.12), and repeat the error analysis.

P18.4 Write a code to implement (18.35) and (18.36) for the one-dimensional propagation case. Use this code to simulate a Gaussian wave packet impinging normally on a half-space medium having a Lorentz dispersion. (See Chapter 9, Sections 9.2.2 and 9.4.4 for examples of the medium parameters.)

P18.5 Repeat the analysis of Section 18.6 for the case of Debye and Drude media. (See Chapter 9, Sections 9.2.1, 9.2.3, and 9.3.4 for examples of the medium parameters.)

Chapter 19

Advances in Hybrid FDTD-FE Techniques

Thomas Rylander, Fredrik Edelvik, Anders Bondeson, and Douglas Riley

19.1 INTRODUCTION

The FDTD method represents a powerful tool for *computational electromagnetics* (CEM). As shown in previous chapters, it can provide accurate solutions to a very wide range of practical engineering applications. However, a fundamental limitation of the basic algorithm is the Cartesian nature of the finite-differencing scheme. Although an FDTD algorithm can be defined for cylindrical and spherical coordinate systems [1, 2], the most widely used form is tied to a purely rectangular grid. As a result, a geometry that does not conform to this rectangular lattice becomes rendered by a so-called "staircased" approximation. While it seems intuitive that curved surfaces should be well approximated in the limit as the space cell size approaches zero, research has shown that staircasing can result in nonconverged solutions, independent of the level of grid refinement [3]. Consequently, the development of geometry-conforming algorithms represents an important area of research in CEM.

In other scientific disciplines such as structural and fluid dynamics, geometry-conforming algorithms based on *finite-element* (FE) and *finite-volume* (FV) techniques have been widely used for several decades [4–7]. The adoption of these methods for CEM has been relatively slow. A possible reason for this was the prominence of research into moment-method techniques as applied to integral-equation representations for Maxwell's equations. Although Yee's original paper on the FDTD technique appeared in 1966 [8], it was not until the early 1980s that widespread interest within the CEM community began to turn toward grid-based differential methods. Nevertheless, research into the application of more advanced techniques for the solution of the differential Maxwell's equations, beyond the basic FDTD method, remained limited, primarily because of the added complexity of these techniques and various theoretical difficulties.

The theoretical problem of "spurious" modes [6] detracted from the application of traditional nodal-based FE methods to the frequency-domain Helmholtz equation or the time-dependent wave equation, even though FE methods were well-established solution approaches for electrostatic and magnetostatic applications based upon the Laplace and Poisson equations. The fundamental problem was that standard nodal-based FE methods incorrectly model the null-space of the curl-curl operator. In addition, experimentation with nonorthogonal FV methods revealed challenges with late-time instabilities when solving the time-dependent differential Maxwell's equations [9].

Despite these technical difficulties, both FE and FV solution techniques have found practical application within the CEM community. This has been largely due to the use of "penalty" methods [6, 10] to mitigate the spurious-mode problem associated with nodal finite-elements when applied to high-frequency applications, and alternative time-integration or spatial-averaging techniques to mitigate the stability problems associated with FV methods [9, 11]. In addition, the subsequent introduction of vector edge-based finite-elements [12, 13] eliminated the spurious-mode problem for FE techniques by properly modeling the null space of the curl-curl operator.

Currently, the primary problem associated with conformal FE and FV techniques for CEM is that these approaches have much greater computational requirements than the basic FDTD algorithm. Specifically, FE methods generally require solving a very large (although sparse) system of equations. *Finite-volume time-domain* (FVTD) techniques generally require many more arithmetic operations per space cell per time-step than FDTD, even though FVTD techniques are usually explicit and involve only the solution of a diagonal matrix.

To reduce the computational burden associated with the use of conformal grids, a number of hybrid gridding concepts have been investigated. Yee et al. introduced an overlapping-grid scheme that utilized both FV and FDTD algorithms [14]. Their FV algorithm was similar to that described by Madsen and Ziolkowski [15]. The basic grid structure for this method is shown in Fig. 19.1(a).

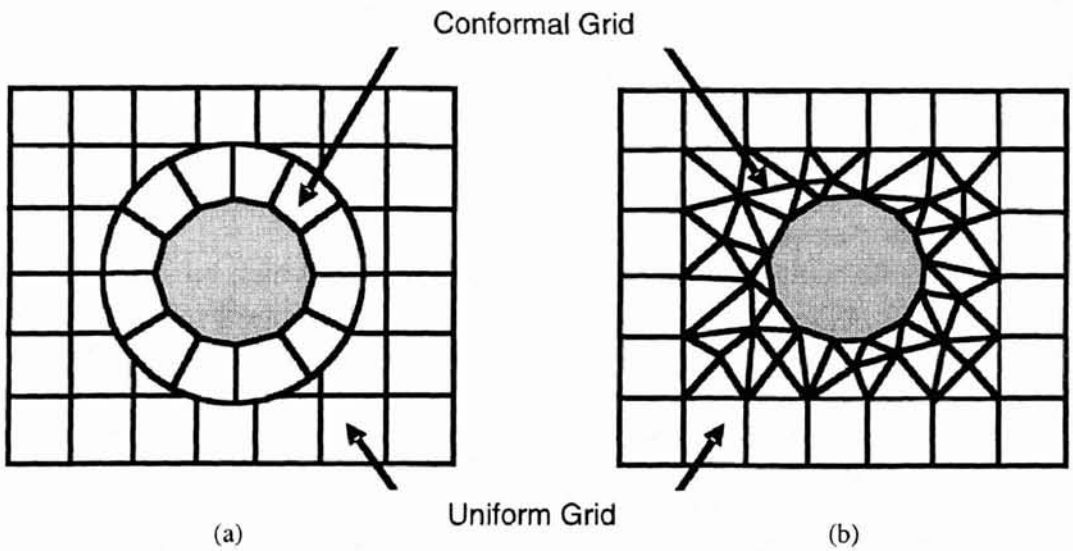


Fig. 19.1 (a) Cross section of the overlapping-grid approach showing a locally conformal mesh overlaid onto a uniform rectangular grid. Spatial interpolations are required at the outer interface of the conformal grid. (b) Local unstructured grid interfaced to a rectangular grid.

The type of hybrid mesh illustrated in Fig. 19.1(a) had been previously widely used for computational fluid dynamics applications, being known as a "Chimera" grid [7]. We note that complicated spatial interpolations are required to connect the grid types at the interface of the locally conformal mesh and the rectangular background mesh. In addition, as proposed, the method was based on a local "mapped-mesh" FV technique. This method extrudes triangular surface elements into three-dimensional prismatic elements. Hence, the locally conforming mesh

is of the "structured" type, which means that the nonorthogonal elements can be referenced using an I-J-K ordering scheme [16] (for comparison, an *unstructured* grid cannot, in general, be referenced in this manner). This technique was shown to provide good results for scattering from a variety of canonical shapes [14, 17, 18]. However, the underlying FV method and grid-interpolation technique rendered the global method weakly unstable.

Riley and Turner [9, 19], and Kalfon and Harran-Klotz [20] reported a generalization of the hybridization concept involving directly connecting an unstructured tetrahedral-element grid to a rectangular FDTD grid, as illustrated in two dimensions in Fig. 19.1(b). The FV technique used in [9, 19] was based on a generalization of the offset-mesh concept proposed by Madsen and Ziolkowski [15], whereas the FV scheme used in [20] was based on a cell-centered, conservative form of Maxwell's equations [21]. Because FV methods typically compute barycentric field values, arbitrary cell shapes are easily accommodated. Consequently, an interface to make the three-dimensional transition between tetrahedra and hexahedra can be accomplished by using a rectangular hexahedron with one quadrilateral face split into two triangles [19]. Alternatively, the transition also can be accomplished by using a more traditional, pyramidal element shape.

An interesting aspect of the offset-mesh, FV method described in [15] is that this technique is a direct generalization of the basic FDTD algorithm to nonorthogonal cells. Hence, the *global* grid can be considered to be purely FV-based, with the only regional difference being the technique used to reference the cells; namely, a structured or unstructured referencing scheme. This concept will be useful later in this chapter when FE hybridizations are discussed. The convergence, accuracy, and stability of tetrahedral-FDTD grid hybridizations using FV methods have been extensively studied by Edelvik et al. [22, 23]. In general, weak late-time instabilities exist for these methods.

Wu and Itoh introduced a two-dimensional hybridization of *finite-element time-domain* (FETD) techniques with FDTD [24]. This type of grid is also illustrated in Fig. 19.1(b). Three-dimensional implementations followed [25, 26]. Here, vector edge-based FE methods were used to avoid the problem of spurious modes associated with nodal elements. In addition, since domains formulated using FE techniques are provably numerically stable [27], the late-time instabilities exhibited by FV techniques on nonorthogonal cells were avoidable. Also, as shown in previous chapters, the basic FDTD algorithm is provably stable as long as the time-step satisfies the Courant condition. Consequently, the hybrid schemes of [24–26] appeared to possess desirable stability characteristics. Unfortunately, even though each constituent algorithm taken individually was provably stable, the resulting hybridization was found to be weakly unstable, similar to that observed for the FV-based hybrids.

In fact, the stability properties of the constituent FE and FDTD algorithms are lost when these techniques are combined in a simple, nonreciprocal manner. As pointed out by Rylander and Bondeson [28], the difficulty can be summarized as follows: At the interface between the two techniques, cell edges exist that are common to both solution domains. The field values on these edges *must* be computed identically by *either* solution method; otherwise, an asymmetry exists that may render the global algorithm unstable. Because FE methods and FDTD both provide a solution to Maxwell's equations, they each provide similar estimates for the field values on these edges, but cannot, in general, be expected to provide numerically identical results. Note that this problem can be avoided by globally using FE techniques on a mixed-element grid, possibly consisting of tetrahedra, hexahedra, and pyramidal elements, since there will be no discontinuity of the solution methods in this case [29, 30]. Hence, a possible approach to obtain a stable FE and FDTD interface is to first explore the potential equivalence of these two solution methods.

It is well-known that by changing the integration method used in FE formulations from Gaussian quadrature to Lobatto quadrature (a generalization of trapezoidal integration), a finite-difference representation can be recovered [4, 31]. In addition, Wong et al. demonstrated that FDTD can be recovered from an edge-based FE formulation on rectangular hexahedra (brick elements) through the use of mass-lumping concepts [32]. Using Lobatto quadrature, Cohen and Monk [33] further demonstrated the equivalence of FDTD and edge-based FE techniques on brick elements. Thus, the basic FDTD scheme is, in fact, a subset of vector edge-based FE formulations when the latter is specialized to rectangular hexahedral elements.

In Reference [28], Rylander and Bondeson reported the successful formulation and testing of a stable, computationally efficient, hybrid FDTD-FE method. The distinguishing feature of this hybrid formulation is that it preserves, by construction, the reciprocity of Maxwell's equations. Reciprocity is preserved by using:

- A curl-conforming representation of the electric field on a space lattice of brick elements and tetrahedrons joined by pyramids;
- A FE treatment of Maxwell's equations based on Galerkin's method;
- Trapezoidal integration on the brick elements to recover the FDTD scheme.

As a consequence, unconditional stability is maintained up to the classic FDTD limit [34, 35] for arbitrary unstructured grids in the FE region. Furthermore, there is zero spectral contamination and nonphysical dissipation, and second-order convergence for smooth bodies.

Topics discussed in this chapter include: an overview of FETD formulations; vector basis functions for tetrahedral, hexahedral, and pyramidal space-lattice elements; proof of the stability of the FDTD-FE hybrid; grid-generation issues; subcell models for arbitrarily oriented thin wires and thin slots; and advanced applications.

19.2 TIME-DOMAIN FINITE ELEMENTS

This section provides an overview of the FE representation of the wave equation, as well as the coupled Maxwell curl equations. The former can be represented in terms of vector edge element basis functions, whereas the latter uses both vector edge elements and vector face, or facet, elements. Although the wave-equation form lends itself to more traditional FE representations, the two coupled curl equations provide a simpler migration from traditional FDTD representations, and is therefore discussed first. For specific details on the FE method, including its history and the development of the various variational representations, the reader is referred to dedicated books on the topic, such as those by Jin [10], Hughes [4], or Zienkiewicz [5].

19.2.1 Coupled Curl Equations

Fig. 19.2 illustrates the problem under consideration. Maxwell's curl equations in the time-domain are given throughout a linear, isotropic, lossless ($\sigma = 0$) volume Ω with boundary $\partial\Omega$ by

$$\nabla \times \left[\frac{1}{\mu} \boldsymbol{B}(\boldsymbol{r},t) \right] = \varepsilon \frac{\partial}{\partial t} \boldsymbol{E}(\boldsymbol{r},t) + \boldsymbol{J}(\boldsymbol{r},t) \qquad \boldsymbol{r} \in \Omega \tag{19.1}$$

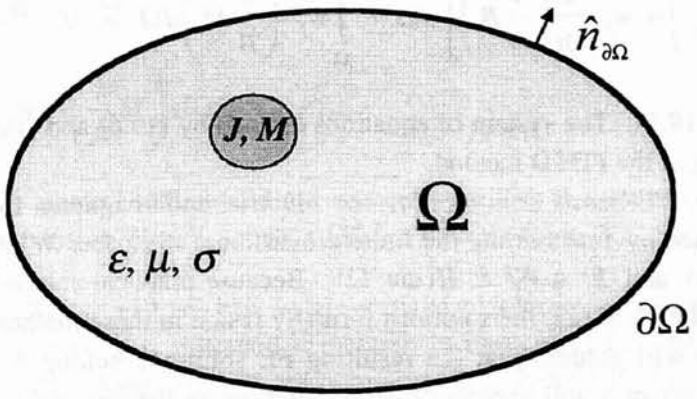


Fig. 19.2 Geometry for the electromagnetic boundary-value problem.

$$\nabla \times \boldsymbol{E}(\boldsymbol{r},t) = -\frac{\partial}{\partial t}\boldsymbol{B}(\boldsymbol{r},t) - \boldsymbol{M}(\boldsymbol{r},t) \qquad \boldsymbol{r} \in \Omega \tag{19.2}$$

where $\boldsymbol{E}(\boldsymbol{r}, t)$ and $\boldsymbol{B}(\boldsymbol{r}, t)$ denote the time-dependent electric and magnetic fields, respectively. For simplicity, a Dirichlet boundary condition on $\partial\Omega$ is assumed:

$$\hat{\boldsymbol{n}}_{\partial\Omega} \times \boldsymbol{E}(\boldsymbol{r},t) = \boldsymbol{M}_{\partial\Omega}(\boldsymbol{r},t) \qquad \boldsymbol{r} \in \partial\Omega \tag{19.3}$$

along with the initial conditions

$$\boldsymbol{E}(\boldsymbol{r},0) = \boldsymbol{E}_0(\boldsymbol{r}) \qquad \boldsymbol{r} \in \Omega \tag{19.4}$$

$$\boldsymbol{B}(\boldsymbol{r},0) = \boldsymbol{B}_0(\boldsymbol{r}) \qquad \boldsymbol{r} \in \Omega \tag{19.5}$$

The finite-element representation of (19.1) to (19.5) is based on the construction of a suitable variational formulation. For this purpose, a weak solution is sought with $\boldsymbol{E} \in \boldsymbol{W}$, $\boldsymbol{B} \in \boldsymbol{W}_f$, where $\boldsymbol{W}$ denotes the space of *curl*-conforming functions defined by $\boldsymbol{W} = \boldsymbol{H}(curl; \Omega) = \{\boldsymbol{w} \in [L^2(\Omega)]^3, \nabla \times \boldsymbol{w} \in [L^2(\Omega)]^3\}$, and $\boldsymbol{W}_f$ denotes the space of *divergence*-conforming functions defined by $\boldsymbol{W}_f = \boldsymbol{H}(div; \Omega) = \{\boldsymbol{w}_f \in [L^2(\Omega)]^3, \nabla \cdot \boldsymbol{w}_f \in [L^2(\Omega)]^3\}$ [12, 36]. The edge (curl-conforming) basis functions provide for tangential continuity between elements, whereas facet (divergence-conforming) basis functions provide for normal continuity. (For the reader familiar with integral-equation formulations, the basis functions typically used in a method of moments setting are those of the divergence-conforming type [37].) The functions $\boldsymbol{E}$ and $\boldsymbol{B}$ are unique solutions to the following variational problem

$$\int_\Omega \left[-(\nabla \times \boldsymbol{w}) \cdot \left(\frac{1}{\mu}\boldsymbol{B} \right) + \varepsilon \boldsymbol{w} \cdot \frac{\partial}{\partial t}\boldsymbol{E} \right] d\Omega + \int_\Omega \boldsymbol{w} \cdot \boldsymbol{J}\, d\Omega = 0 \qquad \forall \boldsymbol{w} \in \boldsymbol{W} \tag{19.6}$$

$$\int_{\Omega}\left[(\nabla\times\boldsymbol{E})\cdot\left(\frac{1}{\mu}\boldsymbol{w}_f\right)+\boldsymbol{w}_f\cdot\frac{\partial}{\partial t}\left(\frac{1}{\mu}\boldsymbol{B}\right)\right]d\Omega+\int_{\Omega}\boldsymbol{w}_f\cdot\left(\frac{1}{\mu}\boldsymbol{M}\right)d\Omega=0\quad\forall\boldsymbol{w}_f\in\boldsymbol{W}_f \tag{19.7}$$

along with (19.3) to (19.5). The system of equations defined by (19.6) and (19.7) can be viewed as a variational analog of the FDTD method.

An approximate, FE-based solution for the electric and magnetic fields, $\boldsymbol{E}^h$ and $\boldsymbol{B}^h$, respectively, is obtained by constructing the finite-dimensional subspaces $\boldsymbol{W}^h$ and $\boldsymbol{W}_f^h$, such that $\boldsymbol{E}^h\in\boldsymbol{W}^h\subset\boldsymbol{H}(\text{curl};\Omega)$ and $\boldsymbol{B}^h\in\boldsymbol{W}_f^h\subset\boldsymbol{H}(div;\Omega)$. Because practical solutions are based on a limited number of finite elements, the solutions formally reside in these mathematical subspaces. For convenience, we will assume that the resulting FE solutions belong to these underlying subspaces. The superscript h will henceforth be omitted unless the distinction between infinite- and finite-dimensional spaces is important.

The semidiscrete (time-continuous) representation for (19.6) and (19.7) is obtained by expanding $\boldsymbol{B}(\boldsymbol{r},t)$ in terms of time-dependent coefficients B_i and spatial (facet) basis functions $\boldsymbol{w}_{f_i}$ as

$$\boldsymbol{B}(\boldsymbol{r},t)=\sum_i B_i(t)\,\boldsymbol{w}_{f_i}(\boldsymbol{r}) \tag{19.8}$$

where $i=1,2,\ldots,N_f$. Similarly, the expansion for $\boldsymbol{E}$ in terms of spatial (edge) basis functions $\boldsymbol{w}_i$ is given by

$$\boldsymbol{E}(\boldsymbol{r},t)=\sum_i E_i(t)\,\boldsymbol{w}_i(\boldsymbol{r}) \tag{19.9}$$

where $i=1,2,\ldots,N_h$. Note that these expansions correspond to a Galerkin formulation for (19.6) and (19.7) due to the equivalence of the basis and testing functions.

Substituting these expansions into (19.6) and (19.7) leads to the following matrix equations:

$$[M]\frac{d}{dt}\{E\}-[A]^{\mathrm{T}}\{B\}+\{f\}=\{0\} \tag{19.10}$$

$$[M_f]\frac{d}{dt}\{B\}+[A]\{E\}+\{g\}=\{0\} \tag{19.11}$$

where $\{E\}=[E_1(t),E_2(t),\ldots,E_{N_h}(t)]^{\mathrm{T}}$ (T denotes transpose), $\{B\}=[B_1(t),B_2(t),\ldots,B_{N_f}(t)]^{\mathrm{T}}$, and

$$M_{ij}=\int_{\Omega}\varepsilon\,\boldsymbol{w}_i\cdot\boldsymbol{w}_j\,d\Omega \tag{19.12}$$

$$M_{f_{ij}}=\int_{\Omega}\frac{1}{\mu}\boldsymbol{w}_{f_i}\cdot\boldsymbol{w}_{f_j}\,d\Omega \tag{19.13}$$

$$A_{ij}=\int_{\Omega}\frac{1}{\mu}\boldsymbol{w}_{f_i}\cdot\left(\nabla\times\boldsymbol{w}_j\right)d\Omega \tag{19.14}$$

$$f_i = \int_\Omega \boldsymbol{w}_i \cdot \boldsymbol{J}\, d\Omega \tag{19.15}$$

$$g_i = \int_\Omega \boldsymbol{w}_{f_i} \cdot \left(\frac{1}{\mu}\boldsymbol{M}\right) d\Omega \tag{19.16}$$

Note that the "mass matrix" $[M]$ should not be confused with the magnetic current density $\boldsymbol{M}$. Matrix $[M]$ is of dimension $N_h \times N_h$, matrix $[M_f]$ is of dimension $N_f \times N_f$, and matrix $[A]$ is of dimension $N_f \times N_h$.

Equations (19.15) and (19.16) can be generalized to accommodate the assumed Dirichlet condition given by (19.3) [33]. Alternative methods to implement Dirichlet conditions are also possible [10]. A Dirichlet condition is appropriate for the hybrid FDTD-FE solution method, which will be discussed in greater detail in Section 19.4. Note that $\{H\}$ is obtained from $\{B\}$ by forming

$$\{H\} = [M_f]\,\{B\} \tag{19.17}$$

This relation will be useful for characterizing thin slots on FE grids, described in Section 19.6.

Because of the fundamental properties relating edge and facet basis functions, which will be discussed in more detail in Section 19.3, (19.11) can be conveniently rewritten as $d/dt\{B\} = -[C]\{E\} - [M_f]^{-1}\{g\}$, where $[C] = [M_f]^{-1}[A]$ is a sparse rectangular (incidence) matrix with nonzero entries consisting simply of ± 1 [32]. Consequently, the magnetic field becomes clearly related to the circulation of the electric field around the cell faces, and a direct analogy with the FDTD representation of Faraday's law is apparent.

The fully discrete representation of (19.10) and (19.11) is obtained through the definition of a suitable time-integration scheme. If a traditional leapfrog method is adopted, then (19.10) to (19.16) form a conditionally stable scheme that can be considered to be a direct generalization of FDTD to a conformal FE-based grid. In this case, the fully discrete system is given by

$$\{B\}|^{n+1/2} = \{B\}|^{n-1/2} - \Delta t\left([C]\{E\}|^n + [M_f]^{-1}\{g\}|^n\right) \tag{19.18}$$

$$\{E\}|^{n+1} = \{E\}|^n + [M]^{-1}\Delta t\left([A]^{\mathrm{T}}\{B\}|^{n+1/2} - \{f\}|^{n+1/2}\right) \tag{19.19}$$

In the following section, matrix $[M]$ is shown to be symmetric positive definite. Consequently, (19.19) can be efficiently solved by using preconditioned conjugate gradient methods [10]. Convergence of the matrix solution is typically obtained with fewer than 10 iterations, provided the mesh is suitably uniform.

19.2.2 Wave Equation

In the absence of magnetic current density sources $\boldsymbol{M}$, the wave equation for the time-dependent total electric field $\boldsymbol{E}(\boldsymbol{r}, t)$ in the volume Ω bounded by the closed surface $\partial\Omega$ is given by

$$\nabla \times \left[\frac{1}{\mu}\nabla \times \boldsymbol{E}(\boldsymbol{r},t)\right] + \varepsilon\frac{\partial^2}{\partial t^2}\boldsymbol{E}(\boldsymbol{r},t) + \sigma\frac{\partial}{\partial t}\boldsymbol{E}(\boldsymbol{r},t) + \frac{\partial}{\partial t}\boldsymbol{J}(\boldsymbol{r},t) = 0 \tag{19.20}$$

Possible boundary conditions on $\partial\Omega$ include:

$$\hat{\boldsymbol{n}}_{\partial\Omega} \times \left[\frac{1}{\mu}\nabla \times \boldsymbol{E}(\boldsymbol{r},t)\right] + Y_{\partial\Omega}\,\hat{\boldsymbol{n}}_{\partial\Omega} \times \left[\hat{\boldsymbol{n}}_{\partial\Omega} \times \frac{\partial}{\partial t}\boldsymbol{E}(\boldsymbol{r},t)\right] = 0 \tag{19.21}$$

$$\hat{\boldsymbol{n}}_{\partial\Omega} \times \boldsymbol{E}(\boldsymbol{r},t) = 0 \tag{19.22}$$

$$\hat{\boldsymbol{n}}_{\partial\Omega} \times [\nabla \times \boldsymbol{E}(\boldsymbol{r},t)] = 0 \tag{19.23}$$

$$\hat{\boldsymbol{n}}_{\partial\Omega} \times \boldsymbol{E}(\boldsymbol{r},t) = \boldsymbol{M}_{\partial\Omega}(\boldsymbol{r},t) \tag{19.24}$$

$$\boldsymbol{E}(\boldsymbol{r},0) = \boldsymbol{E}_0(\boldsymbol{r}) \tag{19.25}$$

along with a suitable initial condition on the time-derivative of the electric field. In (19.21), we note that $Y_{\partial\Omega}$ represents a surface admittance. Defining $Y_{\partial\Omega} = \sqrt{\varepsilon/\mu}$ gives rise to a first-order radiation boundary condition on $\partial\Omega$ (the Silver-Müller condition), which is equivalent to the first-order Mur ABC commonly used with FDTD formulations [10, 38]. Thus, (19.21) can be used to self-consistently terminate the volume Ω. Alternatively, the Dirichlet condition (19.24) can be used to connect a FE domain to an FDTD domain via a tangential electric field on the domain interface. Note that boundary condition (19.22) is appropriate for a PEC, whereas (19.23) is appropriate for a PMC.

The FE representation of (19.20) to (19.25) is again based on the construction of a suitable variational formulation. Since the defined wave equation is only based on the electric field, a weak solution is sought with $\boldsymbol{E} \in \boldsymbol{W}$, where $\boldsymbol{W}$ denotes the space of *curl*-conforming functions defined by $\boldsymbol{W} = \boldsymbol{H}(\text{curl}; \Omega) = \{\boldsymbol{w} \in [L^2(\Omega)]^3,\ \nabla \times \boldsymbol{w} \in [L^2(\Omega)]^3\}$ such that

$$\int_\Omega \left[\frac{1}{\mu}(\nabla \times \boldsymbol{w}) \cdot (\nabla \times \boldsymbol{E}) + \varepsilon \boldsymbol{w} \cdot \frac{\partial^2}{\partial t^2}\boldsymbol{E} + \sigma \boldsymbol{w} \cdot \frac{\partial}{\partial t}\boldsymbol{E}\right] d\Omega +$$

$$\int_{\partial\Omega} \hat{\boldsymbol{n}}_{\partial\Omega} \times \left(\frac{1}{\mu}\nabla \times \boldsymbol{E}\right) \cdot \boldsymbol{w}\, \partial\Omega + \int_\Omega \left(\boldsymbol{w} \cdot \frac{\partial}{\partial t}\boldsymbol{J}\right) d\Omega = 0 \qquad \forall \boldsymbol{w} \in \boldsymbol{W} \tag{19.26}$$

along with the additional boundary conditions (19.21) to (19.25) as necessary.

Similar to Section 19.2.1, the electric field is expanded in terms of vector edge basis functions, $\boldsymbol{w} \subset \boldsymbol{W}$ (formally, $\boldsymbol{w}^h \subset \boldsymbol{W}^h$, $\boldsymbol{E}^h \subset \boldsymbol{W}^h$) such that

$$\boldsymbol{E}(\boldsymbol{r},t) = \sum_i E_i(t)\boldsymbol{w}_i(\boldsymbol{r}) \tag{19.27}$$

Substituting (19.27) into (19.26) leads to the following semidiscrete matrix equation:

$$[M]\frac{d^2}{dt^2}\{E\} + [K]\frac{d}{dt}\{E\} + [S]\{E\} + \{f\} = \{0\} \tag{19.28}$$

where $\{E\} = [E_1(t), E_2(t), \ldots, E_{N_h}(t)]^T$ and

$$M_{ij} = \int_\Omega \varepsilon \, \boldsymbol{w}_i \cdot \boldsymbol{w}_j \, d\Omega \tag{19.29}$$

$$K_{ij} = \int_\Omega \sigma \, \boldsymbol{w}_i \cdot \boldsymbol{w}_j \, d\Omega \tag{19.30}$$

$$S_{ij} = \int_\Omega \frac{1}{\mu}(\nabla \times \boldsymbol{w}_i) \cdot (\nabla \times \boldsymbol{w}_j) \, d\Omega \tag{19.31}$$

$$f_i = \int_\Omega \left(\boldsymbol{w}_i \cdot \frac{\partial}{\partial t} \boldsymbol{J} \right) d\Omega + \int_{\partial\Omega} \left[\hat{\boldsymbol{n}}_{\partial\Omega} \times \left(\frac{1}{\mu} \nabla \times \boldsymbol{E} \right) \right] \cdot \boldsymbol{w}_i \, \partial\Omega \tag{19.32}$$

Dirichlet boundary conditions such as (19.24) can also be directly included into (19.32) [10]. In addition, if the third-kind boundary condition given by (19.21) is applicable, then (19.32) becomes

$$f_i = \int_\Omega \left(\boldsymbol{w}_i \cdot \frac{\partial}{\partial t} \boldsymbol{J} \right) d\Omega + \int_{\partial\Omega} Y_{\partial\Omega} \left(\hat{\boldsymbol{n}}_{\partial\Omega} \times \boldsymbol{w}_i \right) \cdot \left(\hat{\boldsymbol{n}}_{\partial\Omega} \times \frac{\partial}{\partial t} \boldsymbol{E} \right) \partial\Omega \tag{19.33}$$

Further, if, for example $Y_{\partial\Omega} = \sqrt{\varepsilon/\mu}$, then the surface integral in (19.33) effectively provides a contribution to the dE/dt term in (19.28), which physically corresponds to a loss mechanism (the first-order radiation boundary condition previously noted). Contributions to other terms will occur for alternative definitions of $Y_{\partial\Omega}$. For example, including frequency dependence in $Y_{\partial\Omega}$ may lead to a different time dependency for (19.21), with the specific form being obtained through Fourier transformations [39]. The use of boundary integrals to terminate the FE domain is also possible through a suitable definition of (19.21) [10]. Thus, boundary conditions of the third kind provide for a rich set of conditions.

The fully discrete form for (19.28) can be obtained by a finite-difference representation of the time derivatives. The Newmark scheme is particularly useful for this purpose [4]. Specifically:

$$\frac{d^2}{dt^2}\{E\}\Big|^n \cong \frac{1}{(\Delta t)^2}\left(\{E\}\Big|^{n+1} - 2\{E\}\Big|^n + \{E\}\Big|^{n-1}\right) \tag{19.34}$$

$$\frac{d}{dt}\{E\}\Big|^n \cong \frac{1}{\Delta t}\left(\nu\{E\}\Big|^{n+1} + (1-2\nu)\{E\}\Big|^n - (1-\nu)\{E\}\Big|^{n-1}\right) \tag{19.35}$$

$$\{E\}\Big|^n \cong \theta\{E\}\Big|^{n+1} + (0.5 - 2\theta + \nu)\{E\}\Big|^n + (0.5 + \theta - \nu)\{E\}\Big|^{n-1} \tag{19.36}$$

Setting $\nu = 0.5$ and $\theta = 0.25$ leads to the following fully discrete representation for (19.28):

$$\left([M]+\frac{\Delta t}{2}[K]+\frac{(\Delta t)^2}{4}[S]\right)\{E\}\Big|^{n+1} = 2\left([M]-\frac{(\Delta t)^2}{4}[S]\right)\{E\}\Big|^{n} -$$
$$\left([M]-\frac{\Delta t}{2}[K]+\frac{(\Delta t)^2}{4}[S]\right)\{E\}\Big|^{n-1} - \frac{(\Delta t)^2}{4}\left(\{f\}\Big|^{n+1}+2\{f\}\Big|^{n}+\{f\}\Big|^{n-1}\right) \tag{19.37}$$

Several authors have shown that the Newmark-Beta scheme used to obtain (19.37) is an unconditionally stable time-integration scheme [27, 40]. For convenience and completeness, this is shown here by following the discussion by Edelvik [23].

First, we show that mass matrix $[M]$ is symmetric and positive definite, and that mass matrix $[K]$ and stiffness matrix $[S]$ are symmetric and positive semidefinite. Symmetry is easily seen from (19.29) to (19.31). We note that the components of $[M]$ are given by

$$M_{ij} = \int_\Omega \varepsilon\, \boldsymbol{w}_i \cdot \boldsymbol{w}_j \, d\Omega \equiv (\boldsymbol{w}_i, \boldsymbol{w}_j) \tag{19.38}$$

If $\left\{\boldsymbol{w}_i^h\right\}_{i=1,2,\ldots,N^h}$ is a basis for $\boldsymbol{W}^h$, then $\boldsymbol{E}^h = \sum_{i=1}^{N^h} E_i(t)\boldsymbol{w}_i^h(\boldsymbol{r})$. Thus,

$$\{E\}^{\mathrm{T}}[M]\{E\} = \sum_{i,j=1}^{N^h} E_i\, M_{ij}\, E_j = \sum_{i,j=1}^{N^h} E_i\left(\boldsymbol{w}_i^h, \boldsymbol{w}_j^h\right)E_j$$
$$= \left(\sum_{i=1}^{N^h} E_i \boldsymbol{w}_i^h,\ \sum_{j=1}^{N^h} E_j \boldsymbol{w}_j^h\right) = \left(\boldsymbol{E}^h, \boldsymbol{E}^h\right) \geq \{0\} \tag{19.39}$$

with equality to zero only if $\boldsymbol{E}^h \equiv \boldsymbol{0}$ (i.e., only if $\{E\} \equiv \{0\}$). Hence, $[M]$ is positive definite. $[K]$ is also positive definite for $\sigma > 0$, but, in general, positive semidefinite since σ is permitted to be equal to zero. Using a similar procedure, $[S]$ is shown to be positive semidefinite since the case $\boldsymbol{E} = -\nabla\varphi$ implies $\nabla \times \mathbf{E} = \mathbf{0}$.

This result can now be used to establish that the Newmark-Beta scheme is unconditionally stable. For simplicity, assume there are no losses such that $\sigma = 0$. Also assume $\{f\} = \{0\}$. Multiplying (19.37) by $[M]^{-1}$ gives

$$\left([I] + 0.25(\Delta t)^2[M]^{-1}[S]\right)\{E\}\Big|^{n+1} = 2\left([I] - 0.25(\Delta t)^2[M]^{-1}[S]\right)\{E\}\Big|^{n}$$
$$- \left([I] + 0.25(\Delta t)^2[M]^{-1}[S]\right)\{E\}\Big|^{n-1} \tag{19.40}$$

Since $[M]$ is symmetric and positive definite from the previous result, it can be factored as $[M] = [\tilde{R}]^{\mathrm{T}}[\tilde{R}]$. Hence,

$$0.25(\Delta t)^2[M]^{-1}[S] = 0.25(\Delta t)^2[\tilde{R}]^{-1}[\tilde{R}]^{-\mathrm{T}}[S][\tilde{R}]^{-1}[\tilde{R}] \equiv [\tilde{R}]^{-1}[\tilde{P}][\tilde{R}] \tag{19.41}$$

where

$$[\tilde{P}] = 0.25(\Delta t)^2[\tilde{R}]^{-\mathrm{T}}[S][\tilde{R}]^{-1} \tag{19.42}$$

Since $[S]$ is symmetric and positive semidefinite, $[\tilde{P}]$ also possesses these properties. Thus, $[\tilde{P}]$ can be diagonalized by an orthogonal matrix $[Q]$ consisting of the eigenvectors of $[\tilde{P}]$, such that $[Q]^{\mathrm{T}}[\tilde{P}][Q] = [\Lambda]$, where $[\Lambda]$ represents the nonnegative eigenvalues of $[\tilde{P}]$. Since $[Q]$ also diagonalizes $[I]+[\tilde{P}]$ and $[I]-[\tilde{P}]$, (19.40) can be written as

$$\big([I]+[\Lambda]\big)\{z\}\big|^{n+1} = 2\big([I]-[\Lambda]\big)\{z\}\big|^{n} - \big([I]+[\Lambda]\big)\{z\}\big|^{n-1} \tag{19.43}$$

where $\{z\}|^n \equiv [Q]^{\mathrm{T}}[\tilde{R}]\{E\}|^n$. Because of the diagonalization of (19.43), it is sufficient to consider the scalar equation

$$(1+\tilde{\lambda})z_i\big|^{n+1} = 2(1-\tilde{\lambda})z_i\big|^{n} - (1+\tilde{\lambda})z_i\big|^{n-1} \tag{19.44}$$

The corresponding characteristic equation is

$$(1+\tilde{\lambda})\tilde{\mu}^2 - 2(1-\tilde{\lambda})\tilde{\mu} + (1+\tilde{\lambda}) = 0 \tag{19.45}$$

with solutions

$$\tilde{\mu} = \frac{1-\tilde{\lambda}}{1+\tilde{\lambda}} \pm \sqrt{\left(\frac{1-\tilde{\lambda}}{1+\tilde{\lambda}}\right)^2 - 1} \tag{19.46}$$

By defining $\sin\phi \equiv (1-\tilde{\lambda})/(1+\tilde{\lambda})$, (19.46) can be written as

$$\tilde{\mu} = \sin\phi \pm j\cos\phi \tag{19.47}$$

Since these eigenvalues have unit magnitude $|\tilde{\mu}| = 1$, the Newmark-Beta scheme is unconditionally stable for $\sigma = 0$. Similarly, the case $\sigma > 0$ leads to roots that are strictly less than one in magnitude. Consequently, as the solution is marched through time, the basic algorithm is stable independent of the time-step.

We note that, although numerical stability is not dependent upon the choice of the time-step, the accuracy *is* dependent upon the time-step [40], as is the number of required iterations. For applications that involve a wide range of element sizes (greater than 100:1), the use of advanced preconditioners [41] or simply a modest reduction in the time-step can substantially reduce the number of required iterations for a given residual error. Reducing the time-step improves the mass-matrix dominance in the matrix solution.

19.2.3 Equivalences Between Finite Elements and FDTD

Many equivalences exist between the described FE formulations and the traditional FDTD method. A few of these are stated here. Additional details can be found in the references.

- If linear, rectangular hexahedral elements are adopted, and the spatial integrations are evaluated using the trapezoidal rule, then the FDTD algorithm is recovered *identically* from the coupled curl-equation formulation described by (19.18) and (19.19) [32, 33].

