

19980922 052

14th Annual Review of Progress in

APPLIED COMPUTATIONAL ELECTROMAGNETICS

at the

Naval Postgraduate School

Monterey, CA

March 16-20, 1998

QUALITY

CONFERENCE PROCEEDINGS

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

AQI98-12-1171

CONFERENCE PROCEEDINGS

VOLUME II

14th Annual Review of Progress in APPLIED COMPUTATIONAL ELECTROMAGNETICS

at the

Naval Postgraduate School

Monterey, CA

March 16-20, 1998

TECHNICAL PROGRAM CHAIRMAN

Jianming Jin

Sponsored by

The Applied Computational Electromagnetics Society

Naval Postgraduate School, University of Illinois at Urbana-Champaign,

Brigham Young University, University of Kentucky, University of Nevada

THE NAVAL POSTGRADUATE SCHOOL

AQ N98-12-2419

Contents

Table of Contents i

Agenda xiii

VOLUME II

SESSION 1: CEM ANALYSIS AND APPLICATION WITHIN AN ENGINEERING ENVIRONMENT
Chairs: Kenneth R. Siarkiewicz and Donald R. Pflug

"Frameworks - Future Paradigm for CEM Analysis"
K.R. Siarkiewicz 2

"REF CEM Data Dictionary"
T. Wharton and J. Evans 9

"Air Force Research and Engineering Framework Control Panel"
D.M. Hallatt and J.A. Evans 15

"Continued Application of the Research and Engineering Framework (REF) at Raytheon"
J. Labelle, B. Hantman, H. Wright, Y. Chang, and R. Abrams 20

"The Innovative Research Testbed: A PC-Based High Performance Computing and Web-Based Collaboratory for Computational Electromagnetics"
D.M. Leskiw, G. Ingersoll, T. Vidoni, and R. Redmond 28

"Illustrating the Application of Expert Systems to Computational Electromagnetics Modeling and Simulation"
A.L.S. Drozd, T.W. Blocher, C.E. Carroll, Jr. and J.M. Allen 36

" Some Present and Future Aspects of the Quality of Solution in Computational Electromagnetics"
D.R. Pflug 42

"On the Use of Computational Electromagnetics in the Radar Cross Section Measurement Calibration Process"
K.C. Hill and W.D. Wood, Jr. 50

"Radar Calibration at Low Frequencies Using a Triangular Trihedral Corner Reflector"
C.Y. Shen 57

SESSION 2: FINITE-DIFFERENCE TIME-DOMAIN APPLICATIONS
Chairs: John Beggs and Melinda Piket-May

"Time Domain Analysis of Small Multi-Sector Monopole Yagi-Uda Array Antenna Mounted on a Finite Ground Plane Using FDTD Method"
T. Maruyama, K. Uehara, T. Hori and K. Kagoshima 66

"Antenna Performance Calculation in Lossy Media with FDTD Method"
M. Cai and N. Ljepojevic 73

"Far-Zone Transformation in FDTD for VHF-band SAR-image Simulations"
T. Martin and L. Ulander 79

"FDTD Simulations Used to Correct for Ground Effects in Aircraft Testing"
G. Eriksson and U. Thibblin 87

"Validation of FDTD-Computed Handset Patterns by Measurement"
C.W. Trueman, S.J. Kubina, J.E. Roy, W.R. Lauber and M. Vall-llossera 93

"FDTD Analysis of Flip Chip Interconnects"
A.Z. Elsherbeni, V. Rodriguez-Pereyra and C.E. Smith 99

"FDTD Analysis of the Celestron-8 Telescope"
R.R. DeLyser 105

"Modifying a Graphically-Based FDTD Simulation for Parallel Processing"
G. Haussmann, M. Piket-May and K. Thomas 113

"A Time Domain Method for High Frequency Problems Exploiting the Whitney Complex", A. Arkko, T. Tarhasaari and L. Kettunen 121

"Treatment of Boundaries in Multiresolution Based FDTD Multigrid"
K. Goverdhanam, E. Tentzeris and L.P.B. Katehi 127

SESSION 3: APPLICATION OF ANALYTIC AND COMPUTATIONAL METHODS TO THE MODELING OF ELECTROMAGNETIC MATERIALS"
Chairs: Rudy Diaz and David H.Y. Yang

"Computing Dispersion Relations and Radiation Spectra in Photonic Band Gap Materials by Plane Wave Expansion Method"
T. Suzuki and P.K.L. Yu 136

"Waveguides in Photonic Band Gap Materials"
M.M. Sigalas, R. Biswas, K.M. Ho, C.M. Soukoulis and D.D. Crouch 144

"Is There a Relationship Between a Random and an Ordered Composite Mixture?"
- STUDENT PAPER CONTEST -
W.M. Merrill, S.A. Kyriazidou and N.G. Alexopoulos 151

"An Analytic Framework for the Modeling of Effective Media"
R.E. Diaz 159

SESSION 3: APPLICATION OF ANALYTIC AND COMPUTATIONAL METHODS TO THE MODELING (cont)

"Finite Grain Boundary Effects for Magnetic Materials: Tunneling and Intergrain Gaps"
G.G. Bush 167

"Numerical Computation of the Complex Dielectric Permittivity: FFT-Based Hilbert Transform Approximation of the Kramers-Kronig Relations"
- STUDENT PAPER CONTEST - , F. Castro and B. Nabet 173

"Rayleigh Analysis of Novel Dense Medium Exhibiting Narrow-Band Transparency Window", S.A. Kryiazidou, R.E. Diaz and N.G. Alexopoulos 179

SESSION 4: INTERACTIVE POSTER SESSION

"Finite Difference Formulation - Expanding the Capabilities"
K. Davey 188

"Time and Frequency Domain Numerical Modeling of Outbound and Standing Power From Perpendicularly Oriented, Electrically Small TM Dipoles"
G. Liu, C.A. Grimes K.G. Ong, D.M. Grimes 194

"Efficient Analysis of Large Two-dimensional Arbitrarily Shaped Finite Gratings for Quantum Well Infrared Photodetectors"
V. Jandhyala, D. Sengupta, B. Shanker, E. Michielssen, M. Feng, G. Stillman 204

"Forward and Backward Propagation Algorithms Applied to the Electromagnetic Scattering by an Inpenetrable Obstacle: A Progress Report"
G.F. Crosta 210

"Analysis of Broad Wall Slots Excited by Tuning Screws"
T. Azar and R. Coren 216

"Closed-Form Expression of the Arc-Length of the Toroidal/Helical Equilibrium Orbit"
R.A. Speciale 224

"Asymptotic Techniques in Naval Ship Design"
R. Routier and R. Burkholder 243

"A BCG-FFT Solution of Scattering and Radiation by Large Finite Arrays of Microstrip Antennas", C-Fu Wang, F. Ling and J-M. Jin 250

"Analytic Solution for Low-Frequency Electric Induction in an Equatorially Stratified Sphere", T.W. Dawson 258

"Parallelizing Computational Electromagnetics Code Using the Parallel Virtual Machine: Examples"
E. Skochinski and S. Rengarajan 264

"SAF Analysis Codes for Predicting the Electromagnetic Effectiveness of Antenna Enclosed in Composite Structures"
B.J. Cown, J.P. Estrada and R. Routier 270

SESSION 4: INTERACTIVE POSTER SESSION (continued)

"Some Concentrated Solutions of Helmholtz Equation with Nonlocal Nonlinearity"
Y.N. Cherkashin and V.A. Eremenko 278

"Hysteresis and Eddy Currents in Ferromagnetic Media"
J. Fuzi and A. Ivanyi 283

"Electrical Circuit Analysis Considering Hysteresis in Coil Cores"
J. Fuzi 291

"A Paradigm for Proving the Convexity Properties of Slowness Curves"
A.R. Baghai-Wadji (8 pages) 297

"On the Strict-Convexity of the Slowness Surfaces for the Fastest Bulk-Acoustic-Waves in Piezoelectric and Piezoelectromagnetic Media"
A.R. Baghai-Wadji 305

"Solving EMC Problems Using the FDTD Method"
F. Gisin and Z. Pantic-Tanner 313

SESSION 5: TLM MODELING AND APPLICATIONS
Chairs: Wolfgang J.R. Hoefer and Peter Russer

"A Hybrid Time Domain TLM-Integral Equation Method for Solution of Radiation Problems", L. Pierantoni, S. Lindenmeier and P. Russer 320

"Comparison of Symmetric Condensed TLM, Yee FDTD and Integer Lattice Gas Automata Solutions for a Problem Containing a Sharp Metallic Edge"
N. Simons, R. Siushansian, J. LoVetri, G. Bridges and M. Cuhaci 326

"Some Observations on Stubs, Boundaries and Parity Effects in TLM Models"
D. De Cogan and C. Kun 331

"Modelling of Dispersive Media in TLM Using the Propagator Approach"
J. Rebel and P. Russer 336

"A Hybrid Time Domain TLM-Integral Equation Method for Solution of Radiation Problems", L. Pierantoni, S. Lindenmeier and P. Russer 320

"Comparison of Symmetric Condensed TLM, Yee FDTD and Integer Lattice Gas Automata Solutions for a Problem Containing a Sharp Metallic Edge"
N. Simons, R. Siushansian, J. LoVetri, G. Bridges and M. Cuhaci 326

"Some Observations on Stubs, Boundaries and Parity Effects in TLM Models"
D. De Cogan and C. Kun 331

"Modelling of Dispersive Media in TLM Using the Propagator Approach"
J. Rebel and P. Russer 336

"Characterization of Quasiplanar Structures Using the TLM Method"
O. Pertz, U. Muller and A. Beyer 344

SESSION 5: TLM MODELING AND APPLICATIONS (continued)

"Generation of Lumped Element Equivalent Circuits from Time-Domain Scattering Signals", T. Mangold and P. Russer 351

"TLM Analysis of an Optical Sensor"
R.R. DeLyser 359

"TLM Modeling and TDR Validation of Soil Moisture Probe for Environmental Sensing"
G. Tardioli, M. Righi, L. Cascio, W.J.R. Hoefer and R. McFarlane 367

"TLM Analysis of the Celestron-8 Telescope"
R.R. DeLyser 374

"Near-to Far Field Transformation Via Parabolic Equation" - STUDENT PAPER CONTEST -
A.N. Kurokhtin, Y.V. Kopylov, A.V. Popov and A.V. Vinogradov 381

SESSION 6: FREQUENCY-DOMAIN FAST ALGORITHMS
Chairs: Jiming Song and Weng Cho CHew

"Recent Advances in the Numerical Solution of Integral Equations Applied to EM Scattering from Terrain"
P. Cullen and C. Brennan 390

"Solution of Combined-Field Integral Equation Using Multi-Level Fast Multipole Algorithm for Scattering by Homogeneous Bodies"
X.Q. Sheng, J.M. Jin, J.M. Song, W.C. Chew and C.C. Lu 397

"Comparisons of FMM and AIM Compression Schemes in Finite Elements - Boundary Integral Implementations for Antenna Modeling"
K. Sertel, D.S. Filipovic, S. Bindiganavale and J.L. Volakis 405

"High-Order Nystrom Discretization for Faster, More Accurate Scattering Calculations"
L.S. Canino, J.J. Ottusch, M.A. Stalzer, J.L. Visher and S.M. Wandzura 413

"Large Scale Computing with the Fast Illinois Solver Code - Requirements Scaling Properties", J. Song and W.C. Chew 420

"A Fast Technique for Determining Electromagnetic and Acoustic Wave Behavior in Inhomogeneous Media"
M.A. Jensen 428

"Rapid Analysis of Perfectly Conducting and Penetrable Quasi-Planar Structures with the Steepest Descent Fast Multipole Method"
V. Jandhyala, E. Michielssen, B. Shanker and W.C. Chew 434

"Iterative Solution Strategies in Adaptive Integral Method (AIM)"
E. Bleszynski, M. Bleszynski and T. Jaroszewicz 441

"A Fast Moment Method Matrix Solver"
F.X. Canning and K. Rogovin 449

"Vector Parabolic Equation Technique for the RCS Calculations"
A.A. Zaporozhets and M.F. Levy 455

SESSION 7: ELECTROMAGNETICS IN BIOLOGICAL AND MEDICAL APPLICATIONS
Chairs: Cynthia Furse and Maria A. Stuchly

"EM Interaction Evaluation of Handset Antennas and Human Head: A Hybrid Technique", K.W. Kim and Y. Rahmat-Samii 462

"Comparison of RGFM and FDTD for Electromagnetic-Tissue Interaction Problems" M.A. Jensen 470

"Isolated vs. *in situ* Human Heart Dosimetry under Low Frequency Magnetic Exposure" T.W. Dawson, K. Caputa and M.A. Stuchly 476

"Faster than Fourier - Ultra Efficient Time-to-Frequency Domain Conversions for FDTD Applied to Bioelectromagnetic Dosimetry", C.M. Furse 482

"Modelling of Antennas in Close Proximity to Biological Tissues Using the TLM Method" J. Paul, C. Christopoulos and D.W.P. Thomas 490

SESSION 8: ADVANCES IN PERFECTLY MATCHED LAYERS (PML)
Chairs: Weng Cho Chew and Qing Huo Liu

"Conformal Perfectly Matched Layers" F.L. Teixeira and W.C. Chew 500

"Stability Analysis of Cartesian, Cylindrical and Spherical Perfectly Matched Layers" F.L. Teixeira and W.C. Chew 507

"A Unified Approach to PML Absorbing Media" D.H. Werner and R. Mittra 515

"Comparison of the Performance of the PML and the Liao Absorbing Boundary Formulation", M. Vall-llossera and C.W. Trueman 523

"A Uniaxial PML Implementation for a Fourth Order Dispersion-Optimized FDTD Scheme", G. Haussmann and M. Piket-May 531

SESSION 9: VISUALIZATION IN CEM
Chairs: Janice Karty and Stanley J. Kubina

"Plate Scattering Visualization: Images, Near Fields, Currents, and Far Field Patterns" J. Shaeffer and K. Hom 538

"Visualization Aids for Effective Aircraft Antenna Simulations" S.J. Kubina, C.W. Trueman, Q. Luu and D. Gaudine 546

"Visualization of Radiation from a Spiral Antenna Using EM-ANIMATE" R.A. Pearlman, M.R. Axe, J.M. Bomholdt and J.M. Roedder 554

"Evolution of an Antenna Training Aid Using Electromagnetic Visualisation" A. Nott and D. Singh 560

"The NEC-BSC Workbench: A Companion Graphical Interface Tool" G.F. Paynter and R.J. Marhefka 568

SESSION 9: VISUALIZATION IN CEM (continued)

" A New Tool to Assist Use of Legacy Programs"
B. Joseph, A. Paboojian, S. Woolf and E. Cohen 576

"Visual EMag: A 2-D Electromagnetic Simulator for Undergraduates"
D. Garner, J. Lebaric and D. Voltmer 581

"Exploring Electromagnetic Physics Using Thin-Wire Time-Domain (TWTD) Modeling"
E.K. Miller 583

AUTHOR INDEX 589

SESSION 10: ABC'S FOR CEM: THEORETICAL AND IMPLEMENTATION
Chairs: Peter G. Petropoulos and Omar M. Ramahi

"The Concurrent Complementary Operators Method for FDTD Mesh Truncation"
O.M. Ramahi 592

"Accurate Boundary Treatments for Maxwell's Equations and their Computational Complexity", T. Hagstrom, B.K. Alpert, L.F. Greengard and S.I. Hariharan 600

"Perfectly Matched Layer Methods in Spherical Coordinates"
B. Yang and D. Gottlieb 607

"The Unsplit PML for Maxwell's Equations in Cylindrical and Spherical Coordinates"
P.G. Petropoulos 615

"A Comparison of the Grote-Keller and Unsplit PML Absorbing Boundary Conditions for Maxwell's Equations in Spherical Coordinates"
N.V. Kantartzis, P.G. Petropoulos and T.D. Tsiboukis 623

"A Systematic Study of Three PML Absorbing Boundary Conditions Through a Unified Formulation in Cylindrical Coordinates"
J-Q. He and Q-H. Liu 631

"Preconditioned Generalized Minimal Residual (GMRES) Solver for Domains Truncated by Perfectly Matched Layer (PML) Absorbers"
Y.Y. Botros and J.L. Volakis 639

"PML Implementation for the Battle-Lemarie Multiresolution Time-Domain Schemes"
E. Tentzeris, R. Robertson and L.P.B. Katehi 647

"A PML-FDTD Algorithm for General Dispersive Media"
G-X. Fan and Q.H. Liu 655

SESSION 11: CEM AND PARAMETER EXTRACTION FOR PACKAGING ANALYSIS
Chairs: Emilie van Deventer and Jose E. Schutt-Aine

"A New Generalized De-embedding Method for Numerical Electromagnetic Analysis"
Y.O. Shlepnev 664

"A Circuit Extraction Approach in PCB Power-Bus Analysis"
H. Shi and J.L. Drewniak 672

"Modeling of Conductor and Dielectric Losses in Packages"
J. Poltz 680

"Extraction of Effective Capacitance and Inductance of a Power Distribution Structure from Numerical Field Data"
A. Byers, B. Boots, M. Piket-May and R. Gravrok 687

"Extraction of Equivalent Circuit Parameters of Interconnections Using FDTD and PML"
F. Liu and J.E. Schutt-Aine 695

SESSION 12: REDUCED-ORDER MODELING IN ELECTROMAGNETICS
Chairs: Andreas C. Cangellaris and R.F. Remis

"The Use of a Correspondence Principle in Reduced-order Modeling of Electromagnetic Wavefields"
R.F. Remis and P.M. van den Berg 704

"The Spectral Lanczos Decomposition Method for Efficient Time-Domain and Frequency-Domain Finite-Element Solution of Maxwell's Equations"
M. Zunoubi, J-M. Jin, K. Donepudi and W.C. Chew 712

"Passivity of Discrete Electromagnetic Systems"
A.C. Cangellaris and L. Zhao 721

"Rational Krylov Reduced Order Modeling of Multiscreen Frequency Selective Surfaces" - STUDENT PAPER CONTEST -
D.S. Weile, E. Michielssen and K. Gallivan 732

SESSION 13: FINITE ELEMENT METHOD
Chairs: John R. Brauer and Jin-fa Lee

"Comparing High Order Vector Basis Functions"
J.S. Savage 742

"Mesh Refinement for Hybrid FEM in the Analysis of Printed Antennas and Arrays"
J. Gong, D. Bernstein and S. Wedge 750

"A Novel, Efficient Algorithm for Scattering from a Complex BOR Using Vector FEM and PML" - STUDENT PAPER CONTEST - A.D. Greenwood and J-M. Jin 756

"Homogenized Finite Element Model of a Beam Waveguide Resonator Antenna with Over One Hundred Coupling Holes"
J.R. Brauer 763

"A Surface Admittance Formulation for the Transient Modeling of Skin Effect and Eddy Current Problems"
K.N. Wassef and A.F. Peterson 771

SESSION 13: FINITE ELEMENT METHOD (continued)

"Verification of Eddy Current Analysis of Engineering Oriented Loss Model (Problem 21)"
N. Takahashi, K. Fujiwara, K. Sugiyama and J. Takehara 779

"Jacobi-Davidson Algorithm for Modeling Open Domain Lossy Cavities"
C. Liu and J-F. Lee 787

"Analysis of Electromagnetic Penetration Through Apertures of Shielded Enclosure Using Finite Element Method",
B-W. Kim, Y-C. Chung and T-W. Kang 795

"hp-Adaptive Edge Finite Elements for Maxwell's Equations"
L. Demkowicz, L. Vardapetyan and W. Rachowicz 799

SESSION 14: RECENT ADVANCES IN TIME-DOMAIN TECHNIQUES
Chairs: Douglas C. Blake and Douglas J. Riley

"An Analysis of Programming Models for Time-Domain CEM Codes on RISC-Based Computers" D.C. Blake and J.S. Shang 808

"The *VOLMAX* Transient Electromagnetic Modeling System, Including Sub-Cell Slots and Wires on Random Non-Orthogonal Cells"
D.J. Riley and C.D. Turner 816

"Using the Finite Integration Time Domain Technique at Low Frequencies"
R. Ehmann and T. Weiland 825

"Modelling Dispersive Media Using the Finite Integration Technique"
S. Gutschling, H. Kruger and T. Weiland 832

"Transient Analysis of Thin Wire Antennas Mounted on Three-Dimensional Perfectly Conducting Bodies"
K. Aygun, A.A. Ergin, B. Shanker, S.E. Fisher and E. Michielssen 838

"A PSTD Algorithm in Cylindrical Coordinates"
Q.H. Liu and J.Q. He 845

"On the PSTD Method for Large-Scale Problems", Q.H. Liu 852

"Pseudospectral Time-Domain Modeling of Diffractive Optical Elements"
J.S. Hesthaven, P.G. Dinesen and J.P. Lynov 858

"Transient Analysis of Acoustic Scattering Using Marching-on-in Time With Plane Wave Time Domain Algorithm"
A.A. Ergin, B. Shanker, K. Aygun and E. Michielssen 866

"A Plane Wave Time Domain Algorithm for the Fast Analysis of Transient Electromagnetic Scattering Phenomena"
B. Shanker, A.A. Ergin, K. Aygun and E. Michielssen 873

SESSION 15: EMI/EMC
Chairs: Todd Hubing and Bruce R. Archambeault

"EMC Modeling of Shielded Enclosures with Apertures and Attached Wires in a Real-World Environment"
B.R. Archambeault, K. Chamberlin and O. Ramahi 880

"Proposed Standard EMI Modeling Problems for Evaluating Tools which Predict Shielding Effectiveness of Metal Enclosures"
B.R. Archambeault 888

"A Study in the Proper Design of Grounding for SMPS Converters and the Role of CEM"
R. Perez 895

"Expert System Algorithms for EMC Analysis"
T. Hubing, N. Kashyap, J. Drewniak, T. Van Doren and R. DuBroff 905

"The Electromagnetic Compatibility Characteristics of Buildings in Mobile Radio Waves Propagation Channel"
Y. Miyazaki and P. Selormey 911

SESSION 16: HYBRID TECHNIQUES
Chairs: Ulrich Jakobus and William D. Wood, Jr.

"Extension of the MoM/PO Hybrid Technique to Homogeneous Dielectric Bodies"
U. Jakobus 920

"EMAPS: A 3D Hybrid FEM/MoM Code"
Y. Ji and T.H. Hubing 928

"Iterative Coupling of MoM and MMP for the Analysis of Metallic Structures Radiating in the Presence of Dielectric Bodies"
H-O. Ruoss, U. Jakobus and F.M. Landstorfer 936

"A Hybrid Algorithm for Frequency Selective Surface Analysis"
M.J. Walker 944

"Generalized Networks for Waveguide Step Discontinuities"
M. Mongiardo, P. Russer, M. Dionigi and L.B. Felsen 957

SESSION 17: SIGNAL PROCESSING TECHNIQUES IN CEM
Chairs: Douglas H. Werner and Ping L. Werner

"Using Windowed, Adaptive Sampling to Minimize the Number of Field Values Needed to Estimate Radiation and Scattering Patterns"
E.K. Miller 958

"Spectral Domain Interpolation of Antenna Radiation Patterns Using Model-Based Parameter Estimation and Genetic Algorithms"
R.J. Allard, D.H. Werner, J.S. Zmyslo and P.L Werner 964

"An Accurate Algorithm for Nonuniform Fast Fourier Transforms (NUFFT) and Its Applications", Q.H. Liu and N. Nguyen 972

SESSION 17: SIGNAL PROCESSING TECHNIQUES IN CEM (continued)

"A Subspace Approach to Fast Moment Method Scattering Predictions Over Limited Sectors", J. Stach 979

"Application of Biorthogonal B-Spline-Wavelets to Telegrapher's Equations" M. Aidam and P. Russer 983

"NEC Acceleration by the Wavelet Matrix Transform" Y.H. Lee and Y. Lu 991

"Adaptive Segmentation Algorithms for Optimal NEC Modelling of Wire-Grid Structures", Y.H. Lee and Y. Lu 999

SESSION 18: HF-UHF PRACTICAL ANTENNA TOPICS
Chairs: W. Perry Wheless, Jr., and Nathan Cohen

"Comparison of Shipboard HF Transmit Fan Characteristics NEC versus Scale-Model Measurements", K. Lysiak and P. Dombowsky 1006

"NEC Model Results for Shipboard Shielded Crossed Loop Antennas with Scale-Model Range Data", K. Lysiak 1012

"A Near-Earth and Buried HF Antenna Computer Modeling Program" W.P. Wheless, Jr., and LT. Wurtz 1019

"Empirical and Numerical Treatment of Electromagnetic Pulse Induced Currents" M.J. Packer 1028

"Advantages of an Alternate Viewpoint When Designing HF Verticals for 80 and 160 m", R. Severns 1036

"Tower Equivalent Radius", W.F. Cummins 1043

"Simple CP Fractal Loop Array with A Parasitic", N. Cohen 1047

"NEC4 Analysis of a Fractalized Monofilar Helix in an Axial Mode" N. Cohen 1051

"Design of Low Sidelobe Antennas", R.W. Hecht 1058

"EMP Simulations of Near Field Enhancement of Wire Antenna" M. El Hachemi, C. Tosser-Roussey and A. Tosser-Roosey 1066

SESSION 19: OPTIMIZATION TECHNIQUES FOR ELECTROMAGNETICS
Chairs: Eric Michielssen and Randy L. Haupt

"Genetic Algorithm Design of the Conical Interdigitated Log-Periodic Antenna"
P.D. Mannikko, P.J. O'Brien and K.W. Ommodt 1074

"A Comparison of Simple and Complex Genetic Algorithms in Wire Antenna Design"
B.S. Sandlin and A.J. Terzuoli 1080

"Array Failure Correction with a Genetic Algorithm"
B. Beng, K. Yeo and Y. Lu 1087

"Backscattering Synthesis from Tapered Resistive Grids", R.L. Haupt 1095

"Obtaining Linear and Circular Apertures with Smooth Amplitude Distribution and High Efficiency", J.A. Rodriguez and F. Ares 1100

SESSION 20: INTEGRAL EQUATION METHODS AND ERROR CONTROL
Chairs: Goran Eriksson and C.Y. Shen

"Force Calculations and Error Estimates with Boundary Element Methods"
K. Davey and D. Zheng 1106

"Use of Residual Error Bounds to Obtain Stable Numerical Solutions of a Fredholm Integral Equation of the First Kind"
T. Schwengler and E.F. Kuester 1112

"Eigenvalue Studies of Matrices Resulting from EFIE Simulations for Planar Structures"
J.M. Dunn and H. MacMillan 1120

"Iterative Solution of Dense Linear Systems in Electromagnetic Scattering Calculations", J. Rahola 1126

"EMCP2 A Parallel Boundary Element Software Package Using a Novel Parameterisation Technique" G. Eriksson and U. Thibblin 1134

AUTHOR INDEX 1141

FINAL AGENDA

The Fourteenth Annual Review of Progress in Applied Computational Electromagnetics

NAVAL POSTGRADUATE SCHOOL
16-20 MARCH 1998

Jianming Jin, Technical Program Chairman

Michael A. Jensen, Conference Co-Chair & Short Course Chairman

Randy L. Haupt, Conference Co-Chair

Keith Whites, Vendor Chairman

Richard W. Adler, Conference Facilitator

MONDAY MORNING 16 MARCH 1998

0730 - 0820	**SHORT COURSE REGISTRATION**	**Glasgow 103**
0830 - 1630	"Recent Advances in Fast Algorithms for Computational Electromagnetics", Weng Cho Chew, Jiming Song, Eric Michielssen, and Jianming Jin, Center for Computational Electromagnetics, University of Illinois at Urbana-Champaign.	**Full-day course**
0830 - 1630	"An Introduction to the Fast-MoM in Computational Electromagnetics", Ali R. Baghai-Wadji, Department of Applied Electronics, Vienna University of Technology, Vienna, Austria.	**Full-day course.**
0830 - 1630	"Finite Elements for Electromagnetics", John R. Brauer, Ansoft Corporation, Milwaukee, WI.	**Full-day course**
0830 - 1200	"A Survey and Comparison of Computational Electromagnetics Options", Edmund K. Miller, Santa Fe, NM.	**Half-day course**
1300 - 1630	"Using Model-Based Parameter Estimation to Increase the Efficiency and Effectiveness of Computational Electromagnetics", Edmund K. Miller, Santa Fe, NM.	**Half-day course**
0900 - 1800	**CONFERENCE REGISTRATION**	**Glasgow 103**
1200	**BOD MEETING/LUNCHEON**	**Terrace Room, Herrmann Hall**
1630 - 1830	**BOD CONTINUATION MEETING**	**Glasgow 114**

TUESDAY MORNING 17 MARCH 1998

0715 - 0750	**CONTINENTAL BREAKFAST**		**Glasgow Courtyard**
0745 - 0800	**ACES BUSINESS MEETING**	President Hal Sabbagh	**Glasgow 102**
0800-1810	**WELCOME**	Jianming Jin	**Glasgow 102**

SESSION 1: **CEM ANALYSIS AND APPLICATION WITHIN AN ENGINEERING ENVIRONMENT** **(Parallel with Sessions 2 & 3)**
Chair: Kenneth R. Siarkiewicz (Organizer), Co-Chair: Donald R. Pflug (Co-Organizer)

0820	Frameworks--Future Paradigm for CEM Analysis	K.R. Siarkiewicz
0840	REF CEM Data Dictionary	T. Wharton & J.A. Evans
0900	Air Force Research & Engineering Framework Control Panel	D.M. Hallatt & J. A. Evans
0920	Continued Application of the Research and Engineering Framework (REF) at Raytheon	J. LaBelle, B. Hantman H. Wright, Y. Chang, R. Abrams
0940	The Innovative Research Testbed: A PC-Based High Performance Computing and Web-Based Collaboratory for Computational Electromagnetics	D.M. Leskiw, G. Ingersoll T. Vidoni & R. Redmond
1000	**BREAK**	

TUESDAY MORNING 17 MARCH 1998

SESSION 1: CEM ANALYSIS AND APPLICATION WITHIN AN ENGINEERING ENVIRONMENT (cont)

1020	Illustrating the Application of Expert Systems to Computational Electromagnetics Modeling and Simulation	A.L.S. Drozd, T.W. Blocher C.E. Carroll, Jr., & J.M. Allen
1040	Some Present and Future Aspects of the Quality of Solution in Computational Electromagnetics	D.R. Pflug
1100	On the Use of Computational Electromagnetics in the Radar Cross Section Measurement Calibration Process	K.C. Hill & W.D. Wood, Jr.
1120	Radar Calibration at Low Frequencies Using a Triangular Trihedral Corner Reflector	C.Y. Shen
1200	**LUNCH**	

SESSION 2: FINITE-DIFFERENCE TIME-DOMAIN APPLICATIONS (Parallel with Sessions 1 & 3)
Chair: John Beggs (Organizer), Co-Chair: Melinda Piket-May

0820	Time Domain Analysis of Small Multi-Sector Monopole Yagi-Uda Array Antenna Mounted on a Fi inite Ground Plane Using FDTD Method	T. Maruyama, K. Uehara, T. Hori, K.. Kagoshima
0840	Antenna Performance Calculation in Lossy Media with FDTD Method	M. Cai & N. Ljepojevic
0900	Far-Zone Transformation in FDTD for VHF-band SAR-image Simulations	T. Martin & L. Ulander
0920	FDTD Simulations Used to Correct for Ground Effects in Aircraft Testing	G. Eriksson & U. Thibblin
0940	Validation of FDTD-Computed Handset Patterns by Measurement	C.W. Trueman, S.J. Kubina, J.E. Roy, W.R. Lauber, M. Vall-llossera
1000	**BREAK**	
1020	FDTD Analysis of Flip Chip Interconnects	A Z. Elsherbeni, V. Rodriguez-Pereyra, C.E. Smith
1040	FDTD Analysis of the Celestron-8 Telescope	R.R. DeLyser
1100	Modifying a Graphically-Based FDTD Simulation for Parallel Processing	G. Haussmann, M. Piket-May K. Thomas
1120	A Time Domain Method for High Frequency Problems Exploiting the Whitney Complex	A. Arkko, T. Tarhasaari L. Kettunen
1140	Treatment of Boundaries in Multiresolution Based FDTD Multigrid	K. Goverdhanam, E. Tentzeris L.P.B. Katehi
1200	**LUNCH**	

SESSION 3: APPLICATION OF ANALYTIC AND COMPUTATIONAL METHODS TO THE MODELING OF ELECTROMAGNETIC MATERIALS (Parallel with Sessions 1 & 2)
Chair: Rudy Diaz (Organizer), Co-Chair: David H. Y. Yang

0820	Computing Dispersion Relations and Radiation Spectra in Photonic Band Gap Materials by Plane Wave Expansion Method	T. Suzuki & P.K. L. Yu
0840	Waveguides in Photonic Band Gap Materials	M.M. Sigalas, R. Biswas K.M. Ho, C.M. Soukoulis D.D. Crouch
0900	Is There a Relationship Between a Random and an Ordered Composite Mixture? **- STUDENT PAPER CONTEST -**	W.M. Merrill, S.A. Kyriazidou N.G. Alexopoulos
0920	An Analytic Framework for the Modeling of Effective Media	R.E. Diaz
0940	Finite Grain Boundary Effects for Magnetic Materials: Tunneling and Intergrain Gaps	G.G. Bush
1000	**BREAK**	

TUESDAY MORNING 17 MARCH 1998

SESSION 3: APPLICATION OF ANALYTIC AND COMPUTATIONAL METHODS TO THE MODELING OF ELECTROMAGNETIC MATERIALS (cont)

1020	Numerical Computation of the Complex Dielectric Permittivity: FFT-Based Hilbert Transform Approximation of the Kramers-Kronig Relations - **STUDENT PAPER CONTEST -**	F. Castro & B. Nabet
1040	Rayleigh Analysis of Novel Dense Medium Exhibiting Narrow-Band Transparency Window	S.A. Kryiazidou, R.E. Diaz N.G. Alexopoulos
1200	**LUNCH**	

TUESDAY AFTERNOON 17 MARCH 1998

INTERACTIVE POSTER SESSION 1300 - 1530	**Ballroom, Herrmann Hall**
VENDOR EXHIBITS 1300 - 1900	**Ballroom, Herrmann Hall**
WINE AND CHEESE BUFFET 1500 - 1700	**Ballroom, Herrmann Hall**

SESSION 4: INTERACTIVE POSTER SESSION — **Ballroom, Herrmann Hall**

Finite Difference Formulation - Expanding the Capabilities	K. Davey
Time and Frequency Domain Numerical Modeling of Outbound and Standing Power from Perpendicularly Oriented, Electrically Small TM Dipoles	G. Liu, C.A. Grimes K.G. Ong
Efficient Analysis of Large Two-dimensional Arbitrarily Shaped Finite Gratings for Quantum Well Infrared Photo-detectors	V. Jandhyala, D. Sengupta B. Shanker, E. Michielssen M. Feng, & G. Stillman
Forward and Backward Propagation Algorithms Applied to the Electromagnetic Scattering by an Impenetrable Obstacle: A Progress Report	G.F. Crosta
Analysis of Broad Wall Slots Excited by Tuning Screws	T. Azar & R.Coren
Closed-Form Expression of the Arc-Length of the Toroidal/Helical Equilibrium Orbit	R.A. Speciale
Asymptotic Techiques in Naval Ship Design	R. Routier & R. Burkholder
A BCG-FFT Solution of Scattering and Radiation by Large Finite Arrays of Microstrip Antennas	C-Fu Wang, F. Ling, J-M Jin
Analytic Solution for Low-Frequency Electric Induction in an Equatorially Stratified Sphere	T.W. Dawson
Parallelizing Computational Electromagnetics Code Using the Parallel Virtual Machine: Examples	E. Skochinski & S. Regarajan
SAF Analysis Codes for Predicting the Electromagnetic Effectiveness of Antennas Enclosed in Composite Structures	B.J. Cown, J.P. Estrada R. Routier
Some Concentrated Solutions on Hemholtz Equation with Nonlocal Nonlinearity	Y.N. Cherkashin V.A. Eremenko
Hysteresis and Eddy Currents in Ferromagnetic Media	J. Fuzi & A. Ivanyi
Electrical Circuit Analysis Considering Hysteresis in Coil Cores	J. Fuzi
A Paradigm for Proving the Convexity Properties of Slowness Curves	A.R. Baghai-Wadji
On the Strict-Convexity of the Slowness Surfaces for the Fastest Bulk-Acoustic-Waves in Piezoelectric and Piezoelectromagnetic Media	A.R. Baghai-Wadji
Solving EMC Problems Using the FDTD Method	F. Gisin & Z. Pantic-Tanner

TUESDAY EVENING 17 MARCH 1998 **BOD DINNER**

WEDNESDAY MORNING 18 MARCH 1998

0730 - 0800	**CONTINENTAL BREAKFAST**	**Glasgow Courtyard**
SESSION 5:	**TLM MODELING AND APPLICATIONS (Parallel with Sessions 6, 7 & 8))** Chair: Wolfgang Hoefer (Organizer), Co-Chair: Peter Russer	
0820	A Hybrid Time Domain TLM-Integral Equation Method for Solution of Radiation Problems	L. Pierantoni, S. Lindenmeier P. Russer
0840	Comparison of Symmetric Condense TLM, Yee FDTD and Integer Lattice Gas Automata Solutions for a Problem Containing a Sharp Metallic Edge	N. Simons, R. Siushansian J. LoVetri, G. Bridges M. Cuhaci
0900	Some Observations on Stubs, Boundaries and Parity Effects in TLM Models	D. de Cogan & C. Kun
0920	Modelling of Dispersive Media in TLM Using the Propagator Approach	J. Rebel & P. Russer
0940	Characterization of Quasiplanar Structures Using the TLM Method	O. Pertz, U. Muller & A. Beyer
1000	**BREAK**	
1020	Generation of Lumped Element Equivalent Circuits from Time-Domain Scattering Signals	T. Mangold & P. Russer
1040	TLM Analysis of an Optical Sensor	R.R. DeLyser
1100	TLM Modeling and TDR Validation of Soil Moisture Probe for Environmental Sensing	G. Tardioli, M. Righi, L. Cascio W.J.R. Hoefer, & R. McFarlane
1120	TLM Analysis of the Celestron-8 Telescope	R.R. DeLyser
1140	Near to Far Field Transformation via Parabolic Equation - **STUDENT PAPER CONTEST**	A.N. Kurokhtin, Y.V. Kopylov, A.V. Popov, & A.V. Vinogradov
1200	**LUNCH**	
SESSION 6:	**FREQUENCY-DOMAIN FAST ALGORITHMS (Parallel with Sessions 5, 7 & 8)** Chair: Jiming Song (Organizer), Co-Chair: Weng Cho Chew (Co-Organizer)	
0820	Recent Advances in the Numerical Solution of Integral Equations Applied to EM Scattering from Terrain	P. Cullen & C. Brennan
0840	Solution of Combined-Field Integral Equation Using Multi-Level Fast Multipole Algorithm for Scattering by Homogeneous Bodies	X.Q. Sheng, J.M. Jin, J.M. Song W.C. Chew, & C.C. Lu
0900	Comparisons of FMM and AIM Compression Schemes in Finite Element - Boundary Integral Implementations for Antenna Modeling	K. Sertel, D.S. Filipovic S. Bindiganavale, & J.L. Volakis
0920	High-Order Nystrom Discretization for Faster, More Accurate Scattering Calculations	L.S. Canino, J.J. Ottusch M.A. Stalzer, J.L. Visher S.M. Wandzura
0940	Large Scale Computing with the Fast Illinois Solver Code --Requirements Scaling Properties	J. Song & W.C. Chew
1000	**BREAK**	
1020	A Fast Technique for Determining Electromagnetic and Acoustic Wave Behavior in Inhomogeneous Media	M.A. Jensen
1040	Rapid Analysis of Perfectly Conducting and Penetrable Quasi-Planar Structures with the Steepest Descent Fast Multipole Method	V. Jandhyala, E. Michielssen B. Shanker, & W.C. Chew
1100	Iterative Solution Strategies in Adaptive Integral Method (AIM)	E. Bleszynski, M. Bleszynski T. Jaroszewicz
1120	A Fast Moment Method Matrix Solver	F.X. Canning & K. Rogovin
1140	Vector Parabolic Equation Technique for the RCS Calculations	A.A. Zaporozhets & M. F. Levy
1200	**LUNCH**	

WEDNESDAY MORNING 18 MARCH 1998

SESSION 7: **ELECTROMAGNETICS IN BIOLOGICAL AND MEDICAL APPLICATIONS (Parallel with Sessions 5, 6 & 8)**
Chair: Cynthia Furse (Organizer), Co-Chair: Maria A. Stuchly

0820	EM Interaction Evaluation of Handset Antennas and Human Head: A Hybrid Technique	K.W. Kim & Y. Rahmat-Samii
0840	Comparison of RGFM and FDTD for Electromagnetic-Tissue Interaction Problems	M.A. Jensen
0900	Isolated vs. *in situ* Human Heart Dosimetry under Low Frequency Magnetic Exposure	T.W. Dawson, K. Caputa M.A. Stuchly
0920	Faster Than Fourier – Ultra-Efficient Time-to-Frequency Domain Conversions for FDTD Applied to Bioelectromagnetic Dosimetry	C.M. Furse
0940	Modelling of Antennas in Close Proximity to Biological Tissues Using the TLM Method	J. Paul, C. Christopoulos D.W.P. Thomas
1000	**BREAK**	

SESSION 8: **ADVANCES IN PERFECTLY MATCHED LAYERS (PML) (Parallel with Sessions 5, 6, & 7)**
Chair: Weng Cho Chew, Co-Chair: Qing Huo Liu

1020	Conformal Perfectly Matched Layer	F.L. Teixeira & W.C. Chew
1040	Stability Analysis of Cartesian, Cylindrical and Spherical Perfectly Matched Layers	F.L. Teixeira & W.C. Chew
1100	A Unified Approach to PML Absorbing Media	D.H. Werner & R. Mittra
1120	Comparison of the Performance of the PML and the Liao Absorbing Boundary Formulation	M. Vall-llossera, C.W. Trueman
1140	A Uniaxial PML Implementation for a Fourth Order Dispersion-Optimized FDTD Scheme	G. Haussmann & M. Piket-May
1200	**LUNCH**	

WEDNESDAY AFTERNOON 18 MARCH 1998

SESSION 9: **VISUALIZATION IN CEM (Parallel with Sessions 10, 11 & 12))**
Chair: Janice Karty (Organizer), Co-Chair: Stanley J. Kubina

1320	Plate Scattering Visualization: Images, Near Fields, Currents, and Far Field Patterns	J. Shaeffer & K. Hom
1340	Visualization Aids for Effective Aircraft Antenna Simulations	S.J. Kubina, C.W. Trueman Q. Luu, D. Gaudine
1400	Visualization of Radiation from a Spiral Antenna Using EM-ANIMATE	R.A. Pearlman, M.R. Axe J.M. Bomholdt, & J.M. Roedder
1420	Evolution of an Antenna Training Aid Using Electromagnetic Visualisation	A. Nott & D. Singh
1440	The NEC-BSC Workbench: A Companion Graphical Interface Tool	G.F. Paynter and R.J. Marhefka
1500	**BREAK**	
1520	A New Tool to Assist Use of Legacy Programs	B. Joseph, A. Paboojian, S. Woolf, E. Cohen
1540	Visual EMag: A 2-D Electromagnetic Simulator for Undergraduates	D. Garner, J. Lebaric D. Voltmer
1600	Exploring Electromagnetic Physics Using Thin-Wire Time-Domain (TWTD) Modeling	E. K. Miller

SESSION 10: **ABCS FOR CEM: THEORETICAL AND IMPLEMENTATION (Parallel with Sessions 9, 11 & 12)**
Chair: Peter G. Petropoulos (Organizer), Co-Chair: Omar M. Ramahi

1320	The Concurrent Complementary Operators Method for FDTD Mesh Truncation	O.M. Ramahi
1340	Accurate Boundary Treatments for Maxwell's Equations and Their Computational Complexity	T. Hagstrom, B.K. Alpert L.F. Greengard, S.I. Hariharan
1400	Perfectly Matched Layer Methods in Spherical Coordinates	B. Yang & D. Gottlieb

WEDNESDAY AFTERNOON 18 MARCH 1998

SESSION 10: ABCS FOR CEM: THEORETICAL AND IMPLEMENTATION (cont)

1420	The PML for Maxwell's Equations in Cylindrical and Spherical Coordinates	P.G. Petropoulos
1440	A Comparison of the Grote-Keller and Unsplit PML Absorbing Boundary Conditions for Maxwell's Equations in Spherical Coordinates	N.V. Kantartzis, P.G. Petropoulos, T.D. Tsiboukis
1500	**BREAK**	
1520	A Systematic Study of Three PML Absorbing Bo iundary Conditions Through a Unified Formulation in Cylindrical Coordinates	J.-Q. He & Q.-H. Liu
1540	Preconditioned Generalized Minimal Residual (GMRES) Solver for Domains Truncated by Perfectly Matched Layer (PML) Absorbers	Y.Y. Botros & J.L. Volakis
1600	PML Implementation for the Battle-Lemarie Multiresolution Time-Domain Schemes	E. Tentzeris, R. Robertson L.P.B. Katehi
1620	A PML-FDTD Algorithm for General Dispersive Media	G-X. Fan & Q.H. Liu

SESSION 11: CEM AND PARAMETER EXTRACTION FOR PACKAGING ANALYSIS (Parallel with Sessions 9, 10 & 12)
Chair: Emilie van Deventer (Organizer), Co-Chair: Jose E. Schutt-Aine

1320	A New Generalized De-embedding Method for Numerical Electromagnetic Analysis	Y.O. Shlepnev
1340	A Circuit Extraction Approach in PCB Power-Bus Analysis	H. Shi & J.L. Drewniak
1400	Modeling of Conductor and Dielectric Losses in Packages	J. Poltz
1420	Extraction of Effective Capacitance and Inductance of a Power Distribution Structure from Numerical Field Data	A. Byers, B. Boots M. Piket-May, & R. Gravrok
1440	Extraction of Equivalent Circuit Parameters of Interconnections Using FDTD and PML	F. Liu & J.E. Schutt-Aine
1500	**BREAK**	

SESSION 12: REDUCED-ORDER MODELING IN ELECTROMAGNETICS (Parallel with Sessions 9, 10 & 11)
Chair: Andreas C. Cangellaris, Co-Chair: R.F. Remis

1520	The Use of a Correspondence Principle in Reduced-Order Modeling of Electromagnetic Wave Fields	R.F. Remis P.M. van den Berg
1540	The Spectral Lanczos Decomposition Method for Efficient Time-Domain and Frequency-Domain Finite-Element Solution of Maxwell's Equations	M. Zunoubi, J-M. Jin K. Donepudi, & W.C. Chew
1600	Passivity of Discrete Electromagnetic Systems	M. Zunoubi, J.-M. Jin A.C. Cangellaris & L. Zhao
1620	Rational Krylov Reduced Order Modeling of Multiscreen Frequency Selective Surfaces - **- STUDENT PAPER CONTEST -**	D.S. Weile, E. Michielssen K. Gallivan

WEDNESDAY EVENING 18 MARCH 1998

1830	**NO HOST BAR**	**Terrace Room, Herrmann Hall**
1930	**AWARDS BANQUET**	**La Novia Room, Herrmann Hall**

THURSDAY MORNING 19 MARCH 1998

0730 - 0800	**CONTINENTAL BREAKFAST**	**Glasgow Courtyard**
SESSION 13:	**FINITE ELEMENT METHOD (Parallel with Sessions 14, 15 & 16)** Chair: John R. Brauer, Co-Chair: Jin-Fa Lee	
0820	Comparing High Order Vector Basis Functions	J.S. Savage
0840	Mesh Refinement for Hybrid FEM in the Analysis of Printed Antennas and Arrays	J. Gong, D. Bernstein S. Wedge
0900	A Novel, Efficient Algorithm for Scattering from a Complex BOR Using Vector FEM and PML -- **STUDENT PAPER CONTEST--**	A.D. Greenwood & J-M. Jin
0920	Homogenized Finite Element Model of a Beam Waveguide Resonator Antenna with Over One Hundred Coupling Holes	J.R. Brauer
0940	A Surface Admittance Formulation for the Transient Modeling of Skin Effect and Eddy Current Problems - **STUDENT PAPER CONTEST -**	K.N. Wassef & A.F. Peterson
1000	**BREAK**	
1020	Verification of Eddy Current Analysis of Engineering Oriented Loss Model (Problem 21) Sugiyama	N. Takahashi, K. Fujiwara, K. J. Takehara
1040	Jacobi-Davidson Algorithm for Modeling Open Domain Lossy Cavities	C. Liu & J-F Lee
1100	Analysis of Electromagnetic Penetration Through Apertures of Shielded Enclosure Using Finite Element Method	B-W. Kim, .-C. Chung T-W. Kang
1120	hp-Adaptive Edge Finite Elements for Maxwell's Equations	L. Demkowicz, L. Vardapetyan W. Rachowicz
1200	**LUNCH**	
SESSION 14:	**RECENT ADVANCES IN TIME-DOMAIN TECHNIQUES (Parallel with Sessions 13, 15 & 16)** Chair: Douglas C. Blake, Co-Chair: Douglas J. Riley	
0820	An Analysis of Programming Models for Time-Domain CEM Codes on RISC-Based Computers	D.C. Blake & J.S. Shang
0840	The VOLMAX Transient Electromagnetic Modeling System, Including Sub-Cell Slots and Wires on Random Non-Orthogonal Cells	D.J. Riley & C.D. Turner
0900	Using the Finite Integration Time Domain Technique at Low Frequencies	R. Ehmann & T. Weiland
0920	Modelling Dispersive Media Using the Finite Integration Technique	S. Gutschling, H. Kruger, T. Weiland
0940	Transient Analysis of Thin Wire Antennas Mounted on Three-Dimensional Perfectly Conducting Bodies	K. Aygun, A.A. Ergin, B. Shanker S.E. Fisher E. Michielssen
1000	**BREAK**	
1020	A PSTD Algorithm in Cylindrical Coordinates	Q.H. Liu & J. Q. He
1040	On the PSTD Method for Large-Scale Problems	Q.H. Liu
1100	Pseudospectral Time-Domain Modeling of Diffractive Optical Elements	J.S. Hesthaven, P.G. Dinesen J.P. Lynov
1120	Transient Analysis of Acoustic Scattering Using Marching-on-in-Time with Plane Wave Time Domain Algorithm	A.A Ergin, B. Shanker K. Aygun, & E. Michielssen
1140	A Two Level Plane Wave Time Domain Algorithm for Fast Analysis of Transient Electromagnetic Scattering	B. Shanker, A.A. Ergin K. Ayugun, & E. Michielssen
1200	**LUNCH**	

THURSDAY MORNING 19 MARCH 1998

SESSION 15: EMI/EMC (Parallel with Sessions 13,14 & 16)
Chair: Todd Hubing (Organizer), Co-Chair: Bruce Archambeault

0820	EMC Modeling of Shielded Enclosures with Apertures and Attached Wires in a Real-World Environment	B. Archambeault, K. Chamberlin O. Ramahi
0840	Proposed Standard EMI Modeling Problems for Evaluating Tools which Predict Shielding Effectiveness of Metal Enclosures	B. Archambeault & O. Ramahi
0900	A Study in the Proper Design of Grounding for SMPS Converters and the Role of CEM	R. Perez
0920	Expert System Algorithms for EMC Analysis	T. Hubing, N. Kashyap J. Drewniak, T. Van Doren R. DuBroff
0940	The Electromagnetic Comatibility Characteristics of Buildings in Mobile Radio Waves Propagation Channel	Y. Miyazaki & P. Selormey
1000	**BREAK**	

SESSION 16: HYBRID TECHNIQUES (Parallel with Sessions 13, 14 & 15)
Chair: Ulrich Jakobus, Co-Chair: William D. Wood, Jr.

1020	Extension of the MoM/PO Hybrid Technique to Homogeneous Dielectric Bodies	U. Jakobus
1040	EMAP5: A 3D Hybrid FEM/MoM Code	Y. Ji & T. Hubing
1100	Iterative Coupling of MoM and MMP for the Analysis of Metallic Structures Radiating in the Presence of Dielectric Bodies	H-O. Ruoss, U. Jakobus F.M. Landstorfer
1120	A Hybrid Algorithm for Frequency Selective Surface Analysis	M.J. Walker
1140	Generalized Networks for Waveguide Step Discontinuities	M. Mongiardo, P. Russer M. Dionigi & L.B. Felsen
1200	**LUNCH**	

THURSDAY AFTERNOON 19 MARCH 1998

SESSION 17: SIGNAL PROCESSING TECHNIQUES IN CEM (Parallel with Sessions 18, 19 & 20)
Chair: Douglas H. Werner (Organizer), Co-Chair: Ping L. Werner (Co-Organizer)

1320	Using Windowed, Adaptive Sampling to Minimize the Number of Field Values Needed to Estimate Radiation and Scattering Patterns	E.K. Miller
1340	Spectral Domain Interpolation of Antenna Radiation Patterns Using Model-Based Parameter Estimation and Genetic Algorithms	R.J. Allard, D.H. Werner J.S. Zmyslo, & P.L. Werner
1400	An Accurate Algorithm for Nonuniform Fast Fourier Transforms (NUFFT) and Its Applications	Q.H. Liu & N. Nguyen
1420	A Subspace Approach to Fast Moment Method Scattering Predictions over Limited Sectors	J. Stach
1440	Application of Biorthogonal B-Spline-Wavelets to Telegrapher's Equations	M. Aidam & P. Russer
1500	**BREAK**	
1520	NEC Acceleration by the Wavelet Matrix Transform	Y.H. Lee & Y. Lu
1540	Adaptive Segmentation Algorithms for Optimal NEC Modelling of Wire-Grid Structures	Y.H. Lee & Y. Lu

SESSION 18: HF-UHF PRACTICAL ANTENNA TOPICS (Parallel with Sessions 17, 19 & 20)
Chair: W. Perry Wheless, Jr. (Organizer), Co-Chair: Nathan Cohen

1320	Comparison of Shipboard HF Transmit Fan Characteristics: NEC versus Scale-Model Measurements	K. Lysiak & P. Dombowsky
1340	NEC Model Results for Shipboard Shielded Crossed Loop Antennas with Scale-Model Range Data	K. Lysiak

THURSDAY AFTERNOON 19 MARCH 1998

SESSION 18: HF-UHF PRACTICAL ANTENNA TOPICS (cont)

1400	A Near-Earth and Buried HF Antenna Computer Modeling Program	W.P Wheless, Jr. & L.T. Wurtz
1420	Empirical and Numerical Treatment of Electromagnetic Pulse Induced Currents	M.J. Packer
1440	Advantages of an Alternate Viewpoint when Designing HF Verticals for 80 and 160 m	R. Severns
1500	**BREAK**	
1520	Tower Equivalent Radius	W.F. Cummins
1540	Simple CP Fractal Loop Array with Parasitic	N. Cohen
1600	NEC4 Analysis of a Fractalized Monofilar Helix in an Axial Mode	N. Cohen
1620	Design of Low Sidelobe Antennas	R.W. Hecht
1640	EMP Simulation of Near Field Enhancement of Wire Antenna	M. El Hachemi, C. Tosser-Roussey A. Tosser-Roosey

SESSION 19: OPTIMIZATION TECHNIQUES FOR ELECTROMAGNETICS (Parallel with Sessions 17, 18 & 20)
Chair: Eric Michielssen (Organizer), Co-Chair: Randy L. Haupt (Co-Organizer)

1320	Genetic Algorithm Design of the Conical Interdigitated Log-Periodic Antenna	P.D. Mannikko, P.J. O'Brien K.W. Ommodt
1340	A Comparison of Simple and Complex Genetic Algorithms in Wire Antenna Design	B.S. Sandlin & A.J. Terzuoli
1400	Array Failure Correction with a Genetic Algorithm	B. Beng, K. Yeo & Y. Lu
1420	Backscattering Synthesis From Tapered Resistive Grids	R.L. Haupt
1440	Obtaining Linear and Circular Apertures with Smooth Amplitude Distributions and High Efficiency	J.A. Rodriguez & F. Ares
1500	**BREAK**	

SESSION 20: INTEGRAL EQUATION METHODS AND ERROR CONTROL (Parallel with Sessions 17, 18, & 19)
Chair: Goran Eriksson, Co-Chair: C. Y. Shen

1520	Force Calculations and Error Estimates with Boundary Element Methods	K. Davey & D. Zheng
1540	Use of Residual Error Bounds to Obtain Stable Numerical Solutions of a Fredholm Integral Equation of the First Kind	T. Schwengler & E.F. Kuester
1600	Eigenvalue Studies of Matrices Resulting from EFIE Simulations for Planar Structures	J.M. Dunn & H. MacMillan
1620	Iterative Solution of Dense Linear Systems in Electromagnetic Scattering Calculations	J. Rahola
1640	EMCP2 A Parallel Boundary Element Software Package Using a Novel Parameterisation Technique	G. Eriksson & U. Thibblin

FRIDAY 20 MARCH 1998 FULL-DAY SHORT COURSES

0830 - 1630 "Practical Genetic Algorithms," Randy L. Haupt, University of Nevada, Reno, Reno, NV. **Full - day course**

0830 - 1630. "Application of the Finite-Difference Time-Domain Method to Simulation of Electromagnetic Coupling to the Human Body," Cynthia Furse, Utah State University, Logan, UT. **Full - day course**

0830 - 1630 "Using Mathematical Software (MATHCAD and MATLAB) for Computational Electromagnetics," Jovan Lebaric, Naval Postgraduate School, Monterey, CA. **Full day - course**

SESSION 10:

ABCs FOR CEM: THEORETICAL AND IMPLEMENTATION

Chairs:P.G. Petropoulos and O.M. Ramahi

The Concurrent Complementary Operators Method for FDTD Mesh Truncation

Omar M. Ramahi
Digital Equipment Corporation
PKO3-1/R11
129 Parker St.
Maynard, MA 01754, U.S.A.
ramahi@poboxa.enet.dec.com

I. Abstract

A new implementation of the Complementary Operators Method (COM) for FDTD mesh truncation of open-regions is presented. This new implementation, referred to as the Concurrent Complementary Operators Method (C-COM) is based on the simultaneous application of complementary operators in a single computer run. This results in an approximate 50% reduction in the simulation cost over the original COM implementation. Numerical experiments are provided to show the flexibility of applying the C-COM theory to analytic or numerical boundary operators.

II. Introduction

The Complementary Operators Method (COM) was originally introduced as a mesh truncation technique for open region Finite-Difference Time-Domain (FDTD) simulations [1], [2]. The basic premise of the COM is the cancellation of the first-order reflection that arise when the computational domain is terminated with a single-equation boundary operator, or Absorbing Boundary Condition (ABC). This cancellation is made possible by averaging two independent solutions of the problem. The primary strength of the COM is that the cancellation of the first-order reflections takes place for any field independent of the wave number, which implies that effective suppression of the reflections occur whether the fields are composed of evanescent or purely traveling waves.

The COM requires two independent solutions of the problem, which lead to doubling the total operation count in comparison to the traditional implementation of ABCs. Nevertheless, despite the COM effectiveness, it would still be even more desirable to

avoid two independent simulations, since the overhead requirement of the simulation is then reduced by one half, and further allows for effective modeling of non-linear media [2].

In this paper, a new implementation of COM is presented. In this new implementation, instead of applying each of the operators in a separate FDTD simulation, the complementary operators are applied concurrently. This new scheme is referred to as the Concurrent Complementary Operators Method (C-COM). Here, we summarize the theory of complementary operators as was originally implemented in the COM method. Next, we discuss the implementation and performance of the concurrent implementation of the COM in two-dimensional space.

III. Complementary Operators Method

The concept underlying the COM method is the application of two independent boundary operators [2]. Let us denote an ABC by B. Then two complementary operators denoted by B_N^- and B_N^+ can be obtained by applying the ∂t and ∂x operators separately on B to obtain:

$$B_N^-[U] = \partial_x B[U] = 0 \tag{1}$$

$$B_N^+[U] = \partial_t B[U] = 0, \tag{2}$$

where U is the unknown field on which the boundary condition is applied.

It can be shown [2] that for a time-harmonic plane wave, the reflection coefficients for B_N^- and B_N^- are given respectively by:

$$R\{B_N^-\} = (-)R\{B\} \tag{3}$$

$$R\{B_N^+\} = (+)R\{B\} \tag{4}$$

The averaging of the two solutions obtained from applying each of the two operators separately gives a solution containing only second-order reflections, including those that arise from corner regions. The corner reflections, although second-order in nature, can be a significant source of error since the fields impinge at the corners at highly oblique angles which cause the second-order reflection to remain substantial in comparison to second-order reflections coming from the side boundaries. For instance, when using COM4, a wave incident at the corner at an angle 70° comes back into the domain with an approximately 1% reflection.

To cancel corner region reflections in two-dimensional space, four independent simulations, instead of two, would be needed. For each simulation, one needs to impose a unique combination of B_N^- and B_N^+ over the four sides of the outer boundary as shown in Fig. 1, where for brevity, we use $+$ to denote B_N^+, and $-$ to denotes B_N^-.

For further illustration, we show in Table 1. the magnitudes of the first and second order reflections due to the upper-right corner (assuming an incident pulse of unity

magnitude), for each of the four needed solutions. Notice that the average of all the values in the third column eliminates the corner reflections.

In the original implementation of the COM, the focus was on the annihilation of first-order reflections, and thus, only two independent simulations were considered. The four solution scheme was avoided because it was believed to lead to an excessive operation count for practical problems requiring large space and a large number of time steps. The concurrent implementation of COM is intended to achieve two objectives: 1) To implement the complementary operators within one single simulation, and 2) To allow the annihilation of corner region reflections.

IV. Concurrent Complementary Operators Method

The concurrent implementation of the COM involves the application of complementary operators at a distance from the terminal boundary (into the computational domain) such that the first-order reflections are canceled right before they reenter the computational domain. The implementation entails dividing the FDTD computational space into two regions: A boundary layer and an interior region, as shown in Fig. 2. The interior region includes the scattering object and any localized sources. First, we illustrate the application of the C-COM to reduce reflections from side boundaries only. To this end, we assign two storage (memory) locations to each nodal field in the boundary layer. (The following discussion focuses only on the treatment for the TM polarization case; the TE polarization is fully analogous.) We denote the two storage locations assigned to E_z as $E_z^{(1)}$, and $E_z^{(2)}$. Similar assignment is made for H_x and H_y giving $H_x^{(1)}$, $H_x^{(2)}$, and $H_y^{(1)}$, $H_y^{(2)}$, respectively.

Within the interior region, each of the field components is assigned a single storage location, as in typical FDTD implementation. Within the boundary layer, $E_z^{(1)}$, and $E_z^{(2)}$ are updated independently using their associated H fields. Next, we apply the two boundary operators (1) and (2) to $E_z^{(1)}$, and $E_z^{(2)}$ respectively. Notice that 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.

The next step is to connect the solutions in the two regions. This is performed by averaging the two values obtained for each field at the interface lying between the interior region and the boundary layer. The exact location of this interface defines the width of the boundary layer. This width directly impacts the additional memory overhead that will be required in comparison to standard ABC implementation. The width of the boundary layer is required to be at least the width (size) of the stencil needed for the discretization of the ABC in (1) or (2).

The above steps required for the implementation of the C-COM are summarized as follows:

- E_z , H_x and H_y are updated in the interior region according to standard FDTD

equations.

- In the boundary layer, $E_z^{(1)}$ is updated from $H_x^{(1)}$ and $H_y^{(1)}$, and $E_z^{(2)}$ is updated from $H_x^{(2)}$ and $H_y^{(2)}$. Both sets are updated using standard FDTD equations.
- B_N^- is applied to $E_z^{(1)}$ and B_N^+ is applied to $E_z^{(2)}$.
- $E_z^{(1)}$ and $E_z^{(2)}$ are averaged along the interface connecting the two regions. The new values of $E_z^{(1)}$ and $E_z^{(2)}$ along the interface are given the value of the average: $(E_z^{(1)} + E_z^{(2)})/2$.
- Advance time by one half time step.
- Update H_x and H_y in the interior region. At the interface, H_x and H_y in the interior region will use $(E_z^{(1)} + E_z^{(2)})/2$ as calculated in (4).
- In the boundary layer, $H_x^{(1)}$ and $H_y^{(1)}$ are updated using $E_z^{(1)}$, and $H_x^{(2)}$ and $H_y^{(2)}$ are updated using $E_z^{(2)}$.

We mention here that if the averaging is carried out at an interface placed within the stencil of the ABC ((1) or (2)), then the solution becomes catastrophically unstable. This is because averaging within the boundary layer creates a discontinuity that violates the analyticity of the solution.

The procedure outlined above annihilates reflections arising from side boundaries. To extend the annihilation to corner reflections, four storage locations need to be assigned to each field in the boundary layer to account for second-order reflections. For each field set, i.e., $(E_z^{(i)}, H_x^{(i)}, H_y^{(i)}, i = 1,2,3,4)$, one of the ABC combinations shown in Fig. 1 is applied. Then an identical averaging procedure to the one outlined above is performed, with the exception of having four field values to update in the boundary layer and four field values to average at the interface.

In a manner consistent with the nomenclature used for the COM method [2], the C-COM employing a 4th order operator will be denoted as C-COM4. Furthermore, we use two additional parameters to fully identify the methodology used in terms of doubling or quadrupling the fields in the boundary layer and its width. When the fields are doubled in the boundary region, resulting in the cancellation of side reflections only, we refer to the method as C-COM4(2,W), where W indicates the width of the boundary layer. Similarly, when the fields are quadrupled in the boundary region, annihilating corner reflections, we refer to the method as C-COM4(4,W).

The extension of the C-COM implementation to 3D space is performed in an entirely analogous fashion to the implementation in 2D space. To suppress reflections arising from side boundaries (single-reflection), two storage locations need to be reserved for each field in the boundary layer. The annihilation of corner reflections, however, and unlike the 2D space case, would require a total of eight storage locations for each field in the boundary layer. This is because the cancellation of corner reflections requires the imposition of eight possible unique permutations of (1) and (2) at the boundaries (2^M, where M is the number of sides forming a single corner). Notice that in the 3D computational space, there are two types of corners: The first type is a corner formed by two planes, and the second is the one formed by three planes. It can easily be

demonstrated that the cancellation of secondary reflections arising from either of the two types of corners would require 8 storage locations.

In the 3D space, the annihilation of corner reflections levies a heavy memory burden and it is therefore reserved for applications in which substantial computational overhead justifies the desired accuracy [4]

V. Numerical Results

We consider a numerical experiment to show the level of improvement achieved when the terminal boundaries are brought close to the source of radiation. Here, we choose a computational space of size 21 × 21 and uniform cell size in the x and y directions of 0.015m . The boundary layer will then be added to the outside of this domain as will be shown below. A line current source is positioned to coincide with the center of the domain which we indicate by (i_s, j_s), and an observation point is chosen close to the source at $(i_s + 5, j_s + 5)$.

The excitation waveform is a compact pulse given by the convolution $h(t) * h(t)$ where $h(t)$ is defined over the time interval $0 \leq t \leq \tau$ and is given by

$$h(t) = \pi 10^4 (15 \sin(\omega_1 t) - 12 \sin(\omega_2 t) + 3 \sin(\omega_3 t)) \tag{5}$$

where $\tau = 10^{-9}$ and $\omega_i = 2\pi\ i/\tau, i = 1, 2, 3$.

We present the results in terms of the normalized absolute error defined as

$$Error(t) = \frac{|y(t) - y^{ref}(t)|}{max[|y^{ref}(t)|]}, \tag{6}$$

where $y(t)$ is the solution that corresponds to the C-COM solution and $y^{ref}(t)$ is the reference solution (reflection-free solution).

Figure 3 shows the effect of applying the C-COM4 technique when varying the width of the boundary layer from 8 cells to 12 cells. Higdon's 3rd order ABC was used as the basic operator B (see (1) and (2)).

Figure 4 gives a comparison between the PML and C-COM4 solutions, with both having a 12 cell boundary layer. The PML layer chosen was optimized to give the lowest reflection possible. The optimization of the PML was carried out experimentally by trial and error. The PML layer with least amount of reflection for this problem was found to be PML(12,5,1e-8) (following the nomenclature of [3]). The C-COM4 solutions shown in Fig. 4 were obtained by applying complementary operators to Higdon and Liao 3rd order ABCs. Also shown in Fig. 4 is the normalized signal. From Fig. 4, we see that the C-COM procedure yields a substantial reduction of the artificial reflections especially over the portion of the pulse which contains most of the energy.

VI. Summary

A novel implementation of the complementary operators method is presented. This new implementation is based on the application of complementary operators at a distance from the terminal boundaries such that the first order reflections are annihilated before they enter the computational domain. The method is very simple to implement since it is based on the one-way wave equations such as Higdon's boundary operators.

The major accomplishment of the C-COM method is the implementation of complementary operators without the need for two independent simulations as was originally conceived in the COM method. Furthermore, the C-COM theory allows for the annihilation of corner reflections with reasonable efficiency in the 2D space. Finally, we note that unlike the COM method, the C-COM extends to scope and applicability of the complementary operators theory to the efficient treatment of non-linear media.

References

[1] O. M. Ramahi, "Complementary operators: A method to annihilate artificial reflections arising from the truncation of the computational domain in the solution of partial differential equations," *IEEE Trans. Antennas Propagat.*, vol. AP-43, no. 7, pp. 697-704, July 1995.

[2] O. M. Ramahi, "Complementary boundary operators for wave propagation problems," *J. Computat. Phys.*, vol. 133, pp. 113-128, 1997.

[3] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.*, vol. 114, pp. 185-200, 1994.

[4] O. M. Ramahi, "Extension of the concurrent COM mesh truncation technique to 3D FDTD simulation", *Microwave and Opt. Tech. Lett.*, to appear in March 1998.

	1st reflection (R1)	2nd reflection (R2)
solution #1	R	R^2
solution #2	$-R$	R^2
solution #3	R	$-R^2$
solution #4	$-R$	$-R^2$

Table 1. Corner region reflections.

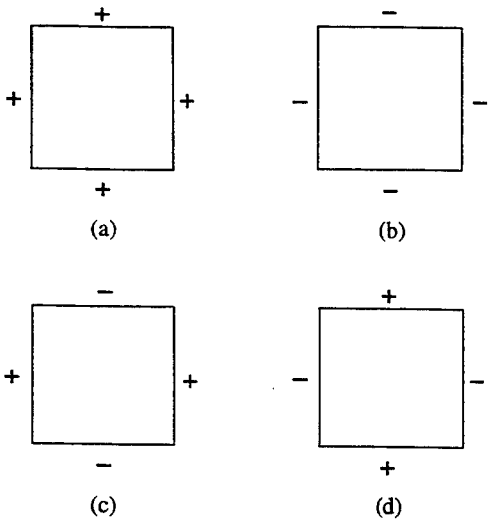


Figure 1

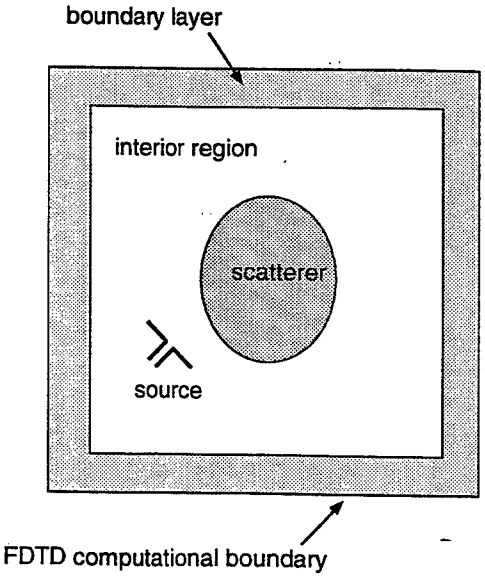


Figure 2

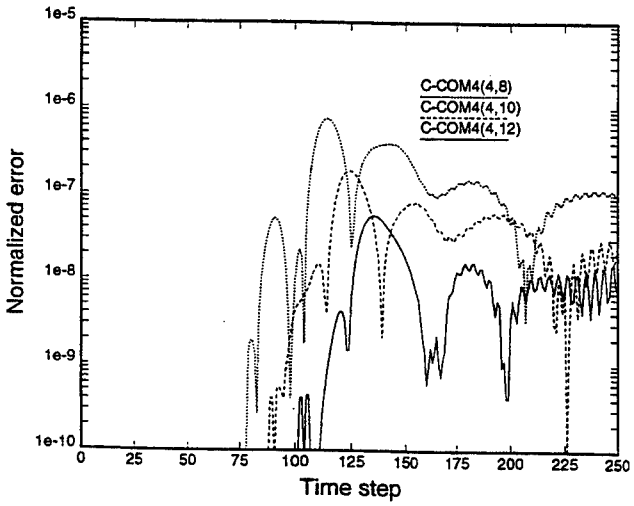


Figure 3. Solutions obtained using different C-COM layer thickness.

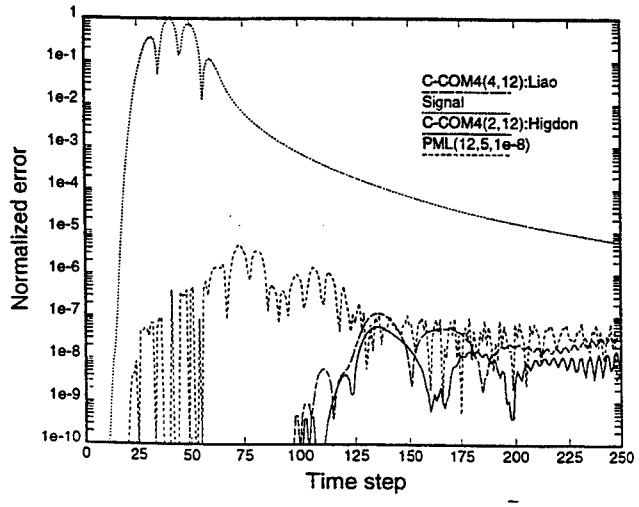


Figure 4. Errors obtained by applying the PML and C-COM techniques.

Accurate Boundary Treatments for Maxwell's Equations and their Computational Complexity

Thomas Hagstrom[†], Bradley K. Alpert[‡], Leslie F. Greengard[§], S. I. Hariharan[¶]

1 Introduction

The problem of accurate boundary treatments has long been an obstacle to the development of efficient and reliable time-domain solvers for electromagnetic wave propagation problems. Ideally, an artificial boundary would be placed immediately adjoining the part of the domain containing any inhomogeneities, and the boundary treatment would be capable of arbitrary accuracy at a cost not exceeding that of the interior solver. In this note we consider a variety of techniques capable of achieving arbitrary accuracy for special boundaries, and estimate the associated cost. For plane boundaries these include direct implementations of the exact condition as a convolution Volterra equation, high-order local boundary conditions deriving from the work of Engquist-Majda-Lindman, and stabilized absorbing layers. For spherical boundaries we consider implementations of the exact condition using local operators, in particular the conditions of Grote-Keller and a new spatially localized equivalent, as well as conditions based on uniform rational approximations.

*Contribution of U.S. Government; not subject to copyright in the United States.

[†]Department of Mathematics and Statistics, The University of New Mexico, Albuquerque, NM 87131. Supported in part by NSF Grant DMS-9600146 and by DARPA/AFOSR Contract F49620-95-C-0075. Work partially carried out while the author was visiting the Courant Institute.

[‡]National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80303. Supported in part by DARPA Appropriation 9780400.

[§]Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012-1110. Supported in part by the U.S. Department of Energy under Contract DEFGO288ER25053 and by DARPA/AFOSR under Contract F49620-95-C-0075.

[¶]Department of Mathematical Sciences, The University of Akron, Akron, OH 44325. Supported in part by NASA Contract NCC-3-283.

We find, in the planar case, that none of the conditions quite meets the goal for long time calculations. From the point of view of work, the implementation of the exact condition is acceptable, but its associated storage cost is high. The absorbing layer requires somewhat less storage, but more work. For the spherical boundary, on the other hand, all the methods presented require acceptable work and storage. Moreover, the introduction of a fast spherical harmonic transformation would make the work associated with the approximate conditions small in comparison with that required by the interior solver.

2 Plane Boundary

2.1 Exact Boundary Condition

Consider the plane boundary, $x = 0$, and suppose that all initial data, inhomogeneities, et cetera, are confined to the region $x < 0$. For $x > 0$ we have Maxwell's equations:

$$\frac{\partial E}{\partial t} = \frac{1}{\epsilon}\nabla \times H, \tag{1}$$

$$\frac{\partial H}{\partial t} = -\frac{1}{\mu}\nabla \times E. \tag{2}$$

Fourier transformation with respect to the tangential variables (dual variables k_2 and k_3) and Laplace transformation with respect to t (dual variable s) leads to a differential-algebraic equation in x. Its solution produces a parametrized representation of exact boundary conditions of the following form:

$$\frac{\partial}{\partial t}\left(\sqrt{\epsilon}E_2 - \sqrt{\mu}H_3\right) + \mathcal{R}\left((1-\alpha)\sqrt{\epsilon}E_2 - \alpha\sqrt{\mu}H_3\right) + \frac{(1-\alpha)}{\sqrt{\epsilon\mu}}\sqrt{\epsilon}\frac{\partial E_1}{\partial y} - \frac{\alpha}{\sqrt{\epsilon\mu}}\sqrt{\mu}\frac{\partial H_1}{\partial z} = 0, \tag{3}$$

$$\frac{\partial}{\partial t}\left(\sqrt{\epsilon}E_3 + \sqrt{\mu}H_2\right) + \mathcal{R}\left((1-\beta)\sqrt{\epsilon}E_3 + \beta\sqrt{\mu}H_2\right) + \frac{(1-\beta)}{\sqrt{\epsilon\mu}}\sqrt{\epsilon}\frac{\partial E_1}{\partial z} + \frac{\beta}{\sqrt{\epsilon\mu}}\sqrt{\mu}\frac{\partial H_1}{\partial y} = 0. \tag{4}$$

Here we have defined:

$$\mathcal{R}u = \mathcal{F}^{-1}\left(\frac{|k|^2}{\epsilon\mu}\mathcal{K} * (\mathcal{F}u)\right), \tag{5}$$

where $\mathcal{F}$ represents Fourier transformation in the tangential variables, $|k|^2 = k_2^2 + k_3^2$, and

$$\mathcal{K}(t) = \mathcal{L}^{-1}\left(s + \sqrt{s^2 + |k|^2/(\epsilon\mu)}\right)^{-1} = \frac{J_1\left(\frac{|k|t}{\sqrt{\epsilon\mu}}\right)}{\left(\frac{|k|t}{\sqrt{\epsilon\mu}}\right)}. \tag{6}$$

The parameters α and β are arbitrary. (For a detailed derivation of these formulas see [5].)

It is possible to directly implement these exact boundary conditions. The main expense is associated with the temporal convolution. To quantify this, suppose we are solving a problem which is 1-periodic in the y and z directions and that the length of the interior domain is also 1. Suppose further that the maximum wavenumber which must be accurately represented is N_{max} and that the time period of interest is T. The work and storage associated with the interior scheme will then scale like:

$$W_I \propto TN_{max}^4, \quad S_I \propto N_{max}^3. \tag{7}$$

For the boundary condition, using the fast method for convolution Volterra equations proposed by Hairer, Lubich and Schlichte [14], and FFT's to compute the Fourier coefficients we find:

$$W_B \propto TN_{max}^3 \ln^2(TN_{max}), \quad S_B \propto TN_{max}^3. \tag{8}$$

We see that, except for extremely large T, i.e. $T \propto e^{N_{max}}$, the work associated with the boundary condition is small compared with the work associated with the interior scheme. However, for T large, the storage required by the exact condition dominates the storage required by the interior scheme. It should be possible to reduce the storage burden by coarsening the past data, making use of the $(|k|t)^{-3/2}$ decay of the kernel. However, we have not yet carried this out. Numerical experiments using this technique for the scalar wave equation will be reported in [4].

2.2 Approximate Conditions

Generally speaking, approximate boundary treatments may be viewed as replacing the temporal convolution operator with some other operator whose action is more easily computed. That is, the operator $\mathcal{R}$ is replaced by a new operator $\mathcal{A}$. To estimate the error in the resulting solution, we must estimate the stability constant, K_A, of the approximate problem, and the difference between $\mathcal{R}$ and $\mathcal{A}$. Following [16], the latter is most easily accomplished in the dual variables. In particular we find:

$$\text{Error}(k) \propto K_A e^{\alpha T} \max_{s=\alpha+i\eta} |\hat{R}(s,k) - \hat{A}(s,k)|, \tag{9}$$

where $\alpha \geq 0$ is chosen so that $\hat{A}$ is analytic in $\Re(s) > \alpha$. Note that for time uniform estimates it is necessary that $\alpha = 0$. We see below that none of the standard approximations on plane boundaries achieve this.

The first improvable sequence of approximate boundary conditions, based on Padé approximants (in s^{-1}) of $\hat{R}$, were suggested by Engquist and Majda [7] and Lindman [15]. These are

chosen to be equivalent to local operators in both space and time. They correspond to:

$$\hat{A} = \frac{|k|^2}{\epsilon\mu}\frac{1}{q+1}\sum_{j=1}^{q}\sin^2\left(\frac{j\pi}{q+1}\right)\frac{s}{s^2+\cos^2\frac{j\pi}{q+1}\frac{|k|^2}{\epsilon\mu}}. \tag{10}$$

Note that the poles of $\hat{A}$ are located on the imaginary axis, precluding time uniform error estimates. Using either the techniques of [16] or the direct time-domain approach of [12] we find $q \propto N_{max}T$. Hence, for these conditions we require:

$$W_B \propto T^2 N_{max}^4, \quad S_B \propto T N_{max}^3. \tag{11}$$

For T large, these estimates suggest that the Padé conditions will cease to be competitive with the exact condition, though for small T they are likely to be reasonably efficient. Some improvement of the T-behavior of the estimates may be possible for non-periodic problems.

An alternative to the use of local approximate boundary conditions is to introduce some sort of absorbing layer. There has been great recent interest in this approach, spurred on by Berenger's introduction of the so-called Perfectly Matched Layer (PML) [6]. It has been shown, however, that the Berenger PML is not strongly well-posed [1]. Moreover, our own numerical experiments [8, 4] have all resulted in long time instabilities, which may be attributable to this. Recently, Abarbanel and Gottlieb [2] have proposed an alternative which avoids this difficulty. It is possible to represent the effect of the layer as an approximate boundary condition, and to apply the preceding theory. In particular, if d denotes the layer width and $\bar{\sigma}$ the average absorption we have:

$$\frac{\hat{A}(s,k)-\hat{R}(s,k)}{\hat{R}(s,k)} = 2\left(\frac{\sqrt{\epsilon\mu s^2+|k|^2}-(\sqrt{\epsilon\mu}s+\sigma(d))}{\sqrt{\epsilon\mu s^2+|k|^2}+(\sqrt{\epsilon\mu}s+\sigma(d))}\right)e^{-2d\sqrt{\epsilon\mu s^2+|k|^2}(1+\bar{\sigma}/(\sqrt{\epsilon\mu}s))}. \tag{12}$$

Clearly this is not small for $s = \pm i|k|/\sqrt{\epsilon\mu}$, precluding time-uniform estimates. Fixing $\bar{\sigma}$, we find that we must take

$$d \propto T^{1/2}, \tag{13}$$

which implies

$$W_B \propto T^{3/2} N_{max}^4, \quad S_B \propto T^{1/2} N_{max}^2. \tag{14}$$

These results are worse than those obtained for the exact condition in terms of work, but better in terms of storage. They are better in all ways (for T large) than those obtained for the Padé approximants. It is possible that the estimates can be significantly improved for non-periodic problems, i.e. that it would then be possible to choose d independent of T.

3 Spherical Boundary

3.1 Exact Condition

We now consider a spherical boundary, $r = R$. Again, using separation of variables, it is possible to derive useful representations of the exact boundary condition [5]. A particularly succinct form is:

$$\frac{\partial}{\partial t}\begin{pmatrix} \sqrt{\epsilon}E_\theta + \sqrt{\mu}H_\phi \\ \sqrt{\epsilon}E_\phi - \sqrt{\mu}H_\theta \end{pmatrix} = -\frac{1}{\epsilon\mu R^2}\sum_{n=0}^{\infty}\sum_{m=0}^{n}\left(w_{nm}^{(e)}\sqrt{\mu}(S_n * \alpha)(t) + w_{nm}^{(h)}\sqrt{\epsilon}(S_n * \beta)(t)\right), \tag{15}$$

where

$$\alpha = \frac{1}{\mu_{nm}}\langle w_{nm}^{(h)}, \begin{pmatrix} H_\theta \\ H_\phi \end{pmatrix}\rangle, \quad \beta = \frac{1}{\mu_{nm}}\langle w_{nm}^{(h)}, \begin{pmatrix} E_\theta \\ E_\phi \end{pmatrix}\rangle, \tag{16}$$

and the vector spherical harmonics are given by:

$$w_{nm}^{(e)} = \begin{pmatrix} \frac{\partial Y_n^m}{\partial\theta} \\ -\frac{1}{\sin\theta}\frac{\partial Y_n^m}{\partial\phi} \end{pmatrix}, \quad w_{nm}^{(h)} = \begin{pmatrix} \frac{1}{\sin\theta}\frac{\partial Y_n^m}{\partial\phi} \\ \frac{\partial Y_n^m}{\partial\theta} \end{pmatrix}. \tag{17}$$

The transform of the temporal convolution kernel is:

$$\hat{S}_n = -z\left(\frac{\left(z^{-1/2}K_{n+1/2}(z)\right)'}{z^{-1/2}K_{n+1/2}(z)} + 1 + \frac{1}{z}\right), \quad z = \sqrt{\epsilon\mu}Rs, \tag{18}$$

with $K_{n+1/2}$ denoting the modified spherical Bessel function.

A remarkable property of $\hat{S}_n(z)$ is that it is a rational function of degree $(n-1, n)$. This implies that convolution by S_n can be localized; that is its equivalent to the solution of an order n ordinary differential equation in t. The first to notice this property of the exact condition and to implement the resulting localized boundary condition were Grote and Keller [9, 10, 11]. Recently we discovered a continued fraction representation of $\hat{S}_n$ which allows some simplification of the Grote-Keller formulation [13]. It is:

$$\hat{S}_n(z) = -\frac{n(n+1)}{2}\cfrac{1}{z+1+\cfrac{n(n+1)-1\cdot 2}{z+2+\cfrac{n(n+1)-2\cdot 3}{z+3+\ldots}}}. \tag{19}$$

The key point is that the index n occurs only in the combination $n(n+1)$ which is the associated eigenvalue of the Beltrami operator. Hence it is possible to formulate the exact condition using

only local operators, that is without spherical harmonic transformations. In either case, it is necessary to introduce roughly N_{max} auxiliary functions at the boundary with the associated work and storage satisfying:

$$W_B \propto TN_{max}^4, \quad S_B \propto N_{max}^3. \tag{20}$$

Here we see that the work and storage required are asymptotically comparable to that of the interior scheme.

4 Approximate Conditions

In order to further reduce the complexity of the boundary condition, one must decrease the number of auxiliary functions required. One simple possibility is to apply the exact condition to $M < N_{max}$ harmonics, and treat the others using some asymptotic approximation. Although in practice this may sometimes be more efficient than the full formulation, it cannot improve the overall scaling of the work and storage. Another approach is to approximate the rational function $\hat{S}_n$ by a rational function $\hat{Q}_n$ of degree $(q-1, q)$ with $q \ll n$. Recently [3], using multipole theory, we have shown that this is possible with:

$$q \propto \ln n. \tag{21}$$

This yields:

$$S_B \propto N_{max}^2 \ln N_{max} \ll S_I. \tag{22}$$

The work, however, still includes spherical harmonic transforms at each time step and thus retains the same order. However, the development of an efficient fast spherical harmonic transform would reduce this to $TN_{max}^3 \ln^p N_{max} \ll W_I$.

References

[1] S. Abarbanel and D. Gottlieb, "A mathematical analysis of the PML method", submitted.

[2] S. Abarbanel and D. Gottlieb, "On the construction and analysis of absorbing layers in C.E.M.", submitted.

[3] B. Alpert, L. Greengard and T. Hagstrom, "On the efficient evaluation of absorbing boundary operators", in preparation.

[4] B. Alpert, L. Greengard and T. Hagstrom, "Accurate solution of the wave equation on unbounded domains", in preparation.

[5] B. Alpert, L. Greengard and T. Hagstrom, "On exact boundary conditions for Maxwell's equations", in preparation.

[6] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves", *J. Comp. Phys.*, 114, (1994), 185-200.

[7] B. Engquist and A. Majda, "Absorbing boundary conditions for the numerical simulation of waves", *Math. Comp.*, 31, (1977), 629-651.

[8] J. Goodrich and T. Hagstrom, "A comparison of two accurate boundary treatments for computational aeroacoustics" , *3rd AIAA/CEAS Aeroacoustics Conference*, (1997).

[9] M. Grote and J. Keller, "Exact nonreflecting boundary conditions for the time dependent wave equation", *SIAM J. Appl. Math*, 55, (1995), 280-297.

[10] M. Grote and J. Keller, "Nonreflecting boundary conditions for time dependent scattering", *J. Comp. Phys.*, 127, (1996), 52-81.

[11] M. Grote and J. Keller, "Nonreflecting boundary conditions for Maxwell's equations", *J. Comp. Phys.*, to appear.

[12] T. Hagstrom On high-order radiation boundary conditions, *IMA Volume on Computational Wave Propagation*, B. Engquist and G. Kriegsmann, eds., Springer-Verlag (1996), 1-22.

[13] T. Hagstrom and S.I. Hariharan, "A formulation of exact boundary conditions for the wave equation using local operators", in preparation.

[14] E. Hairer, C. Lubich and M. Schlichte, "Fast numerical solution of nonlinear Volterra convolutional equations", *SIAM J. Sci. Stat. Comp.* 6 (1985), 532-541.

[15] E. Lindman, "Free space boundary conditions for the time dependent wave equation", *J. Comp. Phys.*, 18, (1975), 66-78.

[16] L. Xu and T. Hagstrom, "On convergent sequences of approximate radiation boundary conditions", in preparation.

Perfectly Matched Layer Methods in Spherical Coordinates *

B. Yang and D. Gottlieb

Division of Applied Mathematics
Brown University, Providence, RI 02912

Abstract

In this paper, we discuss the split-field and the well-posed perfectly matched layer (PML) method in the spherical coordinate system. The PML method admits decaying plane wave solutions in the layer region that match the plane wave solutions in vacuum perfectly. For the split-field PML method, we split the fields in θ and ϕ-direction. For the well-posed spherical PML method, we only need to solve modified Maxwell's equations that is symmetric hyperbolic. Numerical experiments have been done to validate these methods.

1 Introduction

In [1] Berenger proposed the perfectly matched layer (PML) method in the context of truncating the computational domains in the numerical solution of Maxwell's equations. The method is developed for Maxwell's equations in Cartesian coordinates and the absorbing layer is shown to be nonreflecting at the vacuum-layer interface. It was extended into 3-D in [2]. As the 2-D and 3-D PML's designed by Berenger have vacuum-layer interfaces that are rectangular by construction, there have been many efforts in extending the rectangular PML method into other coordinate systems.

In [3], Kuzuoglu and Mittra presented nonplanar perfectly matched absorbers for finite-element mesh truncation. They designed PML's to absorb spherical and cylindrical waves. They also derived the reflection coefficients for the PML's and showed that the coefficients could no longer be made identically zero in general, unlike the rectangular PML method. They showed that the extension was effective when the radius of the vacuum-layer interface was electrically large. The existence of ideal nonreflecting PML methods in spherical or cylindrical coordinate system remained open. It is our purpose here to show that ideal nonreflecting PML method can be obtained in the spherical and the cylindrical coordinate systems.

Note that in [4] we already obtained the polar perfectly matched layer method in polar coordinates (2-D) where the vacuum-layer interface is a circle. Numerical results of the method turned out to be superior to other methods. This method can clearly be extended to the 3-D cylindrical case. In this paper, we concentrate on developing the 3-D spherical PML method by using the same techniques as we used in [4] and [5]. The desired vacuum-layer interface for the 3-D spherical PML method is the surface of a sphere. The method we develop admits plane wave solutions that match perfectly at the vacuum-layer interface, i.e. plane waves of any frequency and any incident angle can pass through the interface without causing any reflection.

The remaining part of the paper is organized as follows. In section 2, we give the non-dimensionalized 3-D Maxwell's equations and formulations of the plane wave solutions in the rectangular and spherical coordinate systems. Section 3 first discusses the rectangular perfectly matched layer method briefly and then presents the new perfectly matched layer methods in the spherical coordinate system. In section 4,

*Research was supported by Air Force Grant F49620-96-1-0426.

numerical results validating the methods will be presented, and concluding remarks are given in section 6.

2 The Non-dimensional Maxwell's Equations

We consider Maxwell's curl equations in free space:

$$\frac{\partial \tilde{H}}{\partial t} = -\frac{1}{\mu_0}\nabla \times \tilde{E} \ , \quad \frac{\partial \tilde{E}}{\partial t} = \frac{1}{\epsilon_0}\nabla \times \tilde{H} \ . \tag{1}$$

Here ε_0 and μ_0 are the free space permittivity and permeability, with the speed of light in free space being $\tilde{c} = (\varepsilon_0\mu_0)^{-\frac{1}{2}}$. To facilitate our analysis of the spherical PML methods, we apply the following transformation to non-dimensionalize the above equations:

$$x = \tilde{x}/L \ , \ y = \tilde{y}/L \ , \ t = \tilde{c}\tilde{t}/L \ ,$$

where L represents a scale length and the fields are normalized as

$$H = \tilde{H} \ , \ E = \sqrt{\frac{\varepsilon_0}{\mu_0}}\tilde{E} = Z_0^{-1}\tilde{E} \ ,$$

where Z_0 represents the free-space impedance. Now we obtain the non-dimensionalized Maxwell's equations in the following:

$$\frac{\partial H}{\partial t} = -\nabla \times E \ , \quad \frac{\partial E}{\partial t} = \nabla \times H \ . \tag{2}$$

The expression of the $\nabla\times$ operation in the spherical coordinate system is

$$\begin{aligned} \nabla \times A = \ & \hat{r}\tfrac{1}{r\sin\theta}[\tfrac{\partial}{\partial\theta}(\sin\theta A_\phi) - \tfrac{\partial A_\theta}{\partial\phi}] + \\ & \hat{\theta}\tfrac{1}{r}[\tfrac{1}{\sin\theta}\tfrac{\partial A_r}{\partial\phi} - \tfrac{\partial}{\partial r}(rA_\phi)] + \hat{\phi}\tfrac{1}{r}[\tfrac{\partial}{\partial r}(rA_\theta) - \tfrac{\partial A_r}{\partial\theta}] \ . \end{aligned} \tag{3}$$

Maxwell's equations admit the following plane wave solutions:

$$E = (l_1\hat{x} + m_1\hat{y} + n_1\hat{z})e^{i\omega(t-lx-my-nz)} \ , \tag{4}$$

$$H = (l_2\hat{x} + m_2\hat{y} + n_2\hat{z})e^{i\omega(t-lx-my-nz)} \ , \tag{5}$$

where

$$l_1\hat{x} + m_1\hat{y} + n_1\hat{z} = (l_2\hat{x} + m_2\hat{y} + n_2\hat{z}) \times (l\hat{x} + m\hat{y} + n\hat{z}) \ , \tag{6}$$

$$l_2\hat{x} + m_2\hat{y} + n_2\hat{z} = (l\hat{x} + m\hat{y} + n\hat{z}) \times (l_1\hat{x} + m_1\hat{y} + n_1\hat{z}) \ . \tag{7}$$

For a plane wave incident in the direction θ_0, ϕ_0, we have

$$e^{i\omega(t-lx-my-nz)} = e^{i\omega(t-r(\cos\theta_0\cos\theta+\sin\theta_0\sin\theta\cos(\phi-\phi_0)))} \ . \tag{8}$$

in spherical coordinates.

Now we write out the plane-wave field components in the spherical coordinate system in the following:

$$E_r = (\cos\phi\sin\theta l_1 + \sin\phi\sin\theta m_1 + \cos\theta n_1)e^{i\omega(t-r(\cos\theta_0\cos\theta+\sin\theta_0\sin\theta\cos(\phi-\phi_0)))} \ , \tag{9}$$

$$E_\theta = (\cos\phi\cos\theta l_1 + \sin\phi\cos\theta m_1 - \sin\theta n_1)e^{i\omega(t-r(\cos\theta_0\cos\theta+\sin\theta_0\sin\theta\cos(\phi-\phi_0)))} \ , \tag{10}$$

$$E_\phi = (-\sin\phi l_1 + \cos\phi m_1)e^{i\omega(t-r(\cos\theta_0\cos\theta+\sin\theta_0\sin\theta\cos(\phi-\phi_0)))} \ . \tag{11}$$

3 Perfectly Matched Layer Methods

Since Berenger presented the split-field PML method, efforts have been seen to modify the method to other coordinate systems and to unsplit-field formulation PML methods. Besides from the latter methods' being computationally more efficient, the efforts were shown to be worthwhile in [8] by Abarbanel and Gottlieb in that the split-field PML equations are only weakly well-posed and may suffer from instability problems.

The unsplit-field PML methods we present in this section modify Maxwell's equations by adding low-order source terms and ordinary differential equations. Hence the governing equations are symmetric hyperbolic and strongly well-posed just like the original Maxwell's equations. We also show that the well-posedness is achieved while keeping all the merits of the split-field PML methods.

3.1 The Discrete Perfectly Matched Uniaxial Medium

In [6] a PML method using an anisotropic lossy uniaxial medium was presented by Sacks et al., and was applied to frequency-domain-based finite-element methods. In [7] Gedney implemented the uniaxial medium as a PML for the FD-TD algorithm. The constitutive parameters of this anisotropic medium are given in terms of the complex permittivity and permeability tensors $\bar{\bar{\epsilon}} = \epsilon_0[\Lambda]$ and $\bar{\bar{\mu}} = \mu_0[\Lambda]$, where $[\Lambda]$ is a diagonal matrix for a uniaxial medium. In a uniaxial medium in the z-direction, non-dimensionalized Ampere's law can be expressed in matrix form as

$$\begin{bmatrix} \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \\ \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \end{bmatrix} = i\omega \begin{bmatrix} 1 + \frac{\sigma_z'(z)}{i\omega} & 0 & 0 \\ 0 & 1 + \frac{\sigma_z'(z)}{i\omega} & 0 \\ 0 & 0 & \frac{1}{1+\frac{\sigma_z'(z)}{i\omega}} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} . \tag{12}$$

One can verify that the above equations admit the following plane wave solutions, as given in [7]:

$$E = (l_1\hat{x} + m_1\hat{y} + n_1(1 + \frac{\sigma_z'(z)}{i\omega})\hat{z})e^{i\omega(t-lx-my-nz)}e^{-\sigma_z(z)n} , \tag{13}$$

$$H = (l_2\hat{x} + m_2\hat{y} + n_2(1 + \frac{\sigma_z'(z)}{i\omega})\hat{z})e^{i\omega(t-lx-my-nz)}e^{-\sigma_z(z)n} , \tag{14}$$

where (l, m, n), (l_1, m_1, n_1), and (l_2, m_2, n_2) are coupled by the relations in Eq. (6)-(7). One can see from the analytical solution that if $\frac{\sigma_z'(z)}{\omega}$ is very large and $\sigma_z(z)n$ is not large, which is possible if ω and n are very small, the magnitude of the solutions could become too large for numerical computation. An analysis of a 2-D case of this problem can be found in [9] where the directional derivative of the magnitude of the plane wave solutions are calculated and analyzed. It was shown that if in a coming pulse there is significant component of moderate frequencies there might be a problem.

However, we can design a layer that has the following plane wave solutions:

$$E = (l_1\frac{i\omega}{\sigma_z'(z) + i\omega}\hat{x} + m_1\frac{i\omega}{\sigma_z'(z) + i\omega}\hat{y} + n_1\hat{z})e^{i\omega(t-lx-my-nz)}e^{-\sigma_z(z)n} , \tag{15}$$

$$H = (l_2\frac{i\omega}{\sigma_z'(z) + i\omega}\hat{x} + m_2\frac{i\omega}{\sigma_z'(z) + i\omega}\hat{y} + n_2\hat{z})e^{i\omega(t-lx-my-nz)}e^{-\sigma_z(z)n} , \tag{16}$$

where (l,m,n), (l_1,m_1,n_1), and (l_2,m_2,n_2) are coupled by the relations in Eq. (6)-(7). The magnitude of this set of plane wave solutions are uniformly bounded. It can be verified that they are indeed solutions of the following equations:

$$\begin{bmatrix} \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \\ \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \end{bmatrix} + \begin{bmatrix} \frac{\sigma_z''(z)}{\sigma_z'(z)+i\omega} H_y \\ -\frac{\sigma_z''(z)}{\sigma_z'(z)+i\omega} H_x \\ 0 \end{bmatrix} = i\omega \begin{bmatrix} 1+\frac{\sigma_z'(z)}{i\omega} & 0 & 0 \\ 0 & 1+\frac{\sigma_z'(z)}{i\omega} & 0 \\ 0 & 0 & \frac{1}{1+\frac{\sigma_z'(z)}{i\omega}} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} . \tag{17}$$

It is interesting to notice that when $\sigma_z''(z) = 0$, Eq. (17) is the same as Eq. (12). In that case, the equation admits the unbounded solutions in Eq. (13)-(14) and the bounded solutions in Eq. (15)-(16), as $\omega \to 0$. Clearly, the physical solutions are the bounded ones.

3.2 Spherical Perfectly Matched Layer Methods

An application of Sacks' anisotropic medium idea in spherical and cylindrical coordinate systems was presented in [3] by Kuzuoglu and Mittra. They gave a full analysis of the reflection and absorption of cylindrical or spherical waves in the medium. Some restrictions and problems with this direct application of Sacks' anisotropic medium idea in those coordinate systems were observed by them. Kuzuoglu and Mittra obtained the spherical and cylindrical wave solutions in the medium and their reflection coefficients at vacuum-layer interface. It was observed that the medium was not ideally nonreflecting anymore.

In our previous work in [4]-[5] and the current work, we found that direct extensions of 2-D and 3-D rectangular PML's to polar (2-D), spherical and cylindrical (3-D) coordinates result in equations for which plane wave solutions could not be easily found. In fact, the equations we propose, which admit plane wave solutions having the same or even better properties than those of the rectangular PML's, are different in formulation compared with other extensions. In our work, our emphasis is on analyzing the analytical plane wave solutions in the layer to make sure that a PML indeed admit analytical plane wave solutions that have the same merits as those of rectangular PML's.

We have obtained the split-field formulations of the PML method in spherical coordinate systems which is ideal and as advantageous as the rectangular PML's.

$$\frac{\partial E_r}{\partial t} = \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta(H_{\phi r}+H_{\phi\theta})) - \frac{1}{r\sin\theta}\frac{\partial(H_{\theta\phi}+H_{\theta r})}{\partial\phi} - \frac{\sigma_r(r)}{r}E_r \ , \tag{18}$$

$$\frac{\partial E_{\theta\phi}}{\partial t} = \frac{1}{r\sin\theta}\frac{\partial H_r}{\partial\phi} - \frac{1}{r}(H_{\phi r}+H_{\phi\theta}) - \frac{\sigma_r(r)}{r}E_{\theta\phi} \ , \ \frac{\partial E_{\theta r}}{\partial t} = -\frac{\partial(H_{\phi r}+H_{\phi\theta})}{\partial r} - \sigma_r'(r)E_{\theta r} \ , \tag{19}$$

$$\frac{\partial E_{\phi r}}{\partial t} = \frac{\partial(H_{\theta\phi}+H_{\theta r})}{\partial r} - \sigma_r'(r)E_{\phi r} \ , \ \frac{\partial E_{\phi\theta}}{\partial t} = -\frac{1}{r}\frac{\partial H_r}{\partial\theta} + \frac{1}{r}(H_{\theta\phi}+H_{\theta r}) - \frac{\sigma_r(r)}{r}E_{\phi\theta} \ . \tag{20}$$

This method admits plane wave solutions that match at the spherical vacuum-layer interface. It can also be shown that these plane waves decay in all directions of propagation. However, since the unsplit-field PML methods may suffer from their drawback of being only weakly well-posed, in the following we want to show that the unsplit-field spherical PML methods could also be derived. We want the layer to admit solutions that are bounded like the solutions in (15)-(16) rather than those solutions in (13)-(14).

The method we propose are symmetric hyperbolic and strongly well-posed by construction. In the layer region, we want the well-posed spherical PML method to satisfy the following plane wave solutions:

$$\tilde{E}_r = E_r D(r,\theta,\phi) \ , \tilde{E}_\theta = \frac{\frac{\sigma_r(r)}{r}+i\omega}{\sigma_r'(r)+i\omega}E_\theta D(r,\theta,\phi) \ , \tilde{E}_\phi = \frac{\frac{\sigma_r(r)}{r}+i\omega}{\sigma_r'(r)+i\omega}E_\phi D(r,\theta,\phi) \ , \tag{21}$$

$$\tilde{H}_r = H_r D(r,\theta,\phi)\,, \tilde{H}_\theta = \frac{\frac{\sigma_r(r)}{r}+i\omega}{\sigma_r'(r)+i\omega} H_\theta D(r,\theta,\phi)\,, \tilde{H}_\phi = \frac{\frac{\sigma_r(r)}{r}+i\omega}{\sigma_r'(r)+i\omega} H_\phi D(r,\theta,\phi)\,, \tag{22}$$

where E_ϕ, E_θ, H_ϕ, and H_θ are plane wave solutions in Eq. (9)-(11) and

$$D(r,\theta,\phi) = e^{-\sigma_r(r)(\cos\theta_0\cos\theta+\sin\theta_0\sin\theta\cos(\phi-\phi_0))} \tag{23}$$

is the decaying factor. Let the vacuum-layer interface be at $r = r_0$. We should require that $\sigma_r(r) = 0$ for $r \le r_0$ for the decaying plane waves in the PML to match incident plane waves perfectly. Following the considerations of the absorbing and reflectionless properties of the polar PML method, $\sigma_r(r)$ is usually chosen as

$$\sigma_r(r) = C(r-r_0)^n,\ n = 1,2,\dots\,,\ r \ge r_0\,, \tag{24}$$

where C is a positive constant, such that the PML has the desired perfectly matching and absorbing properties. For the type of function $\sigma_r(r)$ we use, one notes that

$$\frac{\sigma_r(r)}{r} < \sigma_r'(r) \tag{25}$$

holds for all $r > r_0$. Hence the above plane wave solutions are uniformly bounded.

In obtaining a set of equations that admit the desired solutions, we only want to add complementary source terms to the original Maxwell's equations. The evolution of the source terms can be governed by ordinary differential equations if necessary. Now we first give the equations in the frequency domain that admit the desired plane wave solutions.

$$\frac{(\frac{\sigma_r(r)}{r}+i\omega)^2}{\sigma_r'(r)+i\omega}\tilde{E}_r = \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta(\tilde{H}_\phi)) - \frac{1}{r\sin\theta}\frac{\partial\tilde{H}_\theta}{\partial\phi}\,, \tag{26}$$

$$(i\omega+\sigma_r'(r))\tilde{E}_\theta = \frac{1}{r\sin\theta}\frac{\partial\tilde{H}_r}{\partial\phi} - \frac{\partial\tilde{H}_\phi}{\partial r} - \frac{\tilde{H}_\phi}{r} - \sigma_r''(r)Q_H\,, \tag{27}$$

$$(i\omega+\sigma_r'(r))\tilde{E}_\phi = \frac{\partial\tilde{H}_\theta}{\partial r} - \frac{1}{r}\frac{\partial\tilde{H}_r}{\partial\theta} + \frac{\tilde{H}_\theta}{r} + \sigma_r''(r)R_H\,. \tag{28}$$

Here we just verify one of the equations, Eq. (26). One only needs to notice that

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta(\tilde{H}_\phi)) - \frac{1}{r\sin\theta}\frac{\partial\tilde{H}_\theta}{\partial\phi} = \frac{\frac{\sigma_r(r)}{r}+i\omega}{\sigma_r'(r)+i\omega}(i\omega\tilde{E}_r + \frac{\sigma_r(r)}{r}\tilde{E}_r)\,, \tag{29}$$

by using the relations in Eqs. (21)-(22) and noticing that E_r, H_ϕ, and H_θ satisfy the original Maxwell's equations. Eqs. (27)-(28) are more complicated and lengthy to verify, which we hope to present soon elsewhere.

To obtain our time-domain method, we introduce

$$\bar{E}_r = \frac{\frac{\sigma_r(r)}{r}+i\omega}{\sigma_r'(r)+i\omega}\tilde{E}_r\,, \quad \bar{H}_r = \frac{\frac{\sigma_r(r)}{r}+i\omega}{\sigma_r'(r)+i\omega}\tilde{H}_r\,, \tag{30}$$

and denote

$$D_E = \tilde{E}_r - \bar{E}_r,\ D_H = \tilde{H}_r - \bar{H}_r. \tag{31}$$

We propose the following set of equations in the time domain:

$$\frac{\partial \tilde{E}_r}{\partial t} + (\frac{2\sigma_r(r)}{r} - \sigma_r'(r))\tilde{E}_r = \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta(\tilde{H}_\phi)) - \frac{1}{r\sin\theta}\frac{\partial \tilde{H}_\theta}{\partial\phi} + (\frac{\sigma_r(r)}{r} - \sigma_r'(r))D_E \ , \tag{32}$$

$$\frac{\partial \tilde{E}_\theta}{\partial t} + \sigma_r'(r)\tilde{E}_\theta = \frac{1}{r\sin\theta}\frac{\partial \tilde{H}_r}{\partial\phi} - \frac{\partial \tilde{H}_\phi}{\partial r} - \frac{\tilde{H}_\phi}{r} - \sigma_r''(r)Q_H \ , \tag{33}$$

$$\frac{\partial \tilde{E}_\phi}{\partial t} + \sigma_r'(r)\tilde{E}_\phi = \frac{\partial \tilde{H}_\theta}{\partial r} - \frac{1}{r}\frac{\partial \tilde{H}_r}{\partial\theta} + \frac{\tilde{H}_\theta}{r} + \sigma_r''(r)R_H \ , \tag{34}$$

and the supplementary ordinary differential equations are given in the following:

$$\frac{\partial D_E}{\partial t} = (\sigma_r'(r) - \frac{\sigma_r}{r})\tilde{E}_r - \sigma_r'(r)D_E \ , \tag{35}$$

$$\frac{\partial Q_H}{\partial t} = \tilde{H}_\phi - \sigma_r'(r)Q_H \ , \tag{36}$$

$$\frac{\partial R_H}{\partial t} = \tilde{H}_\theta - \sigma_r'(r)R_H \ . \tag{37}$$

Note that $\tilde{E}_r - D_E = \bar{E}_r$, $\tilde{H}_r - D_E = \bar{H}_r$, and the magnitudes of $\bar{E}_r$, $\tilde{E}_\theta$, and $\tilde{E}_\phi$ are the same, as well as the magnitudes of $\bar{H}_r$, $\tilde{H}_\theta$, and $\tilde{H}_\phi$. This is the property one desires for the multidomain numerical computation purpose, for which the detail will be given elsewhere. Note that the above set of equations are just Maxwell's equations with low order terms that are governed by ordinary differential equations in Eq. (37). The set of equations are still symmetric hyperbolic and strongly well-posed.

4 Numerical Results

To validate our PML methods, numerical experiments of electromagnetic scattering by a perfect electrical conducting (PEC) sphere have been done. The numerical scheme we use is a multidomain pseudospectral scheme. The computational domain is decomposed into a number of subdomains and a $16 \times 16 \times 16$ mesh is used in each subdomain. We have two layers of subdomains, one for the outer domain and the other layer of subdomains are next to the scatterer. In the outer layer of subdomains we apply the well-posed PML method, while in the inner subdomains we still solve the original Maxwell's equations. Detailed description of the 3-D multidomain spectral scheme does not fit in here and we hope to report it in the near future.

In Fig. 1 we present the RCS result of a PEC sphere of electrical size $ka = 5.3$. Here we use the multidomain pseudospectral method with the split-field and the well-posed spherical perfectly matched layer method. The Mie-series RCS result is also plotted for reference. One can barely tell any difference from the two RCS results.

The inner layer of subdomains, next to the sphere, spans one wavelength in the r-direction. In Fig. 2 we plot the E_x field $\frac{\lambda}{2}$ from the scatterer surface in the back scatter region, and the difference between the field and the one obtained in the reference computation using a larger computational domain. It is shown in the figure that the difference between the two fields is within 1×10^{-3} after the initial noise, which is the result of the initial non-smoothness of the type of excitation used.

5 Conclusions

Our emphasis in this paper is to present the perfectly matched layer (PML) methods in the spherical coordinate systems. The reflectionless property at the vacuum-layer interface is guaranteed because the plane wave solutions of Maxwell's equations match perfectly with the decaying plane wave solutions of the equations for the PML methods. The split-field spherical PML method is obtained by splitting

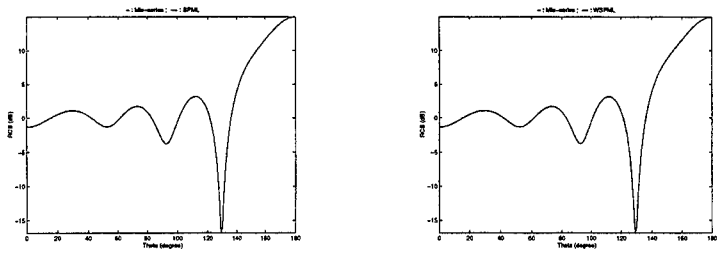


Figure 1: Comparison of RCS's obtained from spectral method with split-field spherical PML (SPML), well-posed spherical PML (WSPML), and Mie-series for a PEC sphere with electrical size $ka = 5.3$.

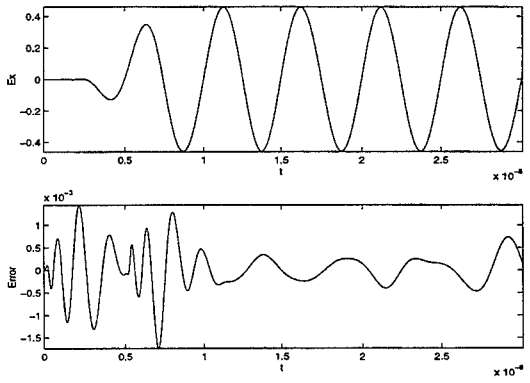


Figure 2: Comparison of field obtained from spectral method with spherical PML and reference for a PEC sphere with electrical size $ka = 5.3$.

the equations for the fields in the θ and the ϕ directions. The equations for the PML methods are obtained by modifying the original Maxwell's equations with low-order terms and O.D.E.'s. Hence they are well-posed by construction.

The PML methods are demonstrated to be effective in numerical experiments, where we compute the electromagnetic wave scattering by a sphere. They indeed carry all the merits of Berenger's rectangular PML method into the spherical coordinate systems. A detailed presentation of the 3-D multidomain spectral method we use for the electromagnetic scattering computation does not fit in the context of this paper and will be reported later.

References

[1] J.-P. Berenger, *A perfectly matched layer for the absorption of electromagnetic waves*, J. Comp. Physics, vol. 114, pp. 185-200, 1994.

[2] J.-P. Berenger, *Three-dimensional perfectly matched layer for the absorption of electromagnetic waves*, J. Comp. Physics, vol. 127, pp. 363-379, 1996.

[3] M. Kuzuoglu and R. Mittra, *Investigation of nonplanar perfectly matched absorbers for finite-element mesh truncation*, IEEE Trans. Antennas and Propagat., vol. 45, pp. 474-486, March 1997.

[4] B. Yang, D. Gottlieb and J. S. Hesthaven, *Spectral simulations of electromagnetic wave scattering*, J. Comput. Phys., vol. 134, pp. 216-230, 1997.

[5] B. Yang, D. Gottlieb and J. S. Hesthaven, *On the use of PML ABC's in spectral time-Domain simulations of electromagnetic scattering*, Proc. of the 13'th Annual Review of Progress in Applied Computational Electromagnetics, 1997.

[6] Z. S. Sacks, D. M. Kingsland, R. Lee, and J.-F. Lee, *A perfectly matched anisotropic absorber for use as an absorbing boundary condition*, IEEE Trans. Antennas and Propagat., vol. 43, pp. 1460-1463, December 1995.

[7] Stephen D. Gedney, *An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices*, IEEE Trans. Antennas and Propagat., vol. 44, pp. 1630-1639, December 1996.

[8] S. Abarbanel and D. Gottlieb, *A mathematical analysis of the PML method*, J. Comput. Phys., vol. 134, pp. 357-363, 1997.

[9] S. Abarbanel and D. Gottlieb, *On the Construction and Analysis of Absorbing Layers in C.E.M*, Proc. of the 13'th Annual Review of Progress in Applied Computational Electromagnetics, 1997.

[10] Carey M. Rappaport, *Perfectly matched absorbing boundary conditions based on anisotropic lossy mapping of space*, IEEE Microwave Guided Wave Lett., vol. 5, no. 3, pp. 94-96, Mar. 1995.

[11] R. W. Ziolkowski, *The design of Maxwellian absorbers for numerical boundary conditions and for practical applications using engineered artificial materials*, IEEE Trans. Antennas and Propagat., vol. 45, pp. 656-671, April 1997.

The Unsplit PML for Maxwell's Equations in Cylindrical and Spherical Coordinates *

Peter G. Petropoulos
Department of Mathematics, SMU
Dallas, TX 75275

1. Introduction

We wish to solve the time-dependent Maxwell equations

$$\begin{aligned} \frac{\partial \mathbf{B}}{\partial t} &= -\nabla \times \mathbf{E} \\ \frac{\partial \mathbf{D}}{\partial t} &= \nabla \times \mathbf{H} \\ \nabla \cdot \mathbf{D} &= 0 \\ \nabla \cdot \mathbf{B} &= 0, \end{aligned} \tag{1.1}$$

closed with constitutive relations

$$\mathbf{D} = \epsilon(\mathbf{x})\mathbf{E}, \quad \mathbf{B} = \mu(\mathbf{x})\mathbf{H}, \tag{1.2}$$

over a domain $\Omega_c \subset \mathcal{R}^3$ that is embedded in an infinite dielectric background medium Ω_m of constant permittivity $\epsilon(\mathbf{x}) = \epsilon$ and permeability $\mu(\mathbf{x}) = \mu$. The initial values of the fields are given functions with compact support in Ω_c. The resulting hyperbolic problem is discretized with a numerical scheme and our work does not depend on its particulars.