- The coupled curl-equation formulation discussed in Section 19.2.1, based on both edge elements and facet elements, is equivalent to the wave-equation formulation of Section 19.2.2 that is based only on edge elements [42]. This equivalence holds provided leapfrog time integration is used for the coupled-curl equations [see (19.18) and (19.19)] and central time-differencing is used for the wave equation. Note that the central time-difference form for the fully discrete wave equation is obtained by setting $\theta = 0$ in (19.36), which leads to

$$\left([M] + 0.5\Delta t[K]\right)\{E\}\big|^{n+1} = \left(2[M] - (\Delta t)^2[S]\right)\{E\}\big|^{n} - \left([M] - 0.5\Delta t[K]\right)\{E\}\big|^{n-1} - (\Delta t)^2\{f\}\big|^{n} \quad (19.48)$$

- On a rectangular hexahedral grid, using edge elements and trapezoidal integration renders the wave-equation representation equivalent to the basic FDTD method, provided central differencing is used for the time integration for the wave equation (19.48). This conclusion follows from the previous results.

These equivalences naturally lead to a symmetric, stable interface between FE methods and FDTD, which is described in greater detail in Section 19.4.

19.3 TETRAHEDRAL, HEXAHEDRAL (BRICK), AND PYRAMIDAL ZEROTH-ORDER EDGE AND FACET ELEMENTS

This section provides a brief overview of the construction of edge and facet basis functions for tetrahedral, hexahedral, and pyramidal finite elements. Only the lowest order (linear) elements are discussed. These basis functions are often referred to as zeroth-order basis functions or first-order mixed basis functions [12]. Higher-order representations are available [43, 44]. However, at the interface with an FDTD grid, the lowest-order linear basis functions are currently required to preserve the symmetry between the FE and FD regions. The future use of hierarchical basis functions [45] may permit mixing the order of the basis functions between these two regions.

It is convenient to construct the mass and stiffness matrices by summing contributions from the individual elements in the FE grid. Therefore, the integrals $\int_{\Omega} \cdot d\Omega$ over the entire computational domain are divided into a sum of integrals $\int_{V_1} \cdot d\Omega + \int_{V_2} \cdot d\Omega + \ldots + \int_{V_{N_e}} \cdot d\Omega$, where $V_1, V_2, \ldots, V_{N_e}$ correspond to the volume occupied by the individual elements. Using FE terminology, this is referred to as the "assembling procedure."

The elemental mass and stiffness matrices on the various elemental shapes are often conveniently computed based on a so-called reference element, where the expressions for the basis functions are relatively simple. The coordinate system $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$ of the reference element is related to the physical element and its coordinate system $\boldsymbol{r} = (x, y, z)$ by a mapping $\boldsymbol{r} = \boldsymbol{r}_1 S_1(\boldsymbol{\xi}) + \boldsymbol{r}_2 S_2(\boldsymbol{\xi}) + \ldots + \boldsymbol{r}_{N_n} S_{N_n}(\boldsymbol{\xi})$, where $\boldsymbol{r}_n$ is the nth node of the physical element, $S_n(\boldsymbol{\xi})$ is an interpolatory shape function associated with the nth node, and N_n denotes the total number of nodes on the element. The Jacobian matrix $[J]$ is then easily constructed given this mapping, and a variable substitution in the integral over the element involves the infinitesimal volume $dx\,dy\,dz = \det[J]\, d\xi_1 d\xi_2 d\xi_3$.

The vector edge basis functions can be expressed in terms of gradients of scalar functions. Consequently, they can be represented in the global domain $\boldsymbol{r}$ by $\boldsymbol{w}_i(\boldsymbol{r}) = [J]^{-1}\boldsymbol{\Omega}_i$, where $\boldsymbol{\Omega}_i$ denotes the edge-element basis function expressed on the reference element [42]. The curl of the edge elements in the global domain is given by $\nabla \times \boldsymbol{w}_i(\boldsymbol{r}) = ([J]^{\mathrm{T}}/\det[J])\nabla \times \boldsymbol{\Omega}_i$ [42]. The facet elements in the global domain $\boldsymbol{r}$ are similarly obtained by forming $\boldsymbol{w}_{f_i}(\boldsymbol{r}) = ([J]^{\mathrm{T}}/\det[J])\boldsymbol{\Lambda}_i$, where $\boldsymbol{\Lambda}_i$ is the facet element basis function on the reference element. The evaluation of the mass and stiffness matrices is completed by using Gaussian quadrature over the elemental shape.

Note that it is also possible to use a direct analytical approach to evaluate the FE matrices for certain elemental shapes. To demonstrate both approaches, the analytical approach will be used for the tetrahedral element discussed in Section 19.3.1, while the reference-element transformation will be used for the hexahedral and pyramidal elements discussed in Sections 19.3.2 and 19.3.3, respectively.

19.3.1 Tetrahedral Finite Elements

The FE matrices for the linear tetrahedron can be constructed using either the reference-element approach in a mapped coordinate system, or a direct approach in the global coordinate system. The latter approach is adopted in this section. A tetrahedron in the global coordinate system $\boldsymbol{r} = (x, y, z)$ is shown in Fig. 19.3.

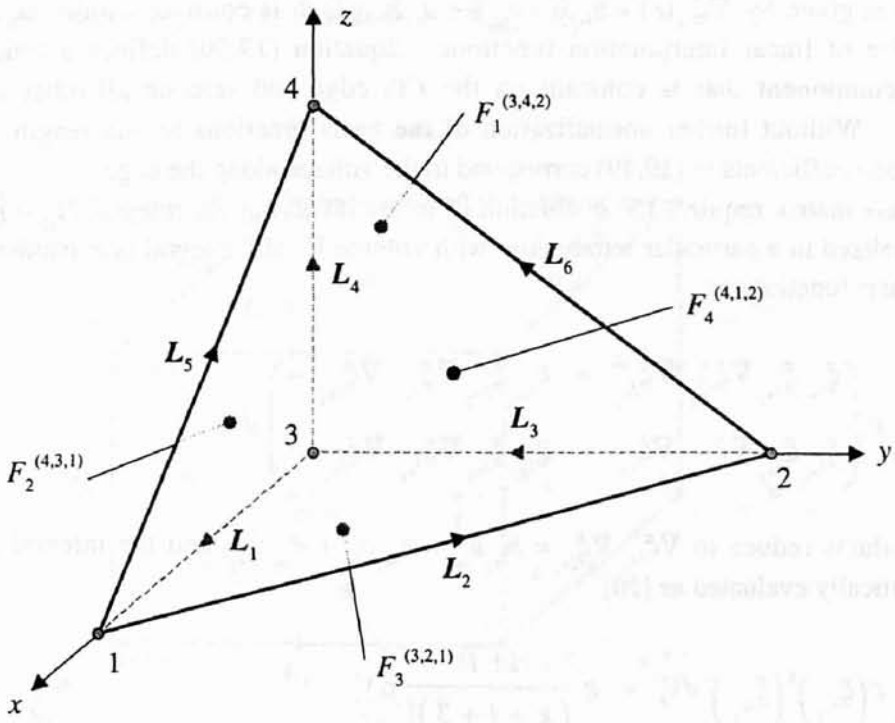


Fig. 19.3 Tetrahedral element. Edges and faces are denoted by L_i and F_i, respectively. The additional superscript on the face index corresponds to the nodes that define the vertices of the particular face.

Within the cell of Fig. 19.3, the electric field is expanded in linear edge elements $\boldsymbol{w}_i(\boldsymbol{r})$ as

$$\boldsymbol{E}(\boldsymbol{r},t) = \sum_{i=1}^{N_{\text{edges}}} E_i(t)\boldsymbol{w}_i(\boldsymbol{r}) \tag{19.49}$$

where $N_{\text{edges}} = 6$ denotes the number of edges on the tetrahedron. Using a general referencing scheme, edge i is defined to point from node n_1 to node n_2, such that the basis function for the i'th edge is defined by

$$\boldsymbol{w}_i(\boldsymbol{r}) = \xi_{i_{n_1}}(\boldsymbol{r})\,\nabla\xi_{i_{n_2}}(\boldsymbol{r}) - \xi_{i_{n_2}}(\boldsymbol{r})\nabla\xi_{i_{n_1}}(\boldsymbol{r}) \tag{19.50}$$

Edge basis functions preserve tangential continuity across element boundaries. The Lagrange interpolation polynomials are defined at the vertices (nodes) of the element, and are denoted by $\xi_{i_{n_1}}$ and $\xi_{i_{n_2}}$ for tetrahedron nodes n_1 and n_2, respectively. Using linear interpolation, the polynomials have the following form for node n_1 on edge i:

$$\xi_{i_{n_1}}(\boldsymbol{r}) = a_{i_{n_1}} + b_{i_{n_1}}x + c_{i_{n_1}}y + d_{i_{n_1}}z \tag{19.51}$$

where $\boldsymbol{r} = (x, y, z)$ denotes spatial position within the element, and $a_{i_{n_1}}$, $b_{i_{n_1}}$, $c_{i_{n_1}}$, and $d_{i_{n_1}}$ denote coefficients. The coefficients are dependent upon the spatial orientation of the element and can be determined by either numerical or analytical means [10]. Note that the gradient of the polynomial is given by $\nabla\xi_{i_{n_1}}(\boldsymbol{r}) = b_{i_{n_1}}\hat{x} + c_{i_{n_1}}\hat{y} + d_{i_{n_1}}\hat{z}$, which is constant within the element due to the choice of linear interpolation functions. Equation (19.50) defines a function with a tangential component that is constant on the i'th edge and zero on all other edges of the tetrahedron. Without further normalization of the basis functions by the length of the edge, the expansion coefficients in (19.49) correspond to the voltage along the edge.

The mass matrix requires the evaluation of terms involving the integral $M_{ij} = \int_\Omega \varepsilon \boldsymbol{w}_i \cdot \boldsymbol{w}_j d\Omega$. When specialized to a particular tetrahedron with volume V_e, the integral is expanded in terms of the linear basis functions as

$$\int_{V_e} \varepsilon \begin{pmatrix} \xi_{i_{n_1}}\xi_{j_{n_1}}\nabla\xi_{i_{n_2}}\cdot\nabla\xi_{j_{n_2}} + \xi_{i_{n_2}}\xi_{j_{n_2}}\nabla\xi_{i_{n_1}}\cdot\nabla\xi_{j_{n_1}} - \\ \xi_{i_{n_1}}\xi_{j_{n_2}}\nabla\xi_{i_{n_2}}\cdot\nabla\xi_{j_{n_1}} - \xi_{i_{n_2}}\xi_{j_{n_1}}\nabla\xi_{i_{n_1}}\cdot\nabla\xi_{j_{n_2}} \end{pmatrix} dV_e \tag{19.52}$$

The dot products reduce to $\nabla\xi_{i_{n_1}}\cdot\nabla\xi_{j_{n_2}} = b_{i_{n_1}}b_{j_{n_2}} + c_{i_{n_1}}c_{j_{n_2}} + d_{i_{n_1}}d_{j_{n_2}}$, and the integral $\int_{V_e}\varepsilon\xi_{i_{n_1}}\xi_{j_{n_1}}dV_e$ can be analytically evaluated as [10]

$$\int_{V_e}\varepsilon\left(\xi_{i_{n_1}}\right)^k\left(\xi_{j_{n_1}}\right)^l dV_e = \varepsilon\frac{k!\;l!}{(k+l+3)!}6V_e \tag{19.53}$$

Note that for the case $i = j$, $k = 2$, and $l = 0$, the integral evaluates to $V_e/10$, whereas when $i \neq j$ and $k = l = 1$, the integral evaluates to $V_e/20$. The evaluation of the stiffness matrix elements (19.31) is similar.

The magnetic field is expanded in linear facet elements $\boldsymbol{w}_{f_l}(\boldsymbol{r})$ as

$$\boldsymbol{B}(\boldsymbol{r},t) = \sum_{l=1}^{N_{\text{faces}}} B_l(t)\boldsymbol{w}_{f_l}(\boldsymbol{r}) \tag{19.54}$$

where $N_{\text{faces}} = 4$ denotes the number of tetrahedron faces. Facet basis functions are represented in terms of the interpolation functions by the following relationship for tetrahedral elements [46]:

$$\boldsymbol{w}_{f_l}(\boldsymbol{r}) = 2\left(\xi_{l_{n_1}} \nabla\xi_{l_{n_2}} \times \nabla\xi_{l_{n_3}} + \xi_{l_{n_2}} \nabla\xi_{l_{n_3}} \times \nabla\xi_{l_{n_1}} + \xi_{l_{n_3}} \nabla\xi_{l_{n_1}} \times \nabla\xi_{l_{n_2}} \right) \tag{19.55}$$

Equation (19.55) is based on the definition that facet l consists of nodes (n_1, n_2, n_3). As previously noted, facet basis functions preserve normal continuity across element faces. Equation (19.55) defines a function with normal component that is constant on the lth face of the tetrahedron and zero on all other faces. Without additional normalization of these basis functions by the area of the face, the expansion coefficients in (19.54) correspond to the magnetic flux passing through the face. The Lagrange interpolation functions are given by (19.51), and the required matrix entries are evaluated by expressions similar to (19.52) and (19.53).

19.3.2 Hexahedral (Brick) Finite Elements

It is often convenient to define a reference hexahedral (brick) element to which arbitrary hexahedral elements are mapped. Fig. 19.4 illustrates our reference brick element, which consists of twelve edges ($N_{\text{edges}} = 12$) and six faces ($N_{\text{faces}} = 6$).

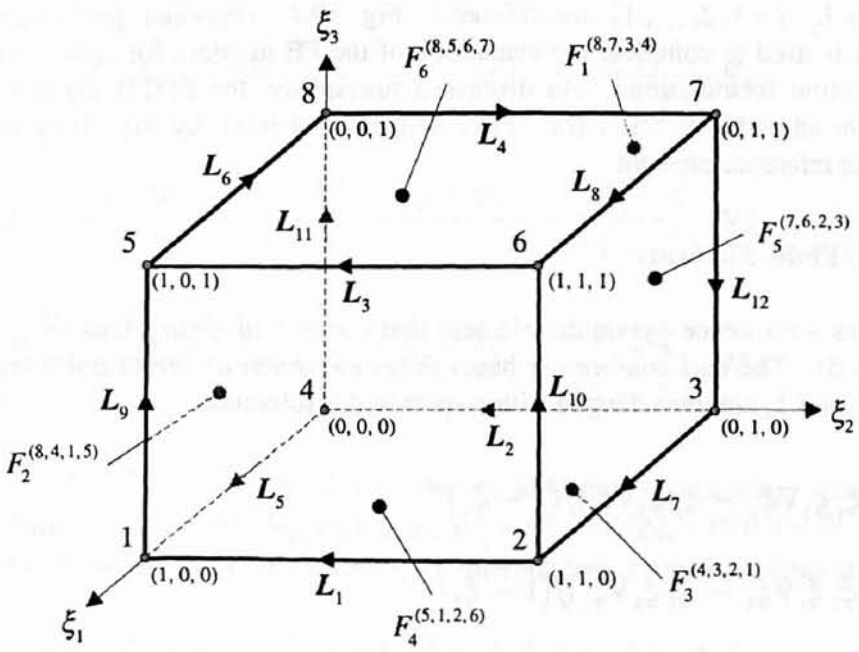


Fig. 19.4 Reference hexahedral element. Edges and faces are denoted by L_i and F_l, respectively. The additional superscript on the face index corresponds to the nodes that define the vertices of the particular face.

Referring to Fig. 19.4, the curl-conforming bases (edge elements) of zeroth order are given by [43]

$$\boldsymbol{\Omega}_1 = \xi_1 \xi_6 \nabla\xi_5 \; ; \quad \boldsymbol{\Omega}_2 = \xi_4 \xi_6 \nabla\xi_5 \; ; \quad \boldsymbol{\Omega}_3 = \xi_1 \xi_3 \nabla\xi_5 \; ; \quad \boldsymbol{\Omega}_4 = \xi_3 \xi_4 \nabla\xi_2 \tag{19.56}$$

$$\boldsymbol{\Omega}_5 = \xi_5 \xi_6 \nabla\xi_1 \; ; \quad \boldsymbol{\Omega}_6 = \xi_3 \xi_5 \nabla\xi_4 \; ; \quad \boldsymbol{\Omega}_7 = \xi_2 \xi_6 \nabla\xi_1 \; ; \quad \boldsymbol{\Omega}_8 = \xi_2 \xi_3 \nabla\xi_1 \tag{19.57}$$

$$\boldsymbol{\Omega}_9 = \xi_1 \xi_5 \nabla\xi_3 \; ; \quad \boldsymbol{\Omega}_{10} = \xi_1 \xi_2 \nabla\xi_3 \; ; \quad \boldsymbol{\Omega}_{11} = \xi_4 \xi_5 \nabla\xi_3 \; ; \quad \boldsymbol{\Omega}_{12} = \xi_2 \xi_4 \nabla\xi_6 \tag{19.58}$$

where ξ_1, ξ_2, and ξ_3 denote independent coordinates, and $\xi_4 = 1 - \xi_1$, $\xi_5 = 1 - \xi_2$, and $\xi_6 = 1 - \xi_3$ denote dependent coordinates. Note that, based on the defined reference element, $\nabla\xi_1 = (1,0,0)$; $\nabla\xi_2 = (0,1,0)$; $\nabla\xi_3 = (0,0,1)$; $\nabla\xi_4 = (-1,0,0)$; $\nabla\xi_5 = (0,-1,0)$; and $\nabla\xi_6 = (0,0,-1)$.

At the vertices of the hexahedron, the interpolatory shape functions can be written as

$$\{S_{\text{hex}}\} = \left[\xi_1\xi_5\xi_6,\ \xi_1\xi_2\xi_6,\ \xi_4\xi_2\xi_6,\ \xi_4\xi_5\xi_6,\ \xi_1\xi_5\xi_3,\ \xi_1\xi_2\xi_3,\ \xi_4\xi_2\xi_3,\ \xi_4\xi_5\xi_3\right]^{\text{T}} \tag{19.59}$$

For the six faces shown in Fig. 19.4 (F_l, $l = 1, 2, \ldots, 6$), the divergence-conforming bases (facet elements) of zeroth order are given, respectively, by

$$\boldsymbol{\Lambda}_1 = (1-\xi_1)\boldsymbol{L}_5; \quad \boldsymbol{\Lambda}_2 = (1-\xi_2)\boldsymbol{L}_4; \quad \boldsymbol{\Lambda}_3 = (1-\xi_3)\boldsymbol{L}_{11} \tag{19.60}$$

$$\boldsymbol{\Lambda}_4 = \xi_1\boldsymbol{L}_6; \quad \boldsymbol{\Lambda}_5 = \xi_2\boldsymbol{L}_2; \quad \boldsymbol{\Lambda}_6 = \xi_3\boldsymbol{L}_{12} \tag{19.61}$$

where the vectors $\boldsymbol{L}_i$, $i = 1, 2, \ldots, 12$ are defined in Fig. 19.4. Gaussian quadrature over the reference element is used to complete the evaluation of the FE matrices for either the coupled-curl or wave-equation formulations. As discussed previously, the FDTD algorithm can be recovered for right-angled hexahedra (i.e., brick-shaped elements), by exploiting trapezoidal integration over the reference element.

19.3.3 Pyramidal Finite Elements

Fig. 19.5 illustrates a reference pyramidal element that consists of eight edges ($N_{\text{edges}} = 8$) and five faces ($N_{\text{faces}} = 5$). The curl-conforming bases (edge elements) of zeroth order are given in [44], noting that ξ_3 and ξ_5 are interchanged with respect to this reference:

$$\boldsymbol{\Omega}_1 = (\xi_1 \xi_2 \nabla\xi_5 - \xi_2 \xi_5 \nabla\xi_1)/(1 - \xi_3) \tag{19.62}$$

$$\boldsymbol{\Omega}_2 = (\xi_2 \xi_5 \nabla\xi_4 - \xi_5 \xi_4 \nabla\xi_2)/(1 - \xi_3) \tag{19.63}$$

$$\boldsymbol{\Omega}_3 = (\xi_5 \xi_4 \nabla\xi_1 - \xi_4 \xi_1 \nabla\xi_5)/(1 - \xi_3) \tag{19.64}$$

$$\boldsymbol{\Omega}_4 = (\xi_4 \xi_1 \nabla\xi_2 - \xi_1 \xi_2 \nabla\xi_4)/(1 - \xi_3) \tag{19.65}$$

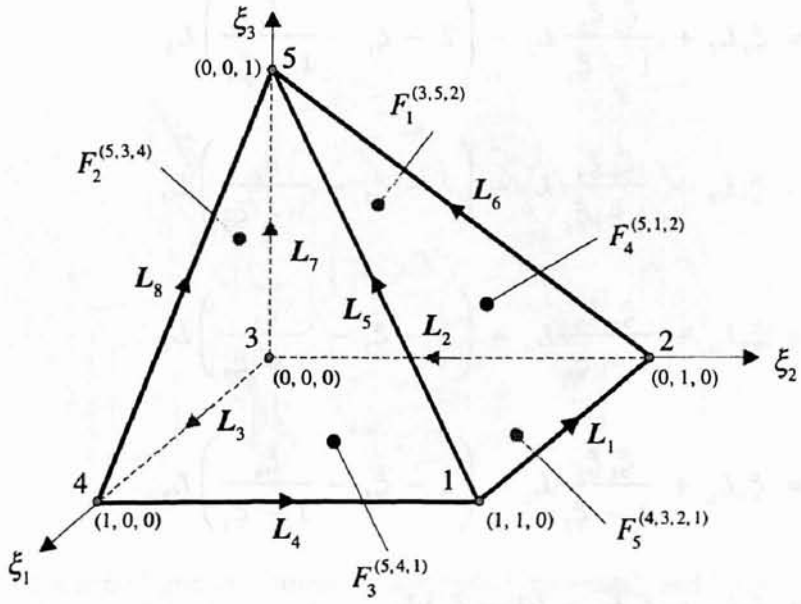


Fig. 19.5 Reference pyramidal element. Edges and faces are denoted by L_i and F_l, respectively. The additional superscript on the face index corresponds to the nodes that define the vertices of the particular face.

$$\boldsymbol{\Omega}_5 = \frac{\xi_1\xi_2\nabla\xi_3 - \xi_2\xi_3\nabla\xi_1 - \xi_1\xi_3\nabla\xi_2}{1-\xi_3} - \frac{\xi_1\xi_2\xi_3}{(1-\xi_3)^2}\nabla\xi_3 \tag{19.66}$$

$$\boldsymbol{\Omega}_6 = \frac{\xi_2\xi_5\nabla\xi_3 - \xi_5\xi_3\nabla\xi_2 - \xi_2\xi_3\nabla\xi_5}{1-\xi_3} - \frac{\xi_2\xi_5\xi_3}{(1-\xi_3)^2}\nabla\xi_3 \tag{19.67}$$

$$\boldsymbol{\Omega}_7 = \frac{\xi_5\xi_4\nabla\xi_3 - \xi_4\xi_3\nabla\xi_5 - \xi_5\xi_3\nabla\xi_4}{1-\xi_3} - \frac{\xi_5\xi_4\xi_3}{(1-\xi_3)^2}\nabla\xi_3 \tag{19.68}$$

$$\boldsymbol{\Omega}_8 = \frac{\xi_4\xi_1\nabla\xi_3 - \xi_1\xi_3\nabla\xi_4 - \xi_4\xi_3\nabla\xi_1}{1-\xi_3} - \frac{\xi_4\xi_1\xi_3}{(1-\xi_3)^2}\nabla\xi_3 \tag{19.69}$$

where $\xi_4 = 1-\xi_2-\xi_3$ and $\xi_5 = 1-\xi_1-\xi_3$ denote dependent coordinates. Based on the defined reference element in Fig. 19.5, $\nabla\xi_1 = (1,0,0)$, $\nabla\xi_2 = (0,1,0)$, $\nabla\xi_3 = (0,0,1)$, $\nabla\xi_4 = (0,-1,-1)$, and $\nabla\xi_5 = (-1,0,-1)$. At the pyramid's vertices, the interpolatory shape functions can be written as

$$\{S_{pyr}\} = \left[\xi_1\xi_2/(1-\xi_3),\ \xi_2\xi_5/(1-\xi_3),\ \xi_5\xi_4/(1-\xi_3),\ \xi_4\xi_1/(1-\xi_3),\ \xi_3\right]^T \tag{19.70}$$

For the five faces shown in Fig. 19.5 (F_l, $l = 1, 2, \ldots, 5$), the divergence-conforming bases (facet elements) of zeroth order are given, respectively, by [44]

$$\boldsymbol{\Lambda}_1 = \xi_3 \boldsymbol{L}_7 + \frac{\xi_2 \xi_3}{1 - \xi_3} \boldsymbol{L}_2 - \left(2 - \xi_1 - \frac{\xi_1}{1 - \xi_3} \right) \boldsymbol{L}_3 \tag{19.71}$$

$$\boldsymbol{\Lambda}_2 = \xi_3 \boldsymbol{L}_8 + \frac{\xi_5 \xi_3}{1 - \xi_3} \boldsymbol{L}_3 - \left(2 - \xi_2 - \frac{\xi_2}{1 - \xi_3} \right) \boldsymbol{L}_4 \tag{19.72}$$

$$\boldsymbol{\Lambda}_3 = \xi_3 \boldsymbol{L}_5 + \frac{\xi_4 \xi_3}{1 - \xi_3} \boldsymbol{L}_4 - \left(2 - \xi_5 - \frac{\xi_5}{1 - \xi_3} \right) \boldsymbol{L}_1 \tag{19.73}$$

$$\boldsymbol{\Lambda}_4 = \xi_3 \boldsymbol{L}_6 + \frac{\xi_1 \xi_3}{1 - \xi_3} \boldsymbol{L}_1 - \left(2 - \xi_4 - \frac{\xi_4}{1 - \xi_3} \right) \boldsymbol{L}_2 \tag{19.74}$$

$$\boldsymbol{\Lambda}_5 = \xi_1 \boldsymbol{L}_3 + \xi_2 \boldsymbol{L}_4 - (1 - \xi_3) \boldsymbol{L}_7 \tag{19.75}$$

where the vectors $\boldsymbol{L}_i$, $i = 1, 2, \ldots, 8$ are defined in Fig. 19.5. Gaussian quadrature over the reference pyramid is used to complete the evaluation of the finite-element matrices for either the coupled-curl or wave-equation formulations [30].

19.4 STABLE HYBRID FDTD-FE INTERFACE

19.4.1 Spatial Discretization

The section is based upon [28, 34, 35] which reported a technique to achieve a stable hybrid of a Cartesian FDTD space lattice with a time-domain FE mesh consisting of tetrahedra and pyramids. The tetrahedral elements permit complex boundaries to be modeled in the FE mesh. These are interfaced with the FDTD "bricks" using pyramidal elements. Each pyramid has triangular faces that can be connected to the unstructured tetrahedrons, and a rectangular base that can be fitted to an FDTD cell. A simple example of such a grid is shown in Fig. 19.6, where the different types of elements are spatially separated to show the interior structure of the grid.

Given such a discretization, the E-field is expressed in terms of edge elements. This yields a curl-conforming representation, as discussed in Section 19.2. Galerkin's method is then used for Maxwell's equations, and, as a consequence of reciprocity, the resulting discretized system is represented in terms of symmetric matrices. In addition to the pyramids, the FDTD-FE interface exploits the basic observation that the FDTD method (viewed as a scheme for the second-order wave equation for the E-field) can be derived from this FE formulation by trapezoidal integration (lumping) for the $\nabla \times \mu^{-1} \nabla \times$ and ε operators on the FDTD bricks. Thus, the vector wave equation $\nabla \times \mu^{-1} \nabla \times \boldsymbol{E} = \omega^2 \varepsilon \boldsymbol{E}$ is represented by $[S]\{E\} = \omega^2 [M]\{E\}$, where the mass matrix $[M]$ is positive definite and the stiffness matrix $[S]$ is positive semidefinite. By virtue of $[M]$ and $[S]$, the numerical eigenvalues ω^2 are guaranteed to be real and nonnegative, which makes it possible to formulate a stable time-stepping scheme [28, 34, 35]. Moreover, the reciprocity of the continuous Maxwell's equations is preserved by this discrete representation.

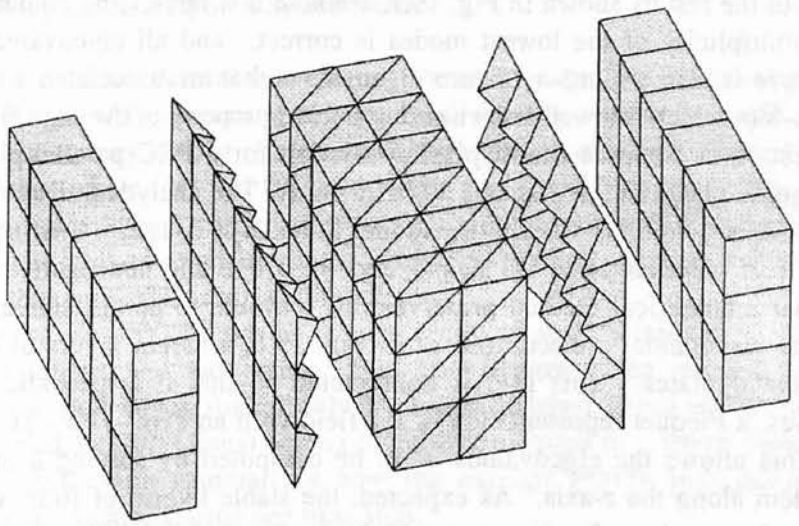


Fig. 19.6 Example of a hybrid grid that combines tetrahedral, pyramidal, and brick-shaped elements. *Source:* Rylander and Bondeson, *J. Comput. Phys.*, 2002, pp. 426–438, © 2002 Elsevier.

The properties of such a spatial discretization are illustrated by calculating the eigenvalues of a cubic cavity having PEC walls of dimension a. The cavity is discretized as in Fig. 19.6, with a layer of tetrahedral cells connected to cubic cells on each side with pyramidal cells. The first 10 normalized cavity eigenvalues $(ka/\pi)^2$ for the numerical model are shown as circles in Fig. 19.7. The analytical eigenvalues $(ka/\pi)^2 = n_x^2 + n_y^2 + n_z^2$ are shown by crosses, where at least two of the mode indices $n_x, n_y, n_z = 0, 1, 2, \ldots$ are nonzero.

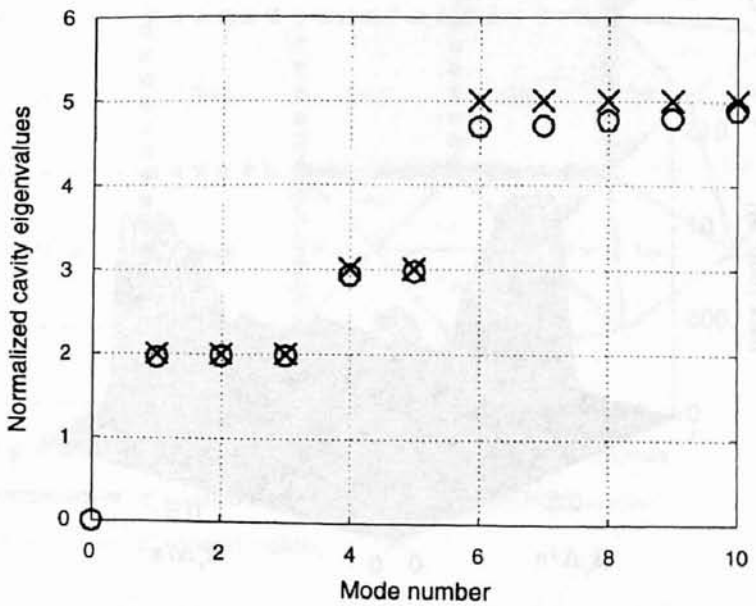


Fig. 19.7 First 10 normalized eigenvalues of a cavity resonator discretized by the hybrid grid of Fig. 19.6: circles – numerical values; crosses – analytical. *Source:* Rylander and Bondeson, *Comput. Phys. Comm.*, 2000, pp. 75–82, © 2000 Elsevier.

With regard to the results shown in Fig. 19.7, we note that there is no contamination of the spectrum, the multiplicity of the lowest modes is correct, and all eigenvalues are real and nonnegative. There is also a number of zero eigenvalues that are associated with electrostatic eigenmodes $\boldsymbol{E} = -\nabla\varphi$, which is a well-known and important property of the edge elements [12].

A similar test is to compute the dispersion relation for a PEC parallel-plate waveguide, where the waveguide plates lie on planes $z = 0$ and $z = d$. The analytical dispersion relation is $k^2 = (\omega/c)^2 = k_x^2 + k_y^2 + (\pi n_z/d)^2$, where the mode index $n_z = 0, 1, 2, \ldots$ is nonzero for TE modes. Since a real wavevector in the xy-plane yields a real and nonnegative eigenvalue ω^2, it is desirable that a numerical method preserves this attribute to permit stable time-stepping. The parallel-plate waveguide is discretized as in Fig. 19.6, wherein layers of cubic cells are parallel to the metal plates. This grid is constructed so that it is periodic in the x- and y-directions. Thus, a Floquet representation of the field with an $\exp(-jk_x x - jk_y y)$ dependence can be used. This allows the eigenvalues ω^2 to be computed by solving a one-dimensional eigenvalue problem along the z-axis. As expected, the stable hybrid of finite differences and finite elements produces only real and nonnegative eigenvalues for this application.

Note that the stable hybrid classifies the *elements*, not the edges, as finite elements or finite differences. Thus, the stable hybrid treats edges on the interface neither as regular FE nor regular FDTD. Although a possibly more intuitive way of constructing the spatial discretization is to use finite differences for *edges* that belong to the bricks and finite elements for all other edges, this edge-based classification gives nonsymmetric matrices and, therefore, stability can no longer be guaranteed. This type of hybrid was applied to the parallel-plate example discussed above. Fig. 19.8 shows the largest imaginary part of the normalized eigenfrequencies $k\Delta$ for $0 \le k_x\Delta \le \pi$ and $0 \le k_y\Delta \le \pi$, where the FDTD cells are cubes of side Δ. Clearly, the approach based on the edges (rather than the elements) yields complex eigenfrequencies ω for real wavenumbers in the xy-plane, and such a scheme is unstable. Instabilities occur for modes that are poorly resolved, and their exponential growth occurs rapidly.

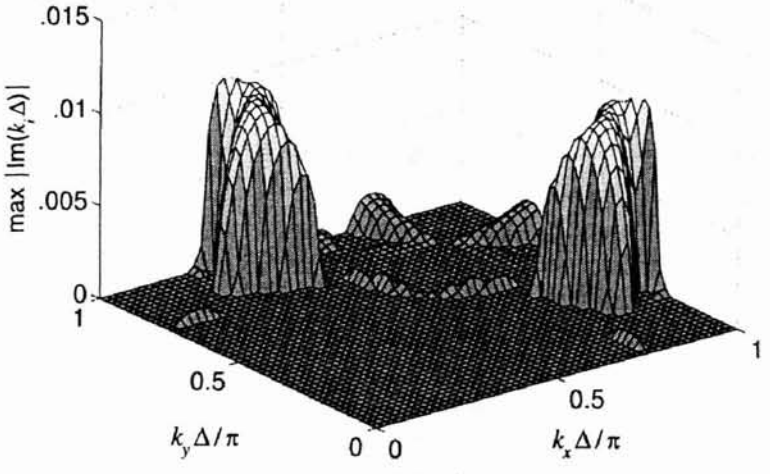


Fig. 19.8 The largest imaginary part of the eigenvalues $k\Delta$ as a function of $k_x\Delta$ and $k_y\Delta$ when finite differences are used for the edges shared by only cubes, and finite elements elsewhere. *Source:* Rylander and Bondeson, *Comput. Phys. Comm.*, 2000, pp. 75–82, © 2000 Elsevier.

19.4.2 Time-Stepping on a Hybrid Space Lattice

Time-stepping on a mixed-element space lattice could, of course, be formulated directly in terms of matrices $[M]$ and $[S]$. However, this is usually computationally inefficient, since large portions of the lattice may consist of structured brick-shaped elements where classic FDTD can be applied, and thus where $[M]$ and $[S]$ need not be stored and processed. This can be exploited by updating the fields in two consecutive substeps. First, every field component that can be treated by FDTD is updated explicitly in the usual Yee manner. Second, every field component not advanced in time by FDTD is updated implicitly.

In this process, the FDTD updates require access to specific tangential E-fields from the unstructured FE lattice, and vice versa. Fig. 19.9 illustrates the relevant geometry in two dimensions. Here, cell edges shared only by bricks (dashed lines) are treated explicitly by FDTD. All other edges (solid lines) belong to either a pyramid or a tetrahedron, and are treated implicitly. The following summarizes how the explicit FDTD and implicit FE schemes interchange field information during one time-step:

1. Field components in the structured part of the lattice are updated explicitly by FDTD. Specifically, $\{H_{\text{FDTD}}\}|^{n+1/2}$ and $\{E_{\text{FDTD}}\}|^{n+1}$ are computed given the previous FDTD values $\{E_{\text{FDTD}}\}|^{n}$ and $\{H_{\text{FDTD}}\}|^{n-1/2}$. Computation of $\{H_{\text{FDTD}}\}|^{n+1/2}$ requires FE values $\{E_{\text{FE}}\}|^{n}$ at the interface of the structured and unstructured parts of the lattice. Cell edges shown by thick solid lines in Fig. 19.9 provide field information from the FE region for FDTD updates.

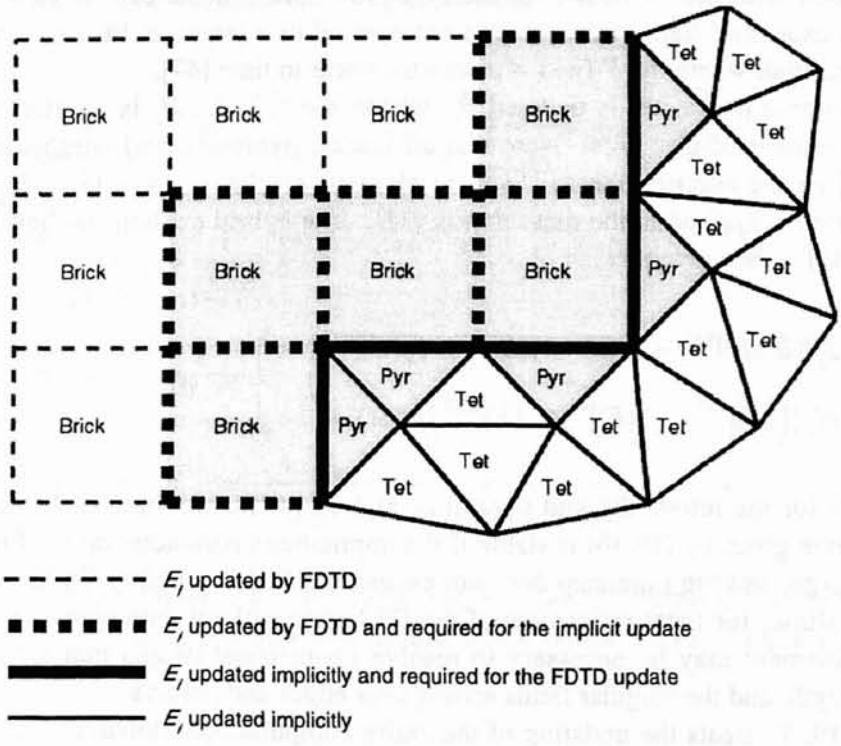


Fig. 19.9 Hybrid space grid. The dashed lines are cell edges where explicit FDTD updates are implemented, and the solid lines involve implicit updates. Thick lines denote cell edges involved in communications between the FDTD and implicit solvers. Shaded bricks and pyramids contribute to the stable hybrid interface, which is neither regular FDTD nor regular FE.

2. The remaining field components belong to the unstructured part of the space lattice. These are advanced in time implicitly by a generalized Newmark scheme. Specifically, $\{E_{\text{FE}}\}|^{n+1}$ is computed given the previous FE values $\{E_{\text{FE}}\}|^{n}$ and $\{E_{\text{FE}}\}|^{n-1}$, together with the FDTD solution $\{E_{\text{FDTD}}\}|^{n+1}$ that is associated with the edges shown as thick dashed lines in Fig. 19.9.

3. One cycle of the time-stepping procedure is now completed. The time-step counter n is advanced to $n+1$ and we return to step 1.

The next section presents the time-stepping scheme in a mathematical form that applies to the entire computational domain. That is, it covers both the explicit FDTD updates on the brick elements and the implicit updates on the unstructured part of the space lattice.

19.4.3 Generalized Newmark Scheme

The unconditional stability of the Newmark scheme with $\theta \geq 0.25$ is very useful for unstructured space grids. On the other hand, the choice $\theta = 0$ gives the standard approximation used by FDTD (viewed as a scheme for the second-order equation for the E-field). Therefore, the Newmark scheme is a good candidate for explicit-implicit time-stepping that interfaces FDTD with unstructured grids. It is attractive to associate the implicitness parameter θ in the Newmark scheme with the finite elements [34, 35]. This follows the successful recipe [28] used for the spatial discretization. Namely, FDTD can be recovered from FE by the spatial integration method associated with the elements. In fact, the Newmark scheme can be viewed as a linear combination of exact and trapezoidal integrations applied to a temporal FE representation based on Galerkin's method, where the E-field is piecewise linear in time [47].

The implicitness parameter is denoted θ_e, where $e = 1, 2, \ldots, N_e$ is an element index and N_e is the total number of elements. Note that all bricks, pyramids, and tetrahedra are indexed by e. Let $[S_e]$ be the contribution to $[S]$ from element e; that is, $[S] = [S_1] + [S_2] + \ldots + [S_{N_e}]$. The same partition is applied to the mass matrix $[M]$. The hybrid explicit-implicit time-stepping algorithm [34, 35] is then given by

$$\sum_{e=1}^{N_e} \begin{pmatrix} [S_e]\big(\theta_e\{E\}|^{n+1} - (2\theta_e - 1)\{E\}|^{n} + \theta_e\{E\}|^{n-1}\big) + \\ [M_e]\big(\{E\}|^{n+1} - 2\{E\}|^{n} + \{E\}|^{n-1}\big)\big/(\Delta t)^2 \end{pmatrix} = \{0\} \tag{19.76}$$

where $\theta_e \geq 0.25$ for the tetrahedra and pyramids, and $\theta_e = 0$ for the lumped bricks. The time-integration scheme given by (19.76) is stable if the implicitness parameter on the FE space lattice is sufficiently large, and the time-step does not exceed the FDTD stability limit. We emphasize that the hybrid allows for local refinement of the FE lattice without reduction of the global time step. Such refinement may be necessary to resolve geometrical details that are much smaller than the wavelength, and the singular fields arising near edges and corners.

Note that (19.76) treats the updating of the entire computational domain (i.e., all tetrahedra, pyramids and bricks). In a computer implementation, the FDTD region of the computational domain is, of course, not explicitly represented by the matrices in (19.76), but instead is treated by a standard FDTD subroutine. Equation (19.76) is mainly useful for constructing the FE matrices and proving stability for the hybrid time-stepping scheme.

19.4.4 Proof of Stability

The hybrid algorithm based on the time-stepping scheme of (19.76) can be rigorously proven to be stable up to the FDTD stability limit [34, 35]. The proof is relatively simple. Further, given the nature of the hybrid method, it also applies to both the FDTD and FE schemes individually. It is based on the von Neumann analysis for the quadratic form of (19.76). For simplicity, the permittivity and permeability are taken to be constants.

Eigenvalues of FDTD for One Brick Element

It is useful to consider a single FDTD cell before proving stability for the hybrid. Specifically, we require a bound on the 12 eigenvalues of a single FDTD element, which has one degree of freedom associated with each of its twelve edges. The eigenvalues are computed from the following:

$$[S_e]\{\tilde{E}\} = \omega^2 [M_e]\{\tilde{E}\} \tag{19.77}$$

with $[S_e]$ and $[M_e]$ lumped by trapezoidal integration. Here, seven eigenvalues are exactly zero. These correspond to the "potential modes" $\tilde{\boldsymbol{E}} = -\nabla\varphi$, where the potential φ is a trilinear function. The potential can be set to zero at one node, and then there is one eigenmode of (19.77) with $\omega = 0$ for each of the remaining seven nodes. For a brick with sides Δx, Δy, and Δz in the three coordinate directions, the remaining eigenvalues are

$$\omega_8^2 = \frac{4}{\varepsilon\mu}\left[\frac{1}{(\Delta y)^2} + \frac{1}{(\Delta z)^2}\right] \qquad \omega_9^2 = \frac{4}{\varepsilon\mu}\left[\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta z)^2}\right]$$

$$\omega_{10}^2 = \frac{4}{\varepsilon\mu}\left[\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2}\right] \tag{19.78}$$

and the largest eigenvalue is the pair

$$\omega_{\max}^2 = \omega_{11}^2 = \omega_{12}^2 = \frac{4}{\varepsilon\mu}\left[\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2} + \frac{1}{(\Delta z)^2}\right] \tag{19.79}$$

The largest eigenvalues equal the eigenvalue for the fastest varying exponential function $\exp[j\pi(x/\Delta x + y/\Delta y + z/\Delta z)]$ on a uniform, infinite space lattice. This mode gives the stability limit for the FDTD time-step:

$$\Delta t \leq \Delta t_{\max} = 2/\omega_{\max} \tag{19.80}$$

If $\Delta x = \Delta y = \Delta z = \Delta$, (19.80) gives the usual stability limit, $\Delta/c\sqrt{3}$. Equation (19.79) for the largest eigenvalues gives the following inequality, valid for any complex array $\{\tilde{E}\}$:

$$\{\tilde{E}\}^{\mathrm{H}}[S_e]\{\tilde{E}\} \le \omega_{\max}^2 \{\tilde{E}\}^{\mathrm{H}}[M_e]\{\tilde{E}\} \tag{19.81}$$

where $\{\tilde{E}\}^{\mathrm{H}}$ is the complex transpose of $\{\tilde{E}\}$.

Quadratic Form for the Hybrid Algorithm

The complex eigenmode $\{\tilde{E}\}$ of the hybrid algorithm (19.76) has a growth factor ρ such that $\{E\}|^n = \rho^n \{\tilde{E}\}$. For this mode, (19.76) gives

$$\sum_{e=1}^{N_e} \left([S_e]\left(\theta_e \rho^2 - (2\theta_e - 1)\rho + \theta_e\right) + [M_e](\rho^2 - 2\rho + 1)/(\Delta t)^2 \right)\{\tilde{E}\} = \{0\} \tag{19.82}$$

Stability is equivalent to $|\rho| \le 1$ for all modes $\{\tilde{E}\}$. With the substitution $\rho = (1 + \vartheta)/(1 - \vartheta)$, this condition becomes $\Re(\vartheta) \le 0$, and (19.82) is transformed into

$$\sum_{e=1}^{N_e} \{\tilde{E}\}^{\mathrm{H}}[S_e]\{\tilde{E}\} = -\vartheta^2 \sum_{e=1}^{N_e} \left(4\{\tilde{E}\}^{\mathrm{H}}[M_e]\{\tilde{E}\}/(\Delta t)^2 + (4\theta_e - 1)\{\tilde{E}\}^{\mathrm{H}}[S_e]\{\tilde{E}\}\right) \tag{19.83}$$

Since all the matrices are Hermitian, ϑ^2 is real. Furthermore, since the left-hand side is nonnegative, stability follows if the sum on the right-hand side is positive for all $\{\tilde{E}\}$. This sum can be split into the contribution from the explicit and implicit elements. The contribution from the implicit elements is nonnegative if $\theta_e \ge 0.25$. By virtue of (19.81), the contribution from the explicit elements is nonnegative if the time-step satisfies FDTD stability condition (19.80). Therefore, explicit-implicit algorithm (19.76) with $\theta_e \ge 0.25$ on the implicit elements is stable for time-steps up to the FDTD stability limit.

19.4.5 Alternative Time-Stepping Schemes

There are alternative versions of the explicit-implicit time-stepping scheme that have similar stability properties. In fact, the original version [28] of the hybrid algorithm used the same implicitness parameter θ for all the edges that are updated implicitly. Consequently, the edges of the pyramids and tetrahedra are updated implicitly by

$$\begin{aligned}[S]\left(\theta\{E\}|^{n+1} - (2\theta - 1)\{E\}|^n + \theta\{E\}|^{n-1}\right) \\ + [M]\left(\{E\}|^{n+1} - 2\{E\}|^n + \{E\}|^{n-1}\right)/(\Delta t)^2 = 0\end{aligned} \tag{19.84}$$

with $\theta \ge 0.25$. Stability of a similar explicit-implicit hybrid in acoustics was proven by Belytschko and Mullen [48]. These authors used the trapezoidal integration rule in time on the implicit grid, with the displacement ξ and velocity $v = \partial\xi/\partial t$ placed on the same time levels. This is equivalent to the Newmark scheme with $\theta = 0.25$. However, accuracy is often somewhat better for the case when the implicitness parameter is associated with the elements instead of the edges [i.e., the time integration scheme described by (19.76)].

19.4.6 Extensions of the Hybrid FDTD-FE Concept

Explicit-implicit time-stepping is also useful for modeling coaxial transmission lines [49]. The cross section of a coaxial cable normally has small dimensions relative to the wavelength as well as the size of the driven geometry. It is appealing to use a grid of prisms that conforms to the cable's cross section, resolves longitudinal wave motion, and connects to an unstructured grid at the feedpoint. The three-dimensional FE solution on the grid of prisms can be reduced to the one-dimensional transmission-line equation by means of "macro elements" with the voltage $V_k^n = V(z_k, t^n)$ as the unknown for the TEM wave [49]. Trapezoidal integration along the cable's axis combined with $\theta_e = 0$ recovers from the FE representation an explicit FDTD scheme for the corresponding transmission-line voltage equation:

$$\frac{V_{k+1}^n - 2V_k^n + V_{k-1}^n}{(\Delta z)^2} - \frac{1}{c^2}\frac{V_k^{n+1} - 2V_k^n + V_k^{n-1}}{(\Delta t)^2} = 0 \tag{19.85}$$

which can be truncated using techniques developed for FDTD [50]. This procedure combines in the same physical region finite differences along the coaxial waveguide with finite elements for the cross section of the waveguide. It enjoys the characteristic features of the stable FDTD-FE hybrid; for example, it preserves reciprocity, and its stability can be rigorously proved for time-steps up the stability limit of the one-dimensional transmission-line equation.

19.4.7 Reflection at the Interface of FDTD and FE Regions of a Hybrid Space Lattice

Different space-cell shapes possess different numerical dispersion characteristics. Therefore nonphysical numerical wave reflections can occur at the interface of FDTD and FE regions of a hybrid space lattice. It is possible that hybrid space lattices could generate nonphysical wave reflections that are comparable to the physical reflections of interest, especially if the latter are weak. A simple arrangement to test reflection at the FDTD-FE interface is illustrated in Fig. 19.10, where a thin layer of tetrahedra and pyramids is embedded in a cubic-cell FDTD space lattice, as was shown in Fig. 19.6. The width of the waveguide is twice its height.

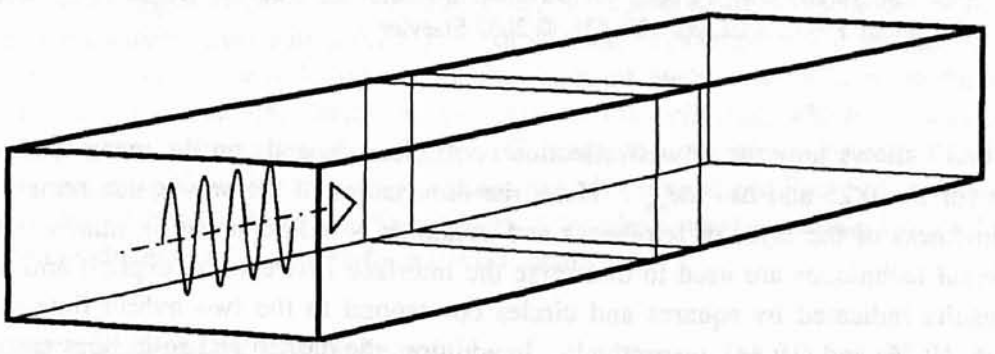


Fig. 19.10 Arrangement to test reflection at the FDTD-FE interface: waveguide with an incident pulse travelling towards the implicit layer. *Source:* Rylander and Bondeson, *Comput. Phys. Comm.*, 2000, pp. 75–82, © 2000 Elsevier.

Referring to Fig. 19.10, the numerical experiment involves calculating the reflection of a TE_{10}-mode pulse from a transverse layer of tetrahedra and pyramids located within the waveguide. The incident pulse has the time dependence $E_y(t) = \exp[-(t-t_0)^2/\tau^2]\sin(2\pi f t)$, where $\tau = 2.5/f_c$, $t_0 = 2.5\tau$, $f = \sqrt{2}f_c$, and f_c is the cut-off frequency for the TE_{10} mode.

Fig. 19.11 provides contour plots which display the power-reflection coefficient for the test geometry of Fig. 19.10 as a function of θ and the normalized Courant number. Here, the cross section of the waveguide is discretized by 10×5 FDTD cells, and the mean spatial resolution is 14 cells per wavelength. From this figure, we see that the power-reflection coefficient is below -46 dB, a level of reflection error that is acceptable for many applications of interest.

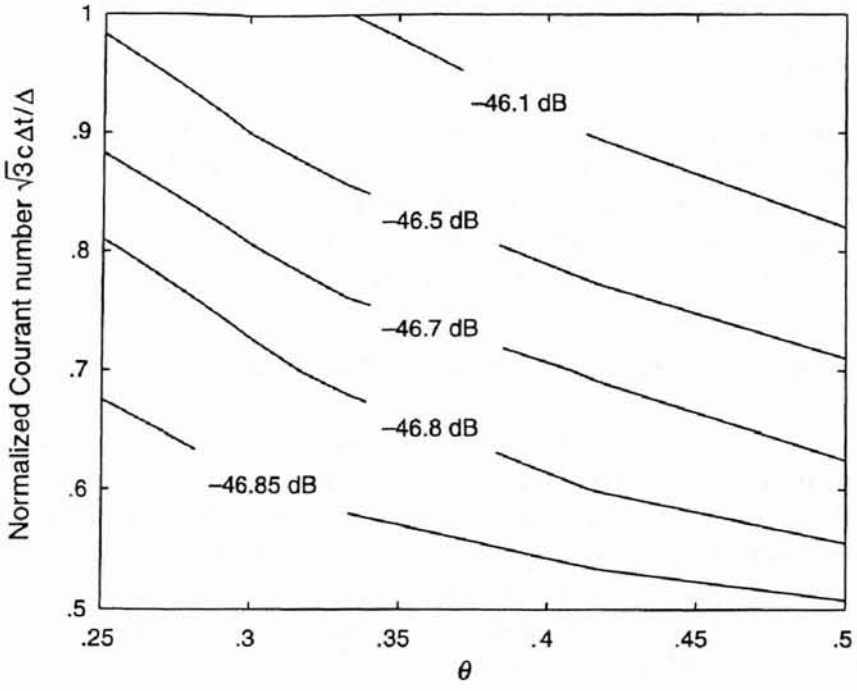


Fig. 19.11 Level contours of the power-reflection coefficient for the waveguide test of the hybrid scheme (see Fig. 19.10) with 14 cells per wavelength resolution. *Source:* Rylander and Bondeson, *J. Comput. Phys.*, 2002, pp. 426–438, © 2002 Elsevier.

Fig. 19.12 shows how the power-reflection coefficient depends on the mean space-lattice resolution for $\theta = 0.25$ and $\Delta t = \Delta t_{max}$. Here, the dimensions of the waveguide remain fixed, and the thickness of the layer of tetrahedra and pyramids is held constant in number of cells. Four different techniques are used to discretize the interface between the explicit and implicit grids. Results indicated by squares and circles correspond to the two hybrid time-stepping schemes of (19.76) and (19.84), respectively. In addition, the dashed and solid lines correspond to exact and trapezoidal integrations over the bases of the pyramids, respectively, as proposed by Riley [51]. A least-squares fit to the results for $\lambda/\Delta \geq 17$ shows that the power-reflection coefficient decreases approximately as $\Delta^{5.7}$ as the mesh resolution becomes finer.

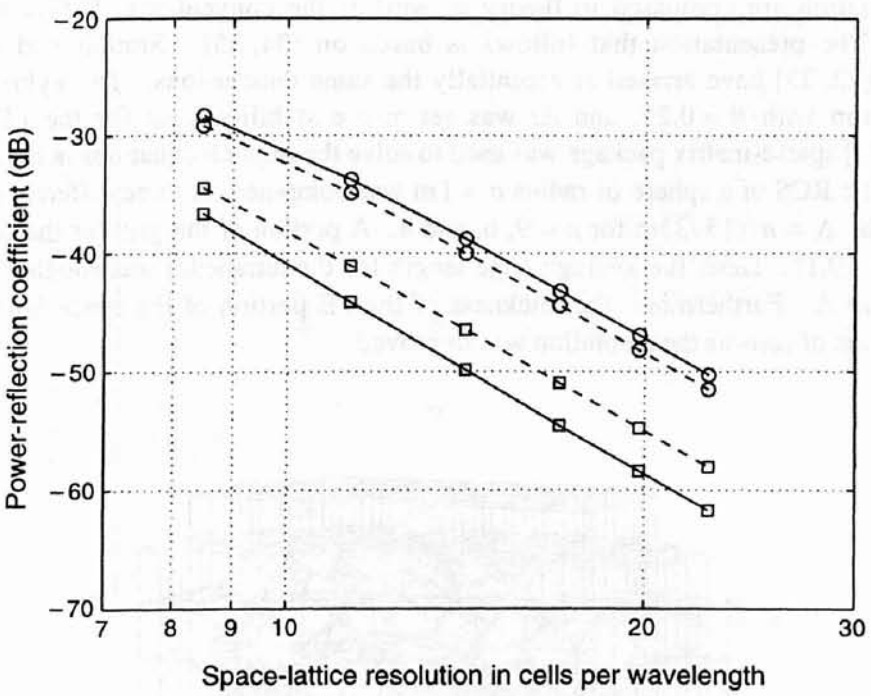


Fig. 19.12 Power-reflection coefficient for $\theta = 0.25$ and $\Delta t = \Delta t_{max}$ versus the mean number of cells per wavelength. The results of time-stepping scheme (19.76) are shown as squares, and those of alternative scheme (19.84) are shown as circles. Dashed lines indicate exact integration over the base of the pyramids [51] while solid lines indicate trapezoidal integration. *Source:* Rylander and Bondeson, *J. Comput. Phys.*, 2002, pp. 426–438, © 2002 Elsevier.

The results depicted in Fig. 19.12 can be interpreted intuitively by considering the space lattice as a dispersive wave-propagation medium. Spatially varying mesh resolutions and element shapes yield spatially varying numerical dispersion characteristics. Given a uniform mesh refinement, the numerical dispersion relations converge towards the analytical result with an error that is, to lowest order, proportional to Δ^2. Thus, the amplitude of the reflection coefficient from an internal meshing interface is expected to decay as the square of the cell size in the asymptotic region of convergence. However, the typical mode of operation for the FDTD-FE hybrid uses a mesh layer with a thickness that is constant in terms of the number of cells and small in relation to the wavelength. The thickness of such a layer is proportional to Δ, and the reflection coefficient decays as the cube of the cell size, which is faster than the discretization errors inherent to linear schemes. Consequently, the power reflection coefficient is expected to decay as Δ^6, which is in good agreement with the results in Fig. 19.12. Similar levels of internal mesh reflection have also been obtained for plane-wave scattering from a free-space block of tetrahedra embedded in a uniform FDTD mesh [51].

19.4.8 Scattering from the PEC Sphere

It is important to characterize and quantify the numerical error of the FDTD-FE hybrid in order to facilitate the reliable prediction of electromagnetic phenomena. Scattering from a PEC sphere is a useful test case since analytical results are available. In this section, results based upon the

hybrid formulation are compared to theory as well as the conventional FDTD scheme with staircasing. The presentation that follows is based on [34, 35]. Similar and independent comparisons [22, 23] have arrived at essentially the same conclusions. The hybrid FDTD-FE scheme was run with $\theta = 0.25$, and Δt was set at the stability limit for the FDTD region. The PETSc [52] sparse-matrix package was used to solve the implicit equations in the FE region.

The bistatic RCS of a sphere of radius $a = 1\,\mathrm{m}$ was computed on three different meshes with FDTD cell size $\Delta = n/(15\sqrt{3})\,\mathrm{m}$ for $n = 9$, 6, and 4. A portion of the grid for the case $n = 9$ is shown in Fig. 19.13. Here, the average edge length for the tetrahedra was roughly equal to the FDTD cell size Δ. Furthermore, the thickness of the FE portion of the space lattice was kept constant in terms of cells as the resolution was improved.

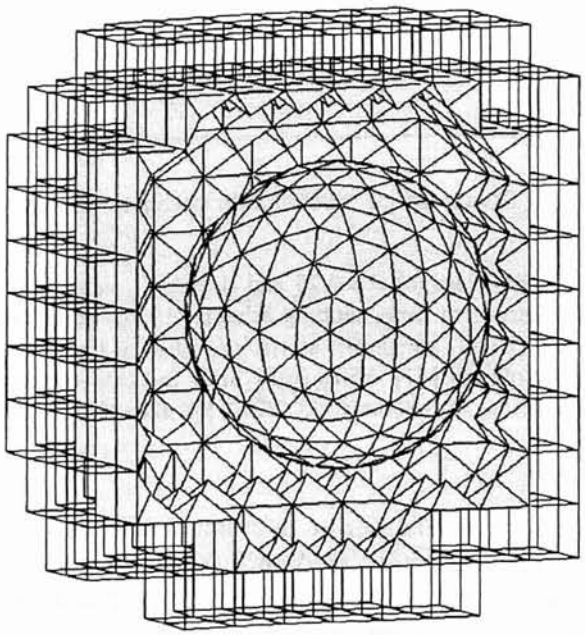


Fig. 19.13 Cutaway view of the hybrid FDTD/FE space lattice for a PEC sphere for the case $\Delta = \sqrt{3}/5$ (coarsest mesh). The discretized surface of the sphere is shown together with some of the pyramids. The first layer of FDTD cubes, connecting to the bases of the pyramids, is indicated by lines. *Source:* Rylander and Bondeson, *J. Comput. Phys.*, 2002, pp. 426–438, © 2002 Elsevier.

Fig. 19.14 shows the relative error $e(\Delta) = \|\sigma_n - \sigma_a\|_2 / \|\sigma_a\|_2$ for the wavelength $\lambda = 4.16\,\mathrm{m}$ ($ka = 1.5$), where σ_n and σ_a are the numerically computed and analytic bistatic RCS, respectively, and $\|\cdot\|_2 = [\int_\Omega (\cdot)^2 d\Omega]^{1/2}$. Squares and circles denote the results of hybrid schemes (19.76) and (19.84), respectively. We see that time-stepping scheme (19.76) reduces the error by 20% to 30% compared to (19.84). Least-squares fits to the model $e(\Delta) \sim \Delta^\alpha$, shown by the dashed lines for $\alpha = 2$, indicate that the FDTD / FE hybrid achieves second-order convergence. This is expected since the hybrid is based on linear approximations. One effect that contributes to deviations from second-order convergence is nonuniform refinement of the unstructured grid.

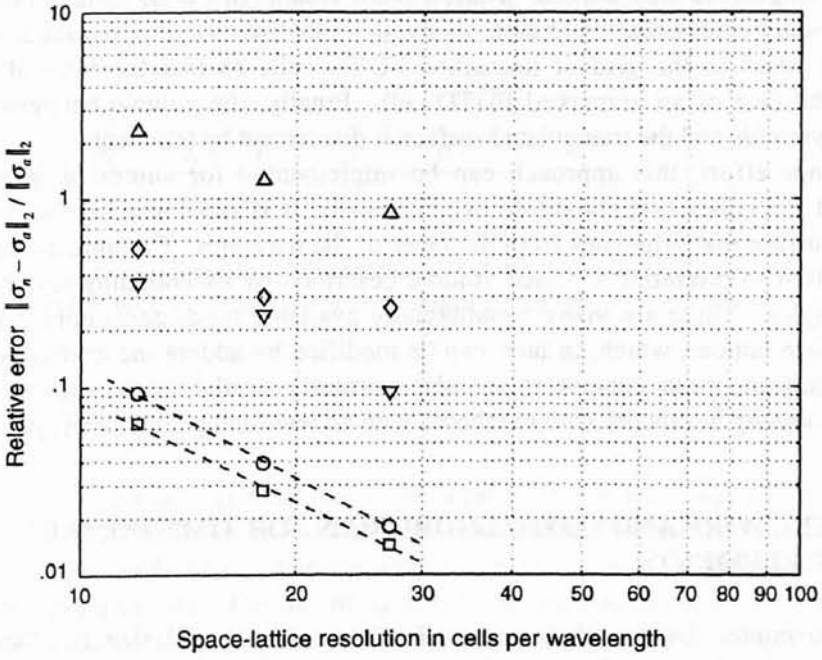


Fig. 19.14 Relative error in the bistatic RCS of a PEC sphere is shown by squares and circles for time-stepping schemes (19.76) and (19.84), respectively. Fits to the model $e(\Delta) \sim \Delta^2$ are shown by the dashed lines. Three different criteria for an FDTD cell to considered PEC were used for the staircased FDTD: (1) at least one corner of the cell is inside the physical sphere (shown by Δ), (2) the midpoint of the cell is inside the physical sphere (shown by $\Diamond$), and (3) all corners of the cell are inside the physical sphere (shown by ∇). *Source:* Rylander and Bondeson, *J. Comput. Phys.*, 2002, pp. 426–438, © 2002 Elsevier.

For the staircased FDTD, some rather arbitrary decisions must be taken for rendering the sphere. Fig. 19.14 shows results obtained using three different criteria for an FDTD cell to be classified as PEC: (1) at least one corner of the cell is inside the physical sphere (shown by Δ), (2) the midpoint of the cell is inside the physical sphere (shown by $\Diamond$), and (3) all corners of the cell are inside the physical sphere (shown by ∇). We see that, at all resolutions, the best FDTD results are obtained with method (3), which yields errors about 5 to 9 times larger than the hybrid. This best staircased FDTD model needs more than twice the linear resolution of the hybrid to achieve the same error level. Moreover, the staircased FDTD schemes do not show a very clear order of convergence, which reduces the predictive power of extrapolation.

19.5 MESH-GENERATION APPROACHES

In principle, the construction of hybrid meshes is relatively simple. One direct approach is to start with a structured FDTD space lattice and a triangulation of a complicated boundary. The triangulated surface is placed in the structured FDTD mesh, and all FDTD cells that contain one or several triangles are marked. The marked FDTD cells could then be removed from the FDTD computation, since that region will be treated using FE. In order to allow for a high-quality unstructured mesh, one or more additional layers of FDTD cells that connect to FDTD

cells cut by triangles are also marked. Thus, a small volume of FDTD cells in the vicinity of the complex boundary is identified, and then, this volume is discretized by pyramids and tetrahedra. Pyramids are fitted to the grid of unmarked FDTD cells so that the base of each pyramid connects to the face of an unmarked FDTD cell. Finally, the volume between the triangular faces of the pyramids and the triangulated surface is discretized by tetrahedra.

With minor effort, this approach can be implemented for automatic generation of the pyramids and the cubes that constitute the FE mesh. It is relatively easy to also extract the triangulated surface that originates from the faces of the pyramids. Commercial mesh generators can then be used to discretize a closed volume described by its bounding surface expressed in terms of triangles. There are many commercially available mesh generators that can generate tetrahedral space lattices, which, in turn, can be modified by adding the necessary pyramids and cubes. In addition, mesh generators are also currently available that directly create hybrid lattices by combining hexahedra with tetrahedra with an interfacing layer of pyramids [53].

19.6 SUBCELL WIRE AND SLOT ALGORITHMS FOR TIME-DOMAIN FINITE ELEMENTS

The ability to model features that are small relative to the cell size is often important in electromagnetic simulations. In principle, an unstructured mesh can be used to resolve these small features. However, the increase in number of unknowns can be prohibitive. Thus, the development of accurate models that characterize the physics of the feature without the need for a highly resolved space lattice is essential to many applications.

Thin wires are often important parts of electromagnetic compatibility and antenna problems. A subcell model for thin wires in FDTD using modified telegraphers' equations was developed by Holland and Simpson [54]. This model has also recently been applied to the FE method [51, 55]. In this setting, the thin-wire equations become a second-order wave equation for the current on the wire. This equation is driven by the E-field along the wire, whereas the current on the wire is a source term for Maxwell's equations. It is shown that a consistent discretization of the wire equation using nodal basis functions and a radial weighting function results in a symmetric spatial coupling between the electromagnetic field and the wire.

Practical systems often possess narrow cracks and gaps that are also challenging to include in an analysis. In [51], a thin-slot model was presented that was based on a dual formulation to thin wires. Consequently, the thin-wire algorithm can be generalized to model thin slots, thereby making a unified theory for both thin wires and thin slots possible. Previous thin-slot models have been susceptible to instabilities. However, by incorporating symmetric coupling between the electromagnetic field and the slot, and between the electromagnetic field and the wire, it is possible to establish the stability of the field-wire-slot system [55, 56].

This section presents algorithms for thin wires and thin slots on an unstructured mesh. The fact that neither the wires nor the slots are required to follow the elemental edges within the volumetric space lattice gives considerable modeling flexibility when including these subcellular features into simulations. Additional details can be found in [55, 56].

19.6.1 Modeling Thin Wires

To derive the wire equation, we follow the approach by Holland and Simpson [54], and assume an infinitely long cylinder of radius a running in the z-direction. For simplicity, we also assume

that $\sigma = 0$ in the neighborhood of the wire. By adopting a cylindrical coordinate system along with the assumption that the electromagnetic fields E_r and H_θ are proportional to $1/r$ local to a thin wire, it follows that [51, 54]

$$L\frac{\partial^2 I}{\partial t^2} + R\frac{\partial I}{\partial t} - \frac{L}{\varepsilon\mu}\frac{\partial^2 I}{\partial z^2} = \frac{\partial E_z}{\partial t} + \frac{\partial \tilde{V}^{\text{inc}}}{\partial t} \tag{19.86}$$

where I is the wire current, R is the wire resistance per unit length, $\tilde{V}^{\text{inc}}$ is a voltage source per unit length, and L is the wire inductance per unit length given by

$$L = \frac{\mu}{2\pi}\log\left(\frac{r_0 + a}{2a}\right) \tag{19.87}$$

where $(r_0 + a)/2$ is an average distance from the wire to the surrounding E-field components used to drive the wire. The focus is on arbitrary wires that do not need to follow mesh lines. Therefore, the E-field needs to be approximated along the wire through interpolation of the surrounding field components. On the other hand, the current density is approximated by distributing the wire current to the surrounding field components, as shown in Fig. 19.15.

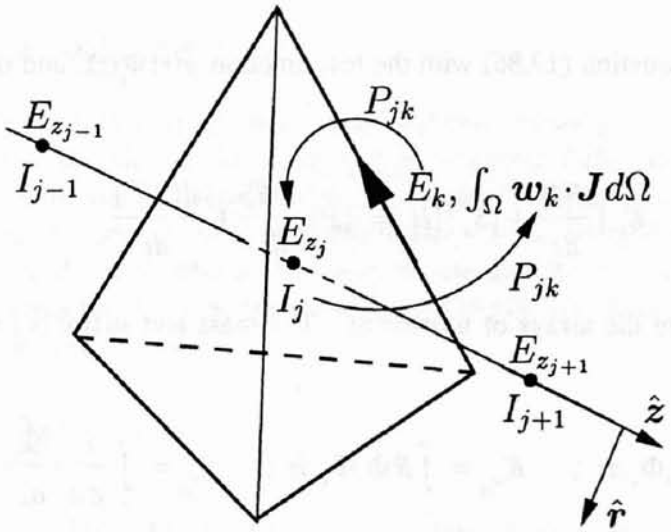


Fig. 19.15 An arbitrary wire described in cylindrical coordinates. Symmetric coupling between the wire variables and the surrounding field variables is realized through the use of the P operator, defined in (19.94), in both directions.

Consistent with the FE framework, the current I along the wire can be expanded in basis functions as

$$I(z,t) = \sum_j I_j(t)\,\Phi_j(z) \tag{19.88}$$

where Φ_j is the standard linear nodal basis function in one dimension [10], and I_j is the unknown current at wire node j. The current density is now expressed as

$$\boldsymbol{J}(r,z,t) \;=\; \sum_j I_j(t)\,\Phi_j(z)\,g(r)\,\hat{z} \tag{19.89}$$

where r is the radial distance from the wire and $g(r)$ is a weighting function satisfying

$$\int_{r\geq a} g(r)\,2\pi r\,dr \;=\; 1 \tag{19.90}$$

and thus has dimension m^{-2}. It is important that this function decrease with r, and equal zero for $r \geq r_0$, which gives a compact support. The function used here is defined by

$$g(r) \;=\; \begin{cases} 0 & r < a \\ \dfrac{1+\cos(\pi r/r_0)}{\pi(r_0^2-a^2)+(2r_0^2/\pi)\left[-1-\cos(\pi a/r_0)-(\pi a/r_0)\sin(\pi a/r_0)\right]} & a \leq r \leq r_0 \\ 0 & r > r_0 \end{cases} \tag{19.91}$$

Multiplying the wire equation (19.86) with the test function $g(r)\Phi_j(z)$, and then integrating over all space yields [55]

$$[M_w]\frac{d^2\{I\}}{dt^2} + [K_w]\frac{d\{I\}}{dt} + [S_w]\{I\} \;=\; [P]\frac{d\{E\}}{dt} + \frac{d\{V^{\text{inc}}\}}{dt} \tag{19.92}$$

where $\{I\}$ and $\{E\}$ are the arrays of unknowns. The mass and stiffness matrices for the wire equation are given by

$$M_{w_{jk}} = \int_z L\,\Phi_j\Phi_k\,dz\;; \qquad K_{w_{jk}} = \int_z R\,\Phi_j\Phi_k\,dz\;; \qquad S_{w_{jk}} = \int_z \frac{L}{\varepsilon\mu}\frac{d\Phi_j}{dz}\frac{d\Phi_k}{dz}\,dz \tag{19.93}$$

and the interpolation operator is defined by

$$P_{jk} \;=\; \int_\Omega \hat{z}\cdot\boldsymbol{w}_k\,g(r)\,\Phi_j(z)\,d\Omega \tag{19.94}$$

For the coupling between the wire and the field, (19.89) is inserted into (19.32), which yields

$$\int_\Omega \boldsymbol{w}_k\cdot\frac{\partial \boldsymbol{J}}{\partial t}\,d\Omega \;=\; \sum_j P_{jk}\frac{dI_j}{dt} \tag{19.95}$$

The interpolation operator is used to approximate the E-field along the wire at each wire node, as illustrated in Fig. 19.15. Due to the compact support of the basis and weighting functions, each wire node is surrounded by an interpolation cylinder of radius r_0 and length given by the two wire beams sharing the node. Only edge-projected fields belonging to tetrahedra that are inside or partly inside this cylinder contribute. A sixth-degree Gaussian quadrature formula is used to calculate the integral in (19.94).

19.6.2 Modeling Thin Slots

The subcell slot model is based on a dual formulation of the thin-wire model. Let a thin slot with length L, width w, and depth d be located in a conducting wall. It is assumed that $L >> w$, and w and d are both electrically small. All fields in the wall are set to zero, and the slot is modeled through the following equation [51]:

$$C_s \frac{\partial^2 V}{\partial t^2} - \frac{C_s}{\varepsilon\mu} \frac{\partial^2 V}{\partial z^2} = \frac{\partial H_z^{\text{diff}}}{\partial t} \tag{19.96}$$

where V is the magnetic current (the voltage across the slot width), H_z^{diff} is the difference in the z-component of the magnetic field on opposite sides of the slot wall, and C_s is the slot capacitance per unit length. The slot capacitance is given by

$$C_s = (2\varepsilon/\pi) \log[(r_0 + a_s)/2a_s] \tag{19.97}$$

where $a_s = (w/4)\exp(-\pi d/2w)$ is an equivalent antenna radius [57], and $(r_0 + a_s)/2$ is an average distance from the slot to the surrounding magnetic fields used to drive the slot. A possible extension of the slot model permits the inclusion of wall loss [51, 58].

Due to the fact that the source term in the slot equation is the difference in the magnetic field on opposite sides of the slot wall, Maxwell's equations are discretized in their coupled-curl form using edge and facet basis functions, as shown in (19.10) to (19.16). Being a flux, $\boldsymbol{M}$ can also be expanded in facet basis functions as

$$\boldsymbol{M}(\boldsymbol{r},t) = \sum_i M_i(t)\, \boldsymbol{w}_{f_i}(\boldsymbol{r}) \tag{19.98}$$

which implies that $\{g\} = [M_f]\{M\}$ in (19.16). The voltage V along the slot can be expanded in linear basis functions as

$$V(z,t) = \sum_j V_j(t)\, \Phi_j(z) \tag{19.99}$$

where V_j is the unknown voltage at slot node j. By proceeding in a manner similar to that for thin wires, slot equation (19.96) is multiplied by the test function $g(r)\Phi_j(z)$. Then, integration over all space yields the following equation for $\{V\}$, the array of nodal voltage unknowns [56]:

$$[M_s]\frac{d^2\{V\}}{dt^2} + [S_s]\{V\} = [P_s]\big(-[C]\{E\} - \{M\}\big) \tag{19.100}$$

In (19.100), the mass and stiffness matrices are given by

$$M_{s_{jk}} = \int_z C_s \Phi_j \Phi_k dz \; ; \qquad S_{s_{jk}} = \int_z \frac{C_s}{\varepsilon\mu} \frac{d\Phi_j}{dz} \frac{d\Phi_k}{dz} dz \tag{19.101}$$

In addition, $[C] = [M_f]^{-1}[A]$ is the incidence matrix for the relation between the edges and the faces, and the interpolation operator $[P_s]$ is given by

$$P_{s_{jk}} = \int_z \frac{1}{\mu} \hat{z} \cdot \boldsymbol{w}_{f_k} \tilde{g}(r) \Phi_j(z) dz \tag{19.102}$$

for $\tilde{g}(r)$ a weighting function that equals $\pm 2g(r)$. As shown in Fig. 19.16, the interpolation between the field and the slot is almost identical to the interpolation between the field and the wire in (19.94). The only differences are that face elements are used for slots while edge elements are used for wires, and the sign in the radial weighting function $\tilde{g}$ is changed. The latter reflects the different normal directions on opposite sides of a slot wall.