On the computational domain boundary $\partial\Omega_c$ an absorbing boundary condition must be imposed to provide field values for the interior solution algorithm. A multitude of such conditions has been derived and implemented by many researchers. An alternative to absorbing boundary conditions is to surround Ω_c with a wave absorbing layer Ω_m of thickness d. Ideally, the transition from Ω_c to Ω_m should not produce wave reflection while the fields that have penetrated into Ω_m should attenuate as they propagate outward. The existence of layers with such properties was shown by Berenger [1] who produced the first split-field PML. Subsequently, the unsplit PML [2] has become popular. Our approach to the derivation of an unsplit perfectly matched layer, which can be viewed as an extension of [2] in cylindrical and spherical coordinates, begins in the frequency-domain, i.e., with (1.1) after applying the Fourier transform in the time direction. Therefore, our work herein can also be used with three-dimensional elliptic solvers for the Maxwell, or Helmholtz equations, in the two coordinate systems. We do not develop our approach in rectangular coordinates since the equations produced are identical to those of [2] in the frequency- and time-domains. The layers herein are to be terminated with a Dirichlet boundary condition but other choices are possible. We will present numerical results elsewhere.

*Supported in part by AFOSR Grant F49620-98-1-0001.

2. The Method

We consider the three-dimensional frequency-domain Maxwell equations ($e^{-i\omega t}$ dependence) in a homogeneous isotropic (lossless) dielectric (with permittivity ϵ and permeability μ) that fills all of $\mathcal{R}'^3$,

$$\begin{aligned} -i\omega\epsilon(\mathbf{x}')\mathbf{E}' &= \nabla' \times \mathbf{H}', \quad \nabla' \cdot \mathbf{E}' = 0 \\ -i\omega\mu(\mathbf{x}')\mathbf{H}' &= -\nabla' \times \mathbf{E}', \quad \nabla' \cdot \mathbf{H}' = 0, \end{aligned} \tag{2.1}$$

to be in *normal form*. Then, we divide space in two parts: the volume Ω_c, identified in applications with the interior computational domain where scatterers are embedded so that $\lim_{\mathbf{x}' \to \partial\Omega_c^-} \epsilon(\mathbf{x}') = \epsilon$ and $\lim_{\mathbf{x}' \to \partial\Omega_c^-} \mu(\mathbf{x}') = \mu$, and the volume Ω_m where $\epsilon(\mathbf{x}') = \epsilon$ and $\mu(\mathbf{x}') = \mu$, which in general extends to infinity and whose presence has to be simulated in a finite-sized scattering computation.

We seek transformations of the independent and dependent variables to rewrite (2.1) in terms of real-valued spatial coordinates, i.e.,

$$\begin{aligned} \mathbf{x}' &= \begin{cases} \mathbf{x};\ \mathbf{x}' \in \Omega_c \\ \mathbf{S}(\mathbf{x},\omega)\cdot\mathbf{x};\ \mathbf{x}' \in \Omega_m \cup \partial\Omega_c \end{cases} \\ \mathbf{E}' &= \begin{cases} \mathbf{E};\ \mathbf{x}' \in \Omega_c \\ \mathbf{A^e}(\mathbf{x},\omega)\cdot\mathbf{E};\ \mathbf{x}' \in \Omega_m \cup \partial\Omega_c, \end{cases} \\ \mathbf{H}' &= \begin{cases} \mathbf{H};\ \mathbf{x}' \in \Omega_c \\ \mathbf{A^m}(\mathbf{x},\omega)\cdot\mathbf{H};\ \mathbf{x}' \in \Omega_m \cup \partial\Omega_c, \end{cases} \end{aligned} \tag{2.2}$$

where $\mathbf{x} \in \mathcal{R}^3$ is an independent variable with units of space, and $\omega \in \mathcal{R}$ is the frequency. We show below that a reflectionless wave-absorbing layer can be achieved in cylindrical and spherical coordinates if the diagonal matrices $\mathbf{S}$, $\mathbf{A^e}$, and $\mathbf{A^m}$ are chosen so that $\mathbf{S} = \mathbf{I}$ for $\mathbf{x} \in \partial\Omega_c$ and coordinate-independent expressions in $\Omega_m \cup \partial\Omega_c$, such as $\nabla' \times \mathbf{E}'$ and $\nabla' \times \mathbf{H}'$, are invariant up to an overall factor that depends on $(\mathbf{x},\omega)$.

The elements of (2.2) will involve the coordinate transformation $u' = \gamma_u(u,\omega)u$ via the function

$$\gamma_u(u,\omega) = \frac{u_0 + \int_{u_0}^{u} \alpha_u(s,\omega)ds}{u}, \tag{2.3}$$

where subscripts indicate the relevant spatial direction and $u \geq u_0 \in \mathcal{R}^+$. The necessary change of variables in $\Omega_m \cup \partial\Omega_c$ will be done with $\frac{\partial}{\partial u'} = \zeta_u \frac{\partial}{\partial u}$, where $\zeta_u = \frac{1}{\alpha_u(u,\omega)}$. Note, $\lim_{u \to u_0^+} \gamma_u = 1$ and $\lim_{u \to u_0^+} \frac{\partial u'}{\partial u} = \alpha_u(u_0,\omega)$. Our method is independent of the choice for $\alpha_u(u,\omega)$ and there are many possibilities.

Herein, we will eventually choose $\alpha_u(u,\omega) \in \mathcal{C}$ with $Im\{\alpha_u(u,\omega)\} \approx O(\frac{1}{\omega}) > 0$, i.e.,

$$\begin{aligned} \alpha_u(s,\omega) &= \xi_u\left(1 + \frac{i\sigma_u(s)}{\omega}\right);\ \ \xi_u \geq 1 \\ \sigma_u(s) &= \sigma_u^{max} s^n;\ \ n \geq 0, \end{aligned} \tag{2.4}$$

with $\xi_u \in \mathcal{R}$, $n \in \mathcal{I}$, $\sigma_u^e(u) = \epsilon\sigma_u(u)$, $\sigma_u^m(u) = \mu\sigma_u(u)$, and $\sigma_u^{max} \in \mathcal{R}^+$. Hence, the independent variables in (2.1) can be thought to be analytically continued into the space of complex numbers in $\Omega_m \cup \partial\Omega_c$ while $\mathbf{x}' \in \mathcal{R}^3$ in Ω_c. As we will show elsewhere, one may also choose $\alpha_u(s,\omega) = \xi_u(1 + \sigma_u(s)/(1 - i\omega))$ which is regular at $\omega = 0$.

2.1 Cylindrical (ρ, ϕ, z) Coordinates.

The volume Ω_c occupies the region $0 \le \rho' < \rho_0,\ 0 \le \phi' < 2\pi,\ |z'| < z_0$, while Ω_m occupies the region $\rho' \ge \rho_0,\ 0 \le \phi' < 2\pi,\ |z'| \ge z_0$. We distinguish three distinct subregions of Ω_m in which PML equations must be derived.

a. Ω_ρ region: $\rho' \ge \rho_0,\ 0 \le \phi' < 2\pi,\ |z'| \le z_0$.

The Maxwell equations (2.1) in *normal form* are

$$\begin{aligned} -i\omega\epsilon(E_{\rho'}, E_{\phi'}, E_{z'}) &= \frac{1}{\rho'}\begin{vmatrix} \hat{a}_{\rho'} & \rho'\hat{a}_{\phi'} & \hat{a}_{z'} \\ \frac{\partial}{\partial\rho'} & \frac{\partial}{\partial\phi'} & \frac{\partial}{\partial z'} \\ H_{\rho'} & (\rho' H_{\phi'}) & H_{z'} \end{vmatrix} \\ -i\omega\mu(H_{\rho'}, H_{\phi'}, H_{z'}) &= -\frac{1}{\rho'}\begin{vmatrix} \hat{a}_{\rho'} & \rho'\hat{a}_{\phi'} & \hat{a}_{z'} \\ \frac{\partial}{\partial\rho'} & \frac{\partial}{\partial\phi'} & \frac{\partial}{\partial z'} \\ E_{\rho'} & (\rho' E_{\phi'}) & E_{z'} \end{vmatrix}, \end{aligned} \tag{2.5}$$

along with $\nabla' \cdot E' = 0,\ \nabla' \cdot H' = 0$.

The presence of $\rho' H_{\phi'}$ and $\rho' E_{\phi'}$ in the r.h.s. of (2.5) leads to the realization that individual entries in these products must scale accordingly. We are naturally led to the scaling

$$\begin{aligned} &(\rho', \phi', z')^T = diag\{\gamma_\rho, 1, 1\} \cdot (\rho, \phi, z)^T \\ &\mathbf{E}' = diag\{\zeta_\rho, \frac{1}{\gamma_\rho}, 1\} \cdot \mathbf{E} \\ &\mathbf{H}' = diag\{\zeta_\rho, \frac{1}{\gamma_\rho}, 1\} \cdot \mathbf{H}. \end{aligned} \tag{2.6}$$

Applying (2.6) to (2.5) we obtain

$$\begin{aligned} -i\omega\epsilon \begin{pmatrix} \gamma_\rho\zeta_\rho & 0 & 0 \\ 0 & \frac{1}{\gamma_\rho\zeta_\rho} & 0 \\ 0 & 0 & \frac{\gamma_\rho}{\zeta_\rho} \end{pmatrix} \cdot \mathbf{E} &= \nabla \times \mathbf{H} \\ -i\omega\mu \begin{pmatrix} \gamma_\rho\zeta_\rho & 0 & 0 \\ 0 & \frac{1}{\gamma_\rho\zeta_\rho} & 0 \\ 0 & 0 & \frac{\gamma_\rho}{\zeta_\rho} \end{pmatrix} \cdot \mathbf{H} &= -\nabla \times \mathbf{E}. \end{aligned} \tag{2.7}$$

Since (2.6) enforces continuity of the unprimed tangential and primed normal electromagnetic field components (recall, $\lim_{u \to u_0^+} \gamma_u = 1$, and $u = \rho$ here), the transition from Ω_c to Ω_ρ is reflectionless. The "material" tensor in the l.h.s. of (2.7) is denoted $\mathcal{T}$.

We must now determine the behavior of waves entering Ω_ρ from Ω_c in the $\hat{\rho}$–direction. First observe that when $\nabla' \cdot \mathbf{E}' = 0$ is employed in the primed cylindrical coordinates, the vector Helmholtz equation for the electric field $\omega^2\epsilon\mu\mathbf{E}' = \nabla' \times \nabla' \times \mathbf{E}'$, obtained after the magnetic field is eliminated from (2.5), decomposes into two coupled equations for $E_{\rho'}$ and $E_{\phi'}$, while the $E_{z'}$ component uncouples and satisfies the scalar Helmholtz equation

$$\frac{1}{\rho'}\frac{\partial}{\partial\rho'}(\rho'\frac{\partial E_{z'}}{\partial\rho'}) + \frac{1}{\rho'^2}\frac{\partial^2 E_{z'}}{\partial\phi'^2} + \frac{\partial^2 E_{z'}}{\partial z'^2} + \omega^2\epsilon\mu E_{z'} = 0, \tag{2.8}$$

and similarly for $\mathbf{H}'$. The solutions of (2.8) and of the corresponding equation for $H'_{z'}$, can be used to define TM-to-$\hat{z}$ ($H'_{z'}=0$) and TE-to-$\hat{z}$ ($E'_{z'}=0$) orthogonal polarizations, respectively; a linear combination of the two polarizations is used to express an arbitrary propagating electromagnetic wave whose $\hat{z}$ wavenumber is κ_z. All components of outgoing waves are represented in the two polarizations in terms of an infinite series in the function (which solves (2.8))

$$U^m_{z'}(\rho',\phi',z') = H^{(1)}_m(\sqrt{\omega^2\epsilon\mu-\tilde{\kappa}_z^2}\rho')e^{im\phi'+i\tilde{\kappa}_z z'}, \tag{2.9}$$

where $\tilde{\kappa}_z = \omega\sqrt{\epsilon\mu}\hat{z}\cdot\mathbf{k} = \omega\sqrt{\epsilon\mu}\cos\theta$ with $\theta>0$ being the angle of propagation w.r.t. the $\hat{z}-$axis, $m\in(-\infty,\infty)$ an integer, and $H^{(1)}_m(\cdot)$ the Hankel function of the first-kind and of order m.

Transforming the zero-divergence condition using (2.6),

$$\nabla'\cdot\mathbf{E}' = 0 \rightarrow \frac{1}{\gamma_\rho\rho}\zeta_\rho\frac{\partial}{\partial\rho}(\rho\gamma_\rho\zeta_\rho E_\rho)+\frac{1}{\rho\gamma_\rho^2}\frac{\partial E_\phi}{\partial\phi}+\frac{\partial E_z}{\partial z}=0, \tag{2.10}$$

we now derive the Helmholtz equation for the E_z field component. Eliminating the magnetic field from (2.7), the vector Helmholtz equation for the electric field in Ω_ρ is

$$\omega^2\epsilon\mu\begin{pmatrix}\gamma_\rho\zeta_\rho & 0 & 0\\ 0 & \frac{1}{\gamma_\rho\zeta_\rho} & 0\\ 0 & 0 & \frac{\gamma_\rho}{\zeta_\rho}\end{pmatrix}\cdot\mathbf{E} = \nabla\times\begin{pmatrix}\frac{1}{\gamma_\rho\zeta_\rho} & 0 & 0\\ 0 & \gamma_\rho\zeta_\rho & 0\\ 0 & 0 & \frac{\zeta_\rho}{\gamma_\rho}\end{pmatrix}\cdot\nabla\times\mathbf{E}. \tag{2.11}$$

The $z-$component of (2.11) is

$$\omega^2\epsilon\mu E_z = \frac{\zeta_\rho}{\rho\gamma_\rho}[\frac{\partial}{\partial\rho}\{\zeta_\rho\rho\gamma_\rho(\frac{\partial E_\rho}{\partial z}-\frac{\partial E_z}{\partial\rho})\}-\zeta_\rho\frac{\partial}{\partial\phi}\{\frac{1}{\rho\gamma_\rho}(\frac{\partial E_z}{\partial\phi}-\frac{\partial(\rho E_\phi)}{\partial z})\}]. \tag{2.12}$$

It is clear from (2.10) that

$$\frac{\partial}{\partial z}[\frac{1}{\rho\gamma_\rho}\zeta_\rho\frac{\partial}{\partial\rho}(\zeta_\rho\rho\gamma_\rho E_\rho)+\frac{1}{\rho\gamma_\rho^2}\frac{\partial E_\phi}{\partial\phi}] = -\frac{\partial^2 E_z}{\partial z^2}, \tag{2.13}$$

so (2.12) becomes

$$\frac{1}{\rho\gamma_\rho}\zeta_\rho\frac{\partial}{\partial\rho}(\rho\gamma_\rho\zeta_\rho\frac{\partial E_z}{\partial\rho})+\frac{1}{\rho^2\gamma_\rho^2}\frac{\partial^2 E_z}{\partial\phi^2}+\frac{\partial^2 E_z}{\partial z^2}+\omega^2\epsilon\mu E_z = 0. \tag{2.14}$$

Observe that (2.14) is really (2.8) written in unprimed variables. These results also apply to the vector Helmholtz equation for the magnetic field.

Now we can show that the solution of (2.14), obtained by applying (2.3)-(2.4) with (2.6) to (2.9), decays in the $\hat{\rho}-$direction independently of the frequency, for a wave of given $\tilde{\kappa}_z$ because $\rho' = \gamma_\rho(\rho,\omega)\rho$ exhibits a positive imaginary part that varies as $O(\frac{1}{\omega})$. Thus, in the asymptotic form of the outgoing Hankel function in (2.9), the desirable frequency-independent exponential decay is achieved in the $\hat{\rho}-$direction since

$$|H^{(1)}_m(\eta\rho')| \approx \frac{|C_m|}{\sqrt{\eta|\rho'|}}e^{-\eta\rho'_i} \tag{2.15}$$

as $|\rho'| \to \infty$, where $\eta = \omega\sqrt{\epsilon\mu}\sin\theta$, and $\rho_i' = \frac{\xi_\rho}{\omega}\int_{\rho_0}^{\rho}\sigma_\rho(s)ds$ is the imaginary part of ρ'. The asymptotic condition $|\rho'| \to \infty$, while clearly achievable if $\rho \to \infty$, can also be achieved if the product $\xi_\rho\sigma_\rho(\rho)$ is sufficiently large in the sponge layer of finite width $d_\rho = \rho_1 - \rho_0$. The decay (2.15) also applies to waves moving towards $\rho = \rho_0$ in the layer. A PEC (Perfect Electric Conductor) condition is applied at $\rho = \rho_1$ $(0 \le \phi < 2\pi,\ |z| \le z_0)$ to terminate the layer in the numerical implementation.

b. $\Omega_{z\rho}$ region: $\rho' \ge \rho_0,\ 0 \le \phi' < 2\pi,\ |z'| \ge z_0$.

In this region we introduce the scaling

$$\begin{aligned}(\rho', \phi', z')^T &= diag\{\gamma_\rho, 1, \gamma_z\} \cdot (\rho, \phi, z)^T \\ \mathbf{E}' &= diag\{\zeta_\rho, \frac{1}{\gamma_\rho}, \zeta_z\} \cdot \mathbf{E} \\ \mathbf{H}' &= diag\{\zeta_\rho, \frac{1}{\gamma_\rho}, \zeta_z\} \cdot \mathbf{H},\end{aligned} \tag{2.16}$$

and find the transformed equations are similar to those in (2.7) with the "material" tensor on each of their l.h.s. now replaced by

$$\mathcal{T} = \begin{pmatrix} \frac{\gamma_\rho\zeta_\rho}{\zeta_z} & 0 & 0 \\ 0 & \frac{1}{\gamma_\rho\zeta_\rho\zeta_z} & 0 \\ 0 & 0 & \frac{\gamma_\rho\zeta_z}{\zeta_\rho} \end{pmatrix}. \tag{2.17}$$

The scaling (2.16) enforces the continuity of the unprimed tangential and primed normal electromagnetic field components across the transition from Ω_z (see (2.20)) to $\Omega_{z\rho}$ at $\rho' = \rho_0$, and across the transition from Ω_ρ (see (2.6)) to $\Omega_{z\rho}$ at $|z'| = z_0$; these transitions are reflectionless.

The divergence-free conditions and (2.16), via a derivation similar to the one that led to (2.14), now lead to the following Helmholtz equation for the $\hat{z}-$component of $\mathbf{E}$ in $\Omega_{z\rho}$

$$\frac{1}{\rho\gamma_\rho}\zeta_\rho\frac{\partial}{\partial\rho}(\rho\gamma_\rho\zeta_\rho\frac{\partial(\zeta_z E_z)}{\partial\rho}) + \frac{1}{\rho^2\gamma_\rho^2}\frac{\partial^2(\zeta_z E_z)}{\partial\phi^2} + \zeta_z\frac{\partial}{\partial z}(\zeta_z\frac{\partial}{\partial z}(\zeta_z E_z)) + \omega^2\epsilon\mu(\zeta_z E_z) = 0. \tag{2.18}$$

Working backwards with (2.16), we observe that (2.18) is really (2.8) written in unprimed variables. These results also apply to the Helmholtz equation for the $\hat{z}-$component of the magnetic field. Thus, the solution of (2.18), and of the corresponding equation for H_z, is again obtained by substituting (2.16) into (2.9); the waves will again decay independently of the frequency, but now in both coordinate directions.

This is because now both $\rho' = \gamma_\rho(\rho, \omega)\rho$ and $z' = \gamma_z(z, \omega)z$ exhibit positive imaginary parts that are $O(\frac{1}{\omega})$ hence, the decay shown on (2.15), is modified in $\Omega_{z\rho}$ and is of the form

$$|H_m^{(1)}(\eta\rho')e^{i\tilde{k}_z z'}| \approx \frac{|C_m|}{\sqrt{\eta|\rho'|}}e^{-\eta\rho_i' - \tilde{k}_z z_i'} \tag{2.19}$$

as $|\rho'| \to \infty$, $z' \ge z_0$, where $z_i' = \frac{\xi_z}{\omega}\int_{z_0}^{z}\sigma_z(s)ds$ is the imaginary part of z'. Again, waves in the layer moving towards $|z| = z_0$ and $\rho = \rho_0$ are also exponentially damped. Again, a PEC condition is applied at $\rho = \rho_1$ for $0 \le \phi < 2\pi$, $z_0 \le |z| \le z_1$, and at $|z| = z_1$ for $\rho_0 \le \rho \le \rho_1$, $0 \le \phi < 2\pi$ to terminate the corner region.

c. Ω_z region: $0 \le \rho' \le \rho_1,\ 0 \le \phi' < 2\pi,\ z_0 \le |z'| \le z_1$.

The scaling now is

$$\begin{aligned}(\rho', \phi', z')^T &= diag\{1, 1, \gamma_z\} \cdot (\rho, \phi, z)^T \\ \mathbf{E}' &= diag\{1, 1, \zeta_z\} \cdot \mathbf{E} \\ \mathbf{H}' &= diag\{1, 1, \zeta_z\} \cdot \mathbf{H},\end{aligned} \tag{2.20}$$

and the resulting $\mathcal{T}$ can be found in [2]. The resulting equations are omited as they can also be obtained by transforming the (x, y) rectangular coordinates version of our layer to polar (ρ, ϕ) coordinates. The PEC condition is now applied at $|z| = z_1$ for $0 \le \rho \le \rho_1$ to terminate the layer.

2.2 Spherical (ρ, θ, ϕ) Coordinates.

The volume Ω_c is now the interior of a sphere of radius ρ_0, and Ω_m is the volume $\rho' \ge \rho_0$. The appropriate scaling is

$$\begin{aligned}(\rho', \theta', \phi')^T &= diag\{\gamma_\rho, 1, 1\} \cdot (\rho, \theta, \phi)^T \\ \mathbf{E}' &= diag\{\zeta_\rho, \gamma_\rho^{-1}, \gamma_\rho^{-1}\} \cdot \mathbf{E} \\ \mathbf{H}' &= diag\{\zeta_\rho, \gamma_\rho^{-1}, \gamma_\rho^{-1}\} \cdot \mathbf{H},\end{aligned} \tag{2.21}$$

and the *normal form* Maxwell's equations (2.1) are transformed in Ω_ρ to the following system

$$\begin{aligned} -i\omega\epsilon \begin{pmatrix} \gamma_\rho^2\zeta_\rho & 0 & 0 \\ 0 & \frac{1}{\zeta_\rho} & 0 \\ 0 & 0 & \frac{1}{\zeta_\rho} \end{pmatrix} \cdot \mathbf{E} &= \nabla \times \mathbf{H} \\ -i\omega\mu \begin{pmatrix} \gamma_\rho^2\zeta_\rho & 0 & 0 \\ 0 & \frac{1}{\zeta_\rho} & 0 \\ 0 & 0 & \frac{1}{\zeta_\rho} \end{pmatrix} \cdot \mathbf{H} &= -\nabla \times \mathbf{E}. \end{aligned} \tag{2.22}$$

The matrix on the l.h.s. of (2.22) is labeled $\mathcal{T}$.

It remains to determine that waves in Ω_ρ decay exponentially in the $\hat{\rho}-$direction independently of ω. The solution of (2.22) can be obtained by applying (2.21) to the solution of the equations in *normal form* which can be expressed as a linear combination of Electric and Magnetic multipole fields in spherical coordinates. Each component of the multipole expansion of the solution in Ω_ρ is proportional to spherical Hankel functions of the first-kind and of order m, so it will suffice to determine how (2.21) alters their behavior. It is a simple matter to determine that the waves decay as required in Ω_ρ since

$$|h_m^{(1)}(\eta\rho')| \approx \frac{|C_m|}{\eta|\rho'|} e^{-\eta\rho_i'} \tag{2.23}$$

as $|\rho'| \to \infty$, where now $\eta = \omega\sqrt{\epsilon\mu}\hat{\rho} \cdot \mathbf{k}$, and ρ_i' is of the form given below (2.15). Waves moving in the layer towards $\rho = \rho_0$ will be similarly damped. In this case too the layer is terminated with a PEC condition applied at $\rho = \rho_1$.

3. Causality and Well-Posedness.

The time-domain formulation is obtained through the inverse Fourier transform. Introduce $\mathbf{D}(\mathbf{x},\omega) = \epsilon\mathcal{T}(\mathbf{x},\omega)\cdot\mathbf{E}(\mathbf{x},\omega)$ and $\mathbf{B}(\mathbf{x},\omega) = \mu\mathcal{T}(\mathbf{x},\omega)\cdot\mathbf{H}(\mathbf{x},\omega)$, where $\mathcal{T}$ is the diagonal matrix appearing in the transformed (2.1) in each coordinate system and sub-region. Simple algebra shows that each $\mathcal{T}$ can be decomposed as $\mathcal{T} = \mathcal{T}_\xi + \mathcal{T}_\omega$, where $\mathcal{T}_\xi$ is a diagonal real matrix independent of ω whose elements depend on the ξ_u and satisfy $\lim_{\xi_u\to 1}\mathcal{T}_\xi = \mathbf{I}$, and $\mathcal{T}_\omega$ is a diagonal complex matrix whose elements are proper rational functions of ω that vanish as $O(\frac{1}{\omega})$ for $\omega\to\infty$. Thus, the inverse Fourier transforms of the electromagnetic variables $\mathbf{P} = \epsilon\mathcal{T}_\omega(\mathbf{x},\omega)\cdot\mathbf{E}(\mathbf{x},\omega)$ and $\mathbf{M} = \mu\mathcal{T}_\omega(\mathbf{x},\omega)\cdot\mathbf{H}(\mathbf{x},\omega)$, which can be viewed as induced polarization functions with a parametric dependence on the spatial variables, satisfy ordinary differential equations in time forced by the $\mathbf{E}$ and $\mathbf{H}$ fields (that is, convolution is not necessary). Further, being lower-order terms their contributions to the hyperbolic systems in each subregion can be dropped for analysis, i.e., causality and well-posedness will depend only on the relevant $\mathcal{T}_\xi$ (which influences the principal part only) in each subregion and coordinate system.

We will provide details using the two dimensional equations in the subregion Ω_z in rectangular coordinates. The causality and well-posedness in all the remaining subregions and coordinate systems can be proven in a similar fashion since the 6×6 Maxwell system that holds in each case can be cast in a coordinate-independent form as

$$\begin{pmatrix} \frac{\partial\mathbf{E}}{\partial t} \\ \\ \frac{\partial\mathbf{H}}{\partial t} \end{pmatrix} + \begin{pmatrix} -\frac{\mathcal{T}_\xi^{-1}\cdot\nabla\times\mathbf{H}}{\epsilon} \\ \\ \frac{\mathcal{T}_\xi^{-1}\cdot\nabla\times\mathbf{E}}{\mu} \end{pmatrix} + \begin{pmatrix} \mathcal{T}_\xi^{-1}\cdot\mathcal{F}^{-1}\{\mathcal{T}_\omega\}*\mathbf{E} \\ \\ \mathcal{T}_\xi^{-1}\cdot\mathcal{F}^{-1}\{\mathcal{T}_\omega\}*\mathbf{H} \end{pmatrix} = 0, \tag{3.1}$$

where the symbol $*$ denotes matrix-vector convolution, and $\mathcal{F}^{-1}$ is the inverse Fourier transform.

The electromagnetic field consists of the vector $\mathbf{U}(x,z,\omega) = (E_x, E_z, H_y)^T$ (Transverse Electric polarization). Applying the inverse Fourier transform in the region $z < z_0$, the vector function $\mathbf{U}(x,z,t) = \mathcal{F}^{-1}\{\mathbf{U}(x,z,\omega)\}$ satisfies

$$\frac{\partial\mathbf{U}}{\partial t} + \begin{pmatrix} 0 & 0 & \frac{1}{\epsilon} \\ 0 & 0 & 0 \\ \frac{1}{\mu} & 0 & 0 \end{pmatrix}\frac{\partial\mathbf{U}}{\partial z} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\mu} \\ 0 & -\frac{1}{\epsilon} & 0 \end{pmatrix}\frac{\partial\mathbf{U}}{\partial x} = 0. \tag{3.2}$$

The eigenvalues of both matrices in (3.2) are the wave speeds $\{-c, 0, c\}$, where $c = \frac{1}{\sqrt{\epsilon\mu}}$ is the wavefront speed in the dielectric, and there is one distinct eigenvector for each eigenvalue; the system is hyperbolic and causal. The change of variables $\mathbf{V} = \mathcal{C}^{-1}\cdot\mathbf{U}$, where

$$\mathcal{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{\frac{\epsilon}{\mu}} \end{pmatrix}, \tag{3.3}$$

shows that (3.2) is symmetric, therefore strongly well-posed.

In the region $z \geq z_0$, the relevant $\mathcal{T}_\xi$ is

$$\mathcal{T}_\xi = \begin{pmatrix} \xi_z & 0 & 0 \\ 0 & \xi_z & 0 \\ 0 & 0 & \frac{1}{\xi_z} \end{pmatrix}. \tag{3.4}$$

The principal part of the system in the layer Ω_z is

$$\frac{\partial \mathbf{U}}{\partial t} + \begin{pmatrix} 0 & 0 & \frac{1}{\xi_z \epsilon} \\ 0 & 0 & 0 \\ \frac{1}{\xi_z \mu} & 0 & 0 \end{pmatrix} \frac{\partial \mathbf{U}}{\partial z} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\frac{\xi_z}{\epsilon} \\ 0 & -\frac{1}{\xi_z \mu} & 0 \end{pmatrix} \frac{\partial \mathbf{U}}{\partial x} = 0. \tag{3.5}$$

Obviously, (3.5) is identical to (3.2) when $\xi_z = 1$, thus causal and strongly well-posed. The choice $\xi_z > 1$ results in anisotropy; the eigenvalues of (3.5) in the $\hat{z}-$direction are now given by the triplet $\{-\frac{c}{\xi_z}, 0, \frac{c}{\xi_z}\}$ (slow-down), while in the $\hat{x}-$direction they are given by $\{-c, 0, c\}$.

Causality is preserved because the maximum wave speed in Ω_z, which occurs in the $\hat{x}-$direction, is c $(> \frac{c}{\xi_z})$. We now establish that (3.5), with $\xi_z > 1$, is strongly well-posed by showing that its principal part (dropping the lower-order terms) is symmetric hyperbolic, i.e., its coefficient matrices can be simultaneously symmetrized by some nonsingular similarity transformation. Applying R_z, the diagonalizer of the first matrix A in (3.5), i.e., $\Lambda_z = diag\{-\frac{c}{\xi_z}, 0, \frac{c}{\xi_z}\} = R_z^{-1} A R_z$, to the second matrix B in (3.5) we obtain

$$R_z^{-1} B R_z = \begin{pmatrix} 0 & -\frac{1}{2\mu\xi_z} & 0 \\ -\frac{\xi_z}{\epsilon} & 0 & -\frac{\xi_z}{\epsilon} \\ 0 & -\frac{1}{2\mu\xi_z} & 0 \end{pmatrix}. \tag{3.6}$$

To show that A and B can be simultaneously symmetrized, we seek a diagonal matrix D, such that $D^{-1} R_z^{-1} B R_z D$ is symmetric, and note $D^{-1} \Lambda_z D$ remains diagonal (hence symmetric). Simple algebra determines that $D = diag\{1, \sqrt{2}\xi_z \sqrt{\frac{\mu}{\epsilon}}, 1\}$ is an appropriate choice. We have just shown the principal part of (3.5) is symmetric hyperbolic, hence strongly well-posed. The equations in Ω_z, including the lower-order term, are a causal, strongly well-posed hyperbolic system.

In order to indicate how our unsplit PML would be implemented in the time-domain in cylindrical coordinates we give the hyperbolic Ampere's Law that results from (2.7). For simplicity we set $\xi_\rho = 1$, and define $\bar{\sigma}_\rho(\rho) = \frac{1}{\rho} \int_{\rho_0}^{\rho} \sigma_\rho(s) ds$. In $\rho_0 \le \rho < \rho_1$ we obtain

$$\begin{aligned} \frac{\partial D_\rho}{\partial t} = (\nabla \times \mathbf{H})_\rho \quad ; \quad & \frac{dD_\rho}{dt} + \sigma_\rho(\rho) D_\rho = \epsilon \frac{dE_\rho}{dt} + \epsilon \bar{\sigma}_\rho(\rho) E_\rho, \\ \frac{\partial D_\phi}{\partial t} = (\nabla \times \mathbf{H})_\phi \quad ; \quad & \frac{dD_\phi}{dt} + \bar{\sigma}_\rho(\rho) D_\phi = \epsilon \frac{dE_\phi}{dt} + \epsilon \sigma_\rho(\rho) E_\phi, \\ \frac{\partial D_z}{\partial t} + \sigma_\rho(\rho) D_z = (\nabla \times \mathbf{H})_z \quad ; \quad & \frac{dD_z}{dt} = \epsilon \frac{dE_z}{dt} + \epsilon \bar{\sigma}_\rho(\rho) E_z, \end{aligned} \tag{3.7}$$

and similarly for Faraday's Law. In a computational setting, the PEC condition would be imposed on E_ϕ and E_z at $\rho = \rho_1$ to truncate the layer. The causality and strong well-posedness of (3.7), along with the corresponding Faraday's Law, can be shown at once by dropping the lower-order terms through setting $\sigma_\rho = 0$. Similar considerations apply to the unsplit time-domain system of equations that are obtained after the inverse Fourier transform is applied to the $\Omega_{z\rho}$ equations (Maxwell in a "material" described by (2.17)), and to (2.22) in spherical coordinates.

References

[1] J.-P. Berenger, "A Perfectly Matched Layer for the Absorption of Electromagnetic Waves," *J. Comput. Physics*, vol. 114, pp. 185-200, 1994.

[2] L. Zhao and A.C. Cangellaris, *IEEE Trans. Microwave Theory Tech.*, vol. 44, no. 12, pp. 2555-2563, 1996.

A Comparison of the Grote-Keller and Unsplit PML Absorbing Boundary Conditions for Maxwell's Equations in Spherical Coordinates

Nikolaos V. Kantartzis*, Peter G. Petropoulos**, and Theodoros D. Tsiboukis*

*Dept. of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki GR 54006, GREECE
E-mail: kant@egnatia.ee.auth.gr; tsibukis@vergina.eng.auth.gr
**Dept. of Mathematics, Southern Methodist University, Dallas TX 75275, USA
E-mail: peterp@golem.math.smu.edu

Abstract

The exact absorbing boundary condition for Maxwell's equations in spherical coordinates, first derived and demonstrated by Grote and Keller, is evaluated against the unsplit perfectly matched layer. The latter approach is a recent generalization of the unsplit PML technique to the cylindrical and spherical coordinate systems. With numerical simulations we compare the convergence properties of both approaches, the evolution of various norms of the error they produce, and their behavior as a function of distance from the scatterer to the computational domain boundary where they are imposed. These results demonstrate that both conditions are remarkably robust, and highly accurate.

I. INTRODUCTION

Recently, Grote and Keller have presented a very promising family of exact nonreflecting boundary conditions for the solution of the time-dependent Maxwell's equations in three space dimensions [1]-[3]. These conditions are formulated on a spherical surface, outside of which the medium is assumed to be homogeneous, isotropic, and free of sources; they are local in time and nonlocal on the spherical surface, and do not involve high-order derivatives in the tangential/normal to the boundary directions. Although the artificial boundary surrounding the computational domain must be a sphere, the technique can be implemented in any coordinate system if appropriate mapping is incorporated in the interior numerical approach.

Another remarkably efficient ABC is the Berenger split-field Perfectly Matched Layer (PML) [4]-[7]. The PML is a constantly developing technique which has provided a major advance in the effort to develop accurate solvers for radiation and scattering problems. Lately, its versatility has been increased via its extension to unsplit formulations and to other coordinate systems apart from the Cartesian one. Split-field PML's were derived for cylindrical and spherical coordinates [8]-[9] based on the complex coordinate stretching approach [5]. An unsplit PML in rectangular coordinates was given in [10] while [11] presented an unsplit frequency-domain PML that combines the anisotropic medium formulation with a geometrical construction involving a particular averaging procedure in the angular direction of the polar coordinate system. An imperfect PML in curvilinear coordinates is given in [12]. In [13]-[14] the application of the PML in nonorthogonal FEM and FDTD meshes was investigated. Finally, an unsplit PML formulation for all three coordinate systems was proposed in [15]. In this technique, a coordinate and field scaling is performed in the frequency domain and is shown to be equivalent to mapping an isotropic dielectric, with certain constitutive parameters which may depend on frequency, to a dielectric that is inhomogeneous, lossy, uniaxial anisotropic, and perfectly matched to the former. By not requiring the arbitrary field splitting of other approaches it maintains the well-posedness and causality properties of Maxwell's equations providing stable numerical boundary closures whose numerical order of accuracy is equal to that of the interior scheme.

In this paper we conduct a comparison of the absorption performance of the Grote-Keller boundary condition [3] and the unsplit PML of [15] in spherical coordinates. Various cases in two (3D reduced to 2D with symmetry arguments) and three dimensions are studied. The convergence of the reflection properties as a function of grid resolution, the behavior of the techniques as their distance from the scattering object is gradually reduced as well as the evolution of certain field components at different locations in the computational domain are some of the tests conducted in the present work. From the results, it is concluded that both ABC's provide very accurate results, enabling us to treat more complex three dimensional electromagnetic problems. Also, we can state that the PML class of ABC's behaves like the exact ABC's in spherical coordinates. In a forthcoming paper we will elaborate on the comparison between PML and exact ABC's in the three commonly used coordinate systems.

II. THE NONREFLECTING GROTE-KELLER BOUNDARY CONDITIONS

A. Derivation. Let us consider time-dependent scattering from a bounded scattering region in a three-dimensional space Ω. We surround this region by a sphere $\Im$ of radius R. In $\Im^{ext}$, the region outside $\Im$, the medium is assumed to be homogeneous, isotropic, linear with constant constitutive parameters ε and μ ($c^2=1/\mu\varepsilon$) and no losses at all. As a consequence, electric field $\mathbf{E}$ and magnetic field $\mathbf{H}$ satisfy the vector wave equation in $\Im^{ext}$

$$\nabla\times\nabla\times\mathbf{E}+\frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2}=0\,,\qquad \nabla\times\nabla\times\mathbf{H}+\frac{1}{c^2}\frac{\partial^2\mathbf{H}}{\partial t^2}=0\,. \tag{1}$$

In order to solve Maxwell's equations in $\Im^{ext}$, the electromagnetic field is decomposed into transverse electric (TE) and transverse magnetic (TM) fields. Hence, the electric component of the TE multipole field of order (n, m) in spherical coordinates (r, θ, φ) is given by

$$\mathbf{E}^{TE}_{nm}(r,\theta,\varphi,t)=f_{nm}(r,t)\mathbf{V}_{nm}(\theta,\varphi)\,, \tag{2}$$

where $\mathbf{V}_{nm}$ and $\mathbf{U}_{nm}$ are the vector spherical harmonics

$$\mathbf{V}_{nm}(\theta,\varphi)=\hat{\mathbf{r}}\times\mathbf{U}_{nm}=\frac{1}{\sqrt{n(n+1)}}\left[-\frac{1}{\sin\theta}\frac{\partial Q_{nm}}{\partial\varphi}\hat{\boldsymbol{\theta}}+\frac{\partial Q_{nm}}{\partial\theta}\hat{\boldsymbol{\varphi}}\right], \tag{3}$$

$$\mathbf{U}_{nm}(\theta,\varphi)=\frac{r\nabla Q_{nm}}{\sqrt{n(n+1)}}=\frac{1}{\sqrt{n(n+1)}}\left[\frac{\partial Q_{nm}}{\partial\theta}\hat{\boldsymbol{\theta}}+\frac{1}{\sin\theta}\frac{\partial Q_{nm}}{\partial\varphi}\hat{\boldsymbol{\varphi}}\right], \tag{4}$$

defined in terms of the orthonormal (according to the L_2 inner product on the unit sphere) nm-th, spherical harmonics

$$Q_{nm}(\theta,\varphi)=\sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}}P_n^{|m|}(\cos\theta)e^{im\varphi}\,,\qquad n\geq 0,\ |m|\leq n\,, \tag{5}$$

while function f_{nm} satisfies

$$L_n[f_{nm}]\equiv\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2}-\frac{\partial^2}{\partial r^2}-\frac{2}{r}\frac{\partial}{\partial r}+\frac{n(n+1)}{r^2}\right)f_{nm}=0\,. \tag{6}$$

Similarly, the magnetic component of the TM multipole field of order (n, m) is given by

$$\mathbf{H}^{TM}_{nm}(r,\theta,\varphi,t)=g_{nm}(r,t)\mathbf{V}_{nm}(\theta,\varphi)\,, \tag{7}$$

with g_{nm} satisfying $L_n[g_{nm}]=0$. It must be mentioned here, that equations (2) and (3) constitute a complete set of solutions for Maxwell's equations in a source free region. Therefore, in $\Im^{ext}$, the total electromagnetic field is a superposition of the above multipole fields, expressed as

$$\mathbf{E}=\sum_{n\geq1}\sum_{|m|\leq n}\mathbf{E}_{nm}=\sum_{n\geq1}\sum_{|m|\leq n}\left\{f_{nm}(r,t)\mathbf{V}_{nm}+\varepsilon^{-1}\nabla\times\left[\mathbf{V}_{nm}\int_0^t g_{nm}(r,s)ds\right]\right\}, \tag{8}$$

$$\mathbf{H}=\sum_{n\geq1}\sum_{|m|\leq n}\mathbf{H}_{nm}=\sum_{n\geq1}\sum_{|m|\leq n}\left\{g_{nm}(r,t)\mathbf{V}_{nm}-\mu^{-1}\nabla\times\left[\mathbf{V}_{nm}\int_0^t f_{nm}(r,s)ds\right]\right\}. \tag{9}$$

By applying the $\hat{\mathbf{r}}\times\nabla\times$ operator to (8) and (9) and performing some mathematical manipulations based principally on calculus and differentiations we conclude to

$$\hat{\mathbf{r}}\times\nabla\times\sum_{n\geq1}\sum_{|m|\leq n}\mathbf{E}_{nm}-\frac{1}{c}\frac{\partial}{\partial t}\sum_{n\geq1}\sum_{|m|\leq n}\mathbf{E}^{\tan}_{nm}=\sum_{n\geq1}\sum_{|m|\leq n}\left\{-\frac{1}{r}\left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}\right)[rf_{nm}]\mathbf{V}_{nm}+\sqrt{\frac{\mu}{\varepsilon}}\frac{1}{r}\left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}\right)[rg_{nm}]\mathbf{U}_{nm}\right\}, \tag{10}$$

$$\hat{\mathbf{r}}\times\nabla\times\sum_{n\geq1}\sum_{|m|\leq n}\mathbf{H}_{nm}-\frac{1}{c}\frac{\partial}{\partial t}\sum_{n\geq1}\sum_{|m|\leq n}\mathbf{H}^{\tan}_{nm}=\sum_{n\geq1}\sum_{|m|\leq n}\left\{-\frac{1}{r}\left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}\right)[rg_{nm}]\mathbf{V}_{nm}-\sqrt{\frac{\varepsilon}{\mu}}\frac{1}{r}\left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}\right)[rf_{nm}]\mathbf{U}_{nm}\right\}, \tag{11}$$

where the superscript *tan* describes the tangential components of the respective fields. Although equations (10) and (11) have the form of a boundary condition, they cannot yet be used for this purpose due to the presence of the radial derivatives of the unknown functions f_{nm} and g_{nm}. Elimination of these derivatives is achieved by accepting that at $r=R$, f_{nm} and g_{nm} satisfy the boundary condition derived in [16]. That is,

$$\left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}\right)[rf_{nm}]\Big|_{r=R}=-\mathbf{d}_n\cdot\boldsymbol{\psi}^E_{nm}(t)\,,\qquad \left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}\right)[rg_{nm}]\Big|_{r=R}=-\mathbf{d}_n\cdot\boldsymbol{\psi}^H_{nm}(t) \tag{12), (13}$$

The vector functions $\boldsymbol{\psi}^E_{nm}(t)=\left\{\psi^{E,j}_{nm}(t)\right\}$ and $\boldsymbol{\psi}^H_{nm}(t)=\left\{\psi^{H,j}_{nm}(t)\right\}$ for $j=1,\ldots,n$, satisfy respectively the following linear first-order ordinary differential equations

$$\frac{1}{c}\frac{d}{dt}\boldsymbol{\psi}^E_{nm}(t)=\mathbf{A}_n\boldsymbol{\psi}^E_{nm}(t)+f_{nm}(R,t)\mathbf{e}_n\,,\quad\text{with}\quad \boldsymbol{\psi}^E_{nm}(0)=0\,, \tag{14}$$

$$\frac{1}{c}\frac{d}{dt}\psi_{nm}^{H}(t)=\mathbf{A}_n\psi_{nm}^{H}(t)+g_{nm}(R,t)\mathbf{e}_n,\quad \text{with}\quad \psi_{nm}^{H}(0)=0. \tag{15}$$

Here, $\mathbf{A}_n$ is a constant $n\times n$ matrix with elements

$$A_n^{ij}=\begin{cases}-n(n+1)/(2R^j) & \text{if } i=1\\ (n+i)(n+1-i)/(2i) & \text{if } i=j+1\\ 0 & \text{otherwise}\end{cases} \tag{16}$$

while the constant n-component vectors $\mathbf{d}_n$ and $\mathbf{e}_n$ are defined as

$$d_n^j=\frac{n(n+1)j}{2R^j},\quad j=1,\ldots,n \qquad \text{and} \qquad \mathbf{e}_n=[1,0,\ldots,0]^{\mathrm{T}}. \tag{17), (18}$$

Substitution of (12)-(18) to (10) and (11) leads to the final form of the nonreflecting boundary conditions at $r=R$

$$\hat{\mathbf{r}}\times\nabla\times\mathbf{E}-\frac{1}{c}\frac{\partial\mathbf{E}^{\tan}}{\partial t}=\frac{1}{R}\sum_{n\geq 1}\sum_{|m|\leq n}\left\{\mathbf{d}_n\cdot\psi_{nm}^{E}(t)\mathbf{V}_{nm}-\sqrt{\frac{\mu}{\varepsilon}}\mathbf{d}_n\cdot\psi_{nm}^{H}(t)\mathbf{U}_{nm}\right\}\Bigg|_{r=R}, \tag{19}$$

$$\hat{\mathbf{r}}\times\nabla\times\mathbf{H}-\frac{1}{c}\frac{\partial\mathbf{H}^{\tan}}{\partial t}=\frac{1}{R}\sum_{n\geq 1}\sum_{|m|\leq n}\left\{\mathbf{d}_n\cdot\psi_{nm}^{H}(t)\mathbf{V}_{nm}+\sqrt{\frac{\varepsilon}{\mu}}\mathbf{d}_n\cdot\psi_{nm}^{E}(t)\mathbf{U}_{nm}\right\}\Bigg|_{r=R}, \tag{20}$$

where the vector functions $\psi_{nm}^{E}(t)$ and $\psi_{nm}^{H}(t)$ satisfy (14) and (15). The unknown functions f_{nm} and g_{nm} can be calculated if we observe from (8) and (9) that the $\nabla\times(\mathbf{V}_{nm}\ldots)$ terms are orthogonal to $\mathbf{V}_{nm}$. Therefore, they can be efficiently replaced by the following inner products which involve integration with respect to θ and φ on the sphere of radius r

$$f_{nm}=\left\langle\mathbf{E}^{\tan}\big|_{r=R},\mathbf{V}_{nm}\right\rangle,\qquad g_{nm}=\left\langle\mathbf{H}^{\tan}\big|_{r=R},\mathbf{V}_{nm}\right\rangle. \tag{21), (22}$$

Equations (19) and (20) are exact and guarantee that no spurious reflections will take place at $\Im$. They only require first-order derivatives of the solution, which makes them robust and easy to implement, thus allowing for the artificial boundary to be brought in (theoretically) as close as desired to the scattering object. In spite of the more complex formulation and their global character over the artificial boundary, they are explicit, well-posed (with respect to perturbations in the initial conditions), local in time and just involve inner products with spherical harmonics. Moreover, the amount of memory needed to store the vector functions $\psi_{nm}^{E}(t)$ and $\psi_{nm}^{H}(t)$ is negligible when compared to the storage required for **E** and **H** fields, while the main computational burden is focused on the calculation of (21) and (22) and the right sides of (19) and (20).

B. Formulation of higher-order exact boundary conditions. In the context of a numerical method such as the FDTD technique, the sums encountered in (19) and (20) cannot apparently be calculated due to the infinite number of terms they incorporate. Consequently, they must be truncated at some finite value N, thus inevitably introducing an error in $n>N$ modes. In order to alleviate this shortcoming, without affecting the accuracy of the $n\leq N$ modes, an alternative formulation is presented in this section. According to this methodology, for the $n>N$ modes it is assumed that the truncated boundary conditions reduce to

$$\hat{\mathbf{r}}\times\nabla\times\mathbf{E}-\frac{1}{c}\frac{\partial\mathbf{E}^{\tan}}{\partial t}=0,\qquad \hat{\mathbf{r}}\times\nabla\times\mathbf{H}-\frac{1}{c}\frac{\partial\mathbf{H}^{\tan}}{\partial t}=0. \tag{23), (24}$$

As stated in [16], (23)-(24) are the time-dependent expressions of the first-order Peterson approximate boundary condition, which suppresses the leading term in a large distance expansion of the electromagnetic field. This condition introduces an error of $O(R^{-3})$ in $n>N$ modes. In an effort to reduce this error significantly, a good idea would be to continue transforming the second-order Peterson condition to the time domain. Thus, for the **E** field we have

$$\left\{\hat{\mathbf{r}}\times(\nabla\times)-\frac{1}{c}\frac{\partial}{\partial t}-\frac{2}{r}\right\}\left\{\hat{\mathbf{r}}\times\nabla\times\mathbf{E}-\frac{1}{c}\frac{\partial\mathbf{E}^{\tan}}{\partial t}\right\}=0, \tag{25}$$

with an error of $O(R^{-5})$. By applying the $(\hat{\mathbf{r}}\times(\nabla\times)-c^{-1}\partial_t-2/r)$ operator to (10) and (11), exact boundary conditions, truncated at $n=N$, are derived in a straightforward way. For illustration (10) becomes

$$\left\{\hat{\mathbf{r}}\times(\nabla\times)-\frac{1}{c}\frac{\partial}{\partial t}-\frac{2}{r}\right\}\left\{\hat{\mathbf{r}}\times\nabla\times\mathbf{E}-\frac{1}{c}\frac{\partial\mathbf{E}^{\tan}}{\partial t}\right\}=\sum_{n\geq 2}^{N}\sum_{|m|\leq n}\left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}+\frac{2}{r}\right)\left\{\frac{1}{r}\left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}\right)[rf_{nm}]\mathbf{V}_{nm}-\sqrt{\frac{\mu}{\varepsilon}}\frac{1}{r}\left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}\right)[rg_{nm}]\mathbf{U}_{nm}\right\}, \tag{26}$$

in which the $\hat{\mathbf{r}}\times\nabla\times(f(r)\mathbf{U}_{nm})=-\frac{1}{r}\frac{\partial(rf(r))}{\partial r}\mathbf{U}_{nm}$ formula has been used.

Taking into account that the exact second-order boundary condition for $f_{nm}(r,t)$ can be written as

$$\left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}+\frac{2}{r}\right)\left(\frac{\partial}{\partial r}+\frac{1}{c}\frac{\partial}{\partial t}\right)[rf_{nm}]\Big|_{r=R}=\mathbf{p}_n\cdot\psi_{nm}^{E}(t), \tag{27}$$

where the elements of the constant vectors $\mathbf{p}_n$ are defined by

$$p_n^j = \frac{n(n+1)j(j-1)}{2R^j}, \quad j=1,\ldots,n. \tag{28}$$

The $n\geq 2$ in the first sum of (25) is attributed to the fact that $p_1^1 = 0$, and thus the $n=1$ terms vanish.

Finally, the expressions for the truncated exact nonreflecting boundary conditions at $r=R$, in terms of the aforementioned algorithm are shown below

$$\left\{\hat{\mathbf{r}}\times(\nabla\times)-\frac{1}{c}\frac{\partial}{\partial t}-\frac{2}{R}\right\}\left\{\hat{\mathbf{r}}\times\nabla\times\mathbf{E}-\frac{1}{c}\frac{\partial\mathbf{E}^{\tan}}{\partial t}\right\}=\frac{1}{R^2}\sum_{n\geq 2}^{N}\sum_{|m|\leq n}\left\{\mathbf{p}_n\cdot\psi_{nm}^{E}(t)\mathbf{V}_{nm}-\sqrt{\frac{\mu}{\varepsilon}}\mathbf{p}_n\cdot\psi_{nm}^{H}(t)\mathbf{U}_{nm}\right\}, \tag{29}$$

$$\left\{\hat{\mathbf{r}}\times(\nabla\times)-\frac{1}{c}\frac{\partial}{\partial t}-\frac{2}{R}\right\}\left\{\hat{\mathbf{r}}\times\nabla\times\mathbf{H}-\frac{1}{c}\frac{\partial\mathbf{H}^{\tan}}{\partial t}\right\}=\frac{1}{R^2}\sum_{n\geq 2}^{N}\sum_{|m|\leq n}\left\{\mathbf{p}_n\cdot\psi_{nm}^{H}(t)\mathbf{V}_{nm}+\sqrt{\frac{\varepsilon}{\mu}}\mathbf{p}_n\cdot\psi_{nm}^{E}(t)\mathbf{U}_{nm}\right\}, \tag{30}$$

C. FDTD implementation. The Grote-Keller boundary condition can efficiently terminate computational domains only in spherical coordinates. So, the curvilinear FDTD method [17]-[20] must be implemented. As this is has been studied in a variety of scientific publications, analysis here will concentrate on special features of the boundary conditions, such as the computation of the inner products appearing in (21) and (22). It must be mentioned though that if appropriate mapping is performed, then this artificial boundary is not restricted to any coordinate system. In every FDTD lattice only one of the two electromagnetic field components must be absorbed at the boundary. Here, we will assume that the **E** fields are located there. Therefore, $\mathbf{E}^{\tan}$ is known at $r=R-\Delta r$ and $r=R$, while $\mathbf{H}^{\tan}$ at $r=R-\Delta r/2$. Since Maxwell's equations at $r=R$, in order to advance $\mathbf{E}^{\tan}$ with the leapfrog scheme, will require radial derivatives of $\mathbf{H}^{\tan}$, whose finite difference stencils involve values outside $\mathfrak{S}$, expression (29) must be used. This can be achieved if we apply (29) at $t=t+\Delta t/2$ and $r=R-\Delta r/2$, and approximate first-order derivatives on the left by centered finite differences. The inner products of (21) and (22) are computed over the sphere $r=R-\Delta r/2$ using the fourth-order Simpson rule, whereas (14) and (15) are solved with the unconditionally stable trapezoidal integration scheme as

$$\left(\mathbf{I}-\frac{\Delta t}{2}\mathbf{A}_n\right)\psi_{nm}^{E}(t_{k+1/2})=\left(\mathbf{I}+\frac{\Delta t}{2}\mathbf{A}_n\right)\psi_{nm}^{E}(t_{k-1/2})+\Delta t\left\langle\frac{\mathbf{E}|_{r=R-\Delta r}^{k}+\mathbf{E}|_{r=R}^{k}}{2},\mathbf{V}_{nm}\right\rangle\mathbf{e}_n, \tag{31}$$

$$\left(\mathbf{I}-\frac{\Delta t}{2}\mathbf{A}_n\right)\psi_{nm}^{H}(t_{k+1/2})=\left(\mathbf{I}+\frac{\Delta t}{2}\mathbf{A}_n\right)\psi_{nm}^{H}(t_{k-1/2})+\frac{\Delta t}{2}\left\langle\mathbf{H}|_{r=R-\Delta r/2}^{k-1/2}+\mathbf{H}|_{r=R-\Delta r/2}^{k+1/2},\mathbf{V}_{nm}\right\rangle\mathbf{e}_n. \tag{32}$$

The complete form of the algorithm has the following steps:

1. Initialization of **E** at $t=0$ and **H** at $t=\Delta t/2$ as well as $\psi_{nm}^{E}(\Delta t/2)=0$ and $\psi_{nm}^{H}(\Delta t/2)=0$.
2. Calculation of **E** at $t_k=t_{k-1}+\Delta t$ in every point of the computational domain Ω according to the usual FDTD technique.
3. Calculation of $\mathbf{E}^{\tan}$ at t_k and $r=R$ in terms of (29) applied at $r=R-\Delta r/2$ and $t_{k-1/2}=t_{k-1}+\Delta t/2$.
4. Calculation of **H** at $t_{k+1/2}$ in every point of the computational domain Ω according to the usual FDTD technique.
5. Calculation of $\psi_{nm}^{E}(t)$ and $\psi_{nm}^{H}(t)$ at $t_{k+1/2}$ in terms of (31) and (32), and return to step 1.

III. THE UNSPLIT PML IN SPHERICAL COORDINATES - REFLECTIONLESS SPONGE LAYERS

The complete methodology for the derivation of the reflectionless sponge layers is fully presented in [15], therefore we will concentrate only on the construction of the absorber in spherical coordinates since this is the case which is going to be compared with the previously described Grote-Keller boundary condition.

We assume that the three-dimensional frequency-domain Maxwell's equations in a homogeneous, isotropic, lossless dielectric that fills all of R^3

$$-j\omega\varepsilon\mathbf{E}'=\nabla'\times\mathbf{H}', \qquad -j\omega\mu\mathbf{H}'=-\nabla'\times\mathbf{E}', \qquad \nabla'\cdot\mathbf{E}'=0, \qquad \nabla'\cdot\mathbf{H}'=0 \tag{33}$$

are in normal form. In spherical coordinates (r, θ, φ) we divide space in two parts, as shown in Fig. 1: volume Ω_c which is a sphere of radius r_0 and volume Ω_r extending from r_0 to infinity ($r'\geq r_0$) whose presence has to be simulated in a finite-sized scattering computation. We also take into account the fact that the independent variables of (33) are analytically continued into the space of complex numbers in $\Omega_r\cup\partial\Omega_c$ while $\mathbf{x}'\in R^3$ in Ω_c. The main objective is to find the appropriate transformations of the independent and dependent variables to rewrite (33) in terms of real-valued spatial coordinates, that is

$$\mathbf{x}'=\begin{cases}\mathbf{x};\ \mathbf{x}'\in\Omega_c\\ \mathbf{S}(\mathbf{x},\omega)\cdot\mathbf{x};\ \mathbf{x}'\in\Omega_r\cup\partial\Omega_c\end{cases}, \qquad \mathbf{E}'=\begin{cases}\mathbf{E};\ \mathbf{x}'\in\Omega_c\\ \mathbf{A}^{\mathbf{e}}(\mathbf{x},\omega)\cdot\mathbf{E};\ \mathbf{x}'\in\Omega_r\cup\partial\Omega_c\end{cases}, \qquad \mathbf{H}'=\begin{cases}\mathbf{H};\ \mathbf{x}'\in\Omega_c\\ \mathbf{A}^{\mathbf{m}}(\mathbf{x},\omega)\cdot\mathbf{H};\ \mathbf{x}'\in\Omega_r\cup\partial\Omega_c\end{cases}, \tag{34}$$

where $\mathbf{x},\omega\in R^3$. Matrices $\mathbf{S}$, $\mathbf{A}^{\mathbf{e}}$ and $\mathbf{A}^{\mathbf{m}}$ are chosen so that $\mathbf{S}=\mathbf{I}$ for $\mathbf{x}\in\partial\Omega_c$ and coordinate-independent expressions in $\Omega_r\cup\partial\Omega_c$, such as the curl operator of a vector field, are invariant up to an overall complex-valued factor. Since, in the primed variables, all components (tangential and normal) of $\mathbf{E}'$ and $\mathbf{H}'$ are continuous across $\partial\Omega_c$, the transition from Ω_c to Ω_r is completely re-

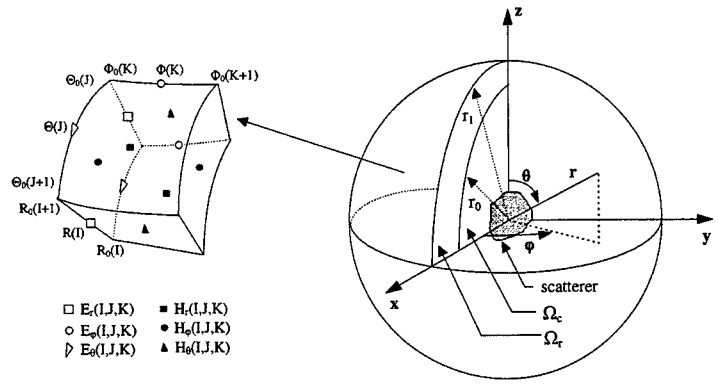


Fig. 1. The geometry of the PML and the FDTD cell in spherical coordinates.

flectionless. Finally, region Ω_r is truncated at some distance d from $\partial\Omega_c$ by imposing a simple Perfectly Electric Conductor (PEC), or a Bayliss-Turkel boundary condition. This allows for the construction of layers with an exponentially small reflection coefficient.

The Maxwell's curl equations (33) in spherical coordinates can now be written as

$$-j\omega\varepsilon(E_{r'},E_{\theta'},E_{\varphi'}) = \frac{1}{r'^2\sin\theta'}\begin{vmatrix} \hat{a}_{r'} & r'\hat{a}_{\theta'} & r'\sin\theta'\hat{a}_{\varphi'} \\ \frac{\partial}{\partial r'} & \frac{\partial}{\partial\theta'} & \frac{\partial}{\partial\varphi'} \\ H_{r'} & (r'H_{\theta'}) & (r'\sin\theta'H_{\varphi'}) \end{vmatrix}, \tag{35}$$

$$-j\omega\mu(H_{r'},H_{\theta'},H_{\varphi'}) = -\frac{1}{r'^2\sin\theta'}\begin{vmatrix} \hat{a}_{r'} & r'\hat{a}_{\theta'} & r'\sin\theta'\hat{a}_{\varphi'} \\ \frac{\partial}{\partial r'} & \frac{\partial}{\partial\theta'} & \frac{\partial}{\partial\varphi'} \\ E_{r'} & (r'E_{\theta'}) & (r'\sin\theta'E_{\varphi'}) \end{vmatrix}. \tag{36}$$

The appropriate transformation (34) is

$$(r',\theta',\varphi')^{\mathrm{T}} = diag\{\gamma_r,1,1\}\cdot(r,\theta,\varphi)^{\mathrm{T}}, \qquad \mathbf{E}' = diag\{\zeta_r,\gamma_r^{-1},\gamma_r^{-1}\}\cdot\mathbf{E}, \qquad \mathbf{H}' = diag\{\zeta_r,\gamma_r^{-1},\gamma_r^{-1}\}\cdot\mathbf{H}, \tag{37}$$

and (35) and (36) become

$$-j\omega\varepsilon\begin{pmatrix} \gamma_r^2\zeta_r & 0 & 0 \\ 0 & \zeta_r^{-1} & 0 \\ 0 & 0 & \zeta_r^{-1} \end{pmatrix}\cdot\mathbf{E} = \nabla\times\mathbf{H}, \qquad -j\omega\mu\begin{pmatrix} \gamma_r^2\zeta_r & 0 & 0 \\ 0 & \zeta_r^{-1} & 0 \\ 0 & 0 & \zeta_r^{-1} \end{pmatrix}\cdot\mathbf{H} = -\nabla\times\mathbf{E}. \tag{38}$$

In the above expressions (also used in [9], [11])

$$\gamma_r(r,\omega) = \frac{r_0 + \int_{r_0}^{r}\alpha_r(s,\omega)ds}{r}, \qquad \alpha_r(r,\omega) = \xi_r\left(1 + j\frac{\sigma_r(r)}{\omega}\right); \quad \xi_r \geq 1, \qquad \sigma_r(r) = \sigma_r^{\max}r^n; \quad n \geq 0, \qquad \zeta_r = \frac{1}{\alpha_r(r,\omega)}, \tag{39}$$

where $\xi_r \in R$, $n \in I$ and $\sigma_r^{\max} \in R^+$. From (38) we can easily derive the unsplit time-domain PML in spherical coordinates (region $r_0 \leq r < r_1$) and we give the hyperbolic Ampere's law, where for simplicity $\xi_r = 1$ and $\bar{\sigma}_r(r) = \frac{1}{r}\int_{r_0}^{r}\sigma_r(s)ds$,

$$\begin{aligned} \frac{\partial D_r}{\partial t} + \bar{\sigma}_r(r)D_r = (\nabla\times\mathbf{H})_r \quad ; \quad & \frac{dD_r}{dt} + \sigma_r(r)D_r = \varepsilon\frac{dE_r}{dt} + \varepsilon\bar{\sigma}_r(r)E_r \\ \frac{\partial D_\theta}{\partial t} = (\nabla\times\mathbf{H})_\theta \quad ; \quad & \frac{dD_\theta}{dt} + \sigma_r(r)D_\theta = \varepsilon\frac{dE_\theta}{dt} \\ \frac{\partial D_\varphi}{\partial t} = (\nabla\times\mathbf{H})_\varphi \quad ; \quad & \frac{dD_\varphi}{dt} + \sigma_r(r)D_\varphi = \varepsilon\frac{dE_\varphi}{dt} \end{aligned} \tag{40}$$

The frequency-independent exponential decay is achieved in the r-direction, in terms of the spherical Hankel functions of the first-kind and of m-order, as follows

$$\left|h_m^{(1)}(\eta r')\right| \approx \frac{|C_m|}{\eta|r'|}e^{-\eta r'_i}; \quad |r'| \to \infty, \qquad \text{with} \quad \eta = \omega\sqrt{\varepsilon\mu}\hat{\mathbf{r}}\cdot\mathbf{k} \quad \text{and} \quad r'_i = \mathrm{Im}\{r'\} = \frac{\xi_r}{\omega}\int_{r_0}^{r}\sigma_r(s)ds. \tag{41}$$

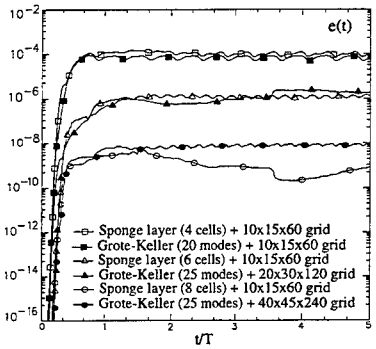


Fig. 1. The evolution of the error (43) as a function of time.

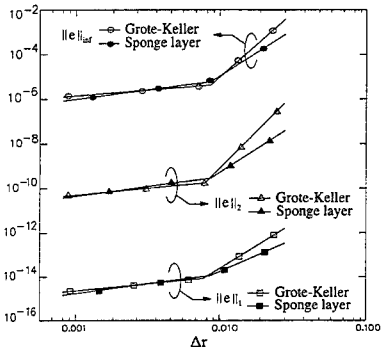


Fig. 2. Convergence of reflection at the rate of the interior scheme.

IV. NUMERICAL RESULTS

We implemented the Grote-Keller (given in (29) and (30)-(31), with the total number of modes as a free parameter) and the unsplit PML (40) ABC's in spherical coordinates to truncate computational domains in which the second-order accurate curvilinear FDTD scheme [17]-[20] is used to solve scattering and radiation problems.

Our initial tests involve a 3D spherical scatterer of permittivity $3\varepsilon_0$, permeability μ_0 and radius r_{sc} illuminated by a wave generated with a pulsed φ-directed magnetic-current point-source, whose time-profile is given by the smooth function

$$g(t) = H_0(10 - 15\cos\omega_1 t + 6\cos\omega_2 t - \cos\omega_3 t) \tag{42}$$

that is compactly supported in $t \in [0, T]$, where H_0 is the maximum source amplitude. The scatterer is centered on the grid at $(r_s, \pi/2, \varphi_s)$, and the point source is placed at $(r', \pi/2, \varphi')$ so that $|r_s - r'| = 2r_{sc}$. Therefore, numerical computations of the reflection errors can be performed along this particular 2D transverse cut ($\theta = \pi/2$ plane). Of course a variety of other source positions could have been implemented in this problem. We assume that $T = 10^{-9}$sec, $\omega_m = 2\pi m/T$ with $m = 1,2,3$, $r_{sc} = 2cT/3$ and $H_0 = (\mu_0/\varepsilon)^{1/2}/320$. This scattering problem is embedded in an infinite three-dimensional free space, and solved numerically in a finite-sized test domain Ω_c with boundary $\partial\Omega_c$ (of radius $3cT/2$), using the curvilinear FDTD method. For the computation of the exact, reflectionless solution, we merely extend the mesh into a much larger domain Ω_L terminated by PEC conditions on $\partial\Omega_L$. Truncation of Ω_c is performed either by the Grote-Keller conditions located at $\partial\Omega_c$, or by the PML (sponge layer) of a certain thickness, which in turn is terminated by PEC condition on the tangential fields. Ω_c is evenly discretized with 10, 20 or 40 intervals in the r-direction, 15, 30 or 45 intervals in the θ-direction, and 60, 120 or 240 intervals in the φ-direction, while the total computation time is $T_{tot} = 5T$. As it is known from the Cartesian coordinates, the stability of the Yee scheme requires the Courant stability condition $\Delta t \le c\Delta h/\sqrt{3}$ to be fulfilled. Here, we set Δt equal to the shortest edge in the mesh multiplied by $c/\sqrt{3}$. For the sponge layers we select $\xi_r = 1$ and a parabolic variation for the conductivity function $\sigma_r(r) = \sigma_r^{\max}(r - r_0)^2$.

The error measure is

$$e(n\Delta t) = \left\| H^{\Omega_c}(r,\varphi,n\Delta t) - H^{\Omega_L}(r,\varphi,n\Delta t) \right\|_2 \quad ; \quad n \in [0, T_{tot}/\Delta t] \tag{43}$$

where the L_2 norm is calculated over $\Omega_c \cup \partial\Omega_c$. Fig. 1 indicates the error (43) versus time computed in the $\theta = \pi/2$ plane. As can be observed, the behavior of a 4-cell thick sponge layer is almost equivalent to a Grote-Keller condition with 20 modes for the small 10×15×60 grid. Further increase of the number of modes leaves the error unaffected at that resolution and this is mainly attributed to the coarse discretization and not to the boundary condition. Then we increased the number of modes to 25 on two successive grids of size 20×30×120 and 40×45×240 for the Grote-Keller ABC, while we maintained the 10×15×60 grid for the unsplit PML. It was found that merely increasing by two the number of cells in the PML was sufficient for the two errors to be comparable.

Also, we calculated $\|e(.)\|_l$, $l = 1,2,\infty$ for $n \in [0, T_{tot}/\Delta t]$, on a progressively refined grid, while keeping the rest of the physical domain parameters the same. In Fig. 2, the rate of convergence for the reflection property of the two ABC's is presented. One can see that for all the cases both the Grote-Keller (20 modes) and PML (6 cells) techniques converge to the exact solution as the grid is successively halved at the same rate. In Fig. 3, the $\|e(.)\|_2$ norm ($\theta = \pi/2$ plane) versus grid resolution over the time interval $[0,5T]$ is shown for the indicated number of modes (in the Grote-Keller ABC) and cells (in the PML). As the number of modes N used for the truncation of the infinite sums in the Grote-Keller boundary conditions is an issue of great interest,we

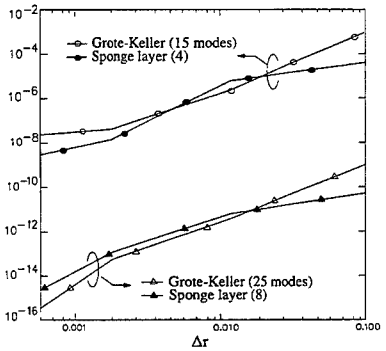


Fig. 3. Convergence of the reflection in the L_2 norm.

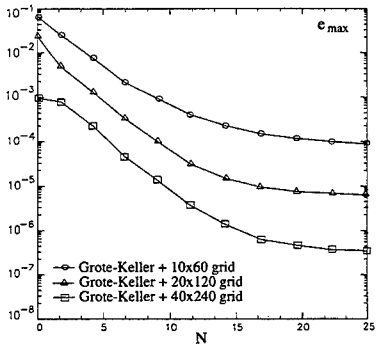


Fig. 4. The effect of N on the maximum error for the Grote-Keller ABC.

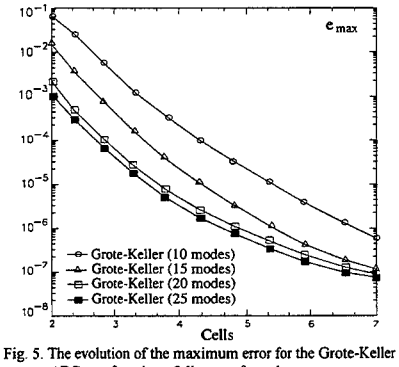


Fig. 5. The evolution of the maximum error for the Grote-Keller ABC as a function of distance from the scatterer.

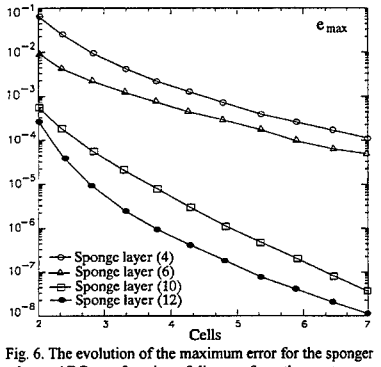


Fig. 6. The evolution of the maximum error for the sponger layers ABC as a function of distance from the scatterer.

tested the method by changing N for three different types of grids. We denote the maximum error over the time interval $[0,5T]$ as $e_{\max} = \|e(.)\|_{\inf}$ and give results in Fig. 4. It must be mentioned here that after 20 modes approximately the reduction of the error ceases and further increase of N seems to be pointless. This, of course, means that the grid resolution must become finer in order to obtain higher levels of accuracy.

In all of the above experiments the distance of the absorbing boundary from the scatterer remained constant. Evidently, since computational resources should be kept at a minimum, we decided to study both methods as function of distance. The results are displayed in Figs 5 and 6 for a 40×45×240 FDTD grid. The PML seems to behave better than the Grote-Keller method since the error levels it introduces are smaller than the ones of the latter.

Finally, we will consider the three dimensional problem presented in [3]. It involves an off-centered radiating electric dipole located a distance z_0=0.4m from the origin. The dipole is aligned along the z-axis thus allowing for the computational domain to be reduced to a 2D one in the (r, θ) plane. Its time dependence is a Gaussian pulse centered at $t=t_0$

$$P(t) = \begin{cases} 0 & t < 0 \\ 10^{-9} e^{-(t-t_0)^2/\sigma^2} & 0 \le t \le 2t_0 \\ 0 & t > 2t_0 \end{cases} \tag{44}$$

and the total computation time is again $T_{tot}=5T$. For the mesh discretization we select a significantly refined 60×360 grid, while the artificial boundaries are imposed at R=1m from the center of the origin. We select two observation points in the computational domain, p_1=(0.75, 85°) and p_2=(0.35, 150°), where we study the evolution of the H_φ waveforms. As can be seen in Figs 7 and 8, both boundary conditions can cope sufficiently with these kind of problems regardless of the location of the observation points. In spite of the axisymmetric character of the previous experiment we also solved it in three dimensions in

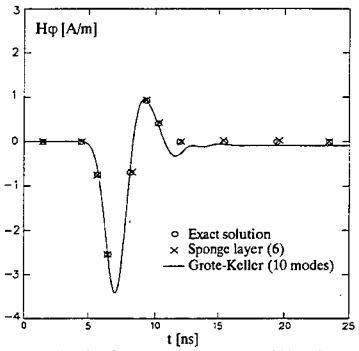


Fig. 7. Solution for the H_φ, computed at grid location p_1.

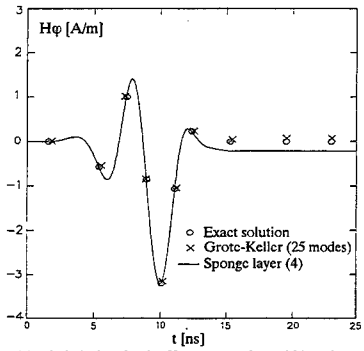


Fig. 8. Solution for the H_φ, computed at grid location p_2.

order to test the efficiency of the ABC's in a realistic scenario. The results of this implementation were identical to the ones presented in Fig. 7 and 8.

V. CONCLUSIONS

In this paper an investigation of the absorption performance of the recently presented Grote-Keller boundary condition versus that of the unsplit PML has been performed using the FDTD technique as the interior scheme in spherical coordinates. We found that both ABC's offer a spectacular reduction in the error due to the artificial domain truncation. Also, both ABC's can be brought very close to the scatterer, thus achieving considerable savings in computational resources. We will give elsewhere more results from extensive numerical tests and computational cost comparisons between the unsplit PML and the exact ABC in all three commonly used coordinate systems.

REFERENCES

[1] M. J. Grote and J. B. Keller, "Exact nonreflecting boundary conditions for the time dependent wave equation," *SIAM J. Appl. Math.*, vol. 55, no. 2, pp. 280-297, 1995.

[2] M. J. Grote and J. B. Keller, "Nonreflecting boundary conditions for time-dependent scattering," *J. Comput. Phys.*, vol. 127, pp. 52-65, 1996.

[3] M. J. Grote and J. B. Keller, "Nonreflecting boundary conditions for Maxwell's equations," *J. Comput. Phys.*, to appear, 1997.

[4] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.*, vol. 114, pp. 185-200, 1994.

[5] W. C. Chew and W. H. Weedon, "A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates," *Microwave Opt. Techol. Lett.*, vol. 7, no. 13, pp. 599-604, 1994.

[6] W. C. Chew and J. M. Jin, "Perfectly matched layers in the discretized space: An analysis and optimization," *Electromagn.*, vol.16, pp. 325-340, 1996.

[7] S. Abarbanel and D. Gottlieb, "A mathematical analysis of the PML method," *J. Comput. Phys.*, vol. 134, pp. 357-363, 1997.

[8] F. L. Teixeira and W. C. Chew, "PML-FDTD in cylindrical and Spherical Grids," *IEEE Microwave Guided Wave Lett.*, vol. 7, no. 9, pp. 285-287, 1997.

[9] F. Collino and P. Monk, "The perfectly matched layer in curvilinear coordinates," *SIAM J. Scientific Computing*, to appear, 1997.

[10] L. Zhao and A. C. Cangellaris, "GT-PML: Generalized theory of perfectly matched layers and its application to the reflectionless truncation of finite-difference time-domain grids," *IEEE Trans. Microwave Theory Tech.*, vol. 44, no. 12, pp. 2555-2563, 1996.

[11] J. Maloney, M. Kesler, and G. Smith, "Generalization of PML to cylindrical geometries," *in Proc. 13th Annu. Rev. of Prog. Appl. Comp. Electromagn. (ACES)*, Monterey, CA, vol. 2, pp. 900-908, 1997.

[12] M. Kuzuoglu and R. Mittra, "Investigation of nonplanar perfectly matched absorbers for finite-element mesh truncation," *IEEE Trans. Antennas Propag.*, vol. 45, no. 3, pp. 474-486, 1997.

[13] J. A. Roden 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 Opt. Techol. Lett.*, vol. 14, no. 2, pp. 71-75, 1997.

[14] D. M. Kingsland, J. Gong, J. L. Volakis, and J.-F. Lee, "Perfectly of anisotropic artificial absorber for truncating finite-element meshes," *IEEE Trans. Antennas Propag.*, vol. 44, no. 7, pp. 975-982, 1996.

[15] P. G. Petropoulos, "Reflectionless sponge layers as absorbing boundary conditions for the numerical solution of the 3-D Maxwell equations," submitted for publication in the *IEEE Trans. Antennas Propag.*

[16] A. F. Peterson, "Absorbing boundary conditions for vector wave equation," *Microwave Opt. Techol. Lett.*, vol. 1, pp. 62-64, 1988.

[17] A. Taflove, *Computational Electrodynamics. The Finite-Difference Time-Domain Method.* Artech House Inc., Norwood MA, 1995.

[18] K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*, CRC Press, Boca Raton FL, 1993.

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

[20] M. Fusco, "FDTD algorithm in curvilinear coordinates," *IEEE Trans. Antennas Propag.*, vol. 38, no. 1, pp. 76-89, 1990.

A SYSTEMATIC STUDY OF THREE PML ABSORBING BOUNDARY CONDITIONS THROUGH A UNIFIED FORMULATION IN CYLINDRICAL COORDINATES

Jiang-Qi He and Qing-Huo Liu

Klipsch School of Electrical and Computer Engineering
New Mexico State University
Las Cruces, New Mexico 88003

Abstract

By using a unified formulation, we compare three perfectly matched layer (PML) absorbing boundary conditions (ABC) in two-dimensional polar coordinates. An improved scheme is proposed to save the number of unknown field variables and computation time. Two-dimensional polar FDTD algorithms are developed to compare the effectiveness and efficiency of these methods. Excellent agreement is found between numerical results and analytical solutions. The formulation is then extended to conductive media in full three-dimensional cylindrical coordinates. We have developed a 3-D nonuniform grid FDTD algorithm using one of these formulations, the quasi-PML formulation, in cylindrical coordinates. Applications of the 3-D program are demonstrated for borehole radar probing.

I. Introduction

The perfectly matched layer (PML) was first introduced by Berenger as a material absorbing boundary condition (ABC) for electromagnetic waves [1]. Because of its extremely low reflections at the computational edge, the PML ABC has enjoyed widespread applications in numerical solutions of multidimensional problems in computational electromagnetics (e.g., [2–7]), elastic and acoustic wave propagation [8–12].

So far, most PML work is focused on Cartesian coordinates. Although there are studies on PML for nonorthogonal grids, notably [13–15], most previous schemes do not admit cylindrical harmonics as the eigensolutions of the modified Maxwell's equations, and hence can give rise to substantial reflections. Only recently, several formulations of PML ABCs have been implemented in cylindrical coordinates [16–21].

In this paper, we use a set of unified equations to introduce three different formulations for PML in cylindrical coordinates, i.e., quasi-PML in [16, 19], complex coordinate system as a generalized absorbing boundary condition [17, 20], and the polar PML presented in [18]. Based on the unified formulation, we propose an improved scheme to save computer memory and computation time. Then, we compare the numerical result of each approach with analytical solutions. For the sake of convenience, these three formulations will be respectively referred to as the QPML, CPML, and PPML in the following discussions.

II. Formulations

Since the formulation of PML in z direction with cylindrical coordinates is the same as that in Cartesian coordinates, for simplicity we first consider Maxwell's equations in two-dimensional polar coordinates for the TE_z

case. The formulation for the TM_z case can be easily derived by using duality.

In the frequency domain, the source-free Maxwell's equations for TE_z case in polar coordinates are

$$-i\omega\epsilon E_r = \frac{1}{r}\frac{\partial H_z}{\partial \theta}, \tag{1a}$$

$$-i\omega\epsilon E_\theta = -\frac{\partial H_z}{\partial r}, \tag{1b}$$

$$i\omega\mu H_z = \frac{1}{r}\frac{\partial (rE_\theta)}{\partial r} - \frac{1}{r}\frac{\partial E_r}{\partial \theta}, \tag{1c}$$

where the time convention $e^{-i\omega t}$ is assumed. Based on this set of equations, one can derive perfectly matched layers. Although several PML formulations have been proposed for cylindrical coordinates [16-18, 21], here we will discuss only on three different PML formulations, as it is shown that the anisotropic PML [21] is equivalent to the complex coordinate formulation [22].

A. A Unified Form for the Three PML Formulations

Here we will adopt the concept of complex coordinates in [17] to present the three PML formulations for polar coordinates. We use the complex coordinate stretching variables e_r and e_θ [2] such that $\frac{\partial}{\partial r} \to \frac{1}{e_r}\frac{\partial}{\partial r}$, $\frac{\partial}{\partial \theta} \to \frac{1}{e_\theta}\frac{\partial}{\partial \theta}$. In general,

$$e_r = a_r + i\omega_r/\omega, \tag{2}$$

while e_θ is different for the three formulations. Then, Maxwell's equations (1) are modified as

$$i\omega\epsilon E_r = -\frac{1}{re_\theta}\frac{\partial H_z}{\partial \theta}, \tag{3a}$$

$$i\omega\epsilon E_\theta = \frac{1}{e_r}\frac{\partial H_z}{\partial r}, \tag{3b}$$

$$-i\omega\mu H_z = -\frac{1}{\tilde{r}e_r}\frac{\partial (\tilde{r}E_\theta)}{\partial r} + \frac{1}{re_\theta}\frac{\partial E_r}{\partial \theta}, \tag{3c}$$

where $\tilde{r} = \tilde{r}(r)$ is in general a complex function which distinguishes the quasi-PML and PML formulations.

(a) PML scheme using complex coordinates (CPML)

If we choose $e_\theta = \tilde{r}/r$, $e_r = a_r(r) + i\omega_r(r)/\omega$, and the complex radial coordinate

$$\tilde{r} = \int_0^r e_r(r')dr' = \int_0^r \left[a_r(r') + i\frac{\omega_r(r')}{\omega}\right]dr' = A_r(r) + i\frac{\Omega_r(r)}{\omega}, \tag{4}$$

equations (3a)–(3c), after splitting $H_z = H_{zr} + H_{z\theta}$, can be cast in the following form:

$$i\omega\tilde{r}\epsilon E_r = -\frac{\partial (H_{zr} + H_{z\theta})}{\partial \theta}, \tag{5a}$$

$$i\omega e_r\epsilon E_\theta = \frac{\partial (H_{zr} + H_{z\theta})}{\partial r}, \tag{5b}$$

$$i\omega e_r\mu \tilde{H}_{zr} = \frac{\partial (\tilde{r}E_\theta)}{\partial r}, \tag{5c}$$

$$i\omega\tilde{r}H_{zr} = i\omega\tilde{H}_{zr}, \tag{5d}$$

$$i\omega\tilde{r}\mu H_{z\theta} = -\frac{\partial E_r}{\partial\theta}. \tag{5e}$$

These are the split equations for the CPML formulations in [20]. This set of equations are appropriate for time-stepping after converted into time domain, as shown in [20]. However, the extra field variables $\tilde{H}_{zr}$ and $\tilde{r}E_\theta$ have to be introduced, making the total number of unknown field variables to increase from 3 to 6 [20]. Furthermore, an extra time-stepping equation (5d) is needed.

(b) Quasi-PML scheme (QPML)

Note that in the unified formulas (3a)–(3c), if we choose $e_\theta = e_r = a_r(r) + i\omega_r(r)/\omega$, and $\tilde{r}(r) = r$, we can rewrite the split equations as

$$-i\omega e_r\epsilon E_r = \frac{1}{r}\frac{\partial H_z}{\partial\theta}, \tag{6a}$$

$$-i\omega e_r\epsilon E_\theta = -\frac{\partial H_z}{\partial r}, \tag{6b}$$

$$i\omega e_r\mu H_z = \frac{1}{r}\frac{\partial(rE_\theta)}{\partial r} - \frac{1}{r}\frac{\partial E_r}{\partial\theta}. \tag{6c}$$

Thus it turns out in 2-D polar coordinates there is no need to split the field in the quasi-PML. Hence, the total number of unknown field variables is 3. It should be noted that, in 2-D polar coordinates, this medium actually reduces to a medium whose relative electric and magnetic conductivities are the same. For 3-D cylindrical problems, however, the splitting is necessary in order to match the interfaces in r and z directions simultaneously.

From (6), the time-domain equations modified from (1) become

$$a_r\epsilon\frac{\partial E_r}{\partial t} = \frac{1}{r}\frac{\partial H_z}{\partial\theta} - \epsilon\omega_r(r)E_r, \tag{7a}$$

$$a_r\epsilon\frac{\partial E_\theta}{\partial t} = -\frac{\partial H_z}{\partial r} - \epsilon\omega_r(r)E_r, \tag{7b}$$

$$a_r\mu\frac{\partial H_z}{\partial t} = -\frac{1}{r}\frac{\partial(rE_\theta)}{\partial r} + \frac{1}{r}\frac{\partial E_r}{\partial\theta} - \mu\omega_r(r)H_z. \tag{7c}$$

Although it can be shown theoretically that this PML is not perfectly matched (thus the name quasi-PML) [16], practically it provides a satisfactory ABC.

(c) Improved CPML scheme

Noting that the computer storage requirement for the original CPML is quite high (6 field variables), we seek an improved formulation to split equation (3). This can be easily done by rewriting (3c) as

$$i\omega\mu(H_{zr} + H_{z\theta}) = \frac{E_\theta}{\tilde{r}e_r}\frac{\partial\tilde{r}}{\partial r} + \frac{1}{e_r}\frac{\partial E_\theta}{\partial r} - \frac{1}{re_\theta}\frac{\partial E_r}{\partial\theta}. \tag{8}$$

Therefore, with $e_\theta = \tilde{r}/r$, and $\tilde{r}$ given by (4), the new split equations to replace (5a)–(5e) are

$$i\omega\tilde{r}\epsilon E_r = -\frac{\partial(H_{zr} + H_{z\theta})}{\partial\theta}, \tag{9a}$$

$$i\omega e_r \epsilon E_\theta = \frac{\partial (H_{zr} + H_{z\theta})}{\partial r}, \tag{9b}$$

$$i\omega \tilde{r} \mu H_{zr} = \frac{\partial E_\theta}{\partial r}, \tag{9c}$$

$$i\omega \tilde{r} \mu H_{z\theta} = E_\theta - \frac{\partial E_r}{\partial \theta}, \tag{9d}$$

Observe that this split set of equations require only 4 unknown field variables: E_r, E_θ, H_{zr}, and $H_{z\theta}$. It saves about 1/3 computer memory and 1/5 computer time compared to the original CPML implementation in equation (5).

(d) Polar PML scheme (PPML)

Finally, by setting $a_r = 1$, $\omega_r = f'(r)$ so that $\tilde{r} = r[1 + if(r)/r\omega]$, equations (9a)–(9b) can be converted into time domain as

$$\epsilon \frac{\partial E_r}{\partial t} = \frac{1}{r} \frac{\partial (H_{zr} + H_{z\theta})}{\partial \theta} - \epsilon \frac{f(r)}{r} E_r, \tag{10a}$$

$$\epsilon \frac{\partial E_\theta}{\partial t} = -\frac{\partial (H_{zr} + H_{z\theta})}{\partial r} - \epsilon f'(r) E_r, \tag{10b}$$

$$\mu \frac{\partial H_{zr}}{\partial t} = -\frac{\partial E_\theta}{\partial r} - \mu f'(r) H_{z\theta}, \tag{10c}$$

$$\mu \frac{\partial H_{z\theta}}{\partial t} = -\frac{E_\theta}{r} + \frac{1}{r} \frac{\partial E_r}{\partial \theta} - \mu \frac{f(r)}{r} H_{zr}, \tag{10d}$$

which are exactly the polar PML scheme presented in [18]. The number of unknown field variables is the same as equation (9). The effectiveness of the QPML, CPML, and PPML will be compared in the numerical results in III.

B. 3-D Cylindrical Coordinates

For 3-D cylindrical coordinates (r, θ, z), the extension is straightforward since the z direction is the same as for the Cartesian coordinates. Furthermore, the extension to conductive media can follow the same procedures as in [4]. Therefore, based on the improved CPML formulation, the time-domain split equations for conductive media can be derived as

$$A_r \epsilon \frac{\partial E_r^{(\theta)}}{\partial t} + (\Omega_r \epsilon + A_r \sigma) E_r^{(\theta)} + \Omega_r \sigma \int_{-\infty}^{t} E_r^{(\theta)}(\tau) d\tau = \frac{\partial H_z}{\partial \theta} - J_r^{(\theta)}, \tag{11a}$$

$$a_z \epsilon \frac{\partial E_r^{(z)}}{\partial t} + (\omega_z \epsilon + a_z \sigma) E_r^{(z)} + \omega_z \sigma \int_{-\infty}^{t} E_r^{(z)}(\tau) d\tau = -\frac{\partial H_\theta}{\partial z} - J_r^{(z)}, \tag{11b}$$

$$a_r \epsilon \frac{\partial E_\theta^{(r)}}{\partial t} + (\omega_r \epsilon + a_r \sigma) E_\theta^{(r)} + \omega_r \sigma \int_{-\infty}^{t} E_\theta^{(r)}(\tau) d\tau = -\frac{\partial H_z}{\partial r} - J_\theta^{(r)}, \tag{11c}$$

$$a_z \epsilon \frac{\partial E_\theta^{(z)}}{\partial t} + (\omega_z \epsilon + a_z \sigma) E_\theta^{(z)} + \omega_z \sigma \int_{-\infty}^{t} E_\theta^{(z)}(\tau) d\tau = \frac{\partial H_r}{\partial z} - J_\theta^{(z)}, \tag{11d}$$

$$a_r \epsilon \frac{\partial E_z^{(r)}}{\partial t} + (\omega_r + a_r \sigma) E_z^{(r)} + \omega_r \sigma \int_{-\infty}^{t} E_z^{(r)}(\tau) d\tau = \frac{\partial H_\theta}{\partial r} - J_z^{(r)}, \tag{11e}$$

$$A_r \epsilon \frac{\partial E_z^{(\theta)}}{\partial t} + (\Omega_r \epsilon + A_r \sigma) E_z^{(\theta)} + \Omega_r \sigma \int_{-\infty}^{t} E_z^{(\theta)}(\tau) d\tau = H_\theta - \frac{\partial H_r}{\partial \theta} - J_z^{(\theta)}. \tag{11f}$$

The other set of equations for updating **H** can be obtained by duality. The equations for the other two PML formulations can be derived similarly.

III. Numerical Results

To show the numerical results of three different PML formulations in polar coordinates, for simplicity, we simulate a line source in free space with the derivative of a Blackman-Harris window time function at a center frequency $f_c = 300$ MHz. The line source is located at $(r, \theta) = (15, 64)$ cells in a cylinder whose computational domain is $N_r \times N_\theta = 80 \times 256$. The radius of the cylinder is 3.2 m which includes 10 PML cells in the radial direction. Fifteen receivers are set uniformly around a semi-circle 20 cells away from the origin, and are 8 cells apart in θ direction. The first receiver is located at $(20, 8)$ in the grid.

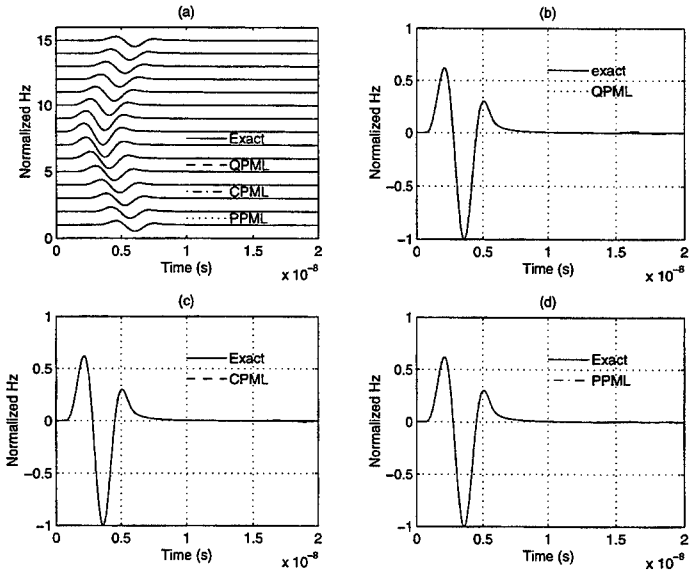


Figure 1. Comparison between analytical and numerical results for three PML schemes in polar coordinates. (a) Array waveforms. Analytical and numerical results at the 8th receiver for (b) QPML, (c) CPML, and (d) PPML.

Figure 1(a) shows the excellent agreement between analytical solution and the numerical result of all three schemes. The comparison is magnified in Figs. 1(b), 1(c), and 1(d) for QPML, CPML. and PPML at the 8th receiver. The reflection is about 1.1%, 0.95%, and 0.91%, respectively. Note that for a fair comparison, we have chosen $a_r = 1$ in the quasi-PML case even though the code allows a profile for a_r. This reflection can be reduced substantially by adjusting a_r.

We model a 3-D case to illustrate the application of the nonuniform cylindrical FDTD using quasi-PML ABC. Figure 2(a) shows the xz cross section of a problem in borehole radar detection of vertical and horizontal fractures. The background medium is conductive with $\epsilon_r = 2.0$, $\mu_r = 1.0$, and $\sigma = 0.001$ S/m. The borehole is located in the middle of the cylinder with radius 16 cm and $\epsilon_r = 4.0$, $\mu_r = 1.0$, and $\sigma = 0.01$ S/m. The horizontal fractures has

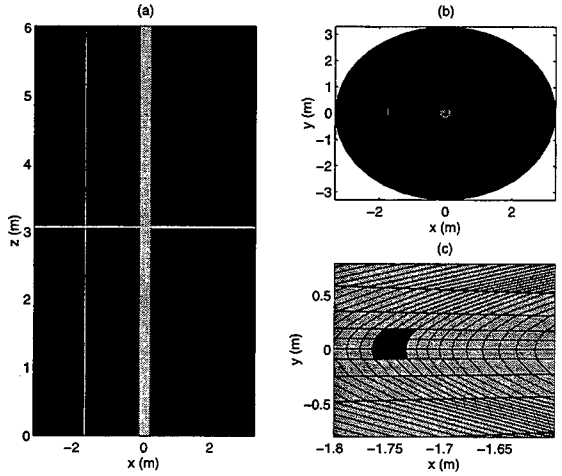


Figure 2. Geometry of a borehole radar detection case using Quasi-PML

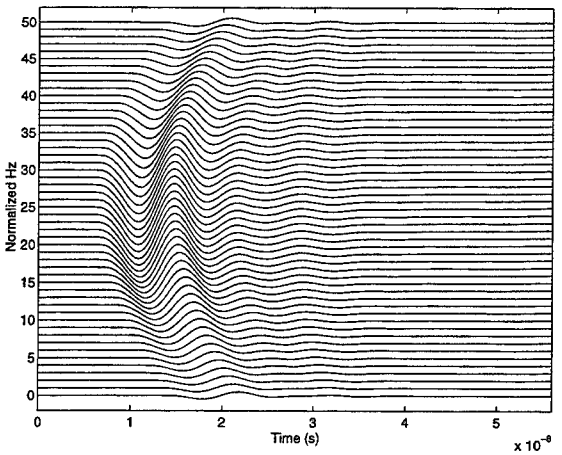


Figure 3. The scattering waveforms of the two fractures in Figure 2

a thickness of 3 cm with $\epsilon_r = 8.0$, $\mu_r = 1.0$, and $\sigma = 0.1$ S/m. The vertical fracture parallel to z axis is about 1.73 m away from the axis and has a thickness of 3 cm and spans 9° in θ direction with the same ϵ_r, μ_r, and σ as the horizontal fracture. Figure 2(b) shows the xy cross section. Figure 2(c) magnifies the cross section in Figure 2(b) to show the grid around the vertical fracture. A magnetic dipole point source is located along the borehole axis and is 3.8 m from the bottom boundary, and a receiver array is located also along the z axis. If the scattering field of the two fractures was calculated by a uniform FDTD, a grid about $N_r \times N_\theta \times N_z = 600 \times 120 \times 350$ should have been used in order to accommodate the small fractures in different directions. In this work, we adopt a nonuniform

grid with $N_r \times N_\theta \times N_z = 140 \times 40 \times 100$, saving about 45 times CPU and memory. The scattering field is shown in Figure 3 with both fractures.

IV. Conclusions

Based on a unified presentation, three different formulations of PML in cylindrical coordinates are compared. Numerical results from 2-D polar FDTD programs based on these methods agree very well with analytical solution, even though the quasi-PML is not a perfectly matched medium. We propose an improved scheme for PML based on the generalized complex coordinates method without introducing extra variables and stepping equations, saving above 1/3 of the computer memory and 1/5 computation time. A 3-D nonuniform FDTD method is developed using the quasi-PML formulation.

Acknowledgment

This work was supported by Environmental Protection Agency under a PECASE grant CR-825-225-010, and by National Science Foundation under a CAREER grant ECS-9702195.

References

[1] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Computational Physics*, vol. 114, pp. 185–200, 1994.

[2] W. C. Chew, and W. H. Weedon, "A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates," *Microwave Opt. Tech. Lett.*, vol. 7, pp. 599–604, 1994.

[3] D. S. Katz, E. T. Thiele, and A. Taflove, "A perfectly matched layer for the absorbing of electromagnetic waves," *IEEE Microwave Guided Wave Lett.*, vol. MGW-4, pp. 268-270, Aug. 1994.

[4] Q. H. Liu, "An FDTD algorithm with perfectly matched layers for conductive media," *Microwave Opt. Tech. Lett.*, vol. 14, pp. 134–137, 1997.

[5] Q. H. Liu, "A spectral-domain method with perfectly matched layers for time-domain solutions of Maxwell's equations," *1996 URSI Meeting*, Baltimore, MD, July 1996.

[6] J. Fang and Z. Wu, "Generalized perfectly matched layer- an extension of Berenger's perfectly matched layer boundary condition," *IEEE Microwave Guided Waves Lett.*, vol. 5, pp. 451-453, Dec. 1995.

[7] Q. H. Liu, "The PSTD algorithm: a time-domain method requiring only two cells per wavelength," *Microwave Opt. Tech. Lett.*, vol. 15, no. 3, pp. 158–165, June 1997.

[8] W. C. Chew, and Q. H. Liu, "Perfectly matched layers for elastodynamics: A new absorbing boundary condition," *Schlumberger-Doll Research Technical Report*, August 1995.

[9] W. C. Chew, and Q. H. Liu, "Using perfectly matched layers for elastodynamics," *Proc. IEEE Antennas Propagat. Soc. Intl. Symp.* (IEEE, New York), vol. 1, pp. 366–369, 1996.

[10] W. C. Chew, and Q. H. Liu, "Perfectly matched layers for elastodynamics: A new absorbing boundary condition," *J. Computational Acoust.*, vol. 4, pp. 72–79, 1996.

[11] F. D. Hastings, J. B. Schneider, and S. L. Broschat, "Application of the perfectly matched layers (PML) absorbing boundary condition to elastic wave propagation," *J. Acoust. Soc. Am.* , vol. 100, pp. 3061–3069, 1996.

[12] Q. H. Liu, and J. P. Tao, "The perfectly matched layers for acoustic waves in absorptive media," *J. Acoust.*

Soc. Am., vol. 102(4), pp. 2072–2082, October 1997.

[13] E. A. Navarro, C. Wu, P. Chung, and J. Litva, "Application of PML superabsorbing boundary condition to non-orthogonal FDTD method ," *Electron. Lett.*, vol. 30, no. 20, pp. 1654-1656, 1994.

[14] W. Chen, E. A. Navarro, P. Y. Chung, and J. Litva, "Modeling of waveguide structures using the nonorthogonal FDTD method with a PML absorbing boundary," *Microwave Opt. Tech. Lett.*, vol. 8, no. 4, pp. 226–228, 1995.

[15] J.A.Roden and S.D.Gedney, "Efficient implementation of the uniaxial-based PML media in three-dimensional nonorthogonal coordinates with the use of FDTD technique ," *Microwave Opt. Tech. Lett.*, vol. 14, no. 2, pp. 71–75, February 1997.

[16] Q. H. Liu and J. Q. He, "Quasi-PML for waves in cylindrical coordinates," *New Mexico State University Technical Report*, July 1997.

[17] W. C. Chew, J. M. Jin, and E. Michelssen, "Complex coordinate system as a generalized absorbing boundary condition," *Intl. IEEE AP-S Symposium Digest*, Montreal, Canada, July 1997.

[18] B. Yang, D. Gottlieb, and J. S. Hesthaven, "Spectral simulations of electromagnetic wave scattering," *J. Comput. Phys.*, vol. 134, pp. 216–230, 1997.

[19] J. Q. He and Q. H. Liu, "An FDTD method with nonuniform cylindrical grid for inhomogeneous conductive media ," *1997 URSI Meeting* , Montreal, Canada, July 1997.

[20] F. L. Teixeira and W. C. Chew, "PML-FDTD in cylindrical and spherical grids," *IEEE Microwave Guided Wave Lett.*, vol. 7, pp. 285-287, sept. 1997.

[21] J. Maloney, M. Kesler, and G. Smith, "Generalization of PML to cylindrical geometries," in *Proc.* 13^{th} *Ann. Rev. Prog. Appl. Comp. Electromag.* , Monterey, CA, Mar. 17-21, 1997.

[22] F. L. Teixeira and W. C. Chew, "Systematic derivation of anisotropic PML absorbing media in cylindrical and spherical coordinates," *IEEE Microwave Guided Wave Lett.*, vol. 7, pp. 371-373, Nov. 1997.

Preconditioned Generalized Minimal Residual (GMRES) Solver for Domains Truncated by Perfectly Matched Layer (PML) Absorbers

Youssry Y. Botros and John L. Volakis
Radiation Laboratory
Department of Electrical Engineering
and Computer Science,
The University of Michigan
Ann Arbor, MI 48109-2212
Email: ybotros@umich.edu and volakis@umich.edu

Abstract

Because of their superior absorption characteristics, the Perfectly Matched Layer (PML) absorbers are used in truncating finite element domains. However, their implementation is equivalent to imposing active elements inside the main mesh. Consequently, the condition number of the resulting systems deteriorates. In this work, an efficient preconditioned generalized minimal residual (GMRES) iterative solver is developed and applied to systems truncated by PML absorbers. This iterative scheme is implemented and tested for different cases representing actual structures.

1 Introduction

The PML layer introduced by Sacks et al [1] is an effective means for truncating finite element domains associated with microwave circuits and packaged networks. In addition to its outstanding absorption performance characteristics, the PML layer is extremely simple to implement without a need to deal with higher order derivatives as is the case with absorbing boundary conditions [2-5]. However, PMLs yield finite element systems which suffer from poor conditioning, thus, deteriorating the convergence of iterative solvers. More specifically, traditional iterative algorithms such as the conjugate gradient and biconjugate gradient methods are very slow to converge and often fail altogether to yield a solution.

In this paper we propose and apply an iterative solver based on the generalized minimal residual method (GMRES) for solving sparse finite element systems. A preconditioning scheme is also proposed and integrated with the flexible-GMRES algorithm. The preconditioner is based on the approximate inverse preconditioning (AIPC) scheme and is typically applied only to those systems which exhibit poor convergence characteristics at the initial iteration steps. It is shown that generally the GMRES algorithms converge very quickly (within a few iterations) provided that a sufficient number of expansion vectors are chosen at the start of the iteration process. Preconditioning also plays a crucial role in the speed and robustness of the solution and in the paper we compare AIPC with other less robust preconditioners. Applications to actual microwave structures are given to provide a measure for the accuracy of the solver and its convergence characteristics when dealing with electromagnetic systems associated with packaging applications.

2 PML Parameters

As shown in Figure 1, the PML layer consists of a metal backed dielectric layer where the medium of the layer has the following permittivity and permeability tensors

$$\overline{\overline{\mu}}_r = \overline{\overline{\epsilon}}_r = \begin{pmatrix} a_2 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & c_2 \end{pmatrix} \tag{1}$$

With the choice $a_2 = b_2 = 1/c_2 = \alpha - j\beta$, where α and β are the phase and attenuation factors repectively, it has been shown that waves impinging at the air dielectric interface are completely non-reflecting for all incidence angles ($0 < \phi < 90$). Since $\frac{\beta}{k_o}$ (where k_o is the free space wave number), is a non-zero attenuation constant, once in the dielectric, the wave decays to small values rather rapidly. Thus the metal backing has a very small effect or nearly no effect on the truncation of the domain. Nevertheless, it simplifies the implementation of finite element simulations.

Throughout the paper and our study, our goal has been the evaluation of the PML performance not only in terms of its absorption effectiveness, but also on its effect on system convergence. There are various ways and parameter choices which can be used for the evaluation of the system convergence. For our case we have used the ratio

$$r = \frac{Number\ of\ Iterations\ before\ Convergence}{FEM\ system\ size} \tag{2}$$

Clearly for small values of r the system is highly convergent. Systems which are associated with values of r that approach unity are considered as poorly conditioned.

2.1 Effect of PML Parameters on Convergence

Previous studies [6]-[11] focused on the optimization and understanding of the PML parameters with respect to the absorption characteristics of the layer. More specifically, in [6], [11] and [13] curves were given for an optimum selection of the attenuation coefficient versus the numerical discretization rates and the desired absorption rate. Basically, it was demonstrated that although the PML provides for a theoretically non-reflecting layer, its numerical counterpart has some given reflectivity which can be controlled by a proper choice of β, layer thickness and discretization rate. It was pointed out that typically a value of $\beta = 1$ provides a good choice for sufficiently thick layers. However, so far the effect of the phase constant $\frac{\alpha}{k_o}$ has not been addressed, although numerical experiments indicated that α has a much smaller effect on the performance of the PML layer. Nevertheless, our initial studies on convergence indicated that α has a noticeable effect on the system convergence rates. Therefore, we begun this study by examining the effect of α and β on convergence for the microstrip line truncation shown in Figure 1(b). The curves shown in Figures 1(c) and 1(d) indicate that although the absorption of the PML for $\alpha < 2$ and $\beta > 1$ is good, the corresponding convergence curves provide a different story. For large values of α with $\beta = 1$, it is seen that the convergence is optimized. However, better convergence is obtained when β is small. From these curves, considerations on both absorption and convergence dictate that a good choice is $\alpha = \beta = 1$.

3 GMRES Solver

We look at the performance of the different solvers by implementing and testing three of these iterative solvers. One was the BiConjugate (BCG) gradient method which is easy to implement and has low CPU and memory costs. However, it lacks the robustness and does not guarantee convergence, [12] and [16]. For badly conditioned systems (as in the PML case), it may not converge at all. The Quasi Minimal Residue (QMR) solver has better convergence features and lower breakdown possibilities [12]. Nevertheless, for the same system, both BCG and QMR converge nearly in the

same number of iterations but typically QMR has better error history. The Generalized Minimal Residual (GMRES) solver is the most robust solver since it guarantees convergence even for poorly conditioned systems. Figure 2 displays the superior convergence characteristics of the GMRES solver over the BCG and QMR for the microstrip line problem shown in Figure 1(b) (terminated by the PML). This type of convergence is typical for most examined FEM systems and therefore GMRES [13] was our choice solver.

In implementing GMRES, one can not ignore the important role of the parameter m which refers to the search vectors used for an estimate of the solution. Although m is arbitrary, it is the main parameter that controls convergence. In general, larger values of m lead to smaller residuals and hence faster convergence. However, CPU and memory costs are directly related to m. For all types of GMRES solvers, the memory cost is $O(mN)$ and the CPU cost is $O(m^2N)$, where N refers to the number of unknowns. Therefore, it is essential to have an estimate for m before executing the GMRES iterations. If this number is lower than the threshold or minimum value, convergence will be extremely slow and may not be achieved at all. On the other hand, if m is too high, storage and CPU are wasted. The optimal value of m is directly related to two main factors, the condition and the size of the matrix. From the studies [10], [13], [14] and [15], we concluded that the system condition has a strong impact on the optimal value of m. One way to reduce the solver dependence on m is by employing a good preconditioner. Therefore, our goal is to apply a strong preconditioner so that the dependence on m is reduced and this will lead to more stable and predictable convergence scheme.

4 Preconditioners

Preconditioners are usually applied to improve the system condition and hence achieve faster convergence. They vary in complexity from the simple diagonal preconditioner (DPC) to the complicated approximate inverse preconditioner (AIPC).

4.1 Diagonal Preconditioner DPC

The Diagonal Preconditioner (DPC) is the simplest of all. It is simply implemented by dividing each row with its largest entry (diagonal element). Thus, it can be implemented with almost no CPU or memory costs. Also, it typically delivers a speed up of 30% to 60%. As shown in Figure 2(b), the DPC achieves substantial convergence improvements without memory or CPU costs. This was already pointed out in previous studies, for example [17].

4.2 Approximate Inverse Preconditioner AIPC

For the general situation, where the FEM matrix is indefinite, standard preconditioning techniques may fail due to code breakdown. Also, when this matrix is not diagonally dominant, most preconditioners (such as diagonal and ILU) may not be effective. The proper preconditioner should have some basic features. It should have low computational and memory costs and should retain robustness even if the FEM matrix is not diagonally dominant. The Approximate Inverse Preconditioning Scheme (AIPC) can achieve these features. The idea behind AIPC is to find a sparse matrix M which minimizes the Frobenius norm of the residual of the matrix

$$R = I - AM \tag{3}$$

where I is the identity matrix, M is the AIPC and A is the FEM matrix in the original system $Ax = b$. According to [12] and [16], this minimization can be achieved in several ways. One of the most efficient methods is the *Column Oriented Algorithms* which minimizes the norm of the individual columns of R

$$R_j = I_j - AM_j \tag{4}$$

where the subscript j denotes the j^{th} column of the corresponding matrix. That is, M_j is found by iteratively minimizing R_j for each j^{th} column. One has the choice of setting higher levels for the residual to speed up the minimization and come up with different levels of accuracy in finding M.

5 GMRES-AIPC Solver

The combined use of the AIPC with the GMRES solver proved quite effective for poorly conditioned FEM systems. To show the significance of the suggested scheme, we implemented and tested an FEM system (approximately 10000 unknowns). The number of search vectors (basis functions) m was scanned from 10-100 and the convergence was examined by recording the total CPU time in seconds for the GMRES with no preconditioning, DPC and AIPC. The results of this study are displayed in Figure 2(c). From this graph, we can conclude the following:

- For all values of m, the total CPU time (code execution time) is less when the AIPC is applied (although the per iteration CPU was of course high). This is due to the fact that this type of preconditioners improves significantly the condition number of the FEM system and thus obtain substantial CPU improvements.
- The CPU dependence on m is dramatically reduced when the AIPC is invoked. This solves the problem of specifying the correct or optimal m to the solver before starting the GMRES iterations. As displayed, wide ranges of m have narrow CPU variations.

6 Microwave Circuit Applications

After presenting the guidelines for the PML implementation, preconditioners and solvers, we proceed to specific applications. In all subsequent examples, we implement the PML absorber with $\alpha = \beta = 1$ and 12-15 samples per wavelength were used for discretization. Also, the GMRES solver with AIPC is applied to solve the resulting linear systems with m in the range of 10 to 40.

6.1 Microstrip Lines and Feed Probes

For the microstrip configuration shown in Figure 1(b), we examined the fields under the microstrip line as the number of feeding probes was varied. A deembedding scheme based on a transmission line analogy is applied to extract the scattering parameters from the FEM simulation [10]. As shown in Figure 2(d), the field under the microstrip line increases with the number of probes. This result is expected and demonstrates that the PML performance does not affect the feeding mechanism since the absorber treats all the feed mechanisms impartially.

6.2 Spiral Inductor with an Air Bridge

We modeled the geometry shown in Figure 3(a) using our FEM simulator. Our goal for this example was to validate the preconditioned GMRES solver using this benchmark geometry. As observed in Figure 3(a), this spiral has fine details to be considered. The results of the scattering parameter S_{11} are shown in Figure 3(b) and as seen good agreement between the measured and calculated data was obtained. To reduce the FEM matrix size, we assumed that the width of the air bridge is equal to that of the microstrip line.

References

[1] Z. S. Sacks, D. M. Kingsland, R. Lee and J. F. Lee, "A perfectly matched anisotropic absorber for use as an absorbing boundary condition," *IEEE Trans. Antennas and Propagation*, December 1994.

[2] B. Engquist and A. Majda, "Absorbing boundary conditions for the numerical simulation of waves," *Math. Comput.* Vol. 31, pp. 629–651, 1977.

[3] J. Webb and V. Kanellopoulos, " Absorbing boundary conditions for the finite element solution of the vector wave equation," *Microwave Opt. Tech. Lett.* Vol. 2, pp. 370–372, 1989.

[4] A. Chatterjee and J.L.Volakis, "Conformal absorbing boundary conditions for the vector wave equation," *Microwave Opt. Tech. Lett.* Vol. 6, pp. 886-889, 1993.

[5] B. Stupfel and R. Mittra, "Efficiency of numerical absorbing boundary conditions for finite element applications," *URSI Radio Science Meeting*, pp. 165, 1994

[6] S. Legault, T.B.A. Senior and J.L. Volakis, "Design of Planar Absorbing Layers for Domain Truncation in FEM Applications," *Electromagnetics J.*, vol 16, no. 4, July-August 1996.

[7] J. P. Berenger, "A Perfectly Matched Layer for the Absorption of Electromagnetic Waves," *J. Comp. Physics*, 114 : 185-200, 1994.

[8] W. C. Chew and W. H. Weedon, "A 3-D Perfectly Matched Medium from Modified Maxwell's Equations with Stretched Coordinates," *Microwave and Optical Technology Letters*, pp. 599–604. September, 1994.

[9] D. M. Kingsland, J. Gong, J.L. Volakis, and J. F. Lee, "Performance of an Anisotropic Artificial Absorber for Truncating Finite Element Meshes," *IEEE Trans. on Ant. and Prop.*, Dec. 1995.

[10] Y.Y. Botros, J. Gong and J.L Volakis, "Perfectly Matched Layer (PML) Absorbers for Truncating Finite Element Meshes," URSI Meeting, Montreal, Canada, July 1997.

[11] J. Gong, S. R. Legault, Y. Y. Botros, J. L. Volakis and P. Petre, "Application and Design Guidelines of the PML Absorber for Finite Element Simulations of Microwave Packages", Microwave Theory and Techniques Symp., San Francisco, June 1996.

[12] R. Barret et. al., *Templetes for the Solution of Linear Systems: Building Blocks for Iterative Solvers*, siam, 1994.

[13] Y. Y. Botros and J. L. Volakis, "A Hybrid Generalized Minimal Residual (GMRES) Algorithm for Solving FEM Systems Generated by the High Frequency Structure Simulator (HFSS) Code," Radiation Laboratory Report (No. 375458-1-T), The University of Michigan, Ann Arbor, November 1997.

[14] A. D. Brown, J. L. Volakis, L. C. Kempel and Y. Y. Botros, "Patch Antenna on Ferromagnetic Substrates," submitted for publication to *IEEE Trans. Ant. and Prop.*.

[15] Y. Y. Botros and J. L. Volakis, "An Iterative Scheme for Large Poorly Conditioned Applications," submitted for publication to *IEEE Trans. Microwave Theory and Tech.*.

[16] Y. Saad *Iterative Methods for Sparse Linear System*, PWS publishing co., 1996.

[17] A. Chatterjee, J. L. Volakis and L. C. Kempel, "Optimization issues in finite element codes for solving 3D electomagnetic problems," Int. J. Numerical Modeling: Elect. Net. Dev. and Fields., Vol. 9, pp. 335–344, Sept-Oct. 1996.