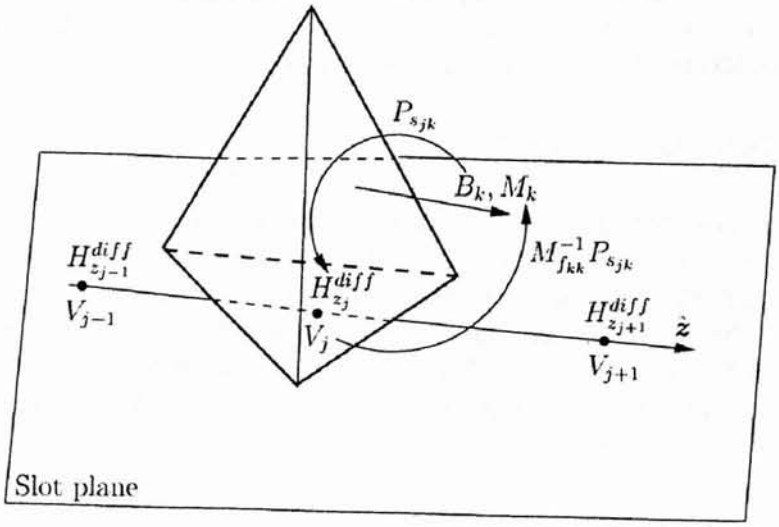


Fig. 19.16 An arbitrary slot located in a conducting plane. Symmetric coupling between the slot variables and the surrounding flux variables is realized through the use of the $[P_s]$ operator defined in (19.102) in both directions. Note that the tetrahedron has its bottom triangle in the slot plane, and that several tetrahedra located on both sides of this plane are involved in the coupling.

The magnetic current density $\boldsymbol{M}$ is calculated from the slot voltage, and can be expressed in cylindrical coordinates as

$$\boldsymbol{M}(r,z,t) = \sum_j V_j(t) \Phi_j(z) \tilde{g}(r) \hat{z} \tag{19.103}$$

Expressions (19.98) and (19.103) for $\boldsymbol{M}$ are set equal in a weak sense, which implies [56]

$$\{M\} = [M_f]^{-1}[P_s]^{\mathrm{T}}\{V\} \tag{19.104}$$

Finally, by using (19.28), (19.92), (19.95), (19.100), and (19.104), as well as the equivalence between the vector wave equation and the coupled-curl equations, the time-continuous field-wire-slot system in matrix form is given by

$$\begin{pmatrix} [M] & [0] & [0] \\ [0] & [M_w] & [0] \\ [0] & [0] & [M_s] \end{pmatrix} \frac{d^2}{dt^2} \begin{pmatrix} \{E\} \\ \{I\} \\ \{V\} \end{pmatrix} + \begin{pmatrix} [K] & [P]^{\mathrm{T}} & [0] \\ -[P] & [K_w] & [0] \\ [0] & [0] & [0] \end{pmatrix} \frac{d}{dt} \begin{pmatrix} \{E\} \\ \{I\} \\ \{V\} \end{pmatrix} +$$

$$\begin{pmatrix} [S] & [0] & [C]^{\mathrm{T}}[P_s]^{\mathrm{T}} \\ [0] & [S_w] & [0] \\ [P_s][C] & [0] & [S_s]+[P_s][M_f]^{-1}[P_s]^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \{E\} \\ \{I\} \\ \{V\} \end{pmatrix} = \begin{pmatrix} \{0\} \\ \{0\} \\ \{0\} \end{pmatrix} \tag{19.105}$$

To obtain an unconditionally stable field-wire-slot algorithm, this system is discretized by the second-order-accurate Newmark-Beta scheme defined in (19.34) to (19.36) [55, 56]. We note that the key to stability is the symmetric coupling between field and wires and field and slots. The preconditioned conjugate gradient (PCG) method can be used to solve the system at each time-step. Since wires and slots are one-dimensional, the additional memory requirements and arithmetic operations to include them in three-dimensional simulations are generally small.

19.6.3 Numerical Results for Thin Wires and Slots

The thin-wire algorithm is first applied to model a 1 m diameter circular loop antenna having a wire radius of 1.05 mm. The loop is discretized using 50 wire segments in an unstructured grid with an average edge length of 67.4 mm, and is excited at one of the wire nodes with an impulsive voltage source. Fig. 19.17 shows that the results of the hybrid FDTD-FE calculation for the loop's driving-point resistance agree well with a method-of-moments benchmark.

The thin-wire and thin-slot algorithms are next applied to model the excitation of a 14-cm-long wire of radius 0.8 mm positioned vertically within a $30 \times 22 \times 14$-cm metal box. The wire is perpendicular to the upper and lower 30×22-cm box surfaces at their centroids. It is fed by a $50\,\Omega$ source at the upper box surface, and is terminated with a $47\,\Omega$ load at the lower box surface. In addition, the lower box surface is penetrated with a slot of length 12 cm, width 1.6 mm, and depth 0.5 mm, which is located parallel to, and 16.5 cm away from, one of the 30-cm edges. An unstructured tetrahedral mesh is used to discretize the box interior and near exterior. Neither the wire nor the slot is aligned with this mesh. Four layers of FDTD cells are located between the unstructured mesh and the absorbing boundary. Fig. 19.18 shows the results for the wire's load power for two different mesh spatial resolutions, 0.5 and 1 cm. There is good agreement between the hybrid FDTD-FE results and measurements [59] over a wide frequency range. A small deviation, in particular for the solution on the coarser mesh, is observed around the resonance at 1.13 GHz. This resonance, as well as the ones at 1.27 and 1.37 GHz, are due to the slot, and disappear in both simulations and measurements when the slot is removed.

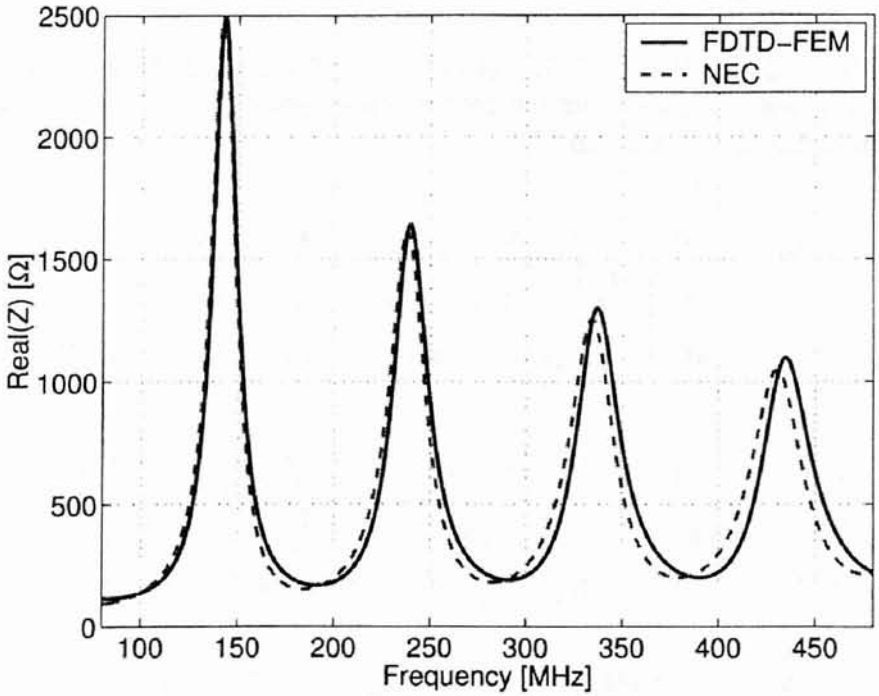


Fig. 19.17 Comparison of hybrid FDTD-FE and Numerical Electromagnetics Code (NEC) calculations of the input resistance of a 1 m diameter circular loop antenna having a wire radius of 1.05 mm. *Source:* Edelvik and Weiland, *IEEE APS Intl. Symp. Dig.*, 2004, pp. 3481–3484, © 2004 IEEE.

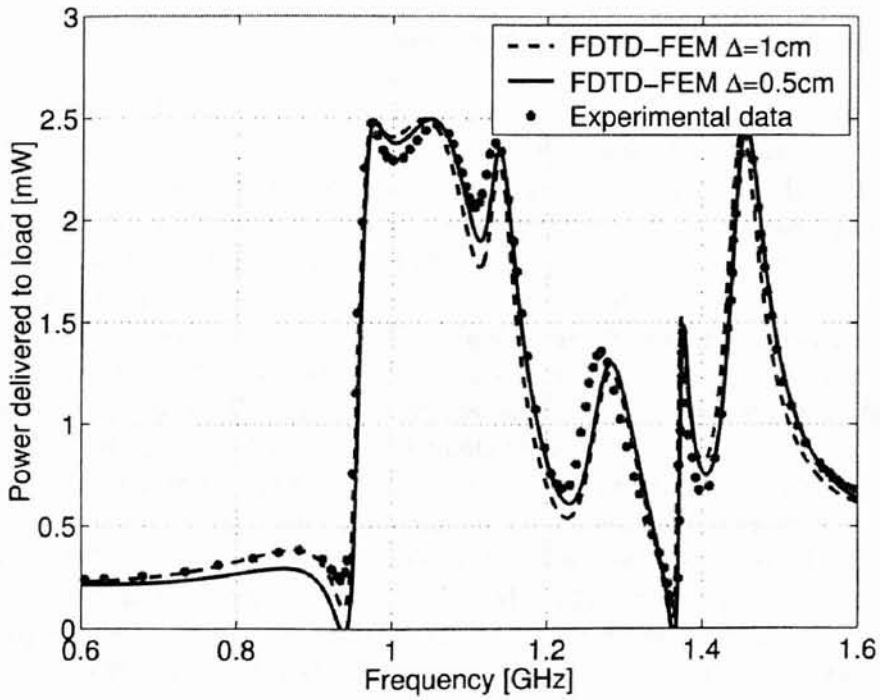


Fig. 19.18 Comparison of hybrid FDTD-FE calculations and measurements for a shielding enclosure. *Source:* Edelvik and Weiland, *Intl. J. Numerical Modeling*, 2004, pp. 365–383, © 2004 Wiley.

19.7 APPLICATION TO ADVANCED SCATTERING AND RADIATION PROBLEMS

This section demonstrates application of the hybrid FDTD-FE method to three advanced modeling problems: monostatic RCS of the NASA "almond," bistatic RCS of the Saab *Trainer* aircraft, and input impedance of the four-arm sinuous antenna. The hybrid FDTD-FE method is found to be much more efficient for such problems than using traditional staircased FDTD on a highly resolved grid [22]. For the same solution accuracy, savings of approximately 10:1 in memory requirements and 20:1 in execution time are possible.

19.7.1 Monostatic RCS of the NASA Almond

A number of electromagnetic wave modeling problems involve structures having sharp corners that generate local field singularities. Other problems involve calculating the RCS of weakly scattering structures having reflections comparable in strength to the background numerical noise. For such modeling challenges, the NASA almond is an excellent test case, since it possesses both a sharp tip and low-RCS behavior. Therefore, it has been used as a popular benchmark for RCS codes [19, 60–62]. A mathematical description of its surface is given by Woo et al. [60] together with data sets for measured and calculated monostatic RCS.

This section discusses a test of the stable FDTD-FE hybrid [62] in calculating the monostatic RCS of the NASA almond for the case when the wavelength equals the length of the almond, 0.2524 m. Results for this case are available in the literature [19, 60, 61]. Fig. 19.19 is a cutaway view of the hybrid mesh for the NASA almond, showing that the unstructured portion of the mesh has increased resolution at the almond's tip. In the discussion to follow, vertical and horizontal polarization refer to cases where the incident E-field is perpendicular and parallel to the x-y plane, respectively.

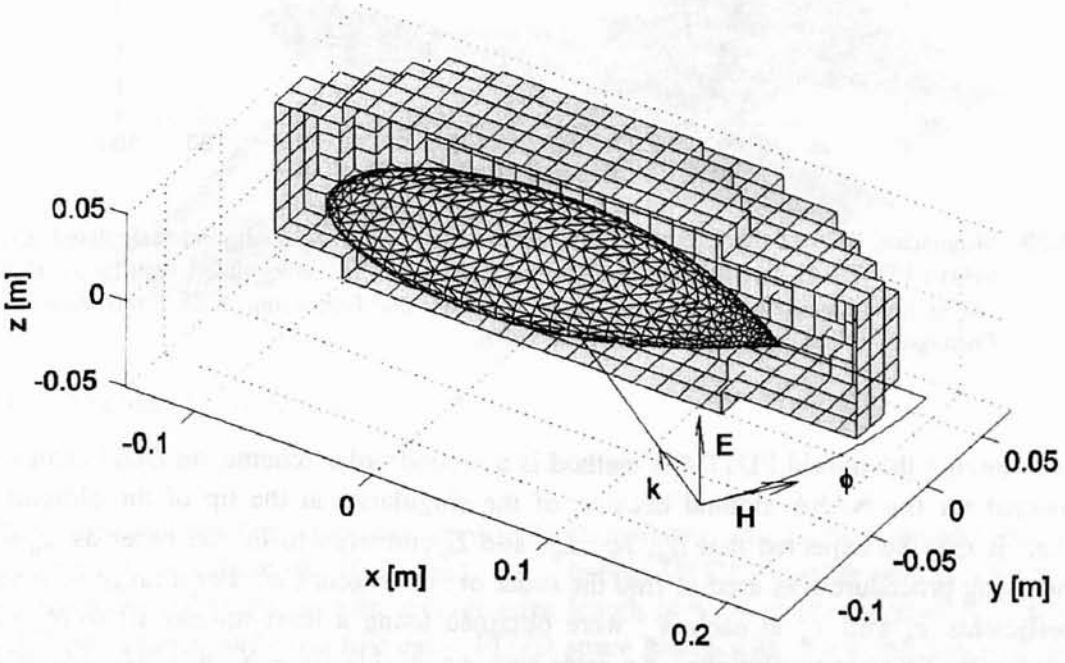


Fig. 19.19 The NASA almond embedded within a hybrid FDTD-FE mesh for the case of a vertically polarized incident plane wave. *Source:* Rylander and Bondeson, *IEEE Trans. Antennas and Propagation*, 2002, pp. 141–144, © 2002 IEEE.

The complex scattering amplitudes N_θ, N_ϕ, L_θ, and L_ϕ for the almond were calculated for azimuth angles $\phi_m = m\pi/16$ (for $m = 0, 1, \ldots, 16$) on six different meshes having decreasing cell size $\Delta = \lambda/(5n)$ (for $n = 3, 4, 5, 6, 8$, and 12). When changing the FDTD cell size, the mesh density was kept as a fixed function of space. The computed quantities were extrapolated to zero cell size. Then, a spline curve was used for interpolation with respect to ϕ, with the slope of the RCS-versus-ϕ curve set to zero for $\phi = 0$ and $\phi = \pi$.

Fig. 19.20 depicts the calculated monostatic RCS, σ, of the almond for horizontal and vertical polarization, as computed on the six different meshes. These results are seen to converge to the extrapolated values, even at the nulls. Relative to horizontal polarization, the vertical polarization RCS is less sensitive to cell size despite its lower levels for most monostatic angles.

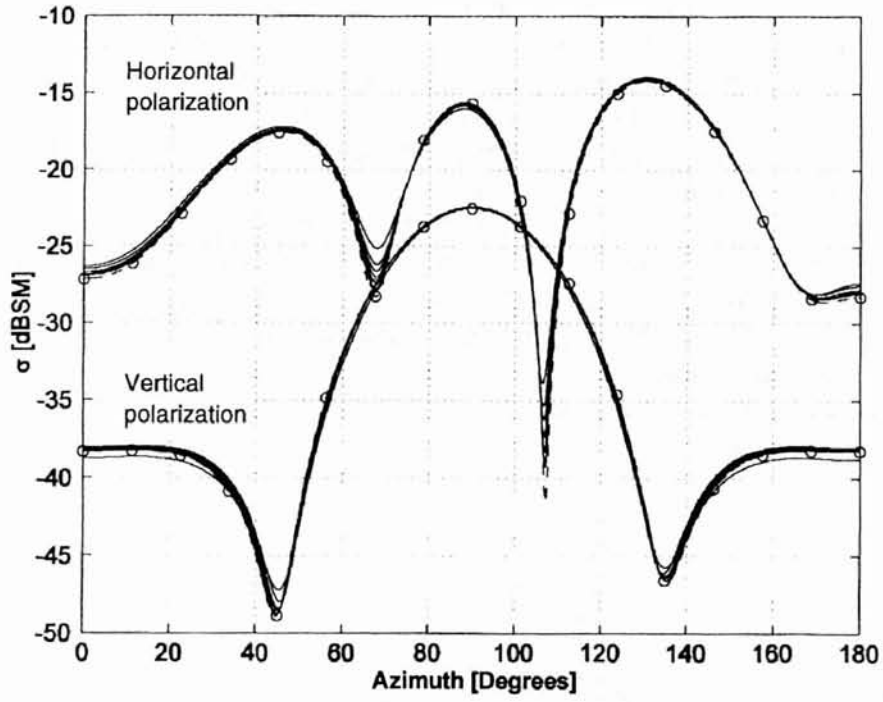


Fig. 19.20 Monostatic RCS of the 0.2524 m (one-wavelength long) NASA almond, calculated using the hybrid FDTD-FE method on six different meshes. The extrapolated results are shown as circles joined by dashed curves. *Source:* Rylander and Bondeson, *IEEE Trans. Antennas and Propagation*, 2002, pp. 141–144, © 2002 IEEE.

Even though the hybrid FDTD-FE method is a second-order scheme, an $O(\Delta^2)$ error cannot be expected for the NASA almond because of the singularity at the tip of the almond [63]. However, it may be expected that N_θ, N_ϕ, L_θ, and L_ϕ converge to lowest order as $c_0 + c_\alpha \Delta^\alpha$. The following procedure was used to find the order of convergence α. For a range of α values, the coefficients c_0 and c_α at each ϕ_m were obtained using a least-squares fit to N_ϕ and L_θ computed with different resolutions. An error sum $e = \sum_{\phi_m} [|Z_0(N_\phi - N_{\phi,\mathrm{f}})|^2 + |(L_\theta - L_{\theta,\mathrm{f}})|^2]$ was then evaluated, and the α that minimized e was chosen, where N_ϕ and L_θ were numerically computed, and $N_{\phi,\mathrm{f}}$ and $L_{\theta,\mathrm{f}}$ denote fitting functions with the form $c_0 + c_\alpha \Delta^\alpha$.

Detailed studies of this type have shown that, for horizontally polarized illumination of the almond (which maximally excites the field singularity at its tip), the order of convergence α of the hybrid FDTD-FE model is approximately 1.2 for meshing densities between 30 and 60 per wavelength [62]. However, for vertical polarization, α is close to 2 (i.e., indicating second-order accuracy of the model). This is reasonable, since the field singularity at the almond's tip is much less excited for vertically polarized illumination; therefore, the model's rate of convergence should approach that of the underlying hybrid algorithm. Overall, for both polarizations, a meshing density of 15 per wavelength yields results that are within 1 dB of the asymptotically converged values for at least 90% of the azimuth angles.

19.7.2 Bistatic RCS of the Saab *Trainer* Aircraft

This section presents results of the hybrid FDTD-FE technique for the bistatic RCS of the conceptual Saab aircraft, *Trainer* [23], shown in Fig. 19.21. The *Trainer* is 11 m long, 3.4 m high, and spans 8 m between its wingtips. It is assumed to be a PEC shell, and includes the jet engine air inlets. Illumination is head-on via a broadband impulse with the incident E-field either parallel to the wing plane (horizontal polarization) or perpendicular to the wing plane (vertical polarization).

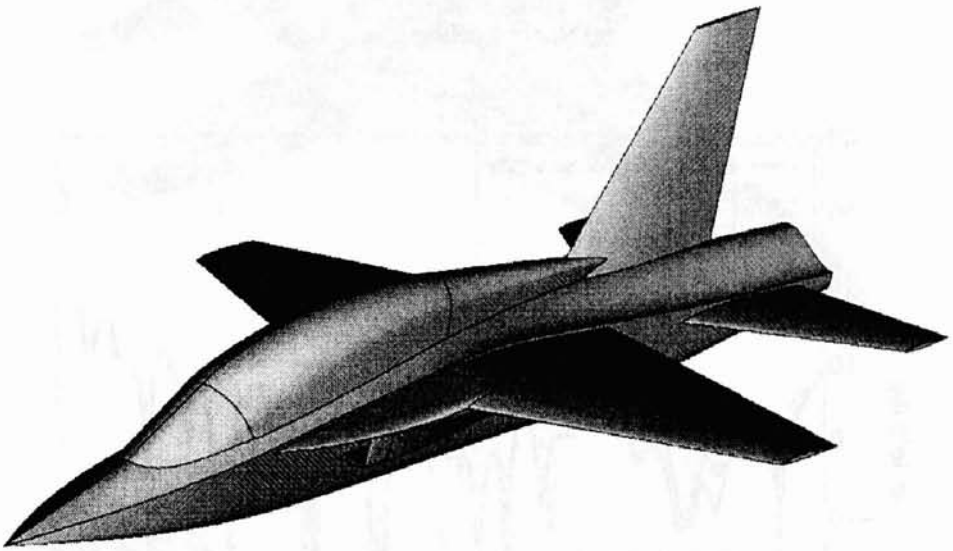
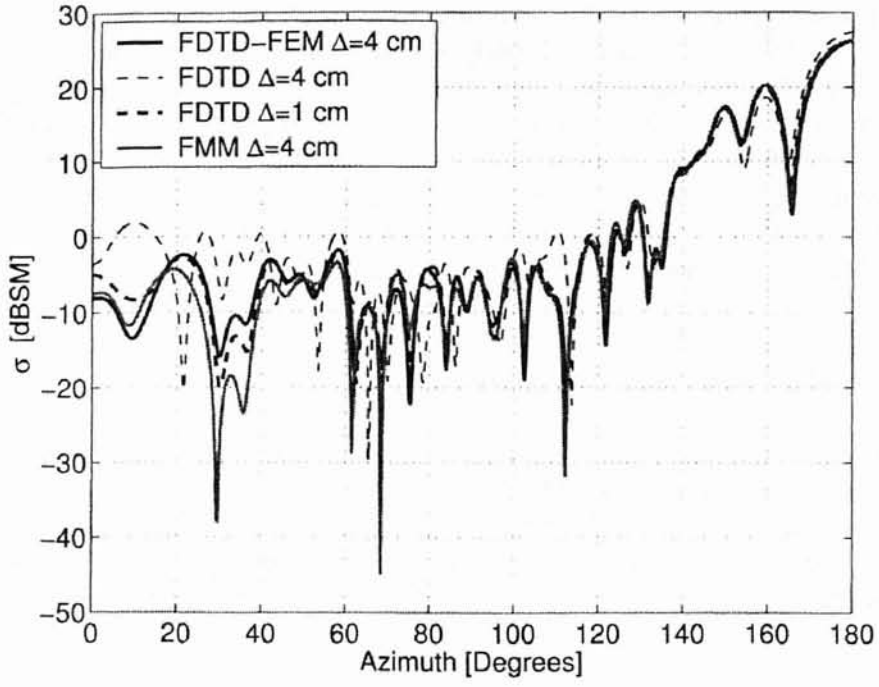
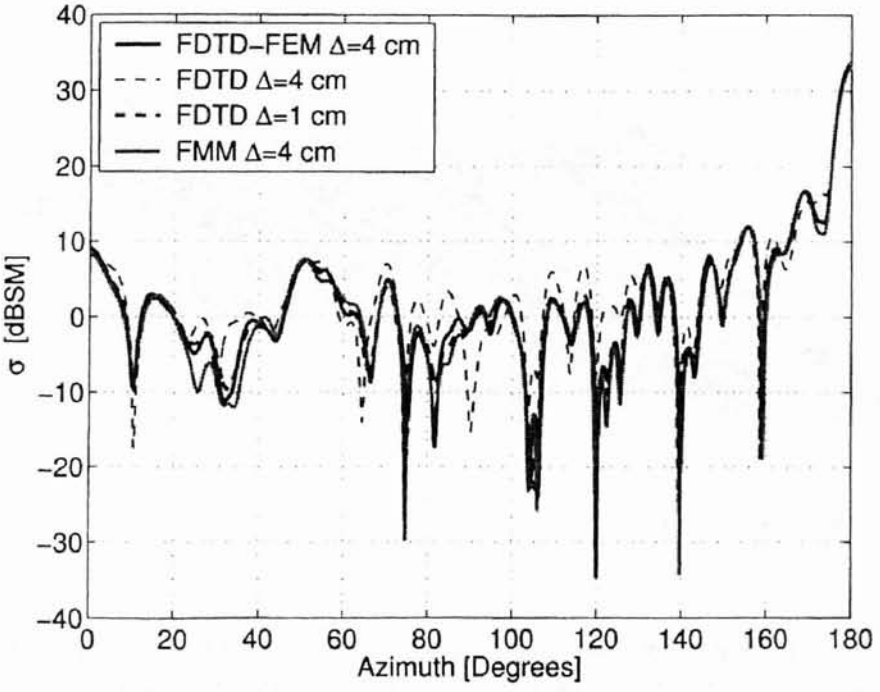


Fig. 19.21 The Saab *Trainer* aircraft.

Three different meshes have been applied to this model: (1) a conventional cubic FDTD space lattice with $\Delta = 4$ cm (9×10^6 cubes); (2) a hybrid FDTD-FE mesh with $\Delta = 4$ cm in the FDTD region (8×10^6 cubes) and an average edge length of 5 cm in the unstructured FE region (1.8×10^6 unknowns); and (3) a fine cubic FDTD space lattice with $\Delta = 1$ cm (400×10^6 cubes). All meshes are terminated by an eight-cell-thick UPML absorbing boundary. The computational domains are each sized such that the closest distance from the aircraft to the absorbing boundary is 10 space cells.



(a) Vertical polarization.



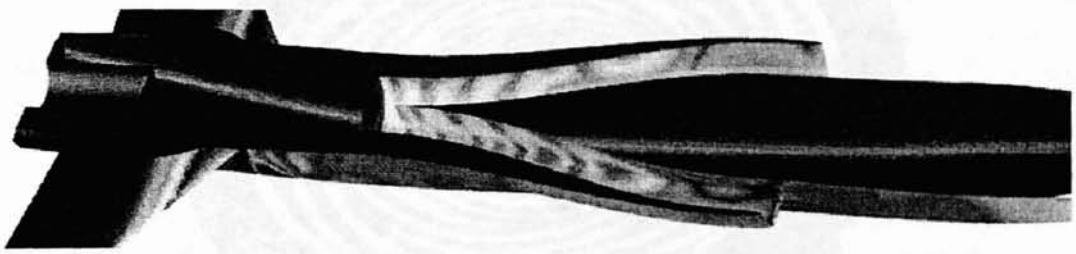
(b) Horizontal polarization.

Fig. 19.22 Comparison of FDTD-FE and phasor-domain fast multipole method calculations of the bistatic RCS of the Saab *Trainer* at 500 MHz.

Fig. 19.22 compares the bistatic RCS of the *Trainer* at 500 MHz with a fast multipole method [64] solution of the phasor-domain electric field integral equation. Here, the azimuth sweep is in the horizontal plane starting from the monostatic angle. Note the close correspondence between the results of the $\Delta = 1$ cm FDTD model (uniform resolution of $\lambda/60$ at 500 MHz), the $\Delta = 4$ cm ($\lambda/15$ resolution) hybrid FDTD-FE model, and the $\Delta = 4$ cm fast multipole method calculation. On the other hand, the $\Delta = 4$ cm FDTD model shows significant deviations, particularly for vertical polarization. Clearly, the staircased FDTD model requires finer resolution than the hybrid FDTD-FE model to obtain results of similar accuracy.

Fig. 19.23 provides snapshot visualizations of the induced surface currents on the *Trainer* calculated by the FDTD-FE technique for head-on illumination by a horizontally polarized Gaussian pulse. The corresponding color visualizations are in Fig. 1.5 of Chapter 1.

(a) Outer surface of aircraft.



(b) Within a jet engine air inlet.

Fig. 19.23 Grayscale visualizations of the induced surface currents on the *Trainer* aircraft calculated by the FDTD-FE technique for head-on illumination by a horizontally polarized Gaussian pulse. The observation time is just as the illuminating pulse reaches the aircraft's tail.

Implementing the $\Delta = 4$ cm hybrid FDTD-FE model discussed above requires 1.9 GB of memory and 8 hr execution time on a single processor of an HP Itanium2 cluster with 900-MHz processors. In comparison, the $\Delta = 4$ cm traditional FDTD model requires 0.56 GB of memory and 1 hr execution time on the same processor, and the high-resolution $\Delta = 1$ cm traditional FDTD model requires 23.4 GB of memory and 15 hr execution time on 12 such processors. It is clear that the hybrid FDTD-FE model is much more efficient than the high-resolution FDTD model.

We note that, for good efficiency of the hybrid FDTD-FE method, it is important to keep the FE region as small as possible. In the present example, approximately three layers of unstructured cells are used to obtain an accurate geometrical representation of the aircraft, while 4-cm FDTD bricks are used throughout the remainder of the computational domain.

19.7.3 Input Impedance of the Four-Arm Sinuous Antenna

The sinuous antenna represents a challenging problem for any analysis method. This is because of the very wide range in length scales from the outer boundary of the antenna into the feed region. The variation of the edge lengths required to mesh this type of antenna is often greater than 100:1, depending upon the spatial extent of the feed region, which controls the high-frequency performance of the antenna.

Sinuous antennas are typically driven to obtain right- or left-hand circular polarization. Coaxial feeds are applied to each arm, and are generally phased in increments of 90° with respect to one another. Four-, six-, and eight-arm sinuous antennas are commonly used in practice. Fig. 19.24 illustrates a four-arm design that has been modeled using the hybrid FDTD-FE technique. The use of a local unconditionally stable solution method is particularly beneficial in this example, because the time-step need not be reduced for numerical stability as the element size progressively becomes smaller into the feed region.

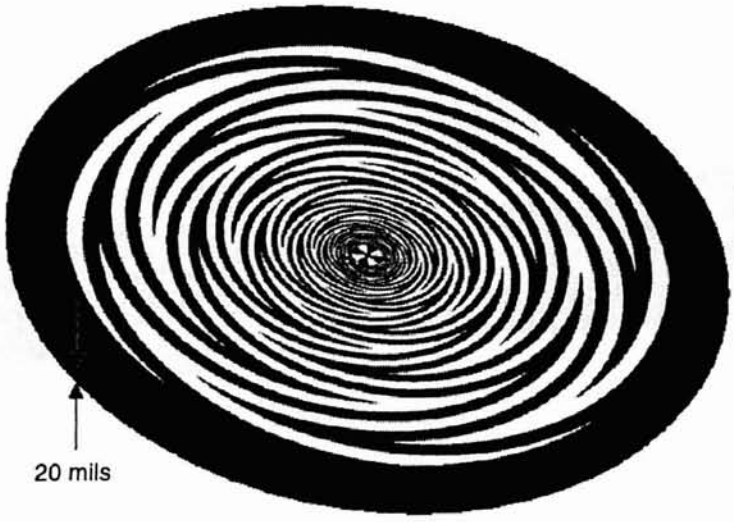


Fig. 19.24 Coaxial-fed four-arm sinuous antenna etched on a 20-mil-thick dielectric substrate.

Over the 4 to 8 GHz operating bandwidth of the antenna of Fig. 19.24, the hybrid FDTD-FE modeling results for the real and imaginary parts of the average active input impedance are in very good agreement with laboratory measurements: 84Ω predicted resistance versus 80Ω measured, and +4.5Ω predicted reactance versus 0Ω measured. We note that, for ideal four-arm self-complementary antennas in free space, the "sum-mode" active impedance is theoretically 133.3Ω. The presence of the thin-dielectric substrate in this example significantly affects the input impedance, and is properly accounted by the hybrid FDTD-FE model.

19.8 SUMMARY

This chapter presented a hybrid FDTD-FE technique for the numerical solution of Maxwell's equations that combines the computational efficiency of FDTD with the body-conforming meshing that is possible using FE. The hybrid method is free from spectral contamination and nonphysical numerical dissipation. It allows arbitrary unstructured meshes in the FE region, and yet is provably stable up to the FDTD stability limit. It also allows local spatial refinement of the FE mesh generally without the need to reduce the global time-step. Such refinement can be needed to resolve geometrical details that are much smaller than the wavelength, or to properly model field singularities near sharp corners, edges, or points.

In the hybrid technique, the FDTD "brick" elements are connected to the FE tetrahedrons by a layer of pyramids, and the E-field is expanded in edge elements to obtain a curl-conforming representation everywhere in the computational domain. The FE formulation is based on Galerkin's method, and trapezoidal integration is used for the structured bricks to recover the FDTD scheme formulated in terms of only the E-field. Consequently, the reciprocity of Maxwell's equations is preserved by construction. This is represented by the symmetric matrices that characterize the formulation of the hybrid method. Reciprocity reflects important properties of Maxwell's equations, and its preservation makes it feasible to formulate stable time-stepping schemes. Furthermore, it poses no restrictions on the materials that exist at the interface between the structured and unstructured grids [65]. Examples of other FDTD-FE hybridizations that preserve reciprocity and are provably stable include the subcell wire/slot algorithms and the coaxial waveguide ports presented in this chapter. These algorithms makes the hybrid FDTD-FE technique a suitable choice for many practical electromagnetic compatibility applications.

The hybrid FDTD-FE method has been tested on a broad range of applications, and has exhibited second-order convergence for problem geometries that possess smooth boundaries. As a specific example, the bistatic RCS of a PEC sphere was shown to converge to the analytical solution with $O(\Delta)^2$ error. However, as with other numerical methods, the order of convergence is lowered for geometries having sharp corners, edges, or points due to the presence of locally singular fields. As an example of applying the FDTD-FE hybrid to model a scatterer with a sharp tip, a computation of the monostatic RCS of the NASA almond demonstrated that a meshing density of 15 per wavelength yields results that are within 1 dB of the asymptotically converged values (i.e., infinitely fine spatial resolution) for at least 90% of the azimuth angles.

The scattering and antenna examples presented in this chapter demonstrate that the FDTD-FE hybrid provides reliable, accurate, and efficient computational results for real-world applications. Practical applications often possess a very wide range in local length scales that can be problematic for purely conditionally stable solution methods. The FDTD-FE hybrid method presented in this chapter can be used to analyze these structures often with no reduction in the time-step prescribed by the FDTD grid that surrounds the body-conforming FE mesh.

However, the convergence rate of an iterative solver is usually reduced in the case of widely varying edge lengths. Therefore, the development of improved matrix preconditioners is an active area of research. In addition, the choice of the time-step can influence the accuracy of the solution. The combination of a conditionally stable FDTD scheme with an unconditionally stable FE formulation provides for a reasonable limitation on the largest selectable time-step that balances both efficiency and accuracy.

ACKNOWLEDGMENTS

The authors acknowledge the assistance of Jonas Hamberg, Saab AB, Dr. Stefan Persson, AerotechTelub AB, Dr. Ulf Andersson, KTH, and Dr. Martin Nilsson, Uppsala University, in the preparation of the *Trainer* example. The assistance of Dr. Norma Riley, Northrop Grumman Corporation, is acknowledged in the preparation of the sinuous antenna example.

REFERENCES

[1] Holland, R., "THREDS: A finite-difference time-domain EMP code in 3D spherical coordinates," *IEEE Trans. Nuclear Science*, Vol. 30, 1983, pp. 4592–4595.

[2] Van Hese, J., and D. De Zutter, "Modeling of discontinuities in general coaxial waveguide structured by the FDTD method," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, 1992, pp. 547–556.

[3] Cangellaris, A., and D. B. Wright, "Analysis of the numerical error caused by the stair-stepped approximation of a conducting boundary in FDTD simulations of electromagnetic propagation," *IEEE Trans. Antennas Propagat.*, Vol. 39, 1991, pp. 1518–1525.

[4] Hughes, T. J. R. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1987.

[5] Zienkiewicz, O. C., and R. L. Taylor, *The Finite Element Method, 5th ed.*, Oxford: Butterworth-Heinemann, 2000.

[6] Jiang, B.-N., *The Least-Squares Finite Element Method*, Berlin: Springer-Verlag, 1998.

[7] Ferziger, J. H., and M. Peric, *Computational Methods for Fluid Dynamics*, Berlin: Springer-Verlag, 1996.

[8] Yee, K. S., "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, Vol. 14, 1966, pp. 302–307.

[9] Riley, D. J., and C. D. Turner, "VOLMAX: A solid-model based, transient volumetric Maxwell solver using hybrid grids," *IEEE Trans. Antennas Propagat. Mag*, Vol. 39, 1997, pp. 20–33.

[10] Jin, J., *The Finite Element Method in Electromagnetics, 2nd ed.*, New York: Wiley, 2002.

[11] Madsen, N., "Divergence preserving discrete surface integral methods for Maxwell's equations using nonorthogonal unstructured grids," *J. Comp. Physics*, Vol. 119, 1995, pp. 34–45.

[12] Nedelec, J. C., "Mixed finite elements in R^3," *Numer. Math.*, Vol. 35, 1980, pp. 315–341.

[13] Bossavit, A., "Whitney forms: A class of finite elements for three-dimensional computations in electromagnetism," *IEE Proc.*, Vol. 135, 1988, pp. 493–500.

[14] Yee, K. S., J. S. Chen, and A. H. Chang, "Conformal finite-difference time-domain (FDTD) with overlapping grids," *IEEE Trans. Antennas Propagat.*, Vol. 40, 1992, pp. 1068–1075.

[15] Madsen, N. K., and R. Ziolkowski, "A three-dimensional modified finite volume technique for Maxwell's equations," *Electromagnetics*, Vol. 10, 1990, pp. 147–161.

[16] Thompson, J. F., B. K. Soni, and N. P. Weatherill, eds, *Handbook of Grid Generation*, Boca Raton, FL: CRC Press, 1999.

[17] Yee, K. S., and J. S. Chen, "Conformal hybrid finite-difference time domain and finite volume time domain," *IEEE Trans. Antennas Propagat.*, Vol. 42, 1994, pp. 1450–1455.

[18] Yee, K. S., and J. S. Chen, "The finite-difference time-domain (FDTD) and the finite-volume time-domain (FVTD) methods in solving Maxwell's equations," *IEEE Trans. Antennas Propagat.*, Vol. 45, 1997, pp. 354–363.

[19] Riley, D. J., and C. D. Turner, "Interfacing unstructured tetrahedron grids to structured-grid FDTD," *IEEE Microwave Guided Wave Lett.*, Vol. 5, 1995, pp. 284–286.

[20] Kalfon, D., and P. Harran-Klotz, "A hybrid finite difference – finite volume method for solving Maxwell equations in time domain," *URSI Meeting*, Newport Beach, CA, 1995.

[21] Rao, S. M., (ed.), *Time Domain Electromagnetics*, London: Academic Press, 1999.

[22] Edelvik, F., and G. Ledfelt, "A comparison of time-domain hybrid solvers for complex scattering problems," *Intl. J. Numerical Modeling*, Vol. 15, 2002, pp. 475–487.

[23] Edelvik, F., *Hybrid Solvers for the Maxwell Equations in Time-Domain*, Ph.D. diss., Uppsala Univ., Uppsala, Sweden, 2002 (ISBN 91-554-5354-6).

[24] Wu, R. B., and T. Itoh, "Hybridizing FDTD analysis with unconditionally stable FEM for objects of curved boundary," *IEEE MTT-S Dig.*, Vol. 2, 1995, pp. 833–836.

[25] Wu, R. B., and T. Itoh, "Hybrid finite-difference time-domain modeling of curved surfaces using tetrahedral edge elements," *IEEE Trans. Antennas Propagat.*, Vol. 45, 1997, pp. 1302–1309.

[26] Hwang, C. T., and R. B. Wu, "Treating late-time instability of hybrid finite-element/finite-difference time-domain method," *IEEE Trans. Antennas Propagat.*, Vol. 47, 1999, pp. 227–232.

[27] Lee, J. F., and Z. Sacks, "Whitney elements time domain (WETD) methods," *IEEE Trans. Magnetics*, Vol. 31, 1995, pp. 1325–1329.

[28] Rylander, T., and A. Bondeson, "Stable FDTD-FEM hybrid method for Maxwell's equations," *Comput. Phys. Comm.*, Vol. 125, 2000, pp. 75–82.

[29] Coulomb, J. L., F. X. Zgainski, and Y. Maréchal, "A pyramidal element to link hexahedral, prismatic, and tetrahedral edge finite elements," *IEEE Trans. Magnetics*, Vol. 33, 1997, pp. 1362–1365.

[30] Zgainski, F. X., J. L. Coulomb, and Y. Maréchal, "A new family of finite elements: The pyramidal elements," *IEEE Trans. Magnetics*, Vol. 32, 1996, pp. 1393–1396.

[31] Thompson, L. L., and P. M. Pinsky, "A Galerkin least squares finite element method for the two-dimensional Helmholtz equation," *Intl. J. Numer. Methods Engineering*, Vol. 38, 1995, pp. 371–397.

[32] Wong, M. F., O. Picon, and V. F. Hanna, "A finite element method based on Whitney forms to solve Maxwell equations in the time domain," *IEEE Trans. Magnetics*, Vol. 31, 1995, pp. 1618–1621.

[33] Cohen, G. and P. Monk, "Gauss point mass lumping schemes for Maxwell's equations," *Numer. Methods Partial Differential Eq.*, Vol. 14, 1998, pp. 63–88.

[34] Rylander, T., and A. Bondeson, "Stability of explicit-implicit hybrid time-stepping schemes for Maxwell's equations," *J. Comput. Phys.*, Vol. 179, 2002, pp. 426–438.

[35] Rylander, T., *Stable FDTD-FEM Hybrid Method for Maxwell's Equations*, Ph.D. diss, Chalmers Univ. of Tech., Gothenburg, Sweden, 2002.

[36] Bossavit, A., and I. Mayergoyz, "Edge elements for scattering problems," *IEEE Trans. Magn.*, Vol. 25, 1989, pp. 1816–1821.

[37] Rao, S. M., D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, Vol. 30, 1982, pp. 409–418.

[38] Mur, G., "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations," *IEEE Trans. Electromagn. Compat.*, Vol. 23, 1981, pp. 377–383.

[39] Riley, D. J., and N. W. Riley, "First order models for thin-material sheets and coatings in the finite-element time-domain method," *IEEE Antennas Propagat. Intl. Symp. Dig.*, Monterey, CA, 2004.

[40] Gedney, S. D., and U. Navsariwala, "An unconditionally stable finite element time-domain solution of the vector wave equation," *IEEE Microwave Guided Wave Lett.*, Vol. 5, 1995, pp. 332–334.

[41] Axelsson, O., and P. Vassilevski, "Algebraic multilevel preconditioning methods, I," *Numer. Math.*, Vol. 56, 1989, pp. 157–177.

[42] Monk, P., and K. Parrott, "Phase-accuracy comparisons and improved far-field estimates for 3D edge element on tetrahedral meshes," *J. Comp. Phys.*, Vol. 170, 2001, pp. 614–641.

[43] Graglia, R. D., D. R. Wilton, and A. F. Peterson, "Higher order interpolatory vector bases for computational electromagnetics," *IEEE Trans. Antennas Propagat.*, Vol. 45, 1997, pp. 329–342.

[44] Graglia, R. D., and I.-L. Gheorma, "Higher order interpolatory vector bases on pyramidal elements," *IEEE Trans. Antennas Propagat.*, Vol. 47, 1999, pp. 775–782.

[45] Webb, J. P., "Hierarchal vector basis functions of arbitrary order for triangular and tetrahedral finite elements," *IEEE Trans. Antennas Propagat.*, Vol. 47, 1999, pp. 1244–1253.

[46] Bossavit, A., "Simplical finite elements for scattering problems in electromagnetism," *Comp. Meth. Appl. Mech. and Engineering*, Vol. 76, 1989, pp. 299–316.

[47] Rylander, T., and J.-M. Jin, "Perfectly matched layer for the time domain finite element method," *J. Comp. Phys.*, Vol. 200, 2004, pp. 238–250.

[48] Belytschko, T., and R. Mullen, "Stability of explicit-implicit mesh partitions in time integration," *Intl. J. Numerical Methods Engrg.*, Vol. 12, 1978, pp. 1575–1586.

[49] Rylander, T., and J.-M. Jin, "Stable coaxial waveguide-port algorithm for the time-domain finite-element method," *Microwave Optical Tech. Lett.*, Vol. 42, 2004, pp. 115–119.

[50] Alimenti, F., P. Mezzanotte, L. Roselli, and R. Sorrentino, "A revised formulation of modal absorbing and matched modal source boundary conditions for the efficient FDTD analysis of waveguide structures," *IEEE Trans Microwave Theory Tech*, Vol. 48, 2000, pp. 50–59.

[51] Riley, D. J., "Transient finite-elements for computational electromagnetics: Hybridization with finite differences, modeling thin wires and thin slots, and parallel processing," *17th Ann. Review Progress Appl. Comp. Electromagn. (ACES)*, Monterey, CA, 2001, pp. 128–138.

[52] Balay, S., W. D. Gropp, L. C. McInnes, and B. F. Smith, *PETSc Home Page*, [online] http://www.mcs.anl.gov/petsc.

[53] *ICEM-CFD*, ICEM CFD Engineering, Berkeley, CA.

[54] Holland, R., and L. Simpson, "Finite-difference analysis of EMP coupling to thin struts and wires," *IEEE Trans. Electromagn. Compat.*, Vol. 23, 1981, pp. 88–97.

[55] Edelvik, F., G. Ledfelt, P. Lötstedt, and D. Riley, "An unconditionally stable subcell model for arbitrarily oriented thin wires in the FETD method, *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 1797–1805.

[56] Edelvik, F., and T. Weiland, "Stable modelling of arbitrary thin slots in the finite-element time-domain method," *Intl. J. Numerical Modelling*, Vol. 17, 2004, pp. 365–383.

[57] Warne, L. K., and K. C. Chen, "Equivalent antenna radius for narrow slot apertures having depth," *IEEE Trans. Electromagn. Compat.*, Vol. 37, 1989, pp. 824–834.

[58] Warne, L. K., and K. C. Chen, "A simple transmission line model for narrow slot apertures having depth and losses," *IEEE Trans. Electromagn. Compat.*, Vol. 34, 1992, pp. 173–182.

[59] Ma, K.-P., M. Li, J. L. Drewniak, T. H. Hubing, and T. P. van Doren, "Comparison of FDTD algorithms for subcellular modeling of slots in shielding enclosures," *IEEE Trans. Electromagn. Compat.*, Vol. 39, 1997, pp. 147–155.

[60] Woo, A. C., H. T. G. Wang, M. J. Schuh, and M. L. Sanders, "Benchmark radar targets for the validation of computational electromagnetics programs," *IEEE Antennas Propagat. Mag.*, Vol. 35, 1993, pp. 84–89.

[61] Bluck, M. J., and S. P. Walker, "Time-domain BIE analysis of large three-dimensional electromagnetic scattering problems," *IEEE Trans. Antennas Propagat.*, Vol. 45, 1997, pp. 894–901.

[62] Rylander, T., and A. Bondeson, "Application of stable FDTD-FEM hybrid to scattering problems," *IEEE Trans. Antennas Propagat.*, Vol. 50, 2002, pp. 141–144.

[63] de Smedt, R., "Electric singularity near the tip of a sharp cone," *IEEE Trans. Antennas Propagat.*, Vol. 36, 1988, pp. 152–155.

[64] Nilsson, M., *Fast Numerical Techniques for Electromagnetic Problems in Frequency Domain*, Ph.D. diss., Uppsala Univ., Uppsala, Sweden, 2003.

[65] Edelvik, F., and B. Strand, "Frequency dispersive materials for 3D hybrid solvers in time domain," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2003, pp. 1199–1206.

Chapter 20

Advances in Hardware Acceleration for FDTD

Ryan Schneider, Sean Krakiwsky, Laurence Turner, and Michal Okoniewski

20.1 INTRODUCTION

Over the past 40 years, the FDTD computational solution of Maxwell's equations has evolved from a fringe concept to an actively pursued, mainstream technique in academic, industrial, and government laboratories. Today, FDTD is a tool of choice for electromagnetic simulations in such diverse areas as wireless technologies, computer hardware, photonics, lithographic mask design, biomedicine, and biophotonics. The current popularity of FDTD has been facilitated by phenomenal advances in *personal computer* (PC) technology. No longer need a designer send a simulation job to a supercomputer; it can be run on a desktop PC equipped with superb graphical and processing capabilities. However, as PCs become more powerful, FDTD modelers are emboldened to simulate ever more complex and realistic structures. It seems that the problems of *real* interest always require 10 times the memory and speed of the best available machines.

Because FDTD models (especially in three dimensions) generally require large allocations of fast memory and significant run times, the use of FDTD to date has been primarily limited to problem analysis, rather than design through optimization. In the basic formulation of FDTD, memory requirements scale as $O(N)$, where $N = N_x \times N_y \times N_z$ space cells; and the required time-step count scales approximately as $O(N^{1/3})$ [1, 2]. This translates to $O(N^{4/3})$ floating-point operations (flops) for a given simulation. It is easy to show that FDTD simulations involving tens of millions of grid cells require gigaflops of processing capability.

When the first edition of this book was published in 1995, supercomputers and high-performance workstations capable of such demands served an important segment of the FDTD simulation market, and the development of software for such systems was an important topic [1, 2]. However, for most industrial users, supercomputers carried prohibitive access costs, limited availability, and lack of commercial software. Therefore, the past decade has seen the migration of most commercial and academic FDTD codes from supercomputers onto generic PCs because of their multipurpose functionality and low cost. However, today's PCs have limited memory capacity (2 GB for 32-bit systems), limited memory bandwidth, and limited speed, even for those machines featuring two processors. In an attempt to address these limitations, researchers have organized PCs into clusters that divide an FDTD simulation among many nodes [3, 4]. Here, the main disadvantage is that performance gains saturate with increasing numbers of processors in the clusters due to the interprocessor communications overhead. Furthermore, FDTD cluster implementations add additional software complexity and portability issues, as well as the need for potentially costly cluster maintenance.

Ultimately, whether running simulations on a generic PC or a multiprocessor cluster, many FDTD users experience lengthy run times. In response, and noting the inherent parallelism, symmetry, and simplicity of the FDTD algorithm, a number of researchers are pursuing run-time reduction via implementing FDTD in custom hardware. This activity leverages favorable industry trends, wherein the speed and complexity of special purpose hardware is rising more rapidly than that of general purpose processors.

In this chapter, we discuss the concept of hardware acceleration of the FDTD method. Two distinct techniques are considered:

- Construction of special-purpose hardware using *field-programmable gate-arrays* (FPGAs). Here, Maxwell's equations are actually rendered in logic gates.
- Use of available *graphics processor units* (GPUs). Here, Maxwell's equations are expressed in a graphical language such as OpenGL.

Both approaches have succeeded in demonstrating significant accelerations of large, nontrivial FDTD simulations. Accelerations approaching 10:1 (as compared with fully optimized FDTD code running on a top-of-the-line PC) should be possible in the near future.

FPGA-based hardware is currently more costly than GPU-based acceleration because special-purpose cards need to be built and populated with large and expensive FPGA chips. While FPGA accelerators are capable of handling large memory, they are limited as to the type of FDTD algorithm features that can be incorporated. On the other hand, GPU accelerators are currently based on mass-produced, consumer graphics cards. While having limited memory, GPUs provide more flexible implementation than FPGA hardware, allowing advanced FDTD features to be readily implemented. The future will likely see a hybrid approach that combines the benefits of both techniques.

20.2 BACKGROUND LITERATURE

Research in hardware-acceleration concepts for FDTD has blossomed since the early 1990s. The following papers mark key developments in this area:

1. As early as 1992, [5] reported the simulation of a *very large-scale integration* (VLSI) circuit that could accelerate FDTD updates by 5:1.
2. References [6–8] reported *hardware description language* (HDL) simulations of two different VLSI designs achieving FDTD updating rates of 6 to 10 million cells per second. Both assumed VSLI *application-specific integrated circuits* (ASICs) designed to connect to the host computer via the *peripheral component interconnect* (PCI) bus.
3. Reference [9] reported a hardware implementation of a two-dimensional FDTD engine. This implementation was a bench prototype, with handwired registers and arithmetic units. Although its throughput was slower than running FDTD software on a general-purpose processor, it was predicted that the implementation could update 6.25 million cells per second for a system clock of 100 MHz.

4. References [10, 11] reported the first FDTD accelerators implemented in commercially available hardware. These accelerators employed bit-serial integer arithmetic and an XESS development board [12] with a Xilinx XCV300 FPGA (3,072 slices). An updating throughput of 8 million two-dimensional grid cells per second was achieved.
5. References [13–15] reported what appear to be the first successful three-dimensional FDTD accelerators to be implemented in physical hardware. Here, front-end software sends appropriate data (such as the mesh size and the number of time-steps) to the hardware via a PCI bus. The FDTD accelerator proceeds to update the fields, periodically sending results back to the host computer for post-processing and visualization. In the most recent work of this group, updating throughputs as high as 30 million three-dimensional cells per second were achieved using a custom circuit board and FPGA implementation.

We see that early research into FDTD hardware accelerators focused on simulation of the hardware architecture, whereas more recent work has achieved successful hardware implementations. Currently, progress is steady in the latter area, with commercial products already available [16, 17].

20.3 FUNDAMENTAL DESIGN CONSIDERATIONS

Reference [14] outlines tradeoffs that must be considered during the design of an FDTD accelerator system. In the analysis, it is assumed that the FDTD accelerator is a peripheral device in a host computer. The host computer is responsible for setting up the FDTD problem and for the communication of the data to and from the accelerator. On the other hand, the accelerator is responsible for the FDTD updates at every time-step and for periodically sending results for postprocessing. Three major areas of concern are identified which must be considered in the design of an FDTD accelerator: (1) the hardware computational engine, (2) the data storage, and (3) the treatment of special nodes.

A hardware computational engine replaces the field-update procedures conventionally implemented in software. The acceleration potential of the hardware is related to the number of FDTD cells that can be updated simultaneously. This yields two extremes. The first is a massively parallel hardware engine that is capable of simultaneously computing all of the field updates in parallel. Achieving this high level of acceleration is not feasible for reasonably sized simulations, since it would require an impractical amount of hardware. At the other extreme is single-field-update (serial-processing) hardware. While simple and inexpensive to implement, this would not make use of the parallelism inherent in the FDTD algorithm, and would suffer from too substantial a communication overhead to provide a significant speedup. Current research trends indicate the desirability of a hybrid approach, wherein enough hardware is provided to exploit the FDTD parallelism to achieve significant acceleration while keeping the hardware cost reasonable.

With respect to data storage, key issues include the amount of data to be stored on the accelerator, the host-accelerator interconnect speed, and the amount of data that must be transferred to and from the host computer. A central problem stems from the fact that while the host computer offers excellent data-storage solutions, the bus speed between the host and the accelerator is slow—as much as an order of magnitude slower than the memory bus—even with the recently introduced PCI Express standard. Ideally, the accelerator would have a large amount of fast memory sufficient to store the entire data space of the problem. This memory would be connected to the update hardware through special-purpose, high-throughput buses. In cases where intermediate data values are not required, as in some steady-state analysis problems, the onboard memory would completely eliminate the overhead of the host-accelerator data transfer. While such large onboard memory is expensive, a compromise solution, wherein a significant fraction of the field data values are stored in the accelerator, would still allow for a useful reduction of the overhead associated with the slower data transfer rate.

The treatment of special FDTD field updates involving sources, subcell models, absorbing boundary conditions, lumped-circuit elements, and so forth, complicates hardware design. Hardware capable of handling all types of field updates would be expensive and difficult to implement, and would eventually rival the complexity of a standard central processing unit. Moreover, it would be very challenging to incorporate new special-purpose field updates introduced by the FDTD research community. We observe, however, that special cells typically occupy only a small fraction of the computational domain. Thus, although the overhead associated with data transfer to and from the accelerator degrades the overall accelerator performance, it may be best to implement most of the special field updates via software in the host computer.

References [10] and [13] explored the "massively parallel" approach to FDTD acceleration. The next section describes the derivation and implementation of such an FDTD accelerator system and computational engine, where every field-update equation is implemented in an individual hardware component. The strong advantage of this approach is that the computation time for a single field update is constant, while the acceleration achieved is directly proportional to the number of cells in the simulation. Thus, larger simulations achieve greater accelerations but require more hardware. In addition, in this particular implementation, the FDTD field updates are implemented using fixed-point arithmetic.

20.4 CONCEPTUAL MASSIVELY PARALLEL FPGA IMPLEMENTATION

Historically, FDTD originated from the field-theory point of view. However, as early as the 1980s, [18] reported an equivalent lumped-circuit representation of FDTD. As shown in Fig. 20.1, for the lossless, two-dimensional TM case, this representation is simply a network of inductors and capacitors similar to a filter. Here, the capacitors and inductors represent energy storage in the electric and magnetic fields, respectively. Through impedance scaling, there are direct relationships between the capacitor and inductor values and the dielectric and magnetic material properties of the FDTD grid.

For discussion purposes, it is easier to analyze the one-dimensional FDTD update equations, a special case of Fig. 20.1. (The same concepts apply to three-dimensional FDTD updates, albeit the interconnected meshes of inductors and capacitors for that case are more difficult to graphically depict.) Fig. 20.2 illustrates the lumped-circuit equivalence and voltage / current signal flow graph for the one-dimensional FDTD grid cell.

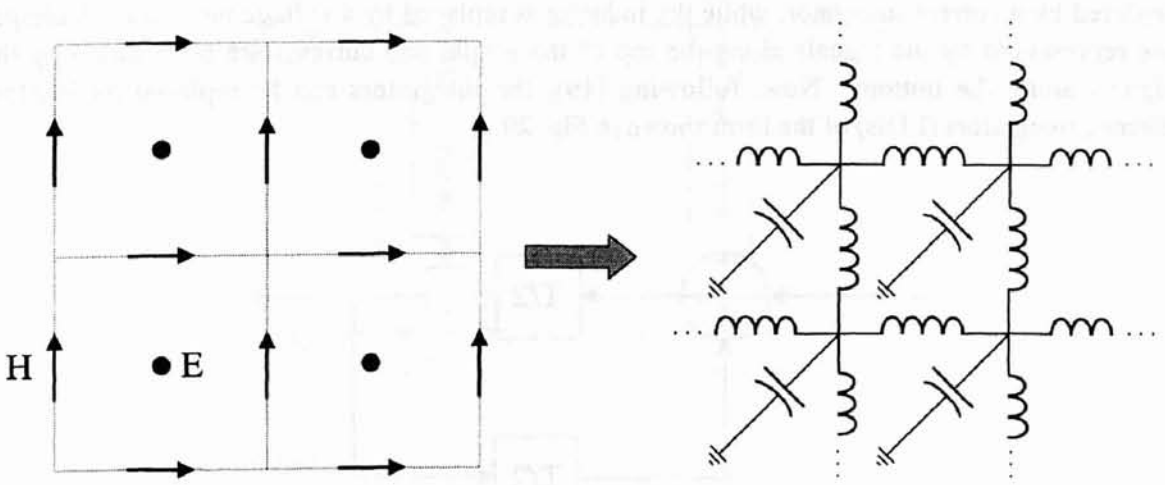
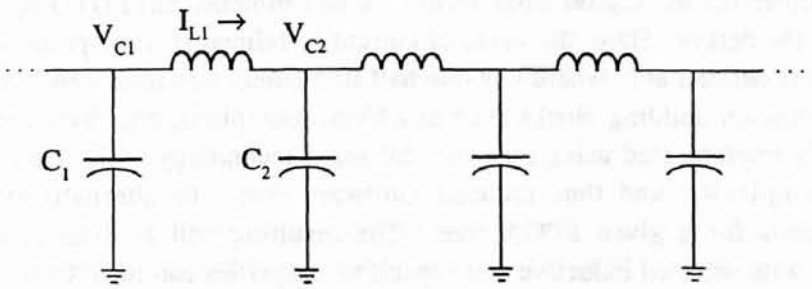
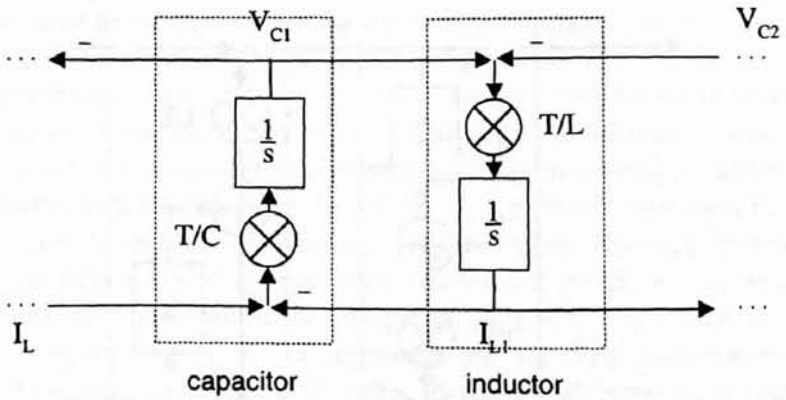


Fig. 20.1 Two-dimensional TM FDTD grid represented as an inductor-capacitor network [18].



(a) Lumped-circuit equivalent of one-dimensional FDTD grid cells.



(b) Voltage / current signal flow graph for a single FDTD grid cell.

Fig. 20.2 One-dimensional FDTD grid-cell representations.

Note that, in Fig. 20.2(b), the $1/s$ term denotes Laplacian integration. Here, the capacitor is replaced by a current integrator, while the inductor is replaced by a voltage integrator. Voltages are represented by the signals along the top of the graph, and currents are represented by the signals along the bottom. Now, following [19], the integrators can be replaced by lossless discrete integrators (LDIs) of the form shown in Fig. 20.3.

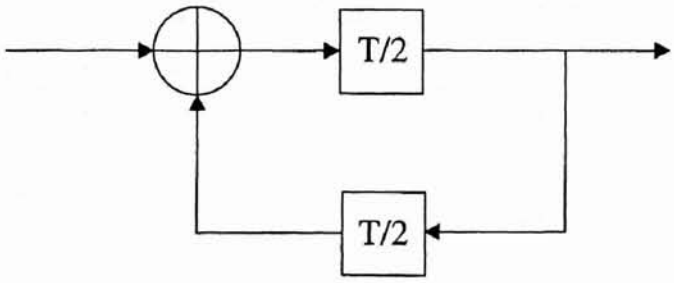


Fig. 20.3 Lossless discrete integrator.

Fig. 20.4 illustrates the digital filter form of a one-dimensional FDTD update after some manipulation of the delays. Here, the value of current is defined at time-point $k + 1$, while the value of voltage is defined at k, where k is one-half of a simulation time-step. We note that this filter contains common building blocks such as adders, multipliers, registers, and delays, which can be efficiently implemented using pipelined bit-serial technology [10]. This has a relatively low level of complexity, and thus reduced hardware cost; or alternatively, allows more computational cells for a given FPGA size. The resulting cell is illustrated in Fig. 20.5. Many such cells with selected inductive and capacitive properties can be interconnected to create a large-scale, massively accelerated FDTD simulation.

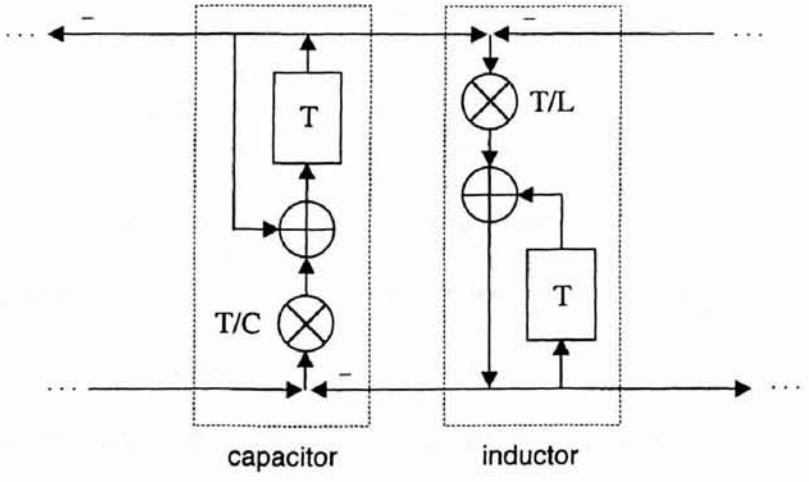


Fig. 20.4 Digital filter form of a one-dimensional FDTD update.

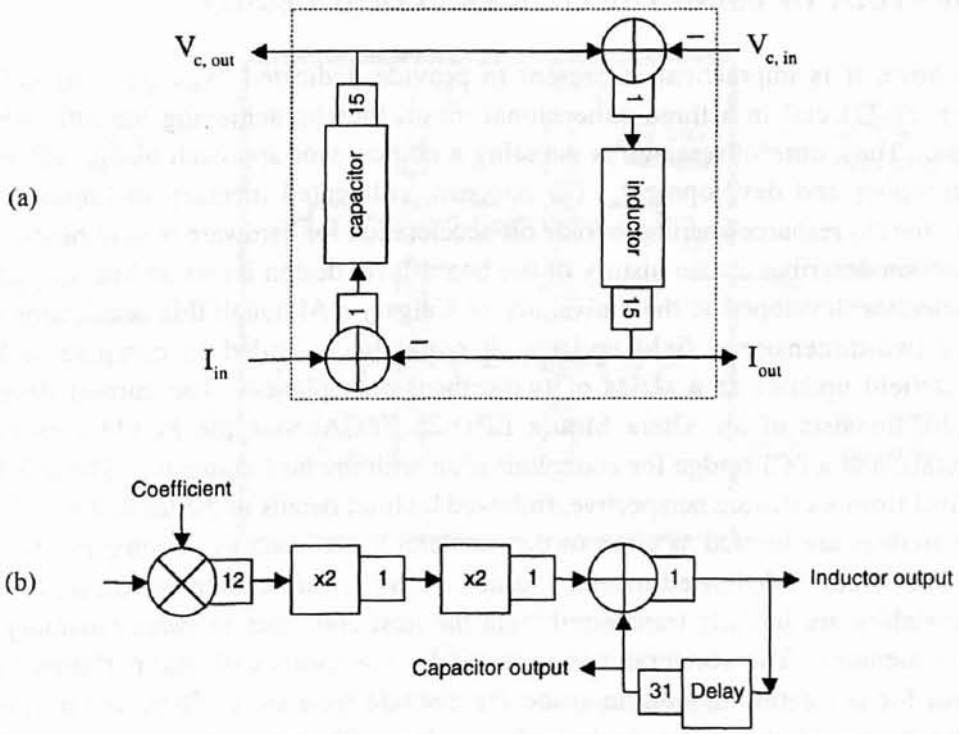


Fig. 20.5 (a) One-dimensional, bit-serial FDTD cell (32-bit system word length); (b) structure for the capacitor and inductor.

The design in Fig. 20.5 uses signed integers with a 32-bit system word length and 12-bit coefficients. The boxes following each block represent the delay through that block. Control signals are distributed around the circuit to mark the arrival of the *least significant bit* (LSB) at each point in the loop. Each delay loop in Fig. 20.5 is 32 bits (system word length) long. The capacitor's delay is distributed between its adder and the rest of the inductor/capacitor loop, requiring 31 bits of delay in the feedback path. The inductor's delay represents the desired system word length delay before it is added back into the data path. The multipliers are followed by a multiply by four, which is used to change the range of coefficients (magnitude larger than one) that can be represented. Due to the symmetrical nature of the design and calculation of the two fields, the structures are identical. It is expected that this is not always the case.

As a conceptual example, consider implementing this design using a current Xilinx FPGA, which has a maximum clock frequency of 37.7 MHz. Each two-dimensional FDTD cell would require 120 Virtex slices (a measure of hardware usage in Xilinx FPGAs). Operating at a serial clock of 32 MHz and with a 40-bit system word length, new results would be available every 1.25 μs (i.e., 100,000 iterations would be completed in 0.125 sec). Each three-dimensional FDTD cell would require 265 slices to represent the six field components, and have a comparable calculation rate. Here, a $100 \times 100 \times 100$-cell FDTD model would require 4,300 FPGAs to update its 6×10^6 field components in parallel. While this scale of hardware is currently impractical, progressive increases in semiconductor densities could make this approach feasible in the future. A nearer-term option could be to implement the required hardware as a smaller collection of ASICs, since the programmability of an FPGA requires dedication of a significant portion of its available silicon area.

20.5 CASE STUDY OF USING THE FPGA AS A COPROCESSOR

As noted above, it is impractical at present to provide dedicated logic gates to individually update each FDTD cell in a three-dimensional mesh, thereby achieving the ultimate parallel computation. Thus, current research is pursuing a coprocessor approach using: (1) FPGAs to speed prototyping and development, (2) onboard, collocated memory to improve memory bandwidth, and (3) resource-sharing to trade off acceleration for hardware requirements.

This section describes a case history of the board-level design issues and results for such an FDTD accelerator developed at the University of Calgary. Although this accelerator currently implements two-dimensional field updates, it could be extended to compute fully three-dimensional field updates as a series of two-dimensional planes. The current development platform [20] consists of an Altera Stratix EP1S25 FPGA, multiple PC133 memory banks (128 MB total), and a PCI bridge for communication with the host computer. The accelerator is now described from a software perspective, followed by brief details of the hardware.

Field variables are located as close to the accelerator hardware as possible by storing these data in the accelerator's dedicated memory banks. A typical simulation is executed as follows. The field variables are initially transferred from the host computer software / memory into the accelerator's memory. The accelerator is activated by a software call, and performs the FDTD field updates for the entire simulation space for a single time-step. Then, the host computer software and the accelerator interactively exchange data. Observations and requested field data are transferred to the host computer, and the local copy of the simulation data is updated. Fields modified by the host computer software are transferred back to the accelerator memory and then another time-step is initiated. This process is repeated until the simulation is complete.

We note that only data accessed or recorded by the host PC is transferred from the accelerator to the host. This saves time between iterations, and works within the limited-bandwidth constraint of the PCI channel. Interaction with the host PC allows important, nonstandard FDTD computations (i.e., excitations, absorbing boundaries, and so forth) to be performed in special regions and updated to the accelerator between time-steps. As long as the regions requiring nonstandard computations are small compared with the overall simulation space, acceleration is still feasible. Furthermore, the host PC is free to perform its own computations in parallel with the hardware accelerator, which provides further acceleration.

Conceptually, the hardware architecture consists of three levels. At the first (lowest) level is the implementation of a single FDTD field-update equation. This is followed by, at the second level, the connection of many FDTD update equations in parallel to form a "window" of $M \times N$ two-dimensional FDTD cells. Finally, at the third (highest) level, a supervisory state machine transfers data between the computation engine and local cache, and between local cache and external memories. Each level is described briefly in the following paragraphs.

At the first level, a generic Yee update for an $\boldsymbol{E}$ or $\boldsymbol{H}$ component (see Chapter 3, Section 3.6.4) is implemented as a hardware circuit. Here, the computer word length (precision) of the field component and the updating coefficients can be specified independently. Data are also double-buffered at the inputs to the computation circuit. This two-register architecture allows simulation data to be concurrently loaded from, and stored to, memory while FDTD computations are being performed. New input values and results are exchanged between iterations by swapping the two registers. Thus, the load / store of the structure is overlapped with the computation of updated fields.

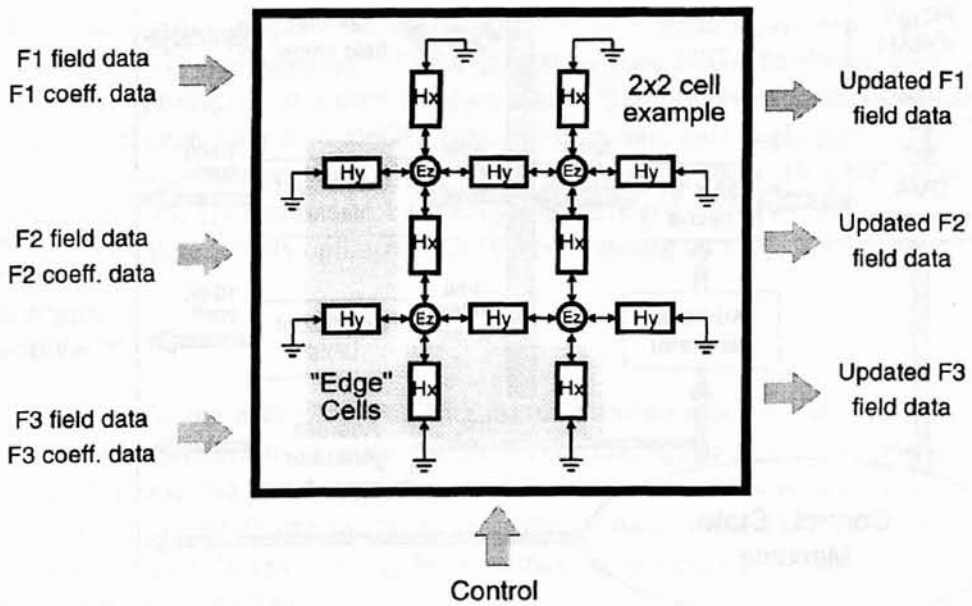


Fig. 20.6 Connection of $M \times N$ two-dimensional FDTD cells to construct a "window."

At the second conceptual level, illustrated in Fig. 20.6, three of the FDTD update equations are combined to form two-dimensional cells, which are then interconnected to form an $M \times N$ cell simulator. This smaller window of cells is moved throughout the larger simulation space until all fields are updated.

The current platform implements a 3×3 window of cells, due to the finite size of the FPGA. Most internal field values are loaded once per update sweep, except for edge values, which are loaded twice. Thus, the larger the accelerator window, the better the ratio of internal cells to edge cells. While an 8×8 window size is more desirable, beyond a crossover point, the load/store time exceeds that for computation. These must be balanced for optimum acceleration.

The third and final conceptual level, illustrated in Fig. 20.7, controls the flow of data into and out of the computation engine. From this figure, we see that data are moved from off-chip, SDRAM memory to an internal cache, which is then accessed to exchange data with the FDTD "window" described earlier. Another strategy of note is the use of coefficient lookup tables located in the accelerator to eliminate the memory bandwidth required for coefficients. The lookup tables are loaded at the start of the simulation and can be updated between time-steps, if necessary. Finally, it is possible to duplicate a number of the described memory channels to exchange data with the computation engine as quickly as possible.

The FDTD accelerator described above has achieved a throughput of 75 million two-dimensional cells per second. This exceeds the performance of FDTD software run on a top-of-the-line Pentium 4 system. However, FPGA-based hardware accelerators such as this one have one key drawback: While they can be extremely well-optimized to perform a certain type of FDTD update, they are rather inflexible in dealing with the special updates that may be required in a given simulation. For example, changing from a simple dielectric update to a Debye dispersion or a thin-wire special cell would require a different hardware configuration. Although on-the-fly hardware reconfiguration is possible, the associated overhead is currently excessive.

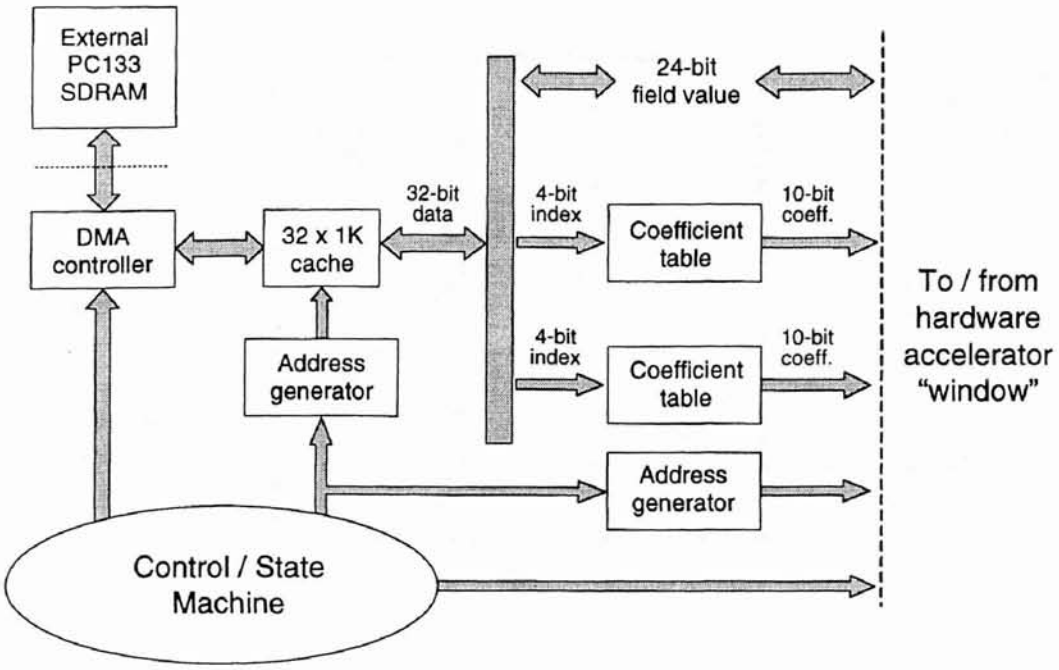


Fig. 20.7 Block diagram of memory channel to transfer simulation data into / out of computation engine.