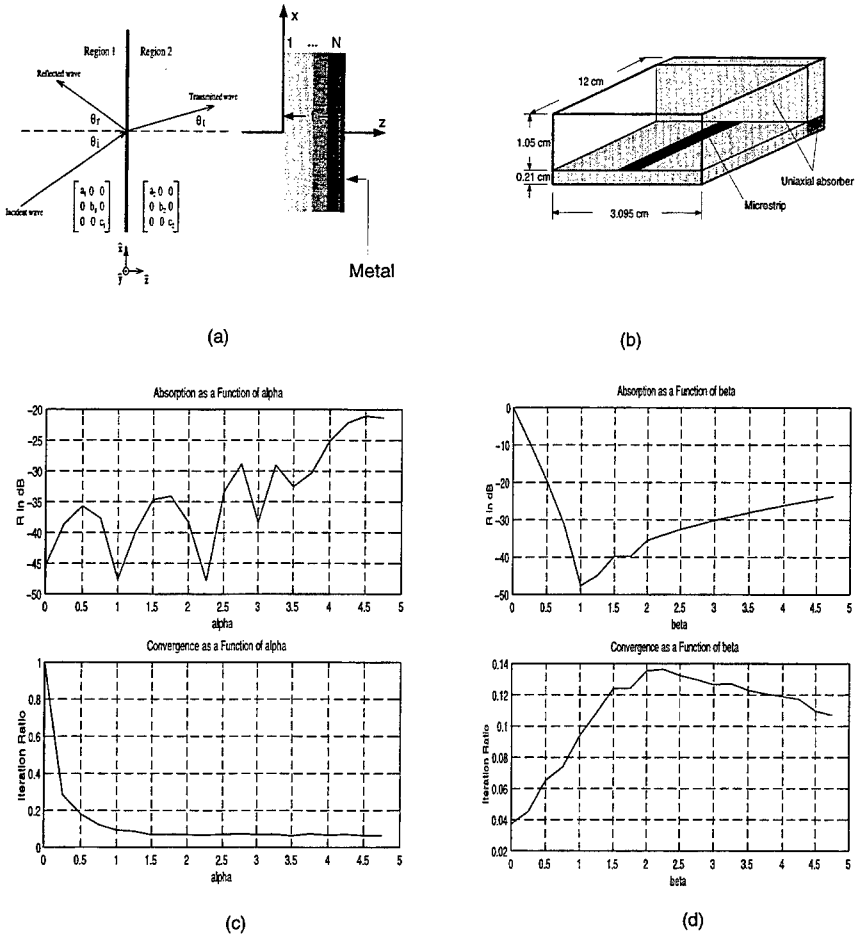


Figure 1: (a) Plane wave incidence on an interface between two diagonally anisotropic half-spaces. (b) Microstrip Line. (c) The dB absorption and Convergence Factor as Functions of α. (d) The dB absorption and Convergence Factor as Functions of β.

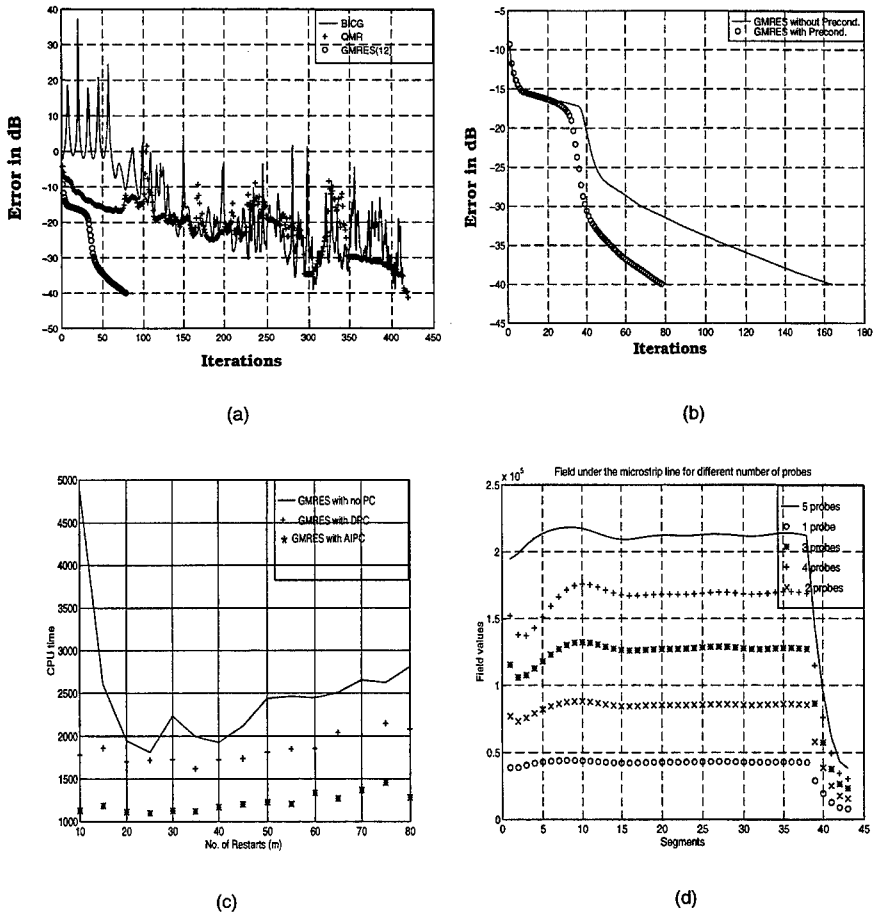
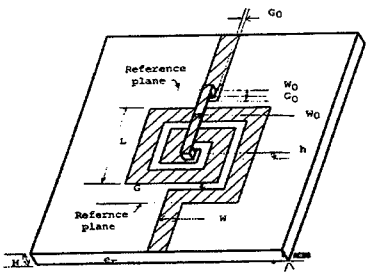
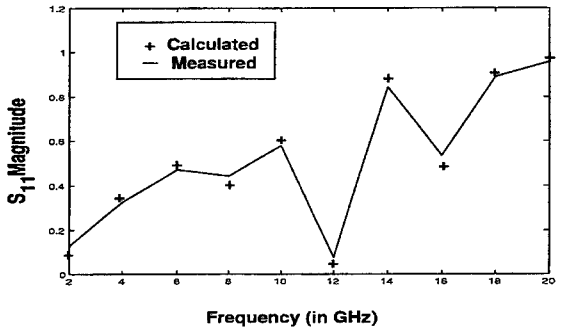


Figure 2: (a) Comparison Between the Convergence of three Different Solvers. (b)Effect of Diadgonal Preconditioning on the GMRES Performance. (c) Effect of the Preconditioning Type on GMRES Performance. (d) Field Magnitude under the Microstrip Line as a Function of the Number of Feeding Probes



DIMENSION:
L = 3.4375 mm
GO = 0.15625 mm
W_O = 0.3125 mm
W = 0.625 mm
G = 0.3125 mm
H = 0.635 mm
h = 0.3125 mm
e_r = 9.8

(a)



(b)

Figure 3: (a) Geometry of the Spiral with an Air Bridge. (b) Comparison Between the Measured and Computed data for the Spiral Antenna.

Invited Paper for the Session Organized by Professor Petropoulos

PML Implementation for the Battle-Lemarie Multiresolution Time–Domain Schemes

Emmanouil Tentzeris, Rob Robertson, Linda P.B. Katehi
Radiation Laboratory, EECS Department, University of Michigan,
Ann Arbor, MI 48109-2122, USA

I Introduction

The Multiresolution Time Domain (MRTD) Technique based on cubic-spline Battle Lemarie scaling and wavelet functions has shown successful application to a variety of microwave problems and has demonstrated unparalleled properties in terms of memory and execution time by one and two orders of magnitude respectively. This technique is used to model open and shielded propagation problems [1, 3] and non-linear optical applications [2]. In addition to time and memory, the most important advantage of this new technique is its capability to provide space and time adaptive meshing without the problems encountered by the conventional Finite Difference Time Domain(FDTD) [4] method. In this paper, an efficient non-split formulation of the PML absorber [5] for the Battle-Lemarie based MRTD scheme is presented. This formulation is validated and applied in the analysis of a two–dimensional parallel–plate waveguide geometry offering a numerical coefficient of reflection below -90dB. Additionally, examples for a three–dimensional patch antenna geometry are given.

II Derivation of the MRTD equations for the PML layer

Without loss of generality, the PML Absorber equations will be presented for a homogeneous medium for TM propagation in 2D. The Absorber formulation for TE propagation is straightforward. Assuming that the PML area is characterized by (ϵ_o, μ_o) and electric and magnetic conductivities (σ_E, σ_H), the TM equations can be written

$$\epsilon_o \frac{\partial E_x}{\partial t} + \sigma_E E_x = -\frac{\partial H_y}{\partial z} \quad (1)$$

$$\epsilon_o \frac{\partial E_z}{\partial t} + \sigma_E E_z = \frac{\partial H_y}{\partial x} \quad (2)$$

$$\mu_o \frac{\partial H_y}{\partial t} + \sigma_H H_y = \frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} \quad (3)$$

PML cells only to the z-direction are considered. Equations for PML cells in the x- and y-directions can be derived in a similar way. For each point z of the PML area, the magnetic conductivity σ^H needs to be chosen as [5]:

$$\frac{\sigma_E(z)}{\epsilon_o} = \frac{\sigma_H(z)}{\mu_o} \tag{4}$$

for a perfect absorption of the outgoing waves. A parabolic spatial distribution of $\sigma_{E,H}$,

$$\sigma_{E,H}(z) = \sigma_{E,H}^{max}(1 - \frac{z}{\delta})^p \quad \text{, with p=2} \qquad \text{for } 0 \leq z \leq \delta = \text{PML thickness} \tag{5}$$

is used in the simulations, though higher order distributions (e.g.Cubic p=3) can give similar results. The PML area is terminated with a PEC and usually has a thickness varying between 4-16 cells. The maximum value σ_E^{max} is determined by the designated reflection coefficient R at normal incidence, which is given by the relationship

$$R = e^{-\frac{2}{\epsilon_o c}\int_o^\delta \sigma_E(z)dz} = e^{-\frac{2\sigma_E^{max}\delta}{\epsilon_o c(p+1)}} \quad . \tag{6}$$

The electric and magnetic field components incorporated in these equations are expanded in a series of Battle-Lemarie scaling and wavelet functions in both x- and z-directions. For example, E_x can be represented as:

$$\begin{aligned} E_x(x,z,t) &= \sum_{k,l,m=-\infty}^{+\infty} {}_kE_{l+1/2,m}^{x,\phi\phi} h_k(t)\phi_{l+1/2}(x)\phi_m(z) \\ &+ \sum_i \sum_{k,l,m=-\infty}^{+\infty} {}_kE_{l+1/2,m}^{x,\phi\psi_i} h_k(t)\,\phi_{l+1/2}(x)\,\psi_{i,m}(z) \\ &+ \sum_i \sum_{k,l,m=-\infty}^{+\infty} {}_kE_{l+1/2,m}^{x,\psi_i\phi} h_k(t)\,\psi_{i,l+1/2}(x)\,\phi_m(z) \\ &+ \sum_{i,j} \sum_{k,l,m=-\infty}^{+\infty} {}_kE_{l+1/2,m}^{x,\psi_i\psi_j} h_k(t)\,\psi_{i,l+1/2}(x)\,\psi_{j,m}(z) \end{aligned} \tag{7}$$

where $\phi_m(x) = \phi(\frac{x}{\Delta x} - m)$ and $\psi_{i,m}(x) = \psi_i(\frac{x}{\Delta x} - m)$ represent the Battle-Lemarie scaling and i-th order resolution wavelet function respectively in space and $h_k(t)$ represent rectangular pulses in time. ${}_kE_{l,m}^{\kappa,\mu\nu}$ and ${}_{k+1/2}H_{l,m}^{\kappa,\mu\nu}$ with $\kappa = x, y, z$ and $\mu, \nu = \phi, \psi$ are the coefficients for the field expansions in terms of scaling and wavelet functions. The indices l, m and k are the discrete space and time indices related to the space and time coordinates via $x = l\Delta x, z = m\Delta z$ and $t = k\Delta t$, where Δx,Δz are the space discretization intervals in x- and z-direction and Δt is the time discretization interval. For an accuracy of 0.1% the above summations are truncated to 16-24 terms. For simplicity, expansion only in scaling functions will be considered. Wavelets are implemented in a similar way. Upon inserting the field expansions, Maxwell's equations

are sampled [3] using pulse functions as time-domain test functions and scaling functions as space-domain test-functions and the following non-split formulation of the fields for the PML region is derived:

$$
\begin{aligned}
{}_{k+1}E^{x,\phi\phi}_{l+1/2,m} &= e^{-\sigma_E^m \Delta t/\epsilon_o}{}_k E^{x,\phi\phi}_{l+1/2,m} \\
&- e^{-0.5\sigma_E^m \Delta t/\epsilon_o}\left(\frac{1}{\Delta z}\sum_{i'=m-9}^{m+8} a(i')_{k+1/2} H^{y,\phi\phi}_{l+1/2,i'+1/2}\right) \\
{}_{k+1}E^{z,\phi\phi}_{l,m+1/2} &= e^{-\sigma_E^{m+1/2} \Delta t/\epsilon_o}{}_k E^{z,\phi\phi}_{l,m+1/2} \\
&+ e^{-0.5\sigma_E^{m+1/2} \Delta t/\epsilon_o}\left(\frac{1}{\Delta x}\sum_{i'=l-9}^{l+8} a(i')_{k+1/2} H^{y,\phi\phi}_{i'+1/2,m+1/2}\right) \\
{}_{k+1/2}H^{y,\phi\phi}_{l+1/2,m+1/2} &= e^{-\sigma_H^m \Delta t/\mu_o}{}_{k-1/2} H^{y,\phi\phi}_{l+1/2,m+1/2} \\
&+ e^{-0.5\sigma_H^m \Delta t/\mu_o}\left(\frac{1}{\Delta x}\sum_{i'=l-9}^{l+8} a(i')_k E^{z,\phi\phi}_{i',m+1/2} - \frac{1}{\Delta z}\sum_{i'=m-9}^{m+8} a(i')_k E^{x,\phi\phi}_{l,i'}\right) \qquad (8)
\end{aligned}
$$

where the terms $\sigma^m_{E,H}$ are given by Eq.(12).

A parallel-plate waveguide of width d=48 mm, terminated at both ends by PML, is used to validate the proposed algorithm. A TM source with a Gabor time variation is excited close to one side of the waveguide. The benchmark MRTD solution with no reflections is obtained by simulating the case of a much longer parallel-plate waveguide of the same width to provide a reflection-free observation area for the time interval of interest. A quadratic variation in PML conductivity is assumed for all cases, with maximum theoretical reflection coefficient of 10^{-5} at normal incidence. Numerical reflection is observed for the frequency range $[0, 0.9 f_c^{TM_1}]$ (TEM propagation) where $f_c^{TM_1} = \frac{c}{2d} = 3.125\ (GHz)$ is the cutoff frequency of the TM_1 mode. It can be seen from Figs.(1)-(2) that for 8 PML cells and σ_E^{max}=0.4 S/m it is $S_{11} \leq$-65 dB and for 16 PML cells and σ_E^{max}=0.2 S/m the reflection is smaller than -91 dB. Thus, the non-split PML absorber can be used effectively in the simulation of antennas and active elements using MRTD.

III Application of PML to the Analysis of Antenna Geometries

MRTD can successfully model both planar circuits [6] and resonating structures [7]. Recently the techniques developed for the simulation of both structures are combined to model a three-dimensional patch antenna geometry [8]. Full three-dimensional MRTD analysis is used, with PML expanded through three coordinate directions. The procedure to derive an equation for the

	Δt	PML cells along z	σ_{max}^{Ex}	σ_{max}^{Ey}	σ_{max}^{Ez}
FDTD ($60 \times 100 \times 16$)	$1.3297 \cdot 10^{-13}s$	6	3.0	3.0	3.0
MRTD($30 \times 50 \times 9$)	$1.6008 \cdot 10^{-13}s$	2–6	3.0	3.0	11.53
MRTD ($20 \times 20 \times 9$)	$1.3297 \cdot 10^{-13}s$	6–10	3.0	3.0	11.53

Table 1: Computational Parameters.

three-dimensional MRTD scheme, with PML along all three coordinate directions is presented in [8].

The patch antenna used in our simulations has the dimensions $12.45mm \times 16mm$, with a microstrip line 20 mm long used as a feed. A Gaussian pulse 4 mm from the PML layer is used to excite the microstrip. The substrate has a thickness of 0.794 mm and a relative dielectric constant equal to 1. An FDTD mesh of $60 \times 100 \times 16$ is compared to MRTD grids of $30 \times 50 \times 9$ and $20 \times 20 \times 9$, which exhibit savings of memory over FDTD on the order of 7.22 and 33 respectively. Note that these values do not include the PML layers. Figure 3 shows a comparison plot of calculated S_{11} data for the three cases listed above. Six cells of PML are added along the $\pm x$, $\pm y$ and $+z$ directions with $\sigma_{max}^{Ex} = \sigma_{max}^{Ey} = 3.0$ and $\sigma_{max}^{Ez} = 11.53$ for all cases. The time discretization interval used for the MRTD $30 \times 50 \times 9$ scheme is $\Delta t = 1.6008 \cdot 10^{-13}s$ while the MRTD $20 \times 20 \times 9$ scheme uses a time discretization interval of $\Delta t = 1.42384 \cdot 10^{-13}s$. FDTD uses a time discretization interval of $\Delta t = 1.3297 \cdot 10^{-13}s$. In all three cases the simulation is performed for 10000 time steps. This information is summarized in Table 1.

Figure 4 shows a comparison of S_{11} data for different numbers of z–directed PML layers for an MRTD discretization of $30 \times 50 \times 9$. Note that the S_{11} values correlate very well even for only 2 PML layers in the z–direction. Figure 4 shows a comparison of S_{11} data for different numbers of z–directed PML layers for an MRTD discretization of $20 \times 20 \times 9$. Once again the values of S_{11} show good correlation.

IV Conclusion

An efficient PML absorber in non-split formulation is presented for the MRTD Scheme based on cubic spline Battle-Lemarie scaling functions. This absorber is used effectively to model an antenna geometry providing extremely small numerical reflections. In comparison to Yee's conventional FDTD scheme, the proposed MRTD scheme coupled with the PML absorber offer memory savings by a factor of 12-30 and execution time savings by a factor of about 3-5 maintaining a better accuracy for S-parameter calculations. For structures where the edge effect is prominent, additional wavelets can be used to improve the accuracy when using a

coarse MRTD mesh.

V Acknowledgments

This work was made possible by ONR contract N00014-95-1-1299 and ARO contract DAAH04-95-1-0321.

References

[1] M.Krumpholz, L.P.B.Katehi, "MRTD: New Time Domain Schemes Based on Multiresolution Analysis", *IEEE Trans. Microwave Theory and Techniques*, vol. 44, no. 4, pp. 555-561, April 1996.

[2] M.Krumpholz, L.P.B.Katehi, "MRTD Modeling of Nonlinear Pulse Propagation", to be published at the *IEEE Trans. Microwave Theory and Techniques.*

[3] E.M.Tentzeris, R.Robertson, M.Krumpholz, L.P.B.Katehi, "Application of the PML Absorber to the MRTD Technique", *Proc. AP-S 1996*, pp. 634-637.

[4] K.S.Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media", *IEEE Trans. Antennas Propagation*, pp.302-307, May 1966.

[5] J.-P. Berenger, "A Perfectly Matched Layer for the Absorption of Electromagnetic Waves", *J.Comput. Physics*, vol. 114, pp. 185-200, 1994.

[6] E. Tentzeris, M. Krumpholz and L.P.B. Katehi, "Application of MRTD to Printed Transmission Lines", *Proc. MTT-S 1996*, pp. 573–576.

[7] R. Robertson, E. Tentzeris, M. Krumpholz and L.P.B. Katehi, "Application of MRTD Analysis to Dielectric Cavity Structures",*Proc. MTT-S 1996*, pp. 1840-1843.

[8] R. Robertson, E. Tentzeris, and L.P.B. Katehi, "Modeling of Membrane Patch Antennas using MRTD Analysis", *Proc. AP-S 1997*, pp. 126–129.

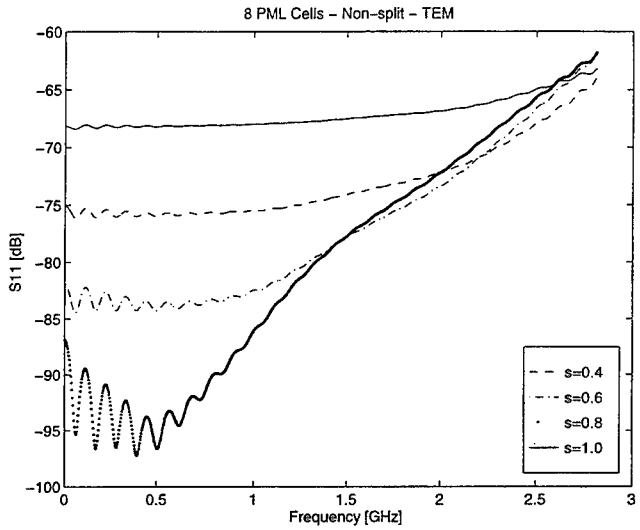


Figure 1: 8 PML cells.

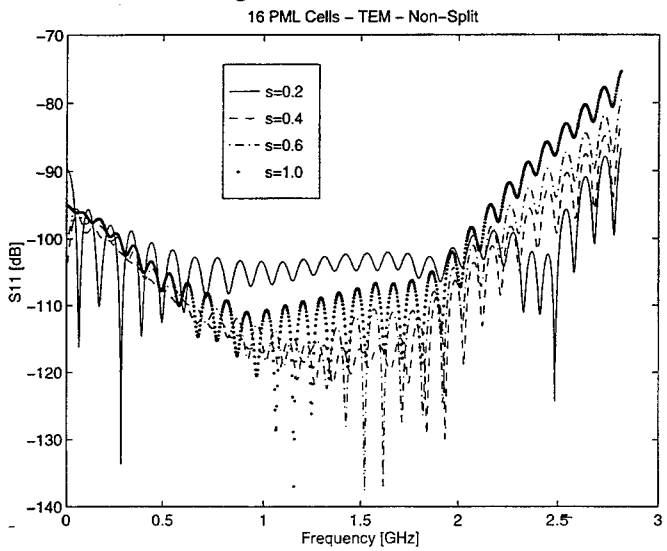


Figure 2: 16 PML cells.

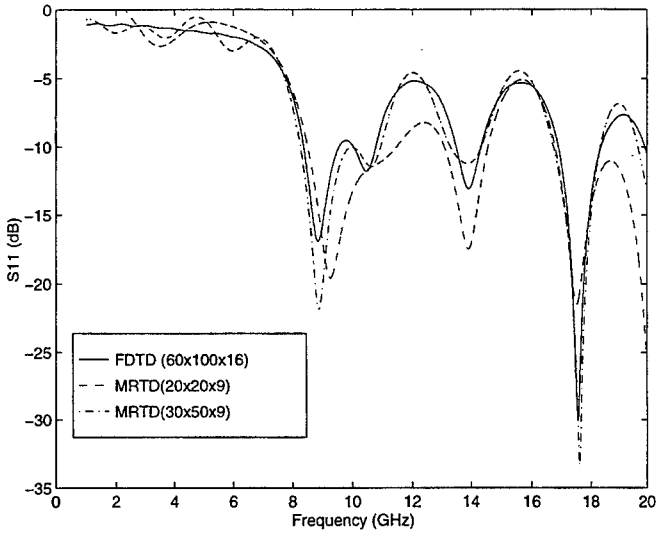


Figure 3: S_{11} **comparison plots for a patch antenna**

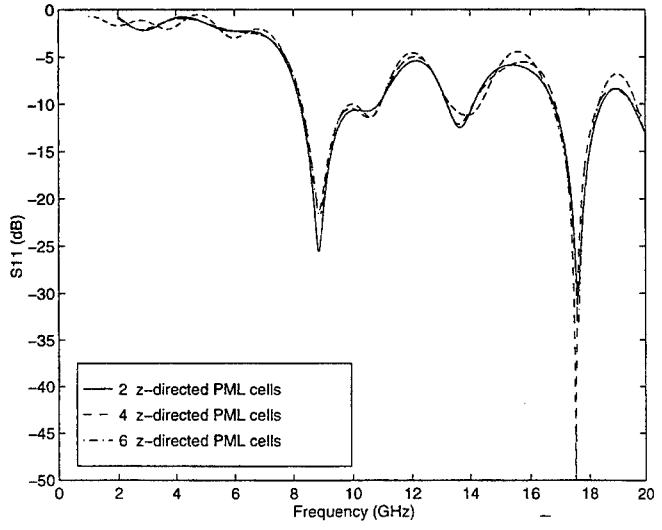


Figure 4: MRTD$(30 \times 50 \times 9)$ S_{11} **plot for varying PML layers in the z–direction**

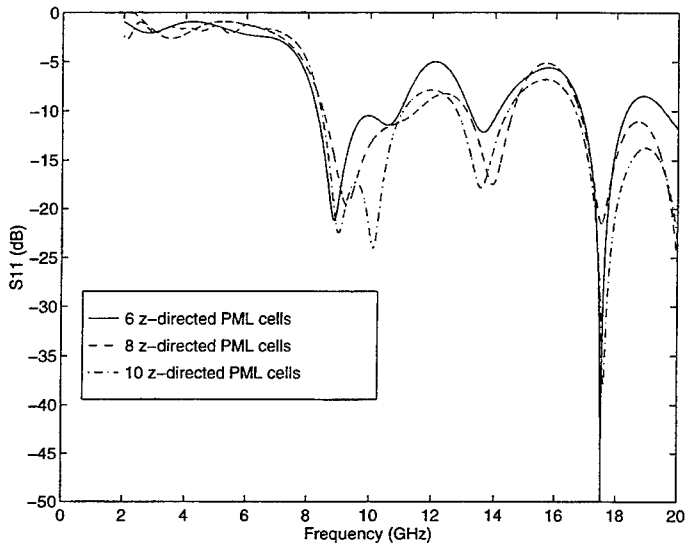


Figure 5: MRTD($20 \times 20 \times 9$) S_{11} plot for varying PML layers in the z–direction

A PML-FDTD Algorithm for General Dispersive Media

Guo-Xin Fan and Qing Huo Liu
Klipsch School of Electrical and Computer Engineering
New Mexico State University
Las Cruces, NM 88003

Abstract A three-dimensional (3D) finite-difference time-domain (FDTD) algorithm with perfectly matched layer (PML) absorbing boundary condition (ABC) is presented for general inhomogeneous, dispersive, conductive media. The modified time-domain Maxwell's equations for dispersive media are expressed in terms of coordinate-stretching variables. The recursive convolution (RC) and piecewise linear recursive convolution (PLRC) approaches are extended to arbitrary dispersive media in a more general form. The algorithm is validated for homogeneous and inhomogeneous dispersive media, and excellent agreement between the FDTD results and analytical solutions is obtained with both RC and PLRC approaches. We demonstrate the applications of the algorithm by several examples in subsurface radar detection of mine-like objects, cylinders and spheres buried in a dispersive half-space, and a three-layer medium with a dipping interface.

I. Introduction

Finite-difference time-domain (FDTD) method, as one of most powerful computational methods in electromagnetics, has been widely used to simulate wave propagation, scattering, and radiation. In the early development and applications of FDTD, the parameters of media are constants independent of frequency. When the media are frequency-dependent, especially for those encountered in the applications involving earth, biological materials, artificial dielectrics, and optical materials, this frequency dispersive property will significantly change the electromagnetic response in the media. In these cases, the original FDTD algorithm needs to be modified to account for the frequency dispersion of the media.

In recent years, three major frequency-dependent FDTD methods have been proposed: recursive convolution (RC) [1,2], auxiliary differential equation (ADE) [3,4], and Z-transform (ZT) [5]. The stability and error analysis for various frequency-dependent FDTD methods is given in [7, 8]. It is reported that among all the above frequency-dependent FDTD methods, the RC and its modified version, piecewise linear recursive convolution (PLRC) methods require least computer storage, and the PLRC, ADE, and ZT approaches have better accuracy than the RC approach [6]. In addition, the RC and PLRC approaches allow to treat a wide variety of dispersive media in a unified form, while the ADE and ZT approaches require the different formulations for different kinds of dispersive media.

As in the FDTD method for non-dispersive media, when applied to an unbounded domain, the frequency dependent FDTD algorithm calls for absorbing boundary conditions (ABCs) to truncate the computational domain. Among all the existing ABCs, the perfectly matched layer (PML) is most effective, which gives zero reflection at the absorbing boundary for all frequencies and all angles of incidence [9, 10]. Moreover, the PML is ideal for parallel computation. In addition, the PML ABC can be applied to the domain where a dipping interface exists. Most previous FDTD algorithms on dispersive media employed non-PML ABCs, such as Mur's ABC and Liao's ABC, and most PML ABCs are limited to lossyless an non-dispersive media. Only recently, the PML ABC has been extented to lossy media and dispersive lossy media [11–14].

In this paper, a 3D FDTD algorithm is presented for general inhomogeneous, dispersive, conductive

media using the coordinate stretching approach, and the RC and PLRC approaches are extended to general dispersive media in a more unified form. Three common types of dispersive media, i.e., Lorentz media, unmagnetized plasma and Debye media, can be treated as special cases of our general formulas. Several validation and application examples are also given.

II. Formulation

A. Modified Maxwell's Equations for Dispersive Media

Consider an isotropic, conductive, inhomogeneous, linear permittivity dispersive medium. Using the coordinate stretching approach [10] and following a similar procedure as in [12], the modified Maxwell's curl equations with the split fields ($\eta = x, y, z$) in the time domain can be written as

$$\frac{\partial}{\partial \eta}(\hat{\eta} \times \mathbf{E}) = -a_\eta \mu \frac{\partial \mathbf{H}^{(\eta)}}{\partial t} - \mu \omega_\eta \mathbf{H}^{(\eta)} - \mathbf{M}^{(\eta)} \tag{1}$$

$$\frac{\partial}{\partial \eta}(\hat{\eta} \times \mathbf{H}) = a_\eta \frac{\partial \mathbf{D}^{(\eta)}}{\partial t} + \omega_\eta \mathbf{D}^{(\eta)} + a_\eta \sigma \mathbf{E}^{(\eta)} + \omega_\eta \sigma \int_{-\infty}^{t} \mathbf{E}^{(\eta)}\, dt + \mathbf{J}^{(\eta)}. \tag{2}$$

Equations (1) and (2) consist of a total of 12 scalar equations, since both $\mathbf{E}^{(\eta)}$ and $\mathbf{H}^{(\eta)}$ have two scalar components perpendicular to $\hat{\eta}$, and $\mathbf{D}^{(\eta)}$ also has the two corresponding components due to the constitutive relations of the medium. These equations are insufficient to solve the total 18 field components. The remaining equations will be given by the constitutive relations.

B. Recursive Convolution Approaches

Noting that the constitutive relations take the same form for all split components, we omit all the superscript (η) in this subsection for simplicity.

For a linear dispersive medium, the relationship between the electric flux density and the electric field intensity in the time domain is described by

$$\mathbf{D}(t) = \epsilon_0 \epsilon_\infty \mathbf{E}(t) + \epsilon_0 \int_{-\infty}^{t} \mathbf{E}(\tau)\, \chi(t-\tau)\, d\tau \tag{3}$$

where ϵ_0 is the free-space permittivity, ϵ_∞ is the relative permittivity at $\omega \to \infty$, and χ is the electric susceptibility.

The frequency domain susceptibility functions, as the transfer function of a linear system, can be generally expressed as a ratio of two polynomials [1] or in a fractional form, i.e.,

$$\chi(\omega) = \sum_{q=1}^{M'} \beta_q\, s^q \Big/ \sum_{q=1}^{M} \zeta_q\, s^q = \sum_{q=1}^{M} \frac{\Gamma_q}{s - s_q}, \qquad (M > M') \tag{4}$$

where $s = -i\omega$, and s_q and Γ_q are the complex poles and the corresponding residues. Then the corresponding time domain susceptibility functions can be written as

$$\chi(t) = \sum_{q=1}^{N} \mathrm{Re}[\hat{\chi}_q(t)] = \sum_{q=1}^{N} \mathrm{Re}[R_q e^{s_q t} U(t)] \tag{5}$$

where $U(t)$ is the unit step function. In (5), $N = M$, $R_q = \Gamma_q$ when all s_q and Γ_q are real; and $N = M/2$, $R_q = 2\Gamma_q$ when there are $M/2$ complex-conjugate pole pairs (such as Lorentz media) which satisfy $\Gamma_q(s_q) = \Gamma_q^*(s_q^*)$ since $\chi(t)$ is a real function. Note that when s_q and Γ_q are real, $\hat{\chi}_q$ and all other derived functions are also real.

To simplify (3), we first introduce a unified piecewise approximation to $\mathbf{E}(t)$ over the time interval $t \in [m\Delta t, (m+1)\Delta t]$ as follows,

$$\mathbf{E}(t) \approx \mathbf{E}(m+1) + K_a \frac{[\mathbf{E}(m+1) - \mathbf{E}(m)]}{\Delta t} [t - (m+1)\Delta t]. \tag{6}$$

It is noted that Equation (6) corresponds to the recursive convolution (RC) [1] when $K_a = 0$, and to the piecewise linear recursive convolution (PLRC) [2] when $K_a = 1$. Combining these two approximations in the form of (6) is convenient for us to compare the numerical accuracy of RC and PLRC approaches in a consistent way.

Using (5) and the unified approximation (6), the convolution integral in (3) is then transformed into the discrete convolution summation,

$$\mathbf{D}(n) = \epsilon_0 \epsilon_\infty \mathbf{E}(n) + \epsilon_0 \sum_{q=1}^{N} \sum_{m=0}^{n-1} \mathrm{Re}\Big\{ \mathbf{E}(n-m) \hat{\chi}_q(m) + [\mathbf{E}(n-m-1) - \mathbf{E}(n-m)] \hat{\xi}_q(m) \Big\} \tag{7}$$

where

$$\hat{\chi}_q(m) = \int_{m\Delta t}^{(m+1)\Delta t} \hat{\chi}_q(\tau)\, d\tau, \qquad \hat{\xi}_q(m) = \frac{K_a}{\Delta t} \int_{m\Delta t}^{(m+1)\Delta t} (\tau - m\Delta t) \hat{\chi}_q(\tau)\, d\tau. \tag{8}$$

It can be shown that

$$\hat{\chi}_q(m) = \hat{\chi}_q(0) e^{s_q m \Delta t}, \qquad \hat{\xi}_q(m) = \hat{\xi}_q(0) e^{s_q m \Delta t} \tag{9}$$

and

$$\hat{\chi}(0) = \begin{cases} R_q \Delta t; \\ \dfrac{R_q}{s_q} (e^{s_q \Delta t} - 1), \end{cases} \qquad \hat{\xi}_q(0) = \begin{cases} \dfrac{K_a R_q \Delta t}{2}; & \text{for } s_q = 0; \\ \dfrac{K_a R_q}{\Delta t s_q^2} \left[1 - (1 - s_q \Delta t) e^{s_q \Delta t}\right], & \text{for } s_q \neq 0 \end{cases} \tag{10}$$

Similar to [1] and [2], we introduce a new variable $\Psi_q(n)$ so that

$$\Psi_q(n) = \sum_{m=0}^{n-1} \left\{ \left[\hat{\chi}_q(0) - \hat{\xi}_q(0) \right] \mathbf{E}(n-m) + \hat{\xi}_q(0) \mathbf{E}(n-m-1) \right\} e^{s_q m \Delta t}. \tag{11}$$

Then (7) can be written as

$$\mathbf{D}(n) = \epsilon_0 \epsilon_\infty \mathbf{E}(n) + \epsilon_0 \sum_{q=1}^{N} \mathrm{Re}[\Psi_q(n)]. \tag{12}$$

Finally, we can obtain

$$\Psi_q(n+1) = [\hat{\chi}_q(0) - \hat{\xi}_q(0)] \mathbf{E}(n+1) + \hat{\xi}_q(0) \mathbf{E}(n) + \Psi_q(n) e^{s_q \Delta t} \tag{13}$$

and

$$\mathbf{D}(n+1) = \epsilon_0 \left\{ \epsilon_\infty + \sum_{q=1}^{N} \mathrm{Re}\left[\hat{\chi}_q(0) - \hat{\xi}_q(0)\right] \right\} \mathbf{E}(n+1) + \epsilon_0 \sum_{q=1}^{N} \mathrm{Re}\left[\hat{\xi}_q(0)\right] \mathbf{E}(n) + \epsilon_0 \sum_{p=1}^{N} \mathrm{Re}\left[\Psi_q(n) e^{s_q \Delta t}\right]. \tag{14}$$

Up to this point, the recursive convolution formulation is derived for general dispersive media. To use it, a set of R_q and s_q need to be determined in advance for a given medium. For example, these parameters for Debye media are given by

$$\chi(\omega) = (\epsilon_s - \epsilon_\infty) \sum_{q=1}^{N} \frac{G_q}{1 - i\omega\tau_q}, \qquad \chi(t) = (\epsilon_s - \epsilon_\infty) \sum_{q=1}^{N} \frac{G_q e^{-t/\tau_q}}{\tau_q} U(t) \tag{15}$$

$$R_q = \frac{(\epsilon_s - \epsilon_\infty) G_q}{\tau_q}, \qquad s_q = -\frac{1}{\tau_q} \tag{16}$$

where ϵ_s is the relative static permittivity, τ_q is the Debye relaxation time constant, and G_q is the pole amplitude.

It is worth pointing out that for an arbitrary linear dispersive medium, when the discrete spectral magnitude data are available for the susceptibility of the medium, the frequency-domain Prony method (FDPM) can be used to find directly the poles and residues, i.e. R_q and s_q [15]. Therefore, for an arbitrary dispersive medium there is no need to fit the dispersive relation with Debye or Lorentz models.

C. Discretization

With the help of the recursive convolution equations (12)-(14), we can proceed to solve Maxwell's equations by using the Yee's algorithm to discretize the split equations (1) and (2). We obtain

$$\left(\frac{a_\eta}{\Delta t} + \frac{\omega_\eta}{2}\right) \mathbf{H}^{(\eta)}(n+\tfrac{1}{2}) = -\frac{\partial}{\partial \eta}\left[\hat{\eta} \times \mathbf{E}(n)\right] - \left(\frac{a_\eta}{\Delta t} - \frac{\omega_\eta}{2}\right) \mathbf{H}^{(\eta)}(n-\tfrac{1}{2}) - \mathbf{M}^{(\eta)}(n) \tag{17}$$

and

$$c_1^{(\eta)} \mathbf{E}^{(\eta)}(n+1) = \frac{\partial}{\partial \eta}\left[\hat{\eta} \times \mathbf{H}(n+\tfrac{1}{2})\right] + \Phi^{(\eta)}(n) + c_0^{(\eta)} \mathbf{E}^{(\eta)}(n) - \sigma\omega_\eta \Delta t \mathbf{E}_I^{(\eta)}(n) - \mathbf{J}^{(\eta)}(n+\tfrac{1}{2}) \tag{18}$$

where

$$\Phi^{(\eta)}(n) = \epsilon_0 \sum_{q=1}^{N} \mathrm{Re}\left\{ \left[\left(\frac{a_\eta}{\Delta t} - \frac{\omega_\eta}{2}\right) - \left(\frac{a_\eta}{\Delta t} + \frac{\omega_\eta}{2}\right) e^{-s_q \Delta t} \right] \Psi_q^{(\eta)}(n) \right\} \tag{19}$$

$$\mathbf{E}_I^{(\eta)}(n) = \mathbf{E}_I^{(\eta)}(n-1) + \frac{7}{8}\mathbf{E}^{(\eta)}(n) + \frac{1}{8}\mathbf{E}^{(\eta)}(n-1) \tag{20}$$

and

$$c_0^{(\eta)} = \left(\frac{a_\eta}{\Delta t} - \frac{\omega_\eta}{2}\right) \epsilon_0 \epsilon_\infty - \frac{a_\eta \sigma}{2} - \left(\frac{a_\eta}{\Delta t} + \frac{\omega_\eta}{2}\right) \epsilon_0 \sum_{q=1}^{N} \mathrm{Re}\left[\hat{\xi}_q(0)\right] \tag{21}$$

$$c_1^{(\eta)} = \frac{\sigma}{2}\left[a_\eta + 2s\omega_\eta \Delta t\right] + \left(\frac{a_\eta}{\Delta t} + \frac{\omega_\eta}{2}\right) \epsilon_0 \left\{ \epsilon_\infty + \sum_{q=1}^{N} \mathrm{Re}\left[\hat{\chi}_q(0) - \hat{\xi}_q(0)\right] \right\}. \tag{22}$$

Table I Parameters for Debye Media

Medium I	Medium II
$\epsilon_\infty = 3$	$\epsilon_\infty = 3.7677$
$\epsilon_s = 4.5$	$\epsilon_s = 20.2677$
$\tau_1 = 6.4 \times 10^{-10}$ s	$\tau_1 = 1.1614 \times 10^{-11}$ s
$G_1 = 1$	$G_1 = 1$
$\sigma = 0.005$ S/m	$\sigma = 0.1165$ S/m

Equations (17)-(20), together with (13), form the FDTD time-stepping equations.

Note that when updating $\mathbf{E}$ field, it appears that the two steps $\mathbf{E}(n)$ and $\mathbf{E}(n-1)$ are needed in the equations (13) and (20). The storage requirement of $\mathbf{E}(n-1)$, actually, can be avoided by means of a temporary variable [2]. At this point, the implementation of the algorithm requires storage for $\mathbf{E}^{(\eta)}$, $\mathbf{H}^{(\eta)}$, $\mathbf{E}_I^{(\eta)}$, and $\Psi_q^{(\eta)}$. Because of introduction of the PML, each of the above quantities has 6 components. While $\mathbf{E}_I^{(\eta)}$ is due to the conductivity of the media, $\Psi_q^{(\eta)}$ results from the frequency dispersion of the media. In addition, in general a complex array $\Psi_q^{(\eta)}$ (except for Debye media and unmagnetized plasma where it is a real array) is needed for each q. Therefore, the treatment of the dispersive media requires more memory than that in non-dispersive media.

III. Numerical Results

Based on the algorithm proposed in the previous section, a 3D-FDTD Fortran program is developed. The PML equations are applied to both the interior region and the matched layers. Ten cells of PMLs are used outside the interior region as the absorbing boundary condition in all computations.

In following examples, an electric dipole directed in $\hat{x}$ direction is used as a source, and the field component E_x is measured at a series of receiver locations. The time function of the source is the first derivative of the Blackman-Harris window function [12]. The central frequency of this function is defined as $f_c = 1.55/T$ where T is the duration of the source function.

A. Validation

To validate the algorithm, we consider two group of testing cases: (i) a homogeneous dispersive medium and (ii) a dispersive sphere embedded in another dispersive or non-dispersive background medium. The analytical solutions are available for a dipole source in both cases. Three typical kinds of media, i.e. Lorentz media, unmagnetized plasma, and Debye media, are under consideration. In these testing examples, the source is located in the origin of coordinates, and the E_x field component at 10 locations is displayed. The field is normalized with respect to the peak value at the fourth receiver in all E_x waveforms. In the calculation of FDTD, the solution region is divided by $64 \times 64 \times 64$ cells. The FDTD numerical results are compared with the analytical solutions. Because of limited space, only the results for Debye media are given below. The parameters for Debye media are given in Table I. The complex permittivity of two Debye media is plotted as a function of frequency in Fig. 1.

Fig. 2 shows the E_x waveforms for a homogeneous Debye medium (I), and Fig. 3 gives the results for a Debye sphere (Medium I) in an unbounded Debye medium II.

In all the testing examples, an excellent agreement between the FDTD numerical results and analytical solutions is observed. It is interesting that both results of RC and PLRC display excellent agreement with

the analytical solutions.

B. Applications

To demonstrate the effectiveness of the algorithm, we consider several applications of subsurface radar. The earth is modeled by Debye dispersive media in all examples. For clarity, only the scattered fields, obtained by subtracting the fields in the absence of buried objects from the total fields, are shown. In these examples, the sources are located in air-ground interface with $(x, y) = (0, 0)$, and the receivers are located on the air-ground interface along x-axis. The central frequency of the source is 80 MHz. The computational region is divided into $200 \times 64 \times 64$ cells or $128 \times 64 \times 64$ cells.

First, we consider two rectangular cylinders and a PEC sphere buried in a half-space of Debye medium I. The cylinders are with air and Debye medium II, respectively. Fig. 4. shows the geometry of the problem and the scattered E_x waveforms received at 181 locations.

Next, we consider the mapping of a dipping interface in ground-penetrating radar detection application. The geometry of the problem is shown in Fig. 5. The upper, middle, and lower media are air, Debye medium I, and Debye medium II. The E_x waveforms are recorded at 109 locations, and the scattered fields are shown in Fig. 5.

In all application examples above, the scattered fields from the buried objects or layers are clearly displayed. For the last problem with a dipping interface, other ABCs will become unstable as soon as the waves propagate to the boundary. The PML ABC provides an unparalleled advantage in this aspect.

IV. Conclusions

We present a 3D FDTD algorithm with the PML absorbing boundary condition for general inhomogeneous, dispersive, conductive media. The modified time-domain Maxwell's equations for dispersive media are expressed in terms of the coordinate-stretching variables. A single formulation is developed to include recursive convolution and piecewise linear recursive convolution for arbitrary dispersive media. We validated the algorithm for both homogeneous dispersive media and a dispersive sphere in another dispersive or non-dispersive background medium for three typical kinds of dispersive media. Excellent agreement between the FDTD results and analytical solutions is obtained for all cases. Several applications are demonstrated for subsurface radar detection of cylinders and a sphere buried in a dispersive half-space. Furthermore, a problem with a dipping interface which cannot be modeled by non-PML ABCs, is simulated. The algorithm proposed is ideal for parallel computation since the same code is shared by both the interior computational region and the outer matched layers. Because of their generality, the algorithm and computer program developed can be used to model biological materials, artificial dielectrics, optical materials, and other dispersive media.

Acknowledgment

This work was supported by Sandia National Laboratories under a SURP grant, and by Environmental Protection Agency under a PECASE grant CR-825-225-010. We thank Drs. David Womble and Scott Hutchinson of Sandia National Laboratories for the suggestions leading to this research.

References

[1] R. Luebbers, F. P. Hunsberger, K. Kunz, R. Standler, and M. Schneider, "A frequency-dependent finite difference time domain formulation for dispersive materials," *IEEE Trans. Electromag. Compat.*, vol. 32, pp. 222-227, 1990.

[2] R. J. Luebbers, D. Steich, and K. Kunz, "FDTD calculation of scattering from frequency-dependent materials," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 1249-1257, 1993.

[3] D. F. Kelley and R. J. Luebbers, "Piecewise linear recursive convolution for dispersive media using FDTD," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 792-797, 1996.

[4] R. M. Joseph, 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 pulse," *Opt. Lett.* vol. 16, pp.1412-1414, 1991.

[5] J. L. Young, "Propagation in linear dispersive media: Finite difference time-domain methodologies," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 422-426, 1995.

[6] D. M. Sullivan, "Z-transform theory and the FDTD method," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 28-34, 1996.

[7] P. G. Petropoulos, "Stability and phase error analysis of FD-TD in dispersive dielectrics," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 62-69, 1994.

[8] J. L. Young, A. Kittichartphayak, Y. M. Kwok, and D. Sullivan, "On the dispersion errors related to $(FD)^2TD$ type schemes," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 1902-1910, 1995.

[9] J. R. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comp. Phys.*, vol. 114, pp. 185-200, 1994.

[10] W. C. Chew and W. H. Weedon, "A 3D perfectly matched medium from modified Maxwell's equation with stretched coordinates," *Microwave Opt. Tech. Lett.*, vol. 7, pp. 599-604, 1994.

[11] J. Fang 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. 44, pp. 2216-2222, 1996.

[12] Q. H. Liu, "An FDTD algorithm with perfectly matched layers for conductive media," *Microwave Opt. Tech. Lett.*, vol. 14, pp. 134-137, 1997.

[13] S. D. Gedney, "An anisotropic PML absorbing media for the FDTD simulation of fields in lossy and dispersive media," *Electromagnetics*, vol. 16, pp. 399-415, 1996.

[14] F. L. Teixeira, W. C. Chew, M. Straka, M. L. Oristaglio, and T. Wang, "3D PML-FDTD simulation of ground penetrating radar on dispersive earth media," *Proc. 17th Intl. Geosci. Remote Sensing Symp. (IGARSS'97)*, Singapore, 1997.

[15] J. N. Brittingham, E.K. Miller, and J. L. Wilows, "Pole extraction from real-frequency information," *Proc. IEEE*, vol. 68, pp. 263-273, 1980.

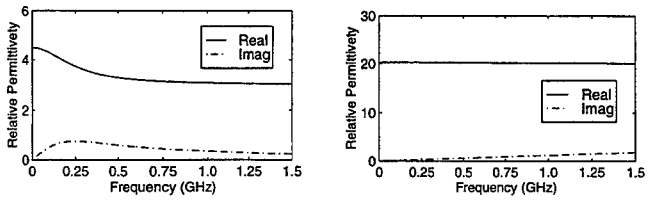


Fig. 1. Complex permittivity of Debye media I (a) and II (b).

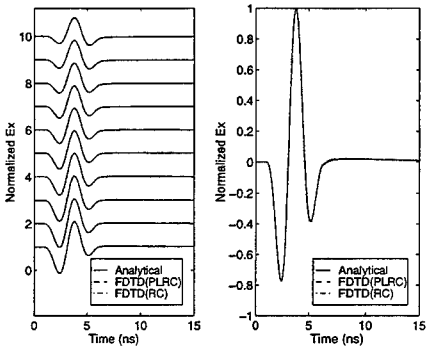


Fig. 2. Comparison of FDTD results with analytical solutions for a homogeneous Debye medium I. (a) The E_x component at the array of receivers. (b) The E_x component at the fourth receiver.

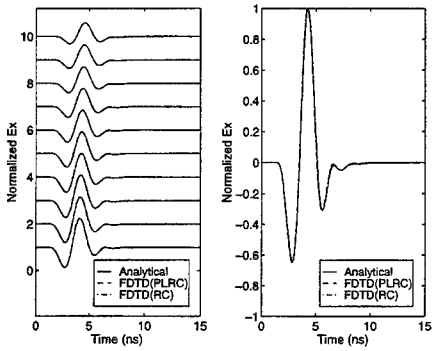


Fig. 3. Comparison of FDTD results with analytical solutions for a Debye sphere (Medium I) in homogeneous Debye medium II. (a) The E_x component at the array of receivers. (b) The E_x component at the fourth receiver.

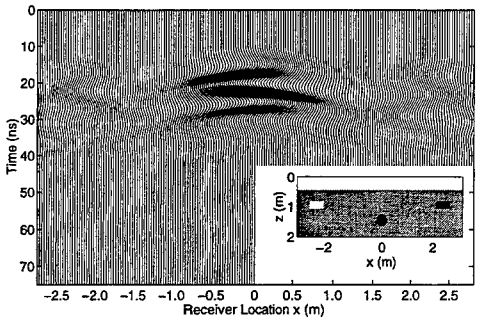


Fig. 4. The E_x waveforms of two rectangular cylinders and a sphere buried in a Debye medium half-space.

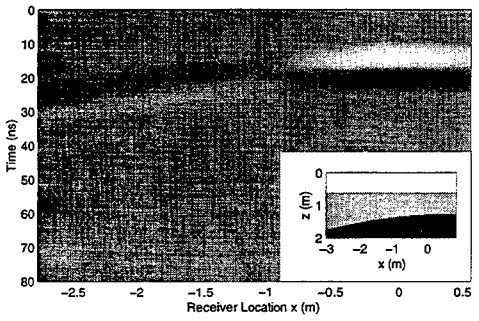


Fig. 5. The E_x field distribution of a three-layer medium with a dipping interface.

SESSION 11:

CEM AND PARAMETER EXTRACTION FOR PACKAGING ANALYSIS

Chairs:E. van Deventer and J.E. Schutt-Aine

A new generalized de-embedding method for numerical electromagnetic analysis.

Yuri O. Shlepnev
Eagleware Corporation
2245 Dillard Rd., Tucker, GA, 30084, USA
tel. 770-939-0156, fax 770-939-0157, e-mail: yuri@eagleware.com

Abstract

A new method of extraction of multimode S-parameters from numerical electromagnetic analysis is presented. The method is based on the idea that any propagated eigenwave of a line can be excited without reflection using a proper system of surface sources. Parameters of the sources, found from analysis of segments of the line, are used to calculate generalized S-parameters of discontinuities in this line. The procedure is named as a method of simultaneous diagonalisation (MoSD). The main advantages of the MoSD are perfect matching of a particular mode and the lack of the necessity of calculating line's eigenwaves. The method is illustrated by analysis of a segment of non-symmetrical coupled microstrip line and an open end in this line. The method of lines and a set of rectangular excitation regions in partial metallization planes are used to solve this problem.

Introduction

A complete electromagnetic (EM) analysis of an entire structure is the best way to obtain the characteristics of a passive microwave structure. Practically, this may be impossible because of undue computer resources requirements. In such cases it is convenient to divide the structure into components that can be analyzed separately and then combine the matrices that describe the separate parts. These matrices (Y, Z, S etc.) are, in general, multimode and can be obtained easily if eigenwaves of transmission lines corresponding to cross-sections formed in the dividing process are involved in the numerical procedure. However, eigenwaves are sophisticated, calculating them is a problem, and direct usage of them leads to unnecessary analytical and numerical difficulties. As an alternative, it is possible to use simpler and more natural functions in the regions where the eigenmodes should be excited or matched. The method of simultaneous diagonalisation (MoSD) has been developed to transform the functions in the excitation regions to the space of eigenwaves and to match each eigenwave perfectly.

The MoSD is based on EM analysis of two (or more) segments of line corresponding to a circuit component port to be de-embedded. These segments have different lengths and have excitation regions at the opposite ends with uncertain boundary conditions. The result of EM analysis is a set of Y-matrices relating excitation functions coefficients of electric and magnetic fields. These matrices transformed from the space of the excitation functions to a space of eigenmodes are set equal to Y-matrices describing independent modes propagated in continuous part of the line segments. It gives the basic non-linear system of equations relating propagation constants and characteristic impedances of the modes, a matrix of transformation from the excitation space to the mode's space (transformation matrix) and some auxiliary matrix that helps to match propagated modes perfectly (compensation matrix). Solution of the system is based on simultaneous diagonalization of Y-matrix blocks. Boundary value problem for the component or discontinuity is formulated in the same way as for the line segments. Each port of the discontinuity is substituted by

excitation regions and could be de-embedded using pre-calculated parameters of the line and transformation and compensation matrices.

The nearest analogue of the proposed technique is numerical de-embedding procedure suggested by J. Rautio [1] and in comparison with it the MoSD makes it possible to separate propagated modes and eliminates completely reflection of modes from simple line segment, that increase accuracy of analysis of discontinuities. As a by-product, the generalization of the "TEM equivalent impedance" [2] on multimode case is obtained. The MoSD was originally proposed for analysis of a single microstrip line discontinuities [3] and then generalized to a multiconductor or multimode line case [4]. The method was used in the computer program TAMIC-I [5] and then in the program =EMstar= [6] for calculation of the characteristics of discontinuities in multiconductor microstrip lines, slotlines, finlines, and coplanar waveguides.

Theory

Let us consider an arbitrary discontinuity formed by a set of semi-infinite transmission lines approaching to it. The structure is bounded by electric or magnetic walls and contains arbitrary number of lossless dielectric or magnetic and (or) metal regions inside. The discontinuity region can also contain some lossy objects. The problem of discontinuity analysis can be reduced to a problem in an enclosed volume. To excite and match propagated eigenwaves of the approaching lines it is possible to use auxiliary sources. The sources can be placed in cross-sections of the lines at the outer boundary of the volume or in some regions of the cross-sections or near the cross-sections. As this take place, the rest of the surfaces (or whole surfaces) of the cross-sections at the outer boundary will be simulated as magnetic or electric walls. Thus, in a general case, we have a 3-D boundary-value problem for the Maxwell's equations with uncertain boundary conditions in the sources regions.

To find characteristics of the sources we need to analyze separately segments of the lines approaching to the discontinuity. Each segment should have the same position of the sources regions at the opposite sides as corresponding input in the discontinuity. Let us find relation between the EM field components in the sources regions for a line segment of length l. To do this, we represent the tangential electric and magnetic field components on the surface of the sources regions as follows:

$$\overline{E}_{1,2}^{N} = \sum_{n=1}^{N} U_{1,2}^{n} \overline{E}_n \qquad \overline{H}_{1,2}^{N} = \sum_{n=1}^{N} I_{1,2}^{n} \overline{H}_n \qquad (1)$$

where $\overline{E}_n$, $\overline{H}_n$ - orthogonal basis functions defined in the sources regions. The subscripts 1 and 2 correspond to the regions at the opposite sides of the line segment. The basis functions are normalized so that unknown coefficients $U_{1,2}^{n}$ and $I_{1,2}^{n}$ have dimensions of voltage and current respectively. Suppose the problem is solved by some numerical method that is best suited to a given line type. As a result of the solution we obtain an admittance matrix $Y(l)$ relating the coefficients $\overline{U}_{1,2} = \left[U_{1,2}^{n}, n = \overline{1,N}\right]$ and $\overline{I}_{1,2} = \left[I_{1,2}^{n}, n = \overline{1,N}\right]$ in the sources regions:

$$\begin{vmatrix} \overline{I}_1 \\ \overline{I}_2 \end{vmatrix} = \begin{vmatrix} Y_{11}(l) & Y_{12}(l) \\ Y_{12}(l) & Y_{22}(l) \end{vmatrix} \begin{vmatrix} \overline{U}_1 \\ \overline{U}_2 \end{vmatrix} \qquad (2)$$

Assuming that there are only K propagated modes in the structure, and it is possible to transform currents and voltages defined in the sources regions to a space of currents and voltages of the line eigenmodes, we obtain the following matrix equations:

$$F \cdot \left[Y_{11}(l) - G \right] \cdot F^{t} = diag\left[Z_k^{-1} \cdot \mathrm{cth}(\beta_k l_k), k = \overline{1,K} \right]$$
$$F \cdot Y_{12}(l) \cdot F^{t} = diag\left[- Z_k^{-1} \cdot \mathrm{cosech}(\beta_k l_k), k = \overline{1,K} \right] \qquad (3)$$

Here F is transformation matrix, ^{t} denotes Hermitian transposed matrices. The right-hand sides of the equations (3) are diagonal matrices of the order K. An eigenmode number k is described by a model line with electric length $\beta_k l_k$ and characteristic impedance Z_k. The order of the matrices Y_{11} and Y_{12} must not be less than K. It is assumed, that the eigenwaves do not interact to each other over a line segment less than or equal l. Matrix G or compensation matrix accounts for an effect of the evanescent waves near the sources regions. The decay of these waves must be much shorter than l. Let us lengthen the investigated line segment by a value Δl so, that the relative positions of the sources regions and the segment butt-ends remain unchanged. Keeping the expansion (1) in the sources regions, we solve the boundary-value problem again and obtain a linear system of equations similar to (2) with the admittance matrix of the lengthen line segment $Y(l + \Delta l)$. It is supposed that the lengthened segment has the same model representation with unchanged matrices F and G and with increased electric lengths of the model lines describing eigenwaves. In this case, the next system of matrix equations for blocks of the matrix $Y(l + \Delta l)$ can be written:

$$F \cdot \left[Y_{11}(l + \Delta l) - G \right] \cdot F^{t} = diag\left[Z_k^{-1} \cdot \mathrm{cth}(\beta_k (l_k + \Delta l)), k = \overline{1,K} \right]$$
$$F \cdot Y_{12}(l + \Delta l) \cdot F^{t} = diag\left[- Z_k^{-1} \cdot \mathrm{cosech}(\beta_k (l_k + \Delta l)), k = \overline{1,K} \right] \qquad (4)$$

Thus, we have obtained the system of nonlinear matrix equations (3,4) with the unknown matrices F and G and parameters of the model lines or eigenwaves. The matrix F could be found as a matrix that transforms the matrices $Y_{12}(l)$ and $Y_{12}(l + \Delta l)$ to diagonal ones simultaneously. In the case when N = K these matrices are symmetric and to diagonalize them and to find F the generalized Jacobi's method may be used. After diagonalization we can find the non diagonal elements of the matrix $F \cdot G \cdot F^{t}$ from the first equations of the system (3) or (4), and after this we derive K systems of four simple trigonometric equations for each eigenwave with unknown variables Z_k, β_k, l_k and g_k (diagonal element of the matrix $F \cdot G \cdot F^{t}$). The solution of these systems may be found analytically [3,4].

After analysis of all different lines approaching to the discontinuity we can analyze the discontinuity itself. As a result of the solution of this problem we obtain an admittance matrix Y_u relating the coefficients of expansion (1) in all sources regions. Now the normalized admittance matrix relating currents and voltages of the eigenwaves of the lines approaching to the discontinuity is determined by:

$$Y = Z_0^{1/2} \cdot F_u \cdot (Y_u - G_u) \cdot F_u^{t} \cdot Z_0^{1/2}$$

Here Fu, Gu are block-diagonal matrices that unite the transformation matrices and the compensation matrices found for each input. Z_0 is a diagonal matrix that unites the characteristic impedances of the eigenmodes of the input lines.

Numerical Examples

As test examples we consider a problem of calculation of the characteristics and scattering parameters of a segment of non-symmetrical coupled microstrip line and a problem of calculation of S-matrix of an open end in this line. To solve corresponding 3D EM boundary-value problem, the impedance-interpreted method of lines is used [4,7-9]. Figure 1 shows a segment of the coupled microstrip line to be analyzed.

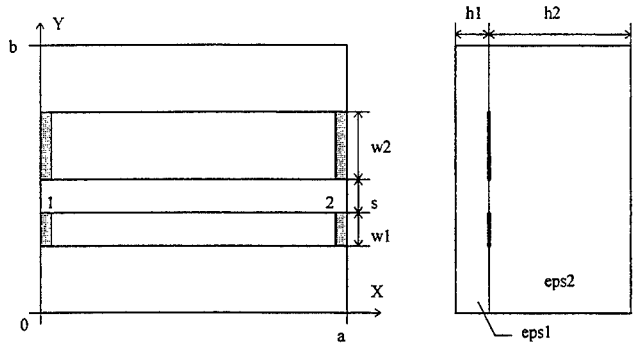


Fig. 1 A segment of non-symmetric coupled line.

To excite two dominant quasi-TEM eigenwaves in the line, we use the surface current sources located between line conductors and electric side-walls (hatched regions in the Fig.1). Size of the regions along line is equal to size of one grid cell. Surface currents are directed along line. The results of EM analysis of the line segment with parameters a=b=1.2mm., w1=s=0.05mm, w2= 0.1mm, h1=0.2mm, eps1=1, h2=2mm, eps2=12.85 at frequency 60 GHz are given in table 1 for different grid parameters.

Table 1.

dx	dy	Z1, Ohm	p1	Z2, Ohm	p2	|S11/11|	|S11/22|	∠S12/11, deg.	∠S12/22, deg.
a/64	w1	84.589	3.1528	32.579	2.6792	6.86e-10	8.09e-09	-272.59	-231.64
	w1/2	88.032	3.1345	37.005	2.6766	2.44e-10	9.23e-09	-271.01	-231.42
	w1/3	89.240	3.1284	38.660	2.6760	5.67e-10	9.66e-09	-270.48	-231.36
	w1/4	89.856	3.1254	39.523	2.6757	7.30e-10	9.88e-09	-270.22	-231.34
	w1/5	90.229	3.1236	40.051	2.6755	8.30e-10	1.00e-08	-270.07	-231.32
	w1/6	90.480	3.1224	40.408	2.6754	8.96e-10	1.01e-08	-269.96	-231.31
a/128	w1	84.546	3.1509	32.567	2.6788	6.14e-10	8.09e-09	-272.427	-231.612
	w1/2	87.988	3.1326	36.991	2.6763	3.17e-10	9.23e-09	-270.847	-231.389
	w1/3	89.196	3.1266	38.646	2.6756	6.39e-10	9.66e-09	-270.326	-231.332
	w1/4	89.812	3.1236	39.508	2.6753	8.03e-10	9.88e-09	-270.067	-231.307
	w1/5	90.186	3.1218	40.036	2.6751	9.02e-10	1.00e-08	-269.912	-231.293
	w1/6	90.437	3.1206	40.392	2.6750	9.69e-10	1.01e-08	-269.809	-231.284

The first two columns of the table (dx and dy) show grid cell sizes in terms of cells per length of the segment and cells per width of the first strip. Z1, p1 and Z2, p2 are characteristic impedance and slowing of the first and second propagating mode. $|S11/11|$ and $|S11/22|$ are absolute values of the reflection coefficients of the first and the second modes. $\angle S12/11$ and $\angle S12/22$ are angles of the transmission coefficients for both modes respectively. As evident from the table, the reflections of the eigenmodes are negligibly small regardless of grid size. Figure 2 gives another illustration of the excitation mechanism. It shows magnitudes and phases of the currents flowing along line when the first mode is excited (plane X=0). Value of current at a grid cell are obtained by integration of current density across line in the cell region. Currents are shown as continuous lines along line that corresponds to the model representation. Incident normalized wave has unit magnitude. Small bumps near the input regions indicate current redistribution that presumably occurs due to absence of y-directed currents in excitation regions. It does not affect scattering parameters of the segment but it should be accounted in a discontinuity analysis problem. A distance between inputs and discontinuity to be analyzed should provide enough space to avoid the influence of the current redistribution.

The next example is open end in the line shown in Fig. 1. Parameters of the problem: a=b=1.2mm., w1=s=0.05mm, w2= 0.1mm, h1=0.2mm, eps1=1, h2=2mm, eps2=12.85, distance between input one and open end is a/2=0.6mm. Some results of calculations at frequency 60 GHz are listed in the table 2 for different grid parameters and show good convergence of the method.

Table 2.

dx	dy	|S12|	∠S12, deg.	time, sec
a/64	w1	0.007326586	70.41083	1
	w1/2	0.010025581	71.42625	3
	w1/3	0.011042284	71.73771	7
	w1/4	0.011556851	71.88237	13
	w1/5	0.011863031	71.96429	23
	w1/6	0.012064706	72.01653	36
a/128	w1	0.0074110984	71.49882	4
	w1/2	0.010250116	72.54727	15
	w1/3	0.011349859	72.87912	39
	w1/4	0.011916206	73.03631	81
	w1/5	0.012256258	73.12625	144
	w1/6	0.012481109	73.18381	233

Here S12 is transmission coefficient from the first to the second mode. The last column of the table gives calculation time on computer Pentium Pro 200 MHz with 64 Mb. Reference plane is placed in the plane of the open end. To verify calculated data we compare them with results of the spectral-domain approach obtained in [10]. Phase of S12 obtained in [10] is about 73 deg., that corresponds to our data. Magnitude calculated in [10] is 0.02, that is slightly different from our result. Figure 3 illustrates interaction of the two dominant modes of the coupled line on the open end. The first mode is excited and has unit normalized magnitude. Current values are obtained through integration in cells regions and are given in amperes.

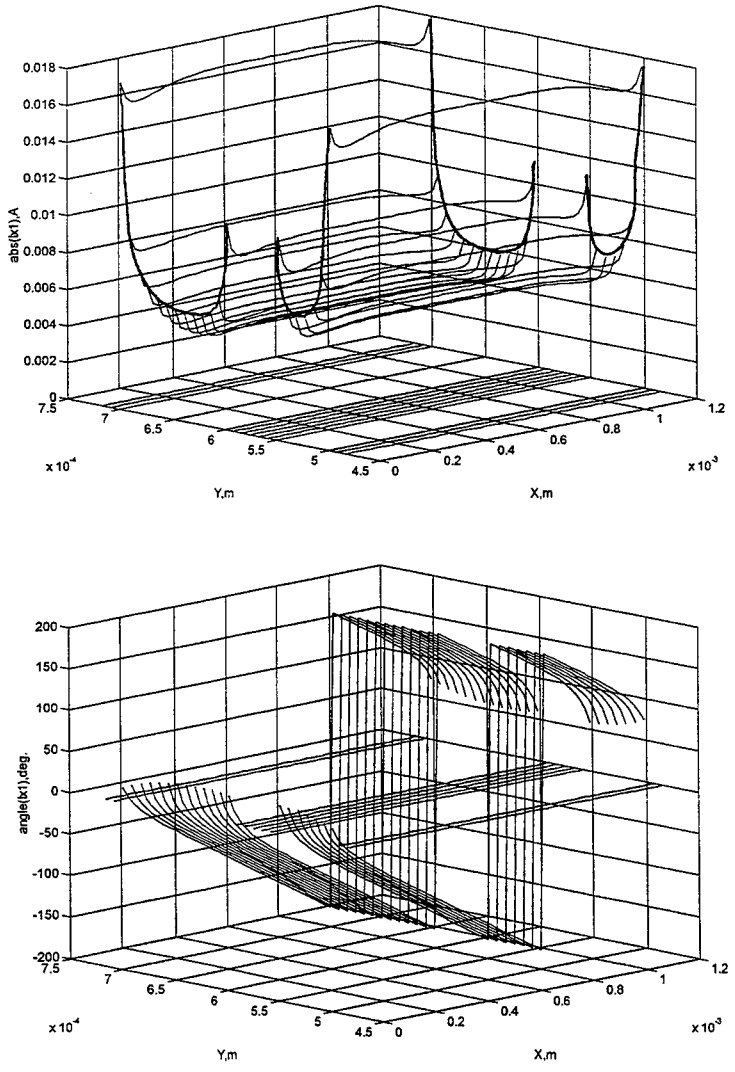


Fig 2. Magnitude (top) and angle (bottom) of the current flowing along segment of the coupled MSL when the first mode is excited at the input one at frequency 60 GHz.

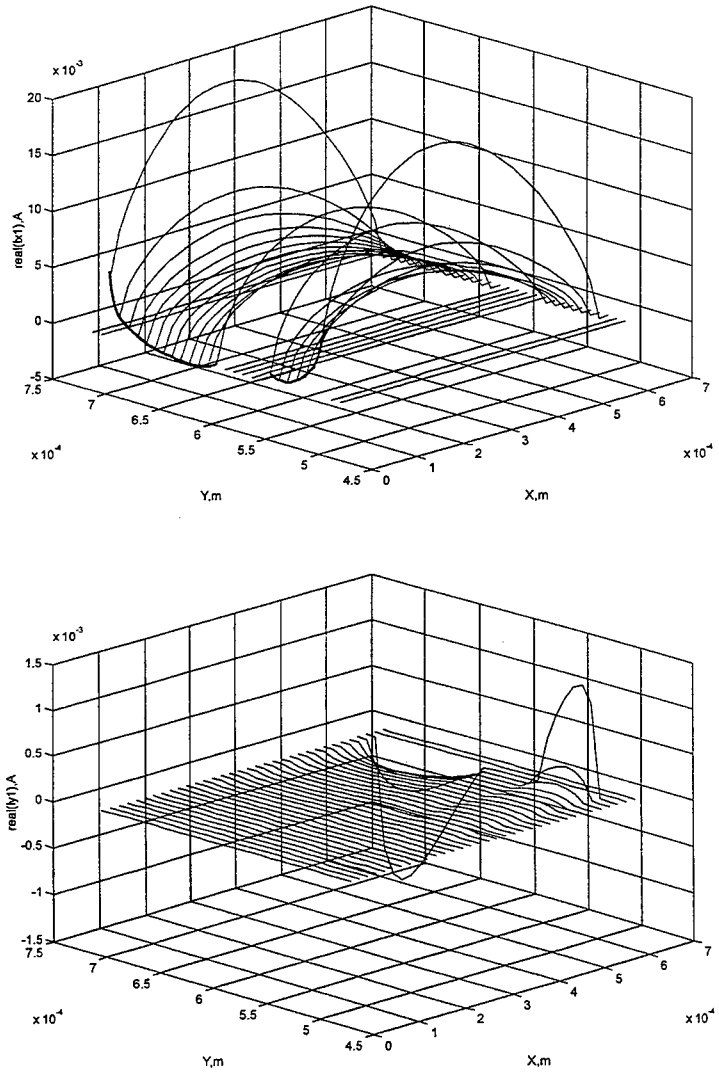


Fig 3. Real parts of the current along line (top) and across line (bottom) in the open end of the coupled MSL when the first mode is excited at the plane X=0 at frequency 60 GHz.

Conclusion

The method of simultaneous diagonalisation is proposed in this paper as a general approach for extraction of multimode S-matrices of discontinuities in arbitrary shielded lines. The main advantage of the method is ideal matching of line eigenmodes in the analysis of line segnment that increases accuracy of discontinuity analysis. The procedure of analysis of a line cross-section is substituted by more natural analysis of 3D structure with extraction of information about eigenwaves indirectly. It develops the idea of "TEM equivalent impedance" [2] on the multimode case. The MoSD is illustrated by coupled microstrip line segment analysis and open end analysis examples. Processes inside structures are explained using current distribution plots. Calculated data are given in the tables and can be used for verification.

References

1. Rautio J.C. A de-embedding algorithm for electromagnetics. - Int. Journal of Microwave and Millimeter-Wave Computer-Aided Engineering. v.1, 1991, N3, p 282-287.
2. Rautio J.C. A new definition of characteristic impedance. - MTT Int. Symposium Digest, June 1991, Boston, p.761-764.
3. Sestroretzkiy B.V., Kustov V.Yu., Shlepnev Yu.O. // *Voprosi Radioelektroniki (in Russian)*, ser. OVR, 1988, N12, p. 26.
4. Shlepnev Yu.O. Method of lines in mathematical modeling of MIC's planar elements (*in Russian*). - Thesis of Candidate of Science. - Novosibirsk, NEIS, 1990.
5. TAMIC-I - A program for 3D electromagnetic analysis of MIC's passive elements of the TAMIC_soft group.
6. =EMstar= - Electromagnetic Analizer of MICs of the Eagleware Corporation.
7. Slobodianskiy M.G. // *Applied Mathematics and Mechanics (in Russian)*, v. 3, 1939, N 1, p. 75.
8. Worm S.B., Pregla R. // IEEE Trans., v. MTT-32, 1984, N 2, p. 191.
9. Kustov V.Yu., Sestroretzkiy B.V., Shlepnev Yu.O. // Proc. of Europ. Symp. on Numerical Methods in Electromagnetics (JEE'93), 1993, Toulouse, France, p. 227.
10. Jansen R.H., Wertgen W. - Alta Frequenza, v. 57, 1988, N 5, p. 203-216.

A circuit extraction approach in PCB power-bus analysis

Hao Shi *
EEsof Division, Hewlett-Packard, haoshi@sr.hp.com
and
James L. Drewniak
University of Missouri-Rolla, drewniak@ee.umr.edu

1 Introduction

The DC power distribution in multi-layered printed circuit boards (PCBs) for high-speed designs is typically achieved with a power bus consisting of at least one pair of ground/power planes. This DC power-bus configuration is known to contribute to electromagnetic interference (EMI) and signal integrity (SI) problems despite its low impedance [1]. A primary concern is the phenomenon of the simultaneous switching noise (SSN) [2].

With appropriate CAD simulations, the effect of SSN can be investigated at a lower cost through various "what-if" scenarios; therefore, a general knowledge can be drawn. The crude model of a lumped parallel-plane capacitor works well up to a certain frequency, typically only in the hundreds of megahertz [3]. The distributed behavior of the power bus neglected by the single-capacitor model can be partially recovered in other models, such as a radial transmission line [4, 5]. Widely applied electromagnetic modeling techniques like FDTD and FEM can also be used, but there are difficulties accommodating device models in a general fashion. Another class of simulation techniques extracts the electromagnetic behavior of a system in terms of a collection of equivalent circuit elements [6, 7, 8]. Because of the accessibility of general purpose SPICE simulators and the availability of standardized device models including IBIS models, a circuit extraction approach is very desirable in power-bus analysis. One well developed circuit extraction technique is the partial element equivalent circuit (PEEC) method proposed by Ruehli [9], and based on an electric field integral equation (EFIE). This study adopts a general circuit extraction technique that employs a conformal mesh [6], and is based on a mixed potential integral equation (MPIE) [10]. The circuit extraction / MPIE technique will be denoted simply CEMPIE. A general purpose DC power-bus modeling tool has been developed that can accommodate multiple layers with scatterers on intervening layer between power and ground layers including via walls and plane segmentations. Surface mount decoupling capacitors are modeled with lumped elements that includes the series inductance and resistanee of the interconnect to the power planes [3].

*Portions of this work was completed during the course of Ph. D. study at the University of Missouri-Rolla

2 Derivation of CEMPIE formulation

Assume S is the domain of a metalization surface in the PCB design. In response to an incident field $\vec{E}^{inc}(\vec{r})$, there will be an induced current $\vec{J}_s$ on S. The scattered electric field is then

$$\vec{E}^s(\vec{J}_s) = -j\omega\vec{A} - \nabla\phi, \tag{1}$$

where the vector and scalar potential are $\vec{A} = \mu \int_S \mathbf{G}_\omega^A(\vec{r},\vec{r}') \cdot \vec{J}_s ds'$, and $\phi = \frac{1}{\epsilon}\int_S \sigma G_\omega^\phi(\vec{r},\vec{r}')ds'$, with $\mathbf{G}_\omega^A$ and G_ω^ϕ the Green's functions for a general medium corresponding to $\mathbf{A}$ and ϕ, respectively. The boundary condition on S requires

$$\hat{n} \times (\vec{E}^{inc} + \vec{E}^s) = Z_s \hat{n} \times \vec{J}_s, \qquad \vec{r} \in S \tag{2}$$

where $\hat{n}$ is the surface normal vector and Z_s is the surface impedance. Substituting expressions for $\vec{A}$ and ϕ into Eqn. 2 yields

$$\hat{n} \times \vec{E}^{inc} = \hat{n} \times (j\omega\mu \int_S \mathbf{G}_\omega^A(\vec{r},\vec{r}') \cdot \vec{J}_s ds' + \nabla\phi + Z_s\vec{J}_s). \tag{3}$$

The unknown current density can be expanded as

$$\vec{J}_s(\vec{r}) = \sum_\gamma^M \frac{i_\gamma}{l_\gamma}\vec{f}_\gamma(\vec{r}) \tag{4}$$

where M is the total number of interior edges, γ is the running index for edges, l_γ is the length of the γth edge, i_γ is the current value (constant) passing perpendicularly across the edge, and $\vec{f}_\gamma(\vec{r})$ is the vector basis function introduced by Rao *et al.* [11]. A set of matrix equations can be derived upon substituting Eqn. 4 into Eqn. 3. Employing charge continuity and with arithmetic simplifications, the discretized system equations are

$$[\mathbf{Q}] = [\mathbf{K}^{-1}][\phi] \tag{5}$$

$$[\mathbf{R} + j\omega\mathbf{L}][\mathbf{i}] - [\mathbf{\Lambda}][\phi] = [\mathbf{v}^e] \tag{6}$$

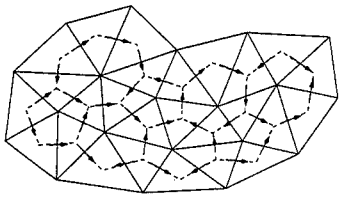
$$-j\omega[\mathbf{Q}] = [\mathbf{I}] + [\mathbf{I}^e], \tag{7}$$

where $R_{\alpha\gamma} \equiv \frac{Z_s}{l_\alpha l_\gamma} < \vec{f}_\alpha, \vec{f}_\gamma >$, $L_{\alpha\gamma} \equiv \frac{\mu}{l_\alpha l_\gamma} < \vec{f}_\alpha, \int_S \mathbf{G}_\omega^A(\vec{r},\vec{r}') \cdot \vec{f}_\gamma ds' >$, $v_\alpha^e = \frac{1}{l_\alpha} < \vec{f}_\alpha, \vec{E}^{inc} >$, and, $K_{mn} = \frac{1}{\epsilon A_m A_n}\int_{T_m}\int_{T_n} G_\omega^\phi(\vec{r},\vec{r}')ds'ds$. The connectivity matrix $\mathbf{\Lambda}_{M\times N}$ is determined by

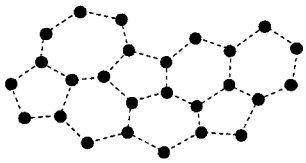
$$\Lambda_{\alpha n} = \begin{cases} 1, & \text{if Node n is the starting point of Edge } \alpha \\ -1, & \text{if Node n is the ending point of Edge } \alpha \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Moreover, the $[\mathbf{I}^e]$ vector represents the external currents injected through the individual cells. Figure 1(a) shows a diagram of all current branches crossing the interior edges. The resultant Voronoi diagram (cell centers as vertices) in Figure 1(b) shows the circuit nodes in the equivalent circuit model.

Eqns. 5 to 7 can be compared to the canonical nodal-based circuit equations yielding an {L,C,R} linear circuit network. The resultant circuit will have N (number of mesh cells) nodes. When losses are neglected, i.e., $Z_s = 0$, there is a parallel LC branch between any two nodes, say Node m, n. A capacitance also exists between each node and the common datum node [7]. A typical circuit element is shown in Figure 2(a) and (b).

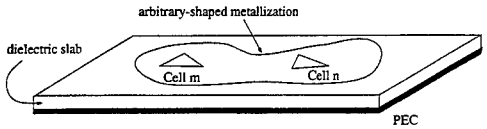


(a)

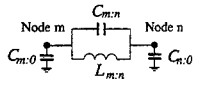


(b)

Figure 1: Discretization of a metal region with (a) Delaunay triangulation and (b) the corresponding Voronoi diagram.



(a)



(b)

Figure 2: Two mesh cells and the corresponding equivalent circuit model.

Table 1: ρ_c (mm) in a grounded dielectric slab configuration. Values enclosed by the square brackets are computed by the empirical formula.

$d\ (mil)$, ϵ_r	$1\ GHz$	$5\ GHz$	$10\ GHz$
40, 9.6	6.46 [6.27]	2.55 [2.40]	1.53 [1.49]
20, 9.6	5.88 [5.97]	1.93 [2.10]	1.27 [1.18]
11, 9.6	5.62 [5.61]	1.46 [1.74]	1.00 [0.82]
11, 4.7	12.98 [12.87]	2.80 [3.38]	1.60 [1.13]
11, 2.55	unclear [22.32]	6.70 [6.70]	3.00 [3.00]

3 Mesh constraint due to quasi-static approximation

A general rule of thumb in the EFIE/MPIE is that the mesh dimension should be at least as small as $0.1\lambda_c = 0.1\sqrt{\epsilon_{eff}}c_0/f_c$ [6], where ϵ_{eff} is the effective dielectric constant. However, a quasi-static approximation is employed here [6, 7, 8], in order to achieve frequency-independent elements in the extracted equivalent circuit. The approximation impacts the mesh density, and the typical $0.1\lambda_c$ guideline must be re-addressed in the context of the quasi-static approximation.

In the quasi-static approximation, the frequency-dependent Green's functions are replaced by their static counterparts, which allows the extracted inductances and capacitances to be frequency independent. Hence, the same circuit model can be implemented easily in simulations of both time and frequency domains. However, as a trade-off, the mesh dimension in CEMPIE needs to be fine enough to capture the high-frequency behavior of the system.

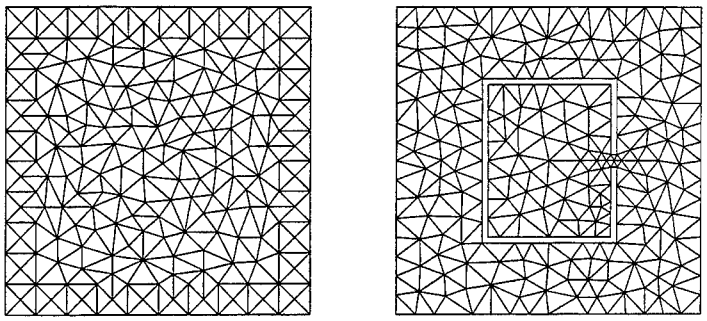
The spatial scalar Green's function is a function of the relative lateral distance between the source and field points, i.e., $G^{\phi}_{\omega>0}(\vec{r}_s, \vec{r}_f) = G^{\phi}(\rho; z_s, z_f)$ where $\rho = \sqrt{(x_s - x_f)^2 + (y_s - y_f)^2}$. It is known that there exists a critical value ρ_c such that $G^{\phi} \sim \rho^{-1}$ when $\rho < \rho_c$ while $G^{\phi} \sim \rho^{-1/2}$ when $\rho > \rho_c$ [12]. In other words, the approximation of G^{ϕ}_{ω} by G^{ϕ}_0 (G^{ϕ}_0 is always behaving ρ^{-1}) becomes very poor when $\rho > \rho_c$.

A two-plane PCB power-bus can be modeled by a power plane over a grounded dielectric slab (or microstrip) with thickness d and dielectric constant ϵ_r. ρ_c values are studied for five combinations of $\{d, \epsilon_r\}$ at three frequencies (1, 5 and 10 GHz), and tabulated in Table 1. Using a trial-and-error approach, an empirical formula for ρ_c is developed based on the tabulated data as

$$\rho_c = \frac{\frac{60.52}{\sqrt{\epsilon_r}} - 1.92\sqrt{\epsilon_r} - 6.58}{\sqrt{f}} - \frac{32.8 - 0.586\epsilon_r - 7.28\sqrt{\epsilon_r}}{\sqrt{d}}, \tag{9}$$

where f is in GHz and d is in mil. The ρ_c's evaluated with the empirical formula are also listed in Table 1. The relative error of ρ_c computed through the empirical formula is less than 20 % for $1 \leq \epsilon_r \leq 12$, $10 \leq d \leq 40$, and $0.5 \leq f \leq 10$, and only applies to this range presently, since beyond ρ_c could be negative. A more general relationship is being pursued.

The mesh dimension can be quantified by the average edge length $\bar{l}_e \equiv \frac{\sum_{\alpha}^{M} l_{e,\alpha}}{M}$, assuming the mesh cells are relatively homogeneous. A "good" quasi-static approximation should assure adequate approximation in the dominant matrix elements of the [**K**] matrix, $K_{mm} =$



(a) (b)

Figure 3: Meshes for (a) *Power-bus* and (b) *Power-island* geometries.

Table 2: Mesh statistics for both *Power-bus* and *Power-island*.

Geometry	No. of cells	Internal edges	edge statistics	l_e (mm)
Power-bus	386	559	77.5 % 3.84 $\sim$ 5.41 mm	4.081
Power-island	416	564	78.4 % 3.06 $\sim$ 4.75 mm	3.686

$\frac{1}{\epsilon A_m^2} \int_{T_m} \int_{T_m} G^{\phi}(\vec{r}, \vec{r}') ds' ds$. It is then necessary to guarantee the quasi-static approximation to be good for any choice of source/field locations within each cell. An "average" cell has the dimension of l_e, thus, it is necessary that $\bar{l}_e < \rho_c$ which leads to an overall CEMPIE meshing criterion

$$\bar{l}_e < \min\{\rho_c,\ 0.1\,\lambda_c\}. \tag{10}$$

4 Applications of CEMPIE in PCB power-bus analysis

Two types of boards are considered to demonstrate the application of the CEMPIE approach: (1) a thin board with $d = 10$ *mil* and $\epsilon_r = 2.99$ and (2) a thick board with $d = 43$ *mil* and $\epsilon_r = 4.7$. The ground plane is free of discontinuities, whose area is intentionally created larger than that of the power plane. The ground plane is treated as if it has infinite extent in the xy-plane; hence, the Green's functions can be computed with a perfect electric conducting plane of infinite extent. Only the power planes then need to be meshed. Two power plane geometries are used for structures without or with discontinuities: (1) a 50 $mm \times 50$ mm power bus or *Power-bus* in short, and (2) a power bus with gaps resembling an island or *Power-island* in short. To be efficient, *Power-buses* or *Power-islands* with different material properties share a common mesh. The meshes for the *Power-bus* and *Power-island* structures are shown in Figure 3(a) and (b), respectively. The mesh characteristics are displayed in Table 2. Assuming the upper frequency

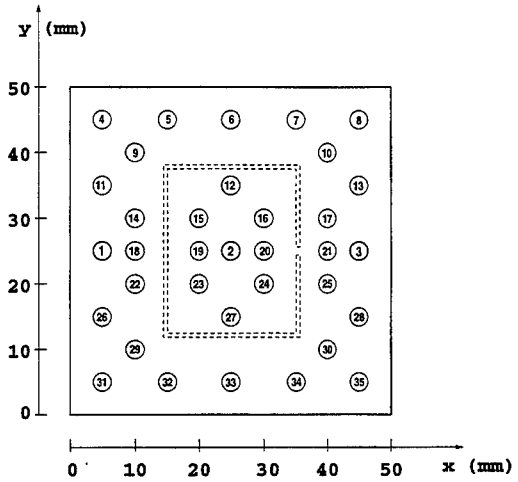


Figure 4: Indexed locations in both *Power-bus* and *Power-island* geometries.

under consideration is $f_c = 5\ GHz$, then $0.1\lambda_c$ is 6 mm, and by Eqn. 9 the critical edge lengths are 5.387 mm and 5.505 mm for the thin and thick board materials, respectively. Thus, the necessary meshing condition in Eqn. 9 is not violated, and the two meshes are likely to provide reliable circuit models.

Since both meshes can be covered by a 50 $mm \times 50$ mm area, a common map is displayed in Figure 4 with indexed locations. The extracted models are used to simulate frequency-domain S_{21} results with corroborating measurements. When the two ports are selected at Locs. 1, 3 referring to Figure 4, both simulated and measured data are presented in Figure 5. When the two ports are selected at Locs. 1, 2, both simulated and measured data are presented in Figure 6. In general, the agreement is good to approximately $3-4\ GHz$. Discrepancies are in part due to measurement artifacts, and variations in the dielectric constant with increasing frequency and skin effect, both of which have been neglected.

5 Discussions

The CEMPIE technique presented here is able to predict the distributed behavior, corroborated by measurements, of a PCB power-bus structure in terms of a LC linear circuit network. It is particularly useful when other devices of concern are also represented by circuit models such as IBIS models. Novel power-plane design such as a power island can be evaluated with simulations. The CEMPIE formulation requires explicit matrix inversions, which is currently implemented with a LU type direct matrix solver. In the future, iterative methods will be explored for matrix computations to improve efficiency. In addition, a circuit reduction technique is being investigated to minimize the number of necessary circuit nodes.

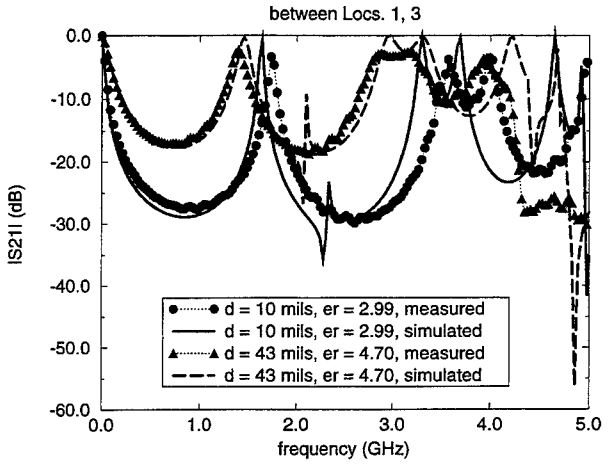


Figure 5: Comparison of simulations and measurement for the *Power-bus* structures.

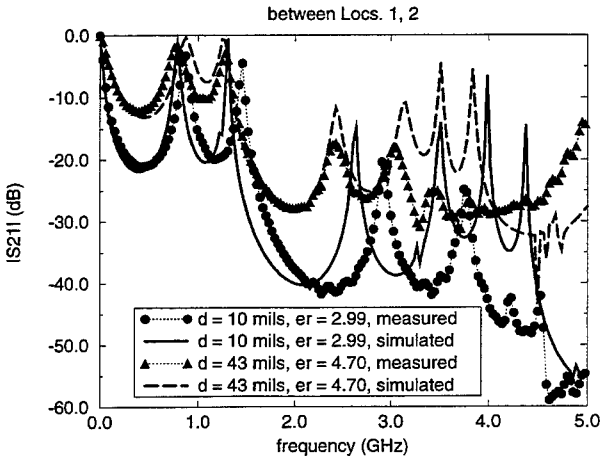


Figure 6: Comparison of simulations and measurement for the *Power-island* structures.

References

[1] Clayton R. Paul, *Introduction to Electromagnetic Compatibility*, John Wiley, New York, 1992.

[2] Ramesh Senthinathan and John L. Prince, *Simultaneous Switching Noise of CMOS Devices and Systems*, Kluwer Academic Publishing, Boston, 1994.

[3] H. Shi, , F. Sha, J. L. Drewniak, T. P. Van Doren, and T. H. Hubing, "An experimental procedure for characterizing interconnects to the DC power bus on a multi-layer printed circuit board," *IEEE Transactions on Electromagnetic Compatibility*, vol. 39, no. 4, 1997.

[4] Jiayuan Fang, Yaowu Liu, Yuzhe Chen, Zhonghua Wu, and Amit Agrawal, "Modeling of power/ground plane noise in high speed digital electronics packaging," *IEEE 2nd topical meeting on Electronical Performance of Electronic Packaging, Monterey, CA*, pp. 206–208, October 1993.

[5] James C. Parker Jr., "Via coupling within parallel rectangular planes," *IEEE Transactions on Electromagnetic Compatibility*, vol. 39, pp. 17–23, no. 1, February 1997.

[6] R. du Cloux, G. P. J. F. M. Maas, and A. J. H. Wachters, "Quasi-static boundary element method for electromagnetic simulation of PCBs," *Phillips Journal of Research*, vol. 48, no. 1–2, pp. 117–144, 1994.

[7] K. J. Scott, *Practical Simulation of Printed Circuit Boards and Related Structures*, John Wiley, New York, 1994.

[8] Frank Y. Yuan, Timothy K. Postel, and Lawrence M. Rubin, "Analysis and modeling of power distribution networks and plane structures in multichip modules and PCB's," *Proceedings of IEEE International Symposium on Electromagnetic Compatibility*, pp. 447–452, 1995.

[9] Albert E. Ruehli, Ulla Miekkala, and Hansruedi Heeb, "Stability of discretized partial element equivalent EFIE circuit models," *IEEE Transactions on Antennas and Propagation*, vol. 43, pp. 553–559, no. 6, June 1995.

[10] Juan R. Mosig, "Arbitrary shaped microstrip structures and their analysis with a mixed potential integral equation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, pp. 314–323, no. 2, February 1988.

[11] Sadasiva M. Rao, Donald R. Wilton, and Allen W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Transactions on Antennas and Propagation*, vol. 30, pp. 409–418, no. 3, May 1982.

[12] J. R. Mosig and F. E. Gardiol, "Analytical and numerical techniques in the Green's function treatment of microstrip antennas and scatters," *IEE PROC. Inst. Elec. Eng. Pt. H*, vol. 130, no. 2, pp. 175–182, March 1983.

Modeling of Conductor and Dielectric Losses in Packages

J. Poltz
OptEM Engineering Inc.

1. Introduction

Interconnects must be included in today's high-frequency circuit simulations. At GHz frequencies skin effect causes reduction of the interconnect inductance and a rapid increase of the interconnect resistance and leakage conductance [1, 3]. Attempts to reduce the propagation delay by lowering the interconnect capacitance (decreasing cross-sectional dimensions) result in an increase in wire resistance which, in turn, increases the rise time and indirectly slows down the response. Therefore, it is impossible to optimize packaging interconnections to maximize the clock rate without analyzing losses (solving Helmholtz equation) and implementing lossy transmission line models.

2. Calculation of inductance and resistance

There is only a range of frequencies where the Helmholtz solution is required [2, 7]. For low frequencies, the dc approximation of current distribution (uniform throughout a conductor cross-section) is accurate enough. For very high frequencies the surface currents screen the interior of the conductors. Both low (DC) and high-frequencies (HF) can be handled well by the Laplace equation - which is used by static field solvers. The transition between DC and HF range can be described as the quasi-stationary region. It is this quasi-stationary region that requires the Helmholtz solution. To estimate the frequency range of the quasi-stationary region one has to calculate the skin depth which is defined as:

$$\delta = 1/\sqrt{\pi f \mu_o \gamma} \qquad (1)$$

where: f - frequency, μ_0- permeability of the free space, γ - conductivity. Strong frequency dependence of the resistance and inductance is expected for frequencies which put the skin depth in the range of conductor cross-sectional dimensions because of a non-uniform current distribution.

Table 1. Skin depth for aluminum conductors as a function of frequency.

f [MHz]	10 MHz	100 MHz	1 GHz	10 GHz
skin depth [μm]	25.76	8.147	2.576	0.815

The Helmholtz equation has to be solved within the conductor region. Equivalent representation of Helmholtz equation is known in the form of Fredholm integral equation of the second kind, which in the three-dimensional space is written as [2, 4]:

$$\underline{J}(M) = -\frac{j\omega\mu\gamma}{2\pi}\int_S \underline{J}(P)\ln\frac{1}{r(M,P)}dS_P - j\omega\gamma\underline{A}(M) \qquad (2)$$

where: $\underline{A}(M)$ - represents external (source) field. P and M represent points in the cross-section.

The inductance and resistance matrices are calculated together as the imaginary and real parts of the impedance matrix. They are both related to the distribution of currents.

The current distribution is calculated from (2) using a combination of Finite Element Method (FEM) and Boundary Element Method (BEM). OptEM software splits bulky conductors in finite elements (bulky conductors) or multilayer boundaries (ground and power planes) as shown in Figure 1. This technique requires combining the finite element method with the boundary element method. Although difficult at the programming stage this method is the most efficient technique of solving the Helmholtz equation for practical interconnect applications.

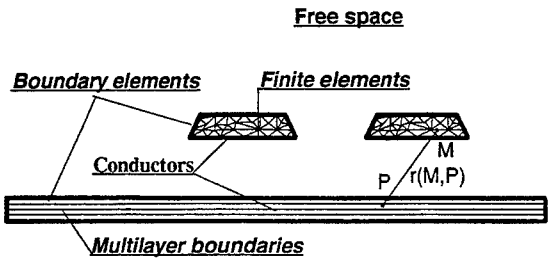


Figure 1 OptEM software calculates L, R values using a combination of BEM and FEM.

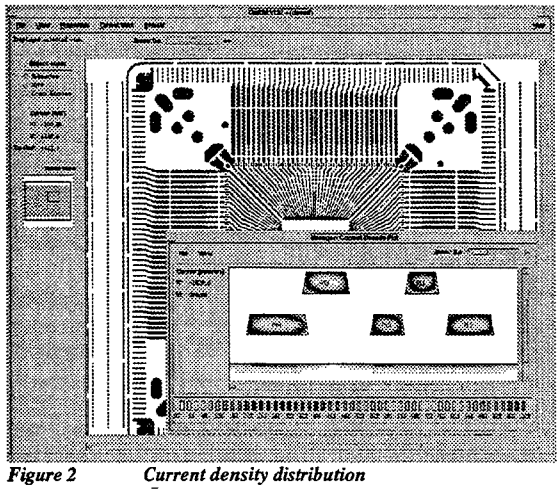


Figure 2 Current density distribution

Numerical solution of the integral equation (2) requires selection of an approximation technique and a generation of linear constrains in the form of algebraic equations. It was found out that current distribution within bulky conductors can be well represented by quadratic elements whereas currents in thin and wide planes (like ground and power layers) are represented by cubic splines based on boundary layers. This combination of finite and boundary elements offers numerical stability and accuracy.

A combination of different approximation techniques requires an adequate procedure for building linear algebraic equations. After analyzing different methods (including simple collocation) the Galerkin method was selected as the only one offering correlation between the approximation accuracy and the numerical effort in setting up and solving algebraic equations.

Before calculating inductance and resistance matrices a complex, frequency dependent current distribution is calculated for a set of independent source currents. The inductance and resistance matrices are calculated together as the imaginary and real parts of the impedance matrix. They are both related to the distribution of currents and both

frequency dependent. For five conductors shown in Figure 2, the current in each conductor is calculated from the current density as:

$$\underline{I}_i = \int_{S_i} \underline{J}(M) dS_M \tag{3}$$

The current calculation has to be repeated for five independent voltage conditions, $[\underline{V}] = j\omega\gamma[\underline{A}]$, in order to assemble entire impedance matrix. In matrix notation:

$$[R + j\omega L][I] = [\underline{V}] \tag{4}$$

In practice matrix [V] is selected as the diagonal unit matrix, therefore the unit inductance and resistance matrices can be calculated by separating the real and imaginary parts of the inverse current matrix:

$$[R] + j\omega[L] = [I]^{-1} \tag{5}$$

Both the inductance and resistance matrices are frequency dependent, as they are based on the realistic (frequency dependent) current distribution. By solving the Helmholtz equation, OptEM software automatically includes eddy-current and proximity effects when calculating current distribution.

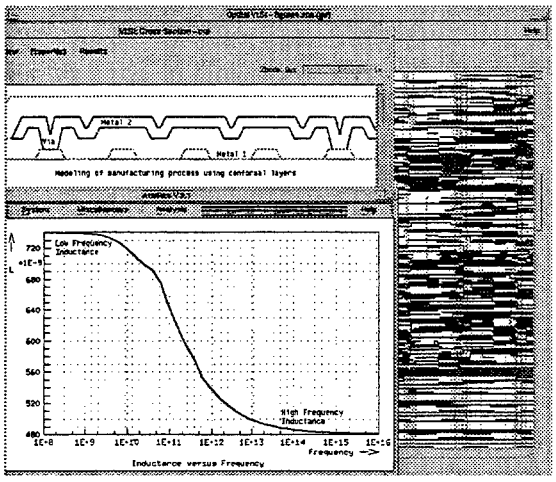


Figure 3 ***OptEM Package can accept GDSII layout and fabrication information and analyze conformal layers (deposition process) to produce frequency dependent parameters and models.***

Since magnetic field distribution outside conductors is almost the same for any frequency of the current, the difference in inductance values is caused by the gradual elimination of the magnetic field from the conductor interior. The magnetic field inside the conductor disappears, as the current distribution is pushed towards the conductor surface at higher frequencies. The loss of the magnetic field inside the conductor results in a gradual reduction of inductance in the quasi-stationary region of frequencies. Finally, at the high frequency range the magnetic field exists only outside of the conductor (Figure 3).

3. Calculating capacitance and conductance

Independent from the inductance and resistance matrices, OptEM Package calculates capacitance and conductance matrices that are related to the distribution of the electric field. Conductance matrices have non-zero entries only if materials included in the cross section have designated non-zero loss tangent.

Conductance matrices calculated for typical packaging materials demonstrate a strong frequency dependent characteristic.

Capacitance and conductance are calculated from the distribution of the transverse electric field. The cross-sectional dimensions of the system are much smaller than the wavelength, therefore it is safe to assume that the transverse electric field is induced only by the charge distribution and is decoupled from the magnetic field. This justifies using a scalar potential V as the representation for the transverse electric field. Scalar potential is defined as:

$$\vec{E} = -\text{grad}\, V \tag{6}$$

Dielectric media used in packaging are frequently lossy. Therefore, the total transverse current through the dielectric has two components known as the conductive current and the displacement current:

$$\vec{J}_t = \vec{J} + \frac{\partial \vec{D}}{\partial t} \tag{7}$$

Using the earlier introduced phasor notation

$$\underline{\vec{J}}_t = \underline{\vec{J}} + j\omega\underline{\vec{D}} \tag{8}$$

Substituting material constrains into (8), one can link the electric field with the total current as:

$$\underline{\vec{J}}_t = \gamma\underline{\vec{E}} + j\omega\varepsilon\underline{\vec{E}} = (\gamma + j\omega\varepsilon)\underline{\vec{E}} = j\omega\underline{\varepsilon}\underline{\vec{E}} \tag{9}$$

where:

$$j\omega\underline{\varepsilon} = \gamma + j\omega\varepsilon \tag{10}$$

Dividing (10) by $j\omega$, converts the complex dielectric constant, into permittivity (the real part) and conductivity scaled by frequency (the imaginary part):

$$\underline{\varepsilon} = \varepsilon + \gamma/j\omega \tag{11}$$

Using the loss tangent notation, the complex dielectric constant can be described as:

$$\underline{\varepsilon} = (1 - j\tan\delta)\varepsilon,$$

where:

$$\tan\delta = \frac{\gamma}{\omega\varepsilon} \tag{12}$$

Both the conductive and the displacement currents are coupled together through the following requirement for the total current:

$$\operatorname{div} \underline{\bar{J}}_t = 0 \tag{13}$$

The source-free behavior of the total current results from the lack of a net space charge in typical interconnect applications. Substituting (6) and (9) into (13) we get the following partial differential equation:

$$\operatorname{div}\left(j\omega \underline{\varepsilon} \operatorname{grad} \underline{V}\right) = 0, \tag{14}$$

which is equivalent to:

$$\operatorname{div}\left(\underline{\varepsilon} \operatorname{grad} \underline{V}\right) = 0 \tag{15}$$

Equation (15) is solved in uniform subregions where:

$$\underline{\varepsilon} = \text{const} \quad \text{and} \quad \operatorname{div} \operatorname{grad} \underline{V} = \Delta \underline{V} = 0 \tag{16}$$

As earlier for the current distribution, the charge distribution on dielectric interface and conductor surface can be calculated from the Fredholm integral equation. In two-dimensional space (cross-section of the transmission system):

$$\underline{V}(M) = \frac{1}{2\pi\underline{\varepsilon}} \int_l \underline{\sigma}_e(P) \ln \frac{1}{r(M,P)} \, dl_P + \underline{C} \tag{17}$$

Additional constraints are applied to the interface. Finally a condition for the total charge is used to calculate the constant $\underline{C}$:

$$\int_l \underline{\sigma}_e \, dl = 0 \tag{18}$$

where:

$\underline{\sigma}$ - charge density,

l - conductor (and dielectric interface) boundaries.

For the individual conductors the charge on the surface is calculated as:

$$\underline{Q}_i = \int_{\text{cond}_i} \underline{\sigma}_i \, dl \tag{19}$$

We must repeat the charge calculations for different voltage boundary conditions $[\underline{V}]$. $\underline{Q}$ and $\underline{V}$ are complex for lossy dielectrics, and when in a matrix notation:

$$[\underline{Q}] = \left[\underline{C} - j\frac{G}{\omega}\right][\underline{V}] \tag{20}$$

Matrix $[\underline{V}]$ can be selected as the diagonal unit matrix and therefore the unit capacitance and conductance matrices can be calculated separating the real and imaginary parts of the complex charge matrix as:

$$[C]-\frac{j}{\omega}[G]=[\underline{Q}] \tag{21}$$

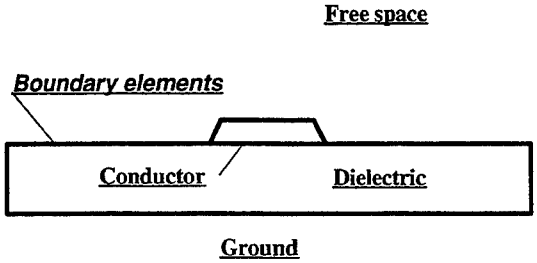


Figure 4 *Boundary Element Method is used to calculate C,G matrices*

A complex charge distribution in the form of cubic splines is used to represent sources of the electric field for lossy dielectrics. The C and G values are calculated from charge distributed on the surface of the conductors as shown in Figure 4.

Any dielectric material becomes a lossy dielectric at a sufficiently high frequency. The phase of the polarization vector lags behind that of the applied field causing a hysteresis loss.

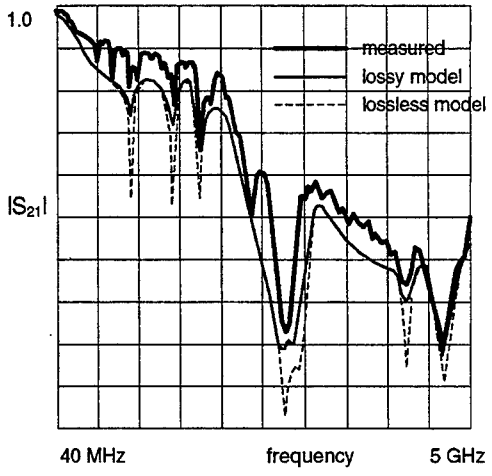


Figure 5. *Comparison of measured and simulated attenuation results for selected trace on an Pin Grid Array package.*

4. Circuit simulation and comparison with measured results

Using OptEM Package software one can assemble a model of a complex 3d interconnect. Segments of wires which are not included in the uniform transmissions lines are automatically interpolated between uniform sections or extrapolated to connection points. Separately, three dimensional non-uniformities like pins, vias and bonding wires are modeled using a three dimensional solver.

Electrically short transmission lines, like most package applications, can be modeled accurately with lumped circuits. To minimize the number of components the program combines segments by integrating their unit parameters along the wire path. However, the length of a ladder section is always carefully calculated to prevent a filtering effect of higher frequencies which may be included in the spectrum of the analyzed signals. Substantial simulation and experimental verification of OptEM software was reported by other authors [5, 6]. Figure 5 presents a comparison of measured

and simulated attenuation results for a selected trace on a Pin Grid Array package. The trace was modeled initially as a lossless transmission line. A good correlation between experimental and simulation results was reached only after conductor and dielectric losses were included in the model.

5. Conclusions

- For realistic simulation of high performance systems, one has to consider adding interconnect models to the already designed circuit.
- Since package interconnects have cross sectional dimensions within the skin depth range, an accurate solution of Helmholtz equation is required. Helmholtz equation allows for the analysis of eddy-currents, and proximity and skin effects for quasi-TEM propagation.
- Conductor and dielectric losses can be efficiently calculated together with inductance and capacitance matrices and included in an optimized package model.
- Inclusion of losses in modeling package interconnect allows very good prediction of experimental results for a wide frequency range.

References

[1] "*Circuit Analysis, Simulation and Design*," from Advances in CAD For VLSI, Vol. 3, Part 2, Ed. A. E. Ruehli, Elsevier Science Publishers B.V., Amsterdam, The Netherlands, 1987.

[2] "*OptEM Package User's Manual*," Version 3.2, OptEM Engineering Inc., Calgary, Alberta, May 1996.

[3] O. A. Palusinski, A. C. Cangellaris, J. L. Prince, J. C. Liao, and L. Vakanas, "Electrical Performance of VLSI Interconnect Systems," from "*Electronics Packaging Forum*," Volume One, Ed. J. E. Morris, Van Nostrand Reinhold, New York, 1990.

[4] J. Poltz, "On Eddy Currents in Thin Plates," *Archif fur Electrotechnik*, Springer-Verlag, Germany, Vol. 66 (4), pp. 225-229, 1983.

[5] T. K. Tang and M. S. Nakhla, "Analysis of High-Speed VLSI Interconnects Using the Asymptotic Waveform Evaluation Technique," *IEEE Transaction on Computer-Aided Design,* vol. 11 (3) , pp. 341-352, March 1992.

[6] G. Weinerth, "Aus Geometriedaten erstellte Netzwerkmodelle fur planare Leitungsgeometrien im GHz-Bereich, verifiziert an einem IC-Gehause," *Ph. D. Dissertation*, Berlin, 1997.

[7] Q. Yu and O. Wing, "Computational models of transmission lines with skin effects and dielectric losses," *IEEE Trans. on Circ. and Syst.*, vol. 41, no. 2, pp. 107-119, Feb. 1994.

Extraction of Effective Capacitance and Inductance of a Power Distribution Structure from Numerical Field Data

Andrew Byers, Bryan Boots, Melinda Piket-May
Department of Electrical and Computer Engineering
Campus Box 425, University of Colorado, Boulder,CO 80309
phone 303-492-7891, fax 303-492-5323
byersa@maori.colorado.edu
boots@ucsub.colorado.edu
mjp@colorado.edu

Roger Gravrok
Sequent Computer Systems
Bldg D2, Suite 219
800 Wisconsin Street
Eau Claire, WI 54703-3611
phone 715-833-8653, fax 715-833-2027
rjg@sequent.com

Abstract

This paper presents numerical methods for extracting the effective lumped-inductance and capacitance of a 3D power-distribution structure. The extraction techniques rely on manipulations of Maxwell's equations and the application of these relations to time domain field data. Data is obtained using the finite-difference time-domain method. Validations of the techniques presented here are provided against simple 3D structures, and are shown to be versatile in the analysis of a complex meshed power-distribution structure.

I. Introduction

DETERMINING the equivalent-circuit lumped inductance and capacitance of a 3D power-distribution structure is of critical importance to successful package- and system-level design. Physical and electrical design trade-offs can be understood and optimal performance can be achieved if these parameters are known before commitment to hardware. Several different methods have been employed to calculate both the effective inductance and capacitance of signal pathways [1] [2]. By design intent, low-inductance power-distribution structures are highly interdigitated structures with globally distributed power and return pathways. The finite inductance of these 3D structures is the result of an inherently complex interaction of 3D electromagnetic field distributions. On the other hand, the geometries of these structures are also designed to maximize system capacitance, this being a desirable attribute of a stable power-distribution structure.

Commonly, techniques that place computational simplicity as a high priority make simplifying assumptions to the solutions of Maxwell's equations. Furthermore, these techniques require either the code or the user to specify the direction(s) and distribution of the return path current(s). Each technique, including those presented in this paper, features inherent advantages and tradeoffs. In the work presented here, a priority was given to 3D modeling accuracy. The techniques explored in this paper employ simple methods that use the characteristic definitions of inductance and capacitance, as derived directly from Maxwell's

equations. These techniques quantitatively determine the natural responses directly from the field data. The user controls the tradeoff between accuracy and computational-resource usage.

II. Method Development

The effective lumped capacitance and inductance of a complex 3D structure can be calculated from the measured voltage and current responses of the system. Both of these responses can be extracted directly from the finite-difference time-domain(FDTD) field data. The FDTD field data is obtained by modeling the physical structure and applying Maxwell's equations[3]. The calculations presented below do not depend on the physical composition of the structure , but rather on the approximation-free data provided by the FDTD solver. This field data is then used to calculate the capacitance and inductance of an arbitrary 3D structure.

A. Capacitance

In this section, the method for determining capacitance will be derived. A Gaussian-shaped voltage pulse provides the input signal to the structure. This excitement sets up electric fields between the conductors and ground planes of the system. The capacitance between a particular conductive element and the ground plane is seen to be the ratio of the leakage current to the time derivative of the voltage, where the leakage current is defined as the current passing between the conductor and the ground plane[1].

$$C = \frac{I}{\frac{\delta V}{\delta t}} \tag{1}$$

Equation 1 is a direct result of the manipulation of Maxwell's Equations, where the resulting capacitance is time independent. The voltage stimulus used in this method must be time varying. The requirement that the system be composed of time independent capacitances yields data for which the leakage current and time-derivative voltage plots are identical in shape.

Intrinsic to a good capacitive structure is that it allows for as little leakage current as possible. For this and other reasons, primarily errors resulting from fringe effects and limitations of the FDTD interface, it is useful to derive an alternative representation of equation 1. Multiplying both sides of equation 1 by δV and integrating yields a relationship for which the product of the capacitance and voltage are equal to the time integral of the current. Dividing both sides by the voltage yields:

$$C = \frac{Q}{V} \tag{2}$$

This derivation takes advantage of the relationship between current and charge. It provides an equation for which the elements are easily arrived at within the confines of the FDTD solver[4]. Another benefit of equation 2 is that, as will be discussed in the validation section, it can easily be used to account for the fringe effects of a capacitive structure.

B. Inductance

In this section, a method for extracting inductance from field data will be described. A Gaussian-shaped current pulse provides the stimulus which propagates through the structure. Shorting the current path at the end of the structure simulates the operation of a PCB load that is drawing current; for example, a CMOS integrated circuit experiencing a current spike. In this state, the conduction current causes flux to develop throughout the structure, thereby emphasizing the characteristic inductance of the structure. A voltage develops across the inputs of the system in response to the dynamic current flowing through the effective inductor. This voltage response to the dynamic current stimulus is directly indicitive of the structure's inductive component, as expressed in equation 3 [1]:

$$L = \frac{V}{\frac{\delta I}{\delta t}} \tag{3}$$

Equation 3 is a simple manipulation of the characteristic equation describing the voltage (V) developed across an inductance (L) due to a $\frac{\delta I}{\delta t}$ stimulus. As in equation 1, equation 3 was arrived via manipulation of Maxwell's Equations and is conditional upon a time independent inductance. The presence of an inductive element can be inferred from the shape of the voltage response by observing the similarity to the shape of the corresponding $\frac{\delta I}{\delta t}$ stimulus.

C. Discussion

The application of Maxwell's equations by the FDTD solver fully accounts for the structure's electrical behavior. Manipulation of the data describes the physical structure's inherent response to the given electrical stimulus. Similar to the use of the transfer function in linear systems, the FDTD resultant data effectively frees the subsequent capacitance and inductance calculations from any internal geometrical considerations . In other words, this approach can be used successfully to calculate the L and C parameters of an arbitrarily-shaped 3D structure. In addition, the FDTD core solver frees the calculation methods from any potentially limiting TEM approximations, thus providing an inherently accurate 3D environment for the required simulation and parameter extraction.

III. Validation

This section provides validation of the capacitance and inductance extraction methods. Relatively simple 3D structures are modeled and simulated using FDTD, and the results are compared to analytical solutions.

A. Capacitance

A parallel plate capacitor structure was modeled with the intent of comparing capacitance values extracted from simulation data to capacitance value calculated from the following analytical equation[1]:

$$C = \frac{\epsilon A}{d} \tag{4}$$

,where ϵ is the permittivity , A is the plate area, and d is the distance between the plates. The dynamic voltage source used to stimulate the parallel plate model is located between two tabs attached at the edges of the plates. Average voltage between the plates is extracted from the FDTD data. The charge that accumulates on the top plate is also extracted. The effective lumped capacitance of the structure is then calculated from the ratio of the charge on the top plate to the voltage between the two plates, as dictated by equation 2. Table 1 below contains data, both analytical and simulated, for several different parallel plate models.

TABLE I
PARALLEL PLATE CAPACITOR RESULTS

dim	ϵ_r	C'_{theor}	C'_{sim}	difference
60x60μm	2.55	40.64fF	43.0fF	5.5%
40x40μm	2.55	18.06fF	19.6fF	7.8%
20x20μm	2.55	4.516fF	5.1fF	18%
60x60μm	100	1.60pF	1.64pF	2.4%
40x40μm	100	0.708pF	0.73fF	3.0%
20x20μm	100	0.177pF	0.19pF	6.8%

An interesting trend is the indirect relationship between plate area and the percent difference in analytical and simulated capacitances. Decreasing the plate area causes the effect of fringing fields on the capacitance value to become more pronounced. These naturally occuring effects are not accounted for in the analytical capacitance equation, due to the difficulty in characterizing the exact nature of the fringing fields from structure to structure. There are analytical methods that approximate the fringing effect, however they are not exact. The application of equation 2 in the FDTD simulation includes the contributions made by all the fields linking the plates by accounting for the total charge on the top plate. Therefore, it is reasonable to assume that capacitance values extracted from simulation data will be slightly higher than analytical values. Table 1 illustrates this general trend toward larger differences between the analytical capacitance and simulated capacitance as fringing becomes more significant. The extent of the fringing fields is also dependent on the value of the dielectric between the plates. A dielectric with higher permittivity will confine the fields between the plates, where they are accounted for by the area parameter in the analytical equation (eq. 4). As the permittivity of the dielectric decreases, the fringing fields tend to increase the effective area of the plates, thus increasing the capacitance. This trend is visible in the data presented in table 2, with the percent difference between the simulated and analytical results generally increasing as the relative permittivity decreases.