20.6 PERFORMANCE OF CUSTOM HARDWARE IMPLEMENTATIONS

We now summarize the reported performance (as of January 2005) of several FDTD hardware accelerator implementations. The figure of merit is the number of Yee cells updated per second, with the caveat that the two-dimensional implementations perform three field updates per cell while the three-dimensional implementations perform six field updates per cell. All comparisons have been obtained when modeling free-space, metal-box resonators of varying sizes, and some implementations may optimize coefficients. The semi-optimized software performance was achieved via compiler optimization without specific tailoring for streaming / vectorized instruction sets or cache performance.

For the two-dimensional implementations, [9] reported an update rate of 6.25 million cells per second (6.25 Mcells/sec). Subsequently, [21] reported 13.79 Mcells/sec, and [22] reported 75 Mcells/sec.

For the three-dimensional implementations, [6–8] reported 6.25 Mcells/sec, [15] reported 30 Mcells/sec, and [17] reported 150 Mcells/sec. By way of comparison, semi-optimized three-dimensional software running on a 2.8-GHz Pentium-4 with DDR 400 memory benchmarked at 22.54 Mcells/sec, and several commercial three-dimensional software packages (including Empire and SEMCAD) benchmarked at 35 to 40 Mcells/sec when run on the same computer .

The results of the research to date on FDTD hardware accelerators are very promising. It is clear that these accelerators can surpass the performance of a top-end Pentium-4 processor. Therefore, it is reasonable to expect that specialized FDTD acceleration processors will become commonplace in the next decade. The next section describes another potential approach to accelerating FDTD with slightly different hardware.

20.7 FUNDAMENTALS OF GRAPHICS PROCESSOR UNITS

In the remainder of this chapter, we discuss the use of *graphics processor units* (GPUs) as accelerators for FDTD computations. Unlike special-purpose FPGA hardware, GPU accelerators are based on mass-produced, consumer graphics cards. Here, Maxwell's equations are expressed in a graphical language such as OpenGL, rather than in hardware logic gates. While currently having more limited memory than FPGAs, GPUs provide more flexible implementation, allowing advanced FDTD features to be readily implemented. This is leading to accelerations approaching 7:1 relative to fully optimized FDTD code running on a top-of-the-line PC [23].

20.7.1 Overview

Over the past decade, the need for processing massive amounts of data has exposed a weakness of conventional von Neumann computer architectures: inadequate memory bandwidth. While in the 1980s and 1990s the number of "instructions per second" was the metric of interest, the current need is to quickly process many gigabytes of data. Mass-market PC architectures are not well-suited for such data processing, because they are designed for generality (i.e., to support operating systems, for example).

GPUs are actually stream computers [24] optimized for parallelism and memory bandwidth. An example is the *n*VIDIA 6800 series GPU with 220 million transistors (the largest mass-marketed processor as of its debut in June 2004). The ability to exploit parallelism comes from constraining the functionality of the processor to support only the needs of graphics applications. Therefore, not only are GPUs the most powerful widely available processors, but because of their constrained functionality, their computational power is concentrated in fewer functional blocks. Currently, GPUs are being designed with dramatically increasing memory bandwidths (greater than 35 GB/sec on high-end boards) to satisfy developers in the graphics community who would otherwise be memory-bound, rather than computation-bound.

Precise architectural descriptions of commercial GPUs are proprietary to the manufacturers. Therefore, exact specifications analogous to what are available for microprocessors are not in the public domain. However, enough of a GPU's architecture is revealed through vendor-neutral program interfaces to understand how it can be used as a fast-streaming FDTD coprocessor.

20.7.2 Graphics Pipeline

The graphics card architecture is referred to as the graphics pipeline. The word "pipeline" here is used at a much higher level than when referred to by microprocessor documentation. It simply refers to how data are processed when streamed from an originating location in memory to a destination in memory, called the render target.

Fig. 20.8 is a high-level view of the graphics pipeline of a typical modern graphics card. The interface between the *central processor unit* (CPU) and the GPU is shown on the left of the figure. The application running on the system CPU is typically a C program using one of two *application programmable interfaces* (APIs): OpenGL or Microsoft's Direct X. These APIs allow the programmer to store data and invoke programs on the graphics card. Currently, the physical interface to the graphics card is the *accelerated graphics port* (AGP) bus. In the near future, the new PCI express bus [25] will constitute a much faster interface, with multiple GPUs per motherboard supported.

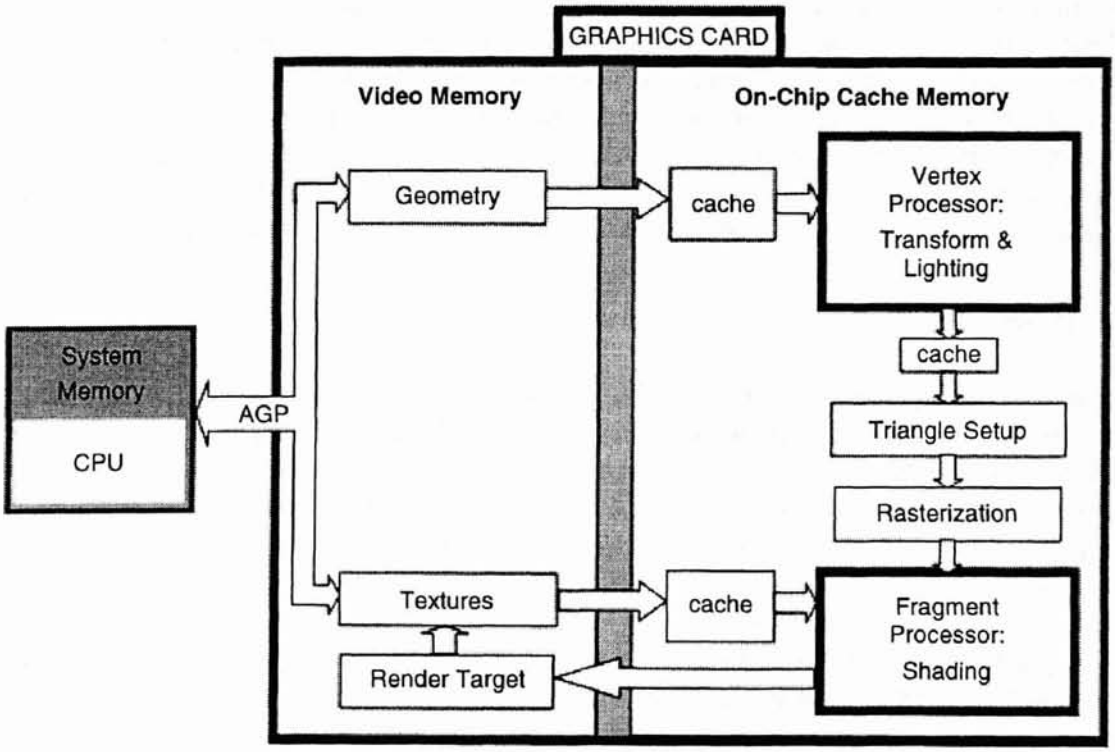


Fig. 20.8 Graphics pipeline.

In a typical graphics program, there are several geometric models consisting of vertices (and their attributes) making up various shapes, such as balls, cubes, cars, tables, and so forth. As these shapes move and change scale, geometric transformations are required. In the graphics card, the vertex processor is responsible for performing these geometric transformations in real time. These transformations tend to be mostly operations such as 4×4 matrix multiplies of four-element vectors. The vertex processor is a programmable *multiple-instruction multiple-data* (MIMD) stream processor that is optimized to do these kinds of instructions rapidly. The vertex processor's job is to accept a stream of vertices, process them, and output a transformed stream of vertices and their attributes, such as primary color, texture coordinate, and fog.

The next stage in the pipeline is the triangle setup. This stage, which is not programmable, accepts the stream of vertices and turns them into triangles. The rasterizer then takes the triangles and turns them into fragments. Fragments are defined as potential pixels with red, green, blue, and translucency components. They are called potential pixels, since they may or may not become real pixels resident in the frame buffer displayed on the screen. Subsequent stages can change the colors of these fragments, or remove them from the scene completely. Fig. 20.9 illustrates the function of the triangle setup and rasterization stages. This figure shows the vertices of an octagon being changed into a set of triangles and ultimately rasterized into fragments.

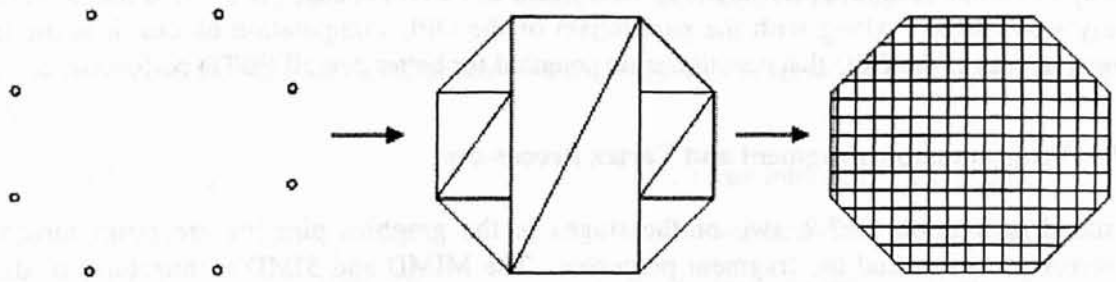


Fig. 20.9 From vertices to triangles then fragments.

The next stage identified in Fig. 20.8 is the programmable fragment processor. This *single-instruction multiple-data* (SIMD) streaming processor accepts the stream of fragments and changes their color, based on instructions of the fragment program that include fetches to textures resident in video memory. The fragment program instructions may be as simple as replacing the fragment color with the color from a particular texture, or may be more complicated to achieve a special visual effect.

There are several other hardware stages in the pipeline such as fog, alpha translucency tests, shadowing, antialiasing, and depth buffering. However, these are not relevant to our nongraphics use of the GPU, and are therefore not discussed. The fragments modified by the fragment processor that survive subsequent stages are written to the render target and can then be displayed on the screen. However, in some applications, these data are not displayed but are instead bound as a texture to be used in subsequent passes as source data. This multipass rendering technique is used in our implementation of FDTD, as discussed in Section 20.8.6.

20.7.3 Memory Interface

The memory interface between onboard DDR chips and the GPU chip is a critical contributor to the fast-streaming capabilities of graphics cards. The memory interfaces for two high-end *n*VIDIA-based cards and one ATI board are given in Table 20.1 below.

TABLE 20.1

Memory Interface Characteristics

Characteristic	*n*VIDIA GeForce 6800	*n*VIDIA GeForce FX5950 Ultra	ATI Radeon 9800XT
Bus width	256 bit	256 bit	256 bit
Memory clock speed	1.1 GHz DDR-3	950 MHz DDR-2	730 MHz DDR-2
Peak memory bandwidth	35.2 GB/s	30.4 GB/s	23.4 GB/s

By way of comparison, the peak memory bandwidth of current 3-GHz CPU-based motherboards is only ≈6 GB/sec. Along with the parallelism of the GPU computation blocks, it is the fast memory access of the GPU that constitutes its potential for better overall FDTD performance.

20.7.4 Programmable Fragment and Vertex Processors

As stated in Section 20.7.2, two of the stages in the graphics pipeline are programmable: the vertex processor and the fragment processor. The MIMD and SIMD architectures of these processors are inspired in part by data-flow diagrams and systolic arrays [24]. Therefore, the focus has been on developing chips with high throughput while sacrificing generality. This approach yields processors that are highly optimized for stream computations. The left side of Fig. 20.10 shows the program model for modern programmable fragment processors.

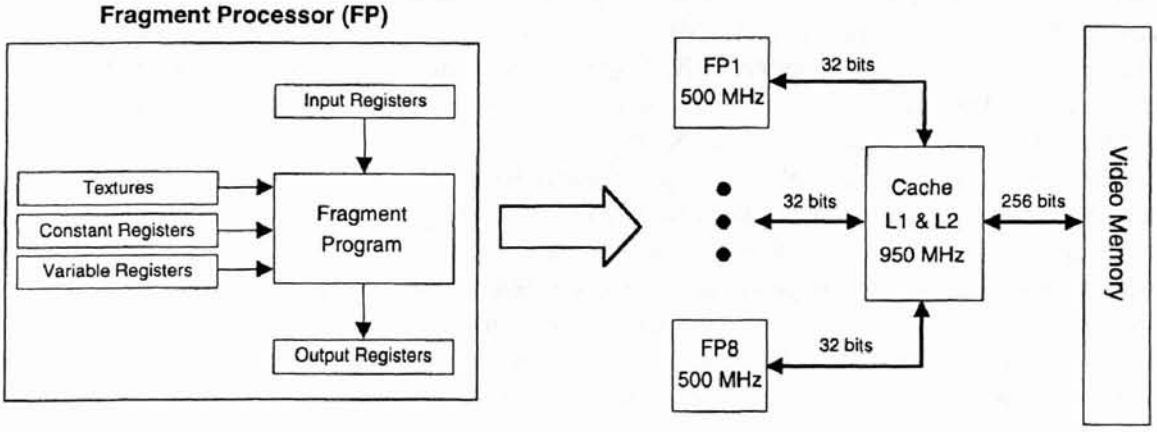


Fig. 20.10 Programming model for fragment processor.

The *n*VIDIA GeForce FX 5950 Ultra has eight fragment processors operating in parallel. Therefore, eight fragments can be processed simultaneously, each one executing the same instructions defined in a fragment program. Fragment program instructions are user-definable via OpenGL or Direct X APIs. A fragment program has access to a set of input registers and must output to at least one output register (shaded color for fragment), and may write to more output registers constituting fragment attributes. In order to calculate the final value for the output register(s), the fragment program can access textures from texture memory (via cache), a set of constants, and variables that can be set from the main C program each pass.

If a fragment program attempts to execute a texture fetch, the cache receives eight requests simultaneously. If the data is in the cache, then all fragment programs receive their data in one clock cycle. However, if the data is not in cache, then a stall occurs. This caching is critical to GPU performance for data processing applications. It is not clear how this caching works and what must be done to exploit it, as the documentation is not available. The standard explanation given by *n*VIDIA and ATI as to why they do not release their caching specifications is that the architecture changes significantly with every release, and they are leery of developers designing for a specific architecture. Therefore, the cache architecture and design needs to be determined by experimentation on a chipset-by-chipset basis [26].

20.8 IMPLEMENTING FDTD ON A GRAPHICS PROCESSOR UNIT

The main tasks in the GPU implementation of FDTD are almost identical with the classic FDTD algorithm:

1. Initialization;
2. Update interior E-field nodes;
3. Update boundary E-field nodes;
4. Update excitation E-field nodes;
5. Update interior H-field nodes;
6. Update boundary H-field nodes;
7. Archive observation nodes.

One nuance must be understood when using a GPU with a frame buffer (i.e., the buffer used for display) that does not support floating-point numbers, but supports only 8-bit fixed-point values. This amounts to clamping all numbers to a range between zero and one. Here, if one wishes to display the physics of wave propagation without any extra processing, it is important to make sure that clamping of the field values does not occur upon display. To avoid clamping: (1) the updating coefficients in the Yee time-stepping algorithm implemented in the GPU should be scaled (i.e., normalized) so that the E- and H-field values occupy a similar numerical range; and (2) the magnitude of the modeled excitation should be adjusted so that any properly occurring physical enhancements of the E- and H- fields do not result in field magnitudes that go outside the range of zero to one. The latter requirement is a bit tricky, since field enhancements (e.g., radial E-fields near PEC edges and corners; looping H-fields near thin wires; both E- and H-fields at standing-wave peaks, especially within resonant structures) may not be quantifiable before the problem is actually run.

20.8.1 Initialization

The first step in implementing FDTD on a GPU is to load all the required fragment and vertex programs used for the field updates and the boundary conditions. This is done with the following function calls which are made available through *n*VIDIA-specific or *Architecture Review Board* (ARB) OpenGL extensions:

```
1    unsigned char *Ez_program_buff = readFile("Ez_fp_float.txt");
2    unsigned int Ez_fpID;
3    glGenProgramsNV(1, &Ez_fpID);
4    glLoadProgramNV(GL_FRAGMENT_PROGRAM_NV, Ez_fpID,
5          strlen((const char*)Ez_program_buff), Ez_program_buff);
```

The first line of code is responsible for reading in a text file called `Ez_fp_float.txt`, containing the assembly instructions constituting the E_z (or $H_{x,y}$) update equations written as fragment programs. Examples of these fragment programs are provided in Sections 20.8.2 to 20.8.4. The second and third lines of code generate a fragment program identifier, `Ez_fpID`, which can be used later to activate or deactivate the program. The fourth line actually loads the fragment program into video memory, but does not invoke or make the program active. That step is done before the rendering pass which constitutes an update of all nodes.

The next step in the initialization is to load the field arrays and material properties into data/texture memory. The following OpenGL function calls are an example of how this is done:

```
1    unsigned int EzHxHyTextureID;
2    glGenTextures(1, &EzHxHyTextureID);
3    GLfloat *pEzHxHyBuffer = new GLfloat[iWidth * iHeight * 4];
4    glBindTexture(GL_TEXTURE_RECTANGLE_NV, EzHxHyTextureID);
5    glTexImage2D(GL_TEXTURE_RECTANGLE_NV, 0, GL_FLOAT_RGBA32_NV,
6                 iWidth,iHeight,0, GL_RGBA, GL_FLOAT, pEzHxHyBuffer);
```

As with the fragment programs, the texture identifiers are managed by using the function call in Line 2, `glGenTextures`, to generate the texture ID. Line 3 is a standard C++ call to build an empty buffer of the correct size. Its factor of four creates enough room to store E_z, H_x, H_y, and one other attribute of the field, such as a split field (E_{zx} and E_{zy}) or coefficient information. In Line 4, the `glBindTexture` call makes the particular texture ID active, such that any future texture-related calls affect the attributes of that particular ID. In Line 5, the `glTexImage2D` call loads the texture data into video memory. All parameters in these functions are standard and extended OpenGL parameters related to texture size and storage format.

In an analogous manner, the coefficients corresponding to material properties for the E- and H-nodes on the FDTD grid can be loaded as textures using the same calls. It is the C program that builds the material properties array before loading the data to video memory with the `glTexImage2D` call. The coefficients are accessed via the E, H, or boundary-fragment program as required in the related FDTD equation.

In addition to loading the required fragment programs and textures, the initialization task involves a series of OpenGL-specific initializations [27], such as opening a window, establishing a coordinate system, and defining the required rendering targets, which in our case are off-screen 32-bit floating point targets [28].

20.8.2 Electric and Magnetic Field Updates

Once all of the initialization tasks are accomplished, the FDTD simulation is executed by invoking a loop of rendering passes. The equations for the interior E_z and $H_{x,y}$ nodes are implemented via short assembly language fragment programs. Consider the following sample code that updates interior E_z nodes (lines beginning with "#" are comments):

```
1    !!FP1.0
2    DEFINE my_offset = (-1.0, -1.0, 0.0, 0.0);
3    # Add offsets to texture coordinate streamed from rasterizer
4    ADD R10, f[TEX0].xyxy, my_offset.xzzy;
5    # Access required H neighbors, old Ez, and CEb coefficient
6    TEX R0, f[TEX0], TEX0, RECT;
7    TEX R1.z, R10.xywz, TEX0, RECT;
8    TEX R1.y, R10.zwxy, TEX0, RECT;
9    # Add the dH components together
10   ADD R3.xyzw, R0.zyxw, R1.yzxw;
11   ADD R4.x, R3.x, -R3.y;
12   # Multiply dH components by CEb coefficient stored in R0.w
13   # and add old Ez value (no loss in this case)
14   MAD o[COLR].xyzw, R0.wxyz, R4.xyzw, R0.xyzw;
15   END
```

Line 1 tells the OpenGL compiler that this is a v1.0 fragment program. The DEFINE instruction declares a constant vector used to access neighbors in subsequent texture fetches. The first ADD instruction in Line 4 adds the offsets previously declared to the fragment coordinates that are streamed through from the rasterizer stage (see Section 20.7.2). These coordinates are calculated via interpolation based on the vertex texture coordinate pairs included in the `glBegin()/glEnd()` block in the C program. In order to improve performance, it is possible to move the DEFINE and first ADD instruction to the vertex processor, and thereby stream the offset coordinates through as vertex attributes. After this first ADD instruction is executed, the coordinates of two neighbors are stored in the `xy` and `zw` elements of the `R10` register.

The next three TEX instructions are used to access texture memory for the neighbors and one coefficient required for the E_z update equation. The next two ADD instructions implement and sum the H_x and H_y finite-differences, and place the result in the x-component of the R4 register. Subsequently, the multiply and add (MAD) instruction in Line 14 performs the final * and + arithmetic operations of the E_z Yee update. Because the three `R4.yzw` elements are assured to be zero [28], H_x, H_y, and CEb or `R2.yzw` are allowed to flow through unmodified along with the updated E_z value. All four values are written to the `o[COLR]` register, which in turn gets written to the floating-point render target.

Once the interior E_z nodes are updated, the E_z boundary conditions must be executed. The interior H_x and H_y nodes are then updated via the following fragment program:

```
1    !!FP1.0
2    DEFINE my_offset = {1.0, 1.0, 0.0, 0.0};
3    # Add offsets to texture coordinate streamed from rasterizer
4    ADD R10, f[TEX0].xyxy, my_offset.zyxz;
5    # Access required E neighbors and old Hxy values
6    TEX R0, f[TEX0], TEX0, RECT;
7    TEX R1.x, R10.xywz, TEX0, RECT;
8    TEX R2.x, R10.zwxy, TEX0, RECT;
9    # Add the dE components together
10   ADD R3.y, R0.x, -R1.x;
11   ADD R3.z, R2.x, -R0.x;
12   # Multiply dE components by hardcoded CHb coefficient
13   # and add old Hx,y values (no loss in this case)
14   MAD o[COLR], {0.0, 0.697622, 0.697622, 0.0}, R3, R0;
15   END
```

The significant difference between the $H_{x,y}$ and E_z fragment programs is that in the $H_{x,y}$ program, two fields are updated simultaneously via the vector arithmetic intrinsic to the fragment program. Once the neighbors are arranged appropriately in the TEX instructions and swizzled (i.e., once their x, y, z, and w components are reordered) in the two ADD instructions, the MAD instruction calculates H_x and H_y at the same time. In the sample program above, the coefficients are hard-coded, rather than being stored in texture memory, as was done in the E_z program. Again, the E_z value stored in `R0.x` is allowed to flow through unmodified.

20.8.3 Boundaries

At some boundaries, such as PEC surfaces, only a single field component requires special treatment. (The PEC boundary is enforced simply by setting its E_z values to zero.) At others, such as within a PML absorber, all three field components require special handling. In the context of the GPU accelerator, all boundary conditions are enforced via fragment programs. The following sections provide sample programs that implement three boundary conditions commonly used by FDTD modelers: the wraparound periodic boundary, the first-order Mur absorbing boundary, and the Berenger split-field PML absorbing boundary.

Wraparound Periodic Boundary Condition

The approach used here entails using the regular E_z FDTD update, except that one of the H_x values is taken not from an adjacent node, which does not exist because we are at the boundary, but from the H_x node on the opposite side of the domain. Hence, the following fragment program is exactly the same as that used for E_z, except that one of the offsets is the height of the domain rather than one. In this example, the periodic boundary is assumed to be implemented at the maximum y-coordinate plane of the grid (the "north boundary"):

```
1    !!FP1.0
2    # Offset is the Height (128 in this case)
3    DEFINE my_offset = {1.0, -128.0, 0.0, 0.0};
4    ADD R10, f[TEX0].xyxy, my_offset.xzzy;
5    # Access required H neighbors and old Ez value
6    TEX R0, f[TEX0], TEX0, RECT;
7    TEX R1.z, R10.xywz, TEX0, RECT;
8    TEX R1.y, R10.zwxy, TEX0, RECT;
9    # Add the dH components together
10   ADD R3.xyzw, R0.zyxw, R1.yzxw;
11   ADD R4.x, R3.x, -R3.y;
12   # Multiply dH components by CEb coefficient stored in R0.w
13   # and add old Ez value (no loss in this case)
14   MAD o[COLR].xyzw, R0.wxyz, R4.xyzw, R0.xyzw;
15   END
```

Of course, the other major difference between this fragment program and the one used for the E_z interior nodes, is that it is executed only on the periodic boundary. This is accomplished in OpenGL by drawing a thin rectangle on the edge of the domain.

First-Order Mur Absorbing Boundary Condition

The first-order Mur absorbing boundary condition is enforced in the same way as a periodic boundary: by rendering a thin rectangle on the boundary region with a fragment program that executes the relevant equation. However, the form of the fragment program is completely different than that used for the periodic boundary. Consider the following sample fragment program for a first-order Mur absorbing boundary implemented at the maximum x-coordinate plane of the grid (the "east" boundary):

```
1    !!FP1.0
2    # Define offset for interior node
3    DEFINE my_offset = (0.0, -1.0, 0.0, 0.0);
4    # Add offsets to texture coordinate streamed from rasterizer
5    ADD R11, f[TEX0], my_offset;
6    # Define pre-calculated Mur 1st-order constant
7    DEFINE K = (-0.222198, 0.0, 0.0, 0.0);
8    # Access the required E neighbors
9    TEX R0.x, f[TEX0], TEX0, RECT;
10   TEX R1.x, R11, TEX0, RECT;
11   TEX R2.w, R11, TEX0, RECT;
12   # Add new interior and old boundary nodes
13   ADD R3.x, R1.x, -R0.x;
14   # Multiply Mur constant by R3 and add old interior node
15   MAD o[COLR], K, R3, R2.wyzx;
16   END
```

Only one neighbor is needed to implement the first-order Mur condition. Hence, the offset register in the DEFINE instruction in Line 3 contains only one value, –1. The other DEFINE instruction in Line 7 declares the first-order Mur constant K. In Lines 9, 10, and 11, three TEX instructions fetch the old E_z value at the boundary and the new and old E_z values at the neighboring interior node, and place these in the R0, R1, and R2 registers, respectively. In Line 13, the ADD instruction sums the new E_z value at the interior node and the old E_z value at the boundary. Subsequently, in Line 15, the MAD instruction performs the K multiplication, and adds to this result the old value of E_z at the interior node, thereby completing the updating of E_z at the boundary. The final result is written to the floating-point render target along with the unmodified H_{xy} values.

Berenger Split-Field PML Absorbing Boundary Condition

Unlike the other boundary conditions discussed, Berenger's split-field PML absorbing boundary condition requires that both E and H field components must be updated (with the exception of the outermost grid boundary planes, which are PEC). The E_{zx} and E_{zy} updates are implemented via a fragment program similar to the one shown below:

```
1    !!FP1.0
2    DEFINE my_offset = (-1.0, -1.0, 0.0, 0.0);
3    # Add offsets to texture coordinate streamed from rasterizer
4    ADD R10, f[TEX0].xyxy, my_offset.xzzy;
5    # Access the required H neighbors and old Ezx and Ezy values
6    TEX R0, f[TEX0], TEX0, RECT;
7    TEX R1.z, R10.xywz, TEX0, RECT;
8    TEX R1.y, R10.zwxy, TEX0, RECT;
9    # Access CEa, CEb, CHa, CHb coefficients
10   TEX R4, f[TEX0], TEX1, RECT;
11   # Ezx: The next four instructions calculate Ezx
12   # [Hy1-Hy2]
13   ADD R2.y, R0.z, -R1.z;
```

```
14    # CEb*dHy
15    MUL R3.y, R4.y, R2.y;
16    # CEa*Ezxold
17    MUL R7.y, R4.x, R0.x;
18    # Ezxnew = CEa*Ezxold + CEb*dHy
19    ADD R5.x, R7.y, R3.y;
20    # Ezy: The next four instructions calculate Ezy
21    # [Hx1-Hx2]
22    ADD R2.z, R1.y, -R0.y;
23    # CEb*dHx: CEb is hard-coded because its sigma is zero
24    MUL R6, (0.0, 0.0, 0.700036, 0.0), R2;
25    # Ezynew = Ezyold + CEb*dHx
26    ADD R5.w, R0.w, R6.z;
27    # Write Ez, Ezx, Ezy to render target
28    MAD o[COLR], (0, 1, 1, 0), R0, R5;
29    END
```

The E_{zx} and E_{zy} fields are calculated and stored in the `x` and `w` components of `R5` (and subsequently written to `o[COLR]` along with the unmodified $H_{x,y}$ fields). The PML boundary regions are not rectangles of single-cell width, as with the Mur and PEC boundaries, but are instead typically a number of grid cells thick. Therefore, in a sense, the PML regions are rectangles analogous to the interior region.

20.8.4 Source Excitation

Source excitations are also implemented via fragment programs. Because current GPUs do not support `glBlend` [28] for 32-bit floating point render targets, it is necessary to recalculate the E-field for the fragments constituting the source region. Then, the source value is added to the E-field. The following code implements a soft source excitation that is modulated in space:

```
1     !!FP1.0
2     # Declaration of source variable set every step by C program
3     DECLARE SourceValue;
4     DEFINE my_offset = (-1.0, -1.0, 0.0, 0.0);
5     # Add offsets to texture coordinate streamed from rasterizer
6     ADD R10, f[TEX0].xyxy, my_offset.xzzy;
7     # Access required H neighbors and old Ez value
8     TEX R0, f[TEX0], TEX0, RECT;
9     TEX R1.z, R10.xywz, TEX0, RECT;
10    TEX R1.y, R10.zwxy, TEX0, RECT;
11    Access texture used for sinusoidal spatial modulation
12    TEX R2.x, f[TEX0], TEX1, RECT;
13    # Add the dH components together
14    ADD R3.xyzw, R0.zyxw, R1.yzxw;
15    ADD R4.x, R3.x, -R3.y;
16    # Multiply by coefficient and add old Ez value
17    MAD R5, (0.697622, 0.0, 0.0, 0.0), R4, R0;
```

```
18   # Modulate the source value in space
19   MUL R6.x, SourceValue.x, R2.x;
20   # Add the modulated source to Ez and write to render target
21   ADD o[COLR], R5, R6;
22   END
```

A variable passed in each time-step via the main C program is stored as a register, in this case called SourceValue. All the instructions used for the E_z update are included in the source fragment program. However, TEX1 contains a texture used for spatial modulation, such as is required with a waveguide. These data are stored in R2.x and multiplied by the SourceValue to constitute the spatially modulated source value. This source value is added to the calculated E_z field and written to the render target. In the future, it will not be necessary to recalculate the E_z field, because it will be possible to blend the source value with the updated E_z value already residing in the render target. Of course, the source fragment program only operates on the fragments corresponding to the nodes at which excitation is desired—typically a very small percentage of the total number of nodes.

20.8.5 Archiving Observation Nodes

In order to archive observation nodes, it is preferable to store them in video memory until the simulation is complete. After the simulation, the data from all the observed time-steps are read back into system memory for analysis. It is possible to read data back into system memory every time-step, but this has a negative impact on performance.

A texture for the storage of observation points is made using the same OpenGL calls made in Section 20.8.1. After the *E*- or *H*-field has been updated into the render target, the results texture is bound as in Line 1 below:

```
1    glBindTexture(GL_TEXTURE_RECTANGLE_NV, ResultsTextureID);
2    glCopyTexSubImage2D(GL_TEXTURE_RECTANGLE_NV, 0, i%4096, i/4096,
3                        ProbePoint[0], ProbePoint[1], 1, 1);
```

Line 2 shows the function used to copy the data from the render target to the archive or results texture. Consider the data for a particular range of pixels corresponding to the observation location(s). These data are copied from the render target to a range of pixels corresponding to the current time-step in the bound archive or results texture. Alternatively, it is possible to render an observation point from a texture containing the field values into a render target that is subsequently released and bound as a texture.

20.8.6 Multipass Rendering

While the rendering sequence varies slightly depending on what kind of boundary conditions are used, in general the "multipass" rendering approach to FDTD is as shown in Fig. 20.11. After initialization, the E_z field is updated by invoking a rendering pass over the geometry corresponding to the entire two-dimensional FDTD domain, or in other words, a rectangle. Typical OpenGL code used to invoke such a rendering pass follows:

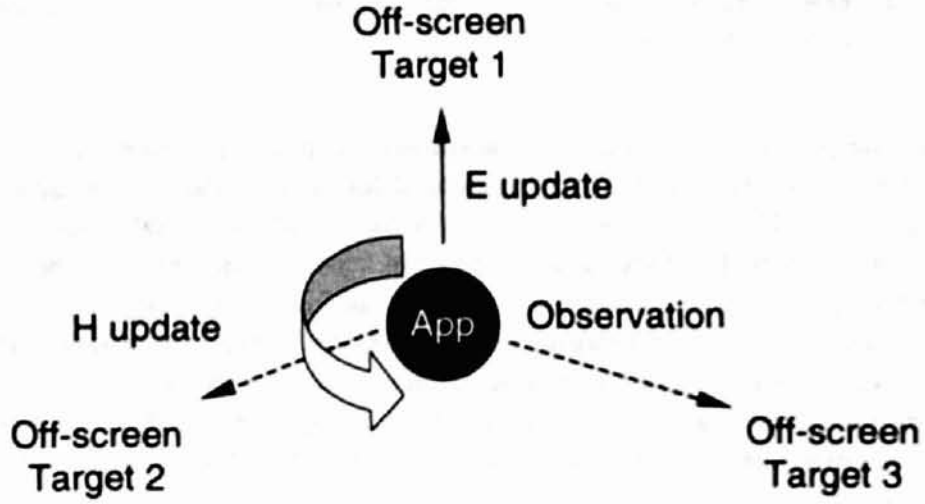


Fig. 20.11 Multipass rendering.

```
1    glBindProgramNV(GL_FRAGMENT_PROGRAM_NV, Ez_fpID_float);
2    glEnable(GL_FRAGMENT_PROGRAM_NV);
3    glActiveTextureARB(GL_TEXTURE0_ARB);
4    glBindTexture(GL_TEXTURE_RECTANGLE_NV, EzHxHyTextureID);
5    glBegin(GL_QUAD);
6          glTexCoord2f(0,iHeight);          glVertex2f(Left, Top);
7          glTexCoord2f(0,0);                glVertex2f(Left, Bottom);
8          glTexCoord2f(iWidth,0);           glVertex2f(Right, Bottom);
9          glTexCoord2f(iWidth, iHeight);    glVertex2f(Right, Top);
10   glEnd();
11   glDisable(GL_FRAGMENT_PROGRAM_NV);
```

The first four lines of code activate and choose the appropriate textures and fragment programs. The `glBegin/End` block tells the GPU to draw a rectangle with dimensions `iWidth` × `iHeight` pixels, and map the four corners of texture unit zero to the four corners of this rectangle. This block of code invokes one rendering pass. During this pass, all of the texture data are streamed through the fragment processors (executing the same fragment program), allowing the active fragment program to execute the equation implemented by its instructions. After this pass, all of the nodes inside the geometry provided are updated, as expected.

Before the next pass, the render target buffer containing the updated E-field values is released as a render target, and bound as a texture to be used as source data for the H-field pass. Subsequently, a separate pass can be used to archive observation points. This triplet of passes constitutes one simulation time-step, and is repeated for as many time-steps as are needed.

20.8.7 Display

The GPU can be used solely as a coprocessor without any display capabilities whatsoever. However, it is possible to make use of the inherent display capabilities with little performance penalty. In order to harvest the intrinsic display capabilities of the graphics card, a useful mapping must be made between the E- and H-field components and the RGB pixel components, which ultimately are the only values that can be displayed. One possible mapping for the two-dimensional case discussed here is $(E_z, H_x, H_x) \rightarrow$ (red, green, blue). Furthermore, before any RGB vectors are displayed, each element should be clamped to the numerical range [0, 1], as previously discussed.

20.9 PERFORMANCE MEASUREMENTS OF THE GPU ACCELERATOR

The following two-dimensional FDTD simulations were performed to validate the accuracy and measure the speed of the GPU FDTD accelerator described in Sections 20.7 and 20.8:

1. Lossless 512×512-cell cavity with PEC walls, excited by a step impulse source. Run duration: 100,000 time-steps.
2. Single 1024×128-cell ($x \times y$) row of an EBG structure comprised of an array of circular dielectric cylinders of permittivity $\varepsilon_r = 4.2$ arranged in a square pattern. Periodic boundary conditions applied at the $\pm y$-edges of the row. First-order Mur absorbing boundaries applied at the $\pm x$-edges of the row. Run duration: 20,000 time-steps.
3. 2381×101-cell ($x \times y$) rectangular waveguide with PEC walls. Small rectangular ceramic "fingers" ($\varepsilon_r = 5.6$, $\sigma = 0.01277$ S/m) placed at intervals along each wall. TE_{10} excitation with eight-cell-thick PMLs terminating the waveguide model at its $\pm x$ outer boundaries. Run duration: 30,000 time-steps.

The GPU used was an *n*VIDIA GeForce FX5950 with 256 MB RAM on a Gainward 5950 Ultra/1800 Golden Sample board. The benchmark system used to compare with the GPU performance was a 1.6-GHz Intel Pentium 4 running Windows XP Professional with 256 MB RAM, 512 KB L2 Cache, and C code compiled by MS VC++ 6.0.

In the three cases studied, the GPU accelerated the FDTD processing speed by 7.78:1, 7.71:1, and 7.74:1, respectively, relative to FDTD software running on the benchmark PC. These accelerations imply measured GPU throughputs of 240 to 250 *million field updates per second* (MFUPS). Note that these very similar performance levels were achieved even though the three applications required implementing very different boundary conditions and materials.

A calculation of the maximum attainable throughput of the GPU (without any cache stalls) indicates that 857 MFUPS should be possible. Thus, it appears that cache-missing may be responsible for the measured performance being lower by a factor of approximately 3.4:1 relative to the theoretical value. It is expected that improved cache and memory models should allow improved performance approaching the theoretical limit. In addition, analysis of the measured GPU performance indicates that the OpenGL overhead factor becomes negligible when the field updating time per complete iteration greatly exceeds the measured 0.3-ms overhead per iteration (independent of grid size) that supports the multipass rendering scheme. This is a factor favoring the processing of large grids.

20.10 SUMMARY AND CONCLUSIONS

This chapter discussed the concept of hardware acceleration of the FDTD method. Two fundamentally different techniques were considered: (1) special-purpose FPGAs, and (2) GPUs. While FPGAs implement Maxwell's equations in logic gates, with GPUs Maxwell equations are expressed in graphical languages such as OpenGL. Both approaches have succeeded in demonstrating acceleration of large, nontrivial FDTD simulations. Order-of-magnitude accelerations relative to optimized (but conventional) FDTD software running on top-of-the-line PCs should be possible in the near future.