B. Inductance

The model that we analyzed in order to validate the inductance extraction method was a coaxial cable, which provided us with a closed-form theoretical solution for comparison. The

TABLE II

PARALLEL PLATE CAPACITOR: FRINGE FIELD TREND

dim	ϵ_r	C'$_{theor}$	C'$_{sim}$	difference
20x20μm	100	177fF	190fF	6.8%
20x20μm	50	88.6fF	93.4fF	5.2%
20x20μm	10	17.7fF	19fF	6.8%
20x20μm	5	8.85fF	9.7fF	8.8%
20x20μm	2.55	4.52fF	5.1fF	18.0%
20x20μm	1	1.77fF	2.2fF	19.5%

theoretical equation for the inductance of a coaxial cable is given by [5]:

$$L'(H/m) = \frac{\mu}{2\pi} * \ln \frac{b}{a} \tag{5}$$

where μ is the dielectric permeability, b is the radius of the outer conductor, and a is the radius of the inner conductor. For the simulations, μ_r was set to one, indicating an air dielectric. The coax was stimulated with a Gaussian current pulse, and the characteristic voltage response due to the inductive element of the line was extracted. Application of the inductance-calculation method to this FDTD field data yielded the results seen in table 3.

TABLE III

INDUCTANCE CALCULATIONS FOR A COAXIAL CABLE

a	b	L'$_{theor}$	L'$_{sim}$
2.5 mm	12.5 mm	0.32μH/m	0.33μH/m
3.5 mm	12.5 mm	0.25μH/m	0.26μH/m
4.5 mm	12.5 mm	0.20μH/m	0.20μH/m

The numerically modeled values of inductance per unit length closely match the corresponding theoretical values. Differences can be attributed to rounding errors or insuffucient resolution in the FDTD model.

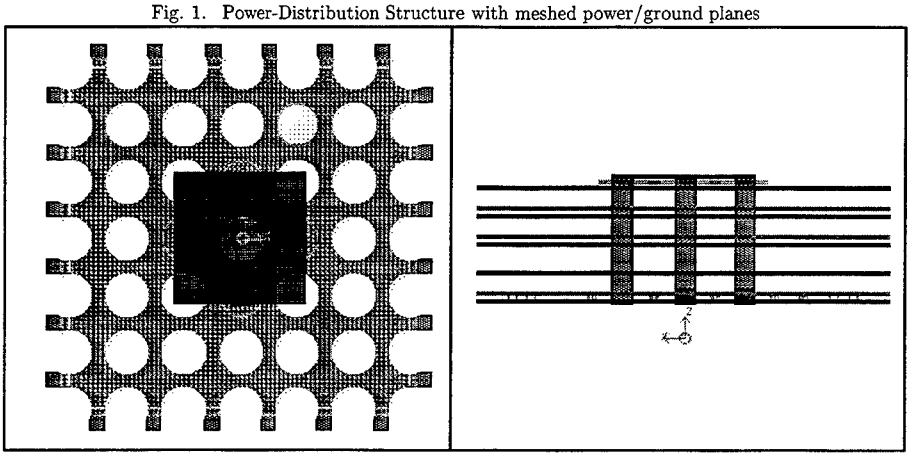
IV. APPLICATION: MESHED PCB MODEL

A. Discription

A significant problem in modeling and designing high speed digital applications has been understanding how to extract system-level parameters from complex power-distribution structures. One such model is shown in figure 1.

This structure is a meshed PCB system of three power planes with current sources at the edges, five ground planes, and nine interdigitated vias connecting the power and ground planes. The model is based on a portion of an actual PCB power-distribution structure underneath a 2499-pin, 30 mm MCM [6] with a pad pitch of 500μm, built with copper/polyimide technology similar to the memory-element MCM published in [7]. The modeled portion of the PCB covers a planar area of 3000μm by 3000μm and a via height of 1020 μm. When

Fig. 1. Power-Distribution Structure with meshed power/ground planes



the planes are meshed due to the signal-pin antipads, approximately 50% of the metal is removed.

B. Inductance extraction

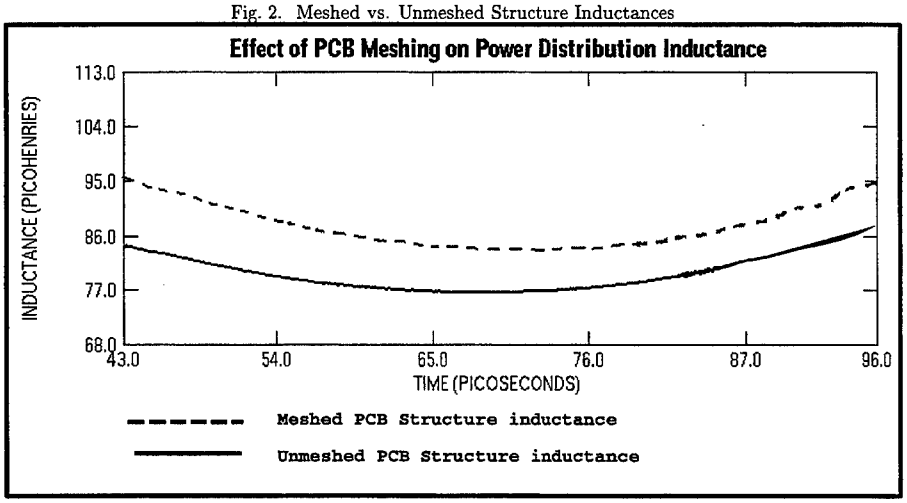
The meshed planes force the current to diverge around the antipad holes, and subsequently flux develops through the holes, increasing the total inductance of the system. It is this increased inductance that is of interest and can be determined from extracted field data. The characteristic voltage response is measured between a power plane and a ground plane. A via short to a ground plane on top of the structure provides a measurement of the dynamic stimulus current that causes a voltage to develop across the structure, essentially simulating a CMOS load. Until now, the inductance calculation for this type of structure has been done on a piecewise scale by examining the effects of the power/ground planes on the vias and vice versa. However, by applying our method based strictly on the structure's stimulus and response, it is possible to represent the entire structure as a lumped inductor with one straightforward calculation. The effective lumped-inductance of the meshed pcb structure was calculated to be approximately 87.5 pH.

The composite physical PCB implementation under the MCM repeats this 3mm by 3mm subsection 8 times in both x and y. Therefore, the composite inductance of the power distribution through the meshed pcb and clustered via arrays feeding the 2499-pin MCM would be approximately 10pH.

A way to qualitatively validate this solution is to quantitatively compare the meshed PCB model to a PCB structure without meshed planes. As was mentioned before, the meshed planes increase the system inductance. Without the meshed holes present in the planes, one would expect the inductance to decrease along with the flux. After running a simulation and extracting the required current and voltage data, the inductance of a solid-plane power-

distribution model was calculated to be approximately 79.2 pH. A comparison between the two modeled inductance results, for the meshed-plane model and unmeshed-plane model, is presented in figure 2. Using unmeshed-planes drops the overall structure inductance 9.5% from 87.5pH to 79.2pH.

Fig. 2. Meshed vs. Unmeshed Structure Inductances



C. Capacitance extraction

In contrast to the inductance, the capacitance between two planes of the structure shown in figure 1 has a relatively straight-forward analytical calculation. Accounting for the decreased planar area due to the meshing, this capacitance can be calculated using equation 4. As an excercise to test the validity of this method, and as a possible step towards the full electrical characterization of this structure (R,L,C,G), the capacitance-extraction method was applied to a sub-system of meshed layers. The sub-system of the structure shown in figure 1 consisted of two ground planes enclosing a single power plane, with the ground planes connected by vias and the power plane stimulated at the edges by Gaussian current sources. The analytical capacitance of the structure was calculated to be approximately 6pF. This value is close to the 5.6pF capacitance value extracted from an FDTD simulation of the same model. The results confirm the versatility of this capacitance extraction method for structures more complex than the parallel plate capacitor model examined in the validation section.

V. Conclusion

This paper has investigated a straightforward method for calculating the equivalent lumped inductance of a 3D power-distribution structure based on the ratio of the characteristic voltage response to a dynamic current stimulus. The subsequent dual approach for calculating

lumped-capacitance based on the ratio of characteristic current response to a dynamic voltage stimulus was also described. Validations of these extraction techniques were provided against theoretical solutions for two simple 3D structures: a parallel plate capacitor and a coaxial cable. The capacitance and inductance of a realistic power- distribution structure were also calculated from FDTD simulation field data. It is clear that this approach can be useful in determining the optimal physical parameters of a power-distribution system while staying within electrically-dictated inductance thresholds that often accompany high-speed digital applications. In addition, the effective capacitance of relatively complex structures, such as meshed parallel planes, can be determined and factored into geometrical considerations. These dual methods can be integral in the development of design guidelines for power-distribution structures similar to figure 1. Future work involving the electrical characterization of an entire 3D power-distribution structure in terms of L,C,G, and R, based on values computed using the methods developed herein, is currently being explored.

References

[1] N.N. Rao, "Elements of Engineering Electromagnetics", third edition, Prentice Hall, 1991.

[2] R. Gravrok, M. Piket-May, and K. Thomas, "LC: An Integrated Methodology to Model and Visualize Complex Electrodynamics of 3D Structures", http://www.cray.com/lc/epep.

[3] A. Taflove, "Computational Electromagnetics: The Finite-Difference Time-Domain Method", Artech House, 1995, chapters 1 - 5.

[4] A.Taflove, "Computational Electromagnetics: The Finite-Difference Time-Domain Method", Artech House, 1995, chapters 13 - 15.

[5] L.C. Shen and J.A. Kong, "Applied Electromagnetism", third edition, PWS Publishing Co., 1995, pg. 482.

[6] R.Gravrok, D. Scheid, J. Ficke, K. VanGoor, V.Rao, "Advanced MCM-Ds Break the I/O Bottleneck with Embedded Area-Array Integrated Circuits", ISHM MCM Conference, Denver, Colorado, April 17, 1996.

[7] J. Ficke, K. Hokanson, C.Berry, and R.Gravrok, "High-Bandwidth, Low-Latency 3D Memory Module", IMAPS International Journal of Microcircuits and Electronic Packaging, volume 19, number 4, pgs. 369-374, 4th quarter, 1996.

Extraction of Equivalent Circuit Parameters of Interconnections Using FDTD and PML

Feng Liu and José E. Schutt-Ainé

Department of Electrical and Computer Engineering
University of Illinois
Urbana, IL 61801

Abstract

In this paper, the finite-difference time-domain (FDTD) method is used to extract the equivalent circuit parameters of multi-conductor interconnects. The perfectly matched layer (PML) is applied to build absorbing boundary conditions (ABCs). Results are included.

1. Introduction

As the speed of high performance VLSI circuits increases, the full-wave nature of the interconnections becomes important. The wave aspects like signal distortion and unwanted signal coupling between different interconnects must be considered. One of the consequences of this is that accurate electrical modeling of these structures is necessary to insure the simulation in the design stage. This requirement can be fulfilled by a full-wave approach, which includes these effects by solving Maxwell's equations.

Among the available full-wave techniques, the finite-difference time-domain (FDTD) method is the most attractive method. The main advantage of the FDTD technique is its ability to model complicated structures. Recently, some work was performed on the extraction of equivalent circuit parameters of multi-conductors using FDTD [1-3]. But the advantages of the method were not fully exploited due to ineffective truncation of the simulation area. The recently introduced "Perfectly Matched Layer" (PML) method in theory provides unmatched performance in yielding reflectionless mesh truncation for FDTD simulation.

In this paper, the FDTD implementation of the modified Maxwell's equations with the PML is first reviewed. Then, the circuit model of multi-conductor interconnects is discussed and an extraction algorithm of equivalent circuit parameters is given. The reliability of the method is verified by comparison with other methods.

2. PML-FDTD Formulation

The finite-difference time domain (FDTD) method is a versatile technique for the full-wave simulation of electromagnetic phenomena governed by Maxwell's equations. Since its introduction by Yee [4], the FDTD method has been applied to many problems. The main challenge for FDTD is in the implementation of the absorbing boundary conditions (ABCs) at the edges of the FDTD grid. Another concern is the size of the computational domain (or the memory requirement for the simulation) for unbounded EM problems, which is determined by the type of mesh truncation.

The recently introduced "Perfectly Matched Layer" (PML) [5] in theory provides reflectionless absorption of EM waves independently of frequency or angle of incidence. In addition, the PML provides unmatched performance in the ability to provide reflectionless mesh truncation for FDTD simulation. The ability to absorb outgoing waves is provided by additional degrees of freedom introduced by a split field formulation with anisotropic electric and magnetic conductivities.

The PML is an artificial lossy medium, which is characterized by electrical conductivity σ and magnetic conductivity σ^*. They are related to each other as follows:

$$\frac{\sigma}{\varepsilon} = \frac{\sigma^*}{\mu} \tag{1}$$

This relationship ensures that the wave impedance of the PML medium is matched to that of the adjacent physical medium. The modified Maxwell's equations are [6]:

$$\begin{aligned}
\mu\frac{\partial \mathbf{H}_{sx}}{\partial t} + \sigma_x^* \mathbf{H}_{sx} &= -\frac{\partial}{\partial x}\hat{x}\times\mathbf{E} \\
\mu\frac{\partial \mathbf{H}_{sy}}{\partial t} + \sigma_y^* \mathbf{H}_{sy} &= -\frac{\partial}{\partial y}\hat{y}\times\mathbf{E} \\
\mu\frac{\partial \mathbf{H}_{sz}}{\partial t} + \sigma_z^* \mathbf{H}_{sz} &= -\frac{\partial}{\partial z}\hat{z}\times\mathbf{E}
\end{aligned} \tag{2}$$

and

$$\varepsilon\frac{\partial \mathbf{E}_{sx}}{\partial t}+\sigma_x\mathbf{E}_{sx}=\frac{\partial}{\partial x}\hat{x}\times\mathbf{H}$$
$$\varepsilon\frac{\partial \mathbf{E}_{sy}}{\partial t}+\sigma_y\mathbf{E}_{sy}=\frac{\partial}{\partial y}\hat{y}\times\mathbf{H}$$
$$\varepsilon\frac{\partial \mathbf{E}_{sz}}{\partial t}+\sigma_z\mathbf{E}_{sz}=\frac{\partial}{\partial z}\hat{z}\times\mathbf{H} \qquad (3)$$

where $\mathbf{E}=\mathbf{E}_{sx}+\mathbf{E}_{sy}+\mathbf{E}_{sz}$ and $\mathbf{H}=\mathbf{H}_{sx}+\mathbf{H}_{sy}+\mathbf{H}_{sz}$. Note that $\mathbf{E}_{si}$, and $\mathbf{H}_{si}$, $i=x,y,z$ are two-component vectors. The above equations contain twelve scalar equations with twelve split field unknowns. Let $\mathbf{H}_{sx}=\hat{y}H_{sxy}+\hat{z}H_{sxz}$, $\mathbf{E}=\hat{x}E_x+\hat{y}E_y+\hat{z}E_z$, and substitute them into (2). By equating the z component, we have

$$\mu\frac{\partial H_{sxz}}{\partial t}+\sigma_x^* H_{sxz}=-\frac{\partial}{\partial x}Ey \qquad (4)$$

Applying the same procedure, we can obtain 12 equations. The FDTD implementation of these equations on a Yee grid is straightforward.

The wave propagation phenomenon in the perfectly matched medium is very similar to that described by Maxwell's equations with the exception that attenuation may be controlled through σ_i, $i=x,y,z$. The degrees of freedom supplied by the anisotropic medium allows one to control the attenuation of individual component of the fields. The absorbing boundaries at the edges of the simulation region can be created by choosing appropriate values of σ_i, $i=x,y,z$. In practice, abrupt changes in conductivity from free-space to the PML medium cause large reflections at air-PML interface due to the errors introduced by the numerical discretization. Thus, the smooth conductivity profiles that increases from 0 at the air-PML interface to $\sigma_{\max}$ at the PEC termination (the fields are nearly zero at the end of the PML) are assigned.

3. Circuit Model

Consider a general multiconductor transmission line structure with N conductors. For a quasi-TEM mode of propagation, this structure can be considered as a guided-wave system and can be described by the distributed circuit parameters of the transmission lines. The voltage and current on the transmission lines satisfy the generalized frequent dependent Telegrapher's equation [7]:

$$-\frac{d}{dz}\mathbf{V}(z,\omega) = j\omega\mathbf{L}(\omega)\mathbf{I}(z,\omega)$$
$$-\frac{d}{dz}\mathbf{I}(z,\omega) = j\omega\mathbf{C}(\omega)\mathbf{V}(z,\omega) \tag{5}$$

where $\mathbf{I}(z,\omega)$ and $\mathbf{V}(z,\omega)$ are current and voltage vectors. $\mathbf{L}(\omega)$ and $\mathbf{C}(\omega)$ are inductance and capacitance matrices respectively. We have neglected the dissipation of the interconnects. The $\mathbf{L}(\omega)$ and $\mathbf{C}(\omega)$ matrices can be obtained by

$$\mathbf{L}(\omega) = -\frac{1}{jw}\left(\frac{d}{dz}\mathbf{V}(z,\omega)\mathbf{I}(z,\omega)^{-1}\right)$$
$$\mathbf{C}(\omega) = -\frac{1}{jw}\left(\frac{d}{dz}\mathbf{I}(z,\omega)\mathbf{V}(z,\omega)^{-1}\right) \tag{6}$$

The current and voltage propagating along each line in the time domain can be calculated from the electromagnetic fields which can be obtained from the FDTD by:

$$i(z,t) = \oint_c \overline{H} \bullet \overline{dl}$$
$$v(z,t) = \int \overline{E} \bullet \overline{dl} \tag{7}$$

where the contour path for $v(z,t)$ extends from the ground plane to the line, while c is the transverse contour of the line. Then, $\mathbf{V}(z,\omega)$ and $\mathbf{I}(z,\omega)$ will be calculated from the FFT of $v(z,t)$ and $i(z,t)$.

Here, we summarize our extraction algorithm. First the EM fields are simulated by the FDTD algorithm, where the PML is used as ABC, next the fields at sample positions are recorded. The currents and the voltages are calculated from the integrations based on the recorded fields. Finally, $\mathbf{L}(\omega)$ and $\mathbf{C}(\omega)$ matrices can be calculated by equation (6).

4. Numerical Results

In this section, the method is used to analyze a microstrip structure with symmetric coupled lines [8]. The geometrical parameters are shown in Fig. 1. The number of grids is $60\times27\times135$ with space $\Delta x = \Delta y = \Delta z = 5\times10^{-5} m$. The time step is $\Delta t = 7.5\times10^{-14} s$, which is restricted by the Courant stability condition. Twelve cells of PML are used to build the absorbing boundary in each direction. The conductivity values were chosen with a parabolic profiles.

First, we validate our FDTD-PML code by exciting the fields with a single frequency source (frequency = 18 GHz). $N_t = 6000$ time steps are computed. From Fig. 2, we observe that the recorded field exhibits an excellent match with the source.

For the extraction, the fields are excited by a Gaussian source:

$$J_x = \frac{1}{\Delta x \Delta y \Delta z} e^{-(t-t_0)^2/T^2} \tag{8}$$

where the Gaussian half-width is $T = 100\,\Delta t$ and the delay time t_0 is set to $500\,\Delta t$. $N_t = 1500$ time steps were computed to insure that the Gaussian pulse pass the sample completely.

The $\mathbf{L}(\omega)$ and $\mathbf{C}(\omega)$ extracted by the method described above are shown in Fig. 3. Our results are in good agreement with those in [3].

5. Conclusion

In this paper, the FDTD and PML are used to simulate the electromagnetic fields in the time domain. Based on the fields recorded, a standard method for the extraction of the equivalent circuit parameters of coupled transmission lines are presented.

6. References:

[1] R. Mittra, W. D. Becker, and P. H. Harms, " A general purpose Maxwell solver for the extraction of equivalent circuits of electric package component for circuit simulation." IEEE Trans. Circuits Syst. I, vol.39, pp. 964-973, Nov. 1992.

[2] T. Dhaene, S. Criel, and D. D. Zutter, " Analysis and modeling of coupled dispersive interconnection lines," IEEE Trans. MTT., vol40, pp. 2103-2105, Nov. 1992

[3] J. Zhao and Z. F. Li, " A time-domain full-wave extraction method of frequency-dependent equivalent circuit parameters of multiconductor interconnection lines", IEEE Trans. MTT, vol. 45, pp. 23-31, Jan. 1997

[4] K. S. Yee, " Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media", IEEE Trans. Antennas Propagat., vol. 14, pp. 302-307, May 1966

[5] J. P. Berenger, " A perfectly matched layer for the absorption of electromagnetic waves", J. Computational Physics, vol. 144, pp. 185-200, Oct. 1994

[6] W. C. Chew and W. H. Weedon, " A 3D perfectly matched medium from modified Maxwell's Equations with Stretched Coordinates", Micro. Opt. Tech. Lett., vol. 7, no. 13, pp. 599-604, Sept. 1994

[7] T. Dhaene and D. De Zutter, " CAD-oriented general circuit description of uniform coupled lossy dispersive waveguide structures, " IEEE Trans. MTT., vol. 40, pp. 1545-1559, July 1992.

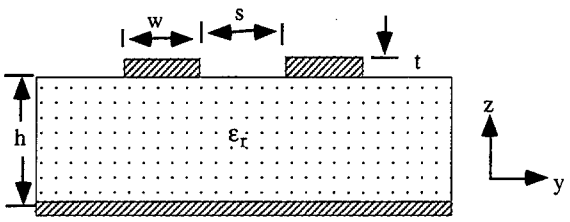


Fig. 1. Cross-section of coupled line microstrip system. The perfectly conducting strips have thickness t=0.05mm, width w= 0.3mm, and separation s=0.3mm. The substrate height is h=0.25mm and has a relative dielectric constant of 4.5. The bottom of the substrate is grounded.

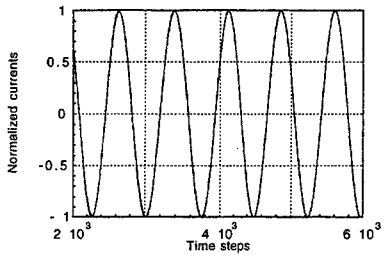


Fig 2. The normalized current at the source (dash line) and sample (solid line) position. The source current has been shifted to compensate for the propagation delay.

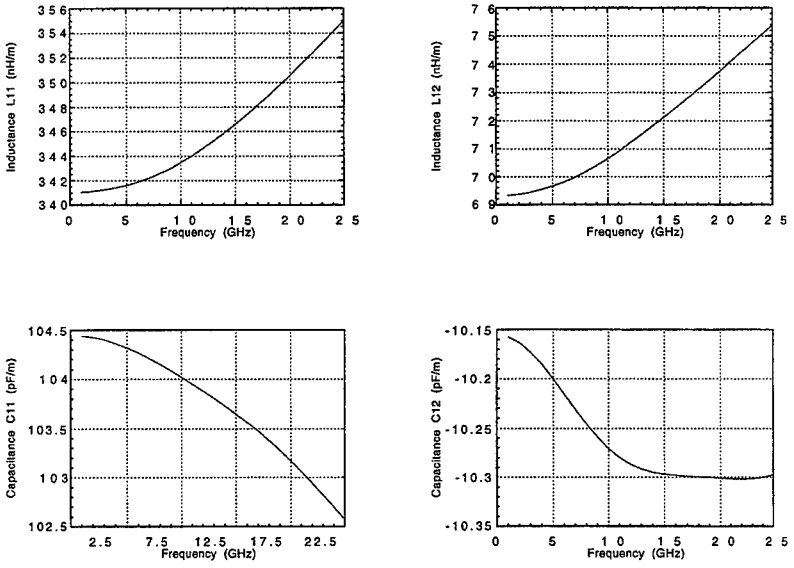


Fig 3. The extracted equivalent circuit parameters of the coupled transmission line system as a function of frequency.

SESSION 12:

REDUCED-ORDER MODELING IN ELECTROMAGNETICS

Chairs: A.C. Cangellaris and R.F. Remis

The Use of a Correspondence Principle in Reduced-order Modeling of Electromagnetic Wavefields

R.F. Remis and **P.M. van den Berg**
Laboratory of Electromagnetic Research, Faculty of Electrical Engineering,
Centre for Technical Geoscience, Delft University of Technology,
P.O. Box 5031, 2600 GA Delft, The Netherlands

1. Introduction

The transient electromagnetic wavefield in an inhomogeneous and lossy medium can be computed efficiently by constructing reduced-order approximations to the electromagnetic wavefield quantities. Since these approximations are based on Maxwell's equations as a system of first-order partial differential equations, they exhibit a specific structure and for the case of a lossless medium we can take advantage of this structure in the following way. Given a reduced-order approximation for the electromagnetic wavefield present in a lossless medium, we can use this approximation to describe the behavior of the electromagnetic field in a corresponding class of lossy media as well. The connection between this class of lossy media and the lossless medium is described by a correspondence principle. If it is satisfied, is not necessary to start the computations all over again. Only a slight modification of the reduced-order approximations is needed. Some numerical examples, for two-dimensional configurations, will illustrate this point.

2. Basic equations

The pointwise behavior of the electromagnetic field, present in an inhomogeneous, anisotropic and lossy medium, is governed by Maxwell's equations written here in the form

$$(\mathcal{D} + \mathcal{M}_1 + \mathcal{M}_2\partial_t)\mathcal{F} = \mathcal{Q}', \tag{1}$$

where $\mathcal{F} = \mathcal{F}(\boldsymbol{x}, t)$ is the field vector consisting of the components of the electric field strength $\boldsymbol{E}$ and the magnetic field strength $\boldsymbol{H}$ as

$$\mathcal{F} = [E_1, E_2, E_3, H_1, H_2, H_3]^T, \tag{2}$$

and $\mathcal{Q}' = \mathcal{Q}'(\boldsymbol{x}, t)$ is the source vector composed of the components of the external electric-current sources $\boldsymbol{J}^{\mathrm{e}}$ and the external magnetic-current sources $\boldsymbol{K}^{\mathrm{e}}$ as

$$\mathcal{Q}' = -[J_1^{\mathrm{e}}, J_2^{\mathrm{e}}, J_3^{\mathrm{e}}, K_1^{\mathrm{e}}, K_2^{\mathrm{e}}, K_3^{\mathrm{e}}]^T. \tag{3}$$

The time-independent matrices $\mathcal{M}_1$ and $\mathcal{M}_2$ are medium matrices given by

$$\mathcal{M}_1 = \begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} & \sigma_{1,3} & 0 & 0 & 0 \\ \sigma_{2,1} & \sigma_{2,2} & \sigma_{2,3} & 0 & 0 & 0 \\ \sigma_{3,1} & \sigma_{3,2} & \sigma_{3,3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{4}$$

and

$$\mathcal{M}_2 = \begin{pmatrix} \varepsilon_{1,1} & \varepsilon_{1,2} & \varepsilon_{1,3} & 0 & 0 & 0 \\ \varepsilon_{2,1} & \varepsilon_{2,2} & \varepsilon_{2,3} & 0 & 0 & 0 \\ \varepsilon_{3,1} & \varepsilon_{3,2} & \varepsilon_{3,3} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_{1,1} & \mu_{1,2} & \mu_{1,3} \\ 0 & 0 & 0 & \mu_{2,1} & \mu_{2,2} & \mu_{2,3} \\ 0 & 0 & 0 & \mu_{3,1} & \mu_{3,2} & \mu_{3,3} \end{pmatrix}. \tag{5}$$

Using energy considerations it can be shown that the permittivity tensor $\varepsilon_{i,j} = \varepsilon_{i,j}(\boldsymbol{x})$ and the permeability tensor $\mu_{i,j} = \mu_{i,j}(\boldsymbol{x})$ are symmetric and positive definite. Moreover, the conductivity tensor $\sigma_{i,j} = \sigma_{i,j}(\boldsymbol{x})$ is positive semidefinite and is taken to be symmetric.

The spatial derivatives are contained in the spatial differential operator matrix $\mathcal{D}$ given by

$$\mathcal{D} = \begin{pmatrix} 0 & 0 & 0 & 0 & \partial_3 & -\partial_2 \\ 0 & 0 & 0 & -\partial_3 & 0 & \partial_1 \\ 0 & 0 & 0 & \partial_2 & -\partial_1 & 0 \\ 0 & -\partial_3 & \partial_2 & 0 & 0 & 0 \\ \partial_3 & 0 & -\partial_1 & 0 & 0 & 0 \\ -\partial_2 & \partial_1 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{6}$$

In addition, we introduce the matrices δ^{E} and δ^{H} as

$$\delta^{\mathrm{E}} = \mathrm{diag}(1,1,1,0,0,0) \tag{7}$$

and

$$\delta^{\mathrm{H}} = \mathrm{diag}(0,0,0,1,1,1). \tag{8}$$

These matrices reveal the structure that is present in Maxwell's equations. For example, from the equations

$$\mathcal{D}\delta^{\mathrm{E}} = \delta^{\mathrm{H}}\mathcal{D} \tag{9}$$

and

$$\mathcal{D}\delta^{\mathrm{H}} = \delta^{\mathrm{E}}\mathcal{D} \tag{10}$$

it follows that when matrix $\mathcal{D}$ operates on a vector proportional to the electric field strength, a vector proportional to the magnetic field strength results and vice versa. Other relations, involving the medium matrices and the matrices δ^{E} and δ^{H}, are

$$\mathcal{M}_1\delta^{\mathrm{E}} = \delta^{\mathrm{E}}\mathcal{M}_1 = \mathcal{M}_1, \tag{11}$$

$$\mathcal{M}_1\delta^{\mathrm{H}} = \delta^{\mathrm{H}}\mathcal{M}_1 = 0 \tag{12}$$

and

$$\mathcal{M}_2\delta^{\mathrm{E}} = \delta^{\mathrm{E}}\mathcal{M}_2, \tag{13}$$

$$\mathcal{M}_2\delta^{\mathrm{H}} = \delta^{\mathrm{H}}\mathcal{M}_2. \tag{14}$$

Now, let the source vector be of the form $\mathcal{Q}'(\boldsymbol{x},t) = w(t)\mathcal{Q}(\boldsymbol{x})$, where $w(t)$ is the source wavelet that vanishes for $t<0$ and $\mathcal{Q}$ is a time-independent vector. The source vector is said to be of the electric-current type if the vector $\mathcal{Q}$ satisfies $\mathcal{Q} = \delta^{\mathrm{E}}\mathcal{Q}$ and of the magnetic-current type if this vector satisfies $\mathcal{Q} = \delta^{\mathrm{H}}\mathcal{Q}$. Since the source wavelet vanishes prior to the time instant $t = 0$ the field vector must do so as well because of causality.

Applying a one-sided Laplace transformation to Eq. (1) with respect to time results in the equation

$$(\mathcal{D} + \mathcal{M}_1 + s\mathcal{M}_2)\hat{\mathcal{F}}(\boldsymbol{x}, s) = \hat{w}(s)\mathcal{Q}(\boldsymbol{x}), \tag{15}$$

with $\mathrm{Re}(s) > 0$. In our further analysis we take s real and positive. Then, Lerch's theorem (see *Widder* [1]) ensures that there is a one-to-one correspondence between a causal time function and its Laplace-transform-domain counterpart, provided that the time function is continuous and is, at most, of exponential growth as $t \to \infty$ and that equality in the definition of the Laplace transform is invoked at the real set of points $\{s_n = s_0 + nh; n = 0, 1, 2, \cdots\}$, where s_0 is sufficiently large and positive and h is positive.

As a next step, we discretize in space in such a way that Eqs. (9)-(14), valid in the continuous context, have a counterpart after discretization. A simple discretization procedure that satisfies this requirement is the standard finite-difference technique of *Yee* [2]. In addition we employ a homogeneous Dirichlet boundary condition. The discrete counterparts of $\mathcal{D}, \mathcal{M}_1, \mathcal{M}_2, \mathcal{F}$ and $\mathcal{Q}$ are given by D, M_1, M_2, F and Q, respectively. The discrete counterparts of the matrices δ^{E} and δ^{H} are denoted by the same symbols. After this discretization procedure, we obtain the algebraic matrix equation

$$(D + M_1 + sM_2)\hat{F}(s) = \hat{w}(s)Q, \tag{16}$$

with s real and positive. All the matrices occurring in this equation are square and of order n; matrix D is real and anti-symmetric and the medium matrices M_1 and M_2 are both symmetric, M_2 being positive definite and M_1 being positive semidefinite.

3. Reduced-order modeling of electromagnetic wavefields in a lossless medium

Matrix M_1 vanishes in case of a lossless medium and Eq. (16) simplifies to

$$(D + sM_2)\hat{F}(s) = \hat{w}(s)Q, \tag{17}$$

with s real and positive. An expression for the field vector $\hat{F}(s)$ may be obtained from Eq. (17) as

$$\hat{F}(s) = \hat{w}(s)(A + sE)^{-1}M_2^{-1}Q, \tag{18}$$

where we have introduced the identity matrix E and matrix A as

$$A = M_2^{-1}D. \tag{19}$$

Because of Lerch's theorem, a unique and causal time-domain counterpart corresponds to the s-domain expression for the field vector $\hat{F}(s)$. Via inspection this time-domain field vector is found as

$$F(t) = w(t) * \chi(t)\exp(-At)M_2^{-1}Q, \tag{20}$$

where $\chi(t)$ is the Heaviside unit step function and $*$ denotes convolution in time.

Computing the field vector by using, for example, the spectral decomposition of matrix A is not feasible due to the large size of this matrix. For example, in a three-dimensional configuration the order of matrix A can easily become as large as 10^6 or even larger. Our approach is therefore to construct approximations to the field vector, the so-called reduced-order approximations, that are all based on a Lanczos algorithm. In the next subsection we will briefly describe this algorithm and we will introduce the reduced-order approximations.

3.1 The Lanczos algorithm and the reduced-order approximations

Let $\langle \cdot, \cdot \rangle$ denote the standard inner product of two vectors from $\mathbb{R}^n$. It is easily verified that matrix A is anti-symmetric with respect to the inner product $\langle M_2 \cdot, \cdot \rangle$, that is, matrix A satisfies

$$\langle M_2 Ax, y \rangle = -\langle M_2 x, Ay \rangle, \quad \text{for all } x, y \in \mathbb{R}^n. \tag{21}$$

This property allows us to carry out the following Lanczos algorithm with matrix A,

$$\begin{aligned} \beta_1 v_1 &= M_2^{-1} Q, \\ \beta_{i+1} v_{i+1} &= A v_i + \beta_i v_{i-1}, \quad \text{for } i = 1, 2, ..., \end{aligned} \tag{22}$$

with $v_0 = 0$. The coefficients $\beta_i \geq 0$ are determined from the condition $\langle M_2 v_i, v_i \rangle = 1$ for $i \geq 1$. After m steps of this algorithm the summarizing equation

$$A V_m = V_m T_m + \beta_{m+1} v_{m+1} e_m^T, \tag{23}$$

holds. In this equation, the n-by-m matrix V_m has the column partitioning $V_m = (v_1, v_2, \cdots, v_m)$, matrix T_m is a real, tridiagonal and anti-symmetric m-by-m matrix containing the recurrence coefficients and is given by $T_m = \text{tridiag}(\beta_i, 0, -\beta_{i+1})$ and e_m is the mth column of the m-by-m identity matrix. We are interested in situations where m is much smaller than the order of matrix A.

Using Eq. (23), and not the orthogonality of the Lanczos vectors v_i with respect to the inner product $\langle M_2 \cdot, \cdot \rangle$, we can construct the reduced-order approximations (see *Remis and Van den Berg* [3])

$$F_m(t) = w(t) * \chi(t) \beta_1 V_m \exp(-T_m t) e_1. \tag{24}$$

It can be shown that the number of iterations needed to obtain an accurate result on the time interval $(0, t_{\text{obs}}]$ is proportional to $\|A\| t_{\text{obs}}$, where $\| \cdot \|$ is the matrix 2-norm.

For source vectors of the electric- or of the magnetic-current type, the Lanczos vectors, as generated by the algorithm described above, are highly structured. For example, assume that the source vector is of the electric-current type. Then, after m steps of the Lanczos algorithm we have

$$v_i = \begin{cases} \delta^{\text{E}} v_i, & \text{when } i \text{ is odd}, \\ \delta^{\text{H}} v_i, & \text{when } i \text{ is even}, \end{cases} \tag{25}$$

and for a source vector is of the magnetic-current type the Lanczos vectors satisfy

$$v_i = \begin{cases} \delta^{\text{H}} v_i, & \text{when } i \text{ is odd}, \\ \delta^{\text{E}} v_i, & \text{when } i \text{ is even}. \end{cases} \tag{26}$$

These properties can be proved by an easy induction using the recursion of Eq. (22) and the observation that matrix A satisfies $A\delta^{\text{E}} = \delta^{\text{H}} A$ and $A\delta^{\text{H}} = \delta^{\text{E}} A$.

From Eq. (25) (Eq. (26)) we infer that in case of a source vector of the electric-current type (magnetic-current type), the odd (even) numbered Lanczos vectors built up the reduced-order approximation to the electric field strength, while the even (odd) numbered vectors built up the reduced-order approximation to the magnetic field strength. Since we want the reduced-order approximations to equally update the electric and the magnetic field strength components, we will always carry out an even number of Lanczos steps.

Now, let us introduce the diagonal m-by-m matrices $\delta_m^{(1,0)}$ and $\delta_m^{(0,1)}$ as

$$\delta_m^{(1,0)} = \text{diag}(1, 0, \cdots, 1, 0) \tag{27}$$

and

$$\delta_m^{(0,1)} = \text{diag}(0, 1, \cdots, 0, 1). \tag{28}$$

Then, because of Eq. (25), we may write

$$\delta^{\text{E}} V_m = V_m \delta_m^{(1,0)}, \tag{29}$$

$$\delta^{\text{H}} V_m = V_m \delta_m^{(0,1)}, \tag{30}$$

for a source-vector of the electric-current type and

$$\delta^{\mathrm{E}} V_m = V_m \delta_m^{(0,1)}, \tag{31}$$

$$\delta^{\mathrm{H}} V_m = V_m \delta_m^{(1,0)}, \tag{32}$$

for a source vector of the magnetic-current type. These equations show that matrix $\delta_m^{(1,0)}$ describes the connection between an electric-current source and the electric field strength or the connection between a magnetic-current source and the magnetic field strength. Similarly, matrix $\delta_m^{(0,1)}$ describes the connection between an electric-current source and the magnetic field strength or the connection between a magnetic-current source and the electric field strength.

4. Reduced-order modeling of electromagnetic fields in a lossy medium using a correspondence principle

In a lossy medium, the electromagnetic field is described by the field vector

$$F(t) = w(t) * \chi(t) \exp(-\tilde{A}t) M_2^{-1} Q, \tag{33}$$

where matrix $\tilde{A}$ is given by

$$\tilde{A} = M_2^{-1}(D + M_1). \tag{34}$$

For general lossy media we can construct reduced-order approximations to this field vector by employing a modified Lanczos scheme, see *Remis and Van den Berg* [4]. However, for a special class of media, that is, media that satisfy a certain correspondence principle, we can compute reduced-order approximations to the electromagnetic field in a lossy medium by using results obtained for the lossless case. The details are as follows.

Consider a lossy medium with a conductivity that satisfies the correspondence principle (*De Hoop* [5])

$$\sigma_{i,j} = \xi \varepsilon_{i,j}, \tag{35}$$

where ξ is a constant. Note that the dimension of ξ is the reciprocal of time. A consequence of this equation is that matrix $\tilde{A}$ can be written as

$$\tilde{A} = A + \xi \delta^{\mathrm{E}}, \tag{36}$$

where matrix A is given by Eq. (19). Using the results of the previous section and assuming a source-vector of the electric-current type we may write

$$\begin{aligned} \tilde{A} V_m &= (A + \xi \delta^{\mathrm{E}}) V_m = A V_m + \xi \delta^{\mathrm{E}} V_m \\ &= V_m T_m + \beta_{m+1} v_{m+1} e_m^T + \xi V_m \delta_m^{(1,0)} \\ &= V_m \tilde{T}_m + \beta_{m+1} v_{m+1} e_m^T, \end{aligned} \tag{37}$$

with $\tilde{T}_m = T_m + \xi \delta_m^{(1,0)}$.

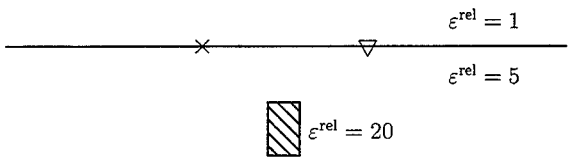


Figure 1. Source and receiver located at the interface in a lossless configuration. The distance between the source and the receiver is 4.84 m. An object of 0.88 m × 1.98 m is present in the lower halfspace. The top of the object is located 1.98 m below the interface.

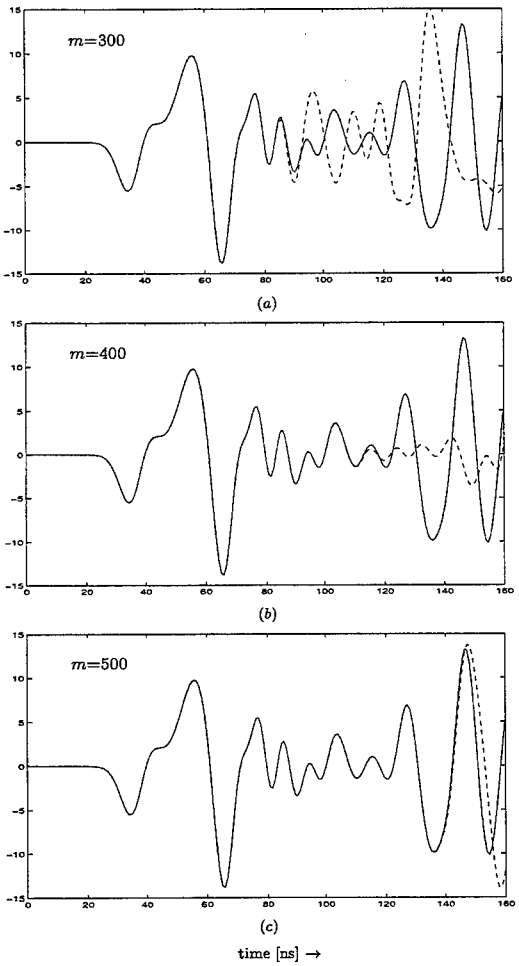


Figure 2. Reduced-order approximations to the electric field strength E_2 at the observation point after 300 iterations (a), 400 iterations (b) and 500 iterations (c). The solid line is the converged reduced-order approximation.

Similarly, for a source vector of the magnetic-current type we have

$$\tilde{A}V_m = V_m\tilde{T}_m + \beta_{m+1}v_{m+1}e_m^T, \tag{38}$$

with $\tilde{T}_m = T_m + \xi\delta_m^{(0,1)}$.

The reduced-order approximation for the electromagnetic field in a medium that satisfies the correspondence principle is now given by

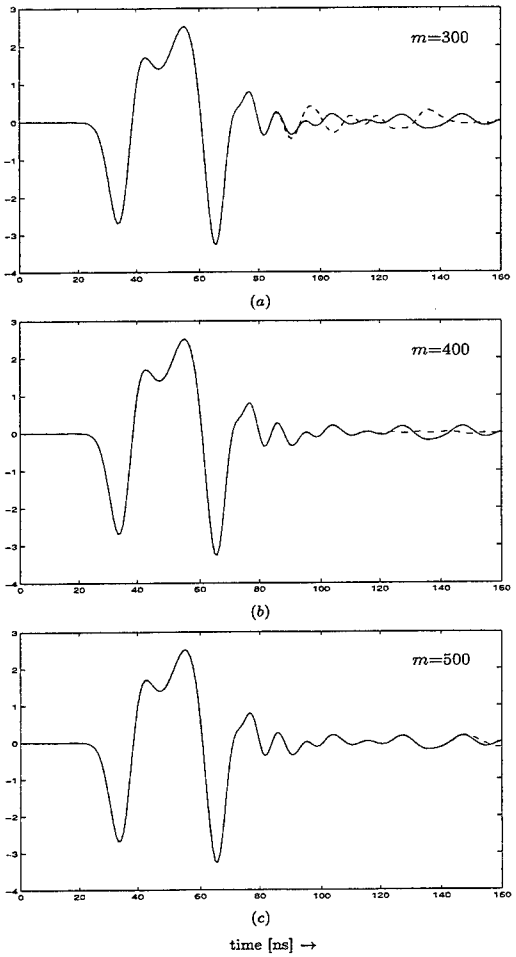


Figure 3. Reduced-order approximations to the electric field strength E_2 at the observation point after 300 iterations (a), 400 iterations (b) and 500 iterations (c). Results were obtained using the correspondence principle. The solid line is the reduced-order approximation obtained by using the modified Lanczos algorithm.

$$F_m(t) = w(t) * \chi(t)\beta_1 V_m \exp(-\tilde{T}_m t)e_1. \tag{39}$$

The only difference between this reduced-order approximation and the one for the lossless case is a difference between the diagonals of the matrices T_m and $\tilde{T}_m$. The reduced-order approximation to the electromagnetic wavefield in a lossy medium follows immediately from the reduced-order

approximation for the corresponding lossless case by adding the appropriate diagonal to matrix T_m, provided that the correspondence principle of Eq. (35) is satisfied. For this special case it is not necessary to start the computations all over again.

5. Numerical Examples

Consider the two-dimensional configuration of Figure 1 in which an inhomogeneous, isotropic and lossless medium is present. The configuration is invariant in the x_2-direction and the x_3-direction points downwardly. An electric-current source excites E-polarized waves and is given by

$$J_1^e(x_1, x_3, t) = 0, \quad J_2^e(x_1, x_3, t) = w(t)\delta(x_1 - x_1^{\text{src}}, x_3 - x_3^{\text{src}}), \quad J_3^e(x_1, x_3, t) = 0. \tag{40}$$

The source wavelet is taken to be a Ricker wavelet and is given by

$$w(t) = \chi(t)\sqrt{\frac{e}{2\theta}}\frac{\mathrm{d}}{\mathrm{d}t}\exp[-\theta(t - t_0)^2]. \tag{41}$$

The parameter θ is chosen such that the peak frequency of this wavelet is 40 MHz.

The solid line in Figure 3 shows the converged reduced-order approximation to the electric field strength E_2 at the observation point. The dashed line in Figure 3*a* is the reduced-order approximation after 300 iterations, in Figure 3*b* after 400 iterations and in Figure 3*c* after 500 iterations. We observe that by increasing the number of iterations, the approximations become more accurate on an increasing time interval.

Now consider a corresponding lossy medium characterized by correspondence constant $\xi = 6 \cdot 10^{-4}\,\varepsilon_0^{-1}\,\mathrm{s}^{-1}$. In this medium, the upper halfspace has a conductivity of $6 \cdot 10^{-4}$ S/m, the lower halfspace a conductivity of $3 \cdot 10^{-3}$ S/m and the object has a conductivity of $1.2 \cdot 10^{-2}$ S/m. Using the reduced-order approximations of the previous example and adding the appropriate diagonal to matrix T_m immediately gives the results as shown in Figure 3. The extra work that is needed to obtain these approximations is negligible. In fact, one can study a whole class of corresponding media by considering several values for ξ at very little extra cost.

Conclusions

Reduced-order approximations to the electromagnetic wavefield in a lossless medium exhibit a particular structure in case electric- or magnetic-current sources are considered. We have shown that, because this structure, these approximations can also be used to describe the behavior of the electromagnetic field in media that satisfy a correspondence principle. Only a slight modification of the reduced-order approximations is necessary requiring very little extra work.

References

[1] D.V. WIDDER, *The Laplace Transform*, Princeton Unviversity Press, 1946.

[2] K.S. YEE, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Transactions on Antennas and Propagation*, vol. 14, no.3, 1966, pp. 302-307.

[3] R.F. REMIS AND P.M. VAN DEN BERG, "A modified Lanczos algorithm for the computation of transient electromagnetic wavefields," accepted for publication in *IEEE Transactions on Microwave Theory and Techniques.*

[4] R.F. REMIS AND P.M. VAN DEN BERG, "Computation of transient electromagnetic wavefields in inhomogeneous media using a modified Lanczos algorithm," *Conference Proceedings of the 13th Annual Review of Progress in Applied Computational Electromagnetics*, Naval Postgraduate School, Monterey, CA., 1997, pp. 141-149.

[5] A.T. DE HOOP, "A general correspondence principle for time-domain electromagnetic wave and diffusion fields," *Geophysical Journal International*, 127, 1996, pp. 757-761.

The Spectral Lanczos Decomposition Method for Efficient Time-Domain and Frequency-Domain Finite-Element Solution of Maxwell's Equations

Mohammad Zunoubi, Jian-Ming Jin, Kalyan Donepudi, and Weng Chow Chew
Center for Computational Electromagnetics
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801-2991
zunoubi@sunchew.ece.uiuc.edu

Abstract

An efficient three-dimensional solver that combines the spectral Lanczos decomposition method (SLDM) and the finite-element method (FEM) is described for the solution of Maxwell's equations in both time and frequency domains. The FEM based on Whitney forms is used to discretize Maxwell's equations and the resultant matrix equation is solved using the SLDM. Our technique is an implicit, unconditionally stable finite-element time- and frequency-domain scheme that requires the implementation of the Lanczos process only at the largest frequency or time of interest. Therefore, a multiple time- and frequency-domain analysis of the electromagnetic fields is performed with minimal amount of extra computing time. We illustrate the efficiency, validity, and accuracy of this new method by considering numerical examples of an air-filled and a partially-loaded lossy dielectric cavity.

1 Introduction

The finite-difference time-domain technique (FDTD) [1] has become the most popular technique for the time-domain analysis of the electromagnetic fields. Although it has been applied to numerous scattering and radiating problems, in its original form, it suffers from staircasing approximation when modeling curvatures. To overcome this problem, hybrid FDTD formulations have been introduced which can conform to the surfaces of all boundaries in the solution domain [2-5]. An alternative and rather robust approach is to employ the time-domain finite-element methods (TD-FEM). The TD-FEM techniques can be categorized by either being explicit [6-8] or implicit [9] time-domain schemes. The TD-FEM explicit methods are only conditionally stable with time steps that are typically equal to or smaller than those imposed by Courant's stability criterion. The implicit schemes, on the other hand, are unconditionally stable while requiring the solution of a matrix equation for every time step. Therefore, for implicit methods to be computationally as efficient as the explicit techniques, either the number of iterations must be very small owing to the fact that a system of equations is solved at each time step, or an efficient procedure for treating the system should be developed.

Recently, Zunoubi *et al* [10,11] have employed the spectral Lanczos decomposition method (SLDM) [12] to analyze the axisymmetric and three-dimensional diffusion problems. It has been illustrated that accurate results can be obtained at many frequencies with a negligible amount of extra computing time while the Lanczos process is implemented only at the lowest frequency of interest.

In this contribution, we introduce an efficient solver for the frequency- and time-domain analysis of the electromagnetic fields in an inhomogeneous lossy medium. The finite-element method based on Whitney forms is used to discretize Maxwell's equations and the resultant matrix equation is solved by the SLDM. We illustrate the efficiency, accuracy, and validity of the present formulation by considering numerical examples of the air-filled and dielectric-loaded microwave cavities.

2 Finite-Element Formulation

Maxwell's equations can be discretized using the finite elements based on Whitney forms. These equations can be written in space-time $\Re^4$ as

$$\begin{aligned} \nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t}, \\ \nabla \times \mathbf{H} &= \mathbf{J} + \epsilon \frac{\partial \mathbf{E}}{\partial t} + \sigma \mathbf{E}. \end{aligned} \tag{1}$$

If a lossless medium is assumed, (1) can be solved for the electric field intensity so that

$$\nabla \times \left(\frac{1}{\mu} \nabla \times \mathbf{E} \right) + \epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} = -\frac{\partial \mathbf{J}}{\partial t} \tag{2}$$

with the boundary conditions

$$\begin{aligned} \hat{n} \times \mathbf{E} &= 0 \text{ on electric walls}, \\ \hat{n} \times \nabla \times \mathbf{E} &= 0 \text{ on magnetic walls}. \end{aligned} \tag{3}$$

The solution of (2) and (3) is equivalent to seeking the stationary point of the functional given by [13]

$$F(\mathbf{E}) = \frac{1}{2} \iiint_V \left[\frac{1}{\mu} (\nabla \times \mathbf{E}) \cdot (\nabla \times \mathbf{E}) + \epsilon \mathbf{E} \cdot \frac{\partial^2}{\partial t^2} \mathbf{E} \right] dV - \iiint_V \frac{\partial \mathbf{J}}{\partial t} \cdot \mathbf{E} dV, \tag{4}$$

with V denoting the volume of interest. We discretize the above functional by first subdividing V into small elements and expanding the electric field as

$$\mathbf{E}(x, y, z) = \sum_{i=1}^{N} \mathbf{N}_i(x, y, z) E_i, \tag{5}$$

where $\mathbf{N}_i$ denotes the expansion function associated with edge i, E_i denotes the associated tangential electric field, and N denotes the total number of edges in V. If we now substitute (5) into (4) and apply the Rayleigh-Ritz procedure, the following matrix equation is obtained

$$\left([C] + \frac{d^2}{dt^2} [T] \right) \{E\} = \frac{d}{dt} \{b\}, \tag{6}$$

where $\{E\} = [E_1, E_2, \ldots, E_N]^T$, and

$$\begin{aligned} C_{i,j} &= \iiint_V \frac{1}{\mu} (\nabla \times \mathbf{N}_i) \cdot (\nabla \times \mathbf{N}_j) dV, \\ T_{i,j} &= \iiint_V \epsilon \, (\mathbf{N}_i \cdot \mathbf{N}_j) \, dV, \\ b_i &= \iiint_V (\mathbf{N}_i \cdot \mathbf{J}) \, dV. \end{aligned} \tag{7}$$

The above matrix equation can be solved by the SLDM.

If we now consider solution domains containing losses, then we obtain the solution of (1) and (3) by exteremizing the functionals

$$\begin{aligned} F(\mathbf{E}) &= -\iiint_V \epsilon \mathbf{E} \cdot \frac{\partial}{\partial t} \mathbf{E} dV - \iiint_V \sigma \mathbf{E} \cdot \mathbf{E} dV + \iiint_V (\nabla \times \mathbf{E}) \cdot \mathbf{H} dV - \iiint_V \mathbf{J} \cdot \mathbf{E} dV, \\ F(\mathbf{H}) &= \iiint_V \mu \mathbf{H} \cdot \frac{\partial}{\partial t} \mathbf{H} dV + \iiint_V \mathbf{H} \cdot (\nabla \times \mathbf{E}) dV. \end{aligned} \tag{8}$$

Again, we discretize above functionals by subdividing the solution domain into small elements and expanding both the electric and the magnetic fields as

$$\mathbf{E}(x,y,z) = \sum_{i=1}^{N_e} \mathbf{N}_i(x,y,z)E_i,$$
$$\mathbf{H}(x,y,z) = \sum_{j=1}^{N_f} \mathbf{U}_j(x,y,z)H_j, \tag{9}$$

where $\mathbf{N}_i$ and $\mathbf{U}_j$ denote the expansion functions associated with edge i and face j, E_i and H_j denote the associated tangential electric field and normal magnetic field, and N_e and N_f denote the total number of edges and faces in V, respectively. If we next substitute these approximations into (8) and impose the stationary condition, we obtain the matrix equation

$$\left\{ \frac{d}{dt} \begin{pmatrix} A_\epsilon & 0 \\ 0 & B \end{pmatrix} + \begin{pmatrix} A_\sigma & C^T \\ C & 0 \end{pmatrix} \right\} \cdot \begin{Bmatrix} E \\ H \end{Bmatrix} = \begin{Bmatrix} f \\ 0 \end{Bmatrix}, \tag{10}$$

where $E = [E_1, E_2, \ldots, E_{N_e}]^T$, $H = [H_1, H_2, \ldots, H_{N_f}]^T$, and

$$A_{\epsilon i,j} = -\iiint_V \epsilon \mathbf{N}_i \cdot \mathbf{N}_j dV,$$
$$A_{\sigma i,j} = -\iiint_V \sigma \mathbf{N}_i \cdot \mathbf{N}_j dV,$$
$$B_{i,j} = \iiint_V \mu \mathbf{U}_i \cdot \mathbf{U}_j dV,$$
$$C_{i,j} = \iiint_V (\nabla \times \mathbf{N}_i) \cdot \mathbf{U}_j dV,$$
$$f_i = \iiint_V \mathbf{J} \cdot \mathbf{N}_i dV. \tag{11}$$

For simplicity, Eq. (10) can be written as

$$\left(\frac{d}{dt}[T] + [D] \right) \{x\} = \{b\}. \tag{12}$$

Equation (12) can also be solved by the SLDM which is discussed in the next section.

3 Spectral Lanczos Decomposition Method

To solve Eq. (6) by the SLDM, we first cast this equation to a form

$$\left(A + \frac{d^2}{dt^2} I \right) x = \frac{d}{dt} u, \tag{13}$$

where I is the identity matrix. Therefore, we convert matrix T in (6) to a diagonal matrix by the lumping procedure to obtain

$$\left(C + \frac{d^2}{dt^2} D \right) E = \frac{d}{dt} b. \tag{14}$$

Note that the brackets for the notation of matrices and vectors are omitted in this section for the sake of convenience. We can further write (14) as

$$\left(D^{-1/2}CD^{-1/2}+\frac{d^2}{dt^2}I\right)E'=D^{-1/2}\frac{d}{dt}b, \tag{15}$$

or

$$\left(A'+\frac{d^2}{dt^2}I\right)E'=\frac{d}{dt}b', \tag{16}$$

where $A'=D^{-1/2}CD^{1/2}$, $E'=D^{1/2}E$, and $b'=D^{-1/2}b$. If we define

$$b'(\mathbf{r},t)=b'(\mathbf{r})[u(t)-u(t-T)], \tag{17}$$

where u denotes a unit step functions, and apply the Laplace transform to (16), we can write

$$E'(s)=b'(\mathbf{r})\left[\frac{1-e^{-Ts}}{s^2I+A'}\right]. \tag{18}$$

Following the inverse Laplace transform of (18) yields

$$E'(t)=\frac{b'(\mathbf{r})}{\sqrt{A'}}\left[-\sin\sqrt{A'}t+\sin\sqrt{A'}(t-T)\right]. \tag{19}$$

By performing an identical procedure, (12) can be written as

$$\left(D'+\frac{d}{dt}I\right)x'=b', \tag{20}$$

with $D'=T^{-1/2}DT^{1/2}$, $x'=T^{1/2}x$, and $b'=T^{-1/2}b$ and the unknown vector x' can be determined in frequency- and time-domain as

$$x'(s)=b'(\mathbf{r})\left[\frac{1-e^{-Ts}}{s(sI+D')}\right]. \tag{21}$$

and

$$x'(t)=b'(\mathbf{r})\left[\frac{-e^{-D't}+e^{-D'(t-T)}}{D'}\right], \tag{22}$$

respectively. It is evident that the electric and magnetic field intensities can be analytically determined in both frequency and time domains from the above expressions. Note also that the time-dependence in (17) is not the only choice; other time functions can be chosen as well provided that their Laplace transforms exist.

The unknown vector E' or x' in the previous equations is approximated by replacing matrices A' or D' with their $M(<N$ or $<(N_e+N_f))$ eigenvalues and eigenvectors which are obtained from a symmetric tridiagonal matrix, H, generated from A' or D' via an orthogonal transformation, or more specifically, the Lanczos process. If we further define Λ and V to be the eigenvalues and their corresponding eigenvectors of matrix H, respectively, then we can write

$$H=V\Lambda V^T, \qquad \Lambda=\mathrm{diag}[\lambda_1,\lambda_2,\ldots,\lambda_M]. \tag{23}$$

The unknown vectors E' and x' can then be determined as

$$E'(s)=\|b'\|QV\left[\frac{1-e^{-Ts}}{s^2I+\Lambda}\right]V^Te_1, \tag{24}$$

$$E'(t) = \|b'\| QV \left[\frac{-\sin\sqrt{\Lambda}t + \sin\sqrt{\Lambda}(t-T)}{\sqrt{\Lambda}} \right] V^T e_1, \tag{25}$$

and

$$x'(s) = \|b'\| QV \left[\frac{1 - e^{-Ts}}{s(sI + \Lambda)} \right] V^T e_1, \tag{26}$$

$$x'(t) = \|b'\| QV \left[\frac{-e^{-\Lambda t} + e^{-\Lambda(t-T)}}{\Lambda} \right] V^T e_1, \tag{27}$$

respectively, where $e_1 = (1, 0, 0, \ldots, 0)^T$ is the first unit M vector and Q is a matrix containing the Lanczos vectors. We further note that matrix A' and D' in the above equations are sparse symmetric real and complex matrices, respectively. Therefore, the Ritz approximation matrix H of A' is a real tridiagonal matrix whereas that of D' is a complex tridiagonal matrix.

For approximate computations of the eigenpairs of A' and D', we implement the PWK and inverse iteration algorithms, and the complex QL and complex inverse iteration algorithms, respectively.

4 Results

The time- and frequency-domain techniques described above have been implemented and applied to several geometries. To validate our formulation, we first consider an air-filled microwave cavity which has unequal side lengths of 4.5 m, 3.5 m, and 2.5 m. The cavity is subdivided into 18, 14, and 10 segments in x, y, and z directions, respectively, resulting in 6,458 unknowns. To excite the cavity modes, a short pulse of duration $T = 1.92ns$ positioned near a corner of the cavity is used. A frequency range from 50 to 120 MHz is considered. First, we calculate the magnitude of the electric field inside the cavity via the SLDM at the largest frequency of interest, 120 MHz, and then we use the same Q and H matrices to evaluate the field at the remaining frequencies. A frequency increment of 0.1 MHz is chosen resulting in 699 frequency samples. The frequency spectrum of the field is plotted in Figure 1(a) indicating the resonant peaks. The total CPU time for above computations is 32.5 seconds while requiring 270 SLDM iterations. The CPU time to compute the field at 120 MHz is on the other hand, 27.4 seconds illustrating that only 5.1 seconds are needed to obtain the results at the remaining 699 frequencies. The above multiple frequency analysis is verified by evaluating the field at each single frequency in the above range and plotting the number of SLDM iterations. Results are depicted in Figure 1(b). It is clearly seen that as the frequency increases the number of iterations increases accordingly. Therefore, we can use the same H matrix generated at $f = 120$ MHz to compute the results at the lower frequencies. Additionally, the computed resonant frequencies are compared with their corresponding analytical values and the results are given in Table 1. As can be seen from this table, an excellent agreement is achieved.

Next we evaluate the electric field intensity at the center of the cavity using expression (19). Since a short pulse contains high frequency components, we use a tapered sine function

$$b'(\mathbf{r}, t) = b'(\mathbf{r}) \left[(1 - e^{-at}) \sin(bt) u(t) \right] \tag{28}$$

with $a = 0.26118 \times 10^9$ and $b = a/5$, to excite the cavity. The results are compared with the corresponding results obtained from the FDTD and the comparison is seen in Figure 2(a). As can be seen from this figure, a good agreement is observed. If we use a low-pass filter on the input signal, more accurate results are obtained as shown in Figure 2(b).

Finally, we present the numerical analysis of a partially-filled lossy microwave cavity as illustrated in Figure 3. We use the source described by (28) to excite the cavity. A time step of $\Delta t = 0.1924ns$ is chosen and the magnitude of the electric field is computed at the center of the cavity at the largest time

of interest, $t = 96.2ns$, while the same H matrix is used to calculate the field in a time period from 0 to $96.2ns$. The results are compared with the corresponding FDTD results and the comparison is given in Figure 4. To minimize the effect of the high frequency components, we use a low-pass filter on the input signal. As can be seen from Figure 4(a), a good agreement is observed. Next, we use Eq. (26) to obtain the frequency spectrum of the field at the cavity center. A short pulse with $T = 0.5772ns$ is used for excitation. Results are computed at the highest frequency of interest, 220 MHz, and the same H matrix is used to obtain results in a frequency range from 0 to 220 MHz. The frequency spectrum of the field is given in Figure 4(b). A good agreement is achieved.

5 Conclusion

A new, efficient time-domain and frequency-domain finite-element solver that can treat various practical electromagnetic problems with a large number of unknowns is introduced. The finite-element method (FEM) is used to discretize Maxwell's curl equations and the resultant matrix equation is solved by the SLDM. Our formulation is capable of obtaining results in both frequency and time domains. The efficiency and accuracy of the present technique are demonstrated by considering numerical analysis of microwave cavities.

References

[1] K. S. Yee, "Numerical Solution of Initial Boundary value Problems Involving Maxwell's Equations in Isotropic Media," *IEEE Trans. Antennas Propagat.*, vol. AP-14, pp. 302-307, May 1966.

[2] T. A. Manteuffel and J. A. White, "The Numerical Solution of the Second-Order Boundary Value Problem on Nonuniform Meshes," *Mathematics of Computation*, vol. 47, pp. 511-535, 1986.

[3] H. Kreiss, T. Manteuffel, B. Shwartz, B. Wendroff, and J. A. White, "Supraconvergent Schemes on Irregular Meshes," *Mathematics of Computation*, vol. 47, pp. 537-554, 1986.

[4] T. G. Jurgens, A. Taflov, K. Umashankar, and T. G. Moore, "Finite-Difference Time-Domain Modeling of Curved Surfaces," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 357-366, 1992.

[5] K. K. Mei, A. Cangellaris, and D. J. Angelakos, "Conformal Time-Domain Finite-Difference Method," *Radio Science*, vol. 19, pp. 1145-1147, 1984.

[6] A. C. Cangellaris, C. C. Lin, and K. K. Mei, "Point-Matched Time Domain Finite-Element Methods for Electromagnetic Radiation and Scattering," *IEEE Trans. Antennas Propagat.*, vol. AP-35, pp. 1160-1173, Oct. 1987.

[7] J. F. Lee, R. Lee, and A. C. Cangellaris, "Time-Domain Finite-Element Methods," *IEEE Trans. Antennas Propagat.*, vol. AP-45, pp. 430-442, March 1997.

[8] D. R. Lynch and K. D. Paulsen, "Time-Domain Integration of Maxwell Equations on Finite-Element," *IEEE Trans. Antennas Propagat.*, vol. AP-38, pp. 1933-1942, Dec. 1990.

[9] S. D. Gedney and U. Navsariwala, "An Unconditionally Stable Finite-Element Time-Domain Solution of the Vector Wave Equation," *IEEE Microwave and Guided Wave Letters*, vol. 5, pp. 332-334, Oct. 1995.

[10] M. Zunoubi, J. M. Jin, and W. C. Chew, "A Spectral Lanczos Decomposition Method for Solving Axisymmetric Low-Frequency Electromagnetic Diffusion by the Finite-Element Method," *J. Electromagn. Waves Applicat.*, vol. 11, pp. 1389-1406, 1997.

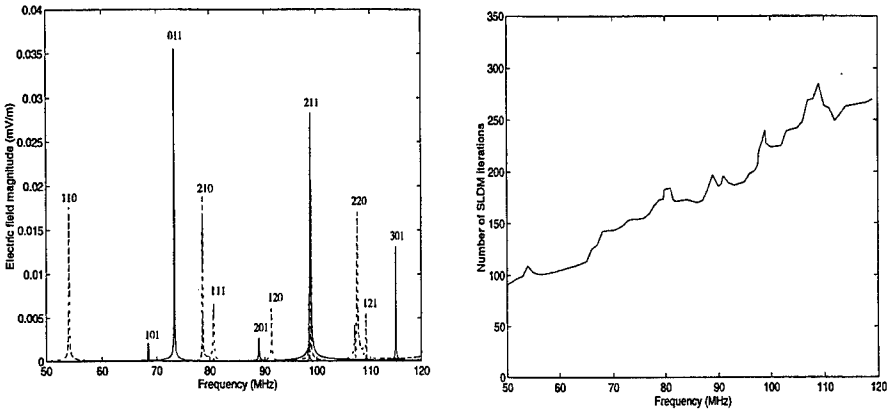
[11] M. Zunoubi, J. M. Jin, and W. C. Chew, "The Spectral Lanczos Decomposition Method for Solving Three-Dimensional Low-Frequency Electromagnetic Diffusion by the Finite-Element Method," *Proc. IEEE APS-URSI Symp.*, p. 39, 1997.

[12] V. Druskin and L. Knizhnerman, "Spectral Approach to Solving Three-Dimensional Maxwell's Diffusion Equations in the Time and Frequency Domains," *Radio Science*, vol. 29, no. 4, pp. 937-953, August 1994.

[13] J. Jin, *The Finite Element Method in Electromagnetics.* John Wiley & Sons, New York, 1993.

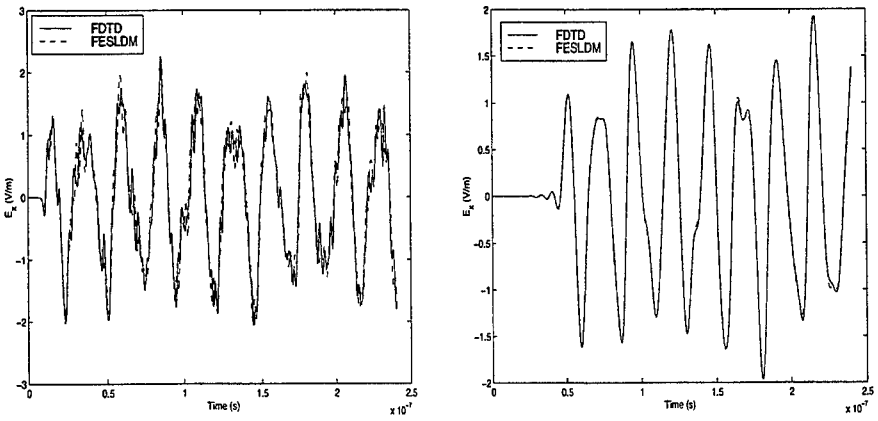
Table 1: Resonant frequencies of a $4.5m \times 3.5m \times 2.5m$ cavity resonator using frequency-domain analysis.

Mode	110	101	011	210	111	201	120	211	220	310	121	301
Exact (MHz)	54.3	68.6	73.7	79.2	80.9	89.7	92.0	99.4	108.6	108.8	109.8	116.6
FEFD (MHz)	54.1	68.4	73.4	78.7	80.7	88.9	91.3	99.0	107.6	108.6	109.2	115.0
% Error	0.37	0.29	0.4	0.63	0.24	0.89	0.76	0.4	0.9	0.18	0.15	1.37



(a) (b)

Figure 1: (a) Typical frequency spectrum of a $4.5m \times 3.5m \times 2.5m$ resonant cavity. (b) Number of SLDM iterations versus frequency.



(a) (b)

Figure 2: A comparison of the FESLDM and FDTD response of the air-filled cavity excited by (a) tapered sine function and (b) tapered sine function with a low-pass filter.

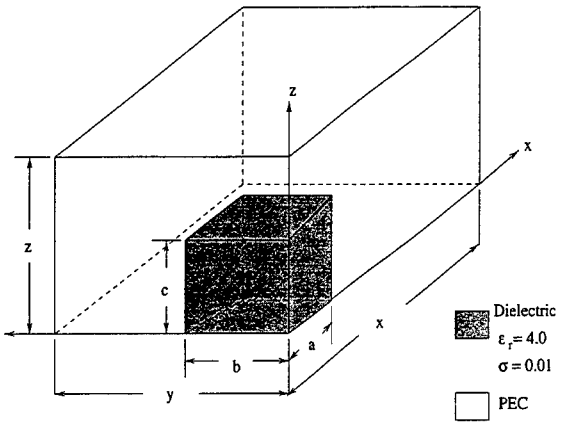
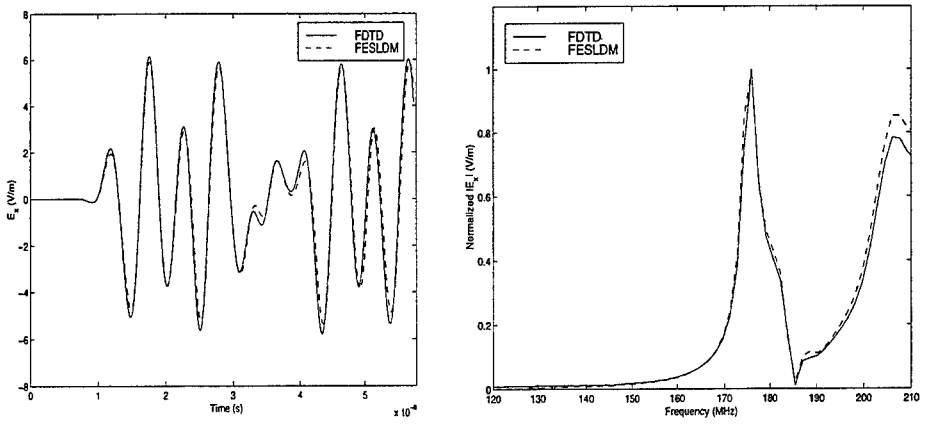


Figure 3: The problem geometry of a lossy dielectric loaded microwave cavity with $x = 2.6m$, $y = 1.6m$, $z = 1.0m$, $a = 0.6m$, $b = 0.4m, c = 0.3m, \epsilon_r = 4.0$, and $\sigma = 0.01S/m$.



(a) (b)

Figure 4: A comparison of the FESLDM and FDTD response of the lossy dielectric loaded cavity. (a) time-domain. (b) frequency-domain.

Passivity of Discrete Electromagnetic Systems

Andreas C. Cangellaris* and Li Zhao**
(*) ECE Department, University of Illinois at Urbana-Champaign
Urbana, IL 61801, U.S.A.
(**) ECE Department, University of Arizona
Tucson, AZ 85721, U.S.A.

Abstract

The issue of passivity of discrete approximations to electromagnetic systems that are passive in their continuous form is examined in this paper. Discrete model passivity is important both for numerical stability purposes and for the development of passive reduced-order models of the system for use in network-oriented circuit simulators. It is shown that the passivity of the semi-discrete model obtained from the numerical discretization of the spatial derivatives in Maxwell's equations can be examined in terms of the properties of the resulting matrix representations of the curl operators. In particular, a set of constraints on these opeartors is derived which, if satisfied, will guarantee the passivity of the state-space representation of the semi-discrete electromagnetic model.

1 Introduction

The process of developing a discrete model for the transient simulation of electromagnetic wave intearctions begins with the development of a semi-discrete approximation to Maxwell's equations. More specifically, the spatial derivatives in the system are approximated through a desired finite method (finite difference, finite element, finite volume, or more sophisticated spectral

and wavelet methods) on a properly constructed numerical grid. This process results in a system of state equations for the degrees of freedom in the numerical approximation which, in the most common cases, turn out to be the unknown time histories of the electric and magnetic field components on the grid. This spatial discretization step is followed by the application of some numerical algorithm for the numerical integration (in time) of the system of state equations (e.g. leap-frog, backward Euler, Runge-Kutta, etc.). In selecting the numerical integration algorithm, the issue of numerical stability is of primary importance. Of relevance here is Lax's equivalence theorem which states that if a linear finite-difference equation is consistent with a properly posed linear initial-value problem then stability guarantees convergence [1]. (A finite-difference scheme is said to be consistent if it has a solution that converges to the solution of the original differential equation as the mesh lengths, both in space and time, tend to zero.) Clearly, numerical stability is highly desirable since it is so tightly coupled to convergence.

Despite the significant attention paid to numerical stability, little attention is given to whether the state system of equations resulting from the spatial discretization of Maxwell's equations maintains the passive character of the continuous system (assuming, of course, that all media present are passive). Nevertheless, the passivity of the resulting semi-discrete system is important since, if guaranteed, numerical solutions of the system, generated by means of a stable integration algorithm, will converge to the correct solution and will not exhibit any (non-physical) instabilities.

Furthermore, passivity of the discrete model is essential when the development of reduced-order models, expressed as Padé approximants, for electromagnetic systems are sought for the purpose of efficient system macromodeling and incorporation in network-oriented circuit simulators. A variety of such model-order reduction methods have been proposed recently both for lumped circuits [2]-[4], and distributed electromagnetic systems [5]-[10]. When macromodels for multiports are connected together, it is important to keep in mind the fact that interconnections of stable systems may not necessarily result in stable systems; however, interconnections of passive circuits always result in systems that are passive and, hence, (asymptotically) stable [11]. This, then, implies that it is not enough for an electromagnetic macromodel to be stable. What matters, when this macromodel is to be connected with other functional blocks, is for the macromodel to be passive.

Even though not done routinely, semi-discrete approximations to Maxwell's

hyperbolic system for electromagnetic propagation in passive media are often checked for conservation of charge and conservation of energy. Such tests are clearly tests of passivity. In this paper, an alternative approach is discussed for the examination of passivity of semi-discrete approximations to Maxwell's equations. The proposed approach involves the matrices resulting from the discretization of the curl operators and (in the general case) any boundary conditions used on surfaces terminating the computational domain. In particular, a relationship between these matrices is derived which, if satisfied, will guarantee the passivity of the discrete model.

2 The Semi-Discrete Electromagnetic Model

While a variety of approaches exist for the spatial discretization of Maxwell's equations, the semi-discrete approximation effected through the use of Yee's lattice [12] is chosen for the purposes of this paper. The reason for this choice will become apparent when the passivity of the semi-discrete model is considered later in the paper.

A uniform, rectangular lattice is assumed, defined by equally spaced nodes along the three axes of a cartesian coordinate system: I along x, J along y, and K along z. The total number of nodes in the grid is $N = I \cdot J \cdot K$. With the definitions $U = 1$, $V = I$, $W = I \cdot J$, the nth electric node, corresponding to node (i, j, k) in the grid, is given by

$$n = 1 + (i-1)U + (j-1)V + (k-1)W \tag{1}$$

where $i = 1, 2, \ldots, I$, $j = 1, 2, \ldots, J$, $k = 1, 2, \ldots, K$, and $n = 1, 2, \ldots, N$.

It is assumed that the media are linear, passive, and time-independent. Thus, Maxwell's curl equations in the Laplace domain assume the form

$$\nabla \times \vec{E} = -s\mu\vec{H} \tag{2}$$

$$\nabla \times \vec{H} = s\epsilon\vec{E} + \sigma E + \vec{J}_s \tag{3}$$

where the electric permittivity, ϵ, electric conductivity σ, and magnetic permeability, μ, are, in general, position dependent. The dependence of the electric and magnetic field vectors $\vec{E}$ and $\vec{H}$, as well as the imposed source current density $\vec{J}_s$, on position, and the Laplace variable s is suppressed for simplicity.

In order to cast the semi-discrete form of Maxwell's equations in a matrix form, we begin with the definition of the following two vectors of discrete unknowns

$$\mathbf{E} = [\mathbf{E_x}, \mathbf{E_y}, \mathbf{E_z}]^T \tag{4}$$

$$\mathbf{H} = [\mathbf{H_x}, \mathbf{H_y}, \mathbf{H_z}]^T \tag{5}$$

where $\mathbf{E_x}$ is a vector of length N, containing the N E_x values on the grid, with similar definitions for the remaining five vectors. Using the definitions in (4),(5) and writing the curl operator in its matrix form

$$\nabla\times = \begin{bmatrix} 0 & -\partial/\partial z & \partial/\partial y \\ \partial/\partial z & 0-\partial/\partial x & \\ -\partial/\partial y & \partial/\partial x & 0 \end{bmatrix} \tag{6}$$

it is straightforward to show that the semi-discrete form of (2),(3) can be written as

$$\begin{bmatrix} \mathbf{0} & \mathbf{A_w}^T & -\mathbf{A_v}^T \\ -\mathbf{A_w}^T & \mathbf{0} & \mathbf{A_u}^T \\ \mathbf{A_v}^T & -\mathbf{A_u}^T & \mathbf{0} \end{bmatrix} \cdot \mathbf{E} = -s\mathbf{D_h}\mathbf{H} \tag{7}$$

$$\begin{bmatrix} \mathbf{0} & -\mathbf{A_w} & \mathbf{A_v} \\ \mathbf{A_w} & \mathbf{0} & -\mathbf{A_u} \\ -\mathbf{A_v} & \mathbf{A_u} & \mathbf{0} \end{bmatrix} \cdot \mathbf{H} = s\mathbf{D_e}\mathbf{E} + \mathbf{D}_\sigma\mathbf{E} + \mathbf{J_s} \tag{8}$$

In the above equations, the matrices $\mathbf{A_u}$, $\mathbf{A_v}$, $\mathbf{A_w}$ are sparse with only two bands having nonzero elements: one band is along the diagonal with all values equal to 1, and the second band at a distance of U, V, W, respectively, to the left of the diagonal with all values equal to -1. The matrices $\mathbf{D_e}$, $\mathbf{D_h}$, $\mathbf{D}_\sigma$ are diagonal matrices with elements dependent on the electromagnetic properties of the media and the grid size.

The system of (7),(8) may be cast in a compact form by defining the vector of state variables $\mathbf{X} = [\mathbf{H}, \mathbf{E}]^T$, and the $3N \times 3N$ matrix, $\mathbf{P}$,

$$\mathbf{P} = \begin{bmatrix} \mathbf{0} & \mathbf{A_w} & -\mathbf{A_v} \\ -\mathbf{A_w} & \mathbf{0} & \mathbf{A_u} \\ \mathbf{A_v} & -\mathbf{A_u} & \mathbf{0} \end{bmatrix} \tag{9}$$

In addition, since the objective is to develop reduced-order macromodels for multiport electromagnetic systems, the source notation is slightly modified

to include the imposed currents used for the excitation of the ports. For this purpose, it is assumed that the electromagnetic system under consideration has p ports, each coinciding with one electric field node, and a constant matrix $\mathbf{B}$ of dimenion $6N \times p$ is introduced, with nonzero elements only in its bottom $3N$ rows associated with the electric field nodes in the state vector $\mathbf{X}$. The specific values of these elements will depend, in general, on the source distribution and numerical grid characteristics. Using the vector $\mathbf{U}(s)$ to denote the Laplace transforms of the current source waveforms at the p ports, the discrete source term may be cast in the form $\mathbf{BU}(s)$. With these definitions, the resulting compact form of (7), (8) is,

$$\begin{bmatrix} \mathbf{0} & \mathbf{P}^T \\ -\mathbf{P} & \mathbf{0} \end{bmatrix} \mathbf{X} = -s \begin{bmatrix} \mathbf{D_h} & \mathbf{0} \\ \mathbf{0} & \mathbf{D_e} \end{bmatrix} \mathbf{X} - \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_\sigma \end{bmatrix} \mathbf{X} - \mathbf{BU}(s) \qquad (10)$$

or, in a yet more compact form,

$$(\mathbf{G} + s\mathbf{C})\mathbf{X} = -\mathbf{BU}(s) \qquad (11)$$

where

$$\mathbf{G} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{P}^T \\ -\mathbf{P} & \mathbf{D}_\sigma \end{bmatrix}, \quad \mathbf{C} \equiv \begin{bmatrix} \mathbf{D_h} & \mathbf{0} \\ \mathbf{0} & \mathbf{D_e} \end{bmatrix} \qquad (12)$$

Because of the assumed passivity of the media, the matrices $\mathbf{D_e}$, $\mathbf{D_h}$ are (symmetric) positive definite and $\mathbf{D}_\sigma$ is (symmetric) non-negative definite. Consequently, $\mathbf{C}$ is also (symmetric) positive definite.

Defining a desired output vector as $\mathbf{Y}(s) = \mathbf{FX}(s)$, where $\mathbf{F}$ is a selector matrix, we have

$$\mathbf{Y}(s) = \mathbf{F}(\mathbf{G} + s\mathbf{C})^{-1}\mathbf{BU}(s) \qquad (13)$$

For the case of multiports, the number of outputs is the same with the number of inputs. More specifically, with the selection of the current density as the excitation at each port, the electric field vector (or, equivalently, the voltage at the port (if the definition of the voltage makes sense) becomes the output quantity. For such cases, the selector matrix $\mathbf{F}$ is a $p \times 6N$ matrix, with constant, non-zero entries at the right half of each of its rows, corresponding to the electric field unknowns in the state vector $\mathbf{X}$. In particular, except for a scaling constant factor associated with the proper definition of

the observed output quantity at the port, the selector matrix $\mathbf{F}$ is simply the transpose of the source matrix $\mathbf{B}$ defined earlier. Thus, the electromagnetic characterization of multiports is effected in terms of a transfer-function matrix, $\mathbf{H}(s)$, which, in view of (13), is given by

$$\mathbf{H}(s) = \mathbf{B}^T(\mathbf{G} + s\mathbf{C})^{-1}\mathbf{B} \tag{14}$$

3 Passivity of the Discrete Model

As already mentioned in the introduction, reduced-order model representations of passive electromagnetic multiports are extremely useful when these multiports constitute parts of more complex functional blocks. For the purposes of design-driven simulation at the functional block level, a network-oriented simulation approach is used. The incorporation of the reduced-order models for the electromagnetic multiports in the overall network-oriented circuit simulator may be effected either through recursive convolution, utilizing the pole-residue representations of the elements of the transfer function matrix [13], or through the direct incorporation of the state-space representation of the reduced system in the circuit simulator [14]. In either case, the passivity of the reduced system needs to be verified in order to avoid non-physical instabilities in the subsequent simulation of the overall circuit. However, it is meaningless to talk about passivity of the reduced-order model without establishing first the passivity of the original discrete model. The investigation of the passivity of the system of (11) is the topic of this section.

Our analysis makes use of the following useful results [15]:

Theorem 1: The transfer function matrix $\hat{\mathbf{H}}(s)$ of a passive (linear, solveable, time-invariant) network is positive-real; that is,

a) Each element of $\hat{\mathbf{H}}(s)$ is analytic for $Re(s) > 0$.

b) $\hat{\mathbf{H}}(s^*) = \hat{\mathbf{H}}^*(s)$ for $Re(s) > 0$.

c) $\hat{\mathbf{H}}_h(s) = \hat{\mathbf{H}}^{*T}(s) + \hat{\mathbf{H}}(s) \geq 0$ for $Re(s) > 0$.

Theorem 2: If a matrix, $\mathbf{A}$, is positive-real, then so is its inverse, $\mathbf{A}^{-1}$, if it exists.

Theorem 3: If $\mathbf{A}$ is positive-real and $\mathbf{A}_h = \mathbf{A}^{*T} + \mathbf{A} > 0$ for $Re(s) > 0$, then $\mathbf{A}^{-1}$ exists.

Theorem 4: If $\hat{\mathbf{B}}$ is a real constant $m \times n$ matrix and $\mathbf{A}(s)$ is an $m \times m$ positive-real matrix, then $\hat{\mathbf{B}}^T\mathbf{A}\hat{\mathbf{B}}$ is an $n \times n$ positive-real matrix.

With the output defined as $\mathbf{Y}(s) = \mathbf{B}^T\mathbf{X}(s)$, the transfer function of (11) is given by

$$\mathbf{H}(s) = -\mathbf{B}^T(\mathbf{G} + s\mathbf{C})^{-1}\mathbf{B} \tag{15}$$

According to Theorem 1, the discrete approximation to the system of Maxwell's equations will be passive if $\mathbf{H}(s)$ is positive-real. However, the matrix $\mathbf{B}$ in (15) is a real constant matrix. Thus, in view of Theorems 2,3, and 4, to prove the passivity of the discrete system of (11) it suffices to show that the matrix $\mathbf{S} = \mathbf{G} + s\mathbf{C}$ is positive-real and $\mathbf{S}_h > 0$.

This takes us back to Theorem 1 in which the three requirements for a matrix to be positive-real are given. First, we note that matrices $\mathbf{G}$ and $\mathbf{C}$ are real. Hence, requirements (a) and (b) are automatically satisfied. To prove that requirement (c) is also satisfied, strengthened by the additional requirement that $\mathbf{S}_h > 0$, we need to show that $\mathbf{z}^{*T}\left(\mathbf{S}^{*T} + \mathbf{S}\right)\mathbf{z} > 0$ for $Re(s) > 0$ and for any complex vector $\mathbf{z}$. Setting $s = a + j\omega$, one obtains after some straightforward matrix algebra

$$\mathbf{z}^{*T}\left(\mathbf{S}^{*T} + \mathbf{S}\right)\mathbf{z} = \mathbf{z}^{*T}\left[\left(\mathbf{G} + \mathbf{G}^T\right) + a\left(\mathbf{C} + \mathbf{C}^T\right)\right]\mathbf{z} \tag{16}$$

where use has been made of the fact that $\mathbf{C}$ is symmetric. Finally, using the fact that $\mathbf{C} + \mathbf{C}^T = 2\mathbf{C}$, the above equation may be cast in the form

$$\mathbf{z}^{*T}\left(\mathbf{S}^{*T} + \mathbf{S}\right)\mathbf{z} = \mathbf{z}^{*T}\left[\left(\mathbf{G} + \mathbf{G}^T\right) + 2a\mathbf{C}\right]\mathbf{z} \tag{17}$$

Since $\mathbf{C}$ is positive definite, it follows immediately that $2a\mathbf{z}^{*T}\mathbf{C}\mathbf{z} > 0$ for $a > 0$. Furthermore, using the fact that $\mathbf{G}$ is skew-symmetric (see (12)), it is straightforward to show that

$$\mathbf{G} + \mathbf{G}^T = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 2\mathbf{D}_\sigma \end{bmatrix} \tag{18}$$

But $\mathbf{D}_\sigma$ is a non-negative definite matrix; hence, $\mathbf{z}^{*T}\left(\mathbf{G} + \mathbf{G}^T\right)\mathbf{z} \geq 0$. Thus, we conclude that the product in (17) is positive definite for any complex vector $\mathbf{z}$ and for $Re(s) > 0$; hence, the discrete system of (11) is passive.

It is important to point out that critical to this proof of passivity of the discrete system was the fact that the matrix $\mathbf{G}$ was skew-symmetric. Clearly, this skew-symmetry of $\mathbf{G}$ is a direct consequence of the uniformity of the (orthogonal) cartesian grid used for the discretization, as well as the staggering of the electric and magnetic field nodes. From a numerical integration point of view, it is extremely useful to be able to validate that the (semi-discrete) system of state equations resulting from the numerical approximations of the spatial derivatives is passive, since passivity guarantees stability and thus a stable numerical solution will always be achieved with a stable integration algorithm. Proof of passivity of the semi-discrete system when unstructured grids are used is rather cumbersome since, in most cases, it can be effected only through an eigenvalue analysis.

Taking a closer look at the structure of the matrix $\mathbf{G}$, it becomes apparent that the specific form of the numerical approximation of the two curl operators in Maxwell's system plays an important role on the passivity of the approximation. In order to account for the general case, let $\mathbf{P_m}$ and $\mathbf{P_e}$ denote the matrix approximations resulting from the discretization of the curl of the magnetic and electric field, respectively, over the entire computational domain. Then, the matrix $\mathbf{G}$ assumes the general form

$$\mathbf{G} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{P_e} \\ -\mathbf{P_m} & \mathbf{D}_\sigma \end{bmatrix} \tag{19}$$

and thus, the requirement for passivity of the approximation centers around the properties of the matrix

$$\mathbf{G}_h = \mathbf{G} + \mathbf{G}^T = \begin{bmatrix} \mathbf{0} & (\mathbf{P_e} - \mathbf{P_m}^T) \\ (\mathbf{P_e} - \mathbf{P_m}^T)^T & 2\mathbf{D}_\sigma \end{bmatrix} \tag{20}$$

More specifically, considering that $\mathbf{D}_\sigma$ is non-negative definite, the passivity of the numerical model depends solely on the properties of the matrix $\mathbf{Q} = \mathbf{P_e} - \mathbf{P_m}^T$. To elaborate, let us expand the term $\mathbf{z}^{*T}(\mathbf{G}_h)\mathbf{z}$, where $\mathbf{z}$ is written in the form

$$\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} \tag{21}$$

It is then

$$\begin{bmatrix} \mathbf{z_1}^{*T} & \mathbf{z_2}^{*T} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{Q} \\ \mathbf{Q}^T & 2\mathbf{D}_\sigma \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} =$$

$$\mathbf{z_1}^{*T}\mathbf{Q}\mathbf{z_2} + \left[\left(\mathbf{z_1}^{*T}\mathbf{Q}\mathbf{z_2}\right)^T\right]^* + 2\mathbf{z_2}^{*T}\mathbf{D}_\sigma\mathbf{z_2} = 2Re\{\mathbf{z_1}^{*T}\mathbf{Q}\mathbf{z_2}\} + 2\mathbf{z_2}^{*T}\mathbf{D}_\sigma\mathbf{z_2} \quad (22)$$

where use has been made of the fact that $\mathbf{Q}$ is real. The second term in the last equation is always non-negative. Considering that it is common to assume loss-free media for numerical wave simulation purposes, the passivity of the numerical model depends on whether the first term in (22) is non-negative for any $\mathbf{z}$. Clearly, this is ensured when $\mathbf{Q} = 0$ or, equivalently, $\mathbf{P_e} = \mathbf{P_m}^T$. As already seen in the first part of this section, this occurs naturally when the (structured) Yee's lattice is used for the discretization. For other discretization choices, the passivity of the discrete model can be evaluated by examining the matrix $\mathbf{P_e} - \mathbf{P_m}^T$. If this matrix is non-negative, the discrete model is certainly passive and its numerical integration (with a numerically stable integration scheme) will lead to a stable numerical solution.

4 Concluding Remarks

The general result of this paper may be stated as follows: "The system of state equations resulting from the discretization of Maxwell's curl equations in passive media is passive if the matrix $\mathbf{P_e} - \mathbf{P_m}^T$ is non-negative definite." In the absence of any boundary conditions, $\mathbf{P_e}$ is the matrix representation of the discretization of $\nabla \times \vec{E}$, and $\mathbf{P_m}$ is the matrix representation of the discretization of $\nabla \times \vec{H}$. Actually, for the general case, boundary conditions on surfaces terminating the computational domain must be taken into account. In their discrete form such conditions lead to modifications in the matrices $\mathbf{P_e}$ and $\mathbf{P_m}$, as well as the matrix $\mathbf{C}$ in (11), and thus impact passivity. Nevertheless, the aforementioned stated requirement for passivity still holds.

For the case where the structured, rectangular Yee's lattice is used for the spatial discretization of Maxwell's equations, proof of passivity is trivial since $\mathbf{P_e} - \mathbf{P_m}^T = 0$. The case of unstructured grids is definitely more difficult to investigate.

The impact of typically encountered boundary conditions (e.g., impedance boundary conditions and absorbing boundary conditions) on the passivity of the discrete model will be discussed in detail in a forthcoming paper. Preliminary work has addressed this issue for the case of one-dimensional distributed

electromagnetic problems, with specific application in the transient analysis of transmission lines [16]. It was shown that, despite their superior computational efficiency and accuracy, spectral Chebyshev approximations of the hyperbolic system of the transmission line equations [8] lead to discrete models which are not passive. The difficulties with the stable numerical integration of such approximations using explicit schemes are well-known [17].

References

[1] R.D. Richtmyer and K.W. Morton, *Difference Methods for Initial-value Problems*, Interscience Publishers, New York, 1967.

[2] S.-Y. Kim, N. Gopal and L.T. Pillage, "Time-domain macromodels for VLSI interconnect analysis," *IEEE Trans. on Computer-Aided Design*, Vol. CAD-13, pp. 1257-1270, 1994.

[3] E. Chiprout and M.S. Nakhla, *Asymptotic Waveform Evaluation And Moment Matching for Interconnect Analysis.* Kluwer Academic Publishers, Dordrecht, 1994.

[4] P. Feldmann and R.W. Freund "Efficient linear circuit analysis by Padé approximation via the Lanczos process," *IEEE Trans. on Computer-Aided Design*, Vol. CAD-14, pp. 639-649, 1995

[5] R.F. Remis and P. Van den Berg, "A modified Lanczos reduction method for the computation of transient wavefields," *Proceedings of the IEEE Antennas and Propagation Society International Symposium*, Vol. 3, pp. 2088-2091, Baltimore, Maryland, 1996.

[6] J.E. Bracken and Z.J. Cendes, "Transient analysis via electromagnetic fast-sweep methods and circuit models," *Proc. 13th Annual Review of Progress in Applied Computational Electromagnetics*, vol. I, pp. 172-179, March 1997.

[7] M.A. Kolbedhari, M. Nakhla, R. Achar and M. Srinivasan, "Solution of EM problems using reduced-order models by complex frequency hopping," *Proc. 13th Annual Review of Progress in Applied Computational Electromagnetics*, vol. I, pp. 165-171, March 1997.

[8] M. Celik and A.C. Cangellaris, "Simulation of dispersive multiconductor transmission lines by Padé approximation via the Lanczos process," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 2525-2535, Dec. 1996.

[9] M. Celik and A.C. Cangellaris, "Simulation of multiconductor transmission lines using Krylov subspace order-reduction techniques," *IEEE Trans. Computer-Aided Design*, vol. 16, pp. 485-496, May 1997.

[10] A.C. Cangellaris and L. Zhao, "Reduced-order modeling of electromagnetic systems with Padé via Lanczos approximations," *Proc. 13th Annual Review of Progress in Applied Computational Electromagnetics*, vol. I, pp. 148-155, March 1997.

[11] R.A. Rohrer and H. Nosrati, "Passivity considerations in stability studies of numerical integration algorithms," *IEEE Trans. on Circuits and Systems*, vol. CAS-28, pp. 857-866, Sept. 1981.

[12] K.S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, vol. 14, pp. 302-307, 1966.

[13] A. Semlyen and A. Dabuleanu, "Fast and accurate switching transient calculations on transmission lines with ground return using recursive convolutions," *IEEE Trans. Power. Appl. Syst.*, vol. PAS-94, pp. 561-571, Mar./Apr. 1975.

[14] P. Feldmann and R.W. Freund, "Reduced-order modeling of large linear subcircuits via a Block Lanczos algorithm," *IEEE/ACM Proc. Design Automation Conference*, pp. 474-479, June 1995.

[15] R.W. Newcomb, *Linear Multiport Synthesis*, McGraw-Hill, New York, 1966.

[16] A.C. Cangellaris and M. Celik, "Order reduction of high-speed interconnect electrical models: the issue of passivity," *Proc. IEEE Symposium on IC/Package Design Integration*, Santa Cruz, California, Feb. 1998.

[17] C. Canuto, M.Y. Hussaini, A. Quarteroni, and T.A. Zhang, *Spectral Methods in Fluid Dynamics*, Springer-Verlag, Berlin, 1988.

Rational Krylov Reduced Order Modeling of Multiscreen Frequency Selective Surfaces

D. S. Weile and E. Michielssen*
Department of Electrical and Computer Engineering
Center for Computational Electromagnetics
University of Illinois at Urbana-Champaign
1406 W. Green St., Urbana, IL 61801

K. Gallivan
Computer Science Department
Florida State University
203 Love Building, Tallahassee, FL 32306

Abstract

Model order reduction of the linear system provided by the application of the spectral Galerkin method to multiple screen Frequency Selective Surfaces (FSSs) is demonstrated. The original spectral Galerkin system has a nonlinear frequency dependence which is approximated by an osculatory polynomial interpolant and subsequently expressed in a linearized companion form. A rational interpolant reduced order model for the reflection coefficient of the FSS as a function of frequency that matches the linearized system reflection coefficient and its derivatives at several select frequencies is generated using rational Krylov techniques. For an FSS originally modeled with a spectral Galerkin system involving hundreds or thousands of unknowns, this procedure results in a system of fewer than twenty unknowns which accurately reproduces the behavior of the original system over a large frequency band.

1. Introduction

Multiscreen Frequency Selective Surfaces (FSSs) (Fig 1.) are useful over a broad part of the electromagnetic spectrum as frequency or angular filters, and they have been especially important in satellite communications where they find use as subreflectors in dish antennas. The most popular analysis method for FSSs is the spectral Galerkin method, which is a derivative of the Method of Moments (MoM) for the analysis of planar periodic structures illuminated by a plane wave. While both full domain and subdomain basis functions have been used to represent the current on the FSS for the MoM analysis, this paper considers the more general subdomain case. When subdomain basis functions are used, the metallized portions of the FSS are partitioned into n basis functions of identical shape over a grid of regularly spaced, equally sized regions of support [1]. The application of the spectral Galerkin method to the discretized FSS structure results in system of equations for $\mathbf{i}_{nl}(f)$, the weighting coefficients associated with the n bases, and a linear relation between the FSS specular reflection coefficient $R_{nl}(f)$ and $\mathbf{i}_{nl}(f)$ which may be expressed as

$$\begin{aligned} \mathbf{A}_{nl}(f)\mathbf{i}_{nl}(f) &= \mathbf{b}_{nl}(f) \\ R_{nl}(f) &= \mathbf{c}_{nl}^{*}(f)\mathbf{i}_{nl}(f), \end{aligned} \tag{1}$$

where $\overline{\mathbf{A}}_{nl}(f)$ is an $n \times n$ system matrix, and $\mathbf{b}_{nl}(f)$ and $\mathbf{c}_{nl}(f)$ are input and output n vectors, respectively, all depending nonlinearly on the frequency f [1]. The asterisk denotes the complex conjugate transpose operator.

While the spectral Galerkin technique is a powerful method for calculating the reflection coefficient of an FSS at a given frequency, the iterative application of this method to characterize FSS behavior over a band of frequencies can be computationally very intensive. For each frequency of interest, the technique involves the summation of doubly infinite series to construct the system, and a solution of an $n \times n$ system of linear equations. The goal of this paper, therefore, is to accelerate the calculation of $R(f)$ by introducing an approximation to system (1) of the form

$$\begin{aligned} \left(f\hat{\overline{\mathbf{E}}} - \hat{\overline{\mathbf{A}}}\right)\hat{\mathbf{i}}(f) &= \hat{\mathbf{b}} \\ \hat{R}(f) &= \hat{\mathbf{c}}^{*}\hat{\mathbf{i}}(f) \end{aligned} \tag{2}$$

where $\hat{R}(f) \approx R_{nl}(f)$, $\hat{\overline{\mathbf{E}}}$ and $\hat{\overline{\mathbf{A}}}$ are small, constant, $m \times m$ matrices, $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$ are constant complex m vectors, and $\hat{\mathbf{i}}$ is an m dimensional state vector. The system size m is typically on the order of ten. This process is known as model order reduction [2], and though it has been used for several years in circuit and interconnect characterization, its application to integral equations is quite recent [3]. This new system will be constructed so that $\hat{R}$ matches R_{nl} and its derivatives at different frequency points in the band of interest. Because of the small size and simple frequency dependence of (2), $\hat{R}$ can be cheaply calculated for a large number of frequencies.

2. Formulation

In this section, the model order reduction of (1) is discussed. The construction proceeds in three steps to be discussed in turn. Section 2.1 discusses the elimination of the frequency dependence of the input and output vectors of (1). Section 2.2 then demonstrates the linearization of the system using an osculatory polynomial interpolant. Finally, Section 2.3 describes the construction of the reduced order model (2) using the Dual Rational Arnoldi (DRA) algorithm [4].

2.1 Constancy of Input and Output Vectors

Because the formulation of the FSS scattering problem has been well documented in the literature [1], the construction of system (1) is not detailed here. However, it is important to note that because (1) results from the application of the spectral Galerkin method with uniformly spaced basis functions, it may be solved for the current weighting coefficient vector $\mathbf{i}_{nl}(f)$ using an iterative method where matrix vector multiplication is accomplished using the Fast Fourier Transform (FFT) [1].

The frequency dependence of the input and output vectors of (1) can be removed with a change of state variables and a matrix multiplication. Define

$$\begin{aligned} \mathbf{b}_{nl}(f) &= \mathbf{B}(f)\mathbf{u} \\ \mathbf{i}_0(f) &= \mathbf{C}^*(f)\mathbf{i}_{nl}(f) \end{aligned} \tag{3}$$

where $\mathbf{B}(f)$ and $\mathbf{C}(f)$ are $n \times n$ diagonal matrices with diagonal elements equal to the elements of $\mathbf{b}_{nl}(f)$ and $\mathbf{c}_{nl}(f)$ respectively, and $\mathbf{u}$ is an n vector of all ones. Rewriting (1) in terms of the new state variables $\mathbf{i}_0(f)$ and premultiplying the state equations by $\mathbf{B}^{-1}(f)$ yields a new system

$$\begin{aligned} \mathbf{A}'_{nl}(f)\mathbf{i}_0(f) &= \mathbf{u} \\ R_{nl}(f) &= \mathbf{u}^*\mathbf{i}_0(f) \end{aligned} \tag{4}$$

where $\mathbf{A}'_{nl}(f) = \mathbf{C}^{-*}(f)\mathbf{A}_{nl}(f)\mathbf{B}^{-1}(f)$, and $\mathbf{C}^{-*}(f) = \left(\mathbf{C}^{-1}(f)\right)^*$. It is important to note that even after this transformation, the $\mathbf{A}'_{nl}(f)$ can still be multiplied with an arbitrary vector using the FFT, so that (4) may be solved for the current as efficiently as (1). This property is of great importance to ensure that the cost associated the construction of the reduced order model does not outweigh its benefits once generated.

2.2 System Linearization

Equations (4) are now in a form where they may be approximated easily by a canonical linear system. To accomplish this, an osculatory polynomial interpolant of order N_{ord} for $\mathbf{A}'_{nl}(f)$ is constructed using a generalized form of the well-known Newton divided difference algorithm [5]. Assuming that $\mathbf{A}'_{nl}(f)$ is to be interpolated at the frequencies f_j, $j = 0,\ldots,N_{ord}$, this process results in an expression of the form

$$\mathbf{A}'_{nl}(f) \approx \sum_{i=0}^{N_{ord}} \mathbf{A}'_i \prod_{j=0}^{i-1} (f - f_j). \tag{5}$$

Using inverse synthetic division [5], (5) can be expressed as a sum over powers of a single monomial as

$$\mathbf{A}'_{nl}(f) \approx \sum_{i=0}^{N_{ord}} \mathbf{A}_i \left(f - f_{exp}\right)^i \tag{6}$$

where f_{exp} is an expansion point chosen near the center of the band to make the calculation of the $\mathbf{A}_i$ from the $\mathbf{A}'_i$ stable. Defining

$$\mathbf{i}_i = \left(f - f_{exp}\right)^i \mathbf{i}_0, \quad i = 1,\ldots N_{ord} - 1, \tag{7}$$

an approximation to system (4) can be written in canonical linear system form as

$$\begin{aligned} \left[\left(f - f_{exp}\right)\mathbf{E} - \mathbf{A}\right]\mathbf{i} &= \mathbf{b} \\ R(f) &= \mathbf{c}^*\mathbf{i} \end{aligned} \tag{8}$$

where $R(f) \approx R_{nl}(f)$,

$$\mathbf{E}=\begin{bmatrix}\mathbf{I} & & & \\ & \mathbf{I} & & \\ & & \ddots & \\ & & & \mathbf{A}_{N_{ord}}\end{bmatrix},\quad \mathbf{A}=\begin{bmatrix} & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \\ -\mathbf{A}_0 & -\mathbf{A}_1 & -\mathbf{A}_2 & -\mathbf{A}_{N_{ord}-1}\end{bmatrix}, \tag{9}$$

$\mathbf{i}^* = \begin{bmatrix}\mathbf{i}_0^* & \mathbf{i}_1^* & \cdots & \mathbf{i}_{N_{ord}-1}^*\end{bmatrix}$, $\mathbf{b}^* = \begin{bmatrix}\mathbf{0}^* & \cdots & \mathbf{0}^* & \mathbf{u}^*\end{bmatrix}$, $\mathbf{c}^* = \begin{bmatrix}\mathbf{u}^* & \mathbf{0}^* & \cdots & \mathbf{0}^*\end{bmatrix}$, and $\mathbf{I}$ is the identity matrix [3]. Note that (8) is in the canonical state space form with a frequency shift that can be ignored if the substitution $f - f_{exp} \rightarrow f$ is made. Like the system matrix $\mathbf{A}'_{nl}(f)$, each $\mathbf{A}_i$ can be multiplied with a vector using an FFT, thus preserving the efficiency of the method.

2.3 Model Order Reduction

Because model reduction algorithms are generally formulated in terms of state space system representations, a reduced order model of (8) can be generated readily [5]. The reduced order system (2) is constructed to match the transfer function of (8) and its derivatives at selected interpolation points $f^{(k)}$, $k = 1,\ldots,K$; such a model is known as a rational interpolant. Specifically, rectangular matrices $\mathbf{Z}$ and $\mathbf{V}$ are sought such that a reduced model of (8) of can be found in the form shown in (2) where $\hat{\mathbf{E}} = \mathbf{Z}^*\mathbf{E}\mathbf{V}$, $\hat{\mathbf{A}} = \mathbf{Z}^*\mathbf{A}\mathbf{V}$, $\hat{\mathbf{b}} = \mathbf{Z}^*\mathbf{b}$, and $\hat{\mathbf{c}} = \mathbf{V}^*\mathbf{c}$. The $\mathbf{V}$ and $\mathbf{Z}$ needed to generate a rational interpolant can be described with the following theorem [4]:

Let $\mathcal{K}_J(\mathbf{G},\mathbf{g}) = \left\{\mathbf{g}, \mathbf{G}\mathbf{g}, \mathbf{G}^2\mathbf{g}, \ldots, \mathbf{G}^{J-1}\mathbf{g}\right\}$ be the Krylov subspace of dimension J generated by the square matrix $\mathbf{G}$ and the vector $\mathbf{g}$. Furthermore, denote the space spanned by the columns of an arbitrary rectangular matrix $\mathbf{H}$ as colsp($\mathbf{H}$). If

$$\bigcup_{k=1}^{K} \mathcal{K}_{J_k^b}\left(\left(\mathbf{A}-f^{(k)}\mathbf{E}\right)^{-1}\mathbf{E}, \left(\mathbf{A}-f^{(k)}\mathbf{E}\right)^{-1}\mathbf{b}\right) \subset \text{colsp}(\mathbf{V}) \tag{10}$$

and

$$\bigcup_{k=1}^{K} \mathcal{K}_{J_k^c}\left(\left(\mathbf{A}-f^{(k)}\mathbf{E}\right)^{-*}\mathbf{E}^*, \left(\mathbf{A}-f^{(k)}\mathbf{E}\right)^{-*}\mathbf{c}\right) \subset \text{colsp}(\mathbf{Z}) \tag{11}$$

then

$$-\mathbf{c}^*\left[\left(\mathbf{A}-f^{(k)}\mathbf{E}\right)^{-1}\mathbf{E}\right]^{j_k-1}\left(\mathbf{A}-f^{(k)}\mathbf{E}\right)^{-1}\mathbf{b} = -\hat{\mathbf{c}}^*\left[\left(\hat{\mathbf{A}}-f^k\hat{\mathbf{E}}\right)^{-1}\hat{\mathbf{E}}\right]^{j_k-1}\left(\hat{\mathbf{A}}-f^{(k)}\hat{\mathbf{E}}\right)^{-1}\hat{\mathbf{b}} \tag{12}$$

for $1 \le j_k \le J_k^b + J_k^c$ and $1 \le k \le K$, where $\hat{\mathbf{E}} = \mathbf{Z}^*\mathbf{E}\mathbf{V}$, $\hat{\mathbf{A}} = \mathbf{Z}^*\mathbf{A}\mathbf{V}$, $\hat{\mathbf{b}} = \mathbf{Z}^*\mathbf{b}$, and $\hat{\mathbf{c}} = \mathbf{V}^*\mathbf{c}$.

The quantities on the left and right hand sides of (12) are seen to be the reflection coefficient and its derivatives at the interpolation points, so the $\mathbf{V}$ and $\mathbf{Z}$ of (10) and (11) generate the required model. While many algorithms exist to construct such $\mathbf{V}$ and $\mathbf{Z}$, this paper uses DRA to ensure that the model is produced in a numerically stable fashion [5].

3. Numerical Results

In this section, the application of the above procedure is demonstrated for six different two-screen FSSs with square periodic cells illuminated by a normally incident wave with electric field vector aligned along the x axis (Fig. 1). The periodic cell for each of the six examples is shown in Figure 2, and both screens of each FSS are identical. The screens are discretized on a 16×16 grid using the procedure given in [1]. If the side of a periodic cell is denoted by Δx, the screens are separated by a distance $\Delta x/10$.

The response of each screen was calculated using the MoM, the polynomial interpolant system (9) discussed in Section 2.2, and a reduced order model generated from it by the DRA. The results are plotted in Figure 3 versus a normalized frequency

$$\hat{f} = \frac{f\Delta x}{c} \tag{13}$$

where f is the frequency of the wave and c is the speed of light. Each FSS was analyzed over a normalized frequency band $0.05 \le \hat{f} \le 0.95$. The plots of Figure (3) were produced using a osculatory polynomial interpolant constructed to match the value of $\mathbf{A}'_{nl}(\hat{f}_i)$ for normalized frequencies $\hat{f}_i$ at the nine Chebyshev nodes

$$\hat{f}_i = .5 - .45\cos\left(\frac{(2i+1)\pi}{20}\right), \quad i = 0,\ldots,9 \tag{14}$$

as well as the derivative of $\mathbf{A}'_{nl}(\hat{f})$ at the interpolation point $\hat{f}_1$. A reduced order model of order 14 was then produced from this interpolant with K=7, $J_k^b = J_k^c = 2$ and interpolation points with normalized frequencies of $\hat{f}^{(k)}$=.1,.3,.4,.5,.6,.7, and .9. While the Chebyshev nodes are chosen for the $\hat{f}_i$ to generate a near minimax approximant as usual , the choice of the $\hat{f}^k$ at a different set of points clustered in the center of the band is motivated by a different observation: The polynomial interpolant is most reliable in the center of the band because the spectral domain Green's function is singular at DC ($\hat{f} = 0$) and at the onset of blazing modes ($\hat{f} = 1$). Notice that both the polynomial approximant and the reduced order model match the MoM results well over the entire band.

Table I contains timing and error results for the graphs plotted in Figure 3. The error values in Table I are the mean absolute error at 101 normalized frequency points spaced evenly between .05 and .95. The timing columns in the table were produced on a 266 MHz DEC Alpha workstation and can be described as follows: The MoM solve time is the number of seconds it takes to calculate the FSS reflection coefficient at a single frequency using the spectral Galerkin method. The polynomial interpolant setup time is the overhead involved in finding the coefficient matrices $\mathbf{A}_i$ of expansion (7), and the polynomial interpolant solve time is the time needed to then setup and solve system (9) at a single frequency. The reduced order model setup time is the overhead involved in constructing a single column of $\mathbf{V}$ and $\mathbf{Z}$ given the polynomial interpolant, and the reduced order model solve time is the time required to calculate the reflection coefficient for a single frequency given the reduced order model. The breakeven frequency is the number of frequencies that need to be calculated before the model reduction algorithm described here becomes cheaper than a straightforward application of the spectral Galerkin method.

Because narrow resonances are not uncommon in FSS systems (see Figure 3) the reflection coefficient of a given FSS must typically be calculated at one hundred or more frequencies to characterize it accurately. Once the breakeven frequency is reached, however, the solution of the reduced order system takes less than .1 ms at each frequency. Thus, the model reduction method presented here represents a large acceleration of the MoM; it is typically an order of magnitude faster if 200 or more frequencies are calculated.

4. Conclusions

A method for the fast calculation of the reflection coefficient of multiscreen FSSs over a large frequency band has been presented. The method is based on a polynomial interpolation of the system generated by the spectral Galerkin method, followed by a model reduction of the resulting system. The method is seen to reduce FSS scattering problems involving hundreds or thousands of unknowns to problems involving tens of unknowns with fairly minimal overhead. The algorithm was demonstrated through its successful application to several scattering problems, and was seen to yield a significant speed up relative to the straightforward spectral Galerkin method with very little loss of accuracy.

5. References

[1] R. Mittra, C. H. Chan and T. Cwik, "Techniques for analyzing frequency selective surfaces-A review," *Proceedings of the IEEE*, vol. 76, no. 12, pp. 1593-1615, 1988.

[2] E. Chiprout and M. S. Nakhla, *Asymptotic Waveform Evaluation and Moment Matching for Interconnect Analysis*. Boston: Kluwer Academic Publishers, 1994.

[3] I. Elfadel, J. Phillips, M. Silveira and J. White. "A Brief Survey of Krylov-subspace Based Model Order Reduction," *North American Radio Science Meeting*, Montreal, 1997, pp. 147.

[4] E. J. Grimme, "Krylov Projection Methods for Model Reduction," Ph. D. Thesis, University of Illinois at Urbana-Champaign, 1997.

[5] S. G. Kellison, *Fundamentals of Numerical Analysis*. Homewood, IL: Richard D. Irwin, Inc., 1975.

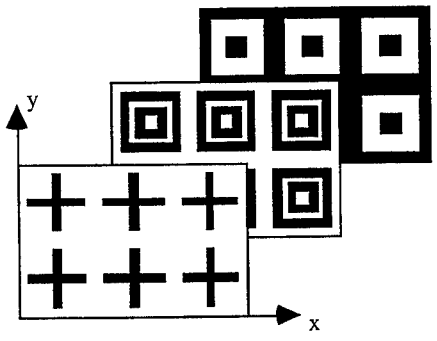


Figure 1. A typical multiscreen FSS.

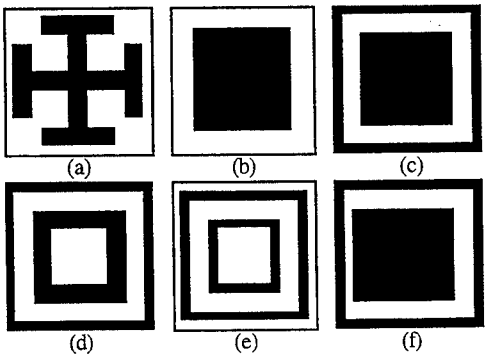


Figure 2. FSS Shape database.