FPGA-based accelerators appear to be considerably more costly than those based on GPUs, since special-purpose cards need to be built with large and expensive FPGA chips. While FPGA-based accelerators are capable of handling large memory, they have very limited adaptability to the special field updates needed to model the advanced sources, boundary conditions, and material properties of importance to FDTD practitioners. On the other hand, GPU accelerators are based on mass-produced consumer graphics cards. While GPUs are currently short on memory, they allow advanced FDTD features to be implemented in a straightforward manner.

The future will likely see a hybrid approach that combines the benefits of both techniques. Industry trends indicate that the hardware acceleration of FDTD should become more attractive with time, because the speed and complexity of special-purpose hardware appear to be increasing faster than those of general CPUs.

REFERENCES

[1] Taflove, A., *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Norwood, MA: Artech House, 1995.

[2] Gedney, S. D. "Efficient implementation of the FDTD algorithm on high-performance computers," in *Time-Domain Methods for Microwave Structures: Analysis and Design*, B. Houshmand and T. Itoh, (eds.), New York: IEEE Press, 1997, pp. 313–322.

[3] Jordan, H., S. Bokhari, S. Staker, J. Sauer, M. El Helbawy, and M. Piket-May, "Experience with ADI-FDTD techniques on the Cray MTA supercomputer," *Proc. SPIE Commercial Applications High-Performance Computing*, Vol. 4528, 2001, pp. 68–76.

[4] Schiavone, G. A., I. Codreanu, R. Palaniappan, and P. Wahid, "FDTD speedups obtained in distributed computing on a Linux workstation cluster," *Proc. IEEE Antennas Propagat. Soc. Intl. Symp.*, Vol. 3, 2000, pp. 1336–1339.

[5] Marek, J. R., M. A. Mehalic, and A. J. Terzuoli, "A dedicated VLSI architecture for finite-difference time-domain calculations," *Proc. 8th Ann. Review Progress Appl. Comp. Electromagn.*, Monterey, CA, Vol. 1, 1992, pp. 546–553.

[6] Placidi, P., L. Verducci, G. Matrella, L. Roselli, and P. Ciampolini, "A custom VLSI architecture for the solution of FDTD equations," *IEICE Trans.*, Vol. E85, 2002, pp. 572–577.

[7] Verducci, L., P. Placidi, G. Matrella, L. Roselli, F. Alimenti, P. Ciampolini, and A. Scorzoni, "A feasibility study about a custom hardware implementation of the FDTD algorithm," *Proc. URSI 2002*, Maastricht, 2002.

[8] Verducci, L., P. Placidi, P. Ciampolini, A. Scorzoni, and L. Roselli, "A standard cell hardware implementation for finite-difference time domain (FDTD) calculation," *Proc. IEEE MTT Intl. Symp.*, Vol. 3, 2003, pp. 2085–2088.

[9] Kawaguchi, H., K. Takahara, and D. Yamauchi, "Design study of ultrahigh-speed microwave simulator engine," *IEEE Trans. Magnetics*, Vol. 38 , 2002, pp. 689–692.

[10] Schneider, R., L. Turner, and M. Okoniewski, "Application of FPGA technology to accelerate the finite-difference time-domain (FDTD) method," *FPGA 2002 Conf.*, Monterey, CA, 2002.

[11] Schneider, R. N., M. M. Okoniewski, and L. E. Turner, "Finite-difference time-domain method in custom hardware?" *IEEE Microwave Wireless Components Lett.*, Vol. 12, 2002, pp. 488–490.

[12] XESS Development Corp. Web site [online]: http://www.xess.com/

[13] Durbano, J. P., M.E.E. thesis, Univ. Delaware, Newark, DE, 2002.

[14] Durbano, J. P., F. E. Ortiz, J. R. Humphrey, M. S. Mirotznik, and D. W. Prather, "Hardware implementation of a three-dimensional finite-difference time-domain algorithm," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 2, 2003, pp. 54–57.

[15] Durbano, J. P., J. R. Humphrey, F. E. Ortiz, P. F. Curt, D. W. Prather, and M. S. Mirotznik, "Hardware acceleration of the 3D finite-difference time domain method," *Proc. IEEE Antennas Propagat. Soc. Intl. Symp.*, Vol. 1 , 2004, pp. 77–80.

[16] EM Photonics Web site [online]: http://www.emphotonics.com/

[17] Acceleware Web site [online]: http://www.acceleware.com

[18] Gwarek, W. K., "Analysis of arbitrarily shaped two-dimensional microwave circuits by finite-difference time-domain method," *IEEE Trans. Microwave Theory Tech.*, Vol. 36, 1988, pp. 738–744.

[19] Bruton, L. T., "Low sensitivity digital ladder filters," *IEEE. Trans. Circuits Syst.*, Vol 22, 1975, pp. 168–176.

[20] Avvida Systems Web site [online]: http://www.avvida.com or SBS Technologies (Canada) Web site [online]: http://www.sbs.com

[21] Chen, W., P. Kosmas, M. Leeser, and C. Rappaport, "An FPGA implementation of the two-dimensional finite-difference time-domain (FDTD) algorithm," *Proc. ACM/SIGDA 12th Intl. Symp. Field Programmable Gate Arrays*, Monterey, CA, 2004, pp. 213–222.

[22] Schneider, R. N., L. E. Turner, and M. Okoniewski, "A software-coupled 2D FDTD hardware accelerator," *Proc. IEEE Antennas Propagat. Soc. Intl. Symp.*, Vol. 2, 2004, pp. 1692–1695.

[23] Krakiwsky, S. E., M. Okoniewski, and L. E. Turner, "Acceleration of finite-difference time-domain (FDTD) using graphics processor units (GPUs)," *Proc. Intl. Microwave Symp.*, Fort Worth, TX, 2004.

[24] Venkatasubramanian, S., "The graphics card as a stream computer," *ACM SIGMOD/PODS Workshop on Management and Processing of Massive Data*, 2003.

[25] PCI Express Web site [online]: http://www.nvidia.com/page/pci_express.html

[26] Fatahalian, K., J. Sugerman, and P. Hanrahan, "Understanding the efficiency of GPU algorithms for matrix-matrix multiplication," *Graphics Hardware*, 2004.

[27] Woo, M., *OpenGL Programming Guide, 3rd ed.*, New York: Addison-Wesley, 1999.

[28] *n*VIDIA OpenGL extension specifications [online]: http://developer.nvidia.com

Acronyms and Common Symbols

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
A	ampere
ABC	absorbing boundary condition
ADE	auxiliary differential equation
ADI	alternating-direction implicit
AFP	analytical field propagation
AGP	accelerated graphics port
AlAs	aluminum arsenide
AlGaAs	aluminum gallium arsenide
API	application programmable interface
AR	autoregressive
ARB	Architecture Review Board
ASIC	application-specific integrated circuit
ATR	attenuated total reflection
B	magnetic flux density
BCGS-FFT	stabilized biconjugate-gradient fast Fourier transform
BJT	bipolar junction transistor
BOR	body of revolution
c	free-space (vacuum) speed of light
C	capacitance
CAD	computer-aided design
CCOM	concurrent complementary operator method
CEM	computational electromagnetics
CFDTD	conformal FDTD
CFS	complex frequency-shifted
CMOS	complementary metal oxide semiconductor
COM	complementary operator method
CP-FDTD	contour-path FDTD
CPML	convolutional perfectly matched layer
CPW	coplanar waveguide
D	electric flux density
DBR	distributed Bragg reflection (or reflector)
DFT	discrete Fourier transform
DGM	discontinuous Galerkin method
DG-PSTD	discontinuous Galerkin pseudospectral time-domain
DNA	deoxyribonucleic acid
DSI	discrete surface integral
E	electric field intensity
EBG	electromagnetic bandgap
ELF	extremely low frequency
EMP	electromagnetic pulse
EXPAR	exponential autoregressive
FCT	fast cosine transform
FD-NTFF	frequency-domain near-to-far-field
FDTD	finite-difference time-domain

FE	finite element
FET	field-effect transistor
FETD	finite-element time-domain
FFT	fast Fourier transform
FMM	fast multipole method
FPGA	field-programmable gate array
fs	femtosecond
FSR	free spectral range
FSS	frequency-selective surface
FV	finite volume
FVTD	finite-volume time-domain
FWHM	full-width at half-maximum
G	conductance
GaAs	gallium arsenide
GB	gigabyte
GC	conductance capacitance
GCL	Gauss Chebyshev Lobatto
GHz	gigahertz
GLL	Gauss Legendre Lobatto
GMT	generalized multipole technique
GNLSE	generalized nonlinear Schrödinger equation
GPR	ground-penetrating radar
GPU	graphics processor unit
GY	generalized Yee
H	magnetic field intensity
HDL	hardware description language
H-H	horizontal-horizontal
HPM	high-power microwave
H-pol	horizontal polarization
I	current, amperes
InGaAs	indium gallium arsenide
InGaAsP	indium gallium arsenide phosphide
InP	indium phosphide
J	electric current density
KB	kilobyte
L	inductance
LC	inductance capacitance
LCD	liquid-crystal display
LCR	inductor-capacitor-resistor
LDI	lossless discrete integrator
$LiNbO_3$	lithium niobate
LSB	least significant bit
LTSA	linear tapered slot antenna
M	magnetic current density
MB	megabyte
MCM	multichip module
MESFET	metal-semiconductor field-effect transistor
MFUPS	million field updates per second
MHz	megahertz
MIMD	multiple-instruction multiple-data
MMP	multiple multipole
MND	matched numerical dispersion

MOCVD	metalorganic chemical vapor deposition
MoM	method of moments
MRI	magnetic resonance imaging
MRTD	multiresolution time-domain
N-FDTD	nonorthogonal FDTD
NASA	National Aeronautics and Space Administration
NEC	Numerical Electromagnetics Code
NLSE	nonlinear Schrödinger equation
nm	nanometer
NPC	nonlinear photonic crystal
ns	nanosecond
NTFF	near-to-far-field
NTNF	near-to-near-field
NUFFT	nonuniform fast Fourier transform
PBG	photonic bandgap
PC	personal computer
PC	photonic crystal
PCB	printed circuit board
PCI	peripheral component interconnect
PCS	personal communications services
PDE	partial differential equation
PEC	perfect electric conductor
PEP	Pauli Exclusion Principle
PGS	periodic-gain structure
PGY	planar generalized Yee
PIC	particle-in-cell
PLRC	piecewise-linear recursive convolution
pm	picometer
PMC	perfect magnetic conductor
PML	perfectly matched layer
PPW	points per wavelength
ps	picosecond
PSTD	pseudospectral time-domain
PWC	personal wireless communications
Q	quality factor
QED	quantum electrodynamics
QPM	quasi phase matching
QW	quantum well
R	resistance
RAM	random-access memory
RC	recursive convolution
RCS	radar cross section
RF	radio frequency
RGB	red green blue
RLC	resistor inductor capacitor
S	scattering (parameter)
SAM	standard anthropomorphic model
SAR	specific absorption rate
SEM	scanning electron microscope
SERS	surface-enhanced Raman spectroscopy
SFT	spatial Fourier transform
SHG	second-harmonic generation

SIBC	surface-impedance boundary condition
SIMD	single-instruction multiple-data
SWR	standing-wave ratio
TD-NTFF	time-domain near-to-far-field
TE	transverse electric
TE_z	transverse electric relative to the z-direction
TEM	transverse electromagnetic
TF/SF	total-field / scattered-field
THz	terahertz
TIR	total internal reflection
TLM	transmission-line matrix
TM	transverse magnetic
TM_z	transverse magnetic relative to the z-direction
TSA	tapered slot antenna
TSCS	total scattering cross section
TSF	thin-slot formulation
UGS	uniform-gain structure
UHF	ultrahigh frequency
ULF	ultralow frequency
UPML	uniaxial perfectly matched layer
UTD	uniform theory of diffraction
UWB	ultrawideband
V	voltage, volts
VCSEL	vertical-cavity surface-emitting laser
VLF	very low frequency
VLSI	very large-scale integration
V-pol	vertical polarization
VSWR	voltage standing-wave ratio
V-V	vertical-vertical
WDM	wavelength-division multiplexing
WG1	waveguide number 1
WG2	waveguide number 2
WGM	whispering-gallery mode
WWW	World Wide Web
Z	impedance
ZCZ	Zheng-Chen-Zhang
λ	wavelength
λ_0	free-space (vacuum) wavelength
λ_B	Bragg wavelength
λ_d	dielectric wavelength
λ_{lasing}	lasing wavelength
μm	micrometer (micron)
μs	microsecond
ω	angular frequency

About the Authors

Allen Taflove has been a professor in the Department of Electrical and Computer Engineering of Northwestern University, Evanston, Illinois, since 1984. Since 1972, he has pioneered basic theoretical approaches and engineering applications of finite-difference time-domain computational electrodynamics. He coined the FDTD acronym in a 1980 IEEE paper, and in 1990 was the first person to be named a Fellow of the IEEE in the FDTD area. In total, he has authored or coauthored 5 books, 20 book chapters and articles, approximately 100 refereed journal papers and 300 conference papers, and 14 U.S. patents. These publications resulted in his being included on ISIHighlyCited.com, the Institute of Scientific Information's list of the most-cited researchers worldwide. He has been the principal investigator on over 40 grants and contracts and the thesis adviser to 20 Ph.D. recipients.

Professor Taflove is currently an elected member of Northwestern's General Faculty Committee, and the graduate program chair of his department. He is also the faculty advisor to McCormick's Undergraduate Design Competition, the Honors Program in Undergraduate Research, and the student chapters of Eta Kappa Nu and Tau Beta Pi. His efforts on behalf of students at all levels were recognized by Northwestern in 2000, when he was named a Charles Deering McCormick Professor of Teaching Excellence.

Professor Taflove's research interests span much of the electromagnetic spectrum. He and his students are modeling electrodynamic phenomena ranging from geophysically induced ultralow-frequency wave propagation about the entire Earth, to the light-scattering behavior of early-stage colon cancer cells. The principle that "Maxwell's equations work from dc to light" is vividly demonstrated in his laboratory every day. For downloads of his papers from 1975 to the present, visit his Web page at http://www.ece.northwestern.edu/ecefaculty/Allen1.html or contact him via e-mail at taflove@ece.northwestern.edu.

Susan C. Hagness is an associate professor in the Department of Electrical and Computer Engineering and a faculty affiliate of the Department of Biomedical Engineering at the University of Wisconsin-Madison. She received B.S. (with highest honors) and Ph.D. degrees in electrical engineering from Northwestern University, Evanston, Illinois, in 1993 and 1998, respectively. While working toward her Ph.D., she was a National Science Foundation Graduate Fellow. Her current research interests include time-domain computational electromagnetics algorithms and applications in photonics and bioelectromagnetics, as well as electromagnetic imaging, sensing, and thermal therapy techniques in biological and medical applications. She has been actively involved in FDTD research since 1990, and has coauthored over 130 journal and conference publications.

In 2000, Professor Hagness was the recipient of the National Science Foundation CAREER award, and the Presidential Early Career Award for Scientists and Engineers. Subsequently, she was an invited participant in the National Academy of Engineering's Seventh Annual Symposium on Frontiers of Engineering. In 2002, she received the Henry G. Booker Fellowship from the United States National Committee (USNC) of the International Union of Radio Science (URSI). That year, she was also named one of the 100 top young innovators in science and engineering in the world by MIT's *Technology Review* magazine. In 2003, she received the University of Wisconsin Emil Steiger Distinguished Teaching Award. In 2004, she received the

IEEE Engineering in Medicine and Biology Society Early Career Achievement Award "for outstanding contributions to the development and application of computational electromagnetics techniques and microwave diagnostic technologies for the detection of breast cancer."

Professor Hagness is an elected member of the IEEE Antennas and Propagation Society (AP-S) Administrative Committee (2003–2005), a Member-at-Large of USNC/URSI (2005–2007), and the IEEE AP-S Representative to the USNC/URSI (2005–2007). She is currently an associate editor of *IEEE Antennas and Wireless Propagation Letters*, a Senior Member of the IEEE, and a member of USNC/URSI Commissions A, B, D, and K. She can be contacted via e-mail at hagness@engr.wisc.edu.

Jeffrey Blaschak (Chapter 12) is a system architect with the Advanced Defensive Systems Technology Center at Northrop-Grumman Corporation, Rolling Meadows, Illinois, where he performs research and development for next-generation electro-optic sensors. Previously, he was vice president, Capital Market Systems, at BankOne, Chicago, Illinois, where he developed trading systems, analytics, and market risk-assessment tools to manage portfolios of interest rate and equity derivative instruments. Dr. Blaschak received an M.S. degree from Purdue University and a Ph.D. degree from Northwestern University, both in electrical engineering. After completion of his Ph.D., he joined the MIT Lincoln Laboratory as a member of the technical staff, and then the U.S. Air Force Research Lab as a branch chief, where he performed modeling and simulation of electromagnetic phenomena with applications to radar and electronic countermeasure systems.

Anders Bondeson (Chapter 19) was a professor at Chalmers University of Technology, Göteborg, Sweden. He suddenly and unexpectedly passed away on March 20, 2004. Professor Bondeson received M.S. and Ph.D. degrees from Chalmers University in 1975 and 1979, respectively. He was with the Laboratory for Fusion and Energy studies, University of Maryland, College Park (1979–1982); Chalmers University (1982–1986); and the Centre de Recherches en Physique des Plasmas, Ecole Polytechnique Fédérale de Lausanne (1986–1993), working mainly in computational magnetohydrodynamics and electron-beam devices. In 1993, he became a professor at Uppsala University, and in 1995 at Chalmers University. Professor Bondeson also made significant contributions to computational electromagnetics for microwave and eddy-current problems. He is greatly missed by his students and colleagues.

Geoffrey Burr (Chapter 16) has been a research staff member at the IBM Almaden Research Center, San Jose, California, since 1999. Previously, he received a Ph.D. degree in electrical engineering from the California Institute of Technology in 1996, and joined IBM Almaden as a postdoctoral researcher. For the first part of his career, Dr. Burr worked extensively on holographic data storage, including systems studies, signal processing and coding, and cutting-edge experimental demonstration platforms. His current areas of interest include nanophotonics, design optimization with FDTD, phase-change nonvolatile electronic memory, and content-addressable holographic storage.

Malgorzata Celuch-Marcysiak (Chapter 10) is an assistant professor at Warsaw University of Technology (WUT), Warsaw, Poland. She received an International Baccalaureate from the United World College of the Atlantic in 1983, and M.Sc. and Ph.D. degrees in electronic engineering from WUT in 1988 and 1996, respectively. She is one of the main authors of the QuickWave and Concerto electromagnetic-simulation softwares, as well as the cofounder and

vice president of QWED, a company that is commercializing these tools. She has contributed to over 80 technical publications, and is a reviewer for *IEEE Transactions on Microwave Theory and Techniques*, *IEEE Transactions on Antennas and Propagation*, and *IEEE Microwave and Wireless Components Letters*. In addition, she has served as a member of the Technical Program Committee of the IEEE Microwave Theory and Techniques Society International Microwave Symposium. Her research interests include numerical modeling of electromagnetic problems.

Shih-Hui (Gilbert) Chang (Chapter 9) is a joint postdoctoral research fellow in the Department of Chemistry of Northwestern University and the Center for Nanoscale Materials of Argonne National Laboratory. Previously, he received a B.S. degree in physics from National Taiwan University (1994), an M.S. degree in physics from the University of California, Riverside (1996), and a Ph.D. degree in electrical engineering from Northwestern University (2004). During his graduate research, he studied optical propagation and lasing processes in random media. His current research interests include nanophotonics, surface plasmons in metal nanostructures, surface-enhanced Raman scattering for single-molecule detection, and parallel-processing multiphysics and multiscale FDTD implementations.

Nicolas Chavannes (Chapters 11 and 14) has been with the Foundation for Research on Information Technologies in Society at the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, since 1999. Here, he is leading the software and computational techniques group. Previously, he received M.S. and Ph.D. degrees in electrical engineering from ETH in 1998 and 2002, respectively. In 1996, he joined the Bioelectromagnetics / EMC Group at ETH, where he was involved in computational electrodynamics and related dosimetric applications. From 1998 to 2002, he was with the Laboratory for Electromagnetic Fields and Microwave Electronics, as well as the Laboratory for Integrated Systems, both located at ETH Zurich. There, his research activities were focused on the development of FDTD local mesh-refinement techniques and their application to numerical near-field analysis. In early 2002, he joined Schmid & Partner Engineering AG (SPEAG) Zurich as head of the software R&D team, focusing on innovations and implementations for SPEAG's experimental assessment and numerical simulation tools. His primary research interests include the development, implementation, and application of computational modeling and simulation techniques in electromagnetics, in particular, antennas and bioelectromagnetic interaction mechanisms.

Hans De Raedt (Chapter 18) is professor of computational physics in the Department of Applied Physics, Materials Science Centre, University of Groningen, the Netherlands. He received his university degree in physics from the University of Antwerp, Belgium, where, in 1976, he received a Ph.D. degree for theoretical work on magnetic phenomena in one dimension. He has held visiting scientist/professor positions at the Max Planck Institute, Munich; IBM Research Laboratory, Zurich; and at the Universities of Osaka and Tokyo, Japan. He served two years as a director of the School of Physics at the University of Groningen. Professor De Raedt has been the thesis advisor of six Ph.D. recipients, has served as principal investigator of 11 grants and EU contracts, and has coauthored more than 150 publications. After contributing to the early development of quantum Monte Carlo methods, his interest shifted towards the simulation of physical phenomena in the time domain. His main current research interests are the development of time-domain algorithms for solving Maxwell's equations, the simulation of physical models for quantum computers, the dynamic properties of nanoscale magnets, and general quantum phenomena. Professor De Raedt can be contacted at h.a.de.raedt@rug.nl.

Fredrik Edelvik (Chapter 19) is with the Fraunhofer-Chalmers Research Centre for Industrial Mathematics, Gothenburg, Sweden. He received an M.S. degree in engineering physics and a Ph.D. degree in scientific computing from Uppsala University, Uppsala, Sweden, in 1997 and 2002, respectively. In 2003–2004, he was a postdoctoral research associate at the Institut für Theorie Elektromagnetischer Felder (TEMF), Technische Universität Darmstadt, Darmstadt, Germany. Dr. Edelvik's current research interests are in computational electromagnetics, with an emphasis on time-domain finite-element methods, hybrid (unstructured / structured) grid methods, and subcell models for thin wires, thin slots, and thin sheets.

Masafumi Fujii (Chapter 9) has been a junior associate professor in the Department of Electrical, Electronic, and System Engineering at Toyama University, Japan, since 2003. He received B.E. and M.E. degrees from Kobe University, Hyogo, Japan, in 1989 and 1991, respectively; and a Ph.D. degree from the University of Victoria, Canada, in 1999. From 2001 to 2003, he was a Humboldt Research Fellow in the Institute for High-Frequency Engineering at the Munich University of Technology, Germany. His current research is in computational electromagnetics, nonlinear optics, and application of wavelets to scientific computation.

Om Gandhi (Chapter 14) is a professor of electrical engineering at the University of Utah, Salt Lake City. He is the author or coauthor of several book chapters and over 200 journal papers on electromagnetic dosimetry, microwave tubes, and solid-state devices. He also edited *Biological Effects and Medical Applications of Electromagnetic Energy* (Prentice-Hall, 1990), and coedited *Electromagnetic Biointeraction* (Plenum, 1989). Professor Gandhi was named a Fellow of the IEEE (1979), and a Fellow of the American Institute for Medical and Biological Engineering (1997). He was the chairman of the IEEE Committee on Man and Radiation (1980–1982); the cochairman of the IEEE SCC 28.IV Subcommittee on RF Safety Standards (1988–1997); and the president of the Bioelectromagnetics Society (1992–1993). He received the Distinguished Research Award from the University of Utah (1979–1980); the d'Arsonval Medal of the Bioelectromagnetics Society for pioneering contributions to the field of bioelectromagnetics (1995); the Microwave Pioneer Award of the IEEE Microwave Theory and Techniques Society (2001); and the State of Utah Governor's Medal for Science and Technology (2002).

Stephen Gedney (Chapters 7 and 11) is a professor in the Department of Electrical and Computer Engineering at the University of Kentucky, where has been a faculty member since 1991. He received a B.Eng.-Honors degree from McGill University, Montreal, Quebec, in 1985, and M.S. and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana-Champaign, Illinois, in 1987 and 1991, respectively. He was a NASA/ASEE faculty fellow at the Jet Propulsion Laboratory, Pasadena, California, in the summers of 1992 and 1993; and was a visiting professor at HRL Laboratories, Malibu, California, in 1996. In 2000, he was named the Reese Terry Professor of Electrical and Computer Engineering at the University of Kentucky. In 2004, he was named an IEEE Fellow for his contributions in the field of computational electromagnetics. His current research is in the field of computational electromagnetics, with specific emphasis on the development of advanced high-order and fast algorithms for the rigorous analysis of microwave circuits and antennas, mixed-signal systems, and electromagnetic scattering. He can be contacted via e-mail at gedney@engr.uky.edu.

Wojciech Gwarek (Chapters 9 and 15) is a professor at Warsaw University of Technology (WUT), Warsaw, Poland. He received a B.Sc. degree from WUT in 1970, an M.Sc. degree from MIT, Cambridge, Massachusetts, in 1974, and a Ph.D. degree from WUT in 1977. Since the

mid-1980s, he has been engaged in research on electromagnetic analysis in the time domain, and is the author of about 150 publications in this area. Since 1990, his activity also includes the development of the commercial software tools QuickWave and Concerto, which have become popular worldwide. Professor Gwarek is the founder and president of QWED, which commercializes electromagnetic modeling software and consults on industrial design using such software. In 1998, QWED won the prestigious European IT Prize. During his entire professional career, he has also been involved in teaching field theory, microwaves, and electromagnetic compatibility. He is a coauthor of an undergraduate textbook on field theory widely used in Poland. Since 2001, he has been a Fellow of the IEEE.

Bijan Houshmand (Chapter 15) is a senior principal engineer with BAE Systems, and an adjunct associate professor in the Electrical Engineering Department of UCLA. He received B.S., M.S., and Ph.D. degrees in electrical and computer engineering from the University of Illinois, Urbana-Champaign, in 1984, 1985, and 1990, respectively. Until 2002, he was with the Jet Propulsion Laboratory. There, he was a principal investigator for the urban remote sensing project. He also contributed to the space shuttle radar topography mission and the hazard-avoidance radar system for the Mars-landing project. He was a visiting professor at Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France, in 2000. He has contributed to two books and more than 100 other publications in the area of computational electromagnetics. His research interests are computational electromagnetics, radar systems, remote sensing, waveguides and antennas, and applications of FDTD in the design of optical waveguides. Dr. Houshmand is a Senior Member of the IEEE.

Tatsuo Itoh (Chapter 15) is a professor of electrical engineering and holder of the TRW Endowed Chair in Microwave and Millimeter Wave Electronics at UCLA, which he joined in 1991. He received a Ph.D. degree in electrical engineering from the University of Illinois, Urbana, in 1969. From 1966 to 1976, he was with the Electrical Engineering Department of the University of Illinois. From 1976 to 1977, he was a senior research engineer in the Radio Physics Laboratory, SRI International, Menlo Park, California. From 1977 to 1978, he was an associate professor at the University of Kentucky, Lexington. In 1978, he joined the University of Texas-Austin, where he became a professor of electrical engineering in 1981, the Hayden Head Centennial Professor of Engineering in 1983, and Director of the Electrical Engineering Research Laboratory in 1984. He received the Shida Award from Japanese Ministry of Post and Telecommunications (1998); the Japan Microwave Prize (1998); the IEEE Third Millennium Medal (2000); and the IEEE Microwave Theory and Techniques Society Distinguished Educator Award (2000). He was elected to the U.S. National Academy of Engineering in 2003.

Professor Itoh is a Fellow of the IEEE, a member of the Institute of Electronics and Communication Engineers of Japan, and a member of Commissions B and D of USNC/URSI. He has served as the editor of *IEEE Transactions on Microwave Theory and Techniques* (1983–1985); the chairman of USNC/URSI Commission D (1988–1990); the vice president and then president of the IEEE Microwave Theory and Techniques Society (1989, 1990); the editor-in-chief of *IEEE Microwave and Guided Wave Letters* (1991–1994); and the chairman of URSI Commission D (1993–1996). Currently, he is the chairman of the Long-Range Planning Committee of URSI, and also serves on the advisory boards and committees of a number of organizations. Dr. Itoh has 350 journal publications and 650 refereed conference presentations, and has written 30 books/book chapters in the area of microwaves, millimeter-waves, antennas, and numerical electromagnetics. He has graduated 64 Ph.D. students.

Thomas Jurgens (Chapter 12) is with Northrop-Grumman Corporation in Rolling Meadows, Illinois. He received B.S. and M.S. degrees in electrical engineering from the University of Illinois, Chicago, and a Ph.D. degree in electrical engineering from Northwestern University, Evanston, Illinois. He is a Senior Member of the IEEE, and has been on the technical review committee of several IEEE symposia. His current research interests include electromagnetic scattering and radiation, synthetic-aperture radar, particle accelerator modeling and design, and visible/infrared image processing.

Morris Kesler (Chapter 13) is a founding partner at Wide Net Technologies, Inc., a high-technology small business providing research, development, and engineering services to government and industry. He received B.S. (1984), M.S. (1984), and Ph.D. (1988) degrees in electrical engineering and computer science from MIT, Cambridge, Massachusetts. Subsequently, he was a postdoctoral fellow at the IBM Zurich Research Laboratory, where he studied semiconductor quantum well lasers. From 1990 to 2000, he was affiliated with the Georgia Tech Research Institute (GTRI), Atlanta, Georgia, where he led research programs in electromagnetic scattering, antenna arrays, novel antenna structures, and photonic bandgap structures; and attained the position of principal research engineer. From 2000 to 2003, he was a consulting engineer at PhotonEx Corporation, where he was responsible for transmission engineering and testing, and played a key role in the development of a 40-Gb/sec optical transport system. Dr. Kesler holds two patents and has published over 40 technical journal and conference papers.

Sean Krakiwsky (Chapter 20) is the cofounder and CEO of Acceleware Inc., which specializes in hardware acceleration. He received a B.Sc. degree in electrical engineering from the University of Calgary in 1994 and an M.Sc. degree in hardware acceleration of FDTD from the University of Calgary in 2005. In addition to Acceleware, he has founded and operated several other high-technology start-up companies.

Faiza Lansing (Chapter 11) has been with the Jet Propulsion Laboratory (JPL), Pasadena, California, since 1983. She received a B.S. degree from Cairo University, Giza, Egypt, in 1970, an M.S. degree from Syracuse University, Syracuse, New York, in 1976, and a Ph.D. degree from the University of Southern California (USC), Los Angeles, California, in 1981. She completed her postdoctoral studies at USC in 1982. Dr. Lansing's current research interests are in the areas of electromagnetic modeling of RF circuits and antennas, and visualization.

Xu Li (Chapter 8) is a postdoctoral fellow in the Department of Biomedical Engineering, Northwestern University, Evanston, Illinois. She received a B.S. degree in biomedical engineering from Tsinghua University, Beijing, China in 1998; and from the University of Wisconsin-Madison, she received an M.S. degree in biomedical engineering in 2000 and a Ph.D. degree in electrical engineering in 2003. Dr. Li received the Wisconsin Alumni Research Foundation Graduate Fellowship from UW-Madison in 1998, and the Second Prize in the Student-Paper Competition of the 2003 IEEE International Microwave Symposium. She has coauthored 17 refereed journal publications, three book chapters, and a number of conference papers. Dr. Li's current research interests include applied and computational electromagnetics, bioelectromagnetics, and biophotonics.

Qing Liu (Chapter 17) is a professor in the Department of Electrical and Computer Engineering of Duke University, having joined the department in 1999. He received a Ph.D. degree in electrical engineering from the University of Illinois, Urbana-Champaign, in 1989. He was a postdoctoral research associate in the Electromagnetics Laboratory of the University of Illinois, Urbana-Champaign (1989–1990); a research scientist and program leader with Schlumberger-Doll Research, Ridgefield, Connecticut (1990–1995); and an associate professor with New Mexico State University (1996–1999).

Professor Liu's research interests include computational electromagnetics and acoustics, inverse problems, geophysical subsurface sensing, biomedical imaging, electronic packaging, and the simulation of photonic and nanodevices. He has published more than 290 papers in refereed journals and conference proceedings. He is a Fellow of the IEEE and a member of the U.S. National Committee of URSI Commissions B and F. Currently, he serves as an associate editor for *Radio Science* and for *IEEE Transactions on Geoscience and Remote Sensing*, for which he also served as a guest editor for a special issue on computational methods. He received the 1996 Presidential Early Career Award for Scientists and Engineers, the 1996 Early Career Research Award from the Environmental Protection Agency, and the 1997 CAREER Award from the National Science Foundation.

James Maloney (Chapters 13 and 14) is a principal research engineer in the Signature Technology Laboratory of the Georgia Tech Research Institute (GTRI), Atlanta, Georgia. Prior to this, he was a principal research engineer and team leader at PhotonEx, Inc., a telecommunications company specializing in 40-Gb/sec ultrawideband optical transport technology. He received B.E.E. and Ph.D. degrees from the Georgia Institute of Technology, Atlanta, in 1987 and 1992, respectively. In 1994, Dr. Maloney was awarded the R.W.P. King Best-Paper Prize by the IEEE Antennas and Propagation Society, and was named honorable mention for the S. A. Schelkunoff Best-Paper Prize. He has authored more than 25 journal publications and conference papers. His current research interests include ultrawideband, transient, conformal, and low-signature antenna and radome systems; advanced electromagnetic materials realized using photonic bandgap structures and composite random materials; adaptive processing; and near-field scanning.

Michal Okoniewski (Chapter 20) is a professor and Canada Research Chair with the Department of Electrical and Computer Engineering, University of Calgary. He is also affiliated with TR-Labs, Calgary. Dr Okoniewski received M.Sc. and Ph.D. degrees from the Gdansk University of Technology. He is a Senior Member of the IEEE, an associate editor of *IEEE Transactions on Antennas and Propagation*, and a member of the IEEE Standard Groups. Professor Okoniewski is interested in many aspects of applied electromagnetics, including computational electrodynamics, reflect-arrays and self-configuring antennas, RF MEMS, and reconfigurable computational hardware for electromagnetics applications. He is also actively involved in bioelectromagnetics, where he works on tissue spectroscopy and cancer detection.

Melinda Piket-May (Chapter 15) is an associate professor in the Department of Electrical and Computer Engineering at the University of Colorado-Boulder, having joined the department in 1993. She received a B.S. degree from the University of Illinois, Urbana-Champaign, in 1988, and M.S. and Ph.D. degrees from Northwestern University, Evanston, Illinois, in 1990 and 1993, respectively. In 1997, she received the NSF CAREER award. From 2001 to 2002, she served as the interim vice-chancellor of research for the University of Colorado-Boulder campus.

Professor Piket-May has an active research program in computational electrodynamics, especially FDTD methods. Her research interests include biomedical engineering, assistive technology, high-speed analog and digital design, signal integrity, electromagnetic compatibility, solar-cell design, and wireless communications. She served on the ADCOM of the IEEE Antennas and Propagation Society (1997–2001), the ADCOM of the IEEE Education Society (2000–2003), the U.S. National Committee of URSI (1997–2001), and as an associate editor of *IEEE Transactions on Antennas and Propagation* (1997–2001).

Professor Piket-May is also active in engineering education, working for an interactive environment where the student is in charge of the learning process. In this regard, she was named a Frontiers in Education (FIE) Sloan New Faculty Fellow, and was honored with the 2000 College of Engineering Peebles Award for teaching. In addition, in 2001–2002, she was selected to participate in the Carnegie Foundation CASTLE program on the scholarship of teaching and learning, and received the 2002 Elizabeth Gee Award for scholarly contributions, distinguished teaching, and advancing women in the academic community. She was the general chair of the 2003 FIE conference, and will serve as the IEEE Education Society Program chair for FIE'2006.

Douglas Riley (Chapter 19) is a technical fellow and associate chief scientist with Northrop-Grumman Corporation. He received a Ph.D. degree in electrical engineering from Virginia Polytechnic Institute in 1986. Prior to joining Northrop-Grumman in 2002, he worked 18 years at Sandia National Laboratories, many as a distinguished member of the technical staff within the Electromagnetics and Plasma Physics Analysis Department. Dr. Riley's experience includes the development and application of both time-domain and frequency-domain computational electromagnetics. He is been actively involved with the development of efficient hybrid algorithms for nearly 20 years. Dr. Riley has published over 80 journal articles, technical reports, and conference presentations. He received a prize paper award at the 1998 IEEE Antennas and Propagation International Symposium, and has also received several research and development awards from Northrop-Grumman. His current research interests include the development and application of novel methods to address antenna design, as well as electromagnetic coupling on electrically large platforms.

Thomas Rylander (Chapter 19) is an assistant professor at Chalmers University of Technology, Göteborg, Sweden. He received an M.S. degree in electrical engineering from the Royal Institute of Technology, Stockholm, Sweden, in 1997 and a Ph.D. degree from Chalmers University in 2002. He was a postdoctoral associate at the Center for Computational Electromagnetics, University of Illinois, Urbana-Champaign, from 2002 to 2004. His research interests are in computational electromagnetics, in particular, finite-element-based methods.

Gregory Saewert (Chapter 12) has been with the Fermi National Accelerator Laboratory, Batavia, Illinois, since 1980, designing equipment for high-energy physics research. His current engineering efforts are involved with the design of systems for the precision control of high-intensity, medium-energy electron beams. He received a B.S. degree in electrical engineering technology from Northern Illinois University, De Kalb, Illinois, in 1973.

John Schneider (Chapter 5) is an associate professor in the School of Electrical Engineering and Computer Science at Washington State University. He received a B.S. degree in electrical engineering from Tulane University and M.S. and Ph.D. degrees in electrical engineering from the University of Washington. His research interests include the use of computational methods to analyze acoustic, elastic, and electromagnetic wave propagation.

Ryan Schneider (Chapter 20) is currently pursuing a Ph.D. degree in electrical engineering at the University of Calgary as a Natural Sciences and Engineering Research Council (NSERC) and Alberta Ingenuity scholar. He received B.S. and M.S. degrees in electrical engineering from the University of Calgary in 2000 and 2003, respectively. Prior to his graduate studies, he worked for two years at Telecommunications Research Labs, Calgary, on spread-spectrum modulation for voice signals and on a digital software radio platform. His current research interests include the novel application of digital hardware and computer architecture design to the acceleration of electromagnetic and other physical simulations.