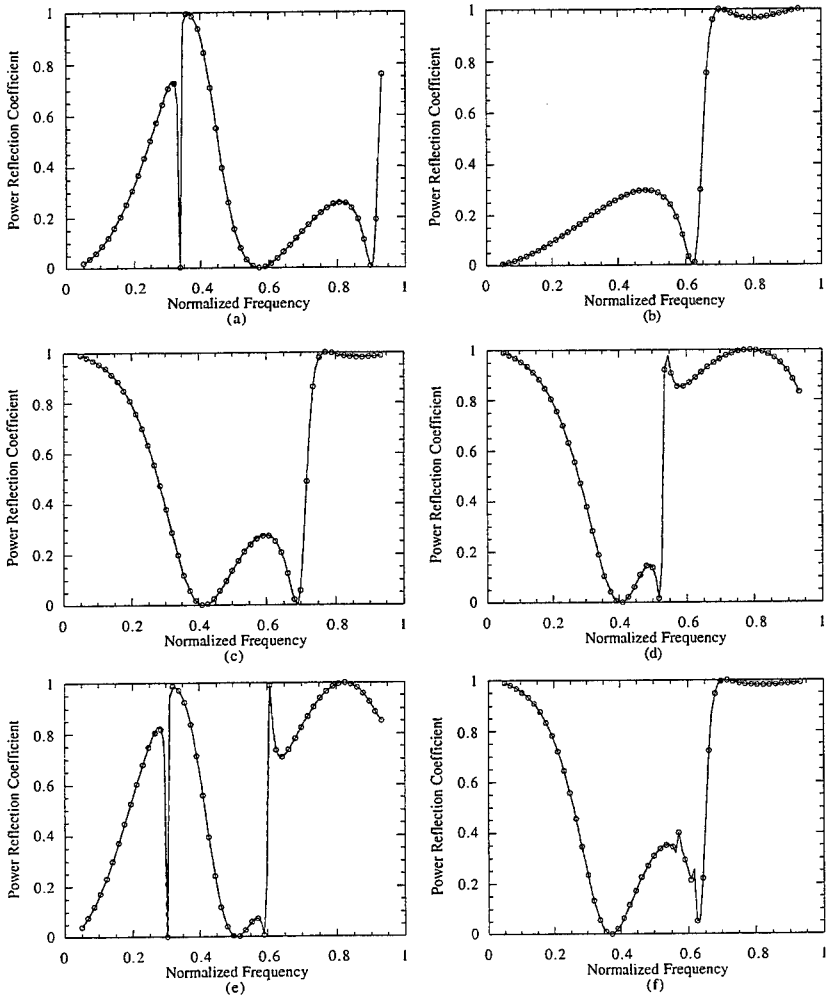


Figure 3. Comparison of MoM, polynomial interpolant and reduced order model solutions for six two-screen FSSs composed of the basic shapes shown in Figure 2. The MoM solution is represented by a solid line, the polynomial interpolant as small circles, and the reduced order model as a dashed line.

Table 1

Timing and Error Results
for Rational Interpolant Approximations to Several Different FSS Screens

Shape	MoM Solve (s/Freq.)	Poly. Int. Setup (s)	Poly. Int. Solve (s/Freq.)	Red. Mod. (s/Order)	Red. Mod. (ms/Freq.)	Breakeven Point	Error
a	2.02	30.3	0.779	1.35	0.416	25	0.00422
b	1.85	29.7	0.579	1.30	0.416	26	0.00145
c	2.23	29.9	0.994	1.80	0.406	25	0.00238
d	2.07	30.0	0.842	1.51	0.416	25	0.00257
e	1.81	29.7	0.522	0.835	0.416	23	0.00338
f	2.63	29.6	1.42	2.29	0.396	24	0.00227

SESSION 13:

FINITE ELEMENT METHOD

Chairs: J.R. Brauer and J-F. Lee

Comparing High Order Vector Basis Functions

J. Scott Savage
Ansoft Corporation
savage@ansoft.com

Abstract – The non–uniqueness of higher–order vector basis functions has led to publication of a variety of basis functions by many authors. This paper identifies important characteristics of these basis functions as applied to finite element analysis and provides for an educated choice of which basis functions are best suited for a particular application. This paper will also quantify the benefits of some previously published functions. As new sets of basis functions appear in the literature, it is imperative that the characteristics of these new functions be discussed, to demonstrate their relative worth.

I. Introduction

Vector basis functions have demonstrated the ability to accurately represent electromagnetic fields, via the finite element method, in complicated, inhomogeneous applications. When applied to the vector Helmholtz equation, the simplest such vector basis functions are often referred to as edge elements, or Whitney elements. More recently, as higher–order vector basis functions have become popular, these lowest order elements have become known variously as $H_0(\text{curl})$ [1,2], constant–tangent / linear–normal (CT/LN) elements [3], zero–order elements [4], and erroneously, first order elements [5]. These descriptions were developed to distinguish between the simplest elements, and elements of higher–order. This paper will use the notation of [1] and [2], which is compact, distinguishes between curl–conforming and divergence–conforming elements, and is consistent with the mathematical literature.

Since simplexes (triangles and tetrahedra) are best suited for modeling arbitrary geometry, this paper will focus on vector functions on simplex meshes. The general class of vector basis functions for finite element applications has been called tangential vector finite elements as well as curl–conforming finite elements. The elements are called "tangential" since they explicitly enforce tangential field continuity, while enforcing normal continuity only in a weak sense. The elements are called "curl–conforming" since the curl of any basis function is well–defined. This property is a direct result of strict tangential continuity.

There is no debate in the literature over the best form of the lowest order elements. The definitions of $H_0(\text{curl})$ elements differ only by an arbitrary scaling parameter, usually the length of the edge on which the function is defined. These elements are called zero–order, since they are complete to polynomial order zero (constant) in both the domain and range space of the curl operator. This completeness property is responsible for the convergence rates of these and all higher–order elements.

Unlike zero–order edge elements, higher–order elements are not uniquely specified. Many $H_1(\text{curl})$ functions can be found in the literature [2–7]. These elements have demonstrated appropriate convergence rates, yet have very different structure. Additionally, various $H_2(\text{curl})$ elements have been published [3–7], although the convergence rates of these elements has yet to be demonstrated.

Two different subcategories of basis functions have emerged in the literature: interpolatory and hierarchical. Interpolatory basis functions are vector analogs of the popular scalar basis functions on simplexes [8]. Each function in the basis set is of equal order, and the field interpolates to the value of individual functions at discrete mesh locations. Hierarchical basis sets, conversely, include all basis functions for every lower order. For example, each $H_0(\text{curl})$ basis function is included in a hierarchical $H_1(\text{curl})$ basis set.

II. Comparison Criteria

To evaluate the relative worth of various sets of vector basis functions, a set of criteria for comparing the functions is needed. The convergence rate of a basis set under mesh refinement is of primary importance. This property alone determines the order of a basis set. A demonstration of the convergence rate should always be included in the presentation of new basis sets. Next, the efficiency of the basis set is critical. The desired convergence rates should be obtained using the fewest possible unknowns. Nedelec has established the minimum number of degrees of freedom per simplex for a given convergence rate [1].

Beyond these two criteria, there is less certainty of the relative weight which should be given to other measures of performance. One important property of basis functions is matrix conditioning. When the basis functions are as nearly orthogonal as possible, the condition number of the global matrix will be reduced. This property greatly affects the performance of iterative matrix solution algorithms. Another important characteristic involves *p*–refinement, that is, using different order approximations in different cells of the same mesh. Hierarchical basis functions lend themselves well to *p*–refinement; although interpolatory functions do not prohibit *p*–refinement, they complicate it.

III. Description of Vector Elements

In this paper, simplex coordinate representations will be used to define vector basis functions. This convenient form allows various basis sets to be analyzed using a symbolic element matrix generation code developed by the author. With this tool, it is a simple task to evaluate any arbitrary basis set which can be written in simplex coordinates. Vector basis functions may be associated with several geometrical entities in a mesh of simplexes. The various incarnations of functions results from the explicit enforcement of tangential continuity. Vector functions may be associated with edges, triangular faces, or tetrahedral cells. When residing on edges, the basis functions are functions of only the two simplex coordinates associated with the nodes at the endpoint of the edge. Similarly, face functions are written in terms of the three simplex coordinates associated with the vertices of the triangle. Tetrahedral, or internal functions include all four simplex coordinate terms. For the purpose of analyzing a given order basis set, it is assumed that any edge function exists in identical form on all edges of the mesh. Likewise for face and internal functions. Thus, a basis set may be defined by listing all the edge functions for one edge, all the face functions for one face, and all the internal functions for one tetrahedron.

IV. Quantifying Basis Sets

The interpolatory basis sets proposed in [4], and a new hierarchical basis set, extended from those proposed in [9], were both analyzed using the symbolic matrix generation code. These elements are defined in Table 1 and Table 2, respectively. Due to the non–uniqueness of vector basis functions, alternative hierarchical $H_1(\text{curl})$ basis functions exist [7,10]. The hierarchical basis functions in Table 2 were chosen since they directly exhibit the proper contributions to the range and null space of the curl operator.

To determine convergence rates, a cavity resonator problem was simulated. The resonant frequencies were predicted using a restarted Lanczos algorithm that fully exploits the sparsity of the eigenvalue equation. As expected, equivalent order basis sets give identical resonant frequency predictions. This is true since interpolatory basis functions and hierarchical basis functions of equal order span exactly the same space. Fig. 1 demonstrates convergence rates for the dominant mode of a 2–D square resonator up to $H_5(\text{curl})$. Fig. 2 illustrates convergence rates for the dominant mode of a 3–D cubic resonator up to $H_6(\text{curl})$. The eigenvalue predictions using $H_p(\text{curl})$ elements for both 2–D and 3–D cavity problems, behave as $O\left(h^{2(p+1)}\right)$. Since the answers for each basis set are identical, and

they each use the minimal number of unknowns prescribed by Nedelec, the only distinguishing quantitative feature is matrix conditioning. Iterative matrix solvers perform best when the condition number of the matrix is low. To reduce the global matrix condition number, the basis functions should overlap as little as possible. One indicator of basis function conditioning relates to the element matrix,

$$T_{ij}^{e} = \iiint_{tet} \mathbf{B}_i \cdot \mathbf{B}_j dV \,. \tag{1}$$

where $\mathbf{B}_i$ represents the *i–th* basis function. This element matrix depends on the shape of the tetrahedron over which the integration is performed. Since most mesh generation packages strive to produce well–shaped cells, an equilateral triangle and equilateral tetrahedron were used to check the conditioning of various 2–D and 3–D basis sets, respectively.

When the basis functions are scaled so that the element matrix, $\mathbf{T}^e$, has uniform diagonal entries, the condition number, $\mathrm{cond}(\mathbf{T}^e)$, is a good indicator of the conditioning of the basis set. The condition number used in this paper is the ratio of the largest eigenvalue to the smallest eigenvalue. Table 3 and Table 4 give the condition numbers for 2–D and 3–D basis sets, respectively. It is apparent that interpolatory vector basis functions are much better conditioned than hierarchical vector basis functions, especially for high–order basis sets. This property should make interpolatory basis functions more desirable when implemented with an iterative matrix solver.

V. Conclusion

Higher–order vector basis functions provide more efficient solutions to Maxwell's equations when high accuracy is desired. Although the vector finite element space for higher–order basis sets has been well defined, the exact form of a particular vector basis set is not unique. This provides the finite element developer freedom in choosing a basis set. To assist in making this choice, this paper has provided a set of criteria for measuring the relative worth of vector basis functions. It was demonstrated that the two most popular vector basis types, interpolatory and hierarchical functions, correctly model Nedelec's finite element spaces with the minimal required degrees of freedom. Furthermore, it was proven that interpolatory vector basis functions are better conditioned than hierarchical vector basis functions.

VI. Acknowledgements

The author wishes to thank John Manges and Sergey Polstyanko for insightful discussions regarding this subject matter.

Table 1.
Interpolatory Vector Basis Functions

Order	Edge Functions	Face Functions	Tetrahedron Functions
$H_0(curl)$	$\lambda_1\nabla\lambda_2 - \lambda_2\nabla\lambda_1 = \Omega_{12}$		
$H_1(curl)$	$(3\lambda_1 - 1)\Omega_{12}$ $(3\lambda_2 - 1)\Omega_{12}$	$\lambda_1\Omega_{23}$ $\lambda_2\Omega_{13}$	
$H_2(curl)$	$(4\lambda_1 - 1)(4\lambda_1 - 2)\Omega_{12}$ $(4\lambda_1 - 2)(4\lambda_2 - 2)\Omega_{12}$ $(4\lambda_2 - 1)(4\lambda_2 - 1)\Omega_{12}$	$\lambda_1(4\lambda_1 - 1)\Omega_{23}$, $\lambda_2(4\lambda_1 - 1)\Omega_{13}$ $\lambda_1(4\lambda_2 - 1)\Omega_{23}$, $\lambda_2(4\lambda_2 - 1)\Omega_{13}$ $\lambda_1(4\lambda_3 - 1)\Omega_{23}$, $\lambda_2(4\lambda_3 - 1)\Omega_{13}$	$\lambda_1\lambda_2\Omega_{34}$ $\lambda_2\lambda_3\Omega_{14}$ $\lambda_3\lambda_4\Omega_{12}$
$H_3(curl)$	$(5\lambda_1 - 1)(5\lambda_1 - 2)(5\lambda_1 - 3)\Omega_{12}$ $(5\lambda_1 - 1)(5\lambda_1 - 2)(5\lambda_2 - 1)\Omega_{12}$ $(5\lambda_1 - 1)(5\lambda_2 - 1)(5\lambda_2 - 2)\Omega_{12}$ $(5\lambda_2 - 1)(5\lambda_2 - 2)(5\lambda_2 - 3)\Omega_{12}$	$\lambda_1(5\lambda_1 - 1)(5\lambda_1 - 2)\Omega_{23}$ $\lambda_1(5\lambda_1 - 1)(5\lambda_2 - 1)\Omega_{23}$ $\lambda_1(5\lambda_1 - 1)(5\lambda_3 - 1)\Omega_{23}$ $\lambda_1(5\lambda_2 - 1)(5\lambda_2 - 2)\Omega_{23}$ $\lambda_1(5\lambda_2 - 1)(5\lambda_3 - 1)\Omega_{23}$ $\lambda_1(5\lambda_3 - 1)(5\lambda_3 - 2)\Omega_{23}$ $\lambda_2(5\lambda_1 - 1)(5\lambda_1 - 2)\Omega_{13}$ $\lambda_2(5\lambda_1 - 1)(5\lambda_2 - 1)\Omega_{13}$ $\lambda_2(5\lambda_1 - 1)(5\lambda_3 - 1)\Omega_{13}$ $\lambda_2(5\lambda_2 - 1)(5\lambda_2 - 2)\Omega_{13}$ $\lambda_2(5\lambda_2 - 1)(5\lambda_3 - 1)\Omega_{13}$ $\lambda_2(5\lambda_3 - 1)(5\lambda_3 - 2)\Omega_{13}$	$\lambda_1\lambda_2(5\lambda_1 - 1)\Omega_{34}$ $\lambda_1\lambda_2(5\lambda_2 - 1)\Omega_{34}$ $\lambda_1\lambda_2(5\lambda_3 - 1)\Omega_{34}$ $\lambda_1\lambda_2(5\lambda_4 - 1)\Omega_{34}$ $\lambda_2\lambda_3(5\lambda_1 - 1)\Omega_{14}$ $\lambda_2\lambda_3(5\lambda_2 - 1)\Omega_{14}$ $\lambda_2\lambda_3(5\lambda_3 - 1)\Omega_{14}$ $\lambda_2\lambda_3(5\lambda_4 - 1)\Omega_{14}$ $\lambda_3\lambda_4(5\lambda_1 - 1)\Omega_{12}$ $\lambda_3\lambda_4(5\lambda_2 - 1)\Omega_{12}$ $\lambda_3\lambda_4(5\lambda_3 - 1)\Omega_{12}$ $\lambda_3\lambda_4(5\lambda_4 - 1)\Omega_{12}$

Table 2.
Hierarchical Vector Basis Functions

Order	Edge Functions	Face Functions	Tetrahedron Functions
$H_0(curl)$	$\lambda_1\nabla\lambda_2 - \lambda_2\nabla\lambda_1 = \Omega_{12}$		
$H_1(curl)$ additional	$\nabla[\lambda_1\lambda_2]$	$\lambda_1\Omega_{23}$ $\lambda_2\Omega_{13}$	
$H_2(curl)$ additional	$\nabla[\lambda_1\lambda_2(\lambda_1 - \lambda_2)]$	$\lambda_1^2\Omega_{23}$, $\lambda_2^2\Omega_{13}$, $\lambda_3^2\Omega_{12}$ $\nabla[\lambda_1\lambda_2\lambda_3]$	$\lambda_1\lambda_2\Omega_{34}$ $\lambda_2\lambda_3\Omega_{14}$ $\lambda_3\lambda_4\Omega_{12}$
$H_3(curl)$ additional	$\nabla[\lambda_1\lambda_2(2\lambda_1 - \lambda_2)(\lambda_1 - 2\lambda_2)]$	$\lambda_1^3\Omega_{23}$, $\lambda_2^3\Omega_{13}$, $\lambda_3^3\Omega_{12}$, $\lambda_1^2\lambda_3\Omega_{12}$ $\nabla[\lambda_1^2\lambda_2\lambda_3]$, $\nabla[\lambda_1\lambda_2\lambda_3^2]$	$\lambda_1\lambda_2\lambda_3\Omega_{14}$, $\lambda_1\lambda_2\lambda_3\Omega_{24}$ $\lambda_1\lambda_2\lambda_4\Omega_{13}$, $\lambda_1\lambda_2\lambda_4\Omega_{23}$ $\lambda_1\lambda_3\lambda_4\Omega_{12}$, $\lambda_1\lambda_3^2\Omega_{24}$ $\lambda_2^2\lambda_3\Omega_{14}$, $\lambda_2\lambda_3^2\Omega_{14}$ $\nabla[\lambda_1\lambda_2\lambda_3\lambda_4]$

Table 3.
Element Matrix Condition Numbers for 2–D Basis Functions

Basis Functions	Interpolatory	Hierarchical
$H_0(curl)$	2.00	2.00
$H_1(curl)$	10.53	65.85
$H_2(curl)$	47.54	750.60
$H_3(curl)$	262.37	31635.95
$H_4(curl)$	1229.60	268730.88

Table 4.
Element Matrix Condition Numbers for 3–D Basis Functions

Basis Functions	Interpolatory	Hierarchical
$H_0(curl)$	2.50	2.50
$H_1(curl)$	25.28	141.50
$H_2(curl)$	237.79	2714.99
$H_3(curl)$	2008.06	189517.27

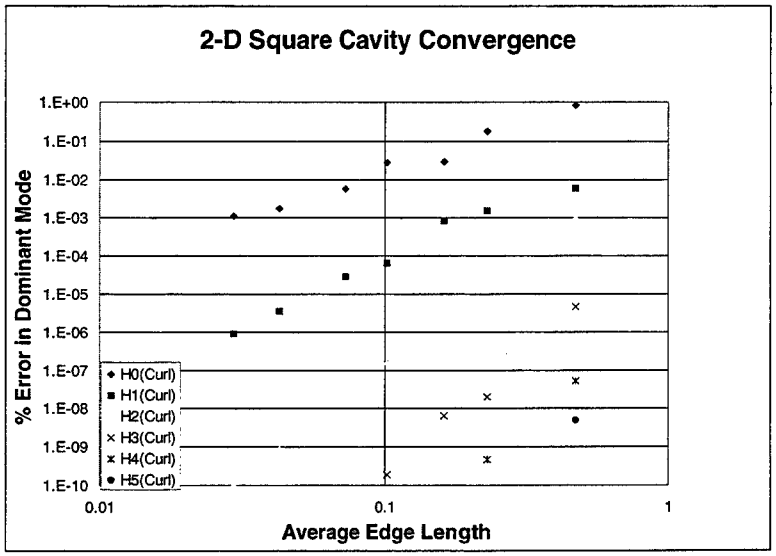


Figure 1.

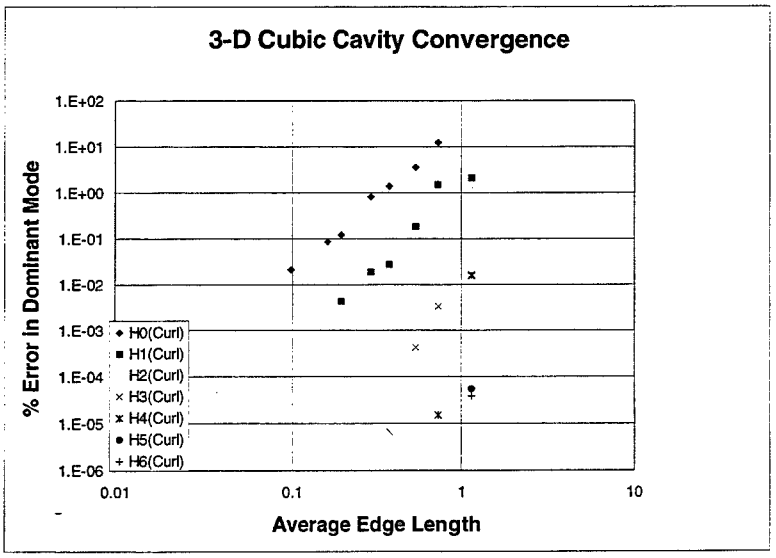


Figure 2.

VI. REFERENCES

[1] J. C. Nedelec, "Mixed finite elements in R3," *Numer. Math.*, vol. 35, pp. 315–341, 1980.

[2] J.–F. Lee, D.–K. Sun, and Z. J. Cendes, "Full–wave analysis of dielectric waveguides using tangential vector finite elements," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 1262–1271, Aug., 1991.

[3] J. S. Savage and A. F. Peterson, "Higher–order vector finite elements for tetrahedral cells," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 874–879, June, 1996.

[4] R. D. Graglia, D. R. Wilton, and A. F. Peterson, "Higher order interpolatory vector bases for computational electromagnetics," *IEEE Trans. Ant. Prop.*, vol. 45, pp. 329–342, Mar. 1997.

[5] T. V. Yioultsis and T. D. Tsiboukis, "Development and implementation of second and third order vector finite elements in various 3–D electromagnetic field problems," *IEEE Trans. Mag.*, vol. 33, pp. 1812–1815, Mar. 1997.

[6] Z. J. Cendes, "Vector finite elements for electromagnetic field computation," *IEEE Trans. Mag.*, vol. 27, pp. 3958–3966, Sept. 1991.

[7] J. P. Webb and B. Forghani, "Hierarchal scalar and vector tetrahedra," *IEEE Trans. Mag.*, vol. 29, pp. 1495–1498, Mar. 1993.

[8] P. P. Silvester and R. L. Ferrari, *Finite Elements for Electrical Engineers*, Cambridge: Cambridge University Press, 1983.

[9] J.–Y. Wu and R. Lee, "Construction of the basis functions for Nedelec's finite element spaces for triangular and tetrahedral meshes," *Abstracts of the URSI Radio Science Meeting*, Montreal, Canada, p. 38, July 1997.

[10] L. S. Andersen and J. L. Volakis, "A novel class of hierarchical higher order tangential vector finite elements for electromagnetics," *Proc. IEEE Ant. Prop. Soc. Inter. Symp.*, Montreal, Canada, pp. 648–651, July 1997.

Mesh Refinement for Hybrid FEM in the Analysis of Printed Antennas and Arrays[†]

Jian Gong, David Bernstein and Scott Wedge
Tanner Research, Inc.
2650 East Foothill Blvd.
Pasadena, CA 91107

Abstract – Techniques are presented to improve the efficiency and accuracy of the hybrid finite element method when applied to printed circuit antennas and arrays. To improve efficiency, we use a prismatic meshing approach that dramatically reduces the number of finite elements required for analyzing thin planar structures as compared to the tetrahedral alternative. For accuracy improvements, we demonstrate how resonant frequencies of narrow band printed antennas can be predicted with much greater accuracy and confidence given a suitable mesh refinement scheme.

I. Introduction

Full wave techniques have been employed for the electromagnetic analysis of antennas and microwave circuits for years [1,2], yet no single approach is known to meet every real-world engineering design need. One major difficulty is the tradeoff between efficiency and generality that must be made when selecting a numerical approach. Moment method approaches generally offer excellent efficiency, but can only handle structure types that can be characterized with Green's functions. Finite element methods offer the most generality, but can be very inefficient when used to analyze open structures.

Few engineering problems face this dilemma more profoundly than the case of printed circuit antennas and arrays. Here, moment method approaches seem more suitable, but the complex feed structures and multi-layer substrate/superstrate combinations commonly used rule out a Green's function approach. Finite element methods, which can address feed structure complexities and substrate inhomogeneity, are poorly suited to efficiently handle the radiation aspect of the problem.

In this work we employ a 3-D hybrid finite element method (FEM) and boundary element method (BEM) take advantage of both MoM and FEM approaches. The hybrid FEM/BEM approach alone, however, does not guarantee adequate speed, efficiency and accuracy. Here we present specific mesh generation, element, and refinement schemes that make possible fast and accurate analyses using a personal computer.

II. Hybrid FEM with 3-D Prismatic Mesh Elements

A goal of our approach is to analyze arbitrary patch shapes for planar antennas in a comparatively general and efficient manner. Since we are dealing with planar layered substrates, 2-D triangulation meshing can be utilized to account for non-rectangular patch shapes. A uniform discretization in the third dimension (extrusion along the z axis) takes into account layered substrate-configurations. The resulting right-angled prism mesh elements allow for general

[†] This work was supported by the US Army Research Office, Research Triangle Park, North Carolina, under contract DAAH04-96-C-0049.

inhomogeneous, lossy and anisotropic material fillings, and printed and non-printed feed line modeling. Prismatic meshing results in a tremendous reduction in the number of finite elements required compared to a tetrahedral meshing approach, especially for thin substrates. In the patch antenna examples presented here, prismatic meshing results in 500-2,000 unknowns, while a comparable tetrahedral approach would need 10,000-100,000 unknowns to match the system condition.

As is conventional with hybrid approaches, our FEM subsystem handles the truncated domain of layered substrates including non-printed feed structures. The FEM analysis is performed over a restricted portion of the substrate within a fictional perfect electric conductor wall (referred to as the *cavity*). When placed certain distance from the antenna the cavity walls have a negligible effect. The BEM subsystem handles the antenna radiating elements with possible printed feed lines. We solve the coupled FEM/BEM system using a BiCG iterative solver. For printed antenna modeling, simple diagonal preconditioning appears sufficient to achieve efficient convergence.

III. Mesh Generation for Efficient Antenna Analysis

Our approach of extruding 3-D meshes from 2-D triangulation was designed specifically for printed circuit antennas and other multi-layer planar structures. Given a layer based design, meshing proceeds in two phases. First, a list of line segments defining the geometry is extracted from the user prescribed physical layout. Second, the extracted 2-D geometric input is triangulated using an algorithm which respects the input geometry (points and line segments).

The 2-D triangulation code uses the standard Bowyer-Watson point insertion method (with implementation along the lines of that found in [3]). Edge-swapping and node relocation smoothing algorithms, similar to those described in [4,5], are included as options. The triangulation code was developed with adaptive mesh refinement in mind; hence it initially computes the coarsest mesh consistent with the geometric input (the locations of line segments) and with an overall triangle quality measure. In general, highly graded triangular meshes are produced with a minimum angle between 30 and 35 degrees. The resulting 2-D triangulation is extruded to the 3-D prismatic mesh as a preprocessing step for the hybrid FEM/BEM engine taking into account substrate material properties. This step is independent of the 2-D mesh code.

As a demonstration of modeling efficiency, we present in Figure 1 the radiation pattern analysis results of a linear patch antenna array with 8 radiating elements. By using the prismatic elements, accurate radiation pattern (far-field) results can be obtained with very efficient sampling of the structure. These calculations took only a few minutes using a Pentium PC (266MHz/64MB). Figure 1(a) shows the linear 8 element array, the antenna feed points, and the outer cavity wall. Figure 1(b) shows the resulting mesh (smoothing was not used in this case). Figure 1(c) shows the predicted radiation pattern of the array using uniform feeding. The corresponding array factor calculation is plotted for comparison. Evident from the comparison, little coupling was present in this example.

IV. Mesh Refinement for Improved Accuracy

An often overlooked difficulty associated with full-wave approaches applied to narrow band patch antennas is the accurate prediction of resonant frequencies. Errors on the order of 5-10% are common even when following the widely suggested sampling rules (20 samples/wavelength for near fields). This problem is not solved in practice by simply refining the overall mesh. This does help approach the exact solution, but rapidly increases the size of the numerical system. We

have observed that the rate of convergence to an exact solution is usually relatively slow compared to the increase in system size.

Because of our ability to handle unstructured meshes, we explored a local mesh refinement scheme to overcome this dilemma. An example of our triangulation code applied to a simple antenna problem is shown in Figure 2. Here we consider a rectangular patch 11.43×7.62 cm housed in a 19.43×15.62 cm cavity. The substrate ($\varepsilon_r = 2.62$) is 0.15875 cm thick. Note that the ratio of the cavity's length to its depth is over 122, making it impractical to use a tetrahedral FEM approach. On the left is the coarsest mesh consistent with the location of the contour of the antenna layout; on the right a locally refined mesh has been produced where the areas close to the patch's longer edges have been refined by the introduction of two artificial lines. Here we have localized the mesh at the two resonant patch edges to improve the modeling of antenna edge effects and electrical length.

As seen in Figure 3, this localized refinement at the patch edges dramatically reduces errors. Accuracy near 1% is obtained by sampling near 50 elements/wavelength. Since the mesh refinement is localized, the increase in size of the numerical system via this refinement is minimal. We have applied a similar edge refinement approach to circular patch antennas and achieved quite consistent results.

As a demonstration, Figure 4 shows the input impedance loci of the rectangular patch antenna at three different feed locations. Figure 5 shows results using mesh refinement, compared with measured data. The analysis time is on the order of a minute/frequency on a 266MHz/64MB Pentium PC.

V. Conclusions

We have presented a hybrid FEM/BEM technique using right-angle prism meshing in conjunction with refinement schemes for efficient and accurate printed antenna modeling. As has been demonstrated, accuracy within 1% can be obtained in predicting patch antenna resonant frequency with minimal increases in computational complexity. We are currently working on automatic refinement schemes and suitable error estimation methods to automate our mesh refinement approach.

References

[1] T. Itoh, G. Pelosi and P. Silvester (ed.), "Finite Element Software For Microwave Engineering," *John Wiley & Sons, Inc.*, 1996

[2] D. Pozar and D. Schaubert (ed.), "Microstrip Antennas: The Analysis and Design of Microstrip Antennas and Arrays," *IEEE Press*, 1995

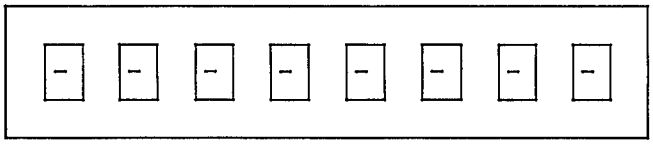
[3] Rebay, S., " Efficient Unstructured Mesh Generation by Means of Delaunay Triangulation and Bowyer-Watson Algorithm," Journal of Computational Physics, **106**, 125-138 (1993).

[4] Zavattieri P.D., E.A. Dari, and G.C. Buscaglia, "Optimization Strategies in Unstructured Mesh Generation," International Journal for Numerical Methods in Engineering, **39**, 2055-2071 (1996).

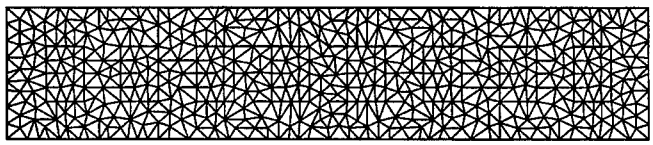
[5] Frey, W.H. and D. A. Field, "Mesh Relaxation: A New Technique for Improving Triangulations," International Journal for Numerical Methods in Engineering, **31**, 1121-1133 (1991).

[6] J.R. James, P.S. Hall and C. Wood, "Microstrip Antenna Theory and Design," *Peter Peregrinus,* London, 1981

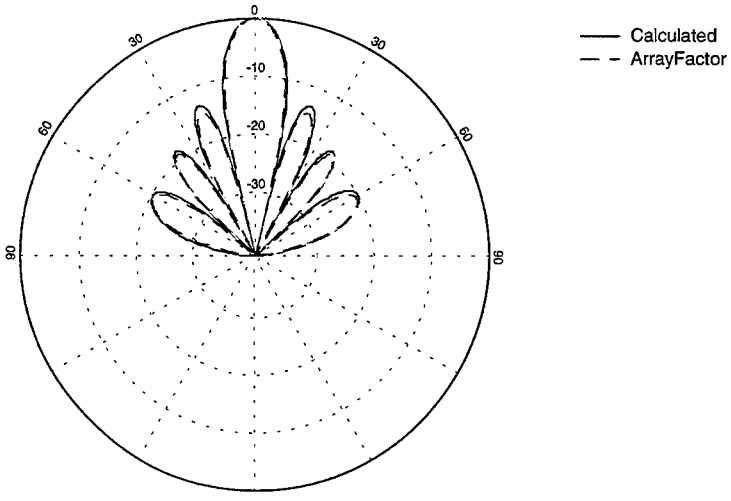
[7] Y.T. Lo, D. Solomon and W.F. Richards, "Theory and Experiment on Microstrip Antennas," *IEEE Trans. Antenna Propagat.*, Vol.27, pp.137-145, March 1979



(a)



(b)



(c)

Figure 1. Layout (a), meshing (b), and radiation patterns (c) of a linear 8-element patch antenna array. Full wave and array factor based calculations are shown for comparison.

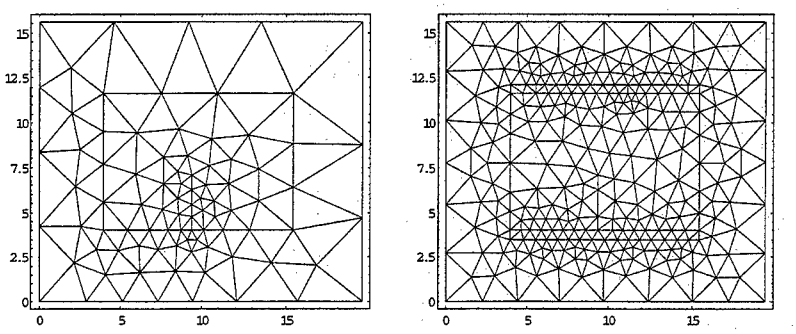


Figure 2. Mesh examples for rectangular patch antennas. On the left is the coarsest mesh consistent with the location of the antenna contour; on the right is the mesh with an artificially introduced lines along the patch's edges.

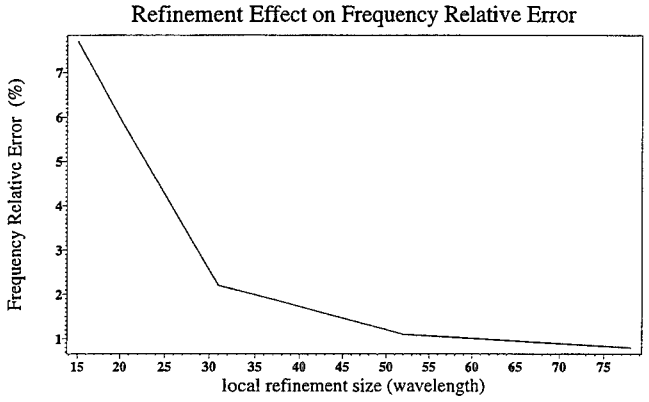


Figure 3. Effect of refinement on resonant frequency relative error.

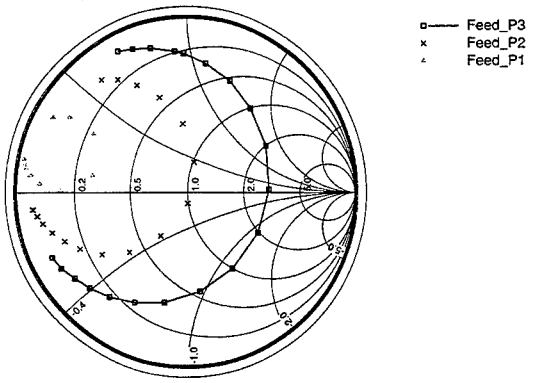


Figure 4. Input impedance loci of a rectangular patch antenna at difference feed locations: 3.05, 2.29 and 0.76 cm from the patch edge, respectively.

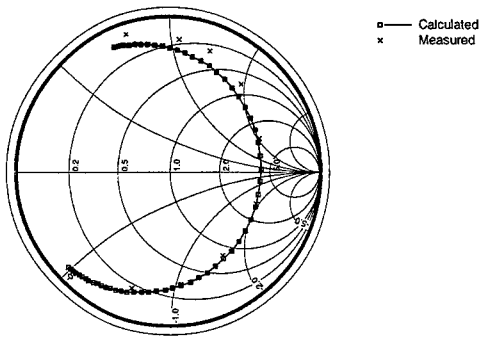


Figure 5. Comparison of analysis and measurement [6,7] results for the rectangular patch.

A Novel, Efficient Algorithm for Scattering from a Complex BOR Using Vector FEM and PML

Andrew D. Greenwood* and Jian-Ming Jin
Center for Computational Electromagnetics
Department of Electrical and Computer Engineering
University of Illinois
Urbana, Illinois 61801

*Also affiliated with Air Force Research Laboratory at Rome/IFSB, Rome, NY 13441

Abstract

An efficient finite element method (FEM) algorithm to compute scattering from a complex body of revolution (BOR) is developed. The method uses edge-based (vector) basis functions to expand the transverse field components and node-based functions to expand the E_ϕ field components. The use of vector basis functions eliminates the problem of spurious solutions suffered by other three component FEM formulations. Perfectly matched layer (PML) absorbers in cylindrical coordinates are used to truncate the mesh. Because PML absorbers are available in cylindrical coordinates, the method is efficient for arbitrarily shaped scatterers. The FEM equations are solved by ordering the unknowns with a reverse Cuthill-McKee algorithm and applying a banded-matrix solution algorithm. The method is capable of handling large radar targets, and good agreement with measured results is achieved for benchmark targets.

1 Introduction

Recent extensions of perfectly matched layer (PML) absorbers to cylindrical coordinates [1–3] make possible the development of an efficient finite element method (FEM) algorithm for scattering from a complex body of revolution (BOR). Past FEM algorithms for scattering from a BOR have employed either a coupled azimuth potential formulation [4–6] or a three-component, node-based formulation [7]. In contrast, this work uses edge-based (vector) basis functions to expand the transverse field components (E_ρ and E_z) and node-based functions to expand the E_ϕ field components. Such an arrangement avoids the problem of spurious modes suffered by the three-component, node-based formulation [7], and, in contrast to the coupled azimuth potential formulation, the unknowns in this arrangement obey the standard vector wave equation, making the application of PML absorbers much simpler. Because PML absorbers are now available in cylindrical coordinates, the method is also very efficient for long, narrow scatterers, which require many extra unknowns if the mesh must be truncated by a spherical absorbing boundary condition.

2 Formulation

The formulation starts with the illumination of an axisymmetric target by a uniform plane wave. A slice of a typical computational domain is depicted in Figure 1. The electric field in this problem obeys the three-dimensional (3-D) vector wave equation with the boundary conditions

$$\hat{n} \times \mathbf{E} = 0 \quad \text{on} \quad S_1 \tag{1}$$

$$\frac{1}{\mu_r}\hat{n} \times (\nabla \times \mathbf{E}) + \gamma_e \hat{n} \times \hat{n} \times \mathbf{E} = 0 \quad \text{on} \quad S_2 \tag{2}$$

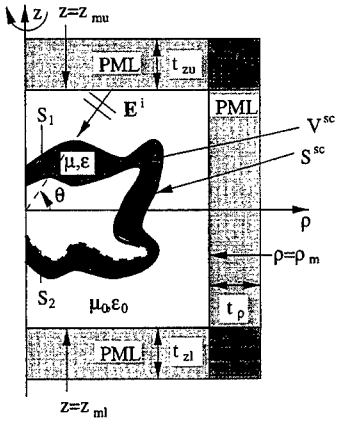


Figure 1: Slice of a typical target.

where $S_1 + S_2$ makes up the surface of an impenetrable scatterer. To solve this problem using FEM, the mesh must be truncated at an artificial boundary. To avoid introducing spurious reflected waves from this artificial boundary, PML is introduced as shown in Figure 1. The constitutive parameters of medium with the PML are [2]

$$\overline{\overline{\mu}} = \mu_0 \mu_r \overline{\overline{\Lambda}}; \quad \overline{\overline{\epsilon}} = \epsilon_0 \epsilon_r \overline{\overline{\Lambda}} \tag{3}$$

where $\overline{\overline{\Lambda}}$ is a diagonal tensor given by

$$\overline{\overline{\Lambda}} = \hat{\rho}\hat{\rho}\Lambda_\rho + \hat{\phi}\hat{\phi}\Lambda_\phi + \hat{z}\hat{z}\Lambda_z \tag{4}$$

with entries

$$\Lambda_\rho = \frac{S_z \tilde{\rho}}{S_\rho \rho}; \quad \Lambda_\phi = \frac{S_z S_\rho \rho}{\tilde{\rho}}; \quad \Lambda_z = \frac{S_\rho \tilde{\rho}}{S_z \rho} \tag{5}$$

in which

$$S_\rho = S_\rho(\rho) = \begin{cases} 1 & 0 \le \rho \le \rho_m \\ 1 - j\alpha\left(\dfrac{\rho - \rho_m}{t_\rho}\right)^2 & \rho > \rho_m \end{cases} \tag{6}$$

$$S_z = S_z(z) = \begin{cases} 1 - j\alpha\left(\dfrac{z_{ml} - z}{t_{zl}}\right)^2 & z < z_{ml} \\ 1 & z_{ml} \le z \le z_{mu} \\ 1 - j\alpha\left(\dfrac{z - z_{mu}}{t_{zu}}\right)^2 & z > z_{mu} \end{cases} \tag{7}$$

$$\tilde{\rho} = \begin{cases} \rho & 0 \le \rho \le \rho_m \\ \rho - j\alpha\dfrac{(\rho - \rho_m)^3}{3t_\rho^2} & \rho > \rho_m. \end{cases} \tag{8}$$

With these constitutive parameters, the wave vector equation is

$$\nabla \times \frac{1}{\mu_r} \overline{\overline{\Lambda}}^{-1} \cdot \nabla \times \mathbf{E} - k_0^2 \epsilon_r \overline{\overline{\Lambda}} \cdot \mathbf{E} = 0 \tag{9}$$

where $k_0 = \omega\sqrt{\mu_0 \epsilon_0}$ is the free space wave number.

According the generalized variational principle, the problem defined by Equations 9, 1, and 2 can be found by extremizing the functional [8]

$$F(\mathbf{E}) = \frac{1}{2} \iiint_V \left[\frac{1}{\mu_r} (\nabla \times \mathbf{E}) \cdot \overline{\overline{\Lambda}}^{-1} \cdot (\nabla \times \mathbf{E}) - k_0^2 \epsilon_r \mathbf{E} \cdot \overline{\overline{\Lambda}} \cdot \mathbf{E} \right] dV + \frac{1}{2} \iint_{S_2} \gamma_e \left[\mathbf{E} \cdot \mathbf{E} - (\hat{n} \cdot \mathbf{E})(\hat{n} \cdot \mathbf{E}) \right] dS. \tag{10}$$

This functional is converted to $F(\mathbf{E}^s)$ by substituting $\mathbf{E} = \mathbf{E}^i + \mathbf{E}^s$ and discarding terms which do not depend on $\mathbf{E}^s$,

$$\begin{aligned} F(\mathbf{E}^s) =& \frac{1}{2} \iiint_V \left[\frac{1}{\mu_r} (\nabla \times \mathbf{E}^s) \cdot \overline{\overline{\Lambda}}^{-1} \cdot (\nabla \times \mathbf{E}^s) - k_0^2 \epsilon_r \mathbf{E}^s \cdot \overline{\overline{\Lambda}} \cdot \mathbf{E}^s \right] dV \\ &+ \frac{1}{2} \iint_{S_2} \gamma_e \left[\mathbf{E}^s \cdot \mathbf{E}^s - (\hat{n} \cdot \mathbf{E}^s)(\hat{n} \cdot \mathbf{E}^s) \right] dS + \iint_{S_2} \gamma_e \left[\mathbf{E}^s \cdot \mathbf{E}^i - (\hat{n} \cdot \mathbf{E}^s)(\hat{n} \cdot \mathbf{E}^i) \right] dS \\ &+ \iiint_{V^{sc}} \left[\frac{1}{\mu_r} (\nabla \times \mathbf{E}^s) \cdot \overline{\overline{\Lambda}}^{-1} \cdot (\nabla \times \mathbf{E}^i) - k_0^2 \epsilon_r \mathbf{E}^s \cdot \overline{\overline{\Lambda}} \cdot \mathbf{E}^i \right] dV - \iint_{S^{sc}} \mathbf{E}^s \cdot (\hat{n} \times \nabla \times \mathbf{E}^i) dS. \end{aligned} \tag{11}$$

To take advantage of the azimuthal symmetry of the problem, both the incident and the scattered fields are expanded in Fourier modes

$$\mathbf{E} = \sum_{m=-\infty}^{\infty} \left[\mathbf{E}_{t,m}(\rho, z) + \hat{\phi} E_{\phi,m}(\rho, z) \right] e^{jm\phi} \tag{12}$$

and the ϕ integration is carried out, yielding a functional which can be extremized in two dimensions,

$$\begin{aligned} F(\mathbf{E}^s) =& 2\pi \sum_{m=-\infty}^{\infty} \left\{ \frac{1}{2} \iint_\Omega \left\{ \frac{\rho}{\mu_r} \left[\frac{1}{\Lambda_\phi} (\nabla_t \times \mathbf{E}^s_{t,-m}) \cdot (\nabla_t \times \mathbf{E}^s_{t,m}) + (\nabla_t E^s_{\phi,-m} + \frac{jm}{\rho} \mathbf{E}^s_{t,-m} + \frac{\hat{\rho}}{\rho} E^s_{\phi,-m}) \right. \right. \right. \\ &\left. \cdot \overline{\overline{\Lambda}}_t'^{-1} \cdot (\nabla_t E^s_{\phi m} - \frac{jm}{\rho} \mathbf{E}^s_{t,m} + \frac{\hat{\rho}}{\rho} E^s_{\phi,m}) \right] - k_0^2 \epsilon_r \rho \left[\mathbf{E}^s_{t,-m} \cdot \overline{\overline{\Lambda}}_t \cdot \mathbf{E}^s_{t,m} + \Lambda_\phi E^s_{\phi,-m} E^s_{\phi,m} \right] \Bigg\} d\Omega \\ &+ \frac{1}{2} \int_{C_2} \gamma_e \rho \left[\mathbf{E}^s_{t,-m} \cdot \mathbf{E}^s_{t,m} - (\hat{n} \cdot \mathbf{E}^s_{t,-m})(\hat{n} \cdot \mathbf{E}^s_{t,m}) + E^s_{\phi,-m} E^s_{\phi,m} \right] d\ell \\ &+ \iint_{\Omega^{sc}} \left\{ \frac{\rho}{\mu_r} \left[(\nabla_t \times \mathbf{E}^s_{t,-m}) - \hat{\phi} \times (\nabla_t E^s_{\phi,-m} + \frac{jm}{\rho} \mathbf{E}^s_{t,-m} + \frac{\hat{\rho}}{\rho} E^s_{\phi,-m}) \right] \cdot \overline{\overline{\Lambda}}^{-1} \cdot (\nabla \times \mathbf{E}^i_m) \right. \\ &\left. - k_0^2 \epsilon_r \rho \left[\mathbf{E}^s_{t,-m} \cdot \overline{\overline{\Lambda}}_t \cdot \mathbf{E}^i_{t,m} + \Lambda_\phi E^s_{\phi,-m} E^i_{\phi,m} \right] \right\} d\Omega \\ &+ \int_{C_2} \gamma_e \rho \left[\mathbf{E}^s_{t,-m} \cdot \mathbf{E}^i_{t,m} - (\hat{n} \cdot \mathbf{E}^s_{t,-m})(\hat{n} \cdot \mathbf{E}^i_{t,m}) + E^s_{\phi,-m} E^i_{\phi,m} \right] d\ell \\ &- \int_{C^{sc}} \frac{\rho}{\mu_r} \left[\mathbf{E}^s_{t,-m} + \hat{\phi} E^s_{\phi,-m} \right] \cdot \left[\hat{n} \times (\nabla \times \mathbf{E}^i_m) \right] d\ell \Bigg\} \end{aligned} \tag{13}$$

where

$$\overline{\overline{\Lambda}}_t = \hat{\rho}\hat{\rho}\Lambda_\rho + \hat{z}\hat{z}\Lambda_z; \quad \overline{\overline{\Lambda}}'_t = \hat{\rho}\hat{\rho}\Lambda_z + \hat{z}\hat{z}\Lambda_\rho. \tag{14}$$

To extremize Equation 13 (subject to Equation 1), FEM expansions are substituted for $\mathbf{E}^s_{t,m}$ and $E^s_{\phi,m}$. The FEM expansions depend on the mode number m because the conditions that the fields must satisfy along the z-axis depend m. Along the z-axis, the boundary conditions are

$$E_\rho = E_\phi = (\nabla \times \mathbf{E})_\rho = (\nabla \times \mathbf{E})_\phi = 0 \qquad \text{for } m = 0; \tag{15}$$

$$E_\rho = \mp j E_\phi, \quad (\nabla \times \mathbf{E})_\rho = \mp j (\nabla \times \mathbf{E})_\phi, \quad E_z = (\nabla \times \mathbf{E})_z = 0 \quad \text{for } m = \pm 1; \tag{16}$$

$$E_\rho = E_\phi = E_z = (\nabla \times \mathbf{E})_\rho = (\nabla \times \mathbf{E})_\phi = (\nabla \times \mathbf{E})_z = 0 \qquad \text{for } |m| > 1. \tag{17}$$

FEM expansions which satisfy these conditions are [9]

$$E_{\phi,0} = \sum_{i=1}^{3} e^e_{\phi i} \rho N^e_i, \quad \mathbf{E}_{t,0} = \sum_{i=1}^{3} e^e_{ti} \mathbf{N}^e_i \qquad \text{for } m = 0; \tag{18}$$

$$E_{\phi,\pm 1} = \sum_{i=1}^{3} e^e_{\phi i} N^e_i, \quad \mathbf{E}_{t,\pm 1} = \sum_{i=1}^{3} \left[\mp j \hat{\rho} e^e_{\phi i} N^e_i + e^e_{ti} \rho \mathbf{N}^e_i \right] \quad \text{for } m = \pm 1; \tag{19}$$

$$E_{\phi,m} = \sum_{i=1}^{3} e^e_{\phi i} \rho N^e_i, \quad \mathbf{E}_{t,m} = \sum_{i=1}^{3} e^e_{ti} \rho \mathbf{N}^e_i \qquad \text{for } |m| > 1 \tag{20}$$

where N^e_i is a standard 2-D nodal-element basis function, and $\mathbf{N}^e_i$ is a standard 2D edge-element basis function. Substituting Equation 18, 19, or 20 into Equation 13, differentiating, setting the result to zero, and taking advantage of symmetry between the m and $-m$ terms yields a system of the form

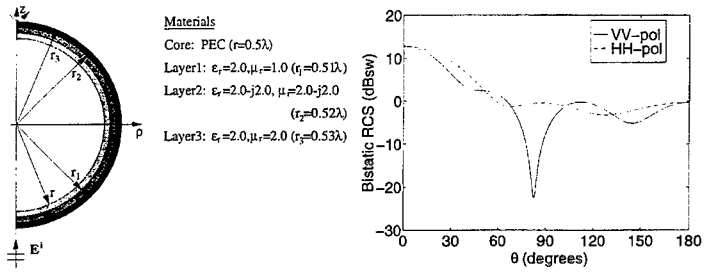
$$\begin{bmatrix} A^m_{tt} & A^m_{t\phi} \\ A^m_{\phi t} & A^m_{\phi\phi} \end{bmatrix} \begin{Bmatrix} e^m_t \\ e^m_\phi \end{Bmatrix} = \begin{Bmatrix} B^m_t \\ B^m_\phi \end{Bmatrix}. \tag{21}$$

Note that the system matrix is sparse and symmetric. Also note that, although the summation in Equation 13 is over all positive and negative numbered modes, it can be shown that

$$\{e^m_t\} = \begin{cases} \{e^{-m}_t\} & \text{V-pol incidence} \\ -\{e^{-m}_t\} & \text{H-pol incidence} \end{cases}; \quad \{e^m_\phi\} = \begin{cases} -\{e^{-m}_\phi\} & \text{V-pol incidence} \\ \{e^{-m}_\phi\} & \text{H-pol incidence} \end{cases} \tag{22}$$

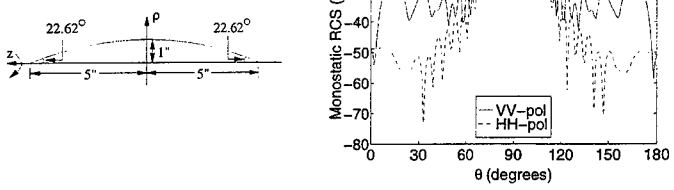
for all m. Use of this relation decreases the computational work by half. A rule of thumb for the number of modes required for a convergent solution is [10] $M_{\max} = k_0 \rho_{\max} \sin\theta + 6$. This rule of thumb is valid for $k_0 \rho_{\max} \sin\theta > 3$.

When the unknowns in Equation 21 are appropriately ordered, the system matrix is highly banded. Thus, to solve Equation 21, the unknowns are first ordered in a reverse Cuthill-McKee ordering [11]. The LDL^T decomposition of the matrix is then computed using a band solver [8]. The solution is finally found by forward and back substitution on the resulting triangular systems. Using this method, the solution for multiple excitation vectors is computed while the LDL^T decomposition of the matrix is computed only once.



(a) Geometry (b) Computed RCS

Figure 2: Computed bistatic RCS of a coated sphere. Compare to results in [7].



(a) Geometry (b) Computed RCS at 9 GHz

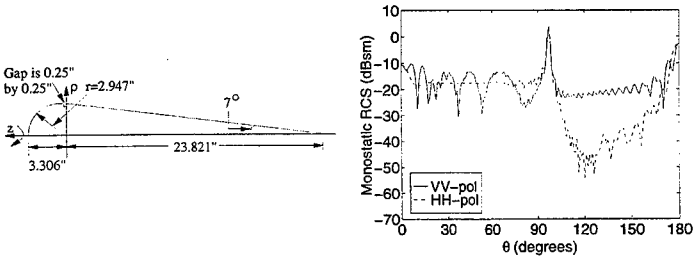
Figure 3: Computed monostatic RCS of a coated sphere. Compare to measured results in [12].

3 Numerical Examples

Three examples of computations by this method are presented here. In each of the examples, the mesh is truncated by a five layer PML with $\alpha = 5.0$. The PML interface is place approximately 0.25 λ from the surface of the scatterer, and unless otherwise noted, the mesh length is $\lambda/20$.

The geometry for the first example is the coated sphere shown in Figure 3. This example is considered in [7] as well. The bistatic radar cross-section (RCS) of the sphere is shown in Figure 3, and agrees well with the both the FEM result and the method of moments (MoM) comparison result presented in [7]. However, the authors of [7] point out that their FEM code does not treat magnetic materials properly, and although their FEM result agrees well with their MoM comparison result, inaccurate results can arise from magnetic materials. The use of edge elements in the current method eliminates this source of potential inaccuracy.

The next example considers the metallic ogive shown in Figure 3. This is one of the benchmark targets presented in [12]. At 9 GHz, the ogive is 7.63λ long, and its computed monostatic RCS is shown



(a) Geometry

(b) Computed RCS at 9 GHz

Figure 4: Computed monostatic RCS of a conesphere with a gap. Compare to measured results in [12].

at this frequency. Measured results are presented in [12]. Good agreement is observed between the computed results presented here and the measured results presented in [12].

The final example is a metallic conesphere with a gap, shown in Figure 3. This target is also one of the benchmark targets in [12], and it is 20.69λ long at 9 GHz. The computed monostatic RCS at this frequency is shown in Figure 3, and measured results are found in [12]. In computing this result, parts of the mesh between the conesphere and the PML are coarsened to mesh length $\lambda/10$ rather than $\lambda/20$, and there is still good agreement between the computed results and the measured results in [12]. Further, because the conesphere is over 20 wavelengths long at 9 GHz, this example also shows that the method is capable of handling large radar targets.

4 Conclusion

A novel, efficient algorithm to compute scattering from a complex BOR using vector FEM and PML is developed. Recent extensions of PML to cylindrical coordinates allow for efficient and accurate mesh truncation. Importantly, the mesh is truncated on a cylindrical boundary rather than a spherical boundary. This is much more efficient for long, narrow scatterers. Further, the use of the electric field components as the unknown values allows simple implementation of the PML. The use of edge elements in the formulation prevents spurious solutions and allows inhomogeneity in both permittivity and permeability. The highly sparse FEM matrix is efficiently solved using banded matrix techniques. Results from the method agree well with previous results and with measurements.

References

[1] W. C. Chew, J. M. Jin, and E. Michielssen, "Complex coordinate stretching as a generalized absorbing boundary condition," in *13th Annual Review of Progress in Applied Computational Electromagnetics*, Monterey, CA, Mar 1997, vol. II, pp. 909–914.

[2] F. L. Teixēira and W. C. Chew, "Systematic derivation of anisotropic PML absorbing media in cylindrical and spherical coordinates," *Microwave Opt. Tech. Lett.*, vol. 7, no. 11, pp. 371–373, Nov 1997.

[3] J. Maloney, M. Kesler, and G. Smith, "Generalization of PML to cylindrical geometries," in *13th Annual Review of Progress in Applied Computational Electromagnetics*, Monterey, CA, Mar 1997, vol. II, pp. 900–908.

[4] R. Mittra and R. Gordon, "Radar scattering from bodies of revolution using an efficient partial differential equation algorithm," *IEEE Trans. Antennas Propagat.*, vol. 37, no. 5, pp. 538–545, May 1989.

[5] M. A. Morgan, S. K. Chang, and K. K. Mei, "Coupled azimuth potentials for electromagnetic field problems in inhomogeneous axially symmetric media," *IEEE Trans. Antennas Propagat.*, vol. 25, no. 3, pp. 413–417, May 1977.

[6] M. A. Morgan and K. K. Mei, "Finite-element computation of scattering by inhomogeneous penetrable bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 27, no. 2, pp. 202–214, Mar 1979.

[7] A. Khebir, J. D'Angelo, and J. Joseph, "A new finite element formulation for RF scattering by complex bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 41, no. 5, pp. 534–541, May 1993.

[8] J. M. Jin, *The Finite Element Method in Electromagnetics*, John Wiley & Sons, Inc., New York, 1993.

[9] M. F. Wong, M. Prak, and V. F. Hanna, "Axisymmetric edge-based finite element formulation for bodies of revolution: Application to dielectric resonators," in *IEEE MTT-S Digest*, May 1995, pp. 285–288.

[10] M. G. Andreasen, "Scattering from bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 13, no. 2, pp. 303–310, Mar 1965.

[11] A. George and J. W. Liu, *Computer Solution of Large Sparse Positive Definite Systems*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1981.

[12] A. C. Woo, 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, no. 1, pp. 84–89, Feb 1993.

Homogenized Finite Element Model of a Beam Waveguide Resonator Antenna with Over One Hundred Coupling Holes

John R. Brauer
Ansoft Corporation
9000 N. Deerbrook Trail, Milwaukee, WI 53223 USA
brauer@ansoft.com, fax 414/357−8487, phone 414/357−8723

Abstract— **A cavity−backed microwave antenna with over one hundred small cylindrical coupling holes is analyzed using a homogenized finite element model. The homogenized model has a single dielectric−filled slot in place of the many coupling holes, thereby greatly reducing the time needed to compute the radiation pattern. The homogenized model of a 9 GHz rectangular beam waveguide resonator antenna is derived to produce the same transmission coefficient as do the small coupling holes. The computed radiation pattern is shown to agree well with the measured pattern.**

INTRODUCTION

Metal boxes with large numbers of holes are commonly used to enclose electrical and electronics devices. The holes are often essential to allow thermal ventilation, and they also reduce overall weight. However, the holes allow electromagnetic radiation to escape, which may produce undesirable electromagnetic interference (EMI).

Holes in metallic walls are also used in some antennas, where they produce desired electromagnetic radiation. Holes are especially useful in cavity−backed antennas, because the hole size can be readily altered to adjust the coupling of electromagnetic energy from inside the cavity to outside radiation.

Modeling enclosures and antennas with many holes can be very difficult, because each hole usually must be modeled in detail. For example, if the finite element method is used, at least 20 or 30 finite elements are typically required to model each hole, so complete finite element models will often require tens of thousands of 3D finite elements. The computer time required for such large models may be prohibitive, especially if many design iterations are to be analyzed.

A technique called *homogenization* can be helpful in reducing model sizes and attendant computer times. The term has been used in computational modeling to signify methods of replacing detailed inhomogeneities with a single homogeneous equivalent [1]−[4].

This paper applies homogenization to finite element analysis of a beam waveguide resonator antenna with a multitude of coupling holes [5]. This cavity−backed antenna is analyzed here for the first time using the finite element method. The computed radiation pattern is compared with measurements for several different sizes of coupling holes.

RECTANGULAR BEAM WAVEGUIDE RESONATOR AND ITS RESONANT FREQUENCIES

Beam waveguide is a type of quasi−optical transmission line that dates from 1961 [6]. Nowadays it is used for feeding large reflector antennas [7], [8] and sometimes for other purposes [9], [10].

A resonator can be created by placing conductive metal walls in a beam waveguide at phase fronts spaced by 180 electrical degrees [5],[11]. One of the phase fronts can be chosen to be a plane, but then the other phase front must be nonplanar. Fig. 1a shows the metal wall shapes for a rectangular beam waveguide resonator designed and built [5] to have its fundamental resonant frequency near 9.0 GHz. Fig. 1a shows only the right half; the other half is symmetrical about the y axis. The 2D quadrilateral finite elements used in Fig. 1b to model the resonator are of thickness equal to its actual height of 19 mm in the z direction. The resonant fields are invariant with z. The actual experimental resonator contains four metal walls: the two phase front walls at y=0 and variable y shown in Fig. 1a, and two planar walls at z = 0 and z = 19 mm. The two cavity resonator ends, at x = +381 (in Fig. 1a) and −381 mm, can be left open because the electromagnetic fields near them are negligibly small [5].

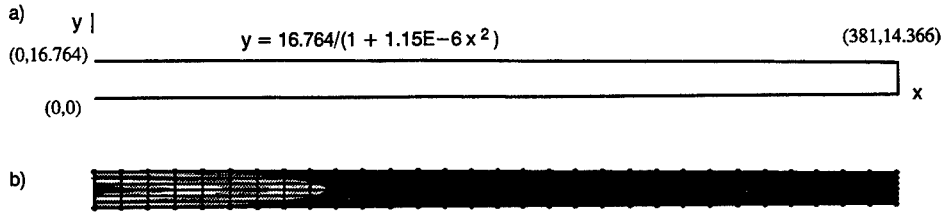


Fig. 1. Right half of beam waveguide resonator.
a) Dimensions in mm, b) Finite element model with computed E for fundamental eigenmode.

The resonant frequencies of the finite element model of Fig. 1b can be computed in two ways. Previously [12] they were computed by introducing an AC exciting current at certain elements or nodes, sweeping the excitation frequency, and noting the frequencies at which large (and equal) magnetic field and electric field energies occur. Instead, in this paper a real eigenvalue solution is used, which does not require any excitation. The eigenvalues (resonant frequencies) and eigenfunction (modes) are computed here using Ansoft's MicroWaveLab™ finite element software [13]. It extracts real eigenvalues using the Lanczos method with Sturm sequencing, which theoretically will always find all modes occurring over a user−selected frequency range [13].

Fig. 1b shows the computed electric field distribution for the fundamental mode of the model of Fig. 1a. The computed resonant frequency is 9.0838 GHz, which agrees closely with the 9.09 GHz computed previously [13] and reasonably well with the measured 8.975 GHz. The mode shape in Fig. 1b appears to follow the theoretical Gaussian distribution. Other modes have shapes that are Hermite polynomials times the Gaussian distribution. Because of the small 19 mm height in the z direction, the fundamental and first dozen or more higher order modes are all invariant in the z direction. The fundamental and several other modes were measured previously [5] using a coaxial probe excitation with the probe pointing in the $-z$ direction and located at $x = 0$ and $y = 8.4$ mm.

A study was next made of the effect of model size in the x direction on the computed resonant frequencies, because the 3D antenna model to be developed should be as small as possible. It was found that reducing the x dimension to 190.5 mm by removing one half of the finite elements from Fig. 1a causes only small changes in the resonant frequencies. The fundamental frequency shifted less than one part per million, and even the frequency of the fifth mode only changed by less than 0.2%. Thus all succeeding models in this paper will have finite elements extending only to $x = 190.5$ mm.

RECTANGULAR BEAM WAVEGUIDE RESONATOR ANTENNA

The cavity resonator of Fig. 1 can be converted into a cavity−backed antenna if proper coupling to free space can be introduced. The chosen coupling method must be manufacturable and must produce the proper coupling coefficient. The coefficient must be small enough that the resonant modes are not perturbed significantly, yet large enough that the antenna has a low reflection coefficient and high gain. Previously [5],[14], a single line of coupling holes along the x axis of Fig. 1 was the design chosen.

Three different sets of coupling holes have been previously fabricated [5],[14]. Their diameters and center−to−center spacings are listed in Table 1. Note that over the entire x axis of Fig. 1 of length 762 mm, all three cases of Table 1 contain well over one hundred holes.

Table 1. Coupling hole configurations, with dimensions in millimeters

Case	Diameter	Spacing	Number of holes
A	2.54	3.175	240
B	3.81	4.445	171
C	4.44	5.08	150

Radiation patterns have been previously measured [5],[14] for all three cases of coupling holes. They have been measured both for the fundamental mode of Fig. 1b, as well as for the mode just above it. Both modes have very low sidelobes and can be used together for monopulse radar and tracking [14]. This paper, however, will compute the radiation pattern only for the fundamental mode.

TRANSMISSION COEFFICIENT OF TWO COUPLING HOLES

Fig. 2 shows views of two holes of Case C in Table 1. A one–half solid model has been made, including the actual coupling wall thickness of 0.64 mm. On both sides of the holes are *conventional* rectangular waveguides of length 10 mm and height 9.5 mm, which is half the actual height of 19 mm, and thus symmetry requires the use of a magnetic wall boundary condition. The width of the waveguide cross section in Fig. 2 is approximately twice the hole spacing, or 10.2 mm. Because 20.4 mm achieves TE10 mode propagation in the region of 9 GHz, symmetry is again imposed through another magnetic wall boundary condition. In other words, only one quarter of the waveguide cross section is modeled.

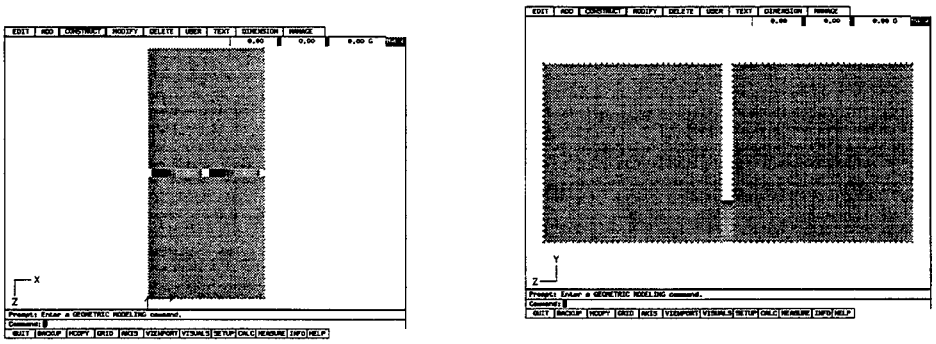


Fig. 2. Views of geometry of waveguides with two holes filled with air, of dimensions of Case C of Table 1.

Fig. 3 shows the finite element model developed for Fig. 2. It consists of 24,545 tetrahedral H1–curl edge finite

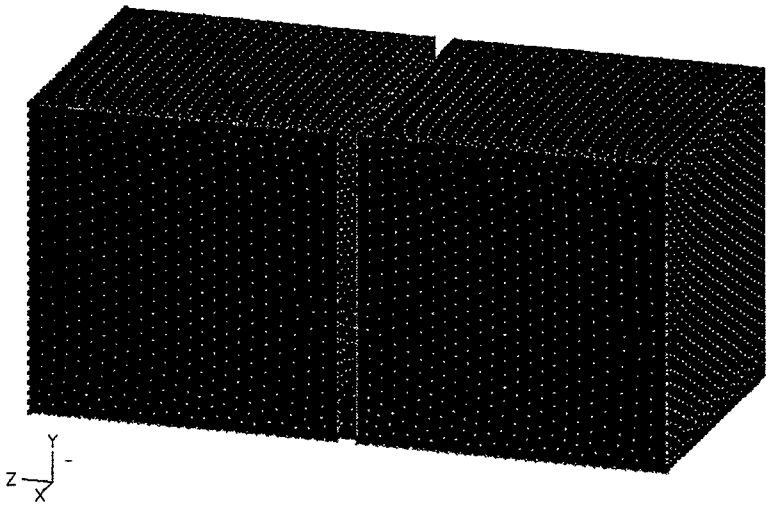


Fig. 3. Finite element model of holes of Fig. 2.

elements and 160,800 edge degrees of freedom. While the model in Figs. 2 and 3 has been made with Ansoft's MicroWaveLab™ software [13], Ansoft's High Frequency Structure Simulator [15] could have been used instead. For port 1 excited with a TE10 mode at 9 GHz, the computed transmission coefficient S21 is −34.28 dB. The computation on an HP 735/125 workstation required 60 Mb memory, 2.8 Gb disk, and 20,270 CPU seconds.

TRANSMISSION COEFFICIENT OF HOMOGENEOUS DIELECTRIC SLOT

To develop a homogenized model, the holes of Figs. 2 and 3 must be replaced by a much simpler aperture. Here the simpler aperture is a slot of uniform width that replaces the holes along the entire x axis of the wall. Such a slot must produce the same −34 dB transmission coefficient as do the holes. The slot width should be fairly large so that the finite elements needed to model it are of reasonably large size and thus produce a small number of finite elements in the required two models (the detailed model and the final overall homogenized model). Therefore it was decided to fill the entire slot with a dielectric material of constant lossless permittivity. By adjusting the permittivity, the transmission coefficient can be adjusted to match the computed S21 of holes of various sizes such as in Table 1.

The slot width chosen is 4 mm, which for the symmetric model of Figs. 2 and 3 is 2 mm. Fig. 4 shows the finite element model developed for the slot aperture in the same waveguide as used in Figs. 2 and 3. It consists of 200 H1−curl hexahedrons, which have been biased to be smaller near the slot aperture.

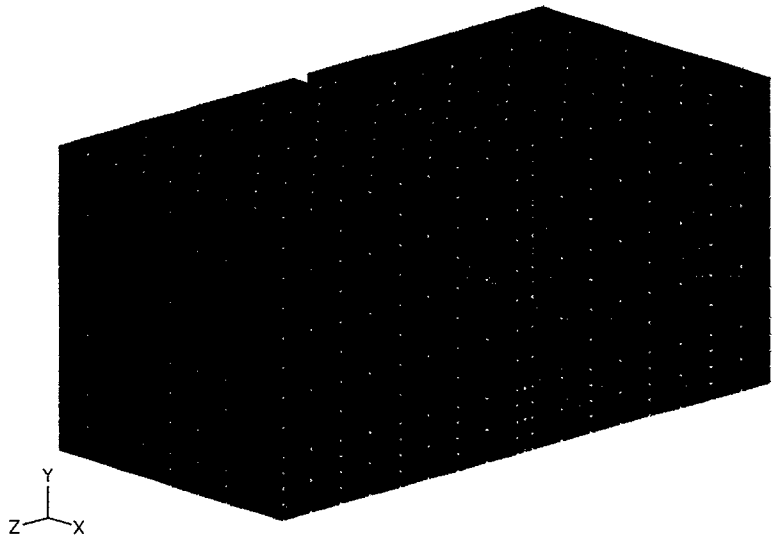


Fig. 4. Finite element model of waveguides with a dielectric−filled slot that replaces the holes of Figs. 2 and 3.

The relative permittivity of the slot in Fig. 4, which has width 2 mm and thickness 0.64 mm (the same as the metal wall) was varied and the resulting transmission coefficient was computed using Ansoft's MicroWaveLab software. Table 2 lists the magnitude of S21 computed for a range of relative permittivities ε_r. Note that $\varepsilon_r = 80$ obtains the desired −34 dB needed to match the S21 computed for the holes of Case C. The computer time is 93 seconds per case.

Table 2. Computed Transmission Coefficient vs. Relative Permittivity of Slot in Fig. 4

Permittivity ε_r	S21 (dB)
10	−20.0
40	−25.3
80	−34.0

HOMOGENIZED FINITE ELEMENT MODEL OF ANTENNA

Using the 2 mm slot of Fig. 4, a homogenized 3D model can now be developed for the actual entire rectangular beam waveguide resonator and antenna. Fig. 5 shows a detail of the solid model developed, which represents one quarter of the overall resonator and antenna. The 2 mm by 0.64 mm slot aperture can be seen in the planar wall of the cavity.

Extending 3 mm above the cavity is one–half of a square model of the coaxial cable feed structure. It approximates the actual cylindrical coax feed [14], which has an outer diameter of approximately 3.2 mm and an inner diameter of approximately 0.7 mm. Here the square coax has an outer width of 3.33 mm and an inner width of 1.11 mm, and is assumed filled with air. The top end of the square coax in Fig. 5 is assumed to be the excited port and only port in the quarter model.

The quarter model of Fig. 5 cannot be used to accurately compute the reflection coefficient S11 seen by the coaxial feed. The main reason is that the actual experimental resonator antenna has only one coaxial feed on its top, and none through its bottom. Thus a half model would be required to compute S11 accurately. A smaller source of error in S11 computations is making the coax square. The model of Fig. 5 is suitable, however, for computing the radiation pattern.

Also visible in Fig. 5 is a small amount of air outside the slot. The outer air is 3 mm thick and allows radiation to escape through the slot.

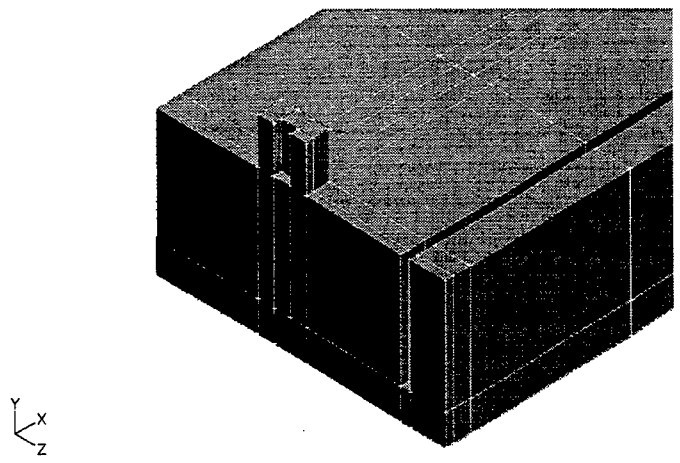
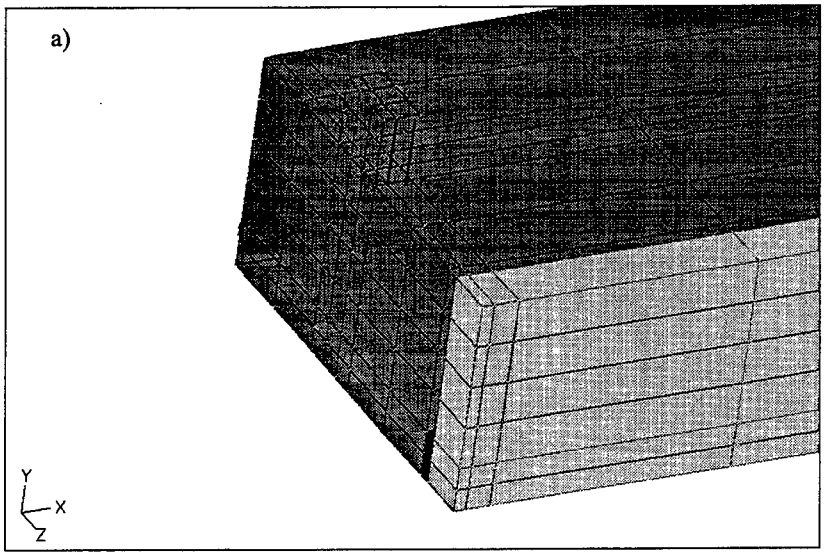


Fig. 5. Detail of solid model of one quarter of the rectangular beam waveguide resonator antenna, showing the x=0 plane through the coax feed and the homogenized dielectric–filled slot.

Fig. 6 shows the finite element model developed for Fig. 5. Fig. 6a is a detailed view similar to Fig. 5, and Fig. 6b is an overall view of the entire model. Because of the assumed square coax, the model of Fig. 6 can consist entirely of hexahedrons. There are 860 H1–curl hexahedrons and 4,224 edge degrees of freedom.

In addition to these finite elements visible in Fig. 6, three layers of PMA (perfectly matched absorber) finite elements are internally generated by Ansoft's MicroWaveLab software to absorb the radiation emitted by the antenna [16]. Based on the electromagnetic fields at the interface of the visible and PMA elements, the radiation pattern is computed.



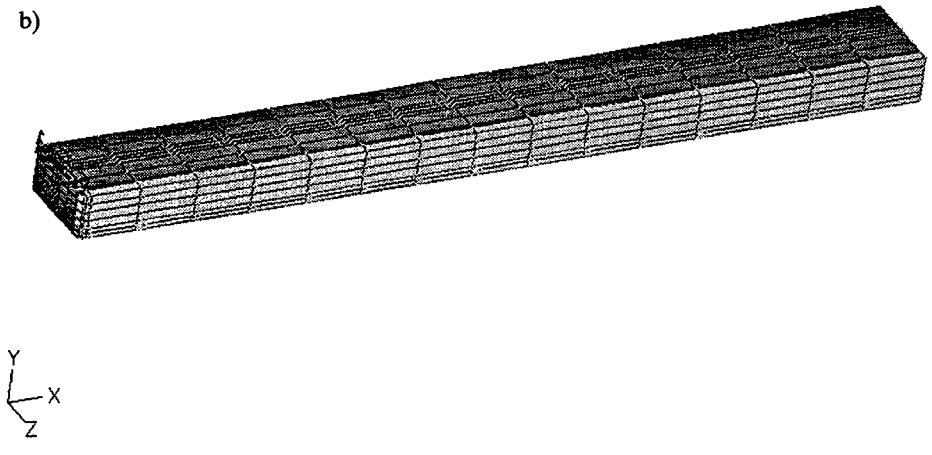


Fig. 6. Finite element model of the homogenized resonator and antenna. The cavity, slot, and air in front of the slot are shown in three different shadings. All finite elements are made of air, except for the slot elements which are given a high permittivity.
a) Detail in the region of Fig. 5. b) View of entire model.

RADIATION PATTERNS OF ANTENNA WITH VARIOUS COUPLING HOLES

Fig. 7 compares computed and measured [14] radiation patterns of the rectangular beam waveguide resonator and antenna. Measurements are shown for all three different cases of coupling holes specified in Table 1. The measured patterns are expected to be accurate only down to about 25 or 30 dB below the peak, due to the incomplete placement of absorptive material in the measurement chamber [14]. All computations are at the fundamental resonant frequency of Fig. 1, which is 9.0838 GHz.

Note that the computed pattern with relative slot permittivity equal to 80 agrees well with the measured pattern for Case C, thereby confirming the validity of the homogenized finite element model that used the S21 computed for Figs. 3 and 4. Also, the computed pattern for slot relative permittivity lowered to 40 shows further broadening of the beam, which is expected based on the measurements.

The computed pattern for relative permittivity equal to 120 agrees well with measured patterns for the smaller coupling holes of Cases A and B. It also agrees well with the theoretical pattern for very small coupling, which has been derived [5], [14] using the unperturbed field of the fundamental cavity mode.

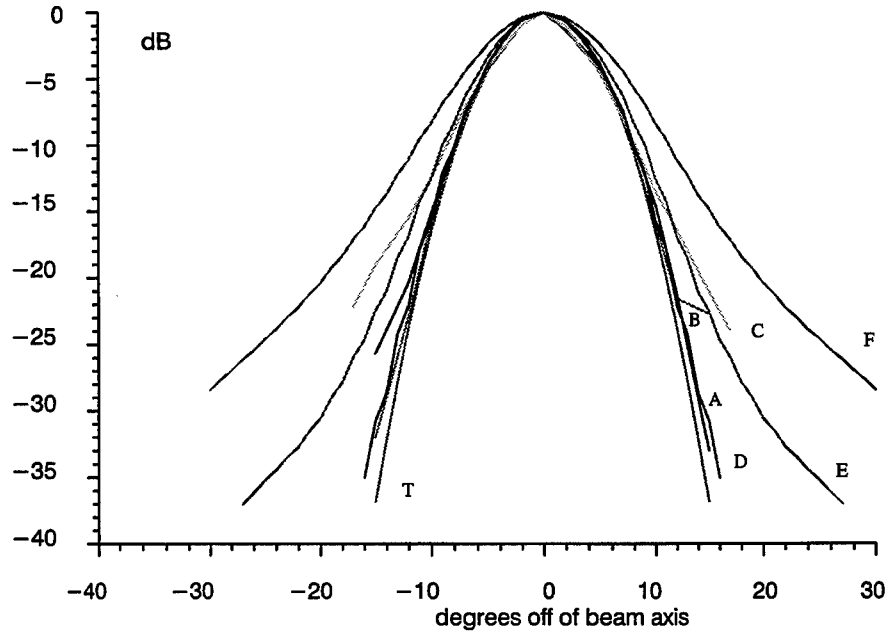


Fig. 7. Radiation patterns of rectangular beam waveguide resonator antenna, denoted by letters: Measured cases A, B, and C.
D−computed for ε_r=120, E−computed for ε_r=80, F−computed for ε_r=40.
T−theoretical pattern for pure unperturbed cavity mode.

CONCLUSION

More than one hundred small coupling holes in an experimental cavity–backed antenna have been replaced by a homogeneous dielectric–filled slot in a finite element model. The slot permittivity has been calculated to produce the same coupling as do the many holes. The radiation pattern has been efficiently computed, and it agrees well with measurements. The analysis has for the first time quantitatively predicted the perturbation of the radiation pattern due to the finite size of the coupling holes.

REFERENCES

[1] Keith W. Whites and Raj Mittra, "An equivalent boundary–condition model for lossy planar periodic structures at low frequencies," *IEEE Trans. Antennas and Propagation*, v 44, Dec. 1996, pp. 1617–1629.

[2] N. Labbe, Y. Marechal, G. Meunier, and H. B. Harara, "2D nonlinear finite element modelling of electromagnetic retarders using time–stepping algorithms, and the Petrov–Galerkin method with homogenization techniques," *IEEE Trans. Magnetics*, v 32, May 1996, pp. 772–775.

[3] E. F. Kuester and C. L. Holloway, "A low–frequency model for wedge or pyramid absorber arrays– I: theory," *IEEE Trans. Electromagnetic Compatibility*, v 36, Nov. 1994, pp. 300–306.

[4] C. L. Holloway, R. R. DeLyser, R. F. German, P. MacKenna, and M. Kanda, "Comparison of electromagnetic absorber used in anechoic and semi–anechoic chambers for emissions and immunity testing of digital devices," *IEEE Trans. Electromagnetic Compatibility*, v 39, Feb. 1997, pp. 33–47.

[5] John R. Brauer, "A rectangular beam waveguide resonator and antenna," *IEEE Trans. Antennas and Propagation*, v 20, Sept. 1972, pp. 595–601.

[6] Georg Goubau and Felix Schwering, "On the guided propagation of electromagnetic wave beams," *IRE Trans. Antennas and Propagation*, v 9, May 1961, pp. 248–256.

[7] T. Hayashi, T. Nishimura, T. Takano, S.–I. Betsudan, and S. Koshizaka, "Japanese deep–space station with 64–m–diameter antenna fed through beam waveguides and its mission applications," *Proc. IEEE*, v 82, May 1994, pp. 646–657.

[8] William A. Imbriale, "Design and applications of beam waveguide systems," *IEEE Aerospace Conf. Proc.*, Feb. 1–8 1997, pp. 121–134.

[9] O. Demokan, "Efficient microwave generation in a beam–waveguide system with an annular plasma sheet," *IEEE Trans. Plasma Science*, v 23, June 1995, pp. 405–408.

[10] H. Hwang, G. P. Monahan, M. B. Steer, J. W. Mink, J. Harvey, A. Paollea, and F. K. Schwering, "A dielectric slab waveguide with four planar power amplifiers," *IEEE Microwave Symposium Digest*, May 16–20, 1995, pp. 921–924.

[11] J. W. Mink and E. H. Scheibe, "A dual mode beam waveguide resonator and frequency stabilizer at millimeter–wave frequencies," *IEEE Trans. Microwave Theory and Techniques*, v 14, May 1966, pp. 222–228.

[12] John R. Brauer, Jeffrey J. Ruehl, and Anthony B. Bruno, "Finite element analysis of microwave cavities," *3rd Annual Review of Progress in Applied Computational Electromagnetics*, Monterey, CA, March 24–26, 1987.

[13] John R. Brauer and Gary C. Lizalek, "Microwave filter analysis using a new 3–D finite element modal frequency method," *IEEE Trans. Microwave Theory and Techniques*, v 45, May 1997, pp. 810–818.

[14] John R. Brauer, "A rectangular beam waveguide resonator and antenna," Ph.D. dissert., University of Wisconsin–Madison, June 1969, no. 69–22353 from University Microfilms, P. O. Box 1346, Ann Arbor, MI 48106.

[15] Din Sun, John Manges, Xingchao Yuan, and Zoltan Cendes, "Spurious modes in finite element methods," *IEEE Antennas and Propagation Magazine*, v. 37, Oct. 1995, pp. 12–24.

[16] J. F. DeFord and G. C. Lizalek, "Use of perfectly–matched absorber boundaries in finite element analysis of patch antennas," *12th Annual Review of Progress in Applied Computational Electromagnetics*, Monterey, CA, March 18–27, 1996, pp. 1190–1195.

A Surface Admittance Formulation for the Transient Modeling of Skin Effect and Eddy Current Problems[1]

Karim N. Wassef and Andrew F. Peterson
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332

1 Introduction

The magnetic vector potential has long been used in finite elements to analyze the magnetic field distribution in regions containing permeable and conductive materials. This method has been applied to both skin-effect and eddy-current problems successfully. However, in cases where the material properties such as conductivity and permeability are several orders of magnitude higher than the surrounding medium, meshing both regions becomes computationally expensive and highly inefficient. In this situation, a surface impedance boundary condition (SIBC) may be applied to model the expected field response at the material interface.

The steady-state application of an SIBC is straightforward. The transient analysis using surface impedances has also been achieved via the fast Fourier transform [1]. Even though an SIBC is defined as a frequency domain parameter, it is desirable to develop a method that is capable of analyzing the transient response at interfaces without transforming into the frequency domain. The proposed formulation incorporates discrete temporal integration using Prony's method in order to create an efficient, recursive procedure. This is similar to the method used in FDTD [2] to implement an SIBC. Since the boundary condition, when applied to the magnetic vector potential formulation, appears in its reciprocal form, it is better viewed as a surface admittace boundary condition (SABC).

2 Theory

2.1 The Magnetic Vector Potential Formulation

Consider a two-dimensional cross section containing conducting, dielectric, and magnetic materials transverse to an applied current and electric field. The magnetic vector potential $\mathbf{A}$ is related to the magnetic field through

$$\mathbf{H} = \frac{1}{\mu}\nabla \times \mathbf{A} = \nu \nabla \times \mathbf{A} \tag{1}$$

Maxwell's curl equations (with low frequency approximations) will be employed:

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \cong \mathbf{J} \tag{2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{3}$$

resulting in the electric field dependence on the magnetic vector potential,

$$\mathbf{E} \cong -\frac{\partial \mathbf{A}}{\partial t} \tag{4}$$

Combining (1) and (2) produces

$$\nabla \times (\nu \nabla \times \mathbf{A}) = \mathbf{J} \tag{5}$$

[1]This research has been supported by the US Office of Naval Research under ONR Grant No. N00014-96-1-0926 and by the National Science Foundation under Grant ECS-9257927

or, since **A** and **J** only have $\hat{\mathbf{z}}$-components,

$$-\nabla \cdot (\nu \nabla A) = J \tag{6}$$

in a region with sources and non-zero conductivity, and

$$-\nabla \cdot (\nu \nabla A) = 0 \tag{7}$$

in regions with zero conductivity.

A formalism is used that breaks the total magnetic vector potential into a forcing term and a response term. The forcing term is due to the imposed currents *in the absence* of all materials. The response term is determined using

$$A^{total} = A^{forced} + A^{response} \tag{8}$$

An SABC is used to exclude all conducting regions from the computational domain; thus (7) is the basic equation to be solved.

2.2 The Surface Admittance Boundary Condition (SABC)

The first order (Leontovich) surface impedance boundary condition relating the tangential field components in the frequency domain is

$$E_z^{response}(\omega) = Z_s(\omega) H_{x,y}^{response}(\omega) \tag{9}$$

where

$$Z_s(\omega) = (1+j)\sqrt{\frac{\omega\mu}{2\sigma}} = \sqrt{\frac{j\omega\mu}{\sigma}} \tag{10}$$

The surface admittance is

$$Y_s(\omega) = \frac{1}{\sqrt{j\omega}}\sqrt{\frac{\sigma}{\mu}} \tag{11}$$

If recast in terms of the primary unknown A, the Surface Admittance Boundary Condition (SABC) becomes :

$$\nu\left(-\frac{\partial A^r(\omega)}{\partial n}\right) = Y_s(\omega)\left(-j\omega A^r(\omega)\right) \tag{12}$$

or, in the time domain,

$$\nu\left(\frac{\partial A^r(t)}{\partial n}\right) = Y_s(t) * \left(\frac{\partial A^r(t)}{\partial t}\right) \tag{13}$$

This condition will be imposed at the surface of good conductors, in order to exclude those regions from the computational domain.

3 The Finite Element Formulation

The finite element formulation is used to find the forced magnetic vector potential by solving

$$-\nabla \cdot \nu_o \nabla A^{forced} = J^{forced} \tag{14}$$

for a given current distribution. Eqn. (14) is converted to a weak form

$$\int_\Delta (\nabla\beta \cdot \nu \nabla A^f) dV - \int_\Omega \beta\nu \frac{\partial A^f}{\partial n} dS = \int_\Delta \beta J^f dV \tag{15}$$

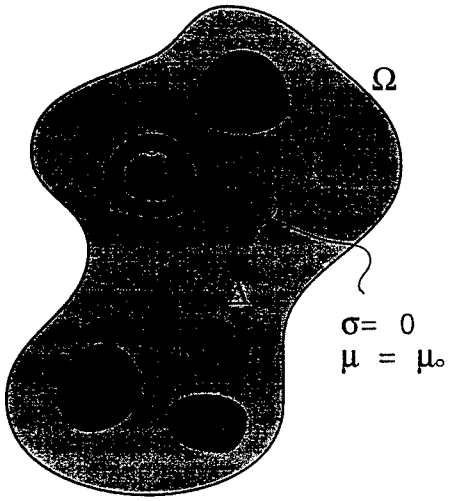


Figure 1: Computational domain used in determining the forced vector potential.

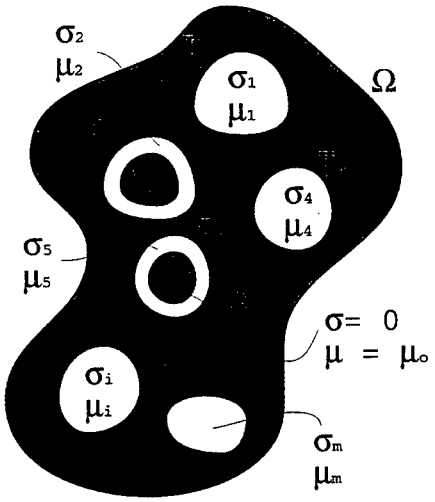


Figure 2: Computational domain used in determining the total vector potential.

where Δ is the computational domain with all materials absent, bounded externally by Ω. An absorbing boundary condition

$$\frac{\partial A}{\partial n} = \frac{1}{R_\infty \ln(R_\infty)} A \tag{16}$$

is imposed on Ω, where R_∞ is the finite distance to Ω from an origin within Δ. Finite Elements are subsequently used to find A^{total} by solving

$$-\nabla \cdot \nu \nabla A^{total} = 0 \tag{17}$$

in the presence of all materials. Eqn. (17) is converted to a weak form

$$\int_\Delta (\nabla\beta \cdot \nu \nabla A^t) dV - \int_\Omega \beta\nu \frac{\partial A^t}{\partial n} dS - \int_\Gamma \beta\nu \frac{\partial A^t}{\partial n} dS = 0 \tag{18}$$

where Ω denotes the external boundary and Γ denotes collectively the surface of conducing regions.
After separating the total magnetic vector potential into its forced and response components,

$$\int_\Delta (\nabla\beta \cdot \nu \nabla A^t) dV - \int_\Omega \beta\nu \frac{\partial A^t}{\partial n} dS - \int_\Gamma \beta\nu \frac{\partial A^f}{\partial n} dS - \int_\Gamma \beta \left(Y_s * \frac{\partial A^r}{\partial t} \right) dS = 0 \tag{19}$$

where the SABC has been employed. Eqn. (22) can be rewritten as

$$\int_\Delta (\nabla\beta \cdot \nu \nabla A^t) dV - \int_\Omega \beta\nu \frac{\partial A^t}{\partial n} dS - \int_\Gamma \beta \left(Y_s * \frac{\partial A^t}{\partial t} \right) dS = - \int_\Gamma \beta \left(Y_s * \frac{\partial A^f}{\partial t} \right) dS + \int_\Gamma \beta\nu \frac{\partial A^f}{\partial n} dS \tag{20}$$

4 Evaluation of the Finite Element and Boundary Element Terms

Using standard first-order linear finite elements, the boundary integral can be implemented along an edge as follows:

$$\begin{aligned}
\int_\Gamma \beta \left(Y_s * \frac{\partial A}{\partial t} \right) dS &= \int_0^1 [\beta]^T [\beta] |\mathcal{J}_\ell| du \left(Y_s * \frac{\partial A}{\partial t} \right) \\
&= \int_0^1 \begin{bmatrix} 1-u \\ u \end{bmatrix} \begin{bmatrix} 1-u & u \end{bmatrix} \ell \, du \left(Y_s * \frac{\partial}{\partial t} \{A(t)\} \right) \\
&= \ell \int_0^1 \begin{bmatrix} 1+u^2-2u & u-u^2 \\ u-u^2 & u^2 \end{bmatrix} du \int_0^t Y_s(\tau) \frac{\partial}{\partial(t-\tau)} \{A(t-\tau)\} \, d\tau \\
&= \ell \begin{bmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} \end{bmatrix} \int_0^t \sqrt{\frac{1}{\tau}} \sqrt{\frac{\sigma}{\mu\pi}} \left(-\frac{\partial}{\partial\tau} A(t-\tau) \right) d\tau \\
&= \frac{\ell}{6} \sqrt{\frac{\sigma}{\mu\pi}} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \xi(t)
\end{aligned} \tag{21}$$

where

$$Y_s(t) = \mathcal{L}^{-1}\{Y_s(\omega)\} = \mathcal{L}^{-1} \left\{ \frac{1}{\sqrt{j\omega}} \sqrt{\frac{\sigma}{\mu}} \right\} = \frac{1}{\sqrt{\pi t}} \sqrt{\frac{\sigma}{\mu}} \tag{22}$$

$$\xi(t) = \int_0^t \sqrt{\frac{1}{\tau}} \left(-\frac{\partial}{\partial\tau} A(t-\tau) \right) d\tau \tag{23}$$

Since a discrete time integration is needed, let $\tau = \alpha\Delta t$ and observe that

$$\int_{m\Delta t}^{(m+1)\Delta t} \frac{d\tau}{\sqrt{\tau}} = \frac{1}{\sqrt{\Delta t}} \int_m^{m+1} \frac{d\alpha}{\alpha} \tag{24}$$

Expanding the unit integration using the Prony Series:

$$\int_m^{m+1} \frac{d\alpha}{\alpha} = \sum_{i=1}^{N} a_i e^{m\alpha_i} = Z_o(m) \tag{25}$$

where a_i and α_i are predetermined Prony expansion coefficients for a given number of series terms N. To evaluate $\xi(t)$, let A^0 be the first (initial) value of A, A^{n-1} be the value of A at the previous time step, and A^n be the new (undetermined) value of A. Then,

$$\begin{aligned}
\xi(t) &= \frac{1}{\sqrt{\Delta t}} \sum_{m=0}^{n-1} Z_o(m) \left(\frac{A^{n-m} - A^{n-m-1}}{\Delta t} \right) \\
&= \frac{1}{\sqrt{\Delta t}} \left[Z_o(0) \left(\frac{A^n - A^{n-1}}{\Delta t} \right) + \sum_{m=1}^{n-1} Z_o(0) \left(\frac{A^{n-m} - A^{n-m-1}}{\Delta t} \right) \right] \\
&= \frac{Z_o(0)}{\sqrt{\Delta t}} \left(A^n - A^{n-1} \right) + \frac{1}{\sqrt{\Delta t}} \sum_{i=1}^{10} \sum_{m=1}^{n-1} a_i e^{m\alpha_i} \left(A^{n-m} - A^{n-m-1} \right)
\end{aligned} \tag{26}$$

Grouping past history terms into a new term ψ_i^n:

$$\psi_i^n = \sum_{m=1}^{n-1} a_i e^{m\alpha_i} \left(A^{n-m} - A^{n-m-1} \right) \tag{27}$$

$$\begin{aligned}
\psi_i^0 &= \psi_i^1 = 0 \\
\psi_i^2 &= a_i e^{\alpha_i} \left(A^1 - A^0 \right) \\
\psi_i^3 &= a_i e^{\alpha_i} \left(A^2 - A^1 \right) + a_i e^{2\alpha_i} \left(A^1 - A^0 \right) \\
&= a_i e^{\alpha_i} \left(A^2 - A^1 \right) + e^{\alpha_i} \psi_i^2 \\
&\Rightarrow \\
\psi_i^n &= a_i e^{\alpha_i} \left(A^{n-1} - A^{n-2} \right) + e^{\alpha_i} \psi_i^{n-1}
\end{aligned} \tag{28}$$

where ψ_i^{n-1}, A^{n-1}, and A^{n-2} are all known terms for the previous time step. The boundary integral can be evaluated:

$$\begin{aligned}
\int_\Gamma \beta \left(Y_s * \frac{\partial A}{\partial t} \right) dS &= \frac{\ell}{6} \sqrt{\frac{\sigma}{\mu\pi}} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \left[\frac{Z_o(0)}{\sqrt{\Delta t}} \left\{ \{A\}^n - \{A\}^{n-1} \right\} + \frac{1}{\sqrt{\Delta t}} \left\{ \sum_{i=1}^{10} \psi_i^n \right\} \right] \\
&= \frac{Z_o(0)\ell}{6} \sqrt{\frac{\sigma}{\mu\pi\Delta t}} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \left\{ \begin{array}{c} A_1 \\ A_2 \end{array} \right\}^n - \frac{Z_o(0)\ell}{6} \sqrt{\frac{\sigma}{\mu\pi\Delta t}} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \left\{ \begin{array}{c} A_1 \\ A_2 \end{array} \right\}^{n-1} \\
&+ \frac{\ell}{6} \sqrt{\frac{\sigma}{\mu\pi\Delta t}} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \left\{ \begin{array}{c} \left(\sum_{i=1}^{10} \psi_i \right)_1 \\ \left(\sum_{i=1}^{10} \psi_i \right)_2 \end{array} \right\}^n
\end{aligned} \tag{29}$$

The interpretation of this result is different based on whether the magnetic vector potential A is the forced or total field term. In the case of the total field matrix components, the first term acts on the unknown new values of A (at time n). This term remains on the LHS of the finite element matrix. The remaining terms act on past or previously computed values and therefore act as forcing terms in the RHS vector of the system of equations. In

the case of the forced field components, all terms are pre-computed and act as RHS forcing vector terms. Define

$$\varrho = \frac{\ell}{6}\sqrt{\frac{\sigma}{\mu\pi\Delta t}} \tag{30}$$

$$\Psi = \sum_{i=1}^{10}\psi_i \tag{31}$$

$$\Delta A^f = A^{f(n)} - A^{f(n-1)} \tag{32}$$

The total field matrix term is:

$$[B]\{A\}^t = Z_o(0)\varrho\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{Bmatrix} A_1^t \\ A_2^t \end{Bmatrix}^n \tag{33}$$

The total field forcing vector term is:

$$\{P\}^t = -Z_o(0)\varrho\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{Bmatrix} A_1^t \\ A_2^t \end{Bmatrix}^{n-1} + \varrho\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{Bmatrix} \Psi_1^t \\ \Psi_2^t \end{Bmatrix}^n \tag{34}$$

$$= \varrho\begin{Bmatrix} -Z_o(0)\left(2A_1^{t(n-1)} + A_2^{t(n-1)}\right) + 2\Psi_1^{t(n)} + \Psi_2^{t(n)} \\ -Z_o(0)\left(A_1^{t(n-1)} + 2A_2^{t(n-1)}\right) + \Psi_1^{t(n)} + 2\Psi_2^{t(n)} \end{Bmatrix} \tag{35}$$

The forced field forcing vector term is:

$$\{R\}^f = Z_o(0)\varrho\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\left\{\begin{Bmatrix} A_1^f \\ A_2^f \end{Bmatrix}^n - \begin{Bmatrix} A_1^f \\ A_2^f \end{Bmatrix}^{n-1}\right\} + \varrho\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{Bmatrix} \Psi_1^f \\ \Psi_2^f \end{Bmatrix}^n \tag{36}$$

$$= \varrho\begin{Bmatrix} Z_o(0)\left(2\Delta A_1^f + \Delta A_2^f\right) + 2\Psi_1^{f(n)} + \Psi_2^{f(n)} \\ Z_o(0)\left(\Delta A_1^f + 2\Delta A_2^f\right) + \Psi_1^{f(n)} + 2\Psi_2^{f(n)} \end{Bmatrix} \tag{37}$$

Therefore,

$$\int_\Gamma \beta\left(Y_s * \frac{\partial A^t}{\partial t}\right) dS = [B]\{A\}^t + \{P\}^t \tag{38}$$

$$\int_\Gamma \beta\left(Y_s * \frac{\partial A^f}{\partial t}\right) dS = \{R\}^f \tag{39}$$

5 Results

To demonstrate the method, consider a copper coaxial cable carrying a transient carrier signal at 50KHz. The conductivity of the copper is 5.9595×10^6 S/m. The geometry of the coaxial cable is such that a=50mm, b=100mm, c=110mm, and the outer boundary Ω is located at d=150mm. The time discretization is 10 divisions per period. The forcing current is 1 Ampere (assuming uniform distribution over the conductor area) and ramped using a cosine waveform over 10 cycles (100 time steps).

Delaunay triangulation was used to mesh the problem space. The results are compared to the analytical solution derived for the case of a perfectly conductive coaxial cable.

These analytical solutions are

$$B_\phi = \frac{\mu_o I}{2\pi\rho} \; ; \; a \leq \rho \leq b \tag{40}$$

$$\frac{\psi_{self}}{l} = \frac{L_{self}}{l} \cdot I = \frac{\mu_o I}{2\pi} \ln\left(\frac{b}{a}\right) \tag{41}$$

and are illustrated in figures (4) and (5).

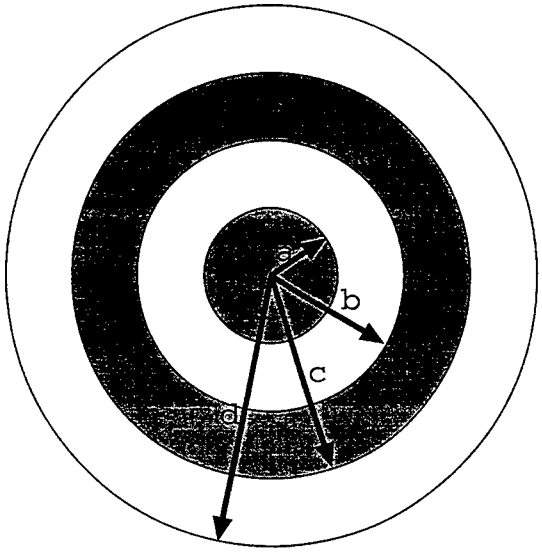


Figure 3: Geometry of the coax.

6 Conclusion

A surface admittance boundary condition is incorporated into a magnetic vector potential formulation for skin effect/eddy current applications. Results indicate the validity of this method. Future work will consider the extension of this approach to incorporate non-linear materials.

References

[1] Kent R. Davey and Larry Turner, "Transient Eddy Current Analysis for Generalized Structures Using Surface Impedances and the Fast Fourier Transform",*IEEE Trans. Magn.*, vol.26, No.3, pp. 1164-1170, May 1990.

[2] Karl S. Kunz and Raymond J. Luebbers, *"The Finite Difference Time Domain Method for Electromagnetics"*,(CRC Press, Inc.,Boca Raton, Florida, 1993).

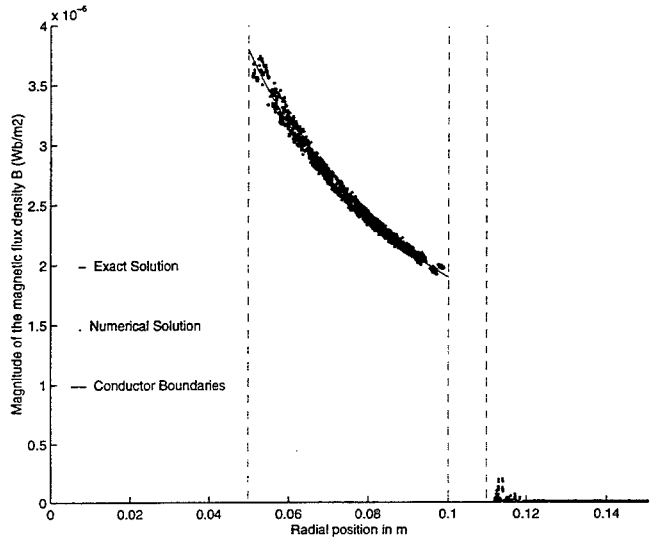


Figure 4: The numerical vs. exact solution of the magnetic flux for the conductive coax.

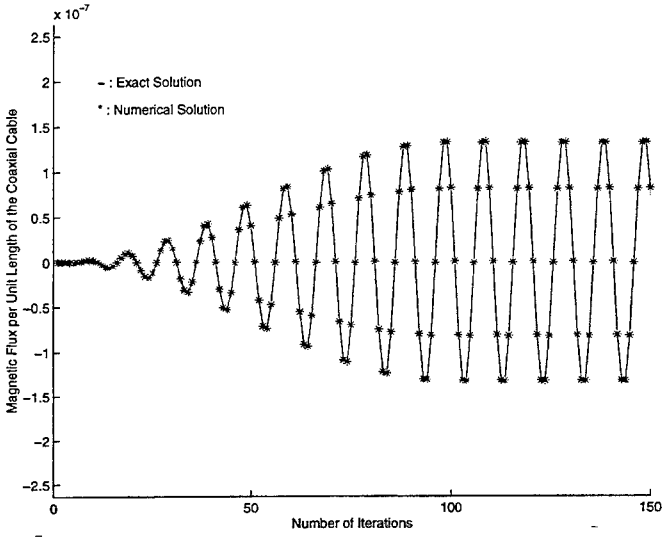


Figure 5: The numerical vs. exact solution of the magnetic flux per unit length for the conductive coax.

Verification of Eddy Current Analysis of Engineering Oriented Loss Model (Problem 21)

N. Takahashi and K. Fujiwara,
Department of Electrical and Electronic Engineering
Okayama University, Okayama 700, Japan

K. Sugiyama and J. Takehara
Technical Research Center, The Chugoku Electric Power Co., Inc.
Higashihiroshima 739, Japan

Abstract

Analysis of a 3-D stray-field loss model (TEAM Workshop Problem 21) is carried out by using the time-periodic finite element method. The flux density and iron loss (eddy current loss and hysteresis loss) are compared with those obtained measurement.

I. INTRODUCTION

A 3-D stray-field loss model (Problem 21) has been proposed as an engineering problem to study eddy current loss distribution in steel plates[1]. The results of linear analysis for Problem 21 have been already reported[2]. It was shown that the eddy current loss of the steel plate is affected by the permeability of steel plate. Therefore, in order to examine the eddy current loss of the steel plate, it is necessary to analyze magnetic fields taking into account the nonlinearity of steel.

In this paper, the nonlinear analysis of Problem 21 is carried out by using the time-periodic finite element method[3-5]. The flux and eddy current distributions and eddy current loss[6] and hysteresis loss are compared with those of the linear analysis and the measured ones.

II. 3-D STRAY-FIELD LOSS MODEL (PROBLEM 21)

Fig.1 shows the analyzed models. Model A consists of two coils of the same dimensions and two steel plates. In the center of one steel plate, there is a rectangular hole. The directions of exciting currents of those coils are different from each other. Model B consists of two coils and a steel plate without hole. Fig.2 shows the outline of flux and eddy current distributions. The ampere-turns of each coil is 3000AT (rms, 50Hz). The conductivity of the steel plate is 5.875×10^6S/m. The B-H curve for the steel plate is shown in Fig.3.

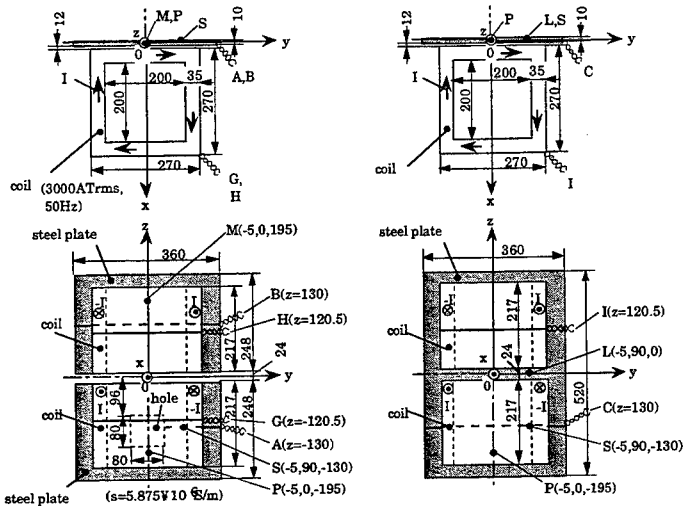
III. METHOD OF ANALYSIS

A. Fundamental Equations

In the 3-D finite element method analysis using edge elements, the residual G_i for the i-th unknown variable is represented as follows[7]:

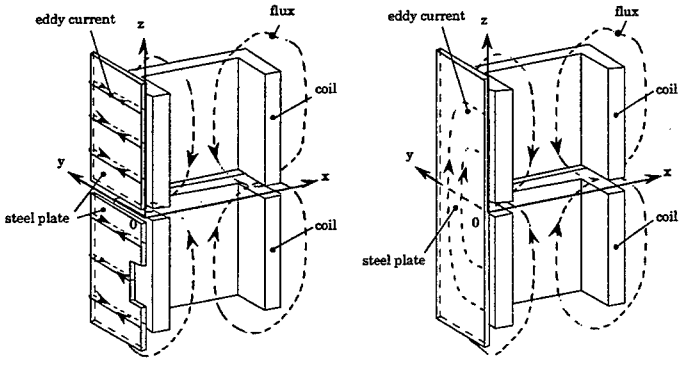
$$G_i = \iiint rot\mathbf{N}_i \bullet (\nu\, rot\mathbf{A})dv - \iiint \mathbf{N}_i \bullet \mathbf{J}_0 dv + \iiint \mathbf{N}_i \bullet \sigma \frac{\partial \mathbf{A}}{\partial t} dv \qquad (1)$$

where $\boldsymbol{A}$ is the magnetic vector potential, $\boldsymbol{J}_0$ is the current density vector in the magnetizing winding, ν and σ are the reluctivity and conductivity, respectively.



(a) model A (b) model B

Fig.1 Analyzed model (problem 21).



(a) model A (b) model B

Fig.2 Schematic diagram of flux and eddy current distributions.

B. Nonlinear Analysis

In the nonlinear analysis using the Newton-Raphson iterative technique, the increments of the unknown variables δA_j are obtained from the following equation[4]:

$$\left[\frac{\partial G_i}{\partial A_j}\right]\{\delta A_j\} = -\{G_i\} \qquad (i,j=1,2,\text{- - -},nu) \tag{2}$$

where *nu* is the number of edges with unknown potentials. The coefficient matrix $[\partial G_i / \partial A_j]$ in (2) is symmetric.

C. Time-Periodic Finite Element Method

When the waveform of a vector potential is symmetric and periodic as shown in Fig.4, the following relationship holds between vector potentials A^t and $A^{t+T/2}$ at the instants t and t+T/2 (T : period):

$$A^t = -A^{t+T/2} \tag{3}$$

In the time-periodic finite element method, the vector potentials A^t, - - - -, $A^{t+T/2-\Delta t}$ (Δt : time interval) are treated as unknown variables, and they are calculated simultaneously taking into account the relationship of (3).

When the potential at each instant is treated as unknown variable, the equations for the nonlinear analysis are as follows[3,4]:

$$\begin{aligned}
&[C^t]\{\delta A_j^{t-\Delta t}\} + [H^t]\{\delta A_j^t\} = -\{G_i^t\} \\
&[C^{t+\Delta t}]\{\delta A_j^t\} + [H^{t+\Delta t}]\{\delta A_j^{t+\Delta t}\} = -\{G_i^{t+\Delta t}\} \\
&\quad\bullet \\
&\quad\bullet \\
&\quad\bullet \\
&[C^{t+T/2-\Delta t}]\{\delta A_j^{t+T/2-2\Delta t}\} + [H^{t+T/2-\Delta t}]\{\delta A_j^{t+T/2-\Delta t}\} = -\{G_i^{t+T/2-\Delta t}\}
\end{aligned} \tag{4}$$

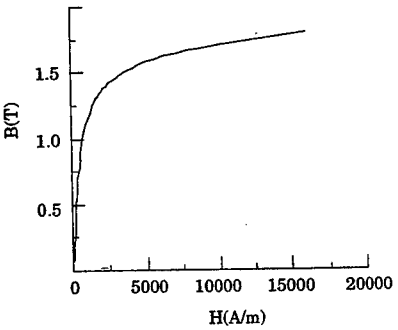


Fig.3 B-H curve.

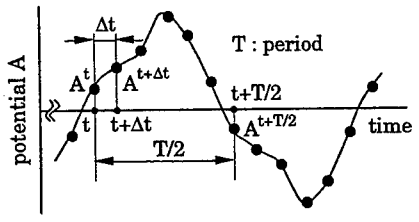


Fig. 4 Periodic waveform.

where [C] and [H] are the same as those of the conventional time-stepping method[8].

By applying the relationship of (3) to $\{\delta A_j{}^{t-\Delta t}\}$ in (4), the following matrix equation is obtained:

$$\begin{bmatrix} [H^t] & \mathbf{0} & \cdots & \mathbf{0} & -[C^t] \\ [C^{t+\Delta t}] & [H^{t+\Delta t}] & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & [C^{t+T/2-\Delta t}] & [H^{t+T/2-\Delta t}] \end{bmatrix} \begin{Bmatrix} \{\delta A_j^t\} \\ \{\delta A_j^{t+\Delta t}\} \\ \vdots \\ \vdots \\ \{\delta A_j^{t+T/2-\Delta t}\} \end{Bmatrix} = \begin{Bmatrix} -\{G_i^t\} \\ -\{G_i^{t+\Delta t}\} \\ \vdots \\ \vdots \\ -\{G_i^{t+T/2-\Delta t}\} \end{Bmatrix} \tag{5}$$

As the coefficient matrix in (5) is very large, considerably long CPU time and large computer memory are required for the solution by conventional method. Therefore, an iterative technique is introduced by dividing (5) into the following equations[3]:

$$[H^{t+m\Delta t}]\{\delta A_j^{t+m\Delta t}\} = -\alpha \cdot \beta_m [C^{t+m\Delta t}]\{\delta A_j^{t+(m-1)\Delta t}\} - \{G_i^{t+m\Delta t}\} \quad (m = 0,1,\cdots,ns-1) \tag{6}$$

where ns is the number of time steps in half a period. β_m is equal to -1 (m=0) and 1 ($m \neq 0$). α is the relaxation factor[3] and is chosen to be zero, because minimum modification is required in the software for the time-stepping method. Although α is chosen to be zero, the relationship of (3) is taken into account in the second term of the right-hand side of (6). The nonlinear iterations are carried out in the outer loop of the time step iterations[4] until A converges as shown in Fig.5. By using this iterative technique, the nonlinear steady-state

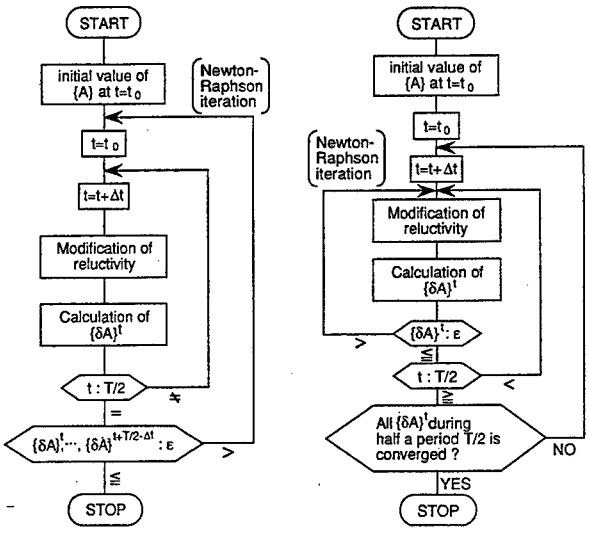


Fig.5 Flowchart.

magnetic fields can be obtained within shorter CPU time than the time-stepping method[8] in which the nonlinear iteration is carried out at each step from the transient state to the steady state.

IV. RESULTS AND DISCUSSION

The magnetic fields of Problem 21 are analyzed by the 3-D finite element method using the 1st order brick edge element. One half of the region is analyzed for model A and one quarter of the region is analyzed for model B. Fig.6 shows the mesh for model A. The steel plate is subdivided into 9 layers. Table I shows the comparison between the discretization data for the linear and nonlinear analyses. The CPU time for the nonlinear analysis is only about 5 times longer than that for the linear analysis.