Jamesina Simpson (Chapters 3 and 15) is a Ph.D. student in the Department of Electrical and Computer Engineering, Northwestern University, Evanston, Illinois. While an undergraduate at Northwestern, she was involved in research in FDTD computational electrodynamics from her freshman year on. This led to her tuition being completely supported by Intel Corporation during both her junior and senior years. Subsequently, she received a B.S. degree in electrical engineering in 2003, and was awarded both Northwestern's Walter P. Murphy Graduate Fellowship and the National Science Foundation Graduate Fellowship for her Ph.D. studies. Her current research interests are wide-ranging, including geophysics and remote sensing (via FDTD analysis of extremely-low-frequency and ultralow-frequency electromagnetic wave propagation throughout the entire Earth-ionosphere waveguide), and synthesis of novel hyperspeed board-level digital interconnects (via FDTD modeling and design of defect-mode electromagnetic bandgap waveguides). Ms. Simpson has been a summer engineering intern at Intel Corporation's Hillsboro, Oregon, facility during 2003, 2004, and 2005.

Glenn Smith (Chapter 14) is the Regents' professor of electrical and computer engineering and the John Pippin Chair in Electromagnetics at the Georgia Institute of Technology, Atlanta, Georgia. He received a B.S. degree from Tufts University, Medford, Massachusetts, in 1967; and S.M. and Ph.D. degrees in applied physics from Harvard University, Cambridge, Massachusetts, in 1968 and 1972, respectively. From 1972 to 1975, he served as a postdoctoral research fellow at Harvard. In 1975, he joined the faculty of Georgia Tech. He is the author of *An Introduction to Classical Electromagnetic Radiation* (Cambridge University Press, 1997) and the coauthor of *Antennas in Matter: Fundamentals, Theory and Applications* (MIT Press, 1981). He is a Fellow of the IEEE and a member of URSI Commissions A and B. His technical interests include electromagnetic theory and measurements, antennas and wave propagation in materials, and the radiation and reception of pulses by antennas.

Eric Thiele (Chapter 14) is a senior research engineer with ATK Mission Research (formerly Mission Research Corporation), Dayton, Ohio, where he been a staff member since 1998. He received a B.S. degree in electrical engineering from the University of Dayton, Ohio, in 1988, and M.S. and Ph.D. degrees in electrical engineering from Northwestern University, Evanston, Illinois, in 1990 and 1994, respectively. Previously, Dr. Thiele was affiliated with MRJ Inc., Oakton, Virginia, as a member of the technical staff; the University of Colorado-Boulder, as a research associate and lecturer; and the Center for PersonKommunikation, Aalborg University, Denmark, as an assistant research professor. Currently, he is managing projects in antenna and antenna-array design and analysis, radar cross section analysis, and signature exploitation/ automatic target recognition. He can be contacted at eric.thiele@atk.com.

Eli Turkel (Appendix 2A) is a professor in the Department of Mathematics at Tel Aviv University, Tel Aviv, Israel. He received a B.A. degree from Yeshiva University in 1965, and a Ph.D. degree from the Courant Institute, New York University, in 1970. He has worked in many fields, including ballistics, plasma physics, computational fluid dynamics, seismology, acoustics, and computational electromagnetics, and he is an associate fellow of the AIAA. He has consulted at the ICASE research center in NASA Langley for over 30 years, and he regularly visits the ICOMP center in NASA Lewis and the RIACS center in NASA Ames. In 1992, he received a group award from NASA as part of the ICASE numerical analysis group. His research in computational electromagnetics includes high-order methods, and the treatment of boundaries both in the far field and at solid bodies and interfaces.

Laurence Turner (Chapter 20) is the associate dean, academic, of the Faculty of Engineering at the University of Calgary, where he is a professor of electrical and computer engineering. He received B.Sc. and Ph.D. degrees in electrical engineering from the University of Calgary in 1974 and 1979, respectively. Since 1979, he has been actively involved as a researcher and teacher in the areas of digital-filter design and implementation, and digital-system design.

Geoff Waldschmidt (Chapter 5) is an RF engineer with Argonne National Laboratory, Argonne, Illinois. In 1991, he received a B.A. degree in computer science and music from Lake Forest College, Lake Forest, Illinois. He received M.S. and Ph.D. degrees in electrical engineering from Northwestern University, Evanston, Illinois, in 2000 and 2002, respectively. During his graduate studies, he worked as a laboratory graduate student at Argonne and as a research engineer with Amphenol T&M Antennas, Inc., Vernon Hills, Illinois. His Ph.D. thesis involved the development of conformal FDTD methods for analyzing a cost-effective linear accelerator capable of accelerating either protons or electrons with only minor adjustments to the accelerator geometry. His research also involved the design of compact cellphone antennas, with emphasis on reducing radiation exposure to users. His current research interests include the modeling, design, and development of high-power RF components such as dampers, input couplers, and multicell cavities for normal and superconducting subatomic-particle accelerator systems.

Christopher Wagner (Chapter 5) is a lecturer in electromagnetics at the University of Idaho. He received B.S. degrees in physics and electrical engineering from the University of Washington, and M.S. and Ph.D. degrees from Washington State University. Prior to his Ph.D. studies, he was a research engineer at the Nuclear Physics Laboratory at the University of Washington; a consultant for Science Applications, Inc., in Bellevue, Washington; and a consultant for International Sensor Technology and Keithley Instruments. His research interests include transducers, electromagnetics, and numerical methods.

Tzong-Lin Wu (Chapter 15) is an associate professor in the Department of Electrical Engineering of National Sun Yat-Sen University (NSYSU), having joined the department in 1998. He received B.S. and Ph.D. degrees from National Taiwan University, Taipei, Taiwan, in 1991 and 1995, respectively. From 1995 to 1996, he was a senior engineer with Microelectronics Technology, Inc., Hsinchu, Taiwan, a wireless/satellite communications company developing RF modules. From 1996 to 1998, he was affiliated with the Central Research Institute, Tatung Company, Taipei, Taiwan, where he was involved in the analysis and measurement of EMC/EMI problems of high-speed digital systems. Dr. Wu received the Excellent Research Award and Excellent Advisor Award from NSYSU in 2000 and 2003,

respectively, and the Outstanding Young Engineer Award from the Chinese Institute of Electrical Engineers in 2002. He is a Senior Member of the IEEE and a member of the Chinese Institute of Electrical Engineers. His research interests include time-domain modeling and measurement of EMC and signal-integrity problems in high-speed digital and optical systems, and numerical modeling of optical waveguides.

Gang Zhao (Chapter 17) is a Ph.D. student at Duke University, having received an M.S. degree in electrical engineering from Duke in 2002. In 2000, he received dual B.S. degrees in electronic engineering and economics administration from Tsinghua University. He has been a consultant to Xtreme Spectrum, Vienna, Virginia, in the area of circuit models for passive RF components; and an intern at RF Micro Devices, Inc., Greensboro, North Carolina, where he designed and modeled RF passive devices and packages. Currently, he is an intern with Intel Corporation, Chandler, Arizona, where he investigates noise-isolation issues and solutions for mixed RF/digital circuits in system-in-package applications. His research interests include computational electromagnetics, modeling of RF passive devices, and high-speed mixed-signal circuit design.

Index

Absorbing boundary conditions (ABCs), 229, 273
  analytical, 229
  perfectly matched layer (PML), 273 (see also Perfectly matched layer absorbing boundary conditions)
Accuracy, order of, 25, 45, 60, 62
Active impedance, antenna array, 599, 644
Alternating-direction implicit (ADI) algorithm, 154, 888
Alternative grids
  Cartesian, pseudospectral time-domain (PSTD), 150
  Cartesian, Yee-like, 80
  curvilinear, 471
  generalized, 487, 494
  geodesic, 99
  hexagonal, 82
  nonorthogonal, 471, 480, 486
  nonuniform, 464
  spherical, 86
Analytical absorbing boundary conditions, 229
  Bayliss-Turkel, 230
  Engquist-Majda, 236
  Higdon, 252
  Liao, 255
  Mur, 240
  one-way wave equations, 236
  Ramahi, 259
  Trefethen-Halpern, 243
Analytical field propagation (AFP), 213, 218
Antennas
  active impedance, antenna array, 599, 644
  arrays, 597, 634, 640
  biomedical applications, 669
  cellphone, 508, 647, 653
  dipole, 418, 649, 650, 652, 662, 664
  dipole array, infinite, 597
  EBG reflector, 662
  effective area, 631
  excitation, 611
  feed models, 612
  gain, 623
  grating lobes, 644
  ground-penetrating radar, 663
  helical, generic cellphone, 900 MHz, 508
  helical, Motorola T250 cellphone, 654
  horn, standard gain, 628
  monopole, PEC, 620
  monopole, resistively loaded, 432
  near-field simulations, 647
  near-to-far-field transformations, 329, 621
  plane-wave source, 204, 625
  radome interaction, 12, 667
  receiving, 609, 616
  reciprocity, 631
  scan blindness, 600
  slotline, 634
  specific absorption rate (SAR), 11, 647, 650, 652, 658
  symmetry, 610, 621
  thin-wire subcell model, 418
  transmitting, 607, 615
  Vivaldi, 634
  voltage standing-wave ratio (VSWR), 647
Autoregressive models, 697
  covariance method, 699
  forward-backward method, 701
  nonlinear predictor, 702
Auxiliary differential equation methods, 361, 376, 387, 394
  four-level two-electron atomic system, 394
  linear dispersive materials, 361
  nonlinear dispersive materials, 376
  optical gain materials, 387, 394

Backscattering from strongly forward-scattering objects, 348
Band diagrams, photonic crystal, 774, 775, 782, 793, 796
Bayliss-Turkel radiation operators, 230
  cylindrical coordinates, 234
  spherical coordinates, 231
Berenger's perfectly matched layer (PML) absorbing boundary condition (ABC), 273, 276, 282
  stretched-coordinate formulation, 282
  three-dimensional case, 281
  two-dimensional TE_z case, 276
  two-dimensional TM_z case, 281
Biophotonics, 822
Blackman window, 779, 788, 801, 866
Bloch boundaries, 776, 797, 808

Bloch equation, 394
Bodies of revolution (BOR), 517
difference equations, off-axis, 519, 525
difference equations, on-axis, 529
field expansion, 517
numerical stability, 535
particle accelerator physics, application, 543
PML ABC, 536
Bootstrapped waveguide source, 224, 745
Born-Oppenheimer approximation, 378
Bragg wavelength, 766, 767
Breast cancer detection, microwave, 11, 669
Brillouin zone, 774, 782, 793, 796, 798, 802

Cancer detection, 11, 669, 824
breast cancer, ultrawideband microwave, 11, 669
cervical cancer, optical scattering, 824
Capacitance, differential, 681
Capacitance, intrinsic space cell, 179
Capacitor, lumped, 709
Cellphone models, 11, 508, 647, 653
Characteristics of FDTD methods
classes of algorithms
dynamic range
scaling to large problems
Chebyshev polynomial algorithm, unconditional stability, 892
Chebyshev pseudospectral time-domain (PSTD) method, multidomain, 861
Chebyshev PSTD method, single domain, 857
Chimera grid, 908
Circuit elements, lumped, 706 (see also Lumped circuit elements)
Circuits, 677 (see also Electronic circuits, high-speed)
Coatings, subcell models, 447
Complementary operator method (COM), 259
Complex-frequency analysis, 40, 130
Courant factor normalization, 133
stability bounds, 132, 154
stable and unstable ranges, 131
Complex frequency-shifted (CFS) tensor, 294
Complex-valued numerical wavenumbers, 32, 120
one-dimensional scalar wave equation, 32
propagation along principal lattice axes, 121
propagation along grid diagonals, 123
Concurrent complementary operator method (CCOM), 261
Contour-path grid-cell technique, 86, 408, 411, 415, 420, 427, 432, 447, 450, 519
average properties model of diagonal and curved dielectric surfaces, 410
bodies of revolution, 519, 521, 523, 528, 529, 532
coatings on PEC surfaces, 447
diagonal / curved-surface models, 410, 420, 424, 528
latitude-longitude grid, 86
material sheets, 427
relativistically moving PEC surfaces, 450
slots in PEC screens, 411
surface impedance, 432
wires, 415
Contravariant and covariant field components, 474
Convolutional PML (CPML), 302, 310, 313
computer implementation, 307
finite-difference expressions, 302
general media, 310
numerical experiments, 313
photonics applications, 794
Courant factor, 30, 109, 133, 142
Courant stability bound, 42, 132, 154
Current, definition, 679
Curved surfaces, subcell models, 410, 420, 424
Curvilinear coordinates, 471, 858

Debye media, 354, 358, 361, 380
Diagonal / curved-surface subcell models, 410, 420, 424
Differential capacitance, 681
Differential inductance, 682
Digital signal processing, 694
Diode, lumped, 714
Dipole antenna models, 418, 649, 650, 652, 662, 664
Dirichlet boundary, 911, 913
Discontinuous Galerkin method (DGM), 873
Discrete Fourier transform (DFT), 335
Disk resonators, photonic, 748, 756
Dispersion (see Numerical dispersion)
Dispersion compensation, TF/SF technique, 213
Dispersive materials, 353
auxiliary differential equation (ADE) methods, 361, 376, 387, 394
Debye media, 354, 358, 361
Drude media, 355, 365
four-level two-electron atomic system, 394
gain materials, optical, 387, 394
Lorentz media, 354, 358, 363
magnetized ferrites, 369
nonlinear dispersive media, 376
piecewise-linear recursive-convolution (PLRC) method, 355

Dispersive numerical wave propagation in one dimension, 29
complex numerical wavenumber regime, 32
example of calculation of numerical phase velocity and attenuation, 34
examples of calculations of wave propagation, 34
general solution, 30
real numerical wavenumber regime, 31
sample values of numerical wavenumber and phase velocity, 29
superluminal propagation, 33, 38
Dispersive numerical wave propagation in two and three dimensions, 107, 110
anisotropy of numerical phase velocity, 111
comparison with ideal dispersion case, 111
complex-valued numerical wavenumbers, 120
derivation of dispersion relation, 107
extension to three dimensions, 110
intrinsic grid velocity anisotropy, 116
numerical phase velocity, sample values, 111
superluminal propagation, 126–129
Distributed Bragg reflection (DBR), 400, 765, 770, 798, 802
Double-Doppler effect, 454
Drude media, 355, 365
Dynamic range, predictive, 7

Earth-ionosphere waveguide, 10, 85
Effective area, antenna, 631
Electric and magnetic current sources, 175
sinusoidal sources, 178
sources and charging, 176
transient (pulse) sources, 178
Electromagnetic-bandgap (EBG) structures, 13, 553, 560, 595, 662, 731, 877
interconnect, 30 to 70 GHz, 13, 731
reflector, antenna, 662
reflector, microwave, 560, 877
woodpile, 595
Electromagnetic pulse (EMP), 2
Electronic circuits, high-speed, 677
capacitance, differential, 681
capacitance, intrinsic space cell, 179
capacitor, lumped, 709
circuit elements, lumped, 706
circuit parameters, basic, 679
current, definition, 679
digital signal processing, 694
diode, lumped, 714
electromagnetic bandgap (EBG) interconnect, 13, 731
impedance, definition, 680
inductance, differential, 682
inductance, intrinsic space cell, 183
inductance, lumped, due to a circuit discontinuity, 682
inductance, power-distribution system, 685
inductor, lumped, 711
interconnect, multilayered, 690
lumped circuit elements, modeling, 706
MESFET amplifier model, 723
microstrips, parallel coplanar, 688
resistive voltage source, lumped, 708
resistor, lumped, 708
S-parameters, definition, 680
S-parameter extraction for general waveguides, 692
signal integrity, 678
spectrum estimation, 694
SPICE, linkage to FDTD, 717
transistor, bipolar junction, lumped, 715
transmission line parameters, definition, 679
two-terminal network, arbitrary linear, lumped, 711
voltage, definition, 679
ELF propagation, 10, 85, 96
Engquist-Majda one-way wave equations, 236
Mur finite-difference scheme, 240
numerical experiments, 247
one- and two-term Taylor series, 237
theoretical reflection coefficient analysis, 245
Trefethen-Halpern generalization, 243
Equivalence theorem, surface, 338
Exponential autoregressive (EXPAR) model, 702
Extended FDTD formulation, lumped circuit elements, 706
Extinction ratio, optical resonators, 755

Fabry-Perot cavities, 390, 765
Fast Fourier transform, nonuniform, 855
FDTD acronym, origination, 4
FDTD modeling applications
antenna, dipole, driving-point characteristics, 418
antenna, dipole, 12 cm, 900 MHz, *E*-field in layered bone-brain half-space, 649
antenna, dipole, 17.4 cm, 840 MHz, SAR in rectangular brain phantom, 650
antenna, dipole, 7.7 cm, 1.9 GHz, SAR in spherical brain phantom, 652
antenna, dipole, infinitesimal, 900 MHz, SAR in spherical brain phantom, 650
antenna, dipole, EBG reflector, 662

antenna feed, one-way signal injector, 618
antenna, helical, generic cellphone, 900 MHz, 508
antenna, helical, Motorola T250 cellphone, 654
antenna, horn, standard gain, 628
antenna, infinite array of dipoles, 597
antenna, monopole, PEC, 620
antenna, monopole, resistively loaded, 432
antenna, Vivaldi slotline, 634, 640
antenna-radome interaction, 12, 667
breast cancer detection, microwave, 11, 669
capacitor, lumped, microstrip termination, 710
cavity, rectangular PEC, dielectric loaded, 705
cavity, twisted elliptical PEC, 421
cellphone, generic, 900 MHz, helical antenna, 508
cellphone, generic, 900 MHz, monopole antenna, 647
cellphone, Motorola T250, 11, 653
cervical cancer detection, optical, 824
circuit-board, meshed multiplane, power-distribution system, 687
coaxial line, transmission dc to 20 GHz, skin effect modeled, 444
computer circuit-board module, 690
dielectric resonator, isolated cylindrical, 697
diode, lumped, microstrip termination, hard clipping, 715
electromagnetic bandgap (EBG) interconnect, 30 to 70 GHz, 13, 731
EBG reflector, antenna, 662
EBG reflector, microwave, 560
EBG, woodpile, 595
extremely low frequency (ELF) propagation, 10, 96
frequency-selective surface (FSS), 597
gain medium, optical, 390
ground-penetrating radar, 663
line current above lossy half-space, 440
laser, 1D Fabry-Perot microcavity, classical gain model, 390
laser, 1D distributed Bragg reflector (DBR) microcavity, semi-classical gain model, 400
laser, 1D vertical-cavity surface-emitting (VCSEL), 765
laser, photonic crystal microcavity, 14, 810
MESFET amplifier, 6 GHz, 723
microstrip bandpass filter, 699
microstrip double-stub structure, 702
microstrip low-pass filter, 469
microstrips, parallel coplanar, 688
particle accelerators, 547
particles, micron size, nanometer-scale refractive-index fluctuations, optical backscattering, 827
periodic screen, lossy, two-dimensional, 563, 566, 577, 583
photonic crystal, square lattice, 774
photonic crystal, triangular lattice, 782
photonic crystal, triangular lattice, finite-thickness slab, 793
photonic switch, cross-waveguide, 15, 812
racetrack, photonic, 752, 761
radar cross section (RCS), closely spaced pair of PEC spheres, 424
radar cross section, PEC winglike object, 422
radome-antenna interaction, 12, 667
relativistically moving / vibrating PEC surfaces, plane-wave scattering, 454
resonator, air-bridge, photonic, quasi-1D DBR structure, 770
resonator, circular disk, photonic, 756
resonator, circular ring, photonic, 750, 752
retinal rod, vertebrate, 822
slotted PEC screen, penetrating fields, 413
solitons, spatial optical, 383
solitons, temporal optical, 381
transistor, bipolar junction, 717
two-terminal linear lumped network, arbitrary, parallel microstrip termination, 712
waveguide, cylindrical PEC, PML ABC, 543
waveguide, defect mode, triangular-lattice photonic crystal, 795
waveguide, dielectric slab, photonic, 746
waveguide, lossy sheet loading, 430
waveguide, lossy walls, 440
waveguides, defect mode, square-lattice and triangular-lattice photonic crystals, PML terminations, 804
Wilkinson power divider, 32 GHz, 499
woodpile EBG structure, 595
FDTD / finite-element (FE) hybrid techniques 907
FDTD / FE interface, stable, 924
finite elements, time domain, 910
hexahedral (brick) finite elements, 918, 921
mesh generation, 935
pyramidal finite elements, 918, 922
subcell thin-slot model, 939
subcell thin-wire model, 936
tetrahedral finite elements, 918, 919
FDTD / FE hybrid techniques, modeling applications
loop antenna, thin wire, 941

radar cross section, NASA almond, 943
radar cross section, Saab *Trainer* aircraft, 12, 945
scattering from PEC sphere, 933
sinuous antenna, four arm, 948
thin-wire / thin-slot coupling, 941
FDTD / FE interface, stable hybrid, 924
alternative time-stepping schemes, 930
extensions of hybrid FDTD-FE concept, 931
generalized Newmark scheme, 928
proof of stability, 929
reflection at FDTD-FE interface, 931
spatial discretization, 924
time stepping, 927
FDTD, hardware acceleration, 955
background literature, 956
design considerations, 957
field-programmable gate-array (FPGA) implementations, 958, 962
graphics processor units (GPUs), fundamentals, 965
GPU, FDTD implementation, 969
performance, custom hardware, 964
performance, GPU, 977
Ferrites, linear magnetized, 369
Field-programmable gate-array (FPGA) hardware acceleration of FDTD, 958, 962
Filter diagonalization method, 780, 800
Fine geometrical features, modeling, 407
Finesse, 747
Finite differences, 23, 24, 60
Finite-difference time-domain (FDTD) methods
characteristics, 3, 6
history, partial, 4
modeling applications (see Examples of FDTD modeling applications)
Finite elements, time domain, 910, 918
coupled curl equations, 910
equivalences between finite elements and FDTD, 917
hexahedral (brick) finite elements, 918, 921
pyramidal finite elements, 918, 922
tetrahedral finite elements, 918, 919
wave equation, 913
Finite-volume time-domain (FVTD) methods, 480, 908
Floquet boundaries, 776, 784, 797, 802, 808
Floquet modes, 555, 600, 926
Four-level two-electron atomic system, 394
Fourier PSTD method, single domain, 150, 854
Fourth-order accurate spatial differences, 139
Free spectral range (FSR), 747
Frequency conversion, 813
Frequency-domain near-to-far-field (FD-NTFF) transformation, 329, 340, 624
Frequency-selective surface (FSS), 553, 597

Gain, antenna, 623
Gain materials, optical, 387, 394
Galerkin's method, 910, 912, 928, 949
Gauss-Chebyshev-Lobatto (GCL) points, 851, 865
Gauss-Legendre-Lobatto (GLL) points, 873
Generalized multipole technique, 424
Generalized Newmark scheme, 928
Generalized Yee algorithm, 487, 494
Geodesic grid, 99
Graphics processor unit (GPU) hardware acceleration of FDTD, 965
Grating lobes, 555, 644
Green functions, 171, 331, 340, 806
Green's theorem, 330
Grids
alternative, 80, 86, 99, 146, 150, 463
bodies of revolution, 517
contour-path subcell models, 408
generalized Yee, 487, 494
nonorthogonal and unstructured, 463
periodic, 553
Yee, 58, 59
Ground-penetrating radar, 663

Hamiltonian, 394
Hard sources, 169
Hardware acceleration of FDTD, 955 (see FDTD, hardware acceleration)
Hexagonal FDTD grids, 82, 99, 146
Higdon radiation operators, 252
High-order methods, 47, 139, 150
High-power microwaves (HPM)
High-speed electronic circuits, 677 (see also Electronic circuits, high-speed)
History of FDTD methods, partial, 4
Horn antenna model, 628
Hybrid FDTD / FE techniques, 907 (see FDTD / FE hybrid techniques)

Impedance, definition, 680
Incident-wave source conditions, 169, 625
dispersion compensation, total-field / scattered-field technique, 213
electric and magnetic current sources, 175
hard sources, 169, 171
plane-wave source condition, 185, 625
scattered-field formulation, 220

total-field / scattered-field (TF/SF) technique, 186, 193, 204, 213, 628
waveguide source conditions, 223
Index-contrast guided-wave structures, 743
Inductance, differential, 682
Inductance, intrinsic space cell, 183
Inductance, lumped, circuit discontinuity, 682
Inductance, power-distribution system, 685
Inductor, lumped, 711
Intrinsic grid velocity anisotropy, 116
Irregular nonorthogonal FDTD grids, 480, 486
structured, 480
unstructured, 486

Kerr nonlinearity, 378, 380, 383

Lasers, microcavity
gain model, classical, 390
gain model, semiclassical, 400
photonic crystal microcavity, 14, 810
vertical-cavity surface-emitting (VCSEL), 765
Latitude-longitude grid, 86
Lattice cell capacitance, intrinsic, 179
Lattice cell inductance, intrinsic, 183
Lax-Richtmyer equivalence theorem, 47
Leontovich impedance boundary condition, 434
Liao absorbing boundary condition, 255
Light line, photonic crystal, 793, 795, 802, 803
Linear dispersive materials, 361
Linear tapered slot antenna, 634
Lobatto quadrature, 910
Lorentz media, 354, 358, 363, 380
Lumped circuit elements, modeling, 708
capacitor, 709
diode, 714
extended FDTD formulation, 706
inductor, 711
resistive voltage source, 708
resistor, 708
SPICE linkage to FDTD, 717
transistor, bipolar junction, 715
two-terminal linear network, arbitrary, 711

Magic time-step, 25, 29
Magnetized ferrites, 369
Matched numerical dispersion (MND), 213, 214
Material sheets, subcell models, 427
Matrix-exponential solution, Maxwell's equations, 884
Maxwell's equations, 1, 51, 858, 871, 873, 884
Ampere's law, 51
conservation form, 858
curvilinear coordinates, 473, 858
electric and magnetic loss, 53
electric and magnetic source currents, 53
Faraday's law, 51
Gauss' law for the electric field, 52
Gauss' law for the magnetic field, 52
matrix-exponential solution, 884
one-dimensional case, 56
Riemann solver, 871
three-dimensional case, 51
two-dimensional cases, 54
weak form, 873, 875
Yee algorithm, 58
MESFET amplifier model, 6 GHz, 723
Mesh generation, hybrid FDTD / FE, 935
Microcavities, photonic, index-contrast, 747, 756, 761
Microcavities, photonic crystal, 14, 807, 810
Microcavity lasers, 14, 390, 400, 765, 810
Modeling applications, FDTD (see FDTD modeling applications)
Modified Yee algorithms, 138
centered phase-velocity distribution, 138
fourth-order spatial differences, 139
generalized, 487, 494
hexagonal grids, 146
pseudospectral time-domain (PSTD), 150, 847
Monopole antenna models, 432, 620
Mur ABC, 240, 914

Nanoplasmonic devices, 820
Near-to-far field (NTFF) transformations, 329, 621
backscattering from strongly forward-scattering objects, 348
frequency domain, three dimensions, 340, 624
frequency domain, two dimensions, 329
obtaining phasor quantities using the discrete Fourier transform, 335
radar cross section, definition, 335, 343
surface equivalence theorem, 338
symmetry, use of, 621
time domain, three dimensions, 343, 622
Newmark-Beta scheme, 916
Newmark scheme, 915, 930
Newmark scheme, generalized, 928
Nonlinear electronic circuit elements, 714, 715, 717
Nonlinear dispersive materials, 376, 387, 394
Nonlinear materials, second-order, optical, 813
Nonlinear photonic crystal (NPC), 817
Nonlinear Schrödinger equation (NLSE), 384

Nonorthogonal and unstructured grids, 463
generalized Yee algorithm, 487, 494
global curvilinear coordinates, 471
irregular nonorthogonal structured grids, 480
irregular nonorthogonal unstructured grids, 486
locally conformal grids, globally orthogonal, 471
nonuniform orthogonal grids, 464
planar generalized Yee (PGY) algorithm, 494
subgrids, Cartesian, 501
Norton equivalent circuit, FDTD space lattice, 719
Numerical dispersion, 27, 107, 213, 856
compensation for total-field/scattered-field technique, 213
complex-valued wavenumbers, 32, 120
fourth-order spatial differences, 139
hexagonal grids, 146
modified Yee algorithms, 138
numerical phase velocity, 111
numerical wavevector, 109
one-dimensional scalar wave equation, 27
periodic-structure models, 559, 569, 574, 581
pseudospectral time-domain method, 150, 856
three-dimensional propagation, 110
two-dimensional propagation, 107
wave-propagation examples, 34, 126
Numerical instability, examples
one-dimensional propagation, 43
two-dimensional propagation, 135, 136
Numerical phase velocity, 30, 111
anisotropy, 111
intrinsic grid anisotropy, 116
sample values, 29, 111
superluminal, 38, 123, 126–129
Numerical stability, 39, 128, 535, 856, 883
alternating-direction implicit (ADI) algorithms, 154
body-of-revolution algorithms, 535
complex-frequency analysis, 40, 130
Courant factor, 30, 109, 132–134, 142
Courant stability bound, 42, 132–134, 142, 154
examples of numerical instability, 43, 135
generalized stability problem, 137
generalized Yee algorithm, 491
irregular nonorthogonal structured grids, 485
nonorthogonal curvilinear grids, global coordinates, 477
periodic-structure models, 570, 573, 580, 589
planar generalized Yee algorithm, 499
pseudospectral time-domain method, 856
subcell models, 457
unconditionally stable techniques, 883
Yee algorithm, 128
Zheng-Chen-Zhang alternating-direction implicit algorithm, 155

One-dimensional scalar wave equation (see Scalar wave equation, one-dimensional)
One-way wave equations, 236
Optical gain media, 387, 394
Optical spatial solitons, 383
Optical temporal solitons, 381
Order-*N* method, photonic crystal band structure, 775, 806
Order of accuracy, 25, 45, 60, 62, 468

Padé approximation, 702
Particle accelerators, application, 547
Particle-in-cell (PIC) algorithms, 546
Pauli Exclusion Principle, 394
Penalty methods, 868, 908
Pencil-of-function technique, 702
Perfectly matched layer (PML) absorbing boundary conditions (ABCs), 273, 536, 590
anisotropic PML, 285
Berenger's PML, 273, 276, 282
body-of-revolution algorithm, 536
complex frequency-shifted (CFS) tensor, 294
convolutional PML (CPML), 302, 310, 313
numerical experiments, 313
periodic structures, 590
photonics, applications, 744, 794, 804
stretched-coordinate formulation, 282
theoretical performance, 291
uniaxial PML (UPML), 285, 297, 590
Periodic structures, 553, 765
absorbing boundary condition, 590
dipole antenna array, infinite, 597
direct field methods, 559
distributed Bragg reflector (DBR) structures, 765, 770
electromagnetic bandgap (EBG) structures, 13, 553, 560, 595
field-transformation techniques, 567
Floquet modes, 555
frequency-selective surface (FSS), 553
grating lobes, 555
lossy materials, 575, 582
multiple-grid approach, 571
numerical dispersion, 559, 569, 574, 581
numerical stability, 570, 573, 580, 589
photonic crystals, 14, 743, 772, 817

scattering from, review of, 555
split-field method, 578, 583
woodpile EBG structure, 595
Photonic bandgap, 772
Photonic crystals, 14, 15, 743, 772, 817
band-structure calculation, 774
mode-pattern calculation, 787
nonlinear, 817
resonators, 14, 807
variational approach, 790
waveguides, defect mode, 15, 791
Photonics, 14, 15, 743
band structure, photonic crystals, 774, 790, 793, 795
biophotonics, 743, 822
cross-waveguide switch, 15, 812
defect-mode waveguides, photonic crystals, 15, 791
disks, laterally coupled, 756
distributed Bragg reflector (DBR) structures, 765, 770
FDTD modeling issues, 744, 748, 778, 784, 820
frequency conversion, 743, 813
lasers, microcavity, 14, 743, 765, 810
microcavities, DBR, 765, 767, 770
microcavities, index-contrast, 747, 756, 761
microcavities, photonic crystal, 14, 807, 810
mode patterns, photonic crystals, 787
nanoplasmonic devices, 743, 820
nonlinear materials, second-order, 743, 813, 817
periodic structures, 743, 765
photonic crystals, 14, 15, 743, 772, 817
plasmons, surface, 820
quasi-phase-matching (QPM) techniques, 813
racetracks, laterally coupled, 752
racetracks, vertically coupled, 761
rings, laterally coupled, 747, 752
second-harmonic generation (SHG), 814
tissue optics, 828
variational approach, computation of band structure of photonic crystals, 790
vertical-cavity surface-emitting lasers (VCSELS), 743, 765
waveguides, index contrast, 744, 750, 760
waveguides, photonic crystals, defect-mode, 15, 791, 795, 798, 803, 806
whispering-gallery modes, disks, 743, 756, 760
Piecewise-linear recursive-convolution (PLRC) method, 355
Planar generalized Yee (PGY) algorithm, 494
Plane-wave source, 185, 201, 625
Plasmons, surface, 820
PML (see Perfectly matched layer absorbing boundary conditions)
Prony's method, 438, 695, 780
Pseudospectral time-domain (PSTD) method, 150, 813, 828, 847
approximation of derivatives, 847, 855, 860, 862
Chebyshev method, multidomain, 861
Chebyshev method, single domain, 857
discontinuous Galerkin method, 873
Fourier method, single domain, 150, 854
numerical stability and dispersion, 150, 856
penalty method, multidomain, 868
PSTD modeling applications
bistatic radar cross section of finite-length lossy dielectric cylinder, 868
electric dipole adjacent to heat sink slab, 878
electric dipole adjacent to PEC cube, 868
electromagnetic bandgap structure, 877
line current source in cylindrical hole within infinite lossy dielectric medium, 866
plane wave impinging upon dielectric cube, 868
point electric dipole in cylindrical hole within infinite lossy dielectric medium, 866
second-harmonic generation (SHG), 814, 817
tissue optics, 828

Quasi-phase-matching (QPM) techniques, 813

Radar, military defense applications, 2
Radar cross section (RCS), 12, 335, 343, 422, 424, 943, 945
Radome-antenna interaction, 12, 667
Raman nonlinearity, 378, 379, 380
Rate equations, 397
Reciprocity, 631
Rejection ratio, optical resonator, 756
Relativistically moving PEC surface, 450
Resistive voltage source, lumped, 708
Resistor, lumped, 708
Riemann solver, Maxwell's equations, 871
Runge-Kutta method, 855, 861, 879

S-parameters, definition, 680
Scalar wave equation, one-dimensional, 21
dispersion relation, 22
finite differences, 23, 24
magic time-step, 25, 29

numerical dispersion, 27
numerical stability, 39
order of accuracy of finite-difference approximation, 25, 45
propagating-wave solutions, 21
Scaling of FDTD models to large sizes
Scan blindness, antenna array, 600
Scattered-field formulation, 220
application to dielectric structures, 221
application to PEC structures, 220
Scattering (see Radar cross section)
Second-harmonic generation (SHG), 814
Second-order nonlinear optical materials, 813
Semi-implicit approximation, 63
Sidewall roughness, optical resonator, 755
Signal integrity, 678
Silver-Müller condition, 914
Slots, subcell models, 411
Slotline antennas, 634
Solitons, optical
spatial, 383
temporal, 381
Sources (see Incident wave source conditions)
Spatial optical solitons, 383
Specific absorption rate, 647, 650, 652, 658
Spectrum estimation, 694, 780
autoregressive models, 697
filter diagonalization method, 780
Padé approximation, 702
Prony's method, 695
SPICE, linkage to FDTD, 717
basic idea, 718
Norton equivalent circuit, FDTD space lattice, 719
Thevenin equivalent circuit, FDTD space lattice, 721
Stability (see Numerical stability)
Subcell models, 407, 936
bodies of revolution, 528
coatings, 447
contour-path models, 408, 411, 415, 420, 427, 432, 447, 450
diagonal and curved surfaces, 410, 420, 424
material sheets, 427
relativistically moving PEC surfaces, 450
surface impedance, 432
thin slots in PEC screens, 411, 939
thin wires, 415, 936
Subgrids, Cartesian, 501
Superluminal propagation, 34, 126
Supraconvergence, 464, 467
Surface equivalence theorem, 338
Surface impedance, subcell models, 432
frequency-dependent, convolution-based, 436
frequency-dependent, equivalent-circuit model, 442
monochromatic, 434
Symmetry, 610, 621, 794

Tapered slot antenna, 634
Temporal optical solitons, 381
Thevenin equivalent circuit, FDTD space lattice, 721
Time-domain near-to-far-field (TD-NTFF) transformation, 343, 622
Tissue optics, 828
Total-field / scattered-field (TF/SF) technique, 186, 193, 204, 213, 628
numerical dispersion compensation, 213
one-dimensional formulation, 188
three-dimensional formulation, 204, 628
two-dimensional formulation, 193
Transistor, bipolar junction, modeling, 715
Transmission line parameters, 679
Trefethen-Halpern absorbing boundary conditions, 243

Unconditionally stable techniques, 154, 883
alternating-direction implicit algorithm, 154, 888
Chebyshev polynomial algorithm, 892
Fourier-space approach, 891
general framework, 883
linear dispersive media, extension, 895
matrix-exponential concepts, 884
product-formula approach, 887
real-space approach, 889
uniaxial PML, extension, 898
Uniaxial PML (UPML) ABC, 285, 297, 590
Unstructured grids (see Nonorthogonal and unstructured grids)

Vertical-cavity surface-emitting lasers (VCSELs), 743, 765
Vivaldi slotline antennas, 634
planar element, 635
planar elements, pair, 637
quad element, 639
quad elements, linear phased array, 640
Voltage, definition, 679
Voltage standing-wave ratio (VSWR), 647

Wake fields, particle accelerator, 545

Wave source conditions (see Incident wave source conditions)
Waveguides, photonic, index contrast, 744, 750, 760
Waveguides, photonic crystal, defect mode, 15, 791
aperiodic, 806
band diagram, 795
band diagram, photonic-crystal slab, 793
cross-waveguide switch, 15
extrinsic scattering loss, 806
intrinsic loss, 798
transmission, 803
Waveguide source conditions, 223
Wavelength-division multiplexing (WDM), 747
Whispering-gallery modes, microcavity disk resonator, 743
Wires, subcell models, 415

Yee algorithm, 58, 887
alternative grids, 80
basic ideas, 58
divergence-free nature, 78
finite-difference expressions, 62, 67, 69, 71, 73
finite-differences and notation, 60
generalized, 487, 494
interpretation as Faraday's and Ampere's laws in integral form, 75
modified for reduced dispersion, 138
numerical dispersion, 107
numerical stability, 128
planar generalized Yee (PGY) algorithm, 494
product-formula approach realization, 887
semi-implicit approximation, 63
space lattice, 59
space region with continuous variation of material properties, 67
space region with finite number of distinct media, 69
space region with nonpermeable media, 71
space-time chart, 61
three dimensions, 62
two-dimensional TM_z and TE_z modes, 73

Zheng-Chen-Zhang (ZCZ) ADI algorithm, 155
additional accuracy limitations, 164
formulation, 155
numerical dispersion, 163
numerical stability, 161
sources, 161