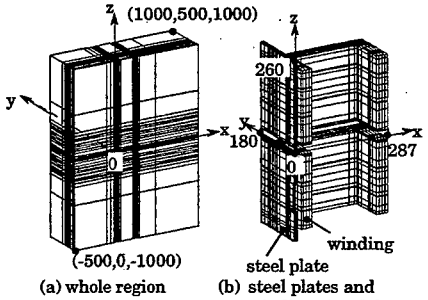
Fig.7 shows flux and eddy current distribution obtained from the nonlinear analysis. ωt=-90° means the instant when the exciting current becomes minimum. Fig.8 shows the comparison of the flux densities near the steel plates obtained from the nonlinear and linear analyses. The results measured by Dr. Z.Cheng[9] are also shown. Tables II and III show fluxes linked with the steel plate and exciting coil. The positions A to C and G to I are shown in Fig.1. The discrepancies between the linear and nonlinear analyses and measurement are small. Fig.9 shows the comparison between the eddy current densities on the surface of the steel plate obtained from the linear and nonlinear analyses.

In order to investigate the discrepancy between the eddy currents obtained from the linear and nonlinear analyses, the effect of the permeability of the steel plate on the flux and eddy current distributions in the steel plate is investigated. Figs.10 and 11 show the results obtained. The flux and eddy current distributions in the steel plates are affected by the permeability due to the difference of the skin depth.

Fig.12 shows the eddy current loss in the steel plate. The eddy current loss *We* is calculated by the following equation:

$$We = \sum_{e=1}^{ne} \frac{1}{2\sigma^{(e)}} |Je^{(e)}|^2 V^{(e)} \tag{7}$$

where *Je* is the maximum value of eddy current density, *ne* is the number of elements in the steel plate and $V^{(e)}$ is the volume of an element *e*. The eddy current loss in the steel plate obtained using the linear analysis is decreased with the permeability. This is because the eddy current in the steel plate is reduced due to the increase of opposing field produced by eddy current when the permeability is increased.



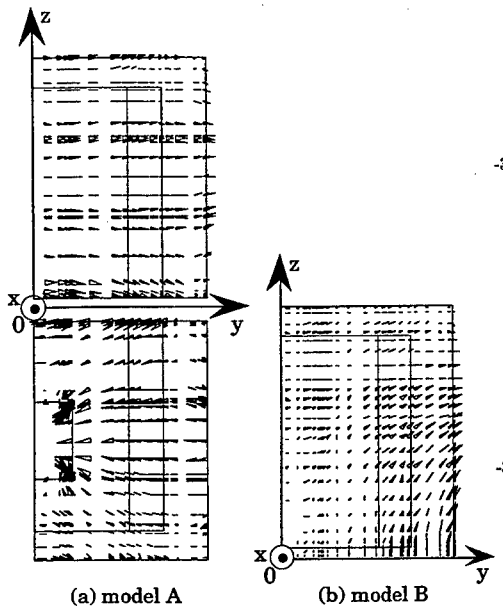
(a) whole region (b) steel plates and windings (model A)

Fig.6 Mesh.

Table I Discretization data and CPU time

model	A (1/2region)		B (1/4region)	
analysis	linear*1	nonlinear*2	linear*1	nonlinear*2
element type	1st-order brick edge			
no. of elements	60,060		33,480	
no. of nodes	65,178		36,704	
no. of unknowns	170,333		94,285	
no. of non-zeros	2,772,105		1,526,165	
memory requirement (MB)	123	78.4	66	43.3
no. of iterations for ICCG method	6,351	60,806	5,124	60,273
CPU time (h)	5.3	25.5	2.4	14.0

computer used : IBM3AT workstation (49.7MFLOPS)
convergence criterion for ICCG method :10^{-7}
*1 time harmonic finite element method (jω method)
*2 time periodic finite element method



(a) model A (b) model B

Fig.7 Eddy current distributions (x=-4.9mm).

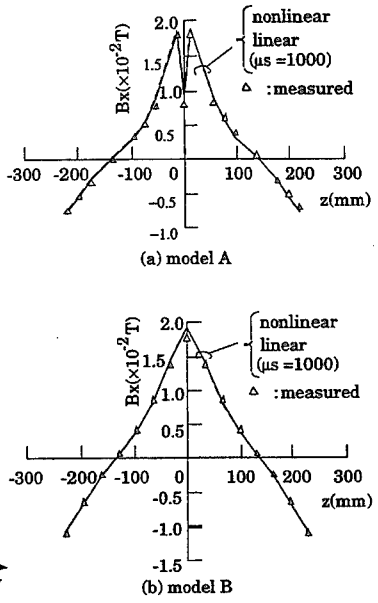


Fig.8 Flux distributions (x=5.76 , y=100).

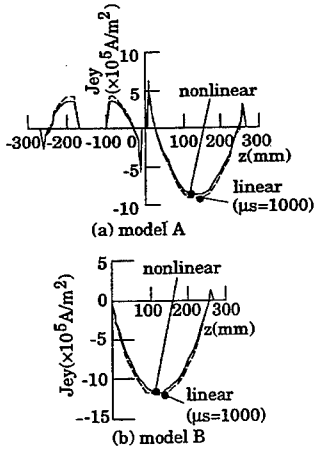


Fig.9 Eddy current distribution (x=4.9,y=0).

Table II Flux linked with steel plate

method		flux(×10-3Wb)		
		A	B	C
calculated	linear(μs=1000)	0.149	0.313	0.314
	nonlinear	0.149	0.312	0.312
measured		0.152	0.311	0.328

Table III Flux linked with exciting coils

method		flux(×10-3Wb)		
		G	H	I
calculated	linear(μs=1000)	0.822	0.825	0.824
	nonlinear	0.823	0.825	0.823
measured		0.836	0.830	0.843

Table IV shows the calculated values of eddy current loss and hysteresis loss. The total iron loss *Wt* in the steel plate measured using watt meter is also shown. Assuming that the hysteresis loss *Wh* is the function of the maximum flux density Bm, Wh is calculated by the following equation:

$$Wh = \sum_{e=1}^{ne} wh(Bm^{(e)})V^{(e)} \tag{8}$$

where *wh* is the dc hysteresis loss. Fig.13 shows *wh*-Bm curve. This is obtained by the following process: Firstly, dc hysteresis loop of steel plate is measured using a permeameter. Then, the area of hysteresis loop is calculated and this value is transferred to 50 Hz. Table IV suggests that the hysteresis loss *Wh* is not negligible even if the flux density in air is small. This is because the flux density near the surface of the steel plate is up to about 0.8T.

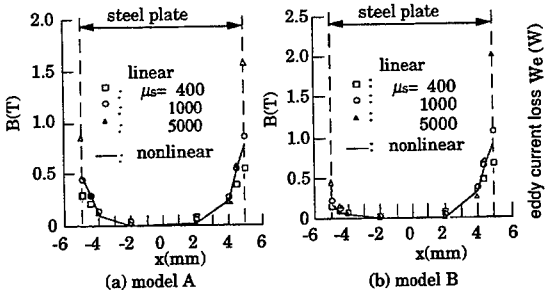


Fig.10 Flux distribution (y=0, z=140).

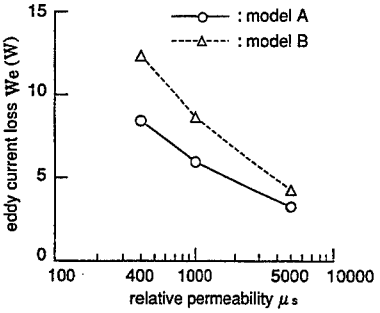


Fig.12 Effect of permeability on eddy current loss.

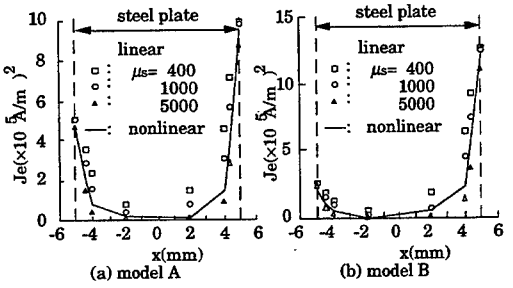


Fig.11 Eddy current distribution (y=0, z=140).

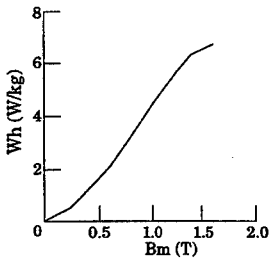


Fig.13 Hysteresis loss of steel plate (50Hz).

Table IV Comparison of calculated and measured iron losses

model	iron loss (W)			
	calculated			measured
	We	Wh	Wt	Wt
A	5.7	3.6	9.3	8.7
B	8.8	4.1	12.9	11.8

V. CONCLUSIONS

The nonlinear analysis of Problem 21 is carried out by using the time-periodic finite element method. The results obtained from the nonlinear analysis are compared with those obtained from the linear analysis and measurement. The results obtained can be summarized as follows:

(1) The CPU time for the nonlinear analysis using the time-periodic finite element method is only about 5 times longer than that for the linear analysis.

(2) The flux and eddy current distributions in the steel plate are affected by the permeability of the plate. Therefore, the nonlinear analysis is obligatory to investigate the eddy current loss in the steel plate.

(3)The hysteresis loss in steel plate is not negligible even if the flux density in air is small in some cases.

REFERENCES

[1] Z.Cheng, Q.Hu, S.Gao, Z.Liu, C.Ye, M.Wu, J.Wang and H.Zhu, "An engineering oriented loss model, Problem 21," *Proceedings of the Fourth International TEAM Workshop*, Miami, pp.137-143, 1993.

[2] N.Takahashi, T.Nakata, K.Fujiwara, H.Ohashi, H.Zhu, "Analysis of flux and eddy current distributions of 3-D stray-field loss model (Problem 21)," *Proceeding of the TEAM Workshop,* Aix-les-Bains, pp.21-24, 1994.

[3] T.Hara, T.Naito and J.Umoto, "Field analysis of corona shield region in high voltage rotating machines by time-periodic finite element method," *Journal of IEE, Japan*, vol. 102-B, no. 7, pp. 423-430, 1982.

[4] T.Nakata, N.Takahashi, K.Fujiwara and A.Ahagon, "3-D non-linear eddy current analysis using the time-periodic finite element method," *IEEE Trans. Magn.*, vol. 25, no. 5, pp. 4150-4152, 1989.

[5] T.Nakata, N.Takahashi, K.Fujiwara, K.Muramatsu, H.Ohashi and H.L.Zhu, " Practical analysis of 3-D dynamic nonlinear magnetic field using time-periodic finite element method," *IEEE Trans. Magn.*, vol. 31, no. 3, pp. 1416-1419, 1995.

[6] N. Takahashi, T. Nakata, K. Fujiwara, K. Muramatsu and T. Torii, " Time-periodic finite element analysis of problem 21," Proceedings of the TEAM Workshop, Berlin, pp.56-61, 1996.

[7] K.Fujiwara, "3-D magnetic field computation using edge element (invited)," *Proceedings of the 5th IGTE Symposium on Numerical Field Calculation in Electrical Engineering,* pp. 185-212, 1992.

[8] T.Nakata and Y.Kawase, "Numerical analysis of nonlinear transient magnetic field using the finite element method," *Journal of IEE, Japan,* vol. 104-B, no. 6, pp. 380-386, 1984.

[9] Z.Cheng, S.Gao, Q.Hu, "Summary of results for problem 21 (3-D stray-field loss model)," *Proceedings of the TEAM Workshop,* Berlin, pp.76-83, 1996.

Jacobi-Davidson Algorithm for Modeling Open Domain Lossy Cavities

Chibing Liu*and Jin-Fa Lee†

Abstract— **This paper presents the application of the newly developed Jacobi-Davidson (JD) algorithm to solve quadratic eigenmatrix equations. The quadratic eigenmatrix equations are resulted from using vector finite element methods to model open domain electromagnetic cavities. The derivation for the JD algorithm presented here uses Newton's method for solving non-linear equation. Consequently, it is intuitive to see the quadratic convergence rate for the basic algorithm when a good initial guess is provided. The complete JD procedure is then derived by combing the basic algorithm with Davidson's subspace method. Numerical examples show superquadratic or cubic convergence even when the correction equations are solved with only 10^{-1} accuracy.**

I. Introduction

In electromagnetics, eigenvalue problems include cavity resonance and wave propagation in both closed and open structures, such as metallic waveguides, open and shielded microstrip transmission lines, optical waveguides, or fibers, etc. These problems can be dealt by finite element methods (FEMs). The resultant system established by the FEM is in the form of:

$$\mathbf{A}(k)\,\widetilde{X} = 0 \tag{1}$$

Where $\mathbf{A}(k)$ is a complex and sparse matrix and is a function of wavenumber, k. The complex wavenumber k and vector $\widetilde{X}$ are the eigenpair to be solved. Here, one aims to find the resonance wavenumber k which makes $\mathbf{A}(k)$ singular; the corresponding notrivial eigenvector $\widetilde{X}$ is the resonant mode. In general, we are interested in only a few dominant modes. For a lossless closed cavity, it is well known that Eq. 1 can be written as:

$$\left(\mathbf{A}_0 + k^2\mathbf{A}_2\right)\widetilde{X} = 0 \tag{2}$$

Where $\mathbf{A}_0$ and $\mathbf{A}_2$ are real symmetric matrices [1]. Equation 2 is a generalized eigenmatrix equation and there are several approaches available to solve such problems [2],[3]. However, for a lossy cavity filled with lossy materials and/or with open domain (modeled by the use of 1st order absorbing boundary condition), the eigenmatrix equation will be of the form:

$$\left(\mathbf{A}_0 + k\mathbf{A}_1 + k^2\mathbf{A}_2\right)\widetilde{X} = 0 \tag{3}$$

Where $\mathbf{A}_0$, $\mathbf{A}_1$ and $\mathbf{A}_2$ are complex square matrices. Equation 3 is a quadratic eigenmatrix equation. The conventional approach to solve 3 is to convert it into a generalized eigenmatrix equation by employing auxiliary matrices and vectors of larger sizes. The drawbacks of such approaches are apparent. Alternatively, this paper investigates the use of the newly developed Jacobi-Davidson algorithm [4], [5], [6] to solve for the lossy and/or open cavity problems.

II. Notations

We shall use, throughout this paper, bold letters to denote matrices.

*This project is sponsored by ERS International.

†The authors are with EMCAD Lab., ECE Dept. Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609. Further correspondances should send to jinlee@ece.wpi.edu and related publications can be found in http://ece.wpi.edu/~jinlee.

C:	The set of complex numbers.
Ω :	The problem domain.
$\partial\Omega$:	The boundary of domain Ω.
Γ_{PEC} :	PEC boundary surface.
Γ_{PMC} :	PMC boundary surface.
μ :	Permeability.
μ_r :	Relative permeability.
ϵ :	Permittivity.
ϵ_r :	Relative permittivity.
σ :	Conductivity.
k :	The wavenumber in free-space.
η :	Intrinsic imepdance of free-space.
$(\mathbf{M})_{ij}$:	The ij entry of the matrix $\mathbf{M}$.
$\overrightarrow{W^{(i)}}$:	i-th vector basis function.
$\overrightarrow{v_\tau}$:	The tangential components of $\overrightarrow{v}$.
$\overrightarrow{E}$:	The electric field.
$\overrightarrow{H}$:	The magnetic field.
$\widetilde{\ }$:	Indicates a column vector.
$\underset{\sim}{\ }$:	Indicates a row vector.
$\widetilde{t}$:	Transpose of the vector.
$\{\widetilde{u}^{\perp}\}$:	$\{\widetilde{v} \mid \widetilde{v} \perp \widetilde{u}\}$.
$A \longleftarrow B$:	Update A with B.
$\oplus$:	Direct sum.
$N(\bullet)$:	The null space of operator $\bullet$.
$R(\bullet)$:	The range space of operator $\bullet$.

III. FEM Formulation for Lossy Cavities

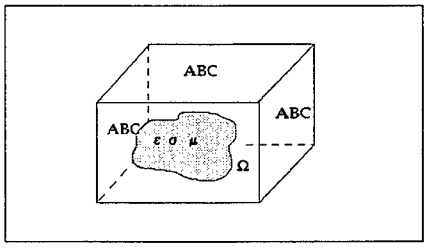


Figure 1 – An open domain lossy cavity.

Shown in Fig. 1 is an open domain lossy cavity. Formulate it in terms of the electric field, $\overrightarrow{E}$, results in the following boundary value problem (BVP):

$$\begin{aligned}
&\nabla \times \tfrac{1}{\mu_r} \nabla \times \overrightarrow{E} - jk\eta\sigma \overrightarrow{E} - k^2 \epsilon_r \overrightarrow{E} = 0 && \text{in } \Omega \\
&\hat{n} \times \nabla \times \overrightarrow{E} = -jk\overrightarrow{E_\tau} && \text{on } \partial\Omega \\
&\hat{n} \times \overrightarrow{E} = 0 && \text{on } \Gamma_{PEC} \\
&\hat{n} \times \nabla \times \overrightarrow{E} = 0 && \text{on } \Gamma_{PMC}
\end{aligned} \tag{4}$$

Applying the vector FEM procedure to Eq. 4 (the hierarchical vector basis functions are described in Refs. [7] and [8]), we obtain the following quadratic eigenmatrix equation:

$$\left(\mathbf{A}_0 + k\mathbf{A}_1 + k^2\mathbf{A}_2\right)\widetilde{X} = 0 \tag{5}$$

with

$$\begin{aligned}
(\mathbf{A}_0)_{ij} &= \int_\Omega \left(\nabla \times \overrightarrow{W^{(i)}}\right) \bullet \frac{1}{\mu_r}\left(\nabla \times \overrightarrow{W^{(j)}}\right) dx^3 \\
(\mathbf{A}_1)_{ij} &= -j\eta \int_\Omega \overrightarrow{W^{(i)}} \bullet \sigma \overrightarrow{W^{(j)}} dx^3 \\
&\quad -j \int_{\partial\Omega} \overrightarrow{W_\tau^{(i)}} \bullet \overrightarrow{W_\tau^{(j)}} dx^2 \\
(\mathbf{A}_2)_{ij} &= -\int_\Omega \overrightarrow{W^{(i)}} \bullet \epsilon_r \overrightarrow{W^{(j)}} dx^3
\end{aligned} \tag{6}$$

IV. Jacobi-Davidson Algorithm

Given a quadratic eigenmatrix equation of the form

$$\left(\mathbf{A}_0 + \mathbf{A}_1\lambda + \mathbf{A}_2\lambda^2\right)\widetilde{X} = 0 \tag{7}$$

where $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2$ are sparse $n \times n$ complex matrices. In this paper, they are generated from the application of FEM method to three-dimensional open domain lossy cavities. The pair $\left\langle \lambda, \widetilde{X} \right\rangle$ which satisfies 7 is referred to as an eigenpair. Practically, n could be a very large number, tens of thousands or hundreds of thousands are not uncommon. Thus, it is desirable to derive an algorithm for solving 7 which does not require solving matrix equations, of order n, exactly or with high accuracy. Lanczos and/or Arnoldi algorithms [9] are already widely used in engineering community for solving standard and generalized eigenmatrix equations. However, without shiftings as preconditioners, Krylov subspace methods typically do require solving matrix equations with high accuracy. This does not necessarily exclude the use of iterative matrix solution techniques, for examples pre-conditioned Conjugate Gradient

(PCCG) method, to solve these matrix equations. But, when the matrix is very ill-conditioned or a good preconditioner is not readily available, direct methods are usually the only choice. In this paper, we shall employ the newly developed Jacobi-Davidson algorithm which is advocated by van der Vorst [5] and his colleague recently. As pointed out in [5], the JD algorithm does not require explicit factorization of a large matrix and the accuracy of solving the correction equation needs not be high. In our experiences, a relative residual of 10^{-1} is sufficient to achieve superquadratic or even faster convergences.

A. *The Basic Algorithm*

Based on 7, we shall first construct a mapping

$$f\left(\lambda,\widetilde{X}\right)=\mathbf{A}_0\widetilde{X}+\lambda\mathbf{A}_1\widetilde{X}+\lambda^2\mathbf{A}_2\widetilde{X}=P(\lambda)\,\widetilde{X} \quad (8)$$

which maps $C\times C^n$ into C^n. In 8, the matrix polynomial $P(\lambda)$ is defined as $P(\lambda)=\mathbf{A}_0+\lambda\mathbf{A}_1+\lambda^2\mathbf{A}_2$. Consequently, the original eigenmatrix equation 7 is transformed into

$$\begin{array}{c} \textit{find } \left\langle\lambda,\widetilde{X}\right\rangle\in C\times C^n \\ \textit{such that} \\ f\left(\lambda,\widetilde{X}\right)=0 \end{array}$$

Assuming at the k-th step of the iterative process, an approximate solution $\langle\theta,\widetilde{u}\rangle$ is obtained and satisfies

$$\underset{\sim}{u}\,\mathbf{A}_0\widetilde{u}+\theta\,\underset{\sim}{u}\,\mathbf{A}_1\widetilde{u}+\theta^2\,\underset{\sim}{u}\,\mathbf{A}_2\widetilde{u}=\underset{\sim}{u}\,P(\theta)\,\widetilde{u}=0 \quad (9)$$

Our objective is to find the next solution $\left\langle\theta',\widetilde{u'}\right\rangle$ which is a *better* approximation than $\langle\theta,\widetilde{u}\rangle$. To do so, we apply the Newton's method for solving nonlinear equation and set

$$\begin{aligned}\theta' &= \theta+\delta \\ \widetilde{u'} &= \widetilde{u}+\widetilde{z}\end{aligned} \quad (10)$$

Substituting 10 into 8 and collecting only up to the first order terms results in

$$\begin{aligned} f\left(\theta',\widetilde{u'}\right) &= \mathbf{A}_0(\widetilde{u}+\widetilde{z})+(\theta+\delta)\,\mathbf{A}_1(\widetilde{u}+\widetilde{z}) \\ &\quad +(\theta+\delta)^2\,\mathbf{A}_2(\widetilde{u}+\widetilde{z}) \\ &\approx (\mathbf{A}_0+\theta\mathbf{A}_1+\theta^2\mathbf{A}_2)\,\widetilde{u} \\ &\quad +\delta(\mathbf{A}_1+2\theta\mathbf{A}_2)\,\widetilde{u} \\ &\quad +(\mathbf{A}_0+\theta\mathbf{A}_1+\theta^2\mathbf{A}_2)\,\widetilde{z} \end{aligned} \quad (11)$$

The procedure to compute the update $\left\langle\theta',\widetilde{u'}\right\rangle$ is to set

$$f\left(\theta',\widetilde{u'}\right)\approx P(\theta)\,\widetilde{u}+\delta\widetilde{w}+P(\theta)\,\widetilde{z}=0 \quad (12)$$

where $\widetilde{w}=(\mathbf{A}_1+2\theta\mathbf{A}_2)\,\widetilde{u}=P'(\theta)\,\widetilde{u}$. We shall use 12 to solve for δ and the correction vector $\widetilde{z}$.

To obtain an update for θ', or to compute δ, we simply left multiply $\underset{\sim}{u}$ to Eq. 12 and note that $\underset{\sim}{u}\,P(\theta)\,\widetilde{u}=0$, we have

$$\delta=-\frac{\underset{\sim}{u}\,P(\theta)\,\widetilde{z}}{\underset{\sim}{u}\,\widetilde{w}}=-\frac{(P(\theta)\,u)^t\,\widetilde{z}}{\underset{\sim}{u}\,\widetilde{w}}=-\frac{\underset{\sim}{r}\,\widetilde{z}}{\underset{\sim}{u}\,\widetilde{w}} \quad (13)$$

Obviously, the residual vector $\widetilde{r}$ equals to $P(\theta)\,\widetilde{u}$. It remains to find an appropriate expression to compute the correction vector $\widetilde{z}$. From Eqs. 12 and 13, it can be shown that

$$\left(I-\frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)P(\theta)\,\widetilde{z}=-\widetilde{r} \quad (14)$$

Also, it will be proven later that for any $\widetilde{r}\perp\widetilde{u}$, there exists an unique complex vector $\widetilde{q}$ such that

$$\left(I-\frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)P(\theta)\left(I-\frac{\widetilde{u}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{u}}\right)\widetilde{q}=-\widetilde{r} \quad (15)$$

Once Eq. 15 is solved for $\widetilde{q}$, the correction vector $\widetilde{z}$ is obtained simply by $\widetilde{z}=\left(I-\frac{\widetilde{u}\underset{\sim}{u}}{\underset{\sim}{u}\widetilde{u}}\right)\widetilde{q}$. Thus, we have a complete numerical scheme to compute the update $\left\langle\theta'=\theta+\delta,\widetilde{u'}=\widetilde{u}+\widetilde{z}\right\rangle$: first solve for $\widetilde{q}$ from 15, obtain the orthogonal correction $\widetilde{z}$ by $\widetilde{z}=\left(I-\frac{\widetilde{u}\underset{\sim}{u}}{\underset{\sim}{u}\widetilde{u}}\right)\widetilde{q}$, and finally compute δ by Eq. 13.

To conclude this section, we present a basic algorithm, *Algorithm 1*, that can be used to compute an approximate eigenpair for the quadratic eigenmatrix equation 7.

Algorithm 1

Initialization :

- Choose a non-trivial initial vector $\widetilde{u}$, and solve

$$\left(\underset{\sim}{u}\,\mathbf{A}_0\widetilde{u}\right)+\left(\underset{\sim}{u}\,\mathbf{A}_1\widetilde{u}\right)\theta+\left(\underset{\sim}{u}\,\mathbf{A}_2\widetilde{u}\right)\theta^2=0 \quad (16)$$

for θ.

- Compute the initial residual vector $\widetilde{r_0}$ by

$$\widetilde{r_0} = P(\theta)\,\widetilde{u} = \left(\mathbf{A}_0 + \theta\mathbf{A}_1 + \theta^2\mathbf{A}_2\right)\widetilde{u} \quad (17)$$

- Set the residual vector $\widetilde{r} = \widetilde{r_0}$.

Iterations :

Iterate through the following steps until converged.

- Update $\widetilde{w}$ by $\widetilde{w} = P'(\theta)\,\widetilde{u} = (\mathbf{A}_1 + 2\theta\mathbf{A}_2)\,\widetilde{u}$.
- Solve for $\widetilde{q}$ from

$$\left(I - \frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right) P(\theta) \left(I - \frac{\widetilde{u}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{u}}\right) \widetilde{q} = -\widetilde{r}$$

- Compute the correction vector $\widetilde{z}$ by

$$\widetilde{z} = \left(I - \frac{\widetilde{u}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{u}}\right) \widetilde{q}$$

and $\widetilde{u} = \widetilde{u} + \widetilde{z}$.

- Evaluate $\delta = -\frac{\underset{\sim}{r}\,\widetilde{z}}{\underset{\sim}{u}\,\widetilde{w}}$ and update θ by $\theta = \theta + \delta$.
- Update the residual vector $\widetilde{r}$ by $\widetilde{r} = P(\theta)\,\widetilde{u} = \left(\mathbf{A}_0 + \theta\mathbf{A}_1 + \theta^2\mathbf{A}_2\right)\widetilde{u}$.
- If $\frac{\|\widetilde{r}\|}{\|\widetilde{r_0}\|} \leq \xi$, where ξ is a prescribed tolerance, then stop and $\langle\theta, \widetilde{u}\rangle$ is a good approximation to $\left\langle\lambda, \widetilde{X}\right\rangle$.

B. Quadratic Convergence Rate

In this section, we shall prove that the basic algorithm outlined in Algorithm 1 will exhibit a quadratic convergence rate when a good initial pair is employed. We will establish this fact through a series of Theorems.

Theorem 1 *For any two complex vectors $\widetilde{u}, \widetilde{w} \in C^n$, $\left(\underset{\sim}{u}\,\widetilde{w} \neq 0\right)$, we define a mapping of the form*

$$M_w^u = \left(I - \frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)$$

Then,

1. the null space $N(M_w^u) = span\{\widetilde{w}\}$
2. the range $R(M_w^u) = \left\{\widetilde{u}^{\perp}\right\}$.

Proof: The proof is trivial.

Theorem 2 *Assume that at $k-th$ step of the iteration of Algorithm 1, we have*

$$\underset{\sim}{u}\left(\mathbf{A}_0 + \theta\mathbf{A}_1 + \theta^2\mathbf{A}_2\right)^{-1}(\mathbf{A}_1 + 2\theta\mathbf{A}_2)\,\widetilde{u} \neq 0 \quad (18)$$

Then for every $\widetilde{y} \in \left\{\widetilde{u}^{\perp}\right\}$, the mapping

$$F(\theta) = \left(I - \frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right) P(\theta) \left(I - \frac{\widetilde{u}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{u}}\right)$$

is non-singular. Namely, $\forall\widetilde{y} \in \left\{\widetilde{u}^{\perp}\right\}, F(\theta)\,\widetilde{y} \neq 0$.

Proof: We shall prove it by contradiction.

- Assume that there exists an $\widetilde{y} \in \left\{\widetilde{u}^{\perp}\right\}$ such that $F(\theta)\,\widetilde{y} = 0$.
- It follows that

$$\begin{aligned} P(\theta)\,\widetilde{y} &= \alpha\widetilde{w} \\ &= \alpha\left(\mathbf{A}_1 + 2\theta\mathbf{A}_2\right)\widetilde{u} \end{aligned} \quad (19)$$

where $\alpha = \frac{\underset{\sim}{u}\left(\mathbf{A}_0 + \theta\mathbf{A}_1 + \theta^2\mathbf{A}_2\right)\widetilde{y}}{\underset{\sim}{u}\,\widetilde{w}}$.

- Since θ $(\theta \neq \lambda)$ is not an eigenvalue of Eq. 7, $\left(\mathbf{A}_0 + \theta\mathbf{A}_1 + \theta^2\mathbf{A}_2\right)$ is non-singular. Also, the coefficient $\alpha \neq 0$ for non-trivial $\widetilde{y}$. For if $\alpha = 0$, then the condition $\left(\mathbf{A}_0 + \theta\mathbf{A}_1 + \theta^2\mathbf{A}_2\right)\widetilde{y} = 0$ implies $\widetilde{y} \equiv 0$.
- Moreover, from Eq. 19, we have

$$\widetilde{y} = \alpha\left(P(\theta)\right)^{-1} P'(\theta)\;\widetilde{u} \quad (20)$$

Equation 20 and the fact that $\widetilde{y} \in \left\{\widetilde{u}^{\perp}\right\}$ yield

$$\alpha\left(\underset{\sim}{u}\left(P(\theta)\right)^{-1} P'(\theta)\,\widetilde{u}\right) = 0 \quad (21)$$

- Because $\alpha \neq 0$, the only possible solution to Eq. 21 is

$$\underset{\sim}{u}\left(\mathbf{A}_0 + \theta\mathbf{A}_1 + \theta^2\mathbf{A}_2\right)^{-1}(\mathbf{A}_1 + 2\theta\mathbf{A}_2)\,\widetilde{u} = 0 \quad (22)$$

which contradicts to the assumption 18. ♯

Additionally, since the matrix $\left(\mathbf{A}_0 + \theta\mathbf{A}_1 + \theta^2\mathbf{A}_2\right)$ becomes more singular when $\langle\theta, \widetilde{u}\rangle$ converges to $\left\langle\lambda, \widetilde{X}\right\rangle$, we should also consider the limit when $\langle\theta, \widetilde{u}\rangle = \left\langle\lambda, \widetilde{X}\right\rangle$ whether or not the mapping $F(\lambda)$ is non-singular for any $\widetilde{y} \in \left\{\widetilde{X}^{\perp}\right\}$. The following Theorem guarantees just that.

Theorem 3 *Suppose* $\left\langle \lambda, \widetilde{X} \right\rangle$ *is an eigenpair of Eq. 7 and with* λ *a simple eigenvalue. Assuming also that* $\underset{\sim}{X}\ \widetilde{X} \neq 0$ *and* $\underset{\sim}{X}\ \widetilde{w} = \underset{\sim}{X}\ (\mathbf{A}_1 + 2\lambda \mathbf{A}_2)\ \widetilde{X} \neq 0$, *then the map*

$$F(\lambda) = \left(I - \frac{\widetilde{w}\ \underset{\sim}{X}}{\underset{\sim}{X}\ \widetilde{w}} \right) P(\lambda) \left(I - \frac{\widetilde{X}\ \underset{\sim}{X}}{\underset{\sim}{X}\ \widetilde{X}} \right)$$

is non-singular for any non-trivial right-hand-side vector $\widetilde{y} \in \left\{ \widetilde{X}^{\perp} \right\}$.

Proof: Again, we shall prove it by contradiction.

- Suppose there exists a non-trivial vector $\widetilde{y} \in \left\{ \widetilde{X}^{\perp} \right\}$ such that $F(\lambda)\widetilde{y} = 0$, then

$$\begin{aligned} P(\lambda)\widetilde{y} &= \frac{\underset{\sim}{X}\ \left(\mathbf{A}_0 + \lambda \mathbf{A}_1 + \lambda^2 \mathbf{A}_2 \right) \widetilde{y}}{\underset{\sim}{X}\ \widetilde{w}} \widetilde{w} \\ &= \frac{\left(P(\lambda)\widetilde{X} \right)^t \widetilde{y}}{\underset{\sim}{X}\ \widetilde{w}} \widetilde{w} = 0 \end{aligned} \tag{23}$$

- Since $P(\lambda)\widetilde{y} = 0$, $\langle \lambda, \widetilde{y} \rangle$ must be an eigenpair of Eq. 7. Namely, both $\widetilde{X}$ and $\widetilde{y}$ are eigenvectors with the same eigenvalue λ. This contradicts to the assumption that λ is a simple eigenvalue. ♯

Theorem 4 *The null space and the range of the mapping* $F(\theta)$ *are: (a)* $N(F(\theta)) = span\{\widetilde{u}\}$; *and (b)* $R(F(\theta)) = \left\{ \widetilde{u}^{\perp} \right\}$.

Proof:

- It is obvious to see that $F(\theta)\widetilde{u} \equiv 0$. Together with Theorem 2 and 3, we have $N(F(\theta)) = span\{\widetilde{u}\}$.
- From theorem 1, we have $R(F(\theta)) \subset \left\{ \widetilde{u}^{\perp} \right\}$. Moreover, the facts

$$\begin{aligned} &\dim(N(F(\theta))) + \dim(R(F(\theta))) = n \\ &\dim(N(F(\theta))) = 1 \end{aligned} \tag{24}$$

  implying that $\dim(R(F(\theta))) = n - 1$.
- Since $\dim\left(\left\{ \widetilde{u}^{\perp} \right\} \right) = n - 1$, therefore, we conclude $R(F(\theta)) = \left\{ \widetilde{u}^{\perp} \right\}$. ♯

Remark 1 *Theorem 4 guarantees that for any* $\widetilde{r} \in \left\{ \widetilde{u}^{\perp} \right\}$, *there exists an unique* $\widetilde{q} \in \left\{ \widetilde{u}^{\perp} \right\}$ *such that* $F(\theta)\widetilde{q} = -\widetilde{r}$.

Theorem 5 *Assume that the correction equation is solved exactly in Algorithm 1, and also* $\underset{\sim}{u}\ \widetilde{w} \neq 0$, $\underset{\sim}{u}\ \widetilde{u} \neq 0$ *in every step. Then, if the initial vector* $\widetilde{u}$ *is close enough to* $\widetilde{X}$, *the sequence of* $\langle \theta, \widetilde{u} \rangle$ *converges to* $\left\langle \lambda, \widetilde{X} \right\rangle$, *and the convergence is quadratic. Namely, given a good initial approximate eigenpair* $\langle \theta, \widetilde{u} \rangle$, *the update* $\left\langle \theta', \widetilde{u}' \right\rangle$ *satisfies*

$$\|\widetilde{X} - \widetilde{u}'\| \propto O\left(\|\widetilde{X} - \widetilde{u}\|^2 \right) \tag{25}$$

Proof:

- Define the error vector of the initial vector $\widetilde{u}$ as $\widetilde{e} = \widetilde{X} - \widetilde{u}$.
- In the iteration, we solve for the correction vector $\widetilde{z}$ by

$$\left(I - \frac{\widetilde{w}\ \underset{\sim}{u}}{\underset{\sim}{u}\ \widetilde{w}} \right) F(\theta)\widetilde{z} = -\widetilde{r} \tag{26}$$

  and update $\widetilde{u}$ by $\widetilde{u}' = \widetilde{u} + \widetilde{z}$ (note $z \in \left\{ \widetilde{u}^{\perp} \right\}$). It follows that $\widetilde{X} - \widetilde{u}' = \widetilde{e} - \widetilde{z}$.
- From the fact $\left(\mathbf{A}_0 + \lambda \mathbf{A}_1 + \lambda^2 \mathbf{A}_2 \right) \widetilde{X} = 0$, we have

$$\begin{aligned} P(\theta)\widetilde{e} &= -\widetilde{r} \\ &\quad + \left(\Delta^2 \mathbf{A}_2 - \Delta P'(\lambda) \right) \widetilde{X} \end{aligned} \tag{27}$$

  with $\Delta = \lambda - \theta$. Neglecting the second order term (Δ^2), we can approximate Δ by

$$\Delta \approx -\frac{\underset{\sim}{u}\ P(\theta)\widetilde{e}}{\underset{\sim}{u}\ (\mathbf{A}_1 + 2\lambda \mathbf{A}_2)\ \widetilde{X}} = -\frac{\underset{\sim}{r}\ \widetilde{e}}{\underset{\sim}{u}\ P'(\lambda)\ \widetilde{X}} \tag{28}$$

  It is easy to see that

$$|\Delta| \leq \frac{\|\widetilde{r}\|}{\left| \underset{\sim}{u}\ P'(\lambda)\ \widetilde{X} \right|} \|\widetilde{e}\| \tag{29}$$

- Multiplying Eq. 27 by $\left(I - \frac{\widetilde{w}\ \underset{\sim}{u}}{\underset{\sim}{u}\ \widetilde{w}} \right)$ and subtracting Eq. 26 yields $\left(\text{note } \left(I - \frac{\widetilde{w}\ \underset{\sim}{u}}{\underset{\sim}{u}\ \widetilde{w}} \right) \widetilde{r} = \widetilde{r} \right)$

$$\left(I - \frac{\widetilde{w}\ \underset{\sim}{u}}{\underset{\sim}{u}\ \widetilde{w}} \right) P(\theta)\left(\widetilde{e} - \widetilde{z} \right)$$

$$= \left(I - \frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)\left(\Delta^2 \mathbf{A}_2 - \Delta P'(\lambda)\right)\widetilde{X}$$

$$= \Delta^2\left(I - \frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)\mathbf{A}_2\widetilde{u} + \left(I - \frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)\left(\Delta^2 \mathbf{A}_2 - \Delta P'(\lambda)\right)\widetilde{e} \tag{30}$$

- Taking L^2 norm on both sides of Eq. 30 results in

$$\begin{aligned}
&\left\|\left(I - \frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)P(\theta)\left(\widetilde{e} - \widetilde{z}\right)\right\| \\
&= \left\| \Delta^2\left(I - \frac{\widetilde{w}\underset{\sim}{u}}{\underset{\sim}{u}\widetilde{w}}\right)\mathbf{A}_2\widetilde{u} + \left(I - \frac{\widetilde{w}\underset{\sim}{u}}{\underset{\sim}{u}\widetilde{w}}\right)\left(\Delta^2\mathbf{A}_2 - \Delta P'(\lambda)\right)\widetilde{e} \right\| \\
&\leq |\Delta^2|\left\|\left(I - \frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)\mathbf{A}_2\widetilde{u}\right\| + \left\|\left(I - \frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)\left(\Delta\mathbf{A}_2 - P'(\lambda)\right)\Delta\widetilde{e}\right\| \\
&\leq \beta_1|\Delta^2| + \beta_2\|\Delta\widetilde{e}\| \leq \gamma\|\widetilde{e}\|^2
\end{aligned} \tag{31}$$

where $\gamma = \beta_1 \frac{\|\widetilde{r}\|^2}{\left|\underset{\sim}{u}P'(\lambda)\widetilde{X}\right|^2} + \beta_2 \frac{\|\widetilde{r}\|}{\left|\underset{\sim}{u}P'(\lambda)\widetilde{X}\right|}$. It is easy to see that $\beta_1, \beta_2 < \infty$, and unless $\underset{\sim}{u}\,P'(\lambda)\,\widetilde{X} = 0$, also $\gamma < \infty$.

- Since $(\widetilde{e} - \widetilde{z}) \in \{\widetilde{u}^{\perp}\}$ and from Theorem 2 and 3, we conclude that there exists $\alpha > 0$ such that

$$\left\|\left(I - \frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)P(\theta)\left(\widetilde{e} - \widetilde{z}\right)\right\| \geq \alpha\|\widetilde{e} - \widetilde{z}\| \tag{32}$$

- By combining Eqs. 31 and 32, we have

$$\|\widetilde{e} - \widetilde{z}\| \leq \frac{\gamma}{\alpha}\|\widetilde{e}\|^2 \tag{33}$$

which proves the quadratic convergence rate of Algorithm 1.

C. *Jacobi-Davidson Algorithm for Quadratic Eigenmatrix Equation*

In examining Algorithm 1 closely, one can observe a major inefficiency. That is, in building the next update, Algorithm 1 uses only the information of $\langle\theta, \widetilde{u}\rangle$ and discards all other information acquired through the entire iterative process. By combining Algorithm 1 and the subspace projection that is commonly adopted in Lanczos and Arnoldi methods, we arrive the complete Jacobi-Davidson algorithm, Algorithm 2, for solving the quadratic eigenmatrix equation. Since Algorithm 2 obtained a better approximation than Algorithm 1 (in fact, the best projection available from the current search vector space) in every step, we should expect better convergence rate than quadratic. In many numerical examples, we observed most of time super-quadratic and sometimes cubic rate of convergence.

Algorithm 2: Jacobi-Davidson Algorithm for Solving the Dominant Eigenpair of a Quadratic Eigenmatrix Equation

Initialization :

- Set $k = 0$, and choose a non-trivial initial vector $\widetilde{v_0}$ and do

$$\widetilde{v_0} \longleftarrow \frac{\widetilde{v_0}}{\|\widetilde{v_0}\|}, \qquad \mathbf{V}_0 = \{\widetilde{v_0}\} \tag{34}$$

Iteration :

- Compute

$$\mathbf{W}_0 \longleftarrow \mathbf{A}_0\mathbf{V}_k, \quad \mathbf{W}_1 \longleftarrow \mathbf{A}_1\mathbf{V}_k, \quad \mathbf{W}_2 \longleftarrow \mathbf{A}_2\mathbf{V}_k \tag{35}$$

and

$$\mathbf{H}_0 \longleftarrow \mathbf{V}_k^t\mathbf{W}_0, \quad \mathbf{H}_1 \longleftarrow \mathbf{V}_k^t\mathbf{W}_1, \quad \mathbf{H}_2 \longleftarrow \mathbf{V}_k^t\mathbf{W}_2 \tag{36}$$

- Solve the eigenpair $\langle\theta = \theta_{\max}, \widetilde{y}\rangle$ from the reduced quadratic eigenmatrix equation

$$\left(\mathbf{H}_0 + \theta\mathbf{H}_1 + \theta^2\mathbf{H}_2\right)\widetilde{y} = 0 \tag{37}$$

- Compute

$$\begin{aligned}
\widetilde{u} &\longleftarrow \mathbf{V}_k\widetilde{y}, \\
\widetilde{w} &\longleftarrow \left(\mathbf{A}_1 + 2\theta\mathbf{A}_2\right)\widetilde{u}, \\
\widetilde{r} &\longleftarrow \left(\mathbf{A}_0 + \theta\mathbf{A}_1 + \theta^2\mathbf{A}_2\right)\widetilde{u}
\end{aligned} \tag{38}$$

- If $\frac{\|\widetilde{r}\|}{\|\widetilde{r}_0\|} < \xi$, where ξ is a prescribed tolerance, then stop and $\langle\theta, \widetilde{u}\rangle$ is good approximation to $\left\langle\lambda_{\max}, \widetilde{X}\right\rangle$.

- Solve for $\widetilde{t}$ approximately

$$\left(I-\frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)P(\theta)\left(I-\frac{\widetilde{u}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{u}}\right)\widetilde{t}\approx-\widetilde{r} \quad (39)$$

and compute the correction vector $\widetilde{z}$ by $\widetilde{z}=\left(I-\frac{\widetilde{u}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{u}}\right)\widetilde{t}$.

- Expand the search space $\mathbf{V}_k$ by modified Gram-Schmidt process, namely

$$\widetilde{v}_{k+1}=\widetilde{z}-\sum_{i=0}^{k}\left(\underset{\sim}{z}\,\widetilde{v}_i\right)\widetilde{v}_i \quad (40)$$

and

$$\widetilde{v}_{k+1}\longleftarrow\frac{\widetilde{v}_{k+1}}{\|\widetilde{v}_{k+1}\|},\quad \mathbf{V}_{k+1}=\mathbf{V}_k\oplus\{\widetilde{v}_{k+1}\} \quad (41)$$

- $k=k+1$.

What remains is how to solve the correction vector $\widetilde{z}$ practically without explicitly forming the map $F(\theta)$. This question can be answered by the following Theorem.

Theorem 6 *The correction vector $\widetilde{z}$ in the $k-th$ iteration of Algorithm 2 can be computed by*

$$\widetilde{w}'=\left(\mathbf{A}_0+\theta\mathbf{A}_1+\theta^2\mathbf{A}_2\right)^{-1}\widetilde{w},\quad \widetilde{z}=-\widetilde{u}+\frac{\underset{\sim}{u}\,\widetilde{u}}{\underset{\sim}{u}\,\widetilde{w}'}\widetilde{w}' \quad (42)$$

Proof:

- First, it is easy to see that $\underset{\sim}{u}\,\widetilde{z}=0$ in Eq. 42, regardless of what $\widetilde{w}'$ is.
- Secondly,

$$\left(I-\frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)P(\theta)\widetilde{z}$$
$$=\left(I-\frac{\widetilde{w}\,\underset{\sim}{u}}{\underset{\sim}{u}\,\widetilde{w}}\right)\left(-\widetilde{r}+\frac{\underset{\sim}{u}\,\widetilde{u}}{\underset{\sim}{u}\,\widetilde{w}'}\widetilde{w}\right)$$
$$=-\widetilde{r} \quad (43)$$

Consequently, $\widetilde{z}$ computed from Eq. 42 is a solution to the correction equation, and it follows from Theorem 2, the only solution.

D. Sample Numerical Examples

In example 1, the matrix dimension is $n=29,980$, and we are to solve for the smallest (modulus) eigenvalue. It corresponds to $\lambda=-0.1467113+j0.1276269$. The second example, we explore the performance of the Jacobi-Davidson algorithm for an interior eigenpair. We randomly choose $\lambda_0=0.5$ and solve for the eigenpair that is closest to it. The results of these two examples are summarized in Table 1. Note in both examples, the correction equations are solved approximately, with the accuracy in terms of relative residual of 10^{-1}, using an PCCG (pre-conditioned Conjugate Gradient) [10] method. For the smallest eigenpair, we observe a superquadratic convergence, whereas for the interior eigenpair, we even observe supercubic convergence.

Iter.	Example 1	
	θ	$\log\left(\frac{\|\widetilde{r}\|}{\|\widetilde{r}_0\|}\right)$
0	$0.2756432-j0.2748964$	0
1	$-0.3003151-j0.3104875$	-0.818
2	$-0.01645236+j0.121385$	-1.481
3	$-0.1422936+j0.1281352$	-3.375
4	$-0.1467113+j0.1276269$	-7.252

Iter.	Example 2	
	θ	$\log\left(\frac{\|\widetilde{r}\|}{\|\widetilde{r}_0\|}\right)$
0	$0.153926+j0.5926128$	0
1	$0.300567+j0.0966729$	-0.4145274
2	$0.47915432-j0.762863$	-1.8485773
3	$0.47343582+j0.238413$	-6.161849

Table 1: Convergence of two eigenpairs for a quadratic eigenmatrix equation with n = 29,980.

Moreover, we have also computed the smallest eigenvalue of this quadratic eigenmatrix equation by using a modified Arnoldi's [3] algorithm which uses $\mathbf{A}_0^{-1}\mathbf{A}_2$ to generate the necessary Krylov subspace from an initial random vector. The performance of this modified Arnoldi algorithm with the said Krylov subspace is compared against the current Jacobi-Davidson's algorithm in Fig. 2. The Krylov based approach, since it is not quite exact for the quadratic eigenmatrix equation, only exhibits linear convergence.

V. Conclusions and Discussions

In closing up this paper, we should point out two major areas regarding the Jacobi-Davidson's algorithm that in our opinions still require extensive re-

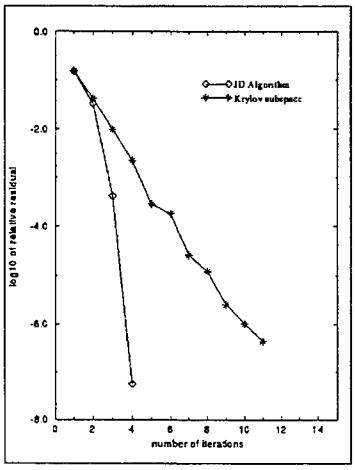


Figure 2 – Convergence comparison between JD algorithm and a modified Arnoldi approach for solving a quadratic eigenmatrix equation.

search work. One is the issue of multiple modes. A possible solution would be to restart with a different initial guess, however, there is no guarantee that this leads to a new eigenpair. Or we may employ the popular *deflation* technique. That is when an eigenvector has converged, then we continue in a subspace spanned by the remaining eigenvectors. But, the benefits of deflation and/or selective orthogonalization type processes are questionable for quadratic (or polynomial) eigenproblems when the matrices do not commute. Secondly, in Algorithm 2, most of the computational effort (CPU time) is spent on solving the correction equation on the last iteration. This is mainly due to the fact when the Ritz pair is almost converged, the matrix equation becomes very ill-conditioned (almost singular). Without proper care, the preconditioners that based on, one form or another, LU factorization will become extremely unstable and subsequently hinder the PCCG convergence even for 10^{-1} accuracy. A matrix solver that resolves this difficulty and removes any unwanted component in the correction vector will improve further the efficiency of the algorithm.

References

[1] S.G. Perepelitsa, R. D. Edlinger, and J.F. Lee. "Finite-Element Analysis of Arbitrarily Shaped Cavity Resonators Using $H_1(curl)$ Elements". *IEEE Trans. Magn.*, 33(2):1176–1179, Mar. 1997.

[2] D.S. Scott. "Solving Sparse Symmetric Generalized Eigenvalue Problems without Factorization". *SIAM J. Numer. Anal.*, 18:102–110, 1981.

[3] D.C. Sorensen. "Implicit Application of Polynomial Filters in a k-step Arnoldi Method". *SIAM J. Matrix Anal. Appl.*, 13(1):357–385, 1992.

[4] G.L.G. Sleijpen and H.A. van der Vorst. "A Jacobi-Davidson iteration Method for linear eigenproblem". *SIAM J. Matrix Anal. Appl.*, 17(2):401–425, April 1996.

[5] G. Sleijpen, H.A. van der Vorst, and M. van Gijzen. "Quadratic Eigenproblems are no Problem". *SIAM News*, Sep. 1996.

[6] G. Sleijpen, J.G.L. Booten, D.R. Fokkema, and H.A. van der Vorst. "Jacobi-Davidson Type Methods for Generalized Eigenproblems and Polynomial Eigenproblems: Part 1.". *Technical Report 923, University Utrecht, Department of Mathematics*, 1995.

[7] R. D. Edlinger and O. Biro. "A Joint Vector and Scalar Potential Formulation for Driven High Frequency Problems Using Hybrid Edge and Nodal Finite Elements". *IEEE Trans. Microwave Theory Tech.*, 44(1):15–23, Jan. 1996.

[8] J.P. Webb and B. Forghani. "Hierarchal Scalar and Vector Tetrahedra". *IEEE Trans. Magn.*, 29:1495–2498, March 1993.

[9] J.F. Lee, D.K. Sun, and Z.J. Cendes. "Full-Wave Analysis of Dielectric Waveguides Using Tangential Vector Finite Elements". *IEEE Trans. Microwave Theory Tech.*, 39(8):1262–1271, Aug. 1991.

[10] O. Axelsson. *"Iterative Solution Methods"*. Cambridge University Press, New York, 1994.

Analysis of electromagnetic penetration through apertures of shielded enclosure using finite element method

Byung-Wook Kim, Yeon-Choon Chung, Tae-Weon Kang

EMC Group, Korea Research Inistitute of Standards and Science, Taejon, 305-606, KOREA

Abstract - The full wave analysis of electromagnetic field penetrated through the apertures of a shielded enclosure is discussed. The electromagnetic field inside a shielded enclosure is dependent on the physical dimension of a metallic enclosure as well as the size, shape and number of apetures. Analysis including the above two considerations by 3-D finite element method and measurement are performed for some shielded enclosure having front panels of different-type apertures. Good agreements were obtained between analytical and empirical results.

I. INTRODUCTION

As the worldwide rules and regulations on electromagnetic immunity start, shielding techniques become more important. For the design of an excellent shielded enclosure, we must account for the two important considerations. One is the choice of good shielding material, the other is the design of a good shielded enclosure including holes and penetrations. The former becomes important in lower frequency range while the latter is important in higher frequency range. This paper deals with electromagnetic penetrations through cavity-backed apertures.

The electromagnetic field inside a shielded enclosure is dependent on the physical dimension of a metallic enclosure itself, as well as the size, shape and number of apetures. Simple predictions are possible for the electromagnetic penetration through the apertures of an infinite shieled plate[1, 2]. But these methods can not provide informations about the geometrical effects of the shielded enclosure, such as cavity resonances. 3-D numerical analysis techniques must be used to improve these weak points. FDTD, TLM methods can be used to analyze these struectures, but because the size of an aperture is relatively small compared with that of a shielded enclosure, it is difficult to apply the straight forward approach of these methods[3,4].

In this paper, finite element method with boundary integral method is presented for the electromagnetic penetrations through the cavity-backed-apertures. The validation of this analytical approach is confirmed by comparing calculation results with measurement results

II. FORMULATION

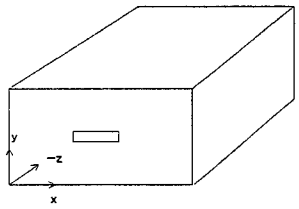


Fig. 1 Shielded enclosure having apertures.

Fig. 1 shows a typical shielded enclosure having aperture(s). Coordinate axes are chosen as shown in Fig. 1. To apply finite elemet method, the shielded enclosure is separated into two regions. One is the outside region of a shielded enclosure, the other is inside of that. Electromagnetic fields inside a shielded enclosure is obtained by FEM, and the outside field is obtained by boundary integral method using Green's function. And field continuity condition is applied at the apetures, electromagnetic filed which is valid everywhere is finally obtained.

1. Boundary integral equation

The electric field in free space outside of shielded enclosure satisfies the vector wave equation

$$\nabla \times \nabla \times E(r) - k_0^2 E(r) = -jk_0 Z_0 J(r) \tag{1}$$

To find the radiated fields we intorduce the dyadic Green's function $\overline{\overline{G}}_e(r, r')$ which satisfies the inhomogeneous differential equation

$$\nabla \times \nabla \times \overline{\overline{G}}_e(r, r') - k_0^2 \overline{\overline{G}}_e(r, r') = \overline{\overline{I}}\delta(r - r') \tag{2}$$

By using this Green's function electic field can be written as

$$E(r) = -jk_0 Z_0 \int\int\int_{V_\infty} J(r') \cdot \overline{\overline{G_e}}(r, r')\, dV' - \oint\oint_{S_\infty} \{[\hat{n} \times E(r')] \cdot [\nabla' \times \overline{\overline{G_e}}(r, r')] + [\hat{n} \times \nabla' \times E(r')] \cdot \overline{\overline{G_e}}(r, r')\} dS' \quad (3)$$

where S_∞ denotes the surface enclosing V_∞ and $\hat{n}$ is the outward unit vector normal to the surface and ∇' has been used to indicate that ∇ operates only on the primed coordinates. If the size of the aperture is smaller than the size of shielded enclosure, half space Green's function may be used to obtain electic filed. By this introduction, the electric field can be expressed as

$$E(r) = -jk_0 Z_0 \int\int\int_{V_\infty} J(r') \cdot \overline{\overline{G_0}}(r, r')\, dV' + jk_0 Z_0 \int\int\int_{V_\infty} J(r') \cdot [\overline{\overline{G_0}}(r, r_i') - 2\hat{z}\hat{z} G_0(r, r')]\, dV' + 2\int\int_{S_a} [\hat{z} \times E(r')] \cdot [\nabla' \times \overline{\overline{G_0}}(r, r')]\, dS' \quad (4)$$

The first term on the right-hand side is the field radiated by J in the free-space environment, thus denoted as E^{inc}. The second term is the field radiated by the image current of J, thus denoted as E^{ref}. The third term is scattered field due to the aperture. With these indentifications, equtiaon (4) can be written as

$$E(r) = E^{inc} + E^{ref} + 2\int\int_{S_a} [\hat{z} \times E(r')] \cdot [\nabla' \times \overline{\overline{G_0}}(r, r')]\, dS' \quad (5)$$

which also can be written as

$$\nabla \times E(r) = -jk_0 Z_0 H^{inc} - jk_0 Z_0 H^{ref(r)} - 2k_0^2 \int\int_{S_a} [\hat{z} \times E(r')] \cdot \overline{\overline{G_0}}(r, r')\, dS' \quad (6)$$

Equations (6) denotes relationship between the electric and magnetic field. By letting z approach zero, we obtain

$$\hat{z} \times [\nabla \times E(r)]_{z=0+} = -jk_0 Z_0 \hat{z} \times H^{inc}(r) - 2k_0^2 \hat{z} \times \int\int_{S_a} [\hat{z} \times E(r')] \cdot \overline{\overline{G_0}}(r, r')\, dS' \quad (7)$$

2. Finite element formulation

By assuming there no source inside the shielded enclosure electric field satisfies the vector wave equation

$$\nabla \times \left(\frac{1}{\mu_r} \nabla \times E\right) - k_0^2 \varepsilon_r E = 0 \quad r \in V \quad (8)$$

At the shielding enclousre wall, the tangential electric field vanished.

$$\hat{n} \times E = 0 \quad (9)$$

At the opening an equivalent boundary condition can be obtained from the integral equation (6) as

$$\hat{z} \times \left[\frac{1}{\mu_r} \nabla \times E\right]_{z=0+} + P(E) = U^{inc} \quad (10)$$

where,

$$U^{inc} = -2jk_0 Z_0 \hat{z} \times H^{inc}(r)$$

$$P(E) = 2k_0^2 \hat{z} \times \int\int_{S_a} [\hat{z} \times E(r')] \cdot \overline{\overline{G_0}}(r, r')\, dS'$$

The equivalent variational problem for this is given by

$$\begin{cases} \delta F(E) = 0 \\ \hat{n} \times E = 0 \quad \text{at cavity wall} \end{cases} \quad (11)$$

where

$$F(E) = \frac{1}{2}\int\int\int_V \left[\frac{1}{\mu_r} (\nabla \times E) \cdot (\nabla \times E) - k_0^2 \varepsilon_r E \cdot E\right] dV - \int\int_{S_a} \left[\frac{1}{2} E \cdot P(E) - E \cdot U^{inc}\right] dS \quad (12)$$

To reduce its singularities some mathematical calcuatioans were performed to obtain

$$F(E) = \frac{1}{2}\int\int\int_V \left[\frac{1}{\mu_r} (\nabla \times E) \cdot (\nabla \times E) - k_0^2 \varepsilon_r E \cdot E\right] dV - k_0^2 \int\int_{S_a} [\hat{z} \times E(r)] \cdot \left\{\int\int_{S_a} [\hat{z} \times E(r')] G_0(r, r')\, dS'\right\} dS + \int\int_{S_a} (\nabla \cdot [\hat{z} \times E(r)]) \left\{\int\int_{S_a} G_0(r, r') \nabla' \cdot [\hat{z} \times E(r')]\, dS'\right\} dS + 2jk_0 Z_0 \int\int_{S_a} [\hat{z} \times E(r)] \cdot H^{inc}(r)\, dS \quad (13)$$

3. Bondary condition of the probe

To prove the validation of our calculation, measurements were performed using probe, details of the probe will be presented in section III. To compare calculated results with measured, the situation of probe inserted is approximated by FEM. In this case we must impose the boundary conditions at the surface of the probe and at the cross section of the coaxial cable which is connected to the probe. The probe of our measurement system is composed of conducting material, so at the surface of the probe equation (9) must be imposed. At the cross section of the cable the fields may be considered as transverse electromagnetic field. The boundary condition of this case is

$$\hat{n} \times E + jk_0 E = 0 \quad (14)$$

To impose this bouandary condition, F_p must be added to equation (13)

$$F_p = \int\int_{S_t} \frac{jk_0}{2} (\hat{n} \times E) \cdot (\hat{n} \times E)\, dS_t \quad (15)$$

where S_t denote the cross section of the coaxial cable.

III. Results

Fig. 2 shows test set-up for measuring the electromagnetic shielding effectiveness of a shielded enclosure. Some test results were obtained by replacing front panels of a shielded enclosure. A monopole probe was injected into shielded enclosure to measure an inside electric field strength. The probe was made in the cylindrical type of 80 mm length and 2 mm diameter. The size of test box, represents shielded enclosure, is 200 mm by 400 mm by 500 mm. The measurements were performed inside full anochonic chamber which its refection loss is below -20 dB from 30 MHz to 1 GHz frequencies ranges. To measure

reference values, test box is removed and isotropic probe was placed at the same position. The square of the electric field measured by isotropic is used as reference values. By this procedure, we may compare measured and calculated values easily. Electromagnetic shielding effectiveness is defined as the differences between reference and measured values. The distance between test box and antenna is maintained 3 m which is typical in the radiated susceptibility measurement in accordance with commercial standards.

To consider the effects of a probe inside shielded enclosure, we approximate a probe using rectangular brick elements to reduce memory demand as shown in Fig. 3. Integraion between two conductor is performed to obtain voltage differences V_d. To compare this calculated values with measured values, the expected values at the receiver (Spectrum analyzer) may be obatined by

$$P = 20\log|V_d| - 10\log 50 + \text{System Gain Loss} \qquad (16)$$

Fig. 4 show the measured and calculated results. Good agreement between measured and calcuated can be found. Fig 5 shows comparison between our calculation and the calculation of Schulz[2]. Good agreement between two results is found. However, Schulz's calculation can't show the geometrical effect of a shielded enclosure at resonant frequencies of the struecture and field distribution inside a shielded enclosure, while those are provided by our calculation.

. References

[1]. Bernhard Keiser, *Principles of electromagnetic compatibility 3rd editions*, Artech House, 1987.

[2]. Richard B. Schulz, "Shielding theory and pratice," IEEE Trans. on EMC., vol 30, No. 3, pp 187-201, August 1988.

[3]. Kuang-Ping Ma, Min Li, James L. Drewniak, Todd H. Hubing and Thomas P. Van Doren, "Comparison of FDTD Algorithms for Subcelluar Modeling of Slots in shielding enclosures", IEEE Trans. EMC., vol. 39, No. 2, pp. 147-155,May 1997.

[4]. J. D. Turner, T.M. Benson, C. Christopoulos and D. W. P. Thomas, "Characterisation of the shielding effectiveness of equipment cabinets containing apetures", EMC '96 Roma International sysmposium on electromagnetic compatibility vol. 2, pp. 574-578, ELLEPI, 1996.

[5]. Jianming Jin, *The finite element method in electromagnetics*, John Wiley & Sons, Inc. 1993.

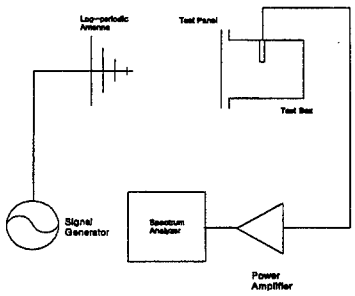


Fig 2. Test setup for measuring the shielding effectiveness of a shielded enclosure.

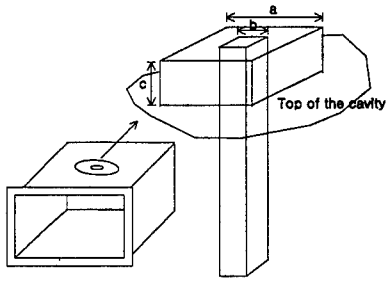


Fig 3. Finite element approximation of probe

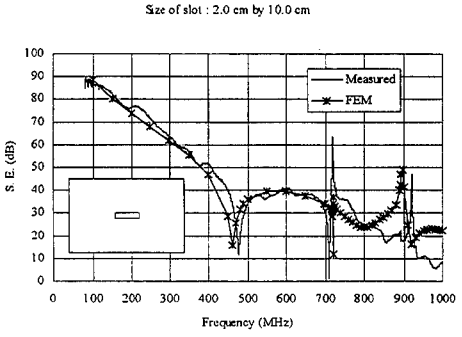


Fig 4. Comparision of calculated and measured results.

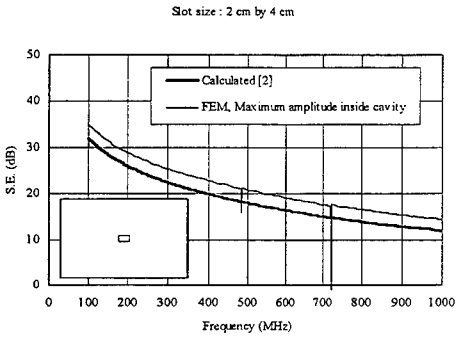


Fig 5. Comparision of our calculated and Schulz's[2]

hp-ADAPTIVE EDGE FINITE ELEMENTS FOR MAXWELL'S EQUATIONS

L. Demkowicz, L. Vardapetyan and W. Rachowicz
The Texas Institute for Computational and Applied Mathematics
The University of Texas at Austin
Taylor Hall 2.400
Austin, Texas 78712, USA

Abstract

A model problem for the steady-state form of Maxwell's equations is considered. The problem is recast into a weak form using a Lagrange multiplier, laying down a foundation for a general class of novel *hp*-adaptive FE approximations. The proposed method is illustrated and verified by a series of 2D experiments which include domains with curved boundaries and nonhomogeneous media.

Key words: Maxwell's equations, *hp* finite elements, covariant projection, error estimates

AMS subject classification: 65N30, 35L15

1 Introduction

The goal of the presented research is to design a stable finite element method for the steady-state Maxwell equations in domains with complex geometries and/or multiple media with varying electromagnetic properties. Singular solutions are expected in both classes of problems as the result of rapidly changing material constants and/or rough geometries.

One of the most powerful methodologies which permit successful modeling of singular solutions is the *hp*-adaptive finite elements [1, 3]. We emphasize that a true *hp* method allows *to vary locally both* element size h and order of approximation p. Only then the exponential rates of convergence are accessible for a wide class of functions with singularities.

The main results of the proposed approach are:

- a formulation for a class of problems with discontinuous material properties, which is uniformly stable with respect to frequency ω, as $\omega \to 0$,
- a discretization with the possibility of varying locally order of approximation p and element size h,
- a curved element to model complex, curvilinear geometry.

In this communication we outline the main points and results of our methodology. For theoretical details on the proposed method we refer to [4, 12] and for details on numerical work to [10].

2 Model Problem. Mixed Variational Formulation. Stability, Existence, and Uniqueness.

We consider the following model problem. A bounded domain Ω consists of two disjoint parts $\Omega_i, i = 1, 2$, filled by possibly lossy media with given parameters $\epsilon_i, \mu_i, \sigma_i, i = 1, 2$, and with an interface Γ_{12}. The boundary Γ of the domain Ω consists of two disjoint parts: the electric wall Γ_1 ($\boldsymbol{E} \times \boldsymbol{n} = 0$ on Γ_1) and the magnetic wall Γ_2 ($(\boldsymbol{n}(j\omega\epsilon+\sigma)\boldsymbol{E}) = 0, \quad \boldsymbol{n} \times (\frac{1}{\mu}\nabla \times \boldsymbol{E}) = 0$ on Γ_2).

We wish to solve for the electric field $\boldsymbol{E}$ excited in Ω by a given time-harmonic $exp(j\omega t)$, divergence-free impressed current $\boldsymbol{J}^{imp}$ subject to the appropriate compatibility conditions [6].

The standard variational formulation for the problem reads as follows:

$$\begin{cases} \boldsymbol{E} \in \boldsymbol{W}, \\ \int_\Omega \frac{1}{\mu}(\nabla \times \boldsymbol{E}) \circ (\nabla \times \bar{\boldsymbol{F}})\, d\boldsymbol{x} \quad - \int_\Omega (\omega^2 \epsilon - j\omega\sigma)\boldsymbol{E} \circ \bar{\boldsymbol{F}}\, d\boldsymbol{x} \\ \qquad\qquad = -j\omega \int_\Omega (\boldsymbol{J}^{imp} \circ \bar{\boldsymbol{F}})\, d\boldsymbol{x}, \quad \forall \boldsymbol{F} \in \boldsymbol{W} \end{cases} \tag{2.1}$$

where $\boldsymbol{W} \stackrel{\text{def}}{=} \{\boldsymbol{E} \in \boldsymbol{H}(\text{curl}, \Omega) : \boldsymbol{n} \times \boldsymbol{E} = \boldsymbol{0} \text{ on } \Gamma_1\}$ is the space of admissible fields equipped with the norm: $\|\boldsymbol{E}\|_{\boldsymbol{W}} = (\|\boldsymbol{E}\|^2_{2,\epsilon,\Omega} + \|\nabla \times \boldsymbol{E}\|^2_{2,\mu,\Omega})^{\frac{1}{2}}$.

Introducing a space of scalar functions $V \stackrel{\text{def}}{=} \{p \in H^1(\Omega) : p = 0 \text{ on } \Gamma_1\}$ and substituting $\boldsymbol{F} = \nabla q \quad (q \in V)$ into (2.1) , we find that the solution $\boldsymbol{E}$ satisfies the continuity equation in the weak sense:

$$\int_\Omega (j\omega\epsilon + \sigma)\boldsymbol{E} \circ \nabla \bar{q} = 0, \quad \forall g \in V \tag{2.2}$$

The main idea of the proposed formulation lies in enforcing (2.2) explicitly at the cost of an extra unknown function - a Lagrange multiplier p. The resulting mixed variational formulation is as follows:

$$\begin{cases} \boldsymbol{E} \in \boldsymbol{W}, \quad p \in V \\ \int_\Omega \frac{1}{\mu}(\nabla \times \boldsymbol{E}) \circ (\nabla \times \bar{\boldsymbol{F}})\, d\boldsymbol{x} \; - \int_\Omega (\omega^2\epsilon - j\omega\sigma)(\boldsymbol{E} + \nabla p) \circ \bar{\boldsymbol{F}}\, d\boldsymbol{x} \\ \qquad\qquad = -j\omega \int_\Omega (\boldsymbol{J}^{imp} \circ \bar{\boldsymbol{F}})\, d\boldsymbol{x}, \quad \forall \boldsymbol{F} \in \boldsymbol{W} \\ \int_\Omega (\omega^2\epsilon - j\omega\sigma)\boldsymbol{E} \circ \nabla \bar{q}\, d\boldsymbol{x} \qquad = 0 \quad \forall q \in V \end{cases} \tag{2.3}$$

Incidentally, we have learned recently that our formulation is equivalent to that of F.Kikuchi's for eigenvalue problems [8].

We consider $\boldsymbol{W}_0 \stackrel{\text{def}}{=} \{\boldsymbol{E} \in \boldsymbol{W} : \nabla \times \boldsymbol{E} = 0\}$ and assume that Ω and the boundary conditions are such [1] that:

$$\boldsymbol{E}_0 \in \boldsymbol{W}_0 \Longleftrightarrow \exists \phi \in V : \boldsymbol{E}_0 = \nabla \phi \tag{2.4}$$

With this compatibility assumption, it can be shown that the variational problem (2.3) has a unique solution and the stability properties of the formulation are frequency independent for $\omega \to 0$. The stability properties of the standard variational formulation (2.1) deteriorate as $\omega \to 0$ [4].

3 Edge Elements of Variable Order

What follows is a brief description of scalar and vector triangular elements. The same approach is valid for rectangles in 2-D and for prisms, cubes, and tetrahedra in 3-D.

We associate with a master triangle $\hat{K}$ a specific order of approximation $p = p_K$ which may vary from element to element. Additionally, with each of its sides $\hat{s}_i$ we associate a possibly different order of approximation $p_i, i = 1, 2, 3, \quad p_i \leq p$. [2]

We introduce two spaces of element shape functions:

[1] We emphasize that Ω need not be simply connected!

[2] In practice, the order of element sides is fixed using the *minimum rule*, i.e. the order of approximation for a side shared by elements K_1, K_2 is set to $min\{p_{K_1}, p_{K_2}\}$.

- the *scalar space* (to approximate the Lagrange multiplier p)

$$\hat{V}(\hat{K}) \stackrel{\text{def}}{=} \{\hat{q} \in \mathcal{P}^{p+1}(\hat{K}) \ : \ \hat{q}|_{\hat{s}_i} \in \mathcal{P}^{p_i+1}(\hat{s}_i), i = 1,2,3\} \tag{3.5}$$

- the *vector space* (to approximate the $\boldsymbol{E}$-field)

$$\hat{\boldsymbol{W}}(\hat{K}) \stackrel{\text{def}}{=} \{\hat{\boldsymbol{F}} \in \mathbf{P}^p(\hat{K}) \ : \ \hat{\boldsymbol{F}}|_{\hat{s}_i} \circ \hat{\boldsymbol{\tau}}_i \in \mathcal{P}^{p_i}(\hat{s}_i), i = 1,2,3\} \tag{3.6}$$

  where $\mathcal{P}^n$ denotes the space of polynomials of order n, and $\hat{\tau}_i$ is a tangent vector to the side $\hat{s}_i, i = 1,2,3$.

Both spaces are constructed as spans of hierarchical shape functions:

- Scalar Shape Functions:

  1. Vertex Shape Functions:

  $\chi_i(\lambda_1, \lambda_2, \lambda_3) = \lambda_i, \quad i = 1,2,3$

  where λ_i are area coordinates $0 \leq \lambda_i \leq 1$

  2. Side Shape Functions:

  $\chi_{s,i} = \chi_s \chi_j \frac{(\chi_j - \chi_s)^{(i-1)}}{2}, \quad s = 1,2,3; \quad j = (s+1) mod(3); \quad i = 1, \ldots, p_s$

  3. Middle Node Shape Functions:

  $\chi_{mid,(i+1)} = \chi_1 \chi_2 \chi_3 \chi_2{}^i \chi_3{}^{j-i}, \quad j = 0, \ldots, (p-2); \quad i = 0, \ldots, j$

- Vector Shape Functions:

  1. "Tangential" Vertex Shape Functions:

  vertex $(0,0)$: $\quad \psi_1 = (\chi_1, 0), \quad \psi_6 = (0, -\chi_1)$

  vertex $(0,1)$: $\quad \psi_2 = (\chi_2, \chi_2), \quad \psi_3 = (0, \chi_2)$

  vertex $(1,0)$: $\quad \psi_4 = (-\chi_3, 0), \quad \psi_5 = (-\chi_3, -\chi_3)$

  2. "Tangential" Side Shape Functions:

  $\psi^t{}_{s,i} = \chi_{s,i} \tau_s; \qquad s = 1,2,3; \quad i = 1, \ldots, p_s - 1$

  3. "Normal" Side Shape Functions:

  $\psi^n{}_{s,i} = \chi_{s,i} \boldsymbol{n}_s; \qquad s = 1,2,3; \quad i = 1, \ldots, p-1$

  where τ_s and $\boldsymbol{n}_s$ are unit tangent and unit normal vectors to the side s.

  4. Middle Node Shape Functions:

  $\psi_{mid,i} = (\chi_{mid,i}, 0); \quad \psi_{mid,i+N} = (0, \chi_{mid,i}); \qquad N = \frac{(p-1)(p-2)}{2}; \quad i = 1, \ldots, N$

These vector elements are flexible enough to model the continuous tangential component of $\boldsymbol{E}$-field and at the same time allow discontinuity in the normal component of the field across the media interfaces.

The corresponding vector-valued FE approximation is $H(\mathrm{curl}, \Omega)$-conforming provided that vector-valued functions defined on the master element are mapped onto the mesh as "gradients" by covariant projection. [9, 5, 2]

The discrete version of compatibility condition (2.4) for $\boldsymbol{W}_h$ and V_h serves the same role as the "inclusion condition" discussed in [2], sorting out non-physical contributions to the numerical solution.

4 Convergence Result

It can be proved that there exists a threshold value h_0 and a constant C, independent of h such that $\forall h \leq h_0$ the following estimate holds:

$$\|\boldsymbol{E} - \boldsymbol{E}_h\|_{\boldsymbol{W}} \leq C \inf_{\boldsymbol{F}_h \in \boldsymbol{W}_h} \|\boldsymbol{E} - \boldsymbol{F}_h\|^2_{\boldsymbol{W}}, \tag{4.7}$$

Explicit inclusion of the constraint on the divergence guarantees that constant C remains bounded when $\omega \to 0$. When combined with the interpolation error estimates for hp-approximations, estimate (4.7) results in standard hp-error estimates with exponential rates of convergence for analytic solutions.

5 Numerical Examples

We illustrate the algorithm with a solution of the benchmark problem of two concentric cylinders of dissimilar media. The radii of the cylinders are 1.0 m and 0.25 m, and the inner cylinder is off-centered by 0.5 m. The wave numbers of the cylinders are k^2 =12.5 m^{-2} and k^2 =125 m^{-2}, respectively. A uniform magnetic field $\boldsymbol{H}$ is imposed along the boundary of the outer cylinder. The larger cylinder is discretized with 128 quadrilateral elements while the smaller one with 128 triangular elements, the order of approximation is uniform, $p = 2$. Figures 1,2 show the contour maps of the x- and y-components of the electric field $\boldsymbol{E}$ together with the finite element mesh.

Finally, Figure 3 presents a preliminary result of hp-mesh optimization for a simple model problem with a polynomial exact solution. At this point we would like to illustrate only the

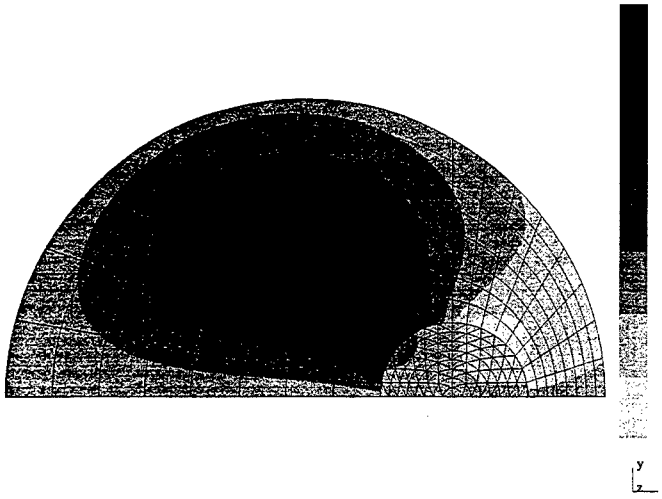


Figure 1: Benchmark solution of the off-centered cylinders problem, the x-component of the electric field.

possibility of both p and h-refinements.

Acknowledgment

The authors would like to thank Professors P.Monk, I. Babuška, D. Neikirk, and J.T. Oden for many interesting discussions on the subject. The work of the first and the third authors has been partially supported by the NSF under Contract DMS 9414480, The support of the second author has been provided from DOD-Air Force Grant F49620-96-1-0032 (P.I. - D. Neikirk).

References

[1] I. Babuška and B. Q. Guo, "Approximation Properties of the *hp* Version of the Finite Element Method", *Computer Methods in Applied Mechanics and Engineering, Special Issue on p and hp- Methods*, eds. I Babuška and J. T. Oden, **133**, 319-346, 1996.

[2] C.W.Crowley , P.P. Silvester, H.Hurwitz Jr, "Covariant Projection Elements for 3D Vector Field Problems", *IEEE Transactions on Magnetics*, **24**,No.1, 397-400, 1988

[3] L. Demkowicz, J. T. Oden, W. Rachowicz, and O. Hardy, "Toward a Universal *hp* Adaptive Finite Element Strategy. Part1: Constrained Approximation and Data Structure", *Computer Methods in Applied Mechanics and Engineering* **77**, 1-2, 79-112, 1989.

[4] L. Demkowicz, L. Vardapetyan , "Modeling of Electromagnetic Absoption/Scattering Problems Using *hp*-adaptive Finite Elements", *Computer Methods in Applied Mechanics and Engineering*, to appear in 1997.

[5] F.Dubois, "Discrete Vector Potential Representation of a Divergence-Free Vector Field in Three-Dimensional Domains: Numerical Analysis of a Model Problem", *SIAM J.Numerical Analysis*,**v.27**,No.5, 1103-1141, 1990

[6] B.N. Jiang, J. Wu, and L.A. Povinelli, "The Origin of Spurious Solutions in Computational Electromagnetics", *Journal of Computational Physics* **125**, 104-123, 1996.

[7] J.Jin, *The Finite Element Method in Electromagnetics*, John Wiley & Sons, Inc., New York 1993.

[8] F.Kikuchi, "Mixed and Penalty Formulations for Finite Element Analysis of an Eigenvalue Problem in Electromagnetism" *Computer Methods in Applied Mechanics and Engineering*,**64**, 509-521, 1987

[9] J.C. Nedelec, "Mixed Finite Elements in $\mathbb{R}^3$", *Numerische Mathematik*, **35**, 315-341, 1980.

[10] W. Rachowicz, L. Demkowicz, "2Dhp90-EM, a Two-Dimensional *hp*-Adaptive FE Package for Electromagnetics", *TICAM Report*, in preparation.

[11] P.P. Silvester, and G. Pelosi (eds.), *Finite Elements for Wave Electromagnetics*, IEEE Press, New York, 1994.

[12] L. Vardapetyan, L. Demkowicz, "*hp*-Adaptive Finite Elements in Electromagnetics", *TICAM Report 97-18*, in review for *Computer Methods in Applied Mechanics and Engineering.*

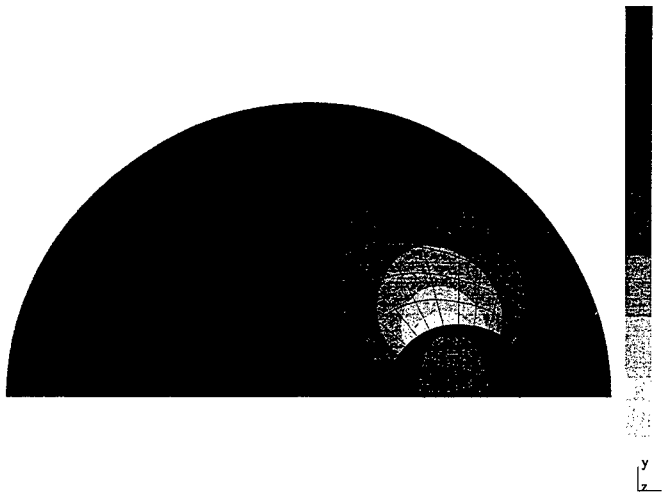


Figure 2: Benchmark solution of the off-centered cylinders problem, the y-components of the electric field.

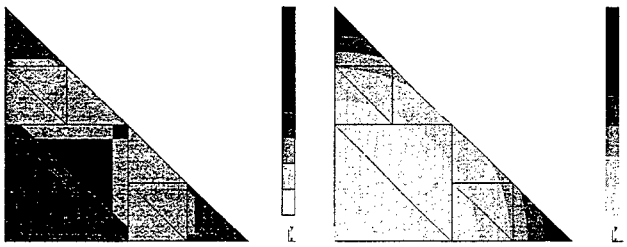


Figure 3: Solution of a 2D Maxwell equations model problem. An optimal *hp* FE mesh (left) and the corresponding contours of the the x- component of the electric field. Final error measured in the energy norm is 0.1 percent of the total energy of the solution.

SESSION 14:

RECENT ADVANCES IN TIME-DOMAIN TECHNIQUES

Chairs:D.C. Blake and D.J. Riley

An Analysis of Programming Models for Time-Domain CEM Codes on RISC-Based Computers

Douglas C. Blake and Joseph S. Shang
Air Force Research Laboratory, Wright-Patterson AFB, OH 45433

Abstract

An analysis of programming paradigms for Reduced Instruction Set Computing (RISC)-based architectures is conducted by examining the performance of three finite-volume time domain (FVTD) computer codes. Each of these codes employs an identical numerical algorithm for solving the Maxwell equations, yet each uses a different program structure to do so. One code utilizes a common program structure tailored to traditional vector computers. The second code was derived from the first by restructuring the control loops in order to modify the data access patterns of the program. Finally, the third code was developed specifically for RISC-based architectures and places particular emphasis on data locality. Details of each code implementation are presented. Single-processor performance is analyzed using the hardware performance counters of the Silicon Graphics, Inc. R10000 processor on an Octane workstation and an Origin2000 supercomputer.

1. Introduction

Today, the largest and most complex electromagnetic scattering and wave propagation simulations are often conducted on massively parallel computing platforms. Nearly all of these machines are characterized first by very fast RISC central processing units (CPUs) which operate at clock speeds in the hundreds of megahertz, and second by large addressable memories on the order of tens to hundreds of gigabytes. Although capacious, the memory systems are quite slow when compared to the processing rate of the CPU. This disparity in speed makes a memory fetch a potentially expensive operation, at times causing the CPU to wait while data is loaded. In order to alleviate this problem, modern machines typically possess a small amount of very fast cache memory which is designed to exploit temporal and spatial locality in data-access patterns. It is possible to intelligently fetch data from main memory and load it into the cache so that a large portion of memory accesses are satisfied from the fast cache rather than the slow main memory. This has the potential of dramatically improving performance.

Before the proliferation of massively parallel machines, supercomputing was dominated by vector processors, and thus, a large number of scientific and engineering computer codes is use today were originally written for vector machines. A well-written vector code can routinely achieve 600-700 million floating point operations per second (Mflops) on a single vector processor rated at roughly 950 Mflops [6,9]. Unfortunately, this same computer code may achieve only 20-50 Mflops when executed on a RISC processor rated at 500+ Mflops [3,12]. Therefore, rather than measure code performance on RISC-based parallel machines merely in terms of parallel speedup and efficiency, it is equally important to assess and optimize the single-processor performance of that code.

Although several RISC processors are currently in use in modern supercomputers, the built-in hardware performance counters of the SGI R10000 make it an ideal test bed for assessing program performance [13]. Using these counters, the present study demonstrates how single-processor performance can be improved through careful program coding. To this end, a finely tuned FVTD vector code is executed on an SGI Octane workstation and an Origin2000 (O2K) supercomputer to simulate the electromagnetic scattering from a simple sphere. A single-zone structured configuration is used, and a variety of grid resolutions are examined in order to assess the effect of problem size on program performance. Several alternative coding schemes designed to improve performance are then implemented and tested. The first of these coding schemes

involves modification to the loop control structures of the vector code. These modifications include loop reordering, loop fusion, and cache blocking [4,6]. While these modifications are designed to improve the data locality of the program, they do not change the basic *operation oriented* programming technique of the code. In this technique one operation (or a small group of operations) is performed on all cells in the computational domain, intermediate results are stored, and the next operation is applied to all cells. This continues until all operations have been performed which are required to update all cells to the next time level of the simulation. In contrast to this approach, the second alternative coding scheme is *cell oriented*. For this technique, operations are performed exclusively on a single cell of the computational domain until that cell has been updated to the next time level. Each of these approaches is discussed in detail in the following sections.

2. *Numerical Methodology*

For the present study, the two Maxwell curl equations are solved using a collocated, cell-centered, explicit FVTD scheme. In general curvilinear (ξ, η, ζ) coordinates, the equations can be written as

$$\frac{\partial \tilde{Q}}{\partial t} + \frac{\partial \tilde{E}}{\partial \xi} + \frac{\partial \tilde{F}}{\partial \eta} + \frac{\partial \tilde{G}}{\partial \zeta} = \tilde{J} \tag{1}$$

where

$$\begin{gathered} \tilde{Q} = \{B_x, B_y, B_z, D_x, D_y, D_z\}^T, \quad \tilde{J} = \{0, 0, 0, -J_x, -J_y, -J_z\}^T \\ \tilde{E} = \xi_x \vec{E} + \xi_y \vec{F} + \xi_z \vec{G}, \quad \tilde{F} = \eta_x \vec{E} + \eta_y \vec{F} + \eta_z \vec{G}, \quad \tilde{G} = \zeta_x \vec{E} + \zeta_y \vec{F} + \zeta_z \vec{G} \end{gathered} \tag{2}$$

The terms ξ_x, ξ_y, ξ_z, η_x, η_y, η_z, ζ_x, ζ_y, and ζ_z are the nine metrics of the coordinate transformation, and the vectors $\vec{E}$, $\vec{F}$, and $\vec{G}$ are the flux vectors in Cartesian coordinates which are given as

$$\begin{gathered} \vec{E} = \{0, -D_z/\varepsilon, D_y/\varepsilon, 0, B_z/\mu, -B_y/\mu\}^T, \\ \vec{F} = \{D_z/\varepsilon, 0, -D_x/\varepsilon, -B_z/\mu, 0, B_x/\mu\}^T, \quad \vec{G} = \{-D_y/\varepsilon, D_x/\varepsilon, 0, B_y/\mu, B_x/\mu, 0\}^T \end{gathered} \tag{3}$$

Integrating equation (1) over a general hexahedral volumetric element yields

$$\frac{\partial \tilde{Q}}{\partial t} + \frac{1}{V} \sum_{k=1}^{6} \tilde{R}_k \cdot \hat{n}_k A_k - \tilde{J} = 0 \tag{4}$$

where $\tilde{R} = \tilde{E}\hat{\xi} + \tilde{F}\hat{\eta} + \tilde{G}\hat{\zeta}$, $\hat{n}_k$ and A_k are the unit surface normal and surface area of cell face k, respectively, and V is the cell volume. Computation of the flux terms appearing in the summation of equation (4) is accomplished via a flux-vector-splitting scheme developed by Steger and Warming [10] in which the flux at a cell face is split into positive and negative components according to the signs of the eigenvalues of the flux Jacobian matrix. The fluxes for the faces oriented in the ξ, η, and ζ directions, respectively, are written as

$$\begin{aligned} \tilde{E}_\xi &= \tilde{E}_\xi^+(\tilde{Q}_\xi^L) + \tilde{E}_\xi^-(\tilde{Q}_\xi^R) \\ \tilde{F}_\eta &= \tilde{F}_\eta^+(\tilde{Q}_\eta^L) + \tilde{F}_\eta^-(\tilde{Q}_\eta^R) \\ \tilde{G}_\zeta &= \tilde{G}_\zeta^+(\tilde{Q}_\zeta^L) + \tilde{G}_\zeta^-(\tilde{Q}_\zeta^R) \end{aligned} \tag{5}$$

The superscripts L and R in equation (5) denote the reconstructed dependent variables at the left and right sides of the cell face, respectively. This reconstruction is performed using a *Monotone Upstream-Centered Schemes for Conservation Laws* (MUSCL) variable extrapolation technique which for an arbitrary face denoted as *i+1/2* can be written as [11]

$$\begin{aligned}\tilde{Q}^{L}_{i+1/2} &= \tilde{Q}_i + 0.25[(1-\kappa)(\tilde{Q}_i - \tilde{Q}_{i-1}) + (1+\kappa)(\tilde{Q}_{i+1} - \tilde{Q}_i)] \\ \tilde{Q}^{R}_{i+1/2} &= \tilde{Q}_i - 0.25[(1+\kappa)(\tilde{Q}_{i+1} - \tilde{Q}_i) + (1-\kappa)(\tilde{Q}_{i+2} - \tilde{Q}_{i+1})]\end{aligned} \tag{6}$$

Once the fluxes have been properly computed, the solution is integrated in time using a two-stage, second-order-accurate Runge Kutta scheme which can be summarized as follows

Stage 1:

$$\tilde{Q}^{*} = \tilde{Q}^{n} - \frac{\Delta t}{V}(\tilde{E}_2(\tilde{Q}^{n}) - \tilde{E}_1(\tilde{Q}^{n}) + \tilde{F}_4(\tilde{Q}^{n}) - \tilde{F}_3(\tilde{Q}^{n}) + \tilde{G}_6(\tilde{Q}^{n}) - \tilde{G}_5(\tilde{Q}^{n})) \tag{7}$$

Stage 2:

$$\tilde{Q}^{n+1} = \frac{1}{2}\left(\tilde{Q}^{*} + \tilde{Q}^{n} - \frac{\Delta t}{V}(\tilde{E}_2(\tilde{Q}^{*}) - \tilde{E}_1(\tilde{Q}^{*}) + \tilde{F}_4(\tilde{Q}^{*}) - \tilde{F}_3(\tilde{Q}^{*}) + \tilde{G}_6(\tilde{Q}^{*}) - \tilde{G}_5(\tilde{Q}^{*}))\right) \tag{8}$$

where it has been assumed that cell faces 1 and 2 are oriented in the $\hat{\xi}$ direction, faces 3 and 4 in the $\hat{\eta}$ direction, and faces 5 and 6 in the $\hat{\zeta}$ direction. A detailed explanation of this characteristic-based FVTD scheme can be found in references 7 and 8.

3. *Programming Paradigms*

The three programming paradigms used in the present effort have been classified by the authors for the purpose of convenience as *array based*, *optimized array based*, and *cell based*. The array-based method of programming has been shown to be highly effective for vector machines. In fact, the array-based code used in the present study (coded in FORTRAN77 and hereafter referred to as MAX3D) has demonstrated a data processing rate of 610 Mflops on a single processor of a Cray C90 [6]. The main data structure in MAX3D is the three-dimensional array, and the primary control structure is the DO loop. The code solves equations (7) and (8) by performing a series of sweeps through the computational grid. The first sweep computes all ξ face fluxes, the second computes all η face fluxes, and the third computes all ζ face fluxes using equations (5) and (6) appropriately. A simple two-dimensional sketch of the process appears in Figure 1. Unfortunately, this approach exploits neither temporal or spatial locality. For example, dependent variable data which is used in the positive flux calculation is used again in the negative flux calculation, but not before other required calculations have been performed. Thus the data has most likely been flushed from the cache before is it needed again. Similarly, depending on the direction of the data traversal through the arrays, data which is moved into the cache as a result of a cache miss may be flushed before it is actually used. For these reasons, this type of program structure, although very common, is expected to have poor cache performance.

The second type of program structure is a modification of the first. The basic data structure remains the three-dimensional array; however, the DO loop control structures are modified to improve data locality. In effect, the loops are constructed so that the sweeps are performed piece-wise through the arrays. This increases the likelihood that previously cached data will be found in cache when required. The drawback to this approach is it is highly dependent on the cache structure of the CPU, and optimizations which improve performance for one CPU may actually degrade performance on another [2]. The techniques for performing this type of modification to existing vector codes are documented in references 1 and 4. The optimized array-based code used in the present study is hereafter referred to as MAX3DO.

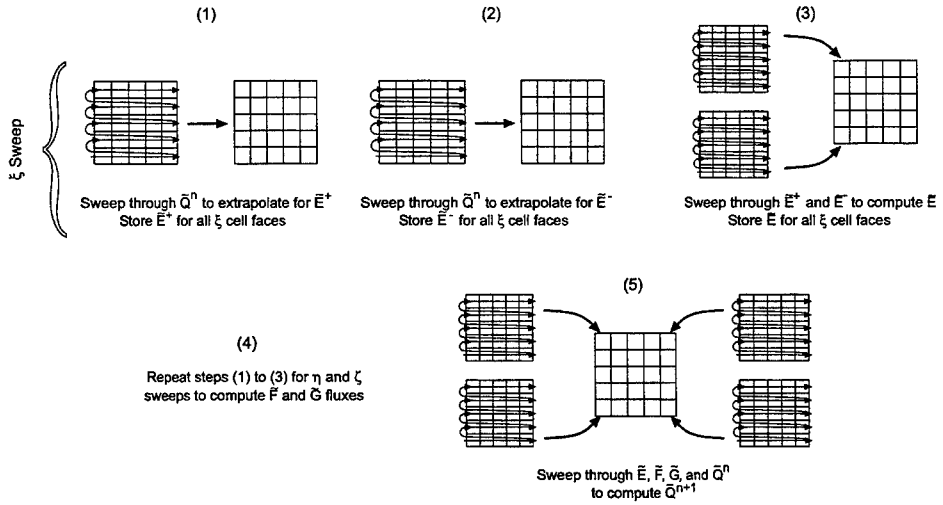


Figure 1: Array-based coding procedure

The third programming technique is fundamentally different than the first two. Instead of performing a series of sweeps through large arrays and storing intermediate results, the code computes the fluxes for all six faces of a given cell using equations (5) and (6), and then updates that cell using equation (7) or (8). This process is then repeated for each cell in the computational domain. A two-dimensional example of this process is depicted in Figure 2. Because all calculations are performed on a single cell, a high degree of temporal locality is achieved. Furthermore, because the six dependent variables for each cell are stored contiguously in memory, the code exhibits good spatial locality as well. This cell-based computer code was developed in C, and is hereafter referred to as CHARGE.

4. *Testing Procedure*

Single-processor performance of all computer codes was assessed on an SGI Octane workstation and the O2K supercomputer. The Octane had two 195 Mhz R10000 processors, each with 1 megabyte of secondary (L2) cache, and access to 1 gigabyte of shared memory. The Origin2000, on the other hand, was a 32-processor distributed-shared memory machine. Each processor had a 4 megabyte L2 cache, and the machine had a total of 16 gigabytes of memory. Both machines utilized the IRIX version 6.4 operating system.

Performance metrics were obtained using the SGI-supplied *perfex* utility in conjunction with the built-in hardware performance counters of the R10000 processor [4,13]. Although an exhaustive study of compiler optimizations was not conducted, several options were examined in order to attempt to maximize performance of each of the three codes. For CHARGE, the options which yielded the best performance were *-O3 -INLINE:must=<fcn list> -OPT:IEEE_arithmetic=3 -64* where *<fcn list>* was the list of functions to be inlined by the compiler. Because the performance counters count the multiply-add instruction of the R10000 as a single floating-point operation, a more accurate Mflop performance assessment was obtained by adding the compiler option *-TARG:MADD=off* to disable the combined multiply-add instruction. Adding this option was found to degrade the speed of the code only slightly. Compilation of MAX3D and MAX3DO was performed using the options *-O3 -r8 -OPT:IEEE_arithmetic=3 -TARG:MADD=off*. Note that the double precision flag was not used

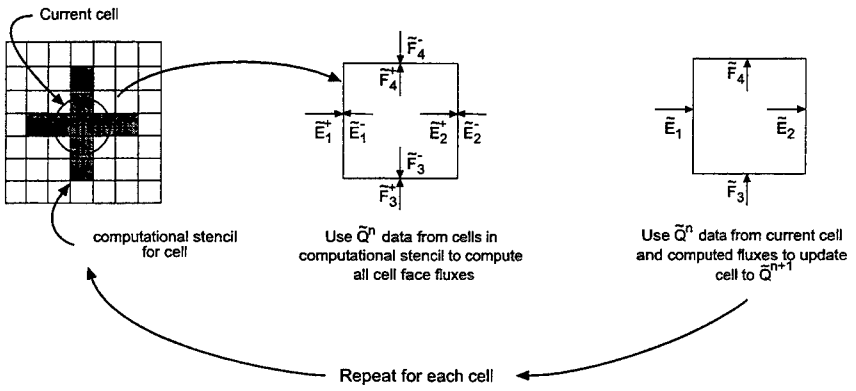


Figure 2: Cell-based coding procedure

for CHARGE since the variables are explicitly declared as type *double* within the code. In order to avoid cache thrashing, care was taken to ensure that no array dimensions nor any product of array dimensions within MAX3D were a power of two.

The perfex utility was invoked using the command *perfex -a -y <executable name>*. The *-a* option enables statistical sampling of all counters. In order to ensure accuracy of the reported results, all timing runs were repeated at least five times, and the best performance was reported. In practice, most performance metrics were found to vary by less than 3 percent. The metrics which are reported here include Mflops, primary (L1) cache hit rate, secondary L2 cache hit rate, and L1 cache line reuse. These metrics are defined by SGI in their on-line man pages as follows:

Mflops:	graduated floating point instructions / program run time
L1 cache hit rate:	1.0 - (L1 cache misses / (graduated loads + graduated stores))
L2 cache hit rate:	1.0 - (L2 cache misses / L1 cache misses)
L1 cache line reuse:	(graduated loads + graduated stores - L1 cache misses) / L1 cache misses

All performance results presented here were obtained on dedicated machines so that timing was completely unbiased by fluctuations in machine loading.

5. *Results*

Before the performance of the three codes was examined, the accuracy of each code was assessed by comparing the computed result of the bistatic radar cross section of the sphere against the analytical Mie series solution [5]. A sample result is depicted in Figure 3 for the case of $ka = 10.47$ where a is the radius of the sphere and k is the wavenumber of the incident field. All three codes produced identical results which agree very well with the theoretical solution.

The floating-point performance of each of the codes is presented in Figure 4. MAX3D fits entirely in cache for problem sizes on the order of a few thousand cells. In this scenario, the code performs well at between roughly 110 and 115 Mflops on both the Octane and the Origin. As the problem size increases, however, and the problem no longer fits entirely in cache, the code suffers a dramatic performance degradation. Furthermore, the drop in performance occurs quite suddenly. For

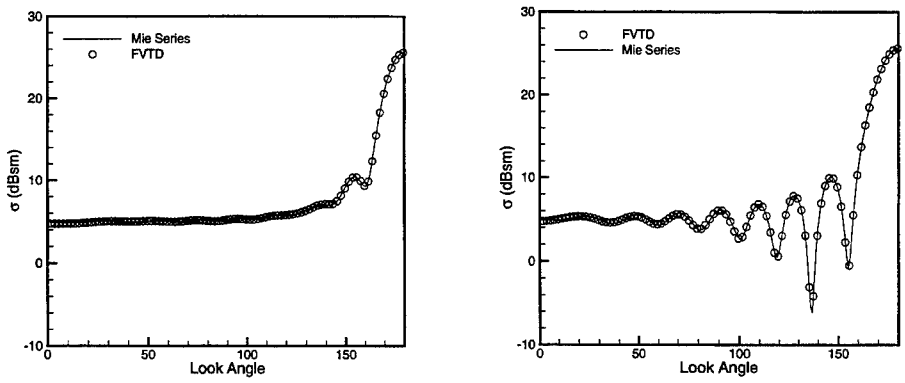


Figure 3: Comparison of computed and theoretical radar cross section of perfectly electrically conducting sphere: L) VV polarization, R) HH polarization

example, on the Octane, code performance drops from approximately 115 Mflops at 2000 cells to less than 50 Mflops at 8000 cells. At 64,000 cells, the code performance has dropped to 30.8 Mflops, only 27 percent of it's peak in-cache performance level. Performance of the code on the Origin is similar with the notable exception that the performance drop-off point is delayed due to the increased size of the secondary cache. The larger cache, however, ultimately does not improve the performance for large problems, and the performance drops to only 27 Mflops for a problem size of 512,000 cells. It should be noted that for a typical three-dimensional problem, cell counts of a million or more cells are not uncommon. It is therefore clear than the standard array-based coding scheme does not provide acceptable performance levels on the R10000.

The abysmal performance of the standard array code is improved dramatically by loop optimizations as shown in Figure 4. Although the performance still decreases as the problem size increases, the reduction is not nearly as dramatic with the code still achieving approximately 80 Mflops for the largest problem size. This represents a three-fold improvement over the unoptimized code. Although the performance of the code has been improved dramatically, the code still exhibits a dependency (although diminished) on the size of the cache. This dependency on cache size can be frustrating when attempting to optimize a code to run on different processor architectures.

The most predictable performance of the three codes was demonstrated by CHARGE. Although not achieving quite the performance of the other codes for the extremely small problems, it showed very little performance degradation with increasing problem size. In fact, performance between the largest and smallest problems varied by only approximately 6 percent. Furthermore, any differences in code performance on the Origin and the workstation were almost negligible. Thus, CHARGE exhibits virtually no dependence on processor cache size. Although additional performance analysis has yet to be completed, it is believed that the code will perform similarly for machines such as the Cray T3E and IBM SP which have caches on the order of a few hundred kilobytes. This belief is supported by examining the cache performance of the three codes more closely.

The data for the L2 cache hit rate of the three codes is presented in Figure 5. The extremely poor performance of MAX3D is clearly evident in the figure. It is apparent from the figure that the data fits entirely into a 1MB cache up to problem sizes of approximately 2000 cells and into a 4MB cache up to a problem size of roughly 8000 cells for both array-based codes. Once the data spills out of cache, however, MAX3D shows a rapid cache-hit-rate reduction to approximately 50 percent on both the Octane and the Origin. This is far below the desired 95 percent hit rate often quoted for good performance

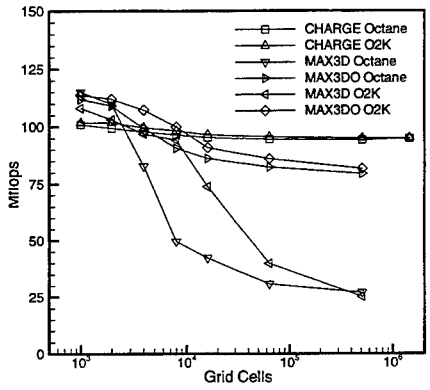


Figure 4: Floating-point performance as a function of problem size

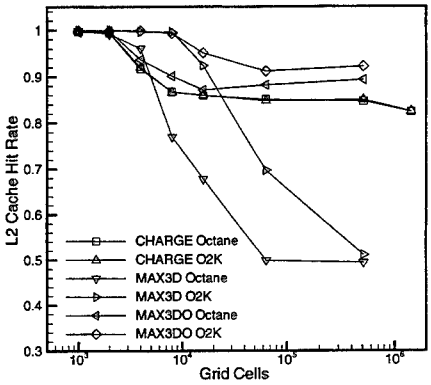


Figure 5: Secondary data cache hit rates

[1]. It is apparent that increasing the size of the cache for the unoptimized code simply delays the onset of the performance reduction and does not change the performance of the code for realistically sized problems.

Although the cache-hit performance of the unoptimized code seems to directly reflect the data in Figure 1, similar conclusions cannot be drawn when comparing the L2 cache hit rates of CHARGE and MAX3DO. Here, MAX3DO outperforms CHARGE for all problem sizes. This would seem to be in contrast with the Mflop results contained in Figure 4; however, the differences can be resolved by examining the L1 cache hit rates of the three codes as presented in Figure 3. Because the L1 cache performance of the codes was nearly identical for both the Octane and the Origin, only the Origin data is presented in the figure. Again, MAX3D exhibits relatively poor performance with an L1 cache hit rate varying between approximately 0.68 and 0.77. MAX3DO performs better with hit rates varying between 0.86 and 0.89. On the other hand, CHARGE exhibits extremely good L1 cache access with a nearly constant 0.99 cache hit rate. This indicates a very high degree of data locality since most accesses can be satisfied by the on-CPU 32K cache, a situation even more desirable than the data residing in L2 cache. This clearly indicates the high degree of data locality in the code, and reinforces the belief that the code should perform similarly on machines having very small caches. A final substantiation of this claim comes from examining the data in Figure 4 which shows the L1 cache line reuse. This metric measures the number of times a piece of data, on average, is used once brought into L1 cache. The data shows that CHARGE utilizes L1 cache data between 71 and 113 times before it is flushed from the cache. Contrast this to the reuse rates of roughly 2 to 3 times for MAX3D and 6 to 9 times for MAX3DO, and it becomes clear why the floating-point performance of CHARGE remains essentially constant over a wide range of problem sizes. This is especially promising given the fact that no processor-specific optimizations were made during code development.

6. *Concluding Remarks*

Three separate FVTD computer codes developed at the Air Force Research Laboratory have been analyzed for performance on the SGI R10000 processor. The traditional vector style of programming was found to underperform two other programming styles more tailored to data locality. The vector style of programming demonstrated a dramatic reduction in performance as the problem size increased. In contrast, the code using a cell-based programming approach was found to have extremely good performance across a large range of problem sizes. This approach also demonstrated virtually no dependence on the processor's cache size. This makes the approach attractive from the perspective of developing a simulation environment which is capable of achieving high levels of performance on a variety of architectures. It is precisely this type of flexibility that is a fundamental requirement for developing a useful tool for conducting complex time-domain electromagnetic simulations.

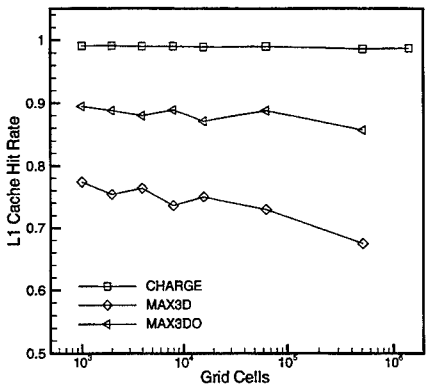


Figure 6: Primary data cache hit rates

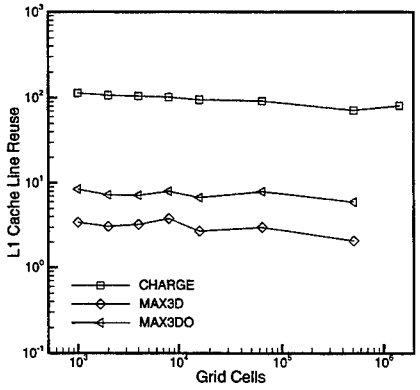


Figure 7: Primary data cache item reuse

Acknowledgments

The Origin2000 program runs were performed on the Aeronautical Systems Center Major Shared Resource Center machine. Our thanks go to Dr. C.J. Suchyta for his help with our many Origin-related questions. In addition, Dr. Yi Pan of the University of Dayton performed the loop modifications on MAX3D to produce MAX3DO. His work is greatly appreciated.

References

1) Dowd, Kevin, *High Performance Computing*, O'Reilly & Associates, Inc., 1993.

2) NAS News, NAS System Division, Ames Research Center, Vol. 2., No. 25, Jul-Aug 1997.

3) NAS News, NAS System Division, Ames Research Center, Vol. 2, No. 27, Nov-Dec 1997.

4) Origin2000 Application Development & Optimization, Student Handbook, Part Number: O2KAPPL-0.7-6.2/3/4-S-SD-W, Silicon Graphics Computer Systems, December 1997.

5) Ruck, G. T., Barrick, D. E., Stuart, W. D. and Kirchbaum, C. K., *Radar Cross Section Handbook, Volume I*, Plenum Press, 1970.

6) Shang, J. S., Wagner, M., Pan, Y. and Blake, D. C., "Strategies for Time-Domain CEM Computations on Multicomputers," AIAA Paper 98-0979.

7) Shang, J. S. "Characteristic-Based Algorithms for Solving Maxwell's Equations in the Time-Domain," *IEEE Antennas and Propagation Magazine*, Vol. 37, No. 3, pp. 15-25, June 1995.

8) Shang, J. S., "Characteristic-Based Methods in Computational Electromagnetics", appearing in *Computational Electromagnetics and Its Applications*, Campbell, T. G., Nicolaides, R. A. and Salas, M. D., editors, ICASE/LaRC Interdisciplinary Series in Science and Engineering, Kluwer Academic Publishers, 1997.

9) Shankar, V., Hall, W. F., Mohammadian, A. and Rowell, C., "Computational Electromagnetics: Development of a Finite-Volume, Time-Domain Solver for Maxwell's Equations," Final Contract Report of Contract N62269-90-C-0257, May 1993.

10) Steger, J. L. and Warming, R. F., "Flux Vector Splitting of the Inviscid Gasdynamics Equations with Application to Finite Difference Methods," Journal of Computational Physics, Vol. 40, pp. 263-93, April 1981.

11) van Leer, B., "Flux-Vector Splitting for the Euler Equations," Technical Report 82-30, ICASE, September 1983.

12) Wang G. and Tafti, D. K., "Performance enhancement on microprocessors with hierarchical memory systems for solving large sparse linear systems, submitted to *International Journal of Supercomputing Applications*, Feb. 1997.

13) Zagha, M., Larson, B., Turner, S. and Itzkowitz, M., "Performance analysis using the mips r10000 performance counters," *Proceedings of Supercomputing '96*, IEEE Computer Society.

The *VOLMAX* Transient Electromagnetic Modeling System, Including Sub-Cell Slots and Wires on Random Non-Orthogonal Cells

Douglas J. Riley and C. David Turner

Radiation and Electromagnetic Analysis Department
Sandia National Laboratories, Albuquerque, New Mexico 87185-1166
Internet: djriley@sandia.gov

1. Introduction

VOLMAX is a three-dimensional transient *vol*umetric *Max*well equation solver that operates on standard rectilinear finite-difference time-domain (FDTD) grids, non-orthogonal unstructured grids, or a combination of both types (hybrid grids) [1-3]. The algorithm is fully explicit. Open geometries are typically solved by embedding multiple unstructured regions into a simple rectilinear FDTD mesh. The grid types are fully connected at the mesh interfaces without the need for complex spatial interpolation. The approach permits detailed modeling of complex geometry while mitigating the large cell count typical of non-orthogonal cells such as tetrahedral elements. To further improve efficiency, the unstructured region carries a separate time step that sub-cycles relative to the time-step used in the FDTD mesh. A cross section of the interface between finite-volume time-domain (FVTD) and FDTD grids is shown in Fig. 1. The "wrapper layer" is a hexahedral region that encloses the unstructured grid and provides nodal connectivity to the surrounding FDTD mesh. The wrapper is constructed automatically based on the unstructured-grid topology. The unstructured region may consist of a single rectangular block, or be of a multiple, block-on-block form.

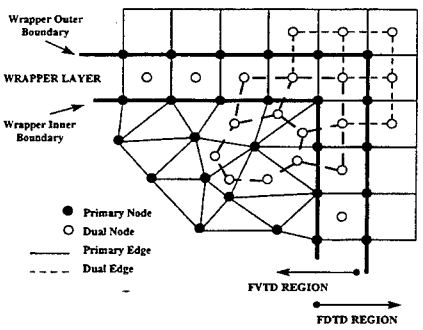


Fig. 1. The hybrid grid interface.

As shown in Fig. 1, *VOLMAX* is based on a staggered grid formulation. Primary and dual grids are used. When the unstructured grid consists exclusively of rectangular hexahedral cells, the field advancement is identically FDTD in nature, although the cells are referenced in an unstructured (indirect) manner. Note that the wrapper layer consists of rectangular cells for its primary grid, but the dual cells on the wrapper inner boundary are generally non-orthogonal. As a consequence, the wrapper layer is common to both the FVTD and FDTD grids. For the case that the unstructured-grid consists of uniform rectangular elements, the algorithm is second-order accurate both in space and time.

The field advancement scheme for the *VOLMAX* hybrid mesh is the following. The electric fields in the FDTD region are initially advanced based on time step, Δt_s. On the outer boundary of the wrapper, the tangential electric fields are second-order time interpolated to provide a Dirichlet boundary condition for the FVTD region. The electric and magnetic fields in the FVTD region are advanced an integral number of sub-time iterations relative to Δt_s. At the completion of the sub-cycling, the tangential electric fields on the inner boundary of the wrapper are used to provide a Dirichlet boundary condition to complete the magnetic-field advancement in the FDTD region. An alternative scheme could map the magnetic fields in the wrapper layer into the respective FDTD locations after the FDTD magnetic fields are advanced in time.

VOLMAX is currently integrated to the commercial CAD package SDRC I-DEAS [4]. Solid model design, mesh generation, and post-processing are all accomplished through the I-DEAS interface. Electromagnetic properties, such as voltage sources, local boundary conditions, current observers, input and output ports, slots, wires, etc., are implemented by assigning nodal attributes to the desired property. The original I-DEAS grid file is input into the *VOLMAX* preprocessor, *PreVol*, which builds the wrapper layer, and the primary and dual grids. Grid construction by *PreVol* is accomplished at the rate of 50,000 to 100,000 cells/minute on a single, high-end processor. Construction time scales

linearly with cell count. The basic user interface for *PreVol* in shown in Fig. 2. Typical inputs include the simulation domain (interior/exterior), node attributes (sources, observers, etc.), and (optionally) the topology of the unstructured region(s).

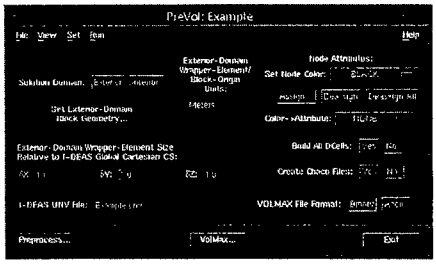


Fig. 2. The basic *PreVol* interface.

The overall design and simulation procedure used in the *VOLMAX* system is outlined in Fig. 3. The closed loop permits an adaptive cycle based on simulation results.

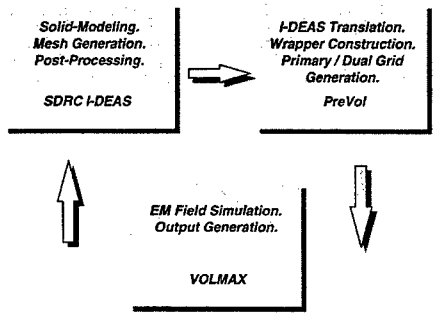


Fig. 3. The simulation cycle.

For demonstration purposes, application of *VOLMAX* is made to a cylindrical resonator and scattering by a simple conducting sphere in Section 2 of the paper. In Section 3, two methods for modeling sub-cell wires on arbitrary non-orthogonal cells are introduced. In Section 4, a generalization of the hybrid thin-slot algorithm (HTSA [5]) to arbitrary cell types is also introduced. EMC/EMI applications are made in Section 5. Concluding remarks are made in Section 6.

2. Application to Canonical Geometries

The hybrid-grid, far back-scattered field from a 0.5 m radius, perfectly conducting sphere gridded with tetrahedra and embedded in FDTD hexahedra is shown in Fig. 4. Note the good agreement with the Mie-series solution even as the resolution of the external FDTD mesh falls below 10 cells/wavelength (λ). A contour rendering of the surface current-density shortly after a Gaussian pulse has hit the sphere is shown in Fig. 5.

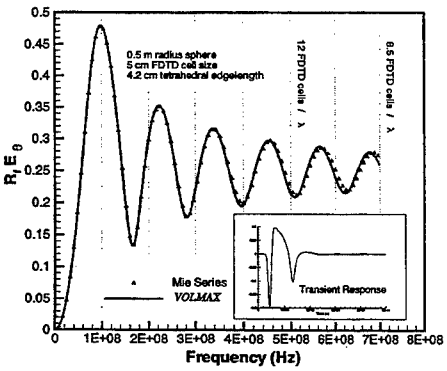


Fig. 4. The far, back-scattered field from a 0.5 m conducting sphere. R_f denotes distance. Hybrid-grid solution. The transient response is inset.



Fig. 5. The early time surface current density on a conducting sphere.

An extruded hexahedral element mesh for a simple cylindrical resonator is shown in Fig. 6. Random edges were selected for the source and observer. A Gaussian pulse excitation was used. The internal transient response demonstrating stability is shown in Fig. 7. The first few TM resonances are shown in Table 1.

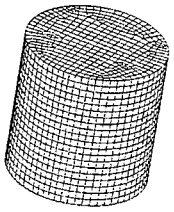


Fig. 6. Cylindrical resonator with an average hexahedral edgelength of 5 cm. The radius is 0.5 m and the height is 1 m.

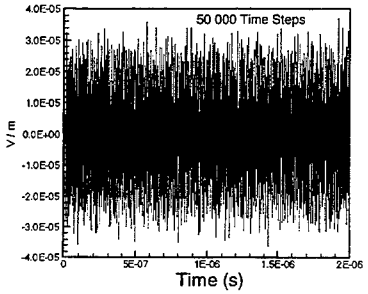


Fig. 7. Internal electric field after 50,000 time steps.

TABLE 1. Resonances of Cylindrical Resonator

Mode	Theory (MHz)	VOLMAX (MHz)
TM011	274.12	274.01
TM012	377.56	376.48
TM111	395.21	393.93
TM112	472.86	471.35
TM013	504.87	500.95

3. Sub-Cell Wire Modeling

The ability to model features that are small relative to the global cell size is important in electromagnetic simulations. By tapering an unstructured mesh, it is possible to resolve small detail; however, the increase in cell count and the reduction in time step can be prohibitive.

Relatively simple algorithms to resolve small wires on rectangular FDTD grids have been developed [6,7]. The algorithms are accurate but require that the wire conforms to the rectangular mesh. This can create problems for applications such as cellular phones that may demand the phone model to be tilted relative to the human head model.

Two algorithms are briefly presented here that enable wires to run arbitrarily along edges of an unstructured mesh. The first method embeds a transient integral equation into the unstructured mesh, whereas the second method is a generalization of the original FDTD scheme to non-orthogonal cells. A similar extension of the FDTD scheme was presented in [8], but the method was only applied to linear wires on prismatic cells. The technique in the present paper further extends and applies the method to curved wires on tetrahedral meshes.

3.1 Integral Equation Thin-Wire Model

A transient integral equation (IE) is used to model the topology of the wire. The wire is defined in the original solid model and is meshed using one-dimensional beam elements. Within *VOLMAX*, the IE operates in one of two modes. The first mode is an exclusive wire mode that is coupled to a *free-space* volumetric mesh. In this mode, *VOLMAX* is similar to a transient version of the frequency-domain *NEC* [9] code, with the added benefit of field visualization into the volumetric region. In the second mode, the IE operates in a field-feedback configuration that enables solid geometry to reside in the unstructured mesh. This algorithm is similar to the hybrid thin-slot algorithm [5] in that local vector fields computed in the volumetric region are injected back into the IE at each time step. These fields correspond to reflections from non-wire geometry and represent additional sources driving the IE. The field-feedback mode has been found to be most effective for *free* wires defined on hexahedral cells, and for wire radii that are a small fraction of the surrounding edge lengths that support the wire.

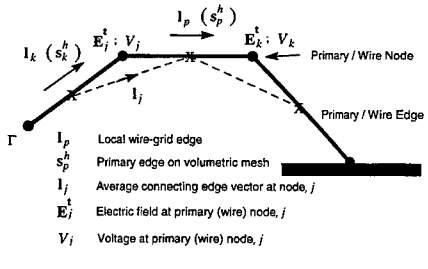


Fig. 8. Relationship of local wire-grid to volumetric primary mesh. Dual cells (not shown) enclose primary nodes. Γ is the wire path.

A section of a simple curved wire is shown in Fig. 8. The IE solution uses overlapping piecewise-linear basis functions that are centered at the nodal positions. Only the governing equations are presented here. Numerical solution details for the IE are similar to Refs. [10,11].

The governing integral equation for the wire system, including the provision for volumetric-mesh feedback, is the following [cf. [11] for the free space case].

$$\varepsilon_o \mathbf{l}\cdot\frac{\partial}{\partial t}\left(\mathbf{E}^{inc}+v\mathbf{E}^{\mathbf{t}}\right)=$$
$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\int_{\Gamma} dl'\,\mathbf{l}\cdot\mathbf{l}'\,I\left(l',\tau\right)\left[G\left(\mathbf{r},\mathbf{r}';a\right)-vG\left(\mathbf{r},\mathbf{r}';a_o\right)\right]-$$
$$\mathbf{l}\cdot\nabla\int_{\Gamma} dl'\,\nabla_{l'}\cdot\left[\mathbf{l}'\,I\left(l',\tau\right)\right]\left[G\left(\mathbf{r},\mathbf{r}';a\right)-vG\left(\mathbf{r},\mathbf{r}';a_o\right)\right] \quad (1)$$

where $\mathbf{r}\in\Gamma$, with Γ the wire path, $\tau=t-\left|\mathbf{r}\text{-}\mathbf{r}'\right|/c$, c denotes the speed of light in vacuum, I denotes the current on the wire, $\mathbf{E}^{inc}$ denotes an impressed source on the wire, and $\mathbf{E}^{\mathbf{t}}$ denotes an average vector electric field from the volumetric grid local to the wire. $v=0$ sets the equation to operate in a free-space (no feedback) mode, and $v=1$ sets the equation to operate in a feedback mode from the volumetric grid. The free-space Green's function is denoted by $G()$ [11], a denotes the wire radius, and a_o denotes an effective radius for matching the integral equation solution to the volumetric solution local to the wire. Note that the volumetric solution for the electric field on the wire will not be identically zero because the solution represents an average value for the electric field over the dual cell containing the wire node; consequently, a_o is typically taken to be ½ the local dual-cell diameter.

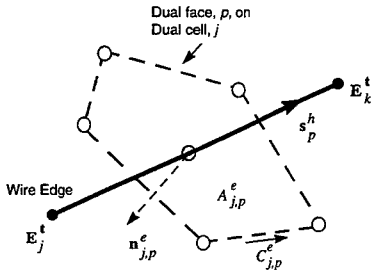


Fig. 9. Wire edge piercing dual face.

The integral equation solves for the current at the wire (primary) nodes. Coupling to the volumetric grid requires the wire current to be defined on primary edges. Let the average wire current on the *p-th* primary edge be denoted by $\tilde{I}_p$. Coupling to the volumetric grid is then approximated through the equation (cf. Fig. 9)

$$\varepsilon_o\frac{\partial}{\partial t}\int_{A_{j,p}^e} E_p\,dA_p=\oint_{C_{j,p}^e}\tilde{\mathbf{H}}^{\mathbf{t}}\cdot d\mathbf{l}-\tilde{I}_p\frac{\mathbf{s}_p^h\cdot\mathbf{n}_{j,p}^e}{\left|\mathbf{s}_p^h\cdot\mathbf{n}_{j,p}^e\right|} \quad (2)$$

To ensure stability, the time-averaging scheme introduced in [1] is applied to the time-integration used for Eq. (2). The spatial integration is over the dual face pierced by primary edge, $\mathbf{s}_p^h$. The normal to this face is denoted by $\mathbf{n}_{j,p}^e$, and the face area is denoted by $A_{j,p}^e$. E_p represents the electric field normal to the dual face, while $\tilde{\mathbf{H}}^{\mathbf{t}}$ denotes average magnetic fields on the dual edges enclosing the face. A more detailed discussion of the grid topology can be found in [1].

The vector electric fields at primary nodes, $\mathbf{E}_j^{\mathbf{t}}$, are approximated using a least-squares fit to the face-normal electric fields (E_p). The average electric field projected in the primary edge direction is defined by

$$\tilde{\mathbf{E}}^{\mathbf{t}}\cdot\mathbf{s}_p^h=\frac{1}{2}\left(\mathbf{E}_j^{\mathbf{t}}+\mathbf{E}_k^{\mathbf{t}}\right)\cdot\mathbf{s}_p^h+$$
$$\left[E_p-\frac{1}{2}\left(\mathbf{E}_j^{\mathbf{t}}+\mathbf{E}_k^{\mathbf{t}}\right)\cdot\mathbf{n}_{j,p}^h\right]\left(\mathbf{n}_{j,p}^h\cdot\mathbf{s}_p^h\right) \quad (3)$$

The integral-equation technique is demonstrated by examining scattering by three curved wires in free space. The simulation is performed two times. In the first case, $v=0$ in Eq. (1) is used, whereas in the second case, $v=1$. Because the geometry involves only wires, the results of the two simulations should be identical. A contour plot for the electric-field distribution local to the wires is shown in Fig. 10. A Gaussian pulse is incident normal to the plane containing the wires. The far, back-scattered field comparing the two simulations is shown in Fig. 11. The wires were locally encapsulated in skewed hexahedral elements that were embedded in tetrahedra. The unstructured-grid block was then embedded in a cubical FDTD mesh out to the grid termination using 5-cm cells.

3.2 Partial Differential Equation Thin-Wire Model

Using a partial differential equation (PDE) model, or equivalently, a transmission-line (TL) model, the wire electric current is defined on primary edges, while the voltage (or charge) is defined at primary nodes. This formulation has a more natural correlation with an FDTD or FVTD volumetric grid than the IE method, and facilitates the connection of wires to solid geometry. In both models, wires are defined using one-dimensional beam elements.

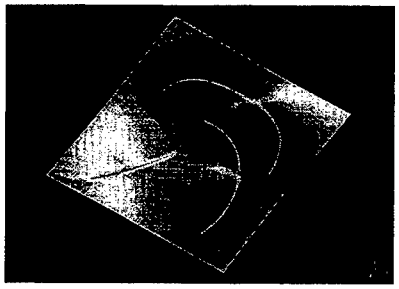


Fig. 10. Scattered electric field surrounding three wires. The wire radius was 2.5 mm and the average edge length was 5 cm. An FDTD grid encloses the unstructured grid.

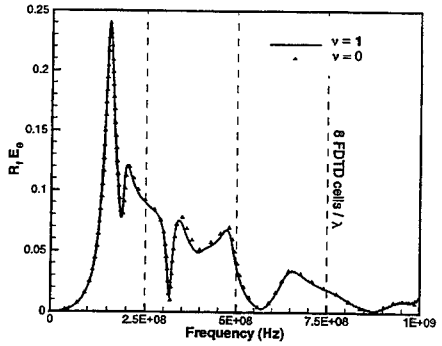


Fig. 11. Normalized, far, back-scattered field from the wire system with feedback on (v=1), and off (v=0). R_f denotes distance.

The governing equations along an arbitrary path defined by the spatial variable, l, are the following [cf. 6,7 for an FDTD implementation]:

$$\frac{\partial}{\partial l} I = -C_w \frac{\partial V}{\partial t} \qquad (4a)$$

$$\frac{\partial}{\partial l} V = -L_w \frac{\partial I}{\partial t} + \mathbf{E}^t \cdot \mathbf{s} + V^{inc} - I R \qquad (4b)$$

I represents current while V denotes voltage. $V = 0$ when the wire terminates on a conductor, whereas $I = 0$ at an open-end termination. The "in-cell" capacitance and inductance are denoted by C_w and L_w, respectively.

With reference to Fig. 8, an explicit algorithm is

$$V_j^{n+1} = V_j^n - \frac{\Delta t_u}{C_{w_j} \left| \mathbf{l}_j \right|} \left(I_p^{n+1/2} \xi_{j,p} - I_k^{n+1/2} \xi_{j,k} \right) \qquad (5a)$$

$$I_p^{n+3/2} = \frac{B_p}{A_p} I_p^{n+1/2} - \frac{1}{A_p} \left[\left(V_k^{n+1} - V_j^n \right) - \tilde{\mathbf{E}}^t \cdot \mathbf{s}_p^h - V_p^{inc} \right] \qquad (5b)$$

$$\xi_{j,p} = \frac{\mathbf{l}_j \cdot \mathbf{s}_p^h}{\left| \mathbf{l}_j \cdot \mathbf{s}_p^h \right|}; \quad \xi_{j,k} = \frac{\mathbf{l}_j \cdot \mathbf{s}_k^h}{\left| \mathbf{l}_j \cdot \mathbf{s}_k^h \right|}$$

$$A_p = \frac{\left| \mathbf{s}_p^h \right| L_{w_p}}{\Delta t_u} + \frac{1}{2} R_p; \quad B_p = \frac{\left| \mathbf{s}_p^h \right| L_{w_p}}{\Delta t_u} - \frac{1}{2} R_p$$

$$C_{w_j} = \frac{2\pi \tilde{\varepsilon}_j}{ln\left(\zeta_j / a\right)}; \quad L_{w_p} = \frac{1}{2} \frac{\left(\tilde{\mu}_k + \tilde{\mu}_j \right) \left(\tilde{\varepsilon}_k + \tilde{\varepsilon}_j \right)}{\left(C_{w_k} + C_{w_j} \right)}$$

V_p^{inc} and R_p denote an impressed voltage source and resistance, respectively, on the *p-th* primary edge. $\tilde{\mu}_j$ and $\tilde{\varepsilon}_j$ denote average permeability and permittivity, respectively, at the *j-th* primary node. a is the wire radius, and Δt_u is the time-step in the unstructured mesh. A superscript, n, denotes time iteration.

ζ_j and $\tilde{\mathbf{E}}^t \cdot \mathbf{s}_p^h$ represent critical quantities that determine the accuracy the of PDE thin-wire method on a random unstructured mesh. ζ_j represents an average distance between the *j-th* primary (wire) node and the (non-wire) nodes that locally surround it. $\tilde{\mathbf{E}}^t \cdot \mathbf{s}_p^h$ represents an average of the non-wire-node vector electric fields surrounding the endpoints of the *p-th* primary edge, projected onto this edge. Figure 12 shows a two-dimensional representation of the geometry.

ζ_j and $\tilde{\mathbf{E}}^t \cdot \mathbf{s}_p^h$ are computed as follows (cf. Fig. 12).

$$\zeta_j = \frac{1}{m} \sum_i s_{j,i}$$

$$\tilde{\mathbf{E}}^t \cdot \mathbf{s}_p^h = \frac{1}{2} \left(\frac{1}{m} \sum_i \mathbf{E}_{j,i}^t + \frac{1}{n} \sum_i \mathbf{E}_{k,i}^t \right) \cdot \mathbf{s}_p^h$$

The summations are taken over the valid primary nodes or edges that support the wire node. A nearly uniform nodal distribution with constant ζ_j (taken at the source) has been found, to date, to provide the best accuracy.

The wire current at each time iteration is obtained by solving Eqs. (5a, b). The current is injected onto the volumetric grid in a manner similar to Eq. (2).

For demonstration, the input admittance of a wire-loop antenna defined on a tetrahedral mesh is examined. The loop diameter is 15 cm and the wire diameter (2 *a*) is 0.5 mm. The loop is modeled on tetrahedral elements with an average edge length of 1.08 cm. The unstructured mesh is embedded in a uniform FDTD mesh with 1-cm cubical elements. A Gaussian-modulated sinusoidal voltage source is impressed on the wire. The planar nodal distribution surrounding the beam elements used to mesh the wire is shown in Fig. 13. The transient driving-point current using the PDE thin-wire algorithm is shown in Fig. 14. A comparison is made with the previous IE thin-wire algorithm for the case $\upsilon = 0$ (no feedback). As seen, the results are virtually identical. The input admittance is shown in Fig. 15.

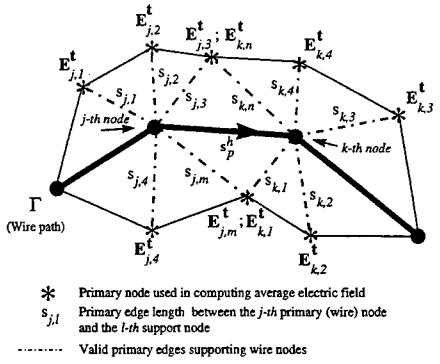


Fig. 12. Primary edges and nodes used in computing average electric fields and edge lengths supporting wire nodes.

4. Sub-Cell Slot Modeling

Several algorithms have been proposed to model narrow apertures on rectangular FDTD grids [5,6,10,12]. However, none of the algorithms have been extended to unstructured grids with non-orthogonal cells. Such an extension is made in this section for the hybrid thin-slot algorithm (HTSA) [5]. The HTSA uses a transient integral equation to model the slot physics.

Similar to the IE thin-wire algorithm described in Section 3.1, the HTSA also uses a field-feedback technique to account for the presence of solid geometry in the neighborhood of the slot. The original algorithm for linear apertures has been shown to be accurate, but long-term stability is dependent on the implementation and the equivalent wire radius [10,13,14]. The generalized HTSA presented in this section improves on stability issues while permitting slots to follow an arbitrary path on a *locally* planar region. The requirement of local planarity is a result of applying the equivalence principle [15] in conjunction with the free-space Green's function.

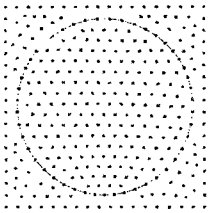


Fig. 13. The nodal distribution in the loop plane of the unstructured grid.

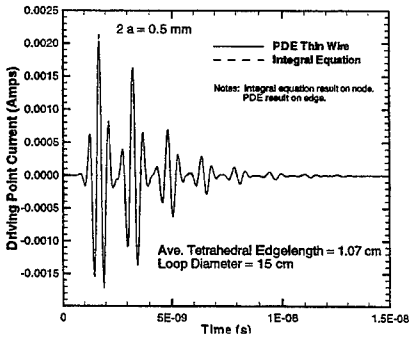


Fig. 14. The loop transient driving-point current.

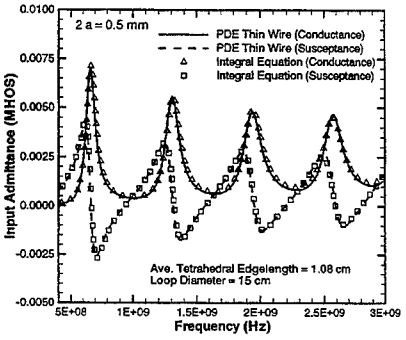


Fig. 15. The loop input admittance.

A slot in a perfectly conducting plane is shown in Fig. 16. Fields are assumed to be incident from both region 1 and region 2. A derivation of an IE for the magnetic current can be found in [10,11]. The HTSA generalizes the standard slot IE by utilizing the total magnetic field from the volumetric grid as a source for the IE. This field includes not only the usual "short-circuit" terms required by the standard IE, but also includes the slot radiation and any additional scattered fields due to finite geometry. The technique is particularly well suited to FDTD, or FVTD formulations that use interleaved grids. The resulting equation is given by [10]

$$\frac{\mu_o}{4}\mathbf{l}\cdot\frac{\partial}{\partial t}\mathbf{H}^{t_1}-\frac{\mu_o}{4}\mathbf{l}\cdot\frac{\partial}{\partial t}\mathbf{H}^{t_2}=$$
$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\int_{\Gamma} dl'\,\mathbf{l}\cdot\mathbf{l}'\,K(l',\tau)\left[G(\mathbf{r},\mathbf{r}';a)-G(\mathbf{r},\mathbf{r}';a_o)\right]-$$
$$\mathbf{l}\cdot\nabla\int_{\Gamma} dl'\,\nabla_{l'}\cdot\left[\mathbf{l}'\,K(l',\tau)\right]\left[G(\mathbf{r},\mathbf{r}';a)-G(\mathbf{r},\mathbf{r}';a_o)\right] \quad (6)$$

where $\mathbf{r}\in\Gamma$, K denotes the magnetic current, and $G()$ denotes the free-space Green's function. The equivalent thin-wire radius, a, for the thin slot is $a=(w/4)\exp\left[-\pi d/(2w)\right]$ [16], where w denotes the slot width and d denotes the slot depth. The total magnetic fields in region 1 and region 2 of the slot plane are denoted by $\mathbf{H}^{t_1}$, and $\mathbf{H}^{t_2}$, respectively. Local to the slot, they are computed by averaging over the vector magnetic fields located at the dual nodes that surround appropriate dual faces in region 1 and region 2. a_o is defined to be an average distance from the slot to the surrounding local magnetic field locations (dual node locations). Other parameters are as defined for Eq. (1). Numerical solution details for the IE can be found in [10,11].

Faraday's law is used to apply the magnetic current onto the volumetric grid. Only the primary faces that have a single edge on the slot plane, and only a single slot node, are used with the appended magnetic current (cf. Fig. 16). For the *l-th* face on the *i-th* primary cell,

$$\mu_o\frac{\partial}{\partial t}\int_{A^h_{i,l}} H_l\,dA_i=-\oint_{C^h_{i,l}}\tilde{\mathbf{E}}^t\cdot d\mathbf{l}\pm\frac{K_j}{2}\frac{\mathbf{l}_j\cdot\mathbf{n}^h_{i,l}}{\left|\mathbf{l}_j\cdot\mathbf{n}^h_{i,l}\right|} \quad (7)$$

The "+" sign is for region 1, whereas the "-" sign is for region 2. Note the ½ scaling factor applied to the magnetic current. This is because the slot is defined to lie along primary edges. Thus, the contribution due to the slot is apportioned to the primary faces that lie "above" and "below" the aperture. This is a distinction relative to previous thin-slot algorithms that assume the aperture falls at the midpoint of the primary edge that passes through the slot. Defining the slot on primary edges enables it to be included in the original solid model and meshed using beam elements. Because beam elements are used for both wires and slots, the nodes associated with the beam elements are given either a slot or wire attribute to activate the appropriate algorithm within *VOLMAX* (cf. *PreVol*, Section 1). Consequently, multiple wires and slots can reside within the same mesh.

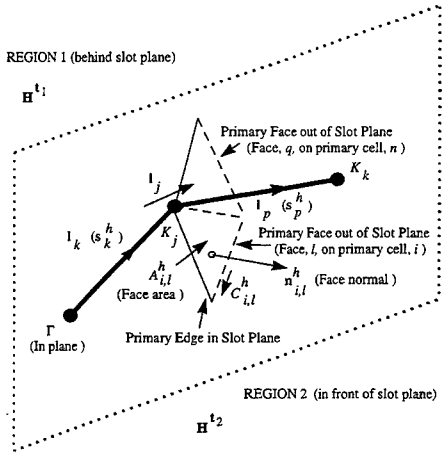


Fig. 16. Arbitrary slot path in a plane.

The vector magnetic field local to the slot is approximated by forming a least-squares fit to the face-normal magnetic fields. The vector field projected along dual edges is defined similar to Eq. (3) [1]. An example of thin-slot/thin-wire coupling is provided in the following section.

5. EMC/EMI Applications

Electromagnetic compatibility (EMC) and electromagnetic interference (EMI) issues are important in system applications. Effective shielding is often crucial to survivability and/or vulnerability requirements. Two shielding enclosure examples are presented in this section. These examples were previously investigated in [13,14] to examine the accuracy of rectilinear FDTD thin-wire and thin-slot algorithms in simplistic, but realistic geometry. The FDTD simulations were compared to measurements with good agreement over the simulation bandwidth. The geometry studied conformed to a rectangular grid. Using rectilinear FDTD

on a rotated geometry, however, can lead to significant errors in slot, wire, and cavity resonance locations [10]. In the following, the rectangular shielding enclosures are modeled using a tetrahedral mesh in conjunction with the generalized thin-wire and thin-slot algorithms. This largely removes FDTD geometrical constraints.

The first example is a closed rectangular resonator that is driven by a 50 Ω source/coaxial line. The geometry, with partial mesh, is shown in Fig. 17. A thin-wire was used with a 50 Ω termination at the top of the resonator and a 47 Ω termination at the bottom of the resonator. The diameter of the wire was 0.16 cm. The entire geometry was built as a solid model and automatically meshed with linear tetrahedral elements. Construction time was approximately 15 minutes using a Sun Ultra SPARC computer. Because the geometry represents an interior problem, there was no need to embed the unstructured grid in an FDTD mesh to form the hybrid-grid configuration.

The power delivered by the source (calculated at the 50 Ω impedance) is shown in Fig. 18. The *VOLMAX* simulation used the tetrahedral mesh, with an average edge length of 1.1 cm, in conjunction with the PDE thin-wire algorithm (Section 3.2). Comparison with measured data is made [13]. The power available from the source was 2.5 mW. The agreement is generally good. A slight (< 1%) shift in cavity resonances at approximately 1.4 GHz and 1.5 GHz is seen. It was noted in [13] that minor changes in the wire radius affect all resonance locations. No effort was made in Fig. 18 to "tune" the results; the physical wire diameter of 0.16 cm was used.

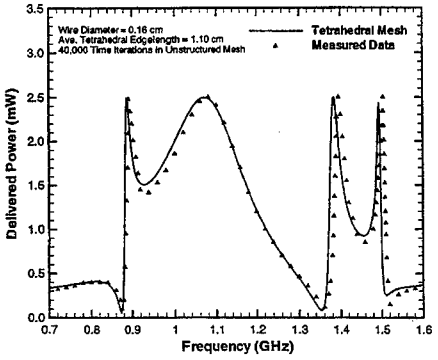


Fig. 18. Power delivered by source for Fig. 17 geometry.

The second example is similar to the first, but adds a narrow slot, with depth, to the shielding enclosure and shifts the wire location (cf. Fig. 19). Because this is now an open geometry, a full hybrid-grid implementation is used in *VOLMAX*. The interior of the enclosure is automatically meshed with linear tetrahedral elements, as well as a 1-cell-layer external to the enclosure. To accomplish this simply requires "partitioning" the enclosure geometry out of a slightly (1-cell) larger rectangular container--a task that is easily done within the CAD system. This extra layer of tetrahedral elements enables the wrapper layer to be constructed by *PreVol* (cf. Section 1) for direct interface to a cubical FDTD grid that is used to terminate the overall mesh.

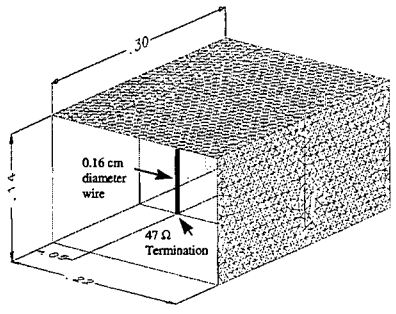


Fig. 17. Closed rectangular shielding enclosure with thin wire. 50 Ω termination at top of wire (not shown). Units in meters unless noted. Tetrahedral meshed.

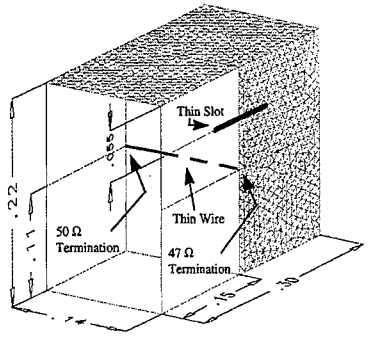


Fig. 19. Enclosure with wire and slot. Wire terminated as in Fig. 17. Slot width, 0.1 cm, slot depth, 0.05 cm, slot length, 12 cm. Wire diameter, 0.16 cm. Units in meters unless noted. Tetrahedral meshed.

The power delivered by the source is shown in Fig. 20. Comparison with measured data is made [14]. The calculation is again made at the 50 Ω load. As in the previous example, there is a slight (< 1%) shift in some resonance locations. The PDE thin-wire model (Section 3.2) and the generalized HTSA model (Section 4) were used in conjunction with the tetrahedral mesh. The resonances at approximately 1.13 GHz, 1.26 GHz, and 1.38 GHz are due to the slot. The transient response ran for 35,000 time iterations in the unstructured mesh (5,000 in the structured-grid portion of the hybrid mesh). No indication of instability was observed when using the standard *VOLMAX* time-averaging scheme on the unstructured mesh [1]. Note that the Q of all resonances is well characterized by the simulation for both examples.

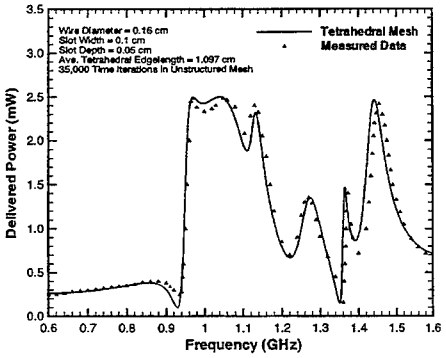


Fig. 20. Power delivered by source for Fig. 19 geometry.

6. Concluding Remarks

VOLMAX is a general-purpose, transient electromagnetic field simulator that operates on hybrid-grid structures. It is coupled to a commercial CAD system that provides advanced solid-modeling, meshing, and post processing. *VOLMAX* has been optimized for shared-memory, multi-processor computer systems (SMP). On a four-processor, Sun Ultra SPARC platform, performance ranges from 0.2 μs/cell-time-step for multi-million element structured grids, to 4 μs/cell-time-step for purely unstructured grids with a few thousand elements. Hybrid-grid problems fall between these limits.

The introduction of sub-cell wire and slot algorithms on unstructured grids significantly extends the application domain. Detailed source modeling, microelectronic packaging, complex aperture coupling, and particle-in-cell (PIC) applications using a *QUICKSILVER-VOLMAX* [17] hybrid are currently being investigated.

Acknowledgment

The authors thank Dr. David Seidel of Sandia National Laboratories for providing valuable input to the original manuscript.

References

1. D.J. Riley and C.D. Turner, "*VOLMAX*: A solid-model-based, transient *vol*umetric *Max*well solver using hybrid grids," *IEEE Antennas Propagat. Mag.*, 39, 1, 1997, pp. 20-33.
2. D.J. Riley and C.D. Turner, "Interfacing unstructured tetrahedron grids to structured-grid FDTD," *IEEE Microwave and Guided Wave Letters*, 5, 9, 1995, pp.284-286.
3. D.J. Riley and C.D. Turner, "Unstructured finite-volume modeling in computational electromagnetics," *11th Annual Review of Progress in Applied Computational Electromagnetics (ACES) Symposium Digest*, 1995, pp.435-444.
4. *I-DEAS Master Series*, Structural Dynamics Research Corporation, Milford, OH, 45150-2789.
5. D.J. Riley and C.D. Turner, "Hybrid thin-slot algorithm for the analysis of narrow apertures in finite-difference time-domain calculations," *IEEE Trans. Antennas Propagat.*, 38, 12, 1990, pp.1943-1950.
6. D.E. Merewether and R. Fisher, *Finite Difference Solution of Maxwell's Equations for EMP Applications*, EMA, Inc., Report: EMA-79-R-4, Albuquerque, NM, April 22, 1980.
7. R. Holland and L. Simpson, "Finite difference analysis of EMP coupling to thin struts and wires," *IEEE Trans. Electromag. Compat.*, 23, 2, 1981.
8. P. Bonnet, X. Ferrieres, F. Issac, F. Paladian, J. Grando, J.C. Alliot, and J. Fontaine, "Numerical modeling of scattering problems using a time domain finite volume method," *J. of Electromag. Waves and Applicat.*, 11, 1997, pp.1165-1189.
9. G.J. Burke, *Numerical Electromagnetics Code—NEC-4*, Report: UCRL-MA-109338, Lawrence Livermore National Laboratory, CA, January, 1992.
10. D.J. Riley and C.D. Turner, *General Implementation of Thin-Slot Algorithms into the Finite-Difference Time-Domain Code, TSAR, Based on a Slot Data File*, Report: SAND91-1061, Sandia National Laboratories, NM, June, 1991.
11. E.K. Reed and C.M. Butler, *Electromagnetic Penetration Through Narrow Slots in Conducting Surfaces and Coupling to Structures on the Shadow Side*, Report: TR-062188-3273F2, Clemson University, SC, July, 1990.
12. J. Gilbert and R. Holland, "Implementation of the thin-slot formalism in the finite-difference EMP code THREDII," *IEEE Trans. Nucl. Sci.*, 28, 1981, pp.4269-4274.
13. M. Li, K.-P. Ma, D.M. Hockanson, J.L. Drewniak, T.H. Hubing, and T.P. Van Doren, "Numerical and experimental corroboration of an FDTD thin-slot model for slot near corners of shielding enclosures," *IEEE Trans. Electromag. Compat.* 39, 3, 1997, pp.225-232.
14. K.-P. Ma, M. Li, J.L. Drewniak, T.H. Hubing, and T.P. Van Doren, "A comparison of FDTD algorithms for subcellular modeling of slots in shielding enclosures," *IEEE Trans. Electromag. Compat.*, 39, 2, 1997, pp. 147-155.
15. R.F. Harrington, Time-Harmonic Electromagnetic Fields, McGraw-Hill, 1961.
16. L.K. Warne and K.C. Chen, "Equivalent antenna radius for narrow slot aperture having depth," *IEEE Trans. Antennas Propagat.*, 37, 1989, pp.824-834.
17. D.B. Seidel, M.L. Kiefer, R.S. Coats, T.D. Pointon, J.P. Quintenz, and W.A. Johnson, "The 3-D electromagnetic, particle-in-cell code, QUICKSILVER," *Proceedings of the CP90 Europhysics Conference on Computational Physics*, 1991, pp.475-482.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-ACO4-94AL85000.

USING THE FINITE INTEGRATION TIME DOMAIN TECHNIQUE AT LOW FREQUENCIES

R.Ehmann[1], T.Weiland[2]

[1]*CST GmbH*
Lauteschlägerstraße 38, 64289 Darmstadt, Germany,
Phone:+49(0)6151/717057, Fax:+49(0)6151/718057,
e-mail: ehmann@temf.tu-darmstadt.de

[2]*University of Technology Darmstadt, FB 18, Schloßgartenstr. 8, 64289 Darmstadt, Germany*

ABSTRACT

The Finite Integration Method in the time domain (equivalent to FDTD) originally designed for high frequency problems will be applied to low frequency problems. A validation example demonstrates the ability to solve these problem types. We show how the computational effort depends on material properties and frequency. Finally we present the computation results of a practical example, a so called 'Shading Ring Sensor', used for distance measurement.

FINITE INTEGRATION ALGORITHM

The method used for our calculations is the FI-method [2], which in the time domain formulation becomes equal to Yee's [6] formulation.
Maxwell's Equations are transformed one to one from the continuous domain to a discrete space by allocating electric fields on a grid G and magnetic fields on a dual grid $\tilde{G}$ [1]. The allocation of the field components on the grid can be seen in Fig.1. The discrete equivalents of Maxwell's equations

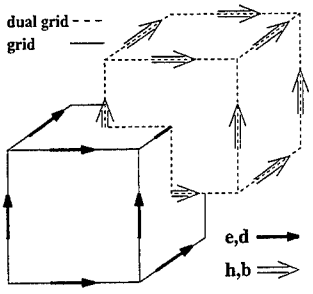


Figure 1: One cell of grid G and dual grid $\tilde{G}$ with electric and magnetic field components

Maxwell's Grid Equations

$$Ce = -\dot{b} \quad (1)$$
$$\tilde{C}h = j + \dot{d} \quad (2)$$
$$Sb = 0 \quad (3)$$
$$\tilde{S}d = q \quad (4)$$

Material Equations

$$d = D_\varepsilon e \quad (5)$$
$$b = D_\mu h \quad (6)$$
$$j = D_\kappa e + j_A \quad (7)$$

are shown in Eq.(1)-(4), where **e** and **h** are the electric voltages between grid points and the magnetic voltages between dual grid points, respectively. **d, b, j** are fluxes over grid or dual grid faces. The discrete analogon of the coupling between voltages and fluxes is represented by the diagonal material matrices $\mathbf{D}_\varepsilon, \mathbf{D}_\mu$ and $\mathbf{D}_\kappa$. Now we have mapped Maxwells Equations on a discrete space. For different

problem types we can simplify these equations now. We are interested in two different approaches, a time domain formulation, which is equivalent to FDTD and a frequency domain formulation with harmonic excitation.
The time integration is performed by using the well known leap-frog-scheme [6], which leads to an explicit scheme to solve the electromagnetic field problem. The scheme for the lossless case is described in Eqs.8-11 [2]. This algorithm is only stable for eigenvalues λ_i of $\mathbf{A}$ lying inside the unit circle. In

$$\mathbf{f}^{i+1} = \mathbf{A}\mathbf{f}^{i} + \mathbf{s}^{i} \quad (8)$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{I} & -\Delta t\mathbf{C} \\ \Delta t\mathbf{D}_{\epsilon}^{-1}\tilde{\mathbf{C}}\mathbf{D}_{\mu}^{-1} & \mathbf{I}-\Delta t^{2}\mathbf{D}_{\epsilon}^{-1}\tilde{\mathbf{C}}\mathbf{D}_{\mu}^{-1}\mathbf{C} \end{pmatrix} \quad (9)$$

$$\mathbf{f}^{i} = \begin{pmatrix} \mathbf{b}^{i} \\ \mathbf{e}^{i+1/2} \end{pmatrix} \quad (10)$$

$$\mathbf{s}^{i} = \begin{pmatrix} \mathbf{0} \\ -\Delta t\mathbf{D}_{\epsilon}^{-1}\mathbf{j}^{i} \end{pmatrix} \quad (11)$$

other words, a maximum stable time step exists, which depends directly on the discretization and the material distribution inside the calculation grid. Instead of solving the eigenvalue problem this limit can be found for regular equidistant grids with homogeneous material distribution by the well known Courant condition:

$$\Delta t \leq \left(c\sqrt{\frac{1}{\Delta x^{2}} + \frac{1}{\Delta y^{2}} + \frac{1}{\Delta z^{2}}} \right)^{-1} \quad (12)$$

In the frequency domain we are interested in fields with harmonic time dependence, so the time derivatives in Eqs.1 and 2 go over into $i\omega$. Within just a few steps we obtain the curl curl equation

$$\left(\tilde{\mathbf{C}}\tilde{\mathbf{D}}\mathbf{C} - \omega^{2}\mathbf{D}\right)\mathbf{e} = -i\omega\mathbf{j}_{A} \quad (13)$$

with

$$\tilde{\mathbf{D}} = \mathbf{D}_{\mu}^{-1} \quad (14)$$

$$\mathbf{D} = \mathbf{D}_{\epsilon} + \frac{1}{i\omega}\mathbf{D}_{\kappa}, \quad (15)$$

which can be solved now with some modern numerical techniques. For the following computations we took into consideration the results of the harmonic solver as well as those of the time domain solver. Among others, both algorithms are implemented in the software package MAFIA[1].

Obtaining Frequency Domain Data out of Time Domain Simulations

For the extraction of harmonic fields out of time domain simulations we can distinguish between two different cases, depending on the excitation. On the one hand an excitation by a harmonic signal containing a specific frequency, on the other hand a broadband excitation, e.g. by a δ-pulse. For these two types the extraction of harmonic fields, namely the real and imaginary part of the field values at a certain frequency, will be discussed.
For the monochromatic excitation we have the general problem of switching on a function. We can illustrate this by a multiplication of the harmonic time signal with a general function $s(t)$.

$$f(t) = s(t) \cdot sin(t) \quad (16)$$

[1]MAxwells Finite Integration Algorithm

Depending on the choice of the function $s(t)$, we get a more or less sharp frequency spectrum. The smoother this function $s(t)$ is chosen, the sharper our frequency spectrum is. According to the choice of $s(t)$ the computation time to reach steady state of a system will increase more or less. In general a couple of periods of the desired frequency are necessary to be computed to get the fields in steady state. The real and imaginary part of the field can then be extracted by two fields at a time distance of a quarter of a period.

$$\begin{aligned} Re\{\underline{E}\} &= E(t_0+T/4) \qquad Im\{\underline{E}\} = E(t_0) \\ Re\{\underline{H}\} &= H(t_0+T/4) \qquad Im\{\underline{H}\} = H(t_0) \end{aligned} \tag{17}$$

As a consequence of the Courant condition the smallest mesh step size in a calculation grid determines the maximum stable time step. This mesh step size is limited by convergence. To get rather accurate results one needs in practice at least 10 meshsteps per wavelength. For that reason the time domain method will usually be applied to high frequency problems. The effort to compute just a single period at a low frequency, e.g. at 50 Hz, is enormously. A homogeneous discretisation and a mesh step size of 10 cm would lead to a time step $\Delta t = 0.19\ ns$, so that the computation of one period would take about 104 million time steps. For the computation of practical structures with some 100000 meshpoints such a computation would last weeks on a modern computer.
In general, the computation time can be reduced by exciting a structure with a broadband pulse. In the following we will use a Gaussian pulse as time excitation. The computational bandwidth for that pulse depends directly on the pulselength of the timesignal (s.Fig.2,3). If we want to extract out of the discrete

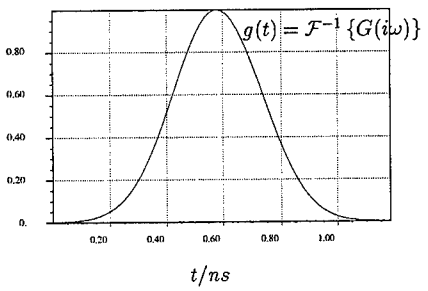


Figure 2: Gaussian Pulse in Time Domain corresponding to Fig.3

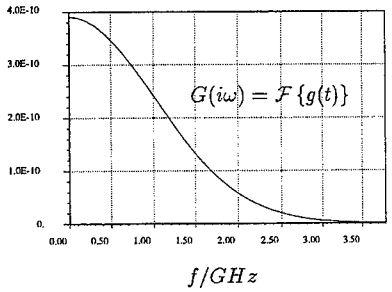


Figure 3: Gaussian Pulse in Frequency Domain corresponding to Fig.2

time domain data the real and imaginary part for one certain frequency, we have to apply a discrete Fourier transform (DFT).

$$\underline{F}(\omega) = \delta t \sum_{n=0}^{N} s(n\delta t) exp(i\omega n\delta t) \tag{18}$$

This DFT was implemented in the time domain algorithm described above. In intervals $\delta t = n_1 \Delta t$ a summation for all field components of interest has to be performed according to Eq.18, where n_1 can be determined by the sampling theorem

$$n_1 \leq \frac{1}{2\Delta t f_{max}}. \tag{19}$$

Δt is the computational time step out of the Courant Condition (Eq.12). Typically broadband excitation is used to determine signal quantities like the input impedance of an antenna or the scattering parameter of a waveguide structure. Moreover for a small number of frequencies the farfield, energies or losses or

just the field pattern of the electric or magnetic field are of interest. With the DFT feature broadband as well as harmonic results, the real and imaginary part of the electric or magnetic field at a certain frequency, can be extracted out of one single broadband computation.
The whole computation time for a structure can be determined by the excitation pulse length, signal propagation times and for resonant structures by decay times.
With an Gaussian pulse excitation we can extract now also low frequency data, since the excitation maximum is at DC. For these frequencies propagation times or resonant effects are often not of importance. As we will see in the following, the diffusion time will be the limiting factor for time the domain simulations.

Validation with a Simple Diffusion-Example

The following structure, a metallic plate with a thickness of 2 cm excited by a current coil at a distance of 2 cm, is investigated. Because of the symmetry of the structure (s. Fig.4 and 5) only a quarter of it is investigated. The frequencies of interest are 50 Hz and 10 kHz. The conductivity of the metallic plate

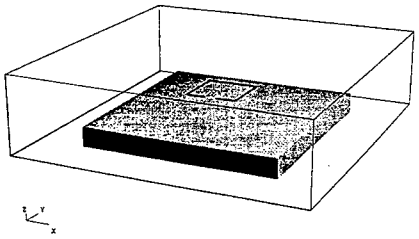


Figure 4: Metallic plate with varying conductivity excited by a coil

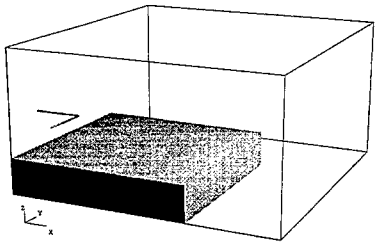


Figure 5: A quarter of the investigated structure Fig.4 using symmetries for the computation

is varied in a range from 1 S/m up to 1e7 S/m. In the figures 6, 7 the results of the frequency domain solver, refered to as F-results (curve with squares) were used as a reference. The curve with the circles show the results of the time domain simulations, in the following called T-results.
Figs.6 (a)-(f) show the z-component of the magnetic flux density on the symmetry axis of our structure, plotted versus z. Figs.6 (a,c,e) show the field at 50 Hz, (b,d,f) at 10 kHz. From top to the bottom we used a conductivity of 1e2 S/m, 1e3 S/m and 1e5 S/m for both simulation frequencies. The T- and F-results shown in Figs.6 (a-d) agree very well, whereas in (e,f) the results differ due to the short simulated time.
Since $\sigma/\omega\epsilon \gg 1$ in our problem, the diplacement current in Maxwell's Equations can be neglected. If we solve Maxwells Equations analytically now, we end up with the diffusion equation, which is written down here for the 1D-case [3].

$$\frac{\partial^2 B_z(x,t)}{\partial x^2} = \mu\sigma \frac{\partial B_z(x,t)}{\partial t} \tag{20}$$

If we solve the diffusion equation for a conducting half space $x > 0$ and an excitation by a unit-step function

$$B_z(x,t) = \begin{cases} 0 & : \quad t < 0 \\ B_0 \cdot \delta(x) & : \quad t \geq 0 \end{cases} \tag{21}$$

we finally obtain the following equation

$$B_z(x,t) = B_0 \cdot erfc\left(\frac{x\sqrt{\mu\sigma}}{2\sqrt{t}}\right) \quad with \quad erfc(z) = \frac{2}{\sqrt{\pi}} \int_x^\infty exp(-u^2)du \tag{22}$$

From this formula we can gain an approximation for the penetration time of an electromagnetic field

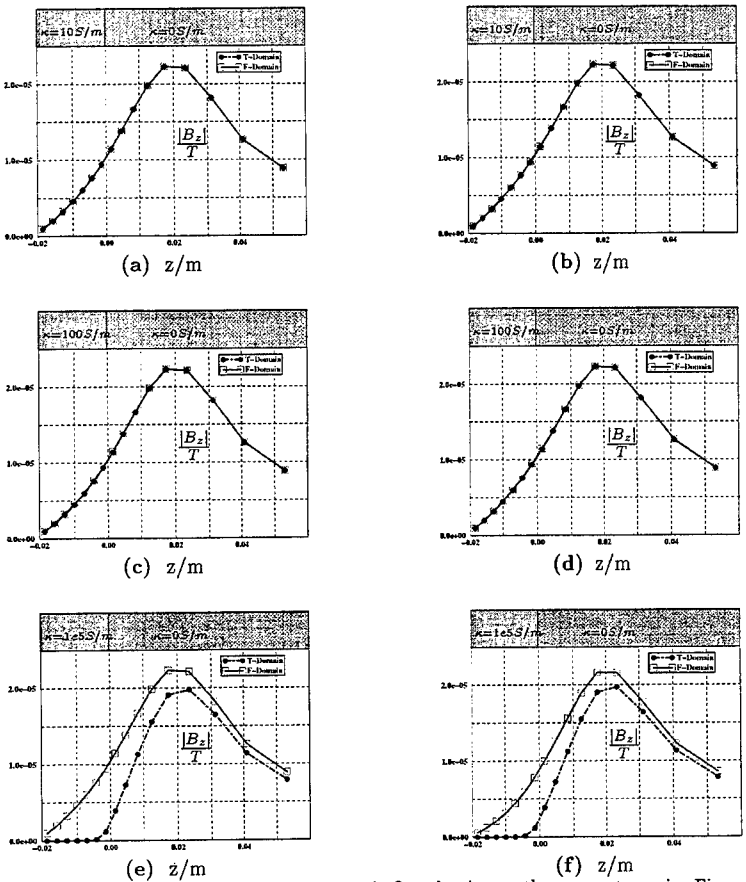


Figure 6: (a)-(f) show the z-component of the magnetic flux density on the symmetry axis. Fig. a,c,e show the field at 50 Hz, Fig. b,d,f at 10kHz for the material parameters shown in the plots

into a conducting material in dependence on the penetration depth. If we determine the half width of the previous result (Eq.22), we will find:

$$erfc\left(\frac{x\sqrt{\mu\sigma}}{2\sqrt{t}}\right) = \frac{1}{2} \quad \Rightarrow \quad \frac{x\sqrt{\mu\sigma}}{2\sqrt{t}} \approx 0.48 \tag{23}$$

$\frac{\sigma}{S/m}$	$\frac{t_{sim}}{\mu s}$	f = 50 Hz				f = 10 kHz			
		δ/m	$\frac{t_{diffus}}{\mu s}$	$\frac{cpu-time}{s}$		δ/m	$\frac{t_{diffus}}{\mu s}$	$\frac{cpu-time}{s}$	
				T	F			T	F
1	0.014	71.2	5.03e-4	131.3	22.04	5.03	5.02e-4	131.0	22.41
10	0.014	41.1	5.03e-3	131.2	22.01	1.59	5.03e-3	131.05	22.5
100	0.051	7.12	0.05	452.26	22.04	0.503	0.05	451.8	20.62
1e5	0.28	0.225	50.3	2488.6	41.2	0.016	31.76	2491.6	36.8
1e7	0.28	0.0225	50.3	2490.8	35.9	0.0016	31.76	2492.6	41.9

Table 1: Skin depths, diffusion times, simulation times, cpu-times for T- and F-results

Approximately we get

$$x \approx \frac{t}{\mu\sigma} \quad and \quad t \approx \mu\sigma x^2. \tag{24}$$

Focussing our attention on the simulation times we realize, that they are for case (e,f) in Fig.6 much too short (s.Tab.1). The electromagnetic field cannot penetrate into the metallic plate. For a computation time of $6\mu s$, 12 % of the diffusion time, we can see in Fig.7 that the T-results converge against the F-results.

The diffusion time increases linear with the conductivity (Tab.1, Eq.24) up to the point where material

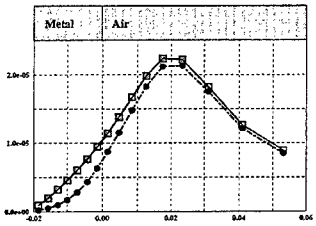


Figure 7: z-component of magnetic flux density for $\sigma = 1e5 S/m$ and a simulated total time of about $6\mu s$ (f=50Hz)

depth and skindepth of the plate coincide. Then the diffusion time becomes independent of the material properties and is inverse proportional to the frequency.

$$t_{diffus} = \begin{cases} \mu\sigma d^2 & : \ \delta \geq d \\ \frac{1}{\pi f} & : \ \delta < d \end{cases} \quad with \quad \delta = \sqrt{\frac{2}{\omega\mu\sigma}}, \quad d = material\ thickness \tag{25}$$

For small conductivities the simulation times (Tab.1) of time and frequency domain solver are comparable, whereas with increasing values of σ the frequency domain solver is obviously preferable.

Although for most practical problems the application of the frequency domain solver is faster, the mere possibilty to obtain the same results with a method originally designed for high frequency problems, is impressing.

Practical Example

The following example may demonstrate the applicability of the method described before to practical applications. Fig.8 shows a quarter of the geometry of a "Shading Ring Sensor" developed by the

company Robert Bosch GmbH, Stuttgart, Germany. It consists of a E-shaped core. On the middle part a coil and a shading ring are located. Depending on the position of that shading ring the impedance of the exciting coil changes. The dependence between inductivity of the coil and the distance is ideally linear, so that the inductivity can be used for distance measurement.

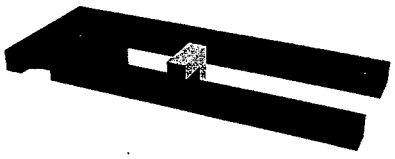


Figure 8: Shading Ring Sensor

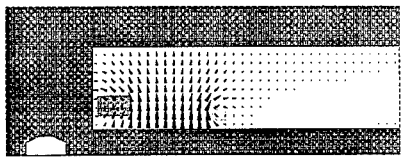


Figure 9: Real part of the magnetic field strength in a symmetry plane

A comparison of the measured inductivites and the computed ones with frequency and time domain solver will be presented at the conference.

Conclusion

The Finite Integration Technique in the time domain, a technique typically used for high frequency applications, was used to solve low frequency problems. The computed results agree very well with reference results obtained from a frequency domain solver. For sure, the computational effort for very low frequency problems can be enourmously, but on the other hand, the frequency range for applying the FI-technique in time domain (or FDTD) is expanded obviously with this approach.

References

[1] Weiland T. : *A Numerical Method for the Solution of the Eigenwave Problem of Longitudinally Homogeneous Waveguides*, Electronics and Communication (AEÜ) Vol.31(1977),pp.308.

[2] Weiland T. : *Time Domain Electromagnetic Field Computation with Finite Difference Methods*, Numerical Modelling, Electronic Networks, Devices and Fields, Vol.9, 295-319 (1996).

[3] Lehner G.:*Elektromagnetische Feldtheorie für Ingenieure und Physiker*, Springer-Verlag, ISBN 0-387-52319-7.

[4] Hahne P. :*Zur Numerischen Berechnung Zeitharmonischer Elektromagnetischer Felder*, Dissertation, TH Darmstadt, 1992.

[5] E.Zabler and F.Heinz : *Shading Ring Sensors as Versatile Position and Angle Sensors in Motor Vehicles*, Sensors and Actuators, 3 (1982,1983), S.315-326

[6] K.S.Yee : *Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media*, IEEE, AP-14, 1966, pp. 302-307.

Modelling Dispersive Media Using the Finite Integration Technique

S. Gutschling, H. Krüger, T.Weiland

Technische Universität Darmstadt, FB 18, Schloßgartenstr. 8, 64289 Darmstadt, Germany
Phone:+49(0)6151/162461, Fax:+49(0)6151/164611,
e-mail: gutschling@temf.tu-darmstadt.de

Abstract

Since simulation of broadband applications have gained in importance in the last years, the dispersive characteristics of various materials must not be neglected anymore. As a result many frequency dependent FDTD methods have been set up which in most cases model special dispersions of low order. On foundation of discrete system analysis we present an algorithm applicable to arbitrary material dispersions up to 2^{nd} order derived from a general approach [1]. The applicability of the presented method is demonstrated with an example using a rectangular waveguide filled with dielectric layers with different dispersion characteristics.

Introduction

The formulation of the *Finite Integration Technique* (FIT) according to Weiland [2] provides a general spatial discretization scheme usable for different electromagnetic applications of arbitrary geometry, e.g. static problems or calculations in frequency and time domain. In our paper we refer to the *Maxwell's Grid Equations* (MGE) (1)-(4) and material relations (5)-(7) given by

$$\mathbf{C}\,\mathbf{D}_s\,\mathbf{e} = -\mathbf{D}_A\,\dot{\mathbf{b}} \quad (1)$$

$$\tilde{\mathbf{C}}\,\tilde{\mathbf{D}}_s\,\mathbf{h} = \tilde{\mathbf{D}}_A\,\dot{\mathbf{d}} \quad (2)$$

$$\tilde{\mathbf{S}}\,\tilde{\mathbf{D}}_A\,\mathbf{d} = \mathbf{q} \quad (3)$$

$$\mathbf{S}\,\mathbf{D}_A\,\mathbf{b} = \mathbf{0} \quad (4)$$

$$\mathbf{d} = \mathbf{D}_\epsilon\,\mathbf{e} \quad (5)$$

$$\mathbf{b} = \mathbf{D}_\mu\,\mathbf{h} \quad (6)$$

$$\mathbf{b} = \mathbf{D}_\kappa\,\mathbf{j}. \quad (7)$$

The geometry is discretized on a dual orthogonal grid system with $\mathbf{e}$, $\mathbf{b}$ located on the normal grid G and $\mathbf{d}$, $\mathbf{h}$ on the dual grid $\tilde{G}$. Correspondent to that the analytical curl operator results in the curl matrices $(\mathbf{C}, \tilde{\mathbf{C}})$ and the divergence operator in the source matrices $(\mathbf{S}, \tilde{\mathbf{S}})$. In the same way the grid resolution is contained in $(\mathbf{D}_s, \tilde{\mathbf{D}}_s)$ representing the grid lines and $(\mathbf{D}_A, \tilde{\mathbf{D}}_A)$ the belonging areas. If the material is assumed to be frequency independent and isotropic, we have diagonal matrices $\mathbf{D}_\epsilon$ and $\mathbf{D}_\mu$ describing the material relations. It can be shown, that the mentioned spatial discretization does not produce any instability since the discrete Maxwell equations fulfil energy and charge conservation [2].
Applying the well-known leap-frog scheme to the FIT formulas we can write (1,2) in form of two recursive update equations with $\mathbf{e}$ and $\mathbf{b}$ as the calculated field variables:

$$\mathbf{b}^{n+1} = \mathbf{b}^n - \Delta t\,\mathbf{D}_A^{-1}\,\mathbf{C}\,\mathbf{D}_s\,\mathbf{e}^{n+1/2} \quad (8)$$

$$\mathbf{e}^{n+3/2} = \mathbf{e}^{n+1/2} + \Delta t\,\mathbf{D}_\epsilon^{-1}\,\tilde{\mathbf{D}}_A^{-1}\,\tilde{\mathbf{C}}\,\tilde{\mathbf{D}}_s\,\mathbf{D}_\mu^{-1}\,\mathbf{b}^{n+1}. \quad (9)$$

Using a homogeneous equidistant grid these equations reduce to the standard finite-difference time-domain ($FDTD$) algorithm according to Yee [3]. Now stability due to time discretization is restricted to a certain interval, namely given by the Courant condition in free-space

$$\Delta t = \left(c_0 \sqrt{\left[\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2} \right]} \right)^{-1} . \tag{10}$$

Since the sofar described time domain algorithm is restricted to non-dispersive materials many efforts have been made in the last years to expand it in a useful way. An important aspect in connection with these extensions is the guarantee of stability, because it is not possible to transfer the criterion (10) to frequency dependent materials in a straight forward manner. Apart from a quite practicable solution for this problem we have proposed in a recent paper [1] a very general time domain algorithm for dispersive materials. There we provide a stability analysis that is applicable to any frequency dependent time-domain method and therefore offers good possibilities for comparisons of the most important $(FD)^2TD$ algorithms [4, 5, 6, 7, 8, 9].

Algorithm for 2^{nd} Order Dispersion Models

Our approach is within the framework of system analysis by first considering a linear time-invariant system of n^{th} order, that can be described in general by a linear ordinary differential equation (ODE) of the same order. Rather then to discretize the n^{th} order ODE directly by replacing time derivatives by the corresponding central difference operator [8], we first apply the *state space formulation* to our system to derive an explicit algorithm for the time-domain simulation[1]. This formulation is chosen, since it employs matrices in its fundamental equations similar to the FIT-method and therefore both methods can easily be combined.
Since this procedure is presented in [1], we skip the derivation of the general approach and we present in the following the derived explicit update equation for a 2^{nd} order dispersion model. We choose a maximum order of two for the dispersion, since it covers the most significant dispersion models like Debye, Drude and Lorentz. Thus in the frequency domain the correspondent permittivity function reads as

$$\epsilon(\omega) = \epsilon_0 \left(\underbrace{\beta_2}_{\epsilon_\infty} + \frac{\beta_0 + j\omega \cdot \beta_1}{\alpha_0 + j\omega \cdot \alpha_1 - \omega^2 \cdot \alpha_2} \right) . \tag{11}$$

The discretization in time is done by using exact integration of the first order ODE's. In general we derive from $dy(t)/dt = A\,y(t) + b(t)$ for the homogeneous case $y_h(t) = C\ exp(A\,t)$ and a special solution $y_s(t) = C(t)\,exp(A\,t)$ with $C(t) = -b/A\,exp(A\,t)$ by variation of parameters. The combination gives us the general solution and finally the expression for a discrete time step Δt

$$y^{n+1} = y^n\, e^{A\,\Delta t} + (e^{A\,\Delta t} - 1)/A\, b^{n+1/2}. \tag{12}$$

Here we like to mention that we assumed the function b as constant over the time step and separated by half a time step, where we choose the allocation of y at full time steps (alternatively

one has to add $\Delta t/2$ to all signals in equation (12) in case that y is allocated at $t = (n+1/2)\,\Delta t)$. Unfortunately this is not the case in the ODE for the first state variable $\mathbf{z_1}$ (the polarization), since it includes the electric field on the right hand side, which is allocated at the same positions in time. In order to ensure a higher accuracy the electric field $\mathbf{e}^{n+1}$ is averaged by its existing neighbour values $\mathbf{e}^{n+1} = (\mathbf{e}^{n+3/2} + \mathbf{e}^{n+1/2})/2$ (see equation (16)). This finally leads to the following set of four coupled equations

$$\mathbf{b}^{n+1} = \mathbf{b}^n - \Delta t\, \mathbf{D}_A^{-1}\, \mathbf{C}\, \mathbf{D}_s\, \mathbf{e}^{n+1/2} \tag{13}$$

$$\mathbf{z_2}^{n+1} = \mathbf{D}_{exp_2}\, \mathbf{z_2}^n + (I - \mathbf{D}_{exp_2})\, \mathbf{D}_{\alpha_1}^{-1}\, (-\mathbf{D}_{\alpha_0}\, \mathbf{z_1}^{n+1/2} + \mathbf{D}_{b_2}\, \mathbf{e}^{n+1/2}) \tag{14}$$

$$\mathbf{e}^{n+3/2} = \mathbf{D}_{exp_1}\, \mathbf{e}^{n+1/2} + (\mathbf{I} - \mathbf{D}_{exp_1})\, (-\mathbf{D}_{b_1}^{-1}\, \mathbf{z_2}^{n+1} + \mathbf{D}_{b_1}^{-1}\, \tilde{\mathbf{D}}_A^{-1} \tilde{\mathbf{C}}\, \tilde{\mathbf{D}}_s\, \mathbf{D}_\mu^{-1}\, \mathbf{b}^{n+1}) \tag{15}$$

$$\mathbf{z_1}^{n+3/2} = \mathbf{z_1}^{n+1/2} + \Delta t\, \mathbf{z_2}^{n+1} + \Delta t\, \mathbf{D}_{b_1}\, \underbrace{\frac{1}{2}\, (\mathbf{e}^{n+3/2} + \mathbf{e}^{n+1/2})}_{\bar{\mathbf{e}}^{n+1}} \tag{16}$$

with the matrices $\mathbf{D}_{b_1} = \mathbf{D}_{\beta_1} + \mathbf{D}_\kappa$; $\mathbf{D}_{b_2} = \mathbf{D}_{\beta_0} - \mathbf{D}_{\alpha_1}\, \mathbf{D}_{\beta_1}$; $\mathbf{D}_{exp_1} = exp(-\mathbf{D}_{\beta_2}^{-1}\, \mathbf{D}_{b_1}\, \Delta t)$ and $\mathbf{D}_{exp_2} = exp(-\mathbf{D}_{\alpha_1}\, \Delta t)$. In this algorithm we have also taken a static conductivity into account, that can easily be added by the extension of the matrix $\mathbf{D}_{b_1} = \mathbf{D}_{\beta_1} + \mathbf{D}_\kappa$, where the diagonal matrix $\mathbf{D}_\kappa$ represents the distribution of the conductivity inside the grid.
For simulating multiple media with different dispersion models up to second order in a single time domain calculation simultaneously, one has to set the dispersion model coefficients accordingly. In Table 1 they are summarized for the most relevant dispersion models, where the not listed coefficients are set to $\alpha_2 = 1$ and $\beta_2 = \epsilon_0\, \epsilon_\infty$ by definition.

Table 1: Permittivity model coefficients of Debye, Drude, Debye 2^{nd} order and Lorentz dispersion for the 2^{nd} order algorithm (13)-(16).

	Debye	Drude	Debye 2^{nd}	Lorentz
α_0	0	0	$1/(\tau_1\, \tau_2)$	ω_0^2
α_1	$1/\tau$	ν_c	$(\tau_1 + \tau_2)/(\tau_1\, \tau_2)$	δ
β_0	0	$\epsilon_0\, \Delta\epsilon\, \omega_p^2$	$\epsilon_0\, (\Delta\epsilon_1 + \Delta\epsilon_2)/(\tau_1\, \tau_2)$	$\epsilon_0\, \Delta\epsilon\, \omega_0^2$
β_1	$\epsilon_0\, \Delta\epsilon/\tau$	0	$\epsilon_0\, (\Delta\epsilon_1\, \tau_2 + \Delta\epsilon_2\, \tau_1)/(\tau_1\, \tau_2)$	0

EXAMPLE

To verify the presented method, the 2^{nd} order algorithm is applied to an S-parameter calculation. In Figure 1 the test structure, a dielectric filled waveguide with different layers in propagation direction, is shown. Two frequency dependent materials with a 2^{nd} order dispersion are present (Debye 2^{nd} order, Lorentz medium; see Figure 2). The rest of the waveguide is filled with vacuum and throughout the waveguide the permeability equals $\mu = \mu_0$.

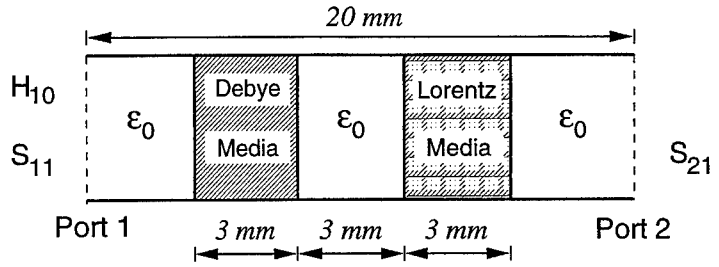


Figure 1: *Rectangular waveguide ($\mu = \mu_0$, port separation is 20 mm, cross-section 20 mm x 5 mm) with layers of different permittivities including ϵ_0 and dispersive permittivities Debye 2nd order: $\epsilon_\infty = 1$; $\epsilon_{s_1} = \epsilon_{s_2} = 2$, $\tau_1 = 1/2/\pi/10e9\ s$, $\tau_2 = 1/2/\pi/20e9\ s$; Lorentz medium: $\epsilon_\infty = 1$; $\epsilon_{s_1} = 2$, $\delta = 20e9\ Hz$; $\omega_0 = 2\ \pi\ 20e9\ Hz$.*

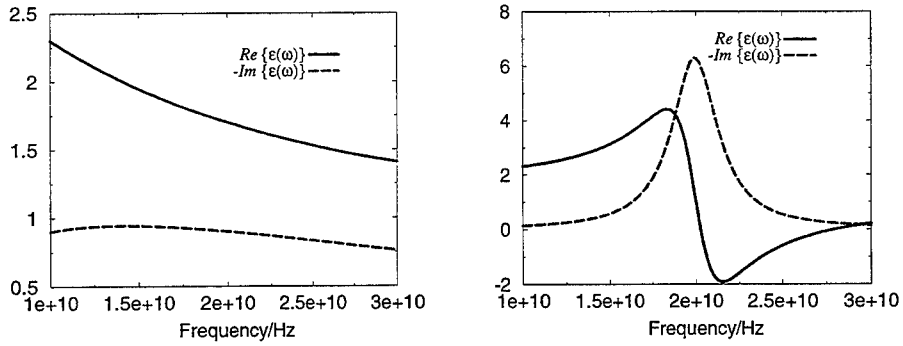


Figure 2: *a) Real and imaginary part of 2nd order Debye material (frequency range 10 GHz-30 GHz) b) Real and imaginary part of Lorentz material (frequency range 10 GHz-30 GHz).*

We want to determine the amplitude and phase of the S_{11}, S_{21} parameters at the given ports separated by 20 mm for the frequency range 10 GHz - 30 GHz. Thus a broadband stimulation with the fundamental mode at port 1 in form of a Gaussian pulse modulated with a carrier frequency of 20 GHz results in the frequency domain in a Gaussian shaped excitation spectrum centred at 20 GHz with a 60dB bandwidth of 10 GHz. At the two ports a special waveguide boundary condition is used [11] that enables the simulation of an infinitely long waveguide ensuring a parasitic reflection of less than -120dB. To minimize grid dispersion the grid resolution is chosen such that it allows for thirty steps per wavelength for the highest frequency.
Thus the S-parameter calculation covers the following steps:

1. 2D-eigenvalue solver: calculation of the propagation modes inside the waveguide by discretizing the cross section of the waveguide ($\epsilon = \epsilon_0, \mu = \mu_0$).

2. 3D time domain simulation: broadband excitation with the fundamental mode at port 1

and monitoring the mode amplitude of the reflected wave at port 1 and the transmitted mode amplitude at port 2.

3. Post-processing: S-parameter calculation from the excitation and the monitored signals in the frequency domain by using the *Fast Fourier Transform* (FFT).

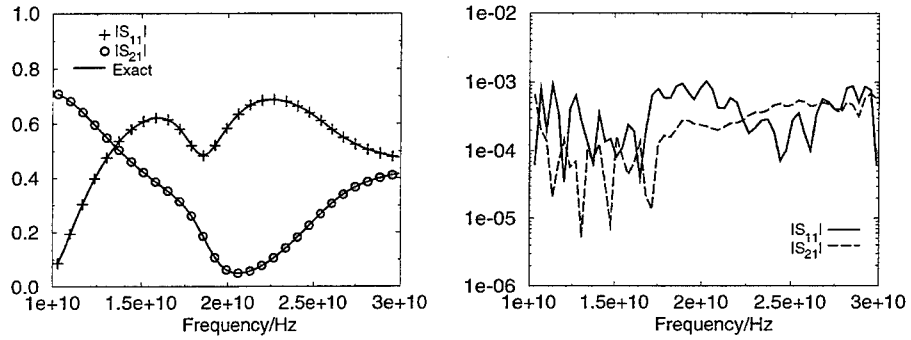


Figure 3: *Comparison of numerical results with analytical solution in the frequency range* $10\,GHz$-$30\,GHz$. *a) Absolute value of S-parameter* S_{11}, S_{21}; *b) amplitude error of S-parameter* $|S_{11}|$, $|S_{21}|$.

Figure 3 presents the absolute value of S-parameter S_{11}, S_{21} compared with the analytical solution and the resulting amplitude error for the frequency range $10\,GHz$-$30\,GHz$. As it can be seen there is an excellent agreement of the numerical results with the exact solution. The absolute amplitude error is well below 10^{-3}.

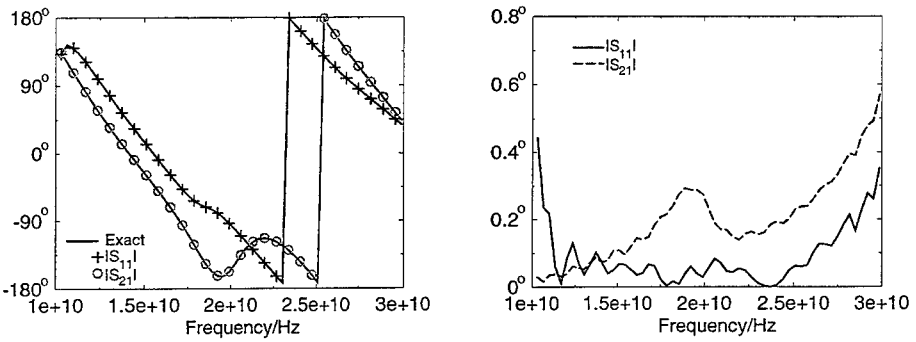


Figure 4: *Comparison of numerical results with analytical solution in the frequency range* $10\,GHz$-$30\,GHz$. *a) Phase of S-parameter* S_{11}, S_{21}; *b) Phase error of S-parameter* $|S_{11}|$, $|S_{21}|$.

A similar good agreement in case of both S-parameter phase results shows Figure 4. Here the maximum absolute phase error is below 0.6^{o}.

Conclusion

In this paper we presented a very general possibility to extend the FIT algorithm for modelling dispersive media with a dispersion of 2^{nd} order. This algorithm was derived from a general approach based on system analysis with a state-space formulation. The additional state-variables correspond to physical properties, the polarisation und the polarisation current density. We demonstrated the good accuracy of our algorithm with an example of a rectangular waveguide filled with two layers of frequency dependent material of second order (Lorentz-Media and a 2^{nd} order Debye-Model).

References

[1] S. Gutschling, H. Krüger, T. Weiland; Time domain simulation of dispersive media with the finite integration technique; submitted to International Journal of Numerical Modelling

[2] T. Weiland; Time domain electromagnetic field computation with finite difference methods; 1996; Int. J. of numerical modelling; Vol. 9; p. 295-319

[3] K.S. Yee; Numerical solution of initial boundary value problems involving Maxwells Equations in isotropic media; 1966; IEEE Transactions on antennas and propagation; Vol. 17; p. 585-589

[4] K.S. Kunz and R.J. Luebbers; The finite difference time domain method for electrodynamics; 1993; Boca Raton, FL

[5] D.M. Sullivan; Z-transform theory and the FDTD method; 1996; IEEE Transactions on antennas and propagation; Vol. 44; p. 28-34

[6] W.H. Weedon and C.M. Rappaport; A general method for FDTD modeling of wave propagation in arbitrary frequency-dispersive media; 1997; IEEE Transactions on antennas and propagation; Vol. 45; p. 401-410

[7] C. Hulse and A. Knoesen; Dispersive models for the finite-difference time-domain method: design, analysis and implementation; 1994; J. Opt. Soc. Am. A; Vol. 11; p. 1802-1811

[8] R.M. Joseph, 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; Opt. Lett. 16; p. 1412-1414 (1991)

[9] J.L. Young; Propagation in linear dispersive media: finite difference time domain methodologies; 1995; IEEE Transactions on antennas and propagation; Vol. 43; p. 422-426

[10] S. Gutschling, T. Weiland; Resonance Effects Inside a Human Head Exposed to a Cellular Phone; 3rd International Conference on Computation in Electromagnetics; CEM'96 Bath, No. 420, pp.62-66

[11] M. Dohlus, P. Thoma, T. Weiland; Broadband Simulation of open Waveguide Boundaries within large frequency ranges; IEEE Workshop; Discrete time domain modelling of electromagnetic fields and networks, Berlin, 1993

Transient Analysis of Thin Wire Antennas Mounted on Three–Dimensional Perfectly Conducting Bodies

Kemal Aygün*, A. Arif Ergin, Balasubramaniam Shanker, Stephen E. Fisher, and Eric Michielssen
Center for Computational Electromagnetics
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
1406 W. Green St., Urbana, IL 61801
e-mail: kaygun@decwa.ece.uiuc.edu

Abstract

An electric field integral equation based marching–on–in–time algorithm is developed for analyzing transient radiation from thin wire antennas mounted on three–dimensional, perfectly conducting bodies. The feed network for the antennas is also included in the analysis using a one–dimensional finite difference time domain scheme. Numerical examples that validate and demonstrate the efficacy of the proposed method are presented.

1 Introduction

In the past, time domain integral equation based methods have been employed for analyzing transient scattering and radiation phenomena. Scattering from three–dimensional perfectly conducting and dielectric bodies has been simulated using both Electric and Magnetic Field Integral Equation (EFIE,MFIE) based solvers [1–4]. Similar studies have been conducted separately on wire antennas [5,6]. This paper describes an EFIE based Marching–On–in–Time (MOT) algorithm that enables the transient analysis of radiation from complex structures that consist of arrays of thin wire antennas mounted on arbitrarily shaped perfectly conducting bodies (Fig. 1). Transient fields on the network feeding the antennas are also computed in conjunction with the currents on the radiating structure.

This paper is organized as follows. Section 2 outlines the formulation of the time domain EFIE and MOT algorithm. Section 3 presents several numerical results obtained using the proposed technique. The last section states the conclusions of this study.

2 Formulation

In this section, an algorithm is outlined for analyzing radiation from wire antennas that are mounted on arbitrarily shaped conducting bodies. Section 2.1 describes the time domain EFIE which relates the excitation field to the electric currents on the radiating structure. Section 2.2 outlines an MOT algorithm for solving this equation. Section 2.3 introduces the feed network which provides the antenna excitation, and describes a finite difference based updating scheme that complements the MOT scheme for modeling the feeds.

2.1 The Time Domain EFIE

Let S denote the surface of a perfectly conducting structure composed of wire antennas mounted on bodies. Assume that an incident electric field $\bar{E}^{inc}(\bar{r},t)$ induces a current $\bar{J}(\bar{r},t)$ on S. The field $\bar{E}^{r}(\bar{r},t)$ radiated by $\bar{J}(\bar{r},t)$ can be computed using a dual potential formulation as

$$\bar{E}^{r}(\bar{r},t) = -\frac{\partial}{\partial t}\bar{A}(\bar{r},t) - \nabla\Phi(\bar{r},t). \tag{1}$$

In Eqn. (1), the magnetic vector potential $\bar{A}$ is

$$\bar{A}(\bar{r},t) = \frac{\mu_0}{4\pi}\int_S dS' \frac{\bar{J}(\bar{r}',\tau)}{R}, \tag{2}$$

and the scalar potential Φ is

$$\Phi(\bar{r},t) = \frac{1}{4\pi\epsilon_0}\int_S dS' \frac{\sigma(\bar{r}',\tau)}{R}. \tag{3}$$

In Eqns. (2) and (3), $R = |\bar{r} - \bar{r}'|$ is the distance between the source and observation points, $\tau = t - R/c$ denotes the retarded time, and μ_0 and ϵ_0 are the free space permeability and permittivity, respectively. The current and charge density on S are related by the continuity equation

$$\nabla \cdot \bar{J}(\bar{r},t) + \frac{\partial}{\partial t}\sigma(\bar{r},t) = 0. \tag{4}$$

Using Eqns. (1) through (4) and enforcing the boundary condition on the tangential electric field on S leads to an integro-differential equation in terms of $\bar{J}(\bar{r},t)$:

$$\hat{s}\cdot\frac{\partial}{\partial t}\bar{E}^{inc}(\bar{r},t) = \hat{s}\cdot\left(\frac{\partial^2}{\partial t^2}\frac{\mu_0}{4\pi}\int_S dS' \frac{\bar{J}(\bar{r}',\tau)}{R} - \frac{1}{4\pi\epsilon_0}\nabla\int_S dS' \frac{\nabla'\cdot\bar{J}(\bar{r}',\tau)}{R}\right), \tag{5}$$

where $\hat{s}$ is a vector tangent to the radiator surface S.

2.2 The MOT Algorithm

The first step in constructing a time-marching procedure to solve Eqn. (5) involves the discretization of S. In what follows, it is assumed that the surfaces belonging to S are approximated in terms of triangular facets, and that wires are modeled by straight wire segments. Three distinct basis functions are used to represent the currents on S. Given a triangular mesh of the surfaces, surface currents are expanded in terms of the well-known Rao-Wilton-Glisson basis functions [8]. One basis function is associated with each edge interior to S:

$$\bar{f}_n^s(\bar{r}) = \begin{cases} \frac{l_n}{2A_n^+}\bar{\rho}_n^+ & ;\bar{r} \text{ in } T_n^+ \\ \frac{l_n}{2A_n^-}\bar{\rho}_n^- & ;\bar{r} \text{ in } T_n^- \\ \bar{0} & ; \text{elsewhere,} \end{cases} \tag{6}$$

where l_n is the length of the edge common to the facets T_n^+ and T_n^-, and $A_n^\pm$ is the area of the triangle $T_n^\pm$ (Fig. 2(a)).

Each wire basis function is associated with a node connecting two wire segments. The current at wire node n located at $\bar{r}_n$ and connecting segments n and $n+1$ is modeled by a triangular basis function given by

$$\bar{f}_n^w(s) = \begin{cases} \hat{s}_n(1 + \frac{s}{s_{n-1}}) & ; -s_{n-1} \le s \le 0 \\ \hat{s}_{n+1}(1 - \frac{s}{s_{n+1}}) & ; 0 \le s \le s_{n+1} \\ \bar{0} & ; \text{elsewhere,} \end{cases} \tag{7}$$

where $\hat{s}_n = (\bar{r}_n - \bar{r}_{n-1})/|\bar{r}_n - \bar{r}_{n-1}|$ is the local tangent unit vector associated with segment n, $s_{n-1} = |\bar{r}_{n-1} - \bar{r}_n|$, $s_{n+1} = |\bar{r}_{n+1} - \bar{r}_n|$, and s is a local length coordinate which measures the distance away from node n in the direction specified by $\hat{s}_n$ for $-s_{n-1} \le s \le 0$ and by $\hat{s}_{n+1}$ for $0 \le s \le s_{n+1}$ (Fig. 2(b)). In what follows, it is assumed that the wire radii are electrically small so that thin wire approximations hold.

A surface–wire junction basis function, $\bar{f}_k^{sw}(\bar{r})$, describes the current flowing into a wire segment from all of the junction triangles that have the junction node as a vertex [9]. The wire portion of $\bar{f}_k^{sw}(\bar{r})$ has exactly the same form as a wire basis. The portion of the basis function associated with the k^{th} junction triangle is given by

$$\bar{f}_k^{sw}(\bar{r}) = \begin{cases} -\dfrac{1-(1-\eta_k)^2}{\alpha\,|\bar{\rho}_k|^2}\bar{\rho}_k & ;\bar{r} \in T_k \\ \bar{0} & ;\text{elsewhere,} \end{cases} \tag{8}$$

where $\bar{\rho}_k$ is the local position vector defined in the junction triangle pointing from the junction node to the observation point and α is the angle in radians subtended by all of the junction triangles. Parameter η_k is defined as $\eta_k = A_1/A$ where A_1 is the area of the triangle determined by $\bar{r}$ and the two nodes of the junction triangle as shown in Fig. 2(c), and A is the total area of the junction triangle.

With the spatial basis functions defined, the current in Eqn. (5) can be approximated as

$$\bar{J}(\bar{r},t) \approx \sum_{n=1}^{N}\sum_{j=-\infty}^{\infty} I_n^j \bar{f}_n^q(\bar{r})T(t-j\Delta t), \tag{9}$$

where N is the total number of unknowns, $\bar{f}_n^q(\bar{r})$ for $q = s, w, sw$ is the corresponding surface, wire, or surface–wire junction basis function, $T(t)$ is a temporal basis function, and Δt is the time step size. $T(t)$ is chosen to be a cubic interpolation function with a piecewise continuous second derivative.

Substituting Eqn. (9) into (5), and applying Galerkin testing at the j^{th} time step yields a system of equations that can be concisely represented in matrix form as

$$\mathcal{Z}_0\mathcal{I}_j = \mathcal{E}_j^{inc} + \sum_{l=1}^{j-1}\mathcal{Z}_l\mathcal{I}_{j-l}, \tag{10}$$

where $\mathcal{I}_j$ is an array of the current coefficients I_n^j, $\mathcal{Z}_l$ is a matrix that accounts for the interactions between the $(j-l)^{th}$ and j^{th} time steps, and the array $\mathcal{E}_j^{inc}$ represents the time derivative of the incident field tested at the j^{th} time step. It is assumed that the incident field is due to delta–gap sources that are located at the surface–wire junctions. In the MOT scheme, current coefficients I_n^j are calculated by starting from the first time step and solving Eqn. (10) at each time step.

2.3 Analysis of the Feed Network

The voltages associated with the delta–gap sources that excite the antennas follow from a transient analysis of the feed network. The feed network is modeled in terms of one–dimensional transmission lines. The currents and voltages on the transmission line are calculated using a one–dimensional finite difference time domain scheme as described in [7]. In this scheme, the update equations for the n^{th} node of the transmission line at the j^{th} time step are given by

$$V_n^{j+0.5} = V_n^{j-0.5} - (Z_0)(\frac{v\Delta t}{\Delta z})[I_{n+0.5}^j - I_{n-0.5}^j], \tag{11}$$

$$I_{n+0.5}^{j+1} = I_{n+0.5}^j - \frac{1}{Z_0}(\frac{v\Delta t}{\Delta z})[V_{n+1}^{j+0.5} - V_n^{j+0.5}], \tag{12}$$

where v and Z_0 are respectively the phase velocity and the characteristic impedance of the line, and Δz is the distance between two adjacent nodes. The voltage update equation at node n_a where an antenna is connected can be found from

$$\mathcal{E}_{j,n_a}^{inc} = \frac{\partial}{\partial t}V_{n_a}^j \approx \frac{V_{n_a}^{j+0.5} - V_{n_a}^{j-0.5}}{\Delta t} = -Z_0(\frac{v}{\Delta z})[I_{n_a+0.5}^j - I_{n_a-0.5}^j] - Z_0(\frac{v}{\Delta z})I_a^j. \tag{13}$$

In Eqn. (13), $V_{n_a}^j$ and I_a^j represent the delta–gap voltage and current at the corresponding surface–wire junction at the j^{th} time step. Hence, the values of $V_{n_a}^{j+0.5}$ and I_a^j can be found by solving Eqns. (10) and (13) simultaneously.

3 Numerical Examples

The first example considered is a $0.5m$ length monopole antenna mounted perpendicularly on a $1m \times 1m$ conducting plate that lies on the xy-plane (Fig. 3). The plate is discretized by 128 triangular facets. The antenna is excited by a delta-gap voltage source that is directly connected to the surface-wire junction. The time dependence of the source is given by

$$V^{inc}(t) = V_0 e^{-(t-t_p)^2/\sigma^2}, \tag{14}$$

where $\sigma = 1.9099 \times 10^{-9}s$, $t_p = 1.1459 \times 10^{-8}s$, and $V_0 = 1V$. All results in this paper are compared to Inverse Fourier Transformed (IFT) frequency domain data that are obtained from a frequency domain Method of Moments (MOM) code that uses spatial basis functions identical to those used in the MOT scheme. For the given problem, the current at the surface-wire junction computed by the MOT algorithm is shown in Fig. 4(a). The agreement between the MOT result and the IFT data is very good. Magnitude of the same surface-wire junction current for the first 10000 time steps is plotted in logarithmic scale in Fig. 4(b). Clearly, the algorithm does not exhibit any late time instabilities. Once all the currents for all time steps are computed, the algorithm calculates the transient radiated far fields according to the procedure outlined in [1]. As an example, the $\hat{\theta}$ directed electric field in the $\theta = 45°, \phi = 0°$ direction is depicted in Fig .5.

The second example consists of two monopole antennas mounted on the square plate of the first example (Fig. 6(a)). The circuit diagram of the feed network is shown in Fig. 6(b). The time dependence of the voltage source is again given by Eqn. (14). The computed transient current at the junction of the shorter monopole and the plate is plotted in Fig. 7. The corresponding IFT results are obtained by post-processing the MOM results at each frequency by a simple frequency domain transmission line network analysis code. Again, the agreement between the two data is very good. Finally, the $\hat{\theta}$ directed far zone electric field in the $\theta = 45°, \phi = 0°$ direction is plotted in Fig. 8.

4 Conclusions

An MOT algorithm has been described for analyzing the transient radiation characteristics of wire antenna arrays mounted on three-dimensional perfectly conducting bodies and excited by a feed network. The accuracy and stability of the algorithm have been demonstrated through numerical examples. Although examples presented in this paper only consist of straight wire antennas mounted on open surfaces, the proposed method can effectively be applied to transient analysis of any arbitrarily shaped wire antenna arrays mounted on either open or closed conducting surfaces.

References

[1] S. M. Rao and D. R. Wilton, "Transient scattering by conducting surfaces of arbitrary shape," *IEEE Trans. Antennas and Propagation*, vol. 39, pp. 56-61, Jan. 1991.

[2] D. A. Vechinski and S. M. Rao, "Transient scattering from three dimensional arbitrarily shaped dielectric bodies," *J. Opt. Soc. Am. A*, vol. 11, no. 4,pp. 1458-1470, Apr. 1994.

[3] B. P. Rynne "Time domain scattering from arbitrary surfaces using the electric field integral equation," *J. Electromagn. Waves Appl.*, vol. 5, no. 1, pp. 93-112, Jan. 1991.

[4] S. P. Walker "Developments in Time-Domain Integral-Equation Modeling at Imperial College," *IEEE Antennas and Propagation Mag.*, vol. 39, pp.7-19, Feb.1997

[5] A. G. Tijhuis, Z. Q. Peng and A. R. Bretones, "Transient excitation of a straight thin-wire segment: A new look to an old problem," *IEEE Trans. Antennas and Propagation*, vol. 40, pp. 1132-1146, Oct. 1992.

[6] A. Salinas, R. G. Martin, A. R. Bretones and I. S. Garcia, "Modeling of straight thin wires using time-domain electric field integral equations," *IEE Proc.-Microwav. Antennas Propag.*, vol. 141, pp. 123-126, Apr. 1994.

[7] J. G. Maloney, K. L. Shlager and G .S. Smith, "A simple FDTD model for transient excitation of antennas by transmission lines," *IEEE Trans. Antennas and Propagation*, vol. 42, pp. 289-292, Feb. 1994.

[8] S. M. Rao, D. R. Wilton and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas and Propagation*, vol. AP-30, pp. 409-419, May 1982.

[9] M. F. Costa and R. F. Harrington, "Electromagnetic radiation and scattering from a system of conducting bodies interconnected by wires," Syracuse University, Syracuse, New York, 1983.

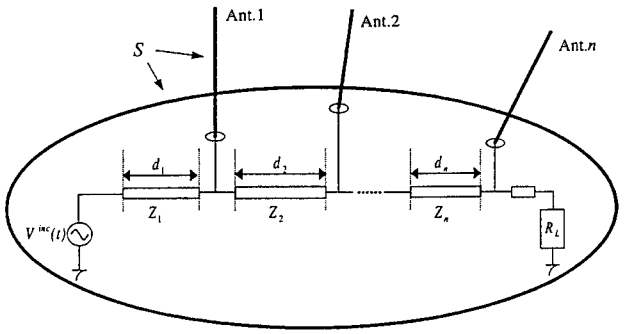


Figure 1: A schematic representation of the radiating structure, where d_i, Z_i for $i = 1, .., n$ are respectively the length and characteristic impedance of different segments of a feed line.

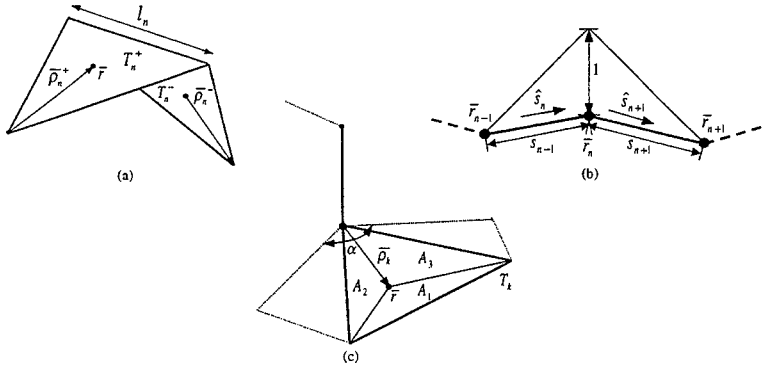


Figure 2: (a) Basis function for surface currents. (b) Basis function for wire currents. (c) Basis function for currents on surface–wire junctions.

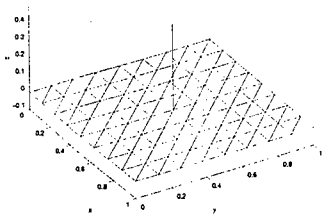
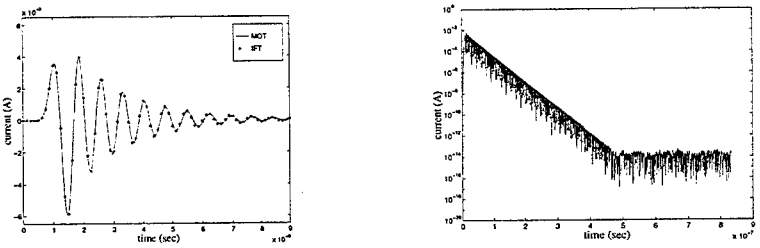
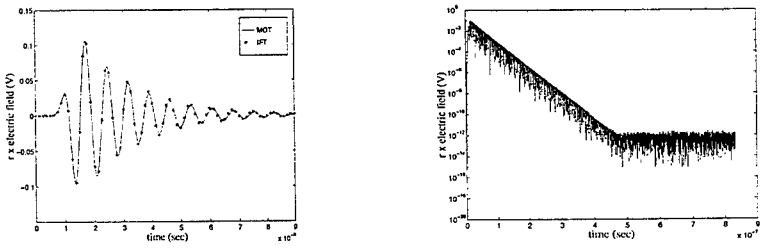


Figure 3: A $0.5m$ length monopole antenna mounted on a square conducting plate.



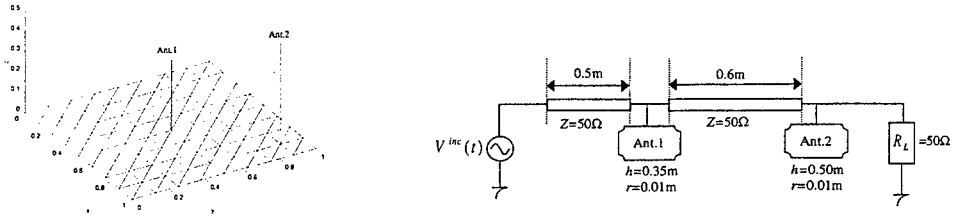
(a) (b)

Figure 4: (a) Transient current at the junction where the monopole is connected to the surface. (b) Magnitude of the current for the first 10000 time steps.



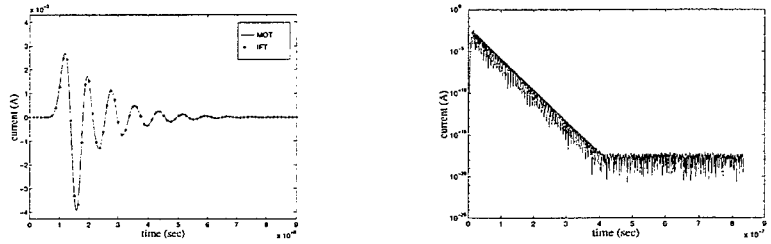
(a) (b)

Figure 5: (a) The $\hat{\theta}$ directed transient radiated electric field in the $\theta = 45°, \phi = 0°$ direction. (b) Magnitude of the electric field for the first 10000 time steps.



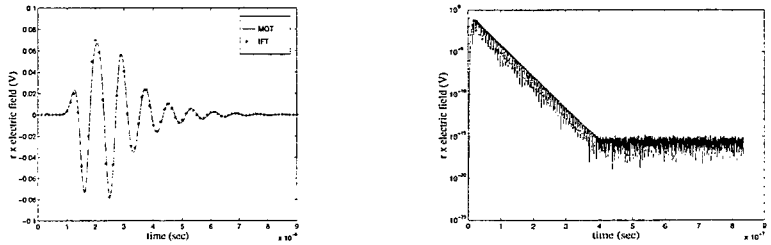
(a) (b)

Figure 6: (a) Two monopole antennas mounted on a square conducting plate. (b) Circuit diagram of the feed network, where h and r are the length and radius of the corresponding wire antenna.



(a) (b)

Figure 7: (a) Transient current at the junction where the shorter monopole is connected to the surface. (b) Magnitude of the current for the first 10000 time steps.



(a) (b)

Figure 8: (a) The $\hat{\theta}$ directed transient radiated electric field in the $\theta = 45°, \phi = 0°$ direction. (b) Magnitude of the electric field for the first 10000 time steps.

A PSTD Algorithm in Cylindrical Coordinates

Q. H. Liu and J. Q. He
Klipsch School of Electrical and Computer Engineering
New Mexico State University
Las Cruces, NM 88003

Abstract

The finite-difference time-domain (FDTD) method is a powerful numerical technique for transient solutions of electromagnetic waves. When applied to cylindrical coordinates in a straightforward way, however, it is limited by the contradictory requirements for accuracy and for numerical stability. These limitations arise because of the nonuniform distribution of cells in the computational domain. Moreover, the staggered grid encounters a singularity problem at the origin. We proposed a new pseudospectral time-domain (PSTD) method for the solution of Maxwell's equations in cylindrical coordinates. It eliminates the singularity problem by using a centered grid. Because of its high accuracy in the spatial derivatives, the PSTD method can employ a much larger cell and time step, making the algorithm far more efficient than the FDTD method.

I. Introduction

The finite-difference time-domain (FDTD) method has been enjoying its widespread applications in the simulations of transient electromagnetic wave propagation and scattering since it was first proposed by Yee in 1966 [1].

However, as the available computer memory and computational speed grow rapidly so that unprecedented large-scale problems can be solved, the FDTD method starts to show its limitation because of its relative large phase dispersion error. As the problem size increases, so does the number of unknowns per wavelength. For example, the standard finite-difference time-domain (FDTD) method requires a *grid density* (number of nodes per minimum wavelength λ) of 10–20 even for a problem of moderate size. For a large-scale problem of 512λ in each direction, for example, a grid density of at least 64 is required in order for the FDTD method to reach an accuracy of about 2%. As a result, with the conventional FDTD method, a large-scale 3-D problem of size $128\lambda \times 128\lambda \times 128\lambda$ requires more than 1.67×10^{10} nodes if a modest grid density of 20 is used. This problem is apparently still beyond the reach of the most powerful supercomputers.

In cylindrical coordinates, the conventional FDTD method encounters yet two more difficulties: (i) the requirement for a very small Δt because of the high concentration of cells near the z axis, and (ii) the singularity at the z axis. Although various remedies have been proposed, the treatment is not straightforward, and requires extra manipulations and computation time.

In this work we propose a pseudospectral time-domain (PSTD) method for 3-D cylindrical and 2-D polar coordinates. Similar to the PSTD algorithm for Cartesian coordinates, it uses the fast Fourier transform (FFT) to represent spatial derivatives, and the PML to remove the wraparound effect in the FFT computation of the non-periodic problem. The PML is based on two different formulations, i.e., the improved PML scheme of complex coordinates formulation, and the simplified quasi-PML formulation. Compared to its Cartesian counterpart, the cylindrical PSTD algorithm is special in that it requires a delicate treatment in the radial direction, as discussed later. The azimuthal direction, on the other hand, is simpler since the problem is naturally periodic in this direction.

Section II first summarizes the equations for the quasi-PML and improved PML using the complex coordinates [2–6]. Then the PSTD algorithm is presented to treat the derivatives in radial and azimuthal directions. Several numerical examples are shown in Section III to demonstrate the efficacy of the cylindrical PSTD algorithm.

II. Formulation

Consider an isotropic, inhomogeneous medium with space-dependent electric permittivity $\epsilon(\mathbf{r})$, magnetic

permeability $\mu(\mathbf{r})$, and conductivity $\sigma(\mathbf{r})$. Maxwell's curl equations governing electromagnetic fields in the medium are given by

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} - \mathbf{M}, \tag{1}$$

$$\nabla \times \mathbf{H} = \epsilon \frac{\partial \mathbf{E}}{\partial t} + \sigma \mathbf{E} + \mathbf{J}, \tag{2}$$

where $\mathbf{J}$ and $\mathbf{M}$ are the imposed electric and magnetic current densities, respectively. Our aim is to solve these two equations in cylindrical coordinates with a new pseudospectral time-domain (PSTD) method [7–9]. The PSTD algorithm will use the fast Fourier transform (FFT) for spatial derivatives, and the cylindrical PML presented below to remove the wraparound effect.

A. Quasi-PML for Cylindrical Coordinates

Using a unified formulation [6], we can derive equations for quasi-PML and true PML. For the quasi-PML formulation, it can be shown that the time-domain split equations in cylindrical coordinates are

$$a_\rho \epsilon \frac{\partial E_\rho^{(\phi)}}{\partial t} + (\omega_\rho \epsilon + a_\rho \sigma) E_\rho^{(\phi)} + \omega_\rho \sigma \int_{-\infty}^{t} E_\rho^{(\phi)}(\tau) d\tau = \frac{\partial H_z}{\partial \phi} - J_\rho^{(\phi)}, \tag{3}$$

$$a_z \epsilon \frac{\partial E_\rho^{(z)}}{\partial t} + (\omega_z \epsilon + a_z \sigma) E_\rho^{(z)} + \omega_z \sigma \int_{-\infty}^{t} E_\rho^{(z)}(\tau) d\tau = -\frac{\partial H_\phi}{\partial z} - J_\rho^{(z)}, \tag{4}$$

$$a_\rho \epsilon \frac{\partial E_\phi^{(\rho)}}{\partial t} + (\omega_\rho \epsilon + a_\rho \sigma) E_\phi^{(\rho)} + \omega_\rho \sigma \int_{-\infty}^{t} E_\phi^{(\rho)}(\tau) d\tau = -\frac{\partial H_z}{\partial \rho} - J_\phi^{(\rho)}, \tag{5}$$

$$a_z \epsilon \frac{\partial E_\phi^{(z)}}{\partial t} + (\omega_z \epsilon + a_z \sigma) E_\phi^{(z)} + \omega_z \sigma \int_{-\infty}^{t} E_\phi^{(z)}(\tau) d\tau = \frac{\partial H_\rho}{\partial z} - J_\phi^{(z)}, \tag{6}$$

$$a_\rho \epsilon \frac{\partial E_z}{\partial t} + (\omega_\rho + a_\rho \sigma) E_z + \omega_\rho \sigma \int_{-\infty}^{t} E_z(\tau) d\tau = \frac{1}{\rho} \frac{\partial (\rho H_\phi)}{\partial \rho} - \frac{\partial H_\rho}{\partial \phi} - J_z, \tag{7}$$

In the above, the split field components are $E_\rho = E_\rho^{(\phi)} + E_\rho^{(z)}$, and $E_\phi = E_\phi^{(\rho)} + E_\phi^{(z)}$. The other set of equations for updating $\mathbf{H}$ can be obtained by duality. Note that in the quasi-PML formulation, there is no need to split E_z and H_z. Therefore, the total number of unknown field components is 10 instead of 12 as for the true PML presented below.

B. An Improved PML for Conductive Media

For 3-D cylindrical coordinates (ρ, ϕ, z), the extension of the improved PML [3, 6] is straightforward since the z direction is the same as for the Cartesian coordinates. Furthermore, the extension to conductive media can follow the same procedures as in [10]. Therefore, based on the improved PML formulation, the time-domain split equations for conductive media can be derived as

$$A_\rho \epsilon \frac{\partial E_\rho^{(\phi)}}{\partial t} + (\Omega_\rho \epsilon + A_\rho \sigma) E_\rho^{(\phi)} + \Omega_\rho \sigma \int_{-\infty}^{t} E_\rho^{(\phi)}(\tau) d\tau = \frac{\partial H_z}{\partial \phi} - J_\rho^{(\phi)}, \tag{8}$$

$$a_z \epsilon \frac{\partial E_\rho^{(z)}}{\partial t} + (\omega_z \epsilon + a_z \sigma) E_\rho^{(z)} + \omega_z \sigma \int_{-\infty}^{t} E_\rho^{(z)}(\tau) d\tau = -\frac{\partial H_\phi}{\partial z} - J_\rho^{(z)}, \tag{9}$$

$$a_\rho \epsilon \frac{\partial E_\phi^{(\rho)}}{\partial t} + (\omega_\rho \epsilon + a_\rho \sigma) E_\phi^{(\rho)} + \omega_\rho \sigma \int_{-\infty}^{t} E_\phi^{(\rho)}(\tau) d\tau = -\frac{\partial H_z}{\partial \rho} - J_\phi^{(\rho)}, \tag{10}$$

$$a_z\epsilon\frac{\partial E_\phi^{(z)}}{\partial t}+(\omega_z\epsilon+a_z\sigma)E_\phi^{(z)}+\omega_z\sigma\int_{-\infty}^{t}E_\phi^{(z)}(\tau)d\tau=\frac{\partial H_\rho}{\partial z}-J_\phi^{(z)}, \tag{11}$$

$$a_\rho\epsilon\frac{\partial E_z^{(\rho)}}{\partial t}+(\omega_\rho+a_\rho\sigma)E_z^{(\rho)}+\omega_\rho\sigma\int_{-\infty}^{t}E_z^{(\rho)}(\tau)d\tau=\frac{\partial(H_\phi)}{\partial\rho}-J_z^{(\rho)}, \tag{12}$$

$$A_\rho\epsilon\frac{\partial E_z^{(\phi)}}{\partial t}+(\Omega_\rho\epsilon+A_\rho\sigma)E_z^{(\phi)}+\Omega_\rho\sigma\int_{-\infty}^{t}E_z^{(\phi)}(\tau)d\tau=H_\phi-\frac{\partial H_\rho}{\partial\phi}-J_z^{(\phi)}. \tag{13}$$

The other set of equations for updating **H** can be obtained by duality. The total number of unknown field components is 12 for this improved PML formulation.

C. The PSTD Algorithm

Since the treatment in z direction in the PSTD algorithm is exactly the same as in Cartesian coordinates, we refer the reader to [7–9] for all aspects of PSTD except for the treatment in ρ and ϕ directions. Hence, below we will discuss on the spatial derivatives of a function $f(\rho,\phi)$.

Instead of using a staggered grid as in the FDTD method, we use a centered grid where all field components are located at the center of each cell. Therefore, if the (ρ,ϕ) cross section $(\rho,\phi)\in\{[0,\rho_{max}]\times[0,2\pi]\}$ is discretized by $N_\rho\times N_\phi$ uniform cells, any field component $f(\rho,\phi)$ is defined at $f[(j_\rho+1/2)\Delta\rho,(j_\phi+1/2)\Delta\phi]\equiv f(j_\rho,j_\phi)$, where $\Delta\rho=\rho_{max}/N_\rho$, $\Delta\phi=2\pi/N_\phi$, $j_\rho=0,\cdots,N_\rho-1$, and $j_\phi=0,\cdots,N_\phi-1$. The first benefit of this centered grid is the removal of the singularity at z axis present in the staggered grid.

In the PSTD algorithm, the spatial derivative $\partial/\partial\phi$ is easily obtained by FFT since there is a natural periodicity in the ϕ direction. Using the discrete Fourier transform (DFT), we can represent $\partial f(j_\rho,j_\phi)/\partial\phi$ as

$$\left[\frac{\partial f(j_\rho,j_\phi)}{\partial\phi}\right]=\frac{1}{2\pi}\sum_{m_\phi=-N_\phi/2}^{N_\phi/2-1}ik_\phi\tilde{f}(j_\rho,m_\phi)e^{ik_\phi\phi_j}, \tag{14}$$

where $k_\phi=m_\phi$, and $\tilde{f}(j_\rho,m_\phi)$ is the Fourier series

$$\tilde{f}(j_\rho,m_\phi)=\Delta\phi\sum_{j_\phi=0}^{N_\phi-1}f(j_\rho,j_\phi)e^{-ik_\phi\phi_j}. \tag{15}$$

This representation is exact up to the spatial Nyquist sampling frequency. Note that the forward and inverse DFT's in (15) and (14) can be achieved by the fast Fourier transform algorithms with a number of arithmetic operations $O(N_\phi\log_2 N_\phi)$.

The treatment of the ρ derivative is more complicated compared to the Cartesian coordinates, simply because that the boundary at $\rho=0$ is not an open boundary. One way of treating this is to use Chebyshev pseudospectral method [11–12] which inevitably increases the number of nodes at $\rho=0$ and has a stringent stability criterion for Δt. Below we present two ways to use the Fourier series for ρ derivatives.

(a) The asymmetric form of PSTD algorithm in ρ direction

The most straightforward way to approximate the ρ derivative $\partial f(j_\rho,j_\phi)/\partial\phi$ is

$$\frac{\partial f}{\partial\rho}\approx\mathcal{D}_\rho f=\mathcal{F}_\rho^{-1}\Big\{ik_\rho\mathcal{F}_\rho[f]\Big\}, \tag{16}$$

where $\mathcal{F}_\rho$ and $\mathcal{F}_\rho^{-1}$ denote the forward and inverse FFT in ρ direction. Since $\rho=0$ is a physical boundary, PML cells have to be placed near the outer boundary $\rho=\rho_{max}$ to remove the wraparound effect due to the periodicity of the DFT.

There are two major disadvantages associated with this approach: (i) More PML cells (usually around 20) are required near the outer boundary since the periodicity applies here (in contrast to a perfect electric conductor for the FDTD). Half of the PML cells have an increasing profile, and the other half have a decreasing profile, i.e.,

$$\omega_\rho(j) = \omega_{\rho,max}\Big(1 - \frac{|N_\rho - M/2 - 1/2 - j|}{M/2}\Big)^p, \qquad (j = N_\rho - M, \cdots, N_\rho - 1), \tag{17}$$

where M is the number of PML cells, p is the order of the PML profile, and $\omega_{\rho,max}$ is the maximum value of ω_ρ. (ii) Because of the periodicity, the negligibly small field at ρ_{max} (due to the PML attenuation) imposes a null-field condition at $\rho = 0$, effectively creating a small ghost source at the z axis. As observed from numerical experiments, this ghost source, although small, produces noticeable spurious fields.

(b) The symmetric form of PSTD algorithm in ρ direction

A much better way to treat the ρ derivatives is to use the symmetric form by assigning a new function for $0 \le j_\phi \le N_\phi/2 - 1$ (assuming N_ϕ is even) such that

$$g(j', j_\phi) = \begin{cases} f(-j' - 1, j_\phi + N_\phi/2), & \text{for } j' = -N_\rho, \cdots, -1 \\ f(j', j_\phi), & \text{for } j' = 0, \cdots, N_\rho - 1 \end{cases} \tag{18}$$

Then the derivative is found by the FFT of these $N_\phi/2$ new arrays of length $2N_\rho$, in a way similar to (16). The total computation burden is reduced from (a) because only half the PML cells are needed. With this approach, both disadvantages in (a) have been removed.

III. Numerical Results

We have implemented the PSTD method for 3-D cylindrical coordinates as well as 2-D polar coordinates for conductive media. Figure 1 shows an example of a line source in a 2-D free space. The source has a Blackman-Harris window time function with center frequency $f_c = 300$ MHz, and is located at $\rho_s = 1.5$ m, $\phi_s = 87.19°$. The computational domain is meshed by $N_\rho \times N_\phi = 32 \times 64$ cells with $\Delta\rho = 0.2$ m (or about 2 cells per wavelength at the frequency $2.5f_c$) and $\Delta t = 12.5$ ps. The snapshots show the effectiveness of the 10-layer PML ABC, while the last sub-plot shows the excellent agreement between the PSTD result and the analytical solution.

For the PSTD code to solve this problem on a SUN Ultra 1 workstation, it takes 140 seconds for the required 4000 time steps. For an acceptable accuracy, the FDTD method needs $N_\rho \times N_\phi = 128 \times 256$ cells, requiring 16 times more computer memory. In addition, a much smaller time step $\Delta t = 1.25$ ps has to be chosen for stability, requiring a total 40,000 time steps for the same problem. As a result, the FDTD code takes about 7 hours CPU time to complete this problem, or roughly 180 times slower.

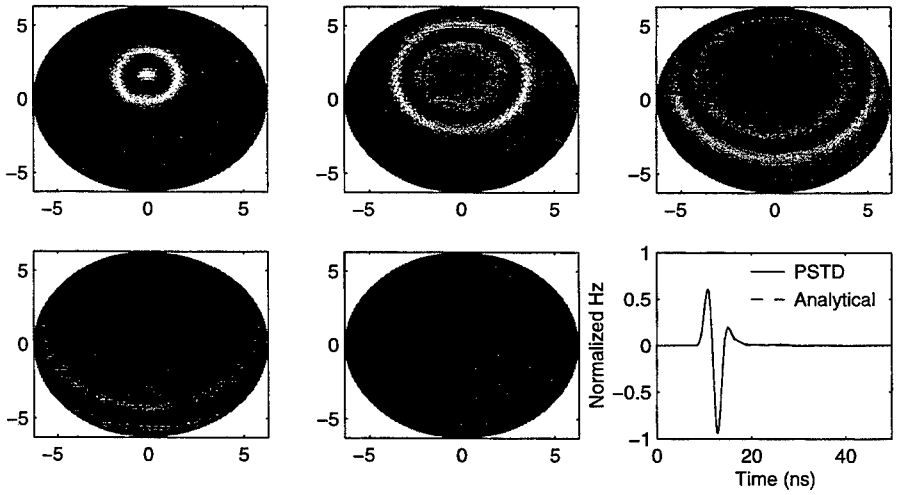


Figure 1. From first to fifth sub-plots, snapshots at time steps $n = 500$, 1000, 1500, 2000, and 2500 ($\Delta t = 12.5$ ps). The last plot compares the PSTD result with the analytical solution at $\rho = 3.1$ m, $\phi = 154.69°$. The source is located at $\rho_s = 1.5$ m, $\phi_s = 87.19°$.

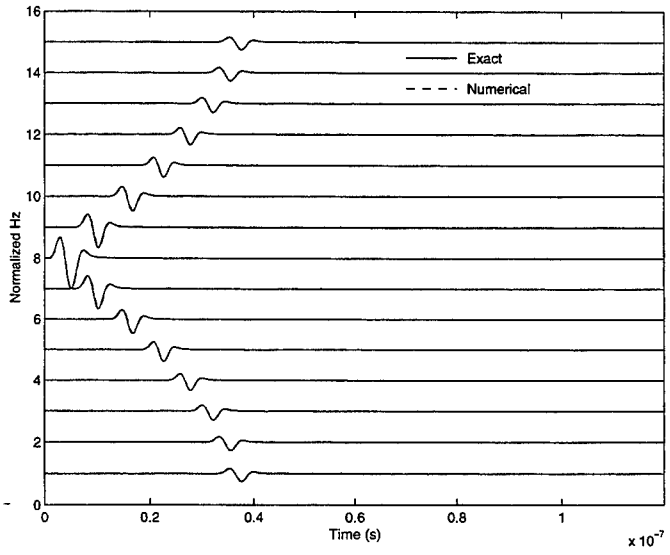


Figure 2. An array waveforms in free space.

We simulate the same source in an even larger problem. The center of source is located at $(\rho, \phi) = (128, 35)$ cells in a computational domain of $N_\rho \times N_\phi = 64 \times 256$ cells ($\rho_{max} = 10$ m). Fifteen receivers are set uniformly around a circle 30 cells away from the origin, and are 16 cells apart in ϕ direction. The first receiver is located at (30, 16). The numerical results agree well with analytical solutions, as shown in Figure 2.

Conclusions

We have developed a pseudospectral time-domain method for 3-D cylindrical and 2-D polar coordinates. It uses the FFT to represent spatial derivatives and the PML to remove the wraparound effect. Excellent agreement between the PSTD results and analytical solutions has been observed even when the sampling is at the Nyquist frequency. Compared with the conventional FDTD method, the PSTD method has the following advantages:

(1) There is no dispersion error related to the spatial derivatives.

(2) The only dispersion error due to temporal derivatives in PSTD is isotropic.

(3) Only 2 nodes/λ are required regardless of the electrical size of the problem.

(4) There is no need for material averaging because of the use of the centered grid.

These advantages are common with those in Cartesian coordinates. For cylindrical coordinates, the additional advantages are:

(5) The singularity at $\rho = 0$ is no longer present.

(6) The largest benefit is that the required number of time steps is reduced by a factor of K^2, where K is the ratio of $\Delta\rho$ in PSTD and in FDTD. For the example shown, $K = 4$; It increases with the electrical size of the problem.

The PSTD algorithm is therefore ideal for large-scale problems. In the near future, we hope to report on our investigation to optimize the PSTD method for problems with mixed scales (large structures with fine details) as well as problems with discontinuous tangential field components such as those near a metallic surface.

Acknowledgment

This work was supported by Environmental Protection Agency through at PECASE grant CR-825-225-010, and by the National Science Foundation through a CAREER grant ECS-9702195.

References

[1] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, AP-14, pp. 302–307, 1966.

[2] W. C. Chew, J. M. Jin, and E. Michelssen, "Complex coordinate system as a generalized absorbing boundary condition," *Intl. IEEE AP-S Symposium Digest*, Montreal, Canada, July 1997.

[3] F. L. Teixeira and W. C. Chew, "Perfectly matched layer in cylindrical coordinates," *Intl. IEEE AP-S Symposium Digest*, Montreal, Canada, July 1997.

[4] Q. H. Liu and J. Q. He, "Quasi-PML for waves in cylindrical coordinates," *New Mexico State University Technical Report*, July 1997.

[5] B. Yang, D. Gottlieb, and J. S. Hesthaven, "Spectral simulations of electromagnetic wave scattering," *J. Comput. Phys.*, vol. 134, pp. 216–230, 1997.

[6] J. Q. He, and Q. H. Liu, "A systematic study of three PML absorbing boundary conditions through a unified formulation in cylindrical coordinates," this conference.

[7] Q. H. Liu, "A spectral-domain method with perfectly matched layers for time-domain solutions of Maxwell's equations," 1996 URSI Meeting, Baltimore, MD, July 1996.

[8] Q. H. Liu, "The PSTD algorithm: a time-domain method requiring only two cells per wavelength," *Microwave Opt. Tech. Lett.*, vol. 10, pp. 158–165, 1997.

[9] Q. H. Liu, "The pseudospectral time-domain (PSTD) method: a new algorithm for solutions of Maxwell's equations," *Intl. IEEE/AP-S Symp. Digest*, Montreal, 1997.

[10] Q. H. Liu, "An FDTD algorithm with perfectly matched layers for conductive media," *Microwave Opt. Tech. Lett.*, **14** (2), 134–137 (1997).

[11] H.-O. Kreiss, and J. Oliger, "Comparison of accurate methods for the integration of hyperbolic equations," *Tellus*, vol. 24, pp. 199–215, 1972.

[12] S. A. Orszag, "Comparison of pseudospectral and spectral approximation," *Stud. Appl. Math.* , vol. 51, pp. 253–259, 1972.

ON THE PSTD METHOD FOR LARGE-SCALE PROBLEMS

Q. H. Liu
Klipsch School of Electrical and Computer Engineering
New Mexico State University
Las Cruces, NM 88003

Abstract

Conventional finite-difference time-domain (FDTD) methods require a large number of unknowns, typically 10–20 nodes per minimum wavelength λ for a problem of medium size. This work makes a theoretical comparison of the pseudospectral time-domain (PSTD) method with the FDTD and MRTD (multi-resolution time-domain) methods. The new PSTD algorithm is based on the fast Fourier transforms and perfectly matched layers. It significantly reduces the number of unknowns to 2 nodes/λ. The method is demonstrated by a three-dimensional problem with an apparently unprecedented size of $128\lambda \times 128\lambda \times 128\lambda$.

I. Introduction

Simulation of electromagnetic wave propagation in large-scale problems is a great challenge to full-wave numerical methods because of the large number of unknowns required. For example, the standard finite-difference time-domain (FDTD) method [1–3] requires a *grid density* (number of nodes per minimum wavelength λ) of 10–20 even for a problem of moderate size. For problems of large scales, the grid density has to increase in order to produce accurate results since the dispersion error accumulates rapidly as the problem size increases. As a result, with the conventional FDTD method, a large-scale 3-D problem of size $128\lambda \times 128\lambda \times 128\lambda$ requires more than 1.67×10^{10} nodes if a modest grid density of 20 is used. Hence, this problem is apparently still beyond the reach of the most powerful supercomputers.

To increase the efficiency and reduce the computer memory requirement, higher order FDTD methods can be used. Alternatively, the multi-resolution time-domain (MRTD) method [4, 5], and spectral-domain methods such as the Fourier method [6] and generalized k-space method [7] have also been developed to reduce the grid density to or close to the Nyquist sampling rate. Unfortunately, the spectral-domain methods in [6, 7] are only valid for spatially-periodic problems.

Recently a pseudospectral time-domain (PSTD) method has been proposed to reduce the grid density to 2 by using the combination of the fast Fourier transform (FFT) algo-

rithm and the perfectly matched layer (PML) [8–9]. The method has been validated for multidimensional inhomogeneous media. In this work, we compare the PSTD algorithm with the FDTD and MRTD methods for the accuracy and efficiency. A large-scale problem of size $128\lambda \times 128\lambda \times 128\lambda$ is shown to demonstrate the power of this new method.

II. Comparison of Methods

Using PML as the absorbing boundary condition, the split Ampere's law, for example, can be written as [9]

$$a_\eta \mu \frac{\partial \mathbf{H}^{(\eta)}}{\partial t} + \mu \omega_\eta \mathbf{H}^{(\eta)} = -\frac{\partial}{\partial \eta}(\hat{\eta} \times \mathbf{E}), \tag{1}$$

where $\eta = x, y, z$ and $\mathbf{H} = \sum_{\eta=x,y,z} \mathbf{H}^{(\eta)}$. The differences among FDTD, MRTD, and PSTD methods are in the approximation of the spatial derivatives

$$\frac{\partial f(\eta)}{\partial \eta} \approx \mathcal{D}_\eta f(\eta), \tag{2}$$

with

$$\mathcal{D}_\eta f(\eta) = \begin{cases} \mathcal{F}_\eta^{-1}\Big\{ik_\eta \mathcal{F}_\eta[f(\eta)]\Big\}, & \text{for PSTD,} \\ \dfrac{1}{\Delta\eta}\displaystyle\sum_{j=1}^{P/2} a_j[f(\eta + j\Delta\eta/2) - f(\eta - j\Delta\eta/2)], & \text{for FDTD/MRTD,} \end{cases} \tag{3}$$

where $P = 2$ and $a_1 = 1$ for FDTD, $P = 18$ and a_j for MRTD are given in [4], and $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the forward and inverse Fourier transforms which are achieved by an FFT algorithm [9]. Under the assumption that $\Delta x = \Delta y = \Delta z$, the dispersion relations in these methods for a plane wave in a homogeneous medium are

$$\sin\frac{\omega\Delta t}{2} = \begin{cases} \dfrac{c\Delta t}{2}\sqrt{k_x^2 + k_y^2 + k_z^2}, & \text{for PSTD;} \\ \dfrac{c\Delta t}{\Delta x}\sqrt{\displaystyle\sum_{\eta=x,y,z}\Big\{\sum_{j=1}^{P/2} a_j \sin^2[k_\eta \Delta\eta(j-1/2)]\Big\}^2}, & \text{for FDTD/MRTD.} \end{cases} \tag{4}$$

Compared with FDTD and MRTD, the PSTD method has the following advantages:

(1) There is no dispersion error related to the spatial derivatives. The only dispersion error is caused by the second-order finite difference approximation in temporal derivatives. This error can always be reduced by using a smaller Δt.

(2) The dispersion error in PSTD is isotropic instead of anisotropic as in other methods.

(3) Since there is no spatial dispersion error in the PSTD method, only 2 nodes/λ are required regardless of the electrical size of the problem. As a result, orders of magnitude saving in computer memory and computation time can be achieved.

(4) The reflection from the PML ABC can be made much smaller in the PSTD method since the larger cell allows a smaller PML attenuation coefficient.

(5) Instead of a staggered grid, a centered grid is used in PSTD. This eliminates the need for material averaging near material discontinuities which distorts the original medium.

The stability criterion of these three methods can be written as $c\Delta t/\Delta x \leq 1/(\alpha\sqrt{3})$ for 3-D problems, where $\alpha = 1$ for FDTD, $\alpha = 1.5684$ for MRTD, and $\alpha = 1.5708$ for PSTD.

III. Numerical Results

Fig. 1(a) compares the dispersion relations in the three methods with the exact dispersion relation for a one-dimensional problem. The relative error in the normalized wavenumber $K = k\lambda$ is shown as a function of the normalized frequency $W = \omega\lambda/c$ (where λ is the minimum wavelength corresponding to the highest frequency). The grid density is 32 for the FDTD algorithm, and 2 for the MRTD and PSTD algorithms. A small Δt is chosen so that the stability condition is satisfied for all algorithms. It is observed that the PSTD algorithm has the highest accuracy in dispersion. The small dispersion error comes from the approximation in temporal derivatives. Indeed, the PSTD algorithm may be considered an optimal time-domain solution in the sense that it requires the minimum spatial sampling rate while maintaining the highest accuracy in its dispersion relation.

This advantage of the PSTD algorithm is important for simulating large-scale problems. As an example, the propagation of electromagnetic waves in a three-layer one-dimensional problem is simulated. An air layer is surrounded by a dielectric background with $\epsilon_r = 4$, and the layer interfaces are at $x = 3$ and $x = 6$ m. A source exciting E_y is located at $x = 9.6$ m, and the propagation of waves is simulated over a long distance from $x = 0$ to $x = 153.6$ m (or about 512λ). Figs. 1(b) and 1(c) show that with a receiver at $x = 150.0$ m (or 468λ away from the source), the FDTD algorithm has a large dispersion error of 8.5% even with a grid density of 32. The PSTD algorithm is accurate to within 0.8% even with a grid density of 2. A similar numerical experiment for the MRTD for a propagation distance of 150λ in a homogeneous medium with a grid density of 2 shows that the dispersion error is up to 15.0%, as in Fig. 1(d). Other examples have been shown

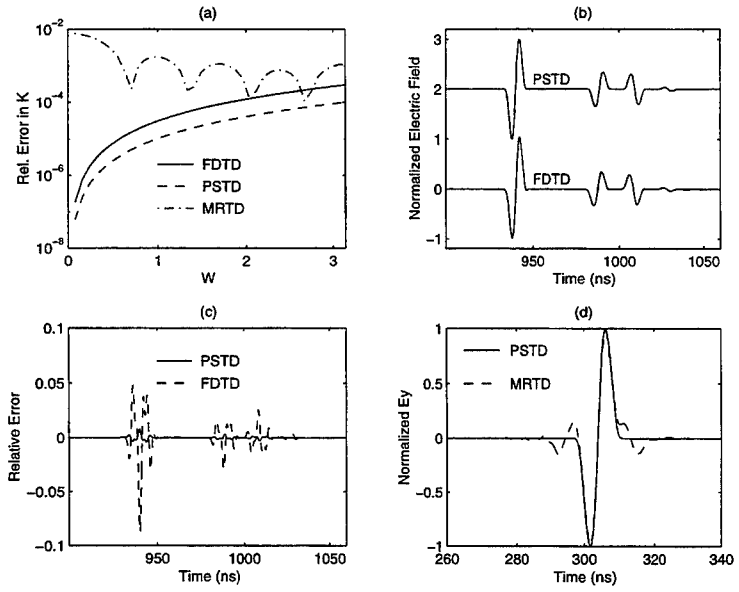


Fig. 1. (a) Relative dispersion errors in K for the FDTD, MRTD, and PSTD algorithms. (b) PSTD (2 nodes/λ) and FDTD (32 nodes/λ) results for E_y at 468λ away for the source in a 3-layer medium. (c) Relative numerical errors in (b). (d) PSTD and MRTD results (2 nodes/λ) for E_y at 150λ away from the source in a homogeneous medium.

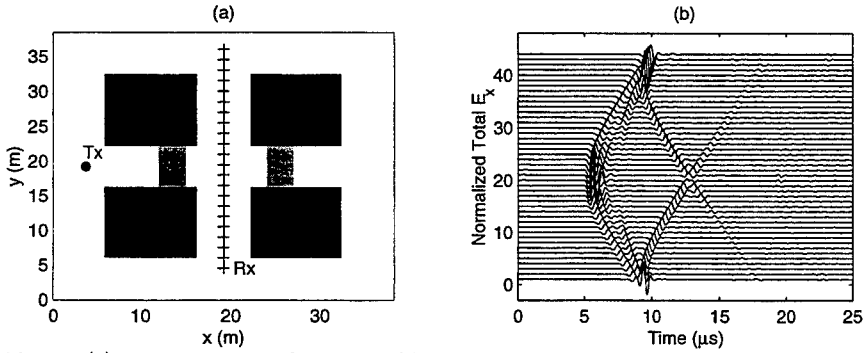


Fig. 2. (a) x-y projection of a 3-D problem with 4 buildings above ground and 2 objects underground. (b) E_z waveforms at the receiver array.

in [9] to validate the PSTD method for multidimensional inhomogeneous media.

To illustrate the applications of the PSTD algorithm to large-scale problems, Fig. 2(a) shows a simple 3-D problem for the study of electromagnetic wave propagation in an urban environment. The dielectric constant is 4 for the earth, and 2 for the four buildings and two underground objects. A vertical electric dipole operates at a central frequency of 166.67 MHz with a Blackman-Harris window time function. At the highest frequency 500 MHz, λ is 0.3 m for $\epsilon_r = 4$, and the problem size is $128\lambda \times 128\lambda \times 128\lambda$. It is simulated by the PSTD with $256 \times 256 \times 256$ nodes. The measured E_z at the receiver array is shown in Fig. 2(b).

This large-scale problem requires 1.359 G-bytes of memory, and 16.36 seconds of CPU time per time step on a 8-CPU HP SPP-2000 Exemplar. In comparison, if this problem is to be solved by the FDTD method with 20 nodes/λ, it would require 1000 times more memory and roughly the same factor more CPU time.

Conclusions

The newly developed pseudospectral time-domain (PSTD) algorithm is compared to the FDTD and MRTD methods. In terms of spatial sampling, the PSTD method can be considered an optimal algorithm since it requires only two nodes per minimum wavelength because the Fourier transform (through an FFT algorithm) is used to represent spatial derivatives exactly up to the Nyquist frequency. The method is applied to solve an apparently unprecedented large-scale problem of size $128\lambda \times 128\lambda \times 128\lambda$. Further investigation is under way to optimize the PSTD method for problems with mixed scales (large structures with fine details) as well as problems with discontinuous tangential field components such as those near a metallic surface.

Acknowledgment

The author is supported by Environmental Protection Agency under a PECASE grant CR-825-225-010, by a SURP grant from Sandia National Laboratories, and by National Science Foundation under a CAREER grant ECS-9702195. He is grateful to Martin Lewis of Hewlett-Packard Co. for the help in parallelizing the PSTD code.

References

[1] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, AP-14, pp. 302–307, 1966.

[2] K. S. Kunz, and R. J. Luebbers, *Finite Difference Time Domain Method for Electromagnetics*, CRC Press Inc., Florida, 1993.

[3] A. Taflove, *Computational Electrodynamics: The Finite Difference Time Domain Method*, Artech House, Inc., Norwood, MA, 1995.

[4] M. Krumpholz, and L. P. B. Katehi, "New prospects for time domain analysis," *IEEE Microwave Guided Wave Lett.*, vol. 5, pp. 382–384, 1995.

[5] M. Krumpholz, and L. P. B. Katehi, "MRTD: new time-domain schemes based on multiresolution analysis," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 555–571, 1996.

[6] S. A. Orszag, "Comparison of pseudospectral and spectral approximation," *Stud. Appl. Math.* , vol. 51, pp. 253–259, 1972.

[7] Q. H. Liu, "Transient electromagnetic modeling with the generalized k-space (GKS) method," *Microwave Opt. Tech. Lett.*, vol. 7, pp. 842–848, 1994.

[8] Q. H. Liu, "A spectral-domain method with perfectly matched layers for time-domain solutions of Maxwell's equations," 1996 URSI Meeting, Baltimore, MD, July 1996.

[9] Q. H. Liu, "The PSTD algorithm: a time-domain method requiring only two cells per wavelength," *Microwave Opt. Tech. Lett.*, vol. 10, pp. 158–165, 1997.

PSEUDOSPECTRAL TIME-DOMAIN MODELING OF DIFFRACTIVE OPTICAL ELEMENTS*

J. S. HESTHAVEN†
DIV. OF APPLIED MATHEMATICS
BROWN UNIVERSITY
PROVIDENCE, RI, USA

P. G. DINESEN AND J. P. LYNOV
OPTICS AND FLUID DYNAMICS DEPT.
RISø NATIONAL LABORATORY
ROSKILDE, DENMARK

Abstract. We develop a pseudospectral multi-domain formulation for the accurate modeling of generic diffractive optical elements, here exemplified by off-plane waveguide holograms for the coupling between guided waves and freely propagating wavefronts.

The individual elements entering the multi-domain formulation for the solution of the time-domain Maxwell equations is described, stressing the ability to accurately and efficiently handling very general geometric complexity and combinations of several materials.

The efficacy of the overall scheme is illustrated by computing the time-domain solution of a plane waveguide problem and an off-plane waveguide coupler.

1. Introduction. The well established and highly reliable fabrication techniques of in-plane semi-conductor laser has spawned intensive research into processes by which electromagnetic energy can be exchanged between the guided waves emerging from the laser and freely propagating wavefronts. Such waveguide couplers have been known for some time although their design have been fairly limited due to shortcomings in fabrication techniques. However, with the present day ability to reliably modify a waveguide surface with an accuracy of about 20 nm, using electron beam techniques and interference exposure processes, it is possible to fabricate very general wavefront modulators in the 1500 nm range of optical communication. The impact of such developments is potentially very large as it in principle allows for realizing integrated optical elements with properties similar to conventional holographic elements, i.e., multiple focal point or beam-shaping elements.

While the fabrication of such off-plane waveguide holograms has become possible, the actual specification of the surface modulation remains a very significant challenge. The analytic tools for problems involving modulations of order of the wavelength are clearly insufficient. However, also computational modeling, using conventional low order schemes, of such elements is nontrivial and in most cases not possible. Since the optical elements are characterized by being composed of several layers of complex materials and often spanning hundreds of free-space wavelengths, low order time-domain as well as frequency domain methods fail to accurately reproduce the details of the out-coupled wavefronts. The main reason for this shortcoming lies in the inability to accurately model the phase behavior which is critical for the correct computation of the out-coupled wavefronts.

The quest for accurate modeling of the phase properties of the waves over several hundred wavelengths suggests that the use of high-order methods, and in particular pseudospectral methods, is not only an option but a necessity as has been shown recently for problems involving scattering by electrically large general objects [1, 2]. These studies has lead us to develop a time-domain pseudospectral multi-domain scheme for the general two-dimensional forward diffraction problem sketched in Fig. 1. The modulation of the element/vacuum interface manipulate the guided wave and allows for a coupling of waveguide energy into free-space. The actual amplitude and phase modification of the scattered fields depends critically on the details of the modulation of the waveguide coupler, hence placing severe constraints on the properties of the numerical scheme. We emphasize that while we shall focus the attention on off-plane waveguide holograms, the computational framework developed here is applicable for the modeling of a much broader range of waveguide phenomena.

The remaining part of this paper is organized as follows. In Sec. 2 we discuss the problem from a electromagnetic point of view by introducing Maxwells equations, boundary and initial conditions. Section 3 is devoted to a detailed development of the various elements that enter the computational framework while Sec. 4 contains a number of test cases. In Sec. 5 we conclude with a few general remarks.

2. The Physical Picture. A typical off-plane waveguide hologram, as depicted in Fig. 1, consists of a number of layers of dielectric material and we shall subsequently assume that all materials can be considered lossless, homogeneous and non-magnetic. For the guided wave to exist we must assume that $1 \leq n_1 \leq n_3 < n_2$ and for simplicity we shall here assume that also $n_1 = 1$ and the modulation takes place directly in the waveguide rather

* The work of the first author was partially supported by DARPA/AFOSR grant F49620-96-1-0426, DOE grant DE-FG02-95ER25239 and NSF grant ASC-9504002.

† Corresponding Author: *jansh@cfm.brown.edu*

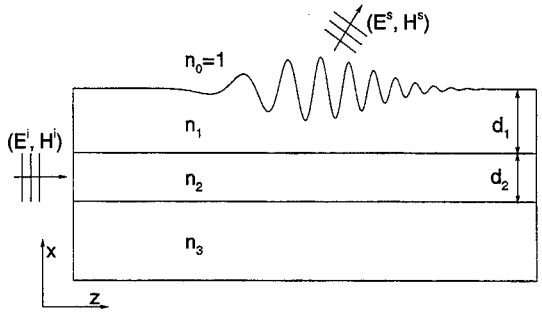


FIG. 1. *Typical configuration of a multi-layer diffractive optical element.*

than in the cladding as illustrated in Fig. 1. The width of the waveguide, d_2, determines whether the waveguide is a single- or a multi-mode waveguide and the width of bulk material, n_3, is assumed to be sufficiently large that the evanescent waves are undisturbed by the lower edge of the element.

We shall restrict the attention to the two-dimensional transverse electrical (TE) case for which Maxwells equations take the form

$$
\begin{aligned}
\frac{\partial \tilde{H}_z}{\partial \tilde{t}} &= -\frac{c}{Z_0}\frac{\partial \tilde{E}_y}{\partial \tilde{x}} \\
\frac{\partial \tilde{H}_x}{\partial \tilde{t}} &= \frac{c}{Z_0}\frac{\partial \tilde{E}_y}{\partial \tilde{z}} \\
\frac{\partial \tilde{E}_y}{\partial \tilde{t}} &= cZ_0\frac{1}{n^2}\left(\frac{\partial \tilde{H}_x}{\partial \tilde{z}} - \frac{\partial \tilde{H}_z}{\partial \tilde{x}}\right) ,
\end{aligned}
\tag{1}
$$

where $\tilde{H}_z$ and $\tilde{H}_x$ represent the dimensional magnetic fields in the plane while $\tilde{E}_y$ refers to the perpendicular component of the electric field. We have also introduced the free space impedance, $Z_0 = \sqrt{\mu_0/\varepsilon_0}$, and the vacuum speed of light, $c = 1/\sqrt{\varepsilon_0\mu_0}$, where ε_0 and μ_0 represents the free space permittivity and permeability, respectively. The index of refraction, $n(z,x)$, is related to the relative permittivity of the dielectric material as $\varepsilon = \varepsilon_r\varepsilon_0 = n^2\varepsilon_0$.

To arrive at the non-dimensional form of the equations we introduce the new variables

$$x = \tilde{x}/\lambda \ , \ y = \tilde{y}/\lambda \ , \ t = c\tilde{t}/\lambda = \tilde{t}\nu \ .$$

Here λ is the free space wavelength of the incoming field, having a frequency, ν. The field components are similarly normalized as

$$H_x = \tilde{H}_x \ , \ H_y = \tilde{H}_y \ , \ E_z = Z_0^{-1}\tilde{E}_z \ ,$$

yielding the non-dimensional TE equations

$$
\begin{aligned}
\frac{\partial H_z}{\partial t} &= -\frac{\partial E_y}{\partial x} \\
\frac{\partial H_x}{\partial t} &= \frac{\partial E_y}{\partial z} \\
\frac{\partial E_y}{\partial t} &= \frac{1}{n^2}\left(\frac{\partial H_z}{\partial z} - \frac{\partial H_z}{\partial x}\right) ,
\end{aligned}
\tag{2}
$$

which we shall consider in what remains.

As the materials are considered to be non-magnetic and lossless, the field components, H_z, H_x and E_y, are subject to the interface conditions

$$E_y^1 = E_y^2 \ , \ \hat{\boldsymbol{n}} \times \boldsymbol{H}^1 = \hat{\boldsymbol{n}} \times \boldsymbol{H}^2 \ , \ \hat{\boldsymbol{n}} \cdot \boldsymbol{H}^1 = \hat{\boldsymbol{n}} \cdot \boldsymbol{H}^2 \ , \tag{3}$$

where the superscripts refers to the field components in two neighboring materials while $\hat{\boldsymbol{n}}$ signifies a unit vector normal to the interface. Hence, the tangential electric field, E_y, as well as both the magnetic field components are continuous for the particular case considered here.

In a typical scenario, the diffractive element is integrated with a laser such that the incoming field itself is a guided wave. Hence, it is only reasonable to model the incoming field as the non-dimensional analytic 3-layer solution given as

$$H_z = -iCe^{i2\pi(t-n_{\text{eff}}z)} \times \begin{cases} qe^{-2\pi qx} & x \geq 0 \\ h\sin(2\pi hx) + q\cos(2\pi hx) & x \in [0, -d_2] \\ -p\left[\cos(2\pi hd_2) + qh^{-1}\sin(2\pi hd_2)\right] e^{2\pi p(x+d_2)} & x \leq -d_2 \end{cases} , \tag{4}$$

and

$$E_y = -\frac{H_x}{n_{\text{eff}}} = Ce^{i2\pi(t-n_{\text{eff}}z)} \times \begin{cases} e^{-2\pi qx} & x \geq 0 \\ \cos(2\pi hx) - qh^{-1}\sin(2\pi hx) & x \in [0, -d_2] \\ \left[\cos(2\pi hd_2) + qh^{-1}\sin(2\pi hd_2)\right] e^{2\pi p(x+d_2)} & x \leq -d_2 \end{cases} , \tag{5}$$

where C signifies an arbitrary amplitude factor. Moreover, we have introduced

$$q = \sqrt{n_{\text{eff}}^2 - n_1^2} \ , \ h = \sqrt{n_2^2 - n_{\text{eff}}^2} \ , \ p = \sqrt{n_{\text{eff}}^2 - n_3^2} \ ,$$

where the effective index of refraction, n_{eff}, is given as the solution to the eigenvalue equation

$$\tan(2\pi hd_2) = \frac{h(p+q)}{h^2 - pq} \ .$$

Provided d_2 is chosen appropriately, this equation only has one solution thereby guaranteeing single-mode operation of the waveguide.

3. The Computational Framework. The construction of the pseudospectral multi-domain scheme for the time-domain solution of Maxwells equations within a general diffractive optical element involves the combination of a number techniques. The key issue in the developments presented here is centered around the spatial approximation scheme while Maxwells equations is advanced in time using a low-storage 5-stage 4th-order Runge-Kutta scheme [3]. Although it requires an extra step for the completion of the step as compared to the standard 4th-order Runge-Kutta scheme it has a slightly larger stability region, implying that the total work is kept about constant. However, only one additional storage level is required for the implementation the scheme. The time-step is chosen such as to obey the CFL-criterion.

In what remains of this section we shall discuss the details of the spatial scheme for the solution of Eq.(2), subject to the prescribed initial and boundary conditions.

3.1. Chebyshev Spectral Methods. The scheme is based on Chebyshev collocation methods, which, due to their superior approximation properties, are widely used for the solution of partial differential equations. Within the two-dimensional unit square, $[-1, 1]^2$, this implies that we seek solutions to Eq.(2) on the form

$$(I_{N,N}f)(z, x) = \sum_{i=0}^{N}\sum_{j=0}^{N} f(z_i, x_j)\, g_i(z) h_j(x) \ ,$$

where the Chebyshev-Gauss-Lobatto grid points are given as

$$(i, j) \in [0, N] : \ z_i = -\cos\left(\frac{i\pi}{N}\right) \ , x_j = -\cos\left(\frac{j\pi}{N}\right) \ ,$$

and the interpolating Chebyshev-Lagrange polynomials are given as

$$g_i(z) = \frac{(1-z^2)T'_N(z)\,(-1)^{i+1}}{c_i N^2(z-z_i)} \ , \ h_j(x) = \frac{(1-x^2)T'_N(x)\,(-1)^{j+1}}{c_j N^2(x-x_j)} \ ,$$

where $c_0 = c_N = 2$ and $c_i = 1$ for $1 \le i \le N-1$ and $T_N(z)$ refers to the Chebyshev polynomial of order N.

To seek approximate solutions to a partial differential equation we ask that the equation is satisfied in a collocation sense, i.e., at the collocation points. Hence, we need to obtain values of the spatial derivatives at the collocation points as is done by approximating the differential operator by matrix operators with the entries given as

$$\mathrm{D}^z_{ij} = g'_j(z_i) \ , \ \mathrm{D}^x_{ij} = h'_j(x_i) \ ,$$

i.e., the computation of derivatives thus becomes matrix-multiplies. The explicit expressions of the entries of the matrix operator and further details on collocation methods, we refer to [4].

To increase the robustness of the scheme we find it useful to introduce a very weak filtering of the solution and employ an exponential filter of the type

$$\sigma_i = \begin{cases} 1 & 0 \le i \le N_c \\ \exp\left[-\alpha\left(\frac{i-N_c}{N-N_c}\right)^\gamma\right] & N_c < i \le N \end{cases} \ ,$$

where N_c is a cut-off mode number, γ is the order of the filter and $\alpha = -\ln \varepsilon_M$ with ε_M being the machine accuracy. To ensure a minimal effect of the filter we choose the order of the filter as $\gamma = N-2$, i.e., it scales with the resolution. The filtering along z may conveniently be expressed as a matrix operator, F, with the entries given as

$$\mathrm{F}_{ij} = \frac{2}{c_j N} \sum_{k=0}^{N} \frac{\sigma_k}{c_k} T_k(z_i) T_k(z_j) \ ,$$

and likewise for filtering along x.

3.2. Maxwells Equations on Curvilinear Form. As mentioned briefly in the previous section, the use of a tensor product approximation requires that $f(z,x)$ is defined on a rectangular grid. The first step towards a geometrically flexible pseudospectral scheme is to circumvent this restriction and extend the use of polynomial expansions to the general curvilinear quadrilateral domain. We assume the existence of a smooth non-singular mapping function, Ψ, relating the (z,x) coordinate system to the general curvilinear coordinate system (ξ,η) like

$$\xi = \xi(z,x) \ , \ \eta = \eta(z,x) \ .$$

We shall return to the actual specification and construction of the smooth map, Ψ, shortly. Adapting this formulation to Eq.(2) yields the hyperbolic system

$$\frac{\partial \boldsymbol{q}}{\partial t} + \mathrm{A}(\nabla\xi)\frac{\partial \boldsymbol{q}}{\partial \xi} + \mathrm{A}(\nabla\eta)\frac{\partial \boldsymbol{q}}{\partial \eta} = 0 \ , \tag{6}$$

where we have the state vector, $\boldsymbol{q} = (H_z, H_x, E_y)^T$. The general operator, $\mathrm{A}(\boldsymbol{n})$, with $\boldsymbol{n} = (n_z, n_x)$ representing the local metric, is given as

$$\mathrm{A}(\boldsymbol{n}) = \begin{bmatrix} 0 & 0 & n_x \\ 0 & 0 & -n_z \\ n_x n^{-2} & -n_z n^{-2} & 0 \end{bmatrix} \ ,$$

where we recall that n refers to the local index of refraction. This operator diagonalizes under the similarity transform, $\mathrm{A}(\boldsymbol{n}) = \mathrm{S}^{-1}(\boldsymbol{n})\,\Lambda(\boldsymbol{n})\,\mathrm{S}(\boldsymbol{n})$, where the diagonal eigenvalue matrix, $\Lambda(\boldsymbol{n})$, has the entries

$$\Lambda(\boldsymbol{n}) = |\boldsymbol{n}| \mathrm{diag}\left[-n^{-1}, 0, n^{-1}\right] \tag{7}$$

corresponding to the characteristic velocities of the waves being counter-, non-, and co-propagating along the normal vector $\boldsymbol{n}$ with the local speed of light. Here we have that $|\boldsymbol{n}|$ represents the length of the vector $\boldsymbol{n}$, such that $\boldsymbol{n} = |\boldsymbol{n}|(\hat{n}_z, \hat{n}_x)$.

The diagonalizing matrices, $\mathrm{S}(\boldsymbol{n})$ and $\mathrm{S}^{-1}(\boldsymbol{n})$, take the form

$$\mathrm{S}(\boldsymbol{n}) = \begin{bmatrix} -\hat{n}_x & \hat{n}_z & -\hat{n}_x \\ \hat{n}_z & \hat{n}_x & \hat{n}_z \\ n^{-1} & 0 & -n^{-1} \end{bmatrix} , \mathrm{S}^{-1}(\boldsymbol{n}) = \frac{1}{2} \begin{bmatrix} -\hat{n}_x & \hat{n}_z & n \\ 2\hat{n}_z & 2\hat{n}_x & 0 \\ -\hat{n}_x & \hat{n}_z & -n \end{bmatrix} ,$$

from which we obtain the characteristic variables

$$\boldsymbol{R} = \mathrm{S}^{-1}(\boldsymbol{n})\boldsymbol{q} = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -\hat{n}_x H_z + \hat{n}_z H_x + n E_y \\ 2\hat{n}_z H_z + 2\hat{n}_x H_x \\ -\hat{n}_x H_z + \hat{n}_z H_x - n E_y \end{bmatrix} . \tag{8}$$

Besides revealing information about the dynamics of the fields, the identification and the use of the characteristic variables takes, as we shall see shortly, an integral role in the specification of the multi-domain scheme, being the topic of the following section.

3.3. The Multi-Domain Formulation. We wish to solve Eq.(6) within a general computational domain, $\Omega \in \mathsf{R}^2$, in the (z,x)-plane. As we have briefly discussed, the most natural and computational efficient way of applying polynomial expansions in several dimensions is through the use of tensor products. However, this procedure requires that the computational domain can be smoothly mapped to the unit square. To surround this limitation, we construct Ω using K non-overlapping general curvilinear quadrilaterals, $\mathsf{D}^k \subset \mathsf{R}^2$, such that $\Omega = \bigcup_{k=1}^{K} \mathsf{D}^k$.

The advantages of such an approach, besides from providing the geometric flexibility, are many. In particular in connection with pseudospectral methods, where the multi-domain framework results in a lower total operation count and higher allowable time-step while providing a very natural data-decomposition, well suited for the implementation on contemporary parallel computers.

Once we have split the global computational domain into K subdomains, we need to construct the map, $\Psi : \mathsf{D} \rightarrow \mathsf{I}$, where $\mathsf{I} \subset \mathsf{R}^2$ is the unit square while we have the Cartesian coordinates, $(z,x) \in \mathsf{D}$, and the general curvilinear coordinates, $(\xi,\eta) \in \mathsf{I}$ being related through the map, $(x,y) = \Psi(\xi,\eta)$. To establish a one to one correspondence between the unit square and the general quadrilateral we construct the local map for each subdomain using transfinite blending functions [5, 6]. This allows for the computation of the metric of the mapping and outward pointing normal vectors at all points of the enclosing edges of the quadrilateral.

Within the multi-domain setting we must solve K independent problems in the individual subdomains. However, to obtain the global solution we must ensure that information is passed between the subdomains in a way consistent with the dynamics of Maxwells equations. In the particular scenario considered here, and illustrated in Fig. 1, we encounter two different types of interfaces, requiring different patching techniques.

The patching across boundaries of domains between regions with different material properties is accomplished by using the physical conditions on the field components, Eq.(3), which are enforced strongly.

For the patching of subdomains having the same material properties we utilize that the system, Eq.(6), is hyperbolic. Hence, it is only natural to transfer information between the various subdomains using the characteristic variables, Eq.(8), which are convected along the normal, $\hat{\boldsymbol{n}}$, with a speed given by the diagonal elements of $\Lambda(\boldsymbol{n})$, Eq.(7). Once the outward normal vector at the enclosing boundary of the subdomain is known, as it is once the map, Ψ, is constructed, we may uniquely determine which characteristics are leaving the subdomain and which are entering and thus needs specification. Indeed, we observe from the eigenvalues of $\mathrm{A}(\boldsymbol{n})$, Eq.(7), that while R_3 is always leaving the domain and therefore needs no boundary conditions, R_1 is always entering the computational domain and requires specification to ensure well-posedness. Thus, R_3, leaving a domain, supplies the sought after boundary conditions for R_1 in the neighboring domain and reversely for R_1 in the first domain. For the non-propagating R_2 we simply use the average across the interface. Once the characteristic variables have been adjusted, the physical fields are recovered through the relation $\mathrm{S}(\boldsymbol{n})\boldsymbol{R} = \boldsymbol{q}$. This procedure is applied along all interface points, including the vertices where it is done dimension-by-dimension, to arrive at the global solution at each times-step. As we shall see shortly, this procedure of patching hyperbolic systems is stable as well as accurate. Moreover, in a parallel setting the communication between subdomains grows only like the surface of the geometric building block rather than the volume.

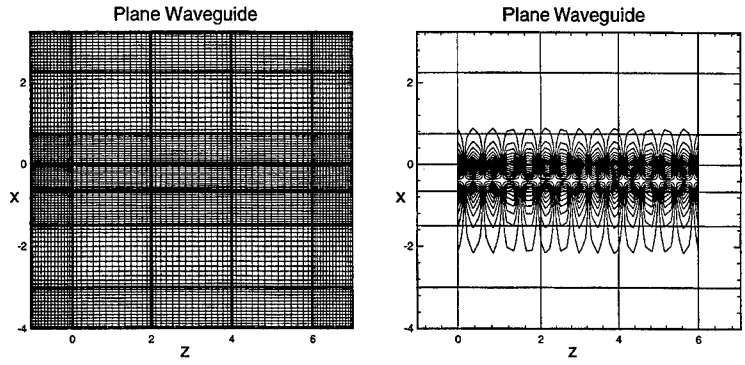


FIG. 2. *Illustration of the plane waveguide test case. On the left is the grid illustrated, with the high-index waveguide just below* $x = 0$. *On the right is the* H_z *shown at an arbitrary time.*

3.4. Far Field Boundary Conditions. The introduction of the perfectly matched layer (PML) methods [7] has spawned significant research into such methods. However, serious problems with these types of boundary conditions has recently been exposed [8] for the two-dimensional PML methods and a number of alternatives have subsequently appeared in the literature.

In [9], a well-posed PML scheme was introduced and shown to perform well in connection with pseudospectral multi-domain schemes. We have chosen to use that particular scheme after adapting it to the situation of general dielectric media. The implementation of the scheme is done in the total-field/scattered-field formulation to enhance the efficiency of the layers, being given as

$$
\begin{aligned}
\frac{\partial H_z}{\partial t} &= -\frac{\partial E_y}{\partial x} - 2\sigma_x H_z - \sigma_x P_x \\
\frac{\partial H_x}{\partial t} &= \frac{\partial E_y}{\partial z} - 2\sigma_z H_x - \sigma_z P_z \\
\frac{\partial E_y}{\partial t} &= \frac{1}{n^2}\left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}\right) - \sigma_z' Q_z + \sigma_x' Q_x \\
\frac{\partial P_z}{\partial t} &= \sigma_z H_x \qquad \frac{\partial Q_z}{\partial t} = -\sigma_z Q_z - n^{-2} H_x \\
\frac{\partial P_x}{\partial t} &= \sigma_x H_z \qquad \frac{\partial Q_x}{\partial t} = -\sigma_x Q_x - n^{-2} H_z \ .
\end{aligned}
$$

Note that the additional degrees of freedom, making possible the perfectly matched layer property, is introduced through a number of auxiliary fields rather than an unphysical splitting of the electromagnetic fields.

The PML assumes a rectangular interface bounded by $|z| \leq z_0$ and $|x| \leq x_0$ and the absorption profiles takes the polynomial form

$$
\sigma_z(z) = C_z(|z| - z_0)^p \ , \ \sigma_x(x) = C_x(|x| - x_0)^p \ ,
$$

where the constants, C_z and C_x, are tunable for optimal performance. We have found that using $p = 4$ in the profiles yields a good balance between smoothness and damping efficiency.

4. Numerical Experiments. Combining all of the elements described in the previous section into a general computational framework yields a model with significant geometric flexibility. Moreover, since high-order schemes are used in space as well as time one can hope to be able to accurately model electrically large waveguide structures in an reliable and efficient manner.

In the following we shall present two test cases with the purpose of validating the performance of the scheme and illustrate the prospects for addressing general problems of interest to science and engineering.

4.1. The Plane 3-Layer Waveguide Problem. To validate the accuracy and overall performance of the scheme, we consider a 3-layer plane waveguide as depicted in Fig. 1, assuming only that $n_0 = n_1 = 1$, i.e., vacuum is used as the cladding material. The waveguide itself is composed of a layer with $n_2 = 1.2$ of a thickness $d_2 = 0.6637\lambda$ while the bulk-material is assumed to be infinite with an index of refraction of $n_3 = 1.1$. These dimensions and materials ensure that only the fundamental mode can exist in the waveguide. The exact solution to this problem is given in Eqs.(4)-(5) with the effective index of refraction being $n_{\text{eff}} = 1.1369$.

As our test case we choose a 6λ long fragment of the otherwise infinite waveguide and assume that the wave pattern are fully developed, i.e., the fields given by Eqs.(4)-(5) are assumed to fill the waveguide. In Fig. 2 we show the grid for $N = 16$ in each subdomain and the total field region, in which the computation is conducted, as well as the scattered field PML layer is shown.

TABLE 1
Error in the computation of the plane waveguide solution at $t = 10$.

N	N_{ppw}	Δt	$L_\infty(H_z)$	$L_\infty(H_x)$	$L_\infty(E_y)$
16	7.3	1.83E-2	1.27E-2	4.53E-2	3.55E-2
20	9.1	1.19E-2	1.28E-4	3.31E-4	3.23E-4
24	10.9	8.27E-3	1.53E-5	7.94E-5	7.45E-5
32	14.5	5.05E-3	5.57E-6	3.12E-5	2.68E-5

The exact solution is advanced 10 periods after which the global L_∞ error of the three field components is measured. In Table 1 we give the results, clearly illustrating the spectral convergence when increasing the number of modes, N. One also notes that using as little 9 points per wavelength, N_{ppw}, in the waveguide results in very accurate field computations, illustrating the superior numerical properties of the pseudospectral time-domain methods.

4.2. Off-Plane Waveguide Holograms. As a second, and more realistic, test case we have chosen to consider a situation in which the waveguide/vacuum interface of the plane waveguide considered above is modulated as

$$0.25 \exp\left[-\left(\frac{z-10}{5}\right)^2\right] \sin(2\pi z) \ ,$$

i.e., a tapered periodic pattern. We shall also assume that the waveguide now spans 20λ. The full grid, including the surrounding PML layer region, consists of 96 elements and is shown in Fig. 3 with the resolution being taken to $N = 20$. To ensure that the initial conditions are divergence free we assume that the waveguide is empty at $t = 0$ at which point a guided wave is entering from the left.

In Fig. 3 we show the H_z component of the field after about 66 wave periods, clearly illustrating the exchange of energy from the guided wave mode into a wavefront propagating freely away from the holographic element. Although not shown, the situation for the remaining field components is similar, hence confirming the ability of the developed scheme to handle problems of interest for the construction and fabrication of integrated optical elements.

5. Concluding Remarks. It has been the purpose of this paper to develop a pseudospectral time-domain framework suitable for the accurate and efficient modeling of very general waveguide couplers. It also constitutes the first of example of a pseudospectral time-domain method for problems in computational electromagnetics involving general geometries and several different materials.

As we have confirmed through computational tests, the proposed pseudospectral time-domain framework is well suited for a variety of generic problems involving guided waves, coupling phenomena and wavefront modulation. Not only does the very accurate modeling of the phase behavior allow for a detailed computation of problems evolving over many periods of time and spanning many wavelengths. The requirement of only 7-9 points per wavelengths also significantly reduces the computational requirements and allows for addressing large problems at even moderately sized workstations.

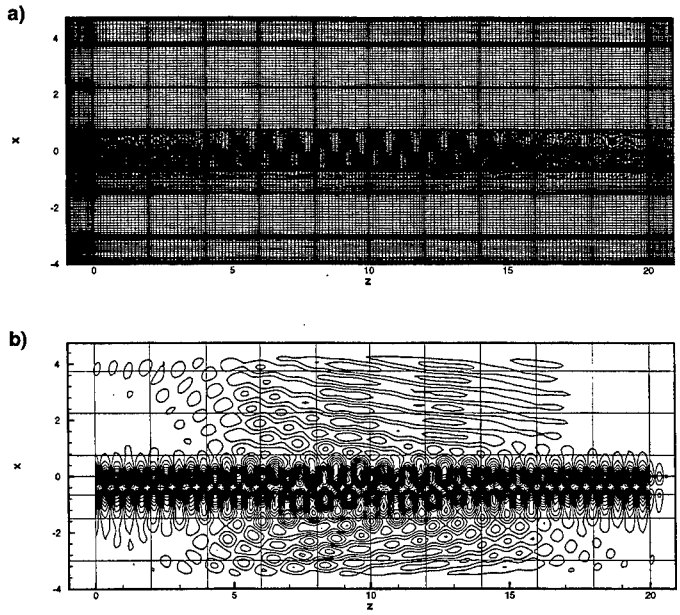


FIG. 3. *Example of a pseudospectral time-domain modeling of an off-plane waveguide hologram. a) The computational grid. b) The computed H_z component at $t = 66$.*

Acknowledgments. The authors would like to express their gratitude to Mr. B. Yang and Prof. D. Gottlieb, Brown University, for many helpful discussions.

REFERENCES

[1] B. Yang, D. Gottlieb and J. S. Hesthaven, *Spectral Simulation of Electromagnetic Wave Scattering*, J. Comput. Phys. **134**(1997), pp. 216-230.

[2] B. Yang and J. S. Hesthaven, *A Pseudospectral Method for Time-Domain Computation of Electromagnetic Scattering by Bodies of Revolution*, IEEE Trans. Antennas Propaga. 1997 - submitted.

[3] M. H. Carpenter and C. A. Kennedy, *Fourth-Order 2N-Storage Runge-Kutta Schemes.* NASA-TM-109112, 1994.

[4] D. Funaro. *Polynomial Approximation of Differential Equations*, volume m8 of *Lecture Notes in Physics.* Springer Verlag, Berlin, 1992.

[5] W. J. Gordon and C. A. Hall, *Transfinite Element Methods: Blending-Function Interpolation over Arbitrary Curved Element Domains*, Numer. Math. **21**(1973), pp. 109–129.

[6] J. S. Hesthaven, *A Stable Penalty Method for the Compressible Navier-Stokes Equations. III. Multi Dimensional Domain Decomposition Schemes*, SIAM J. Sci. Comp. 1997 - accepted.

[7] J. P. Berenger, *A Perfectly Matched Layer for the Absorption of Electromagnetic Waves*, J. Comput. Phys. **114**(1994), pp. 185–200.

[8] S. Abarbanel and D. Gottlieb, *A Mathematical Analysis of the PML Method*, J. Comput. Phys. **134**(1997), pp. 357–363.

[9] B. Yang D. Gottlieb and J. S. Hesthaven, *On the Use of PML ABC's in Spectral Time-Domain Simulations of Electromagnetic Scattering.* In *Proceedings of the 13'th Annual Review of Progress in Applied Computational Electromagnetics*, volume II, pages 926–934, Naval Postgraduate School, Monterey, CA, March 1997.

Transient Analysis of Acoustic Scattering Using Marching-on-in-Time With Plane Wave Time Domain Algorithm

A. Arif Ergin*, Balasubramaniam Shanker, Kemal Aygün, and Eric Michielssen

Center for Computational Electromagnetics,
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
1406 W. Green St., Urbana, IL 61801, USA.
aergin@decwa.ece.uiuc.edu

Abstract

This paper presents novel Plane Wave Time Domain (PWTD) algorithms which accelerate the computational analysis of transient surface scattering phenomena. The cost of performing a transient analysis of scattering from a body modeled by N_s spatial basis functions for a duration of N_t time steps scales as $O(N_t N_s^2)$ if classical time domain integral–equation based methods are used. This cost can be reduced to $O(N_t N_s^{4/3} \log N_s)$ and $O(N_t N_s \log N_s)$ using the proposed two–level and multilevel PWTD schemes, respectively. These algorithms are the time domain counterparts of frequency domain fast multipole methods and make feasible the practical broadband analysis of scattering from large and complex bodies. The two-level PWTD algorithm is described in the context of acoustic scattering and numerical examples validating the approach are presented.

1. Introduction

Recently, the scientific community has expressed a renewed interest in the analysis of short–pulse radiation and transient scattering phenomena. The characterization of transient wave phenomena is of paramount importance in disciplines ranging from electromagnetics to geophysics. Efficient computational analysis of these phenomena hinges upon the availability of fast time domain algorithms.

Time Domain Integral–Equation (TDIE) based solvers have been applied to the analysis of acoustic and electromagnetic scattering problems [1-4]. The TDIE techniques exhibit a number of desirable characteristics: (*i*) they only require a discretization of the scatterer surface, (*ii*) they implicitly impose the radiation condition, and (*iii*) they are devoid of grid dispersion errors. Unfortunately, these techniques have long been conceived as intrinsically unstable and computationally expensive when compared to their differential–equation counterparts. However, recently, progress towards stable TDIE based schemes has been reported [4]. In contrast, literature on techniques for reducing the computational complexity of these TDIE techniques is virtually nonexistent. This is in spite of the fact that the last decade has witnessed significant acceleration in frequency domain integral equation solvers with the advent of the Fast Multipole Method (FMM) [5], the impedance matrix localization technique, the multilevel matrix decomposition algorithm, etc. Although the structure of transient wave fields has been well studied, to our knowledge no TDIE algorithms with reduced computational complexity have been reported.

Classical TDIE based schemes for analyzing transient wave phenomena suffer from a high computational cost. Consider a surface scatterer of area S which resides in a homogeneous three–dimensional space and which is excited by a pulse whose temporal spectrum vanishes for $\omega > \omega_{\max}$. Integral–equation based approaches model the fields scattered from the surface as those produced by induced

surface sources. Since the sum of the incident and scattered fields satisfies certain boundary conditions on the surface of the scatterer, an integral equation relating the incident field on the scatterer to the field produced by all current and past sources can be constructed. Assuming that the induced surface sources reside on the scatterer surface for a duration T, after which they become vanishingly small, the source distribution can be discretized and represented in terms of $N_s \propto S(\omega_{\max}/c)^2$ spatial and $N_t \propto T\omega_{\max}$ temporal samples. Here, c denotes the wave speed in the medium. Discretization of the integral–equation results in a time marching procedure for computing the induced sources. Updating the source distribution requires the computation of the field at N_s observers due to all N_s sources. Since this computation has to be performed for all time steps, the computational complexity associated with the classical TDIE algorithms scale as $O(N_t N_s^2)$.

This paper introduces diagonalized time domain translation operators which permit the rapid evaluation of transient fields produced by surface–bound source distributions. These diagonalized translation operators can be used in tandem with classical integral–equation based techniques for analyzing transient scattering phenomena. The computational complexities associated with the solution of large-scale surface scattering problems using the proposed two–level and multilevel fast PWTD algorithms scale as $O(N_t N_s^{4/3} \log N_s)$ and $O(N_t N_s \log N_s)$, respectively.

The organization of this manuscript is as follows: Section 2 introduces a diagonalized translation operator for time domain fields that satisfy the scalar wave equation, both for continuous and sampled field representations. An implementation of two–level PWTD scheme for analyzing acoustic scattering problems together with a complexity analysis and some numerical examples are presented in Section 3. Section 4 presents our conclusions.

2. A Diagonalized Time Domain Translation Operator

In this section, a plane wave representation for transient wave fields is derived together with space–time constraints that ensure its validity and applicability in a time marching algorithm. It is shown that the plane wave representation has a diagonalized translation operator. A closed–form expression for the translation function for sampled field representations is also presented.

Consider a source distribution $q(\mathbf{r},t)$ residing in a source sphere of radius R_s centered around $\mathbf{r}^{c(s)}$ and radiating in an unbounded, nondispersive, and homogeneous medium. The field $u(\mathbf{r},t)$ produced by $q(\mathbf{r},t)$ is to be evaluated at observers distributed throughout an observation sphere of radius R_o centered around $\mathbf{r}^{c(o)}$. Let $\mathbf{R}_c = \mathbf{r}^{c(o)} - \mathbf{r}^{c(s)}$ denote the vector connecting the source and observation sphere centers. Without loss of generality, it is assumed that $R_o = R_s$, and that $\mathbf{R}_c = R_c\hat{\mathbf{z}}$, where $R_c = |\mathbf{R}_c|$.

The field $u(\mathbf{r},t)$ satisfies the wave equation

$$\nabla^2 u(\mathbf{r},t) - \frac{\partial_t^2}{c^2} u(\mathbf{r},t) = -q(\mathbf{r},t), \tag{1}$$

where ∂_t^2 denotes second derivative with respect to time. The field at an observer located at $\mathbf{r}$ can be succinctly expressed as

$$u(\mathbf{r},t) = \int_{V_s} d\mathbf{r}' q(\mathbf{r}',t) * \frac{\delta(\mathbf{r}', t - |\mathbf{r}-\mathbf{r}'|/c)}{4\pi|\mathbf{r}-\mathbf{r}'|} \tag{2}$$

where V_s is the volume enclosing the source, * denotes time convolution, and $\delta(\cdot)$ is a Dirac impulse. If $q(\mathbf{r},t)$ consists of a point source located at $\mathbf{r}^s$ in the source sphere with a time signature $f(t)$, i.e., if $q(\mathbf{r},t) = f(t)\delta(\mathbf{r}-\mathbf{r}^s)$, then the field observed at a point $\mathbf{r}^o$ in the observation sphere is given by

$$u(\mathbf{r}^o,t)=\frac{f\left(t-r_s^o/c\right)}{4\pi r_s^o}, \tag{3}$$

where $r_s^o=\left|\mathbf{r}_s^o\right|$ and $\mathbf{r}_s^o=\mathbf{r}^o-\mathbf{r}^s$. Henceforth, to simplify the notation, the positions of the source and observation points relative to their respective sphere centers are denoted by $\tilde{\mathbf{r}}^s=\mathbf{r}^s-\mathbf{r}^{c(s)}$ and $\tilde{\mathbf{r}}^o=\mathbf{r}^o-\mathbf{r}^{c(o)}$.

The computational cost associated with the evaluation of transient fields via Eqn. (3) due to multiple sources at multiple observers can be reduced significantly provided that the fields are represented in terms of a plane wave basis. As a first step towards representing the field $u(\mathbf{r}^o,t)$ as a superposition of plane waves, the source signal $q(\mathbf{r},t)$ of duration T is artificially broken up into L subsignals $q_l(\mathbf{r},t)$, $l=0,...,L-1$, of duration $T_s=T/L$ such that

$$q(\mathbf{r},t)=\sum_{l=0}^{L-1}q_l(\mathbf{r},t)=\delta(\mathbf{r}-\mathbf{r}^s)\sum_{l=0}^{L-1}f_l(t), \tag{4}$$

where $f_l(t)=0$ for $t<lT_s$ and $t\geq(l+1)T_s$. Let $u_l(\mathbf{r}^o,t)$ denote the field at $\mathbf{r}^o$ due to $q_l(\mathbf{r},t)$. Then, $u(\mathbf{r}^o,t)=\sum_{l=0}^{L-1}u_l(\mathbf{r}^o,t)$.

To arrive at a plane wave representation of $u_l(\mathbf{r}^o,t)$, consider the field $\tilde{u}_l(\mathbf{r}^o,t)$ defined by

$$\tilde{u}_l\left(\mathbf{r}^o,t\right)=-\frac{\partial_t}{8\pi^2c}\int_0^{2\pi}d\phi\int_0^{\theta_{\text{int}}}d\theta\sin\theta\,\delta\left(t-\hat{\mathbf{k}}\cdot\tilde{\mathbf{r}}^o/c\right)*\delta\left(t-\hat{\mathbf{k}}\cdot\mathbf{R}_c/c\right)*\tilde{q}_l\left(\hat{\mathbf{k}},t\right), \tag{5}$$

where $\hat{\mathbf{k}}=\hat{\mathbf{x}}\sin\theta\cos\phi+\hat{\mathbf{y}}\sin\theta\sin\phi+\hat{\mathbf{z}}\cos\theta$, the integration is over a cap of the unit sphere for which $\theta\leq\theta_{\text{int}}$, and $\tilde{q}_l(\hat{\mathbf{k}},t)$ is the Slant Stack Transform (SST) of the source distribution $q_l(\mathbf{r},t)$ defined by

$$\tilde{q}_l\left(\hat{\mathbf{k}},t\right)=\int_{V_s}d\mathbf{r}'\,\delta\left(t+\hat{\mathbf{k}}\cdot\tilde{\mathbf{r}}^s/c\right)*q_l\left(\mathbf{r}',t\right). \tag{6}$$

The SST maps the source distribution into a time dependent plane wave emanating form the source sphere and propagating along $\hat{\mathbf{k}}$. Henceforth, plane waves obtained by an SST will be termed as outgoing rays. For the point source configuration specified above, the SST reduces to $\delta(t+\hat{\mathbf{k}}\cdot\tilde{\mathbf{r}}^s/c)*f_l(t)$ and the integral in Eqn. (5) can be evaluated explicitly to yield [6,7]

$$\tilde{u}_l(\mathbf{r}^o,t)=\frac{1}{4\pi r_s^o}f_l\left(t-\frac{r_s^o}{c}\right)-\frac{1}{8\pi^2r_s^o}\int_0^{2\pi}d\phi' f_l\left(t-\frac{r_s^o}{c}\cos\theta'_{\text{int}}(\phi',\mathbf{r}^o,\mathbf{r}^s)\right). \tag{7}$$

where $\theta'_{\text{int}}(\phi',\mathbf{r}^o,\mathbf{r}^s)$ is the angle between the vector $\mathbf{r}_s^o$ and the vector extending to the boundary of the integration cap from $\mathbf{r}^s$. In Eqn. (7), the first term on the right hand side corresponds to the true observed field $u_l(\mathbf{r}^o,t)$. Note that, were it not for the second term, which will be referred to as the ghost signal, $\tilde{u}_l(\mathbf{r}^o,t)$ would be identical to $u_l(\mathbf{r}^o,t)$. If the ghost signal can somehow be removed from $\tilde{u}_l(\mathbf{r}^o,t)$, Eqn. (5) implies that the true observed field can be constructed as a superposition of plane waves using techniques that are akin to those underlying the frequency–domain FMM.

To derive a scheme that retains only the true observed field by time gating $\tilde{u}_l(\mathbf{r}^o,t)$, the following observations are in order. From Eqn. (7) it follows that the ghost signal present in $\tilde{u}_l(\mathbf{r}^o,t)$ vanishes after

$$t_l^{ghost} = \frac{r_s^o}{c}\cos\theta'_{\min} + (l+1)T_s \;<\; (R_c\cos\theta_{\text{int}} + 2R_s)/c + (l+1)T_s \tag{8}$$

where $\theta'_{\min} = \min\left[\theta'_{\text{int}}(\phi', \mathbf{r}^o, \mathbf{r}^s)\right]$, and the upper bound follows from geometrical considerations. The fields in the observation sphere coincide with the true fields after the ghost signal has vanished, i.e. $t > t_l^{ghost}$. Also, the true field does not reach the observation sphere before

$$t_l^{trans} = (R_c - 2R_s)/c + lT_s. \tag{9}$$

Therefore, provided that $t_l^{trans} > t_l^{ghost}$, all ghost fields in the observation sphere cease to exist before the true signal arrives. In addition, if $t_l^{trans} > (l+1)T_s$, all source activity related to the l^{th} time interval ends before the true signal reaches any observer. Hence, it is possible to obtain a ghost–free solution via Eqn (5) by choosing T_s and θ_{int} such that the conditions $t_l^{trans} > t_l^{ghost}$ and $t_l^{trans} > (l+1)T_s$ are satisfied.

The above observations are the basis for the construction of the PWTD algorithm in which evaluation of Eqn. (5) is interpreted as a three–stage process. In the first stage, outgoing rays are formed by calculating the SST of the source distribution using Eqn. (6). The second stage consists of translating the outgoing rays from the source sphere to the observation sphere. This is accomplished by convolving the outgoing rays with the translation function $-\partial_t\delta(t - \hat{\mathbf{k}}\cdot\mathbf{R}_c/c)/8\pi^2 c$. In the PWTD algorithm, the removal of the ghost is achieved by performing the translation at $t = t_l^{trans}$. The savings in computational cost are due to the fact that the translation operation maps an outgoing plane wave to another plane wave that travels in the same direction, i.e., the translation operator is diagonal. In the third stage, the translated rays are projected onto the observer location yielding the true field $u_l(\mathbf{r}^o, t)$.

In practice, Eqn. (5) needs to be evaluated numerically in a discrete setting. Hence one needs to work with sampled field representations which has been a topic of ongoing research [8]. This leads to an expression for $\tilde{u}_l(\mathbf{r}^o, t)$ of the form

$$\tilde{u}_l(\mathbf{r}^o, t) = \sum_{n=0}^{M'} \sum_{m=-M_n}^{M_n} w_{nm}\,\delta\left(t - \hat{\mathbf{k}}_{nm}\cdot\tilde{\mathbf{r}}^o/c\right) * \mathscr{T}_{nm}(t) * \tilde{q}_l(\hat{\mathbf{k}}_{nm}, t), \tag{10}$$

where w_{nm} are the integration weights, $\hat{\mathbf{k}}_{nm}$ denote the integration directions, the translation function $\mathscr{T}_{nm}(t)$ is given by

$$\mathscr{T}_{nm}(t) = -\frac{\partial_t}{4\pi R_c(2M_n+1)}\Psi_n\left(\cos^{-1}\frac{ct}{R_c}\right) \quad \text{for} \quad \frac{R_c}{c}\cos\theta_{\text{int}} \le t \le \frac{R_c}{c}, \tag{11}$$

and $\Psi_n(\theta)$ is a spatially bandlimited version of $\delta(t - \hat{\mathbf{k}}\cdot\mathbf{R}_c/c)$. In [6] it is shown that $\mathscr{T}_{nm}(t)$ can be expressed as a polynomial in t and that the error incurred in the numerical evaluation of Eqn. (10) can be reduced to arbitrary precision.

3. Implementation and Complexity Analysis of an Algorithm for Fast Analysis of Acoustic Scattering

To illustrate the use of the PWTD algorithm in conjunction with the TDIE based schemes, a two–level algorithm for the analysis of acoustic scattering from rigid bodies is presented next. Consider a rigid body bounded by a surface S. Let this body be illuminated by an incident pressure field $P^{inc}(\mathbf{r}, t)$. Then the total pressure field $P(\mathbf{r}, t)$ on S satisfies the integro–differential equation

$$P(\mathbf{r}, t) = P^{inc}(\mathbf{r}, t) + \int_{S^*} d\mathbf{r}'\, P(\mathbf{r}', t) * \hat{\mathbf{n}}'\cdot\nabla'\frac{\delta(\mathbf{r}', t - |\mathbf{r} - \mathbf{r}'|/c)}{4\pi|\mathbf{r} - \mathbf{r}'|}, \tag{12}$$

where $\hat{\mathbf{n}}'$ denotes the outward normal to S at $\mathbf{r}'$, and S^+ indicates that the integral is evaluated as $\mathbf{r}' \to S$ from outside the body. Conventionally, a solution to Eqn. (12) is obtained numerically using the Marching–On–in–Time (MOT) scheme. In this method, S is discretized by a suitable triangulation and $P(\mathbf{r},t)$ is expanded as a weighted superposition of N_s spatial and N_t temporal basis functions. Applying spatial Galerkin testing to the resulting equation at the j^{th} time step yields a matrix equation which can be solved for the unknown expansion coefficients for that time step. Hence, all the expansion coefficients can be found by starting at $j=1$ and solving a matrix equation of the form $\mathbf{A}\mathbf{x}_j = \mathbf{b}_j$ at each time step. The most expensive part of this algorithm is synthesizing the part of the right hand side vector $\mathbf{b}_j$ obtained by evaluating the fields over N_s testing functions due to past pressure fields represented by N_s basis functions, at each time step. This process is equivalent to evaluating the integral in Eqn. (12) and has a computational complexity of $O(N_s^2)$ per time step. However, the similarity of the integral in Eqn. (12) to the one in Eqn. (2) suggests that evaluation of this integral can be accelerated by utilizing the PWTD algorithm.

The first step in arriving at a reduced complexity algorithm is to divide the scatterer into N_g subscatterers, each of which can be enclosed in a sphere of radius R_s. The sizes of the subscatterers are chosen such that approximately $M_s = N_s/N_g$ spatial basis functions are contained in each sphere. Then, the interactions between sufficiently remote subscatterers are computed using the PWTD algorithm and all other interactions are accounted for by the conventional MOT scheme.

Although the kernels of the integrals in Eqns. (2) and (12) are different, the PWTD part of the accelerated scheme still consists of three–stages: forming the outgoing rays for all spheres, translating the outgoing rays to observation spheres at $t = t_l^{trans}$, and projecting the translated rays onto the observers. However, it can be shown that for the kernel of the integral in Eqn. (12), the SST takes the form

$$\tilde{P}(\hat{\mathbf{k}},t) = \int_{V_s} d\mathbf{r}' \left(\hat{\mathbf{n}}' \cdot \hat{\mathbf{k}}\right) \delta\left(t + \hat{\mathbf{k}} \cdot \tilde{\mathbf{r}}^s / c\right) * P(\mathbf{r}',t), \tag{13}$$

and the translation function becomes $\partial_t \mathcal{T}_{nm}(t)/c$. It should also be noted that applying the PWTD algorithm in a multiple sphere setting permits further savings since (*i*) the outgoing rays from a source sphere can be reused to calculate the fields at different observation spheres, and (*ii*) rays that are translated to an observation sphere from different source spheres are superimposed before they are projected onto the observers.

To derive the computational complexity of a surface scattering analysis using the resulting two–level algorithm, the total cost C_T is identified as the sum of two components. The first component C_{NF} is due to the direct analysis of the interactions between nearby subscatterers using the MOT algorithm whose computational complexity scales as $O(N_t M_s^2)$ per group. Since, only a few nearby groups are associated with each group, it is found that

$$C_{NF} \propto N_t M_s^2 N_g \propto N_t N_s M_s. \tag{14}$$

The second cost component is associated with computation of interactions between remote groups using the three-step PWTD algorithm and is considered next. The cost of forming outgoing rays, C_{FF}^1, is proportional to the number of sources in a group, the number of ray directions, the number of time step, and the number of groups. It can be shown that the number of ray directions associated with each sphere is proportional to M_s. Hence,

$$C_{FF}^1 \propto N_t M_s^2 N_g \propto N_t N_s M_s. \tag{15}$$

To derive an expression for the cost of translation step, it is assumed that a T_s which is proportional to R_s/c and which is of duration $M_t = N_t/L$ time steps is chosen. Then, the length of both the outgoing rays and the translation functions are proportional to M_t. Hence, the translation process, which is the convolution of outgoing rays with translation functions, can be accomplished in $O(M_t \log M_t)$ operations using fast Fourier

transforms. Furthermore, noting that for surface scatterers $M_t \propto T_s \propto R_s \propto M_s^{1/2}$, the cost of performing translations between N_g^2 group pairs, for L time intervals is seen to be

$$C_{FF}^2 \propto N_g^2 L M_t \log M_t \propto N_t \left(\frac{N_s}{M_s}\right)^2 \log M_s \tag{16}$$

Since projecting incoming rays onto observer locations is the reverse process of generating outgoing rays, the cost of the last step C_{FF}^3 is also proportional to $N_t N_s M_s$.

The total cost is given by $C_T = C_{NF} + C_{FF}^1 + C_{FF}^2 + C_{FF}^3$. It can be verified that C_T scales as $O(N_t N_s^{4/3} \log N_s)$ if M_s is chosen to be proportional to $N_s^{1/3}$. Furthermore, it can be shown that casting this algorithm in a multilevel framework yields a computational complexity of $O(N_t N_s \log N_s)$ [6].

In order to validate the applicability of the above algorithm, surface pressure on a square cylinder illuminated by a Gaussian plane wave propagating along the $-z$ direction is computed using both the classical and accelerated MOT methods. The scattered surface pressure observed at a point on the top of the cylinder is plotted in Figure 2(a) for both solutions. Figure 2(b) shows the percent difference between the two solutions normalized to the peak amplitude of the scattered surface pressure. As a second example, scattering of a similar incident field from an almond modeled with 4900 spatial basis functions is analyzed. Figures 3(a) and 3(b) compare the total surface pressure at a point on the almond and the normalized backscattered far field, respectively. It is seen that the results are in excellent agreement. It can be shown that the accuracy of the simulation can be controlled by the various approximations invoked in the PWTD algorithm.

4. Conclusions

This paper presented a PWTD scheme that permits the fast computation of transient fields radiated by surface bound sources. This scheme relies on diagonalized translation operators and can be considered the time domain counterpart of the frequency domain FMM. The PWTD scheme complements integral–equation based source updating methods and reduces the computational complexity associated with the analysis of surface scattering phenomena from $O(N_t N_s^2)$ to $O(N_t N_s^{4/3} \log N_s)$ for two–level and to $O(N_t N_s \log N_s)$ for multilevel algorithms. The practical implementation of the PWTD algorithm has been elucidated, and examples illustrating its accuracy and application to acoustic scattering problems have been presented.

Acknowledgments: This research was supported by the Air Force Office for Scientific Research via the MURI program under contract no. F49620–96–1–0025, the NSF under the grant ECS 95–02138, and the Gebze Institute of Technology, Turkey.

References

[1] M.J. Bluck and S.P. Walker, "Analysis of three-dimensional transient acoustic wave propagation using the boundary integral equation method," *International Journal for Numerical Methods in Engineering,* vol. 39, pp. 1419-1431, 1996.

[2] B. P. Rynne, "Time domain scattering from arbitrary surfaces using the electric field integral equation," *Journal of Electromagnetic Waves and Applications,* vol. 5, no. 1, pp. 93-112, Jan. 1991.

[3] S. M. Rao and D. A. Wilton, "Transient scattering by conducting surfaces of arbitrary shape," *IEEE Trans. Antennas Propag.* vol. 39, no. 1, pp. 56-61, Jan. 1991.

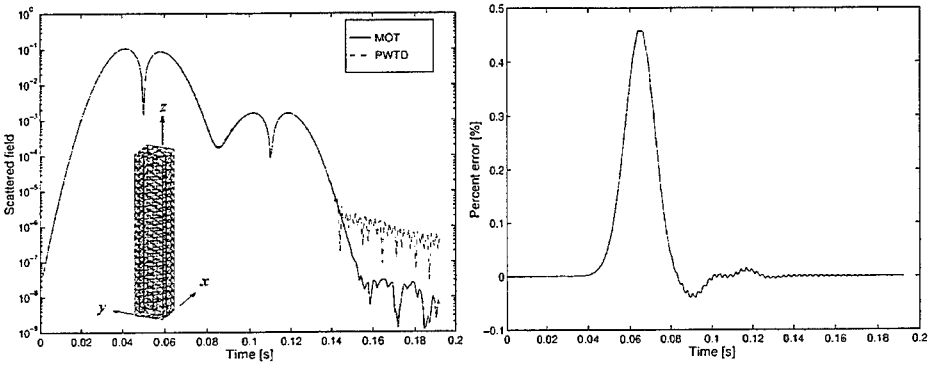
[4] M. J. Bluck and S. P. Walker, "Time-domain BIE analysis of large three-dimensional electromagnetic scattering problems," *IEEE Trans. Antennas Propag.,* vol. 45, no. 5, pp. 894-901, May 1997.

[5] R. Coifman, V. Rokhlin and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian approach," *IEEE Antennas and Propagation Magazine,* vol. 35, no. 3, pp. 7-12, June 1993.

[6] A.A. Ergin, B. Shanker, and E. Michielssen, "Fast evaluation of transient wave fields using diagonal translation operators," Center for Computational Electromagnetics, Univ. of Illinois at Urbana-Champaign, Res. Rep. CCEM-30-97, Sept. 1997. (to appear in *J. Comput. Phys.*)

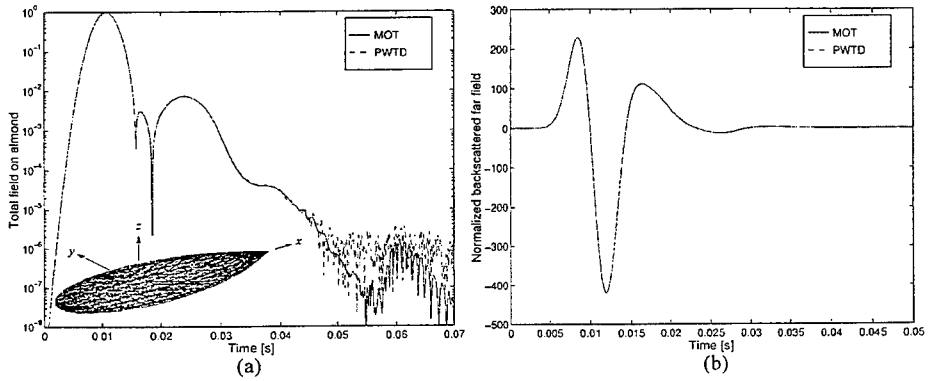
[7] E. Heyman, "Time-dependent plane-wave spectrum representations for radiation from volume source distributions," *Journal of Mathematical Physics,* vol. 37, no. 2, pp. 658-681, Feb. 1996.

[8] O.M. Bucci, C. Genneralli, and C. Savarese, "Optimal interpolation of radiated fields over a sphere," *IEEE Trans. Antennas Propag.*, vol. 39, no. 11, pp. 1633-1643, Nov. 1991.



(a) (b)

Figure 1. Transient scattering of a Gaussian plane wave with significant spectral content up to $f = 115Hz$ by a $1\times1\times10m$ square cylinder modeled by 612 unknowns. ($c = 343m/s$) (a) Scattered field observed at the top. (b) Percent difference between the two solutions normalized to maximum scattered field amplitude.



(a) (b)

Figure 2. Transient scattering of a Gaussian plane wave with significant spectral content up to $f = 575Hz$ by a 4900 unknown almond that fits in a box of $5\times2\times0.5m$. ($c = 343m/s$) (a) Total pressure observed near the front end. (b) Normalized backscattered far field $|\mathbf{r}|P(\mathbf{r},t)$.

A plane wave time domain algorithm for the fast analysis of transient electromagnetic scattering phenomena

Balasubramaniam Shanker*, A. Arif Ergin, Kemal Aygün and Eric Michielssen
Center for Computational Electromagnetics
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
1406 W. Green St., Urbana, IL 61801
Email: shanker@socrates.ece.uiuc.edu

Abstract

This paper presents a fast algorithm for solving time–domain surface integral equations commonly encountered in electromagnetics. The proposed algorithm is based on plane wave time–domain expansions of radiated fields [1] and augment classical Marching–On–in–Time (MOT) method for solving the magnetic field surface integral equation. The computational cost associated with this algorithm scales as $O(N_t N_s^{4/3} \log N_s)$ as opposed to $O\left(N_t N_s^2\right)$, where N_t is the number of time steps and N_s denotes the number of spatial samples used in discretizing the current on the scatterer. Numerical results demonstrating the applicability of the proposed solvers to the analysis of transient scattering from electrically large structures are presented.

1 Introduction

Development of techniques for the analysis of transient wave phenomena is a topic of renewed interest in a number of disciplines, including computational electromagnetics. Most popular analysis tools are differential equation based. However, these methods depend on volumetric discretization of the scatterer, and when, applied to surface scatterers the computational cost associated with these techniques scales unfavorably. On the other hand, time–domain integral equation (TDIE) techniques only require a discretization of the scatterer surface. Until recently, these methods were thought to be intrinsically unstable and computationally expensive. In the last few years a considerable research effort has been devoted to the stabilization of the TDIE techniques [2, 3, 4], and recently Walker [4] introduced an implicit scheme which alleviates the instability problem to a large extent. Thus, reducing the computational complexity of TDIE schemes would yield an even more viable approach for analyzing transient scattering from electrically large structures.

The problems associated with TDIE methods are not altogether a-similar to those overcome by the Fast Multipole Method (FMM) in frequency domain computations. Recently, an algorithm was introduced which enables fast computation of transient scalar wave fields by relying on the decomposition of the radiated field into transient plane waves [1]. It was theoretically shown that use of this algorithm in conjunction with classical time–stepping integral equation schemes significantly reduces the computational complexity associated with the analysis of scattering from large surface structures. Furthermore, this algorithm has been implemented in conjunction with the MOT scheme for the electric field integral equation, and is being presented elsewhere [5].

This paper presents an algorithm designed to accelerate the analysis of transient electromagnetic scattering from large perfectly electrically conducting surface scatterers residing in free space. This algorithm complements classical MOT scheme for solving Magnetic Field Integral Equations (MFIE). Also, computational results illustrating the usefulness of this algorithm in analyzing transient scattering from electrically large structures are presented.

2 Formulation

In this section, the MFIE for analyzing transient electromagnetic surface scattering phenomena is introduced. An MOT scheme for solving these equations is presented, and the Plane Wave Time Domain (PWTD) Method for accelerating the solution process is elucidated.

2.1 Integral Equations

Let S denote the surface of a conducting body excited by an electromagnetic field $\mathbf{E}^i(\mathbf{r},t), \mathbf{H}^i(\mathbf{r},t)$, and let $\mathbf{J}(\mathbf{r},t)$ denote the surface current on S. An integro–differential equation can be derived by using the well known relation

between the surface current density and incident and scattered magnetic fields. The resulting MFIE is

$$\hat{\mathbf{n}} \times \mathbf{H}^i(\mathbf{r},t) = \mathbf{J}(\mathbf{r},t) - \hat{\mathbf{n}} \times \mathbf{H}^s(\mathbf{r},t) \tag{1a}$$

where $\hat{\mathbf{n}}$ denotes the normal to S, and

$$\mathbf{H}^s(\mathbf{r},t) = \frac{1}{4\pi}\nabla \times \int_s dS' \left[\frac{\delta\left(t - R/c\right)}{R} * \mathbf{J}(\mathbf{r}',t)\right] \tag{1b}$$

where $R = |\mathbf{r} - \mathbf{r}'|$ is the distance between the source and observation points, and $\delta(\cdot)$ is a Dirac delta. Henceforth, ∂_t is used to denote a time derivative and c denotes the speed of light.

2.2 Marching–on–in–Time formulation

The MFIE (eqn.(1)) can be solved using the MOT method. The spatial and temporal variation of the current $\mathbf{J}(\mathbf{r},t)$ on S are represented using the basis functions, $\mathbf{j}_n(\mathbf{r})$ for $n = 1, \cdots, N_s$, and $T_j(t)$ for $j = 0, \cdots, N_t$, respectively. The basis functions $\mathbf{j}_n(\mathbf{r})$ are chosen to be the Rao–Wilton–Glisson (RWG) functions which have been extensively used in both frequency and time domain analysis. The reader is referred to Ref. [6] for a complete description. Rao and Wilton [2] use triangular functions to represent the temporal variation. However, higher order interpolants can also be used to improve stability and accuracy of the MOT scheme. Manara *et al.* [7] suggest choosing the order of the temporal interpolants to match the highest temporal derivative that appears in the equations. Thus, the current on S is represented by

$$\mathbf{J}(\mathbf{r},t) = \sum_{j=0}^{N_t}\sum_{n=1}^{N_s} I_{n,j}\mathbf{j}_n(\mathbf{r})T_j(t) \tag{2}$$

where $I_{n,j}$ is the weight associated with the space–time basis function $\mathbf{j}_n(\mathbf{r})T_j(t)$.

Using eqn. (2) in eqns. (1), and applying Galerkin testing, an MOT scheme can be constructed for the MFIE and is succinctly represented in matrix form as

$$\mathcal{Z}_0\mathcal{I}_j = \mathcal{F}_j^{inc} - \sum_{l=1}^{j-1}\mathcal{Z}_l\mathcal{I}_{j-l} \tag{3}$$

More specifically,

$$\mathcal{Z}_{l,mn} = \left\langle \mathbf{j}_m(\mathbf{r}), \left\{\mathbf{j}_n(\mathbf{r})T_{j-l}(t) - \hat{\mathbf{n}} \times \mathbf{H}^s_{n,j-l}(\mathbf{r},t)\right\}\right\rangle\Big|_{t=t_j} \quad , \tag{4a}$$

$$\mathbf{H}^s_{n,j-l}(\mathbf{r},t) = \frac{1}{4\pi}\nabla \times \int_{s_n} dS' \left[\mathbf{j}_n(\mathbf{r}')\frac{T_{j-l}\left(t - R/c\right)}{R}\right] \quad . \tag{4b}$$

In the above equations,

$$\langle \mathbf{j}_m(\mathbf{r}), \mathbf{\Phi}(\mathbf{r})\rangle = \int_s dS \mathbf{j}_m(\mathbf{r}) \cdot \mathbf{\Phi}(\mathbf{r}) \tag{5}$$

for any function $\mathbf{\Phi}(\mathbf{r})$, $t_j = j\Delta_t$, and Δ_t is the time step size. Also, $\mathcal{F}_j^{inc}$ represents the incident field tested by a basis function, and $\mathcal{I}_j$ is an array of the weights $I_{n,j}$, $n = \mathbf{1}, \cdots, N_s$.

Equation (3) is the basis of the classical MOT scheme. In the past, this algorithm has been observed to be unstable. But recent research has shown that it can be stabilized by adopting implicit schemes [4]. The stability can be further improved by using backward differencing in time, and by adopting more accurate numerical integration rules and a center of patch testing procedure. The evaluation of the right hand side of eqn. (3) requires $O(N_s^2)$ CPU resources; hence, the cost associated with the computation of all currents for all time steps scales as $O(N_s^2 N_t)$.

2.3 Plane Wave Time Domain algorithms

It was theoretically shown in Ref. [1] that using a two-level PWTD algorithm in conjunction with the MOT reduces the computational complexity of the analysis from $O(N_s^2 N_t)$ to $O(N_t N_s^{4/3} \log N_s)$. This reduced complexity is a consequence of expressing scattered fields as a superposition of plane waves.

To implement the PWTD algorithm in the framework of the MOT algorithm, the scatterer is divided into sub-scatterers which are confined to boxes defined on a rectangular grid. Two boxes are said to be in each other's far–field if the distance between their centers is larger than a prescribed distance. All other boxes are said to reside in each other's near–field. All near–field interactions are accounted for using the classical MOT scheme while all far–field interactions are computed using the PWTD algorithm. It is to be noted that the boxes should be scaled such that the number of spatial unknowns per box is of $O(N_s^{1/3})$.

Consider a source and an observation box which are in each other's far–field. Assume that each box is contained in a circumscribing sphere of radius R_s, and denote the center of the source and observation boxes by $\mathbf{r}_s$ and $\mathbf{r}_o$, respectively. Let $\mathbf{R}_c = \mathbf{r}_o - \mathbf{r}_s$ denote the vector connecting the source and observation box centers. Without loss of generality, it is assumed that $\mathbf{R}_c = \hat{\mathbf{z}}\,|\mathbf{R}_c| = \hat{\mathbf{z}} R_c$.

If the current $\mathbf{J}(\mathbf{r},t)$ in the source box is divided into L consecutive sub-signals each occupying a time slice $(l-1)T_s \leq t < lT_s$ for $l = 1, \cdots, L$, then the field at any point $\mathbf{r}$ in the source box due to these sources can be expressed as

$$\langle \mathbf{j}_m(\mathbf{r}), -\hat{\mathbf{n}} \times \mathbf{H}_n^s(\mathbf{r},t)\rangle = \sum_{l=1}^{L-1} \sum_j \langle \mathbf{j}_m(\mathbf{r}), -\hat{\mathbf{n}} \times \mathbf{H}_{n,j}^s(\mathbf{r},t)\rangle \quad \text{for } (l-1)T_s \leq j\Delta_t < lT_s - \Delta_t \tag{6}$$

It can be shown that $\langle \mathbf{j}_m(\mathbf{r}), -\hat{\mathbf{n}} \times \mathbf{H}_n^s(\mathbf{r},t)\rangle = 0$ for $t < t_{trans}$ where

$$t_{trans} = (R_c - 2R_s)/c + (l-1)T_s \tag{7}$$

and for $t \geq t_{trans}$

$$\langle \mathbf{j}_m(\mathbf{r}), -\hat{\mathbf{n}} \times \mathbf{H}_{n,j}^s(\mathbf{r},t)\rangle = \frac{\partial_t^2}{8\pi^2 c^2} \int_0^{2\pi} d\phi \int_0^{\theta_{int}} d\theta \sin\theta \left[\mathbf{S}_m^{h,-}\left(\hat{\mathbf{k}},t,\hat{\mathbf{n}}\right)\right]^T * \delta\left(t - \hat{\mathbf{k}}\cdot\mathbf{R}_c/c\right) * \left[\mathbf{S}_n^{h,+}\left(\hat{\mathbf{k}},t,\hat{\mathbf{k}}\right)\right] * T_j(t) \tag{8a}$$

where

$$\mathbf{S}_m^{h,\pm}\left(\hat{\mathbf{k}},t,\hat{\mathbf{v}}\right) = \int_s dS' \hat{\mathbf{v}} \times \mathbf{j}_m(\mathbf{r}') \delta\left(t \pm \hat{\mathbf{k}}\cdot(\mathbf{r}' - \mathbf{r}_c)/c\right) \tag{8b}$$

$\mathbf{r}_c$ is the center of the box which includes the basis $\mathbf{j}_m(\mathbf{r})$, the superscript T denotes the transpose, and $\hat{\mathbf{k}} = \hat{\mathbf{x}}\sin\theta\cos\phi + \hat{\mathbf{y}}\sin\theta\sin\phi + \hat{\mathbf{z}}\cos\theta$. Equation (8) holds provided that

$$\frac{cT_s}{R_s} \leq \frac{R_c}{R_s} - 2 \tag{9a}$$

$$\frac{cT_s}{R_s} \leq \frac{R_c}{R_s}(1 - \cos\theta_{int}) - 4 \tag{9b}$$

Equations (6) - (8) imply that if for a given source and observation sphere pair a cT_s/R_s and θ_{int} are chosen such that they satisfy eqns. (9), then the scattered field at the observer can be reconstructed as a superposition of plane waves.

The task of computing the current distribution at each time step is divided into computing near–field interactions using the usual MOT scheme, and far–field interactions using the PWTD algorithm. In order to do so, T_s is selected such that the constraints (eqn. (9)) are satisfied for the closest box pair which are in each other's far–field. The algorithm then follows the three step procedure:

1. Compute the projection of all the sources in a box on to outgoing rays from the box. This involves computing $\mathbf{S}_n^{h,+}(\hat{\mathbf{k}},t,\hat{\mathbf{k}})$. This is done for all boxes and all ray directions.
2. Project the rays from the source box to the rays entering an observation box when $t = t_{trans}$. Analogous to FMM, this operation is called translation, and it involves the convolution of $\delta(t - \hat{\mathbf{k}}\cdot\mathbf{R}_c/c)$ with $\mathbf{S}_n^{h,+}(\hat{\mathbf{k}},t,\hat{\mathbf{k}})$.
3. Finally, the rays entering all the spheres are projected on to the observers.

3 Results

PWTD augmented MOT schemes have been implemented for the MFIE. Here, we compare numerical results obtained using these PWTD based fast solvers to those obtained using classical MOT schemes. In all comparisons presented herein, plane wave Gaussian pulses traveling along $-z$ with an electric field polarized along $+x$ excite the scatterer. Figure 1 shows the cross–section of an arbitrary scatterer placed on a cartesian grid. Denoting the length of the largest side by a, it can be prescribed that if the distance between the centers of two boxes $d \leq 2a$ then they lie in each other's near–field. From fig. 1, it can be seen that the box pairs $(1,2)$, (1,3), (2,4) are in each other's near–field while (2,3), (1,4) and (3,4) are in each other's far–field. In all other examples that follow, a similar subdivision of the scatterer is used. In fig. 2, the current at a specific location is compared. The cylindrical structure was modeled using 918 spatial unknowns. Figs. 3(a) and 3(b) compare the current on a specific location on an almond and far scattered fields obtained using both approaches. The almond was modeled in terms of 2610 spatial unknowns. Finally, Figs. 4(a) and 4(b) compare similar data for a larger almond modeled in terms of 4680 spatial unknowns. The numerical results obtained using the PWTD based solver are in perfect agreement with those from the classical MOT solver. It should be noted that no instability is observed. In all our numerical experiments, we have observed break-even points of $N_s = 2000$; for larger problems, the PWTD accelerated schemes outperform the classical MOT algorithms.

4 Summary

This paper presented an algorithm that permits the fast analysis of transient elecromagnetic scattering phenomena. The computational complexity of this algorithm scales as $O(N_t N_s^{4/3} \log N_s)$ as opposed to $O(N_t N_s^2)$ complexity of the conventional MOT algorithm. The PWTD algorithm has been derived and implemented in the framework of the conventional MOT for the MFIE. It is seen that the agreement between the solution obtained using classical and PWTD accelerated MOT schemes is excellent. Multilevel schemes with a computational complexity of $O(N_t N_s \log N_s)$ are currently being implemented.

Acknowledgments: This work was supported in part by a grant from AFOSR via the MURI Program under contract no F49620-96-1-0025, the NSF grant ECS 95–02138, and the Gebze Institute of Technology.

References

[1] A. A. Ergin, B. Shanker, and E. Michielssen, "Fast evaluation of transient wave fields using diagonal translation operators," accepted for publication in *J. Comput. Phys.*.

[2] S. M. Rao and D. R. Wilton, "Transient scattering by conducting surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 56–61, 1991.

[3] B. P. Rynne and P. D. Smith, "Stability of time marching algorthms for the electric field integral equations," *J. Electromagn. Waves Applicat.*, vol. 12, pp. 1181–1205, 1990.

[4] M. J. Bluck and S. P. Walker, "Time–domain BIE analysis of large three dimensional electromagnetic scattering problems," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 894–901, 1997.

[5] B. Shanker, A. A. Ergin, K. Aygün and E. Michielssen, "Computation of transient scattering from electrically large structures using the plane wave time domain algorithm," *Submitted to IEEE Antennas Propagat. Symposium*, 1998.

[6] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. 30, pp. 408–418, 1982.

[7] G. Manara, A. Monorchio, and R. Reggiannini, "A space–time discretization criterion for a stable time–marching solution of the electric field integral equation," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 527–532, 1997.

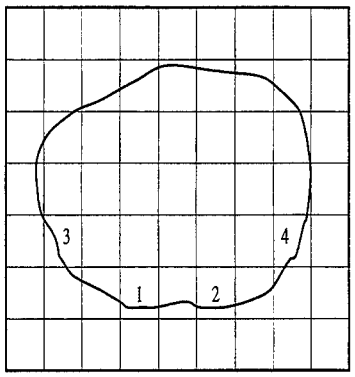


Figure 1: A cross–section of the scatterer placed on a cartesian grid. Box pairs (1,2), (1,3), (2,4) are in each other's near–field while (3,4), (2,3) and (1,4) are in each other's far–field.

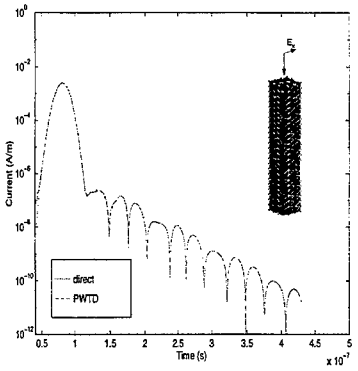
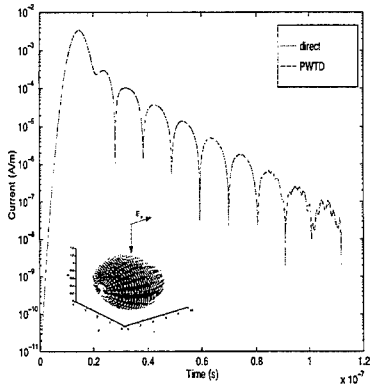
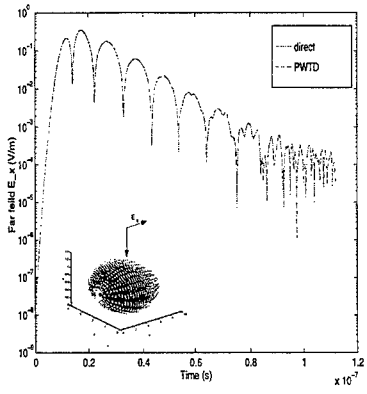


Figure 2: Scattering from an almond computed using the MFIE. The scatterer is discretized using 918 unknowns. The dimensions of the cylinder are $1 \times 1 \times 10\ m^3$, and the incident pulse has significant spectral content up to $f = 200$Mhz.

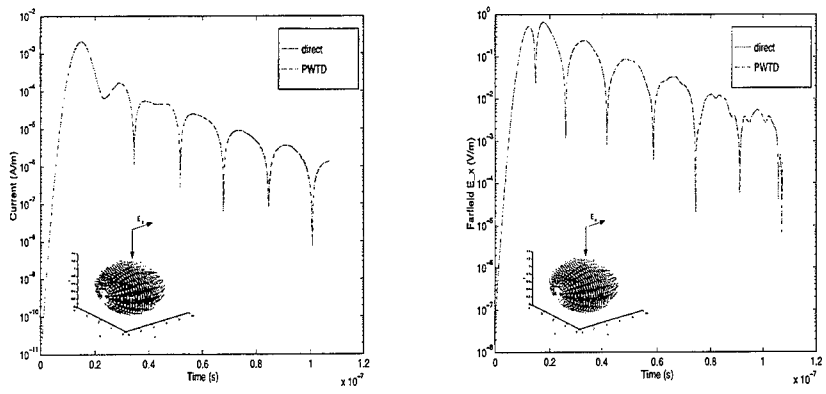


(a)



(b)

Figure 3: Scattering from an almond computed using the MFIE. The scatterer is discretized using 2610 unknowns. The largest linear dimension of the almond is $3m$, and the incident pulse has significant spectral content up to $f = 503$Mhz. (a) Current at a location on the almond; (b) The backscattered far–field E_x



(a) (b)

Figure 4: Scattering from an almond computed using the MFIE. The scatterer is discretized using 4680 unknowns. The largest linear dimension of the almond is $5m$, and the incident pulse has significant spectral content up to $f = 404$Mhz. (a) Current at a location on the almond; (b) The backscattered far–field E_x

SESSION 15:

EMI/EMC

Chairs: T. Hubing and B. Archambeault

EMC MODELING OF SHIELDED ENCLOSURES WITH APERTURES AND ATTACHED WIRES IN A REAL-WORLD ENVIRONMENT

Bruce R. Archambeault
IBM

Kent Chamberlin
University of New Hampshire at Durham

Omar Ramahi
Digital Equipment Corp.

Background

In 1981, the Federal Communications Commission (FCC) established limits on the strength of electromagnetic radiation allowable from computing devices sold in the United States. In the past few years, various other countries have developed similar standards, mostly to control these devices' potential to interfere with data communications systems, broadcast radio and television, and emergency systems. The European Community has recently standardized these requirements in Europe, and expanded them to include not only computing devices, but nearly every product containing digital electronics.

The result of these various regulations is that all manufacturers, not only computer manufacturers, must pay close attention to the electromagnetic interference (EMI) levels that their products produce. Pressures to shorten design cycle times, reduce product costs, and meet EMI regulations has served to increase the interest in using modeling and simulation to help ensure optimum hardware designs. An optimum design will ensure only required EMI features are included, since it is no longer acceptable in industry to simply use excessive shielding, costly filters, small aperture air vents, and other such fixes to meet EMI regulations.

Typically, electromagnetic radiation testing is performed as a *system*, that is, all the various parts that are generally used together must be tested together. In the case of a most personal computers, this would include the computer system box, monitor, keyboard, mouse, and printer. Also, any other cables that might be connected to the units must be included, for example, modem cables, speaker cables, etc.

The work presented here focuses on modeling the entire problem, that is, the radiation from a source within a shielded enclosure, the coupling of that energy to the outside via apertures in the enclosure, the effect of wires connected nearby those apertures, and the test environment itself. Because of the complex nature of the problem, it is impractical to accurately model this problem without using multiple modeling stages and different modeling approaches. The method to be presented here makes possible the modeling of configurations that were previously considered impractical. This method also makes possible direct comparisons of the simulation results to the regulatory limits to predict pass/fail of the device.

Practical EMI/EMC Problem and Test Environment

Although there is a large number of different types of products that must meet EMI/EMC regulations, most fall into the general class of products with shielded enclosures containing apertures and having long wires attached to the enclosure. Plastic enclosures are often shielded either by a metal internal coating or by metal fragments imbedded in the plastic during the molding process. Computer products, consumer electronics products, and communications devices all fit this category.

The source of the radiated emissions is usually a high-speed (fast rise time) clock or data signal on the printed circuit board within the shielded enclosure. The source creates a complex electric and magnetic field structure within the enclosure. Some of this energy 'leaks' out through the apertures (e.g. air vents, slots between option cards, shielded enclosure seams) and creates RF currents on the outside of the shielded enclosure. These currents are then distributed over the entire outside structure

(including wires, cables, etc.), and radiate into the outside environment. The fields are then measured 10 meters away in the presence of a ground reference plane, as described earlier.

The products under test typically have long wires attached to different connectors (power cords, modem lines, printer cables, etc.) which will greatly affect the radiated emissions from the product. RF currents that have leaked out from an aperture and are on the outside of the metal shield will couple onto the wires and cables. The wires will greatly increase the effective aperture of the 'antenna' (the equipment under test or EUT) since the overall size of the EUT with wires is typically increased by more than an order of magnitude by the presence of the wires.

Using contemporary computational technology and techniques, there is no practical way to model the entire problem described above with a single model. Earlier work successfully modeled certain aspects of the overall problem, (e.g. radiation from printed circuit boards (PCB) with a microstrip near a reference plane edge [1][2], PCB via's [2], decoupling capacitor placement [3], or shielding through apertures [4][5][6][7][8]). However, these earlier efforts have addressed only specific facets of the overall problem, and, therefore, were not adequate to predict compliance with regulatory standards. For example, in [1] and [2], emissions from an unshielded printed circuit board (PCB) with a microstrip line was modeled. No attempt to include a shielded enclosure with apertures was made. In [4]-[8] emissions through apertures in a infinite metal sheet were modeled, but no attempt was made in these previous studies to include a PCB as the source, nor to include the required measurement environment. These studies were useful to help understand specific phenomena, but did not include all the parts of the overall problem to allow for a comparison to the regulatory limits.

The strengths of the two modeling approaches implemented in the hybrid technique allow a source, a shielded enclosure with apertures, and the required measurement environment. Thus the results of the overall problem can now be compared to the regulatory limits for pass/fail analysis. Other internal features, such as partial shielding walls, extra cables, etc. can be included as required. This hybrid technique uses the Finite-Difference Time-Domain (FDTD) method to model the source and the inside of the shielded enclosure, including the effects of the apertures. The Method of Moments (MoM) approach is used to model the outside of the shielded enclosure, including attached wires, and the test environment.

The Stage-One FDTD Aperture Model

The electric field strength of a particular location within the aperture is frequency and position dependant. The FDTD technique was used to model the aperture in the infinite metal plate by extending the metal plate to ABC. A diagram of the FDTD computational space used for single-aperture modeling is shown in Figure 1. The electric field within the aperture was found to vary across the aperture with the maximum value at the center of the aperture for frequencies below the first resonant frequency of the aperture.

The Stage-Two MoM Aperture Model

Creating an infinite metal plane with an aperture using the MoM technique is not as practical as in the FDTD technique. In order to simulate an infinite metal plane, a single, electrically large, wire mesh plate is created. Wire mesh can be used successfully to simulate a solid plane [11][12], especially when another wire is to be connected to the structure, as is discussed below. The wire mesh should be small compared to the wavelength of interest in order to insure the currents flow over the sheet as if it was a solid sheet of metal.

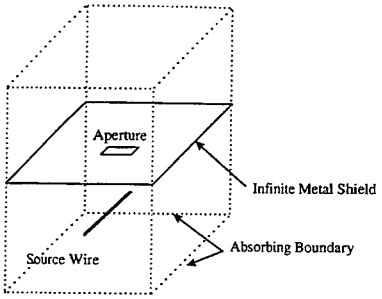


Figure 1 FDTD Model for Single Aperture

A 10 mm x 2 mm aperture is modeled as a hole in the wire mesh of a large plate using the MoM technique. Since it is desired for the model to be accurate to about 15 GHz (the first resonant frequency of this size aperture), the segment sizes are selected to be no larger than about 2 mm for the MoM model. Figure 2 shows this wire mesh model. The radial wires at the corners and sides are used to increase the effective size of the plate to simulate an infinite plate.

The original FDTD 10 mm x 2 mm aperture problem is repeated with a larger computational domain in the outside area. The electric field level at a distance of 65 mm away from the aperture is found using FDTD. This distance is selected to create a reasonably sized FDTD computational domain while allowing the fields to be found in the far field at frequencies above about 770 MHz (one-sixth lambda from the aperture to the observation point).

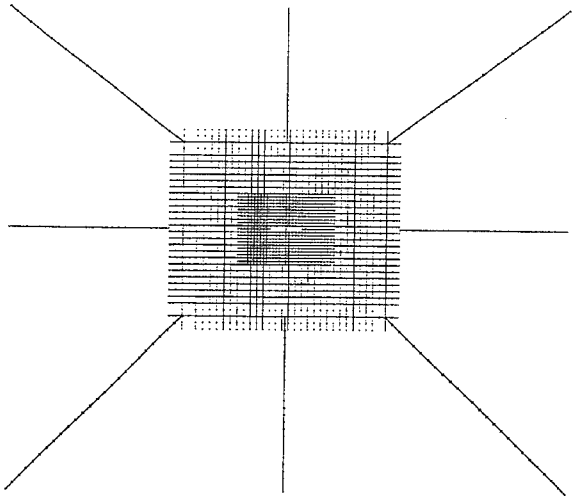


Figure 2 MoM Wire Frame Model with Single 10x2 mm Aperture

The electric field level is found in the aperture in the Stage One Model and then is corrected at low frequencies using the Herzian dipole impedance method. [13] The corrected electric field is then used as the voltage source across the aperture in the MoM (Stage Two) model. A single electric field source is placed across the center of the aperture and set to the maximum value (center location) of the Stage One model results. The radiated electric field results are compared in Figure 3, and show a good agreement between modeling techniques between about 4 - 15 GHz, thus validating the use of MoM as a second stage in the hybrid approach to modeling apertures.

The lower frequencies (typically below 4 GHz) in the FDTD model results may be affected by the absorbing boundary condition. Since it is not always practical to increase the FDTD computational domain to a point where there are no ABC effects, the results in the FDTD case may be limited in that the low frequency information may be in error. However, since the hybrid FDTD/MoM model results at those same low frequencies use the corrected aperture fields, the hybrid FDTD/MoM results show the correct low frequency electric field values.

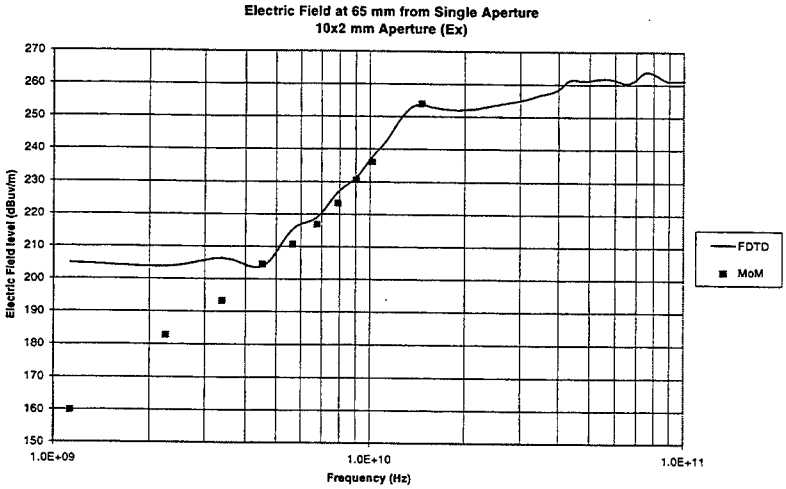


Figure 3 Comparison Between MoM and FDTD at 65mm from Single Aperture

Comparisons between FDTD-only and Hybrid FDTD/MoM models for apertures in a metal plane show very good agreement over the range of frequencies where the FDTD-only model is valid (above the frequencies where the ABC introduces errors). The hybrid approach demonstrates its strength by improving upon the FDTD-only approach at low frequencies.

Hybrid models showing the effects of real-world test environment configurations demonstrate the importance of including these features in the model, and show how effectively the hybrid approach can include them.

Example of the Hybrid Modeling Technique

The first step (Stage One) to use this hybrid modeling technique is to create the FDTD model of the enclosure, aperture(s), the internal source and whatever internal structure is considered important. In the case of an empty shielded enclosure (100 mm cube) with a 10 x 2 mm aperture, the FDTD cells must be small enough to describe the aperture correctly. For this example, a FDTD cell size of .5

mm is selected. This size is also small enough to provide at least 10 cells per wavelength up to the highest frequency of interest, as per FDTD approach requirements.

The source is selected to be a simple current on a wire and placed near the aperture. As described earlier, this is representative of a PCB ground reference plane edge. The wire is oriented perpendicular to the aperture to ensure maximum possible emissions coupled through the aperture.

Figure 4 shows a diagram of the FDTD model. The aperture is placed on the top face of the enclosure for convenience, but could be on any side desired. The top part of the enclosure is extended beyond the enclosure walls to restrict any external resonances from affecting the fields in the aperture. The internal structure of the enclosure is maintained to allow any internal resonances to occur. Both the electric and magnetic field at the center of the aperture is saved as the output from this Stage One model.

The time domain electric and magnetic field results are then converted to the frequency domain using a Fast Fourier Transform (FFT). The frequency domain impedance (E/H) is examined to determine if errors occurred due to the ABC's close proximity to the aperture at low frequencies which are of interest. The electric field is then corrected using the Herzian dipole technique.

The same enclosure and aperture is modeled in MoM for Stage Two using a wire mesh frame with the openings in the wire mesh small compared to the shortest wavelength of interest. The corrected electric field is then applied across the center of the aperture in the MoM model for each frequency desired.

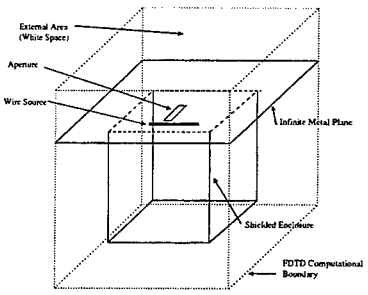


Figure 4 FDTD Example Model of Shielded Enclosure and Aperture

Hybrid Model Comparison Between Free Space and Real-World Test Environment

As stated earlier, it is important to model the test environment correctly. The following examples demonstrate the effects of the environment on the final results. As mentioned earlier, EMI emissions measurements are required to be made over a ground plane. The receive antenna must be 10 meters away, and it must be scanned (for maximum receive level) over a one to four meter height while rotating 360 degrees. The scanning of the antenna height ensures there is no chance of a destructive interference path artificially lowering the measured emissions levels. The rotation of the EUT through the 360 degrees ensures the maximum emissions are received, regardless of any possible directionality of the EUT's radiation pattern.

Figures 5 and 6 show the model results for the shielded enclosure EUT with and without the ground plane present for both the horizontal and vertical polarizations. The presence of the ground plane, and the effect of scanning over the height range, can greatly increase the measured emissions level due to the reflected wave adding in phase to the direct wave.

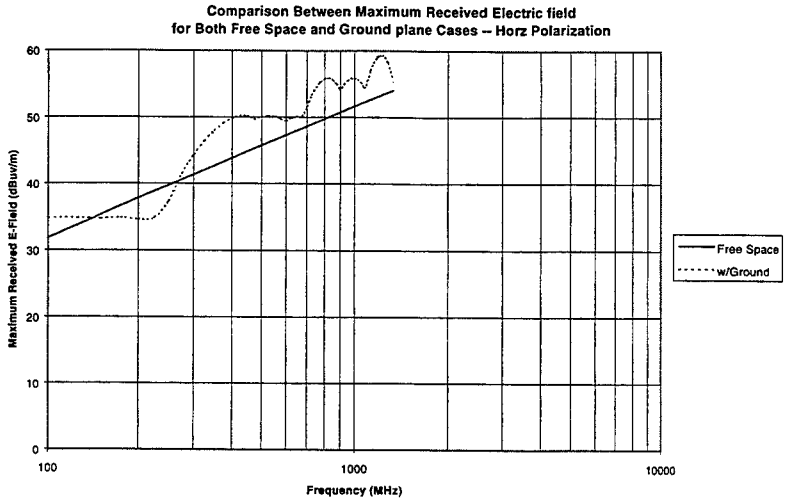


Figure 5 Maximized Electric Field Comparison with and without Ground Plane (Horizontal Polarization)

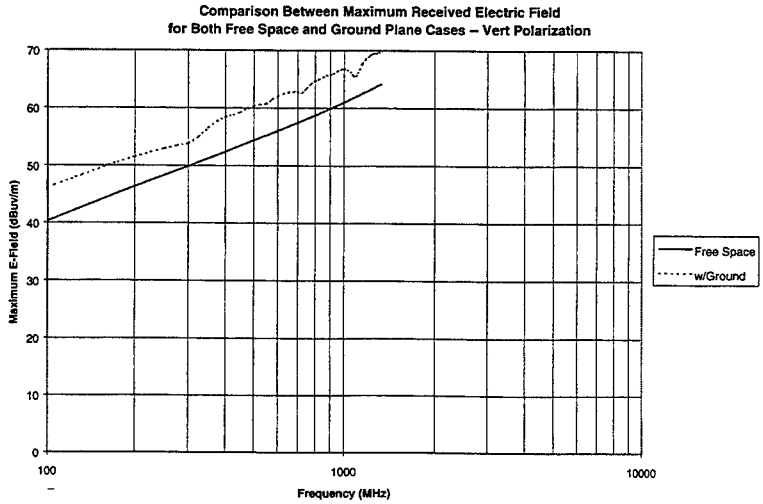


Figure 6 Maximized Electric Field Comparison with and without Ground Plane (Vertical Polarization)

EMI Emissions test standards also require that all cables be attached to the EUT. This effectively increases the EUT's electrical size, and typically increases the emissions levels significantly at some frequencies. A single cable, one meter long, is now attached to the initial enclosure model, as shown in Figure 7. This cable is attached directly to the enclosure shield, as in the case of a cable shield being 'grounded' to the case.

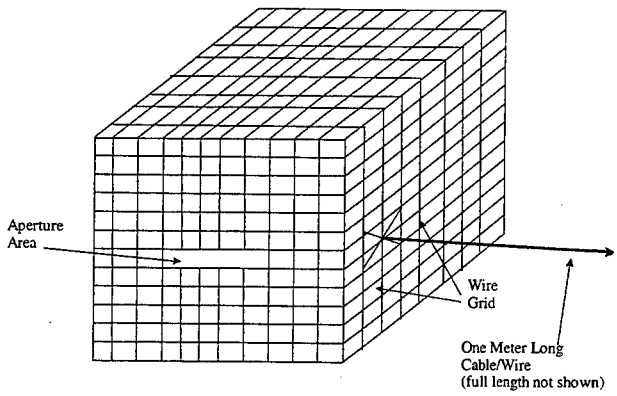


Figure 7 MoM Model of Shielded Enclosure with 1 meter Cable Attached

The same electric field is applied across the aperture for this new configuration. The maximum received emissions are greatly increased, as seen in Figure 8, due to the addition of the cable. This demonstrates the importance of the hybrid model including all of the test environment features.

Summary

A hybrid approach to modeling a complex real-world configuration has been shown to provide accurate results. The inside of a shielded enclosure is modeled using the FDTD approach. The electric fields in the aperture are found during the FDTD simulation and then used as the source for the second stage MoM model. In cases where close proximity to the ABCs during the FDTD simulation resulted in electric field errors at low frequencies, a method to correct those errors using Herzian dipole impedances is provided. Results are then shown for a variety of real-world test configurations showing the importance of including all the configuration details in the model.

References:

[1] B.R Archambeault, "Modeling of EMI Emissions from Microstrip Structures with Imperfect Reference Planes," *Applied Computational Electromagnetics Society Symposium*, March 1997, pp. 1058-1063.

[2] S. Poh, "Electromagnetic Emissions from Microstrip Discontinuities in Interchip Packaging," Internal Digital Equipment Corporation Report, 13 Feb 91.

[3] S. Daijavad, H. Heeb, "On the Effectiveness of Decoupling Capacitors in Reducing Electromagnetic Radiation from PCBs," *1993 IEEE International Symposium on Electromagnetic Compatibility*, pp. 330-333.

[4] B.R. Archambeault, C. Brench, "Shielded Air Vent Designs Guidelines from EMI Modeling," *1993 IEEE International Symposium on Electromagnetic Compatibility*, pp. 195-199.

[5] D.M. Hockanson, J.L. Drewniak, T.H. Hubing, T.P. VanDoren, "Application of the Finite Difference Time Domain Method to Radiation from Shielded Enclosures," *1994 IEEE International Symposium on Electromagnetic Compatibility*, pp. 83-88.

[6] B.L. Brench, C.E. Brench, "Shield Degradation in the Presence of External Conductors," *1994 IEEE*

International Symposium on Electromagnetic Compatibility, pp. 269-273.
[7] B.R. Archambeault, C.E. Brench, "Modeled and Measured Results from Two Proposed Standard EMI Modeling Problems," *1995 IEEE International Symposium on Electromagnetic Compatibility*, pp. 349-352
[8] B. R. Archambeault and K. Chamberlin, "Modeling the EMI Performance of Various Seam Shapes", *Applied Computational Electromagnetics Society International Symposium*, 1995
[9] K. Chamberlin and L. Gordon, "Modeling Good Conductors Using the Finite-Difference Time-Domain Technique, "*IEEE Trans. Electromagnetic Compatibility*, vol. 37, No. 2, May 1995.
[10] J.D. Kraus, *Antennas*, McGraw-Hill Book Company, Second Edition 1988
[11] Ludwig, A. C., "Wire grid modeling of surfaces," *IEEE Trans. Antennas and Propagation*, vol. AP-35, No. 9, September 1987, pp. 1045-1048
[12] Archambeault, B. R., R. Mellitz, "EMI Prediction using wire mesh boxes in NEC," *Applied Computational Electromagnetics Society Symposium*, March 1993.
[13] Archambeault, B.R., "Modeling of the Electromagnetic Radiation from Shielded Enclosures with Apertures and Attached Wires in a Real-World Environment," Ph.D Dissertation, Electrical and Computer Engineering Department, University of New Hampshire, 1997.

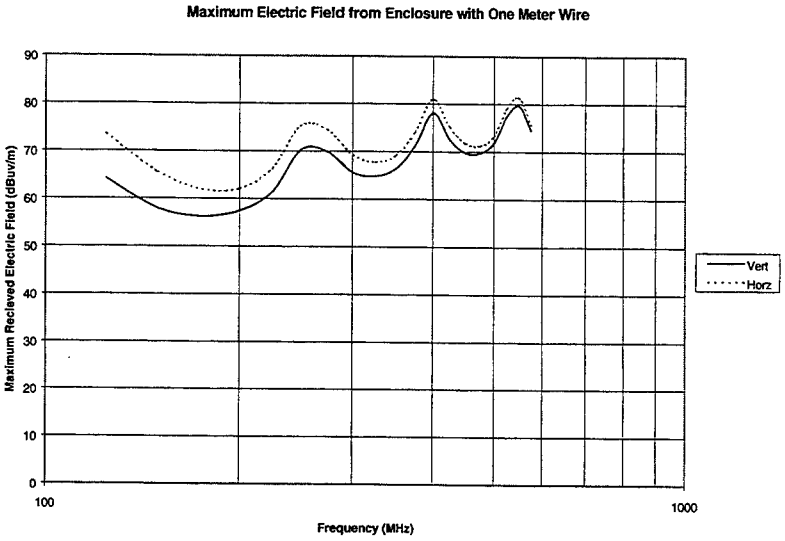


Figure 8 Maximum Received Electric Field from EUT with One Meter Wire Attached

Proposed Standard EMI Modeling Problems for Evaluating Tools which Predict Shielding Effectiveness of Metal Enclosures

Bruce Archambeault
IBM

Omar Ramahi
Digital Equipment Corporation

Introduction

Numerical modeling tools are becoming very popular for a variety of EMI/EMC applications. Metal shields around printed circuit boards remain one of the primary techniques used to control emissions and provide immunity. Predicting the shielding effectiveness of these metal shields is complex, but certain full wave modeling techniques can be used to predict the shielding performance, for both near-field and far-field emissions.

Not all numerical modeling techniques are equal. Every technique has strengths, that is, certain types of applications where it excels, as well as weakness where it can not efficiently perform the modeling necessary. The Method of Moments (MoM) and the Finite-Different Time-Domain (FDTD) technique are the two most commonly used modeling techniques for EMC shielding applications. This paper presents the results of modeling the same shielding configurations with a number of different modeling tools, using both MoM and FDTD, and demonstrates the limitations of these techniques against this application.

Since shielding effectiveness is very dependent upon the actual test/model configuration, a set of standard shielded enclosures is proposed and then each is evaluated using each of the modeling techniques. Two different size enclosures are included, and each with two different size apertures. The enclosures were also modeled without apertures to show the dynamic range limitation of both modeling techniques.

Shielding Effectiveness

The term 'shielding effectiveness' is somewhat misleading. Most EMC engineers understand that a shielding effectiveness test consists of a comparison between the radiated emissions with an enclosure and without the enclosure. A simple antenna is typically used as the source antenna, with a receive antenna at some distance away. Once the source antenna is placed within an enclosure, the source antenna's characteristics are changed, and so the comparison is not really consistent. However, this test is commonly used, and the results are consistent, as long as the test configuration is maintained. If the test configuration is changed in any way, then the comparisons between tests are not valid, hence the importance of a standard set of shielding effectiveness configurations.

Proposed Standard Shielding Effectiveness Modeling Configurations

Two different sized rectangular enclosures were developed for use as the standard shielding effectiveness model applications. Figure 1 shows a diagram of the basic model, and Table 1 lists the various dimensions for each of the different models.

Figure 1 Standard Shielding Box Diagram

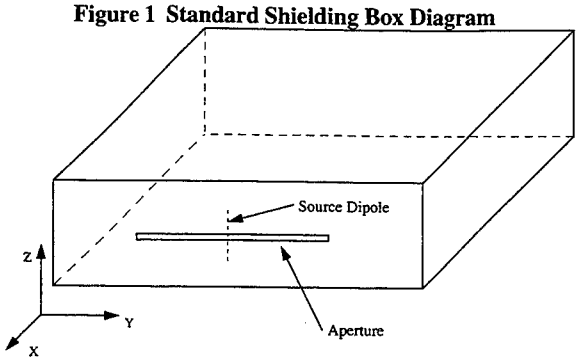


Table 1 Standard Enclosure Sizes

Enclosure	X-size (cm)	Y-size (cm)	Z-size (cm)	Aperture size (cm)
Large Box (1)	40	50	30	25 x 1
Large Box (2)	40	50	30	10 x 1
Small Box (1)	25	30	15	25 x 1
Small Box (2)	25	30	15	10 x 1

In all cases, the shielded enclosure was a rectangular metal box with an aperture centered in the 'front' face (z-y plane) of the enclosure. The source was a 4.5 cm dipole placed 5 cm inside the aperture, and centered. The polarization of the source was selected to be perpendicular to the long dimension of the aperture. The 4.5 cm source dipole consisted of two 2.0 cm long wires, separated by 0.5 cm gap.

The MoM Model

The MoM model was constructed using surface patches. The surface patches were constrained to be no larger than 1/10th lambda at the highest frequency (in this case, 1 GHz). The source dipole was driven using a delta-gap voltage source.

The FDTD Model

The FDTD model was constructed using cubical cells with 1/2 cm side dimension. The dipole was driven using a constant electric field source positioned between the dipole halves. The temporal waveform employed was a differentiated Gaussian function with a width sufficient to give a frequency range up to 10 GHz.

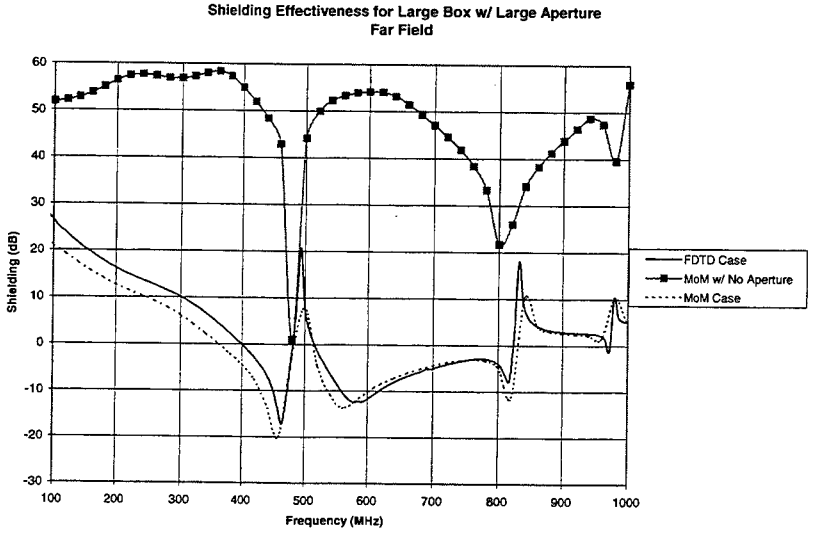
Model Results

The results from the various model configurations were normalized to the case of the source antenna without any shielded enclosure. Far field results were taken 10 meters from the enclosure and directly in front of the aperture. Near field results were taken 1 cm in front of the aperture. -In the case of the MoM models, a no-aperture case (that is, completely closed shielded enclosure) was modeled to show the dynamic range limitation when using this technique. There is no corresponding dynamic range limitation with the FDTD technique, so no closed box case was modeled in FDTD.

Figures 2 and 3 show the shielding effectiveness for the large box with the large aperture for the far field and the near field cases, respectively. The results from the FDTD model and the MoM model agree well within typical EMC tolerances. Work has been done in the past to show closer agreement between FDTD and MoM but would require careful manipulation of the source in the two modeling techniques. This was considered unnecessary for this application, since FDTD and MoM agreement was not the primary purpose of this work.

Figures 2 and 3 also show the maximum amount of shielding possible from the MoM model. This serves as the dynamic range limitation of the MoM model (for shielding effectiveness use) for this configuration. Note that the resonance frequencies excited by this source are clearly shown in this dynamic range curve, and that in the areas of resonance, there is little or no dynamic range available in the MoM model.

Figure 2



Figures 4 and 5 show the shielding effectiveness for the large box with the small aperture for the far field and near field cases, respectively. Again, the normalized results agree between the FDTD and MoM models. The amount of shielding provided with the small aperture was greater than with the large aperture.

Figures 6 and 7 show the shielding effectiveness for the small box configuration with the large aperture in the far field and near field, respectively. As in the large box case, the resonant frequency due to box dimensions is clear. In this size box the additional resonance due to the aperture length is also clearly visible. In the large box example, the aperture resonance was overshadowed by the box resonance.

Figure 3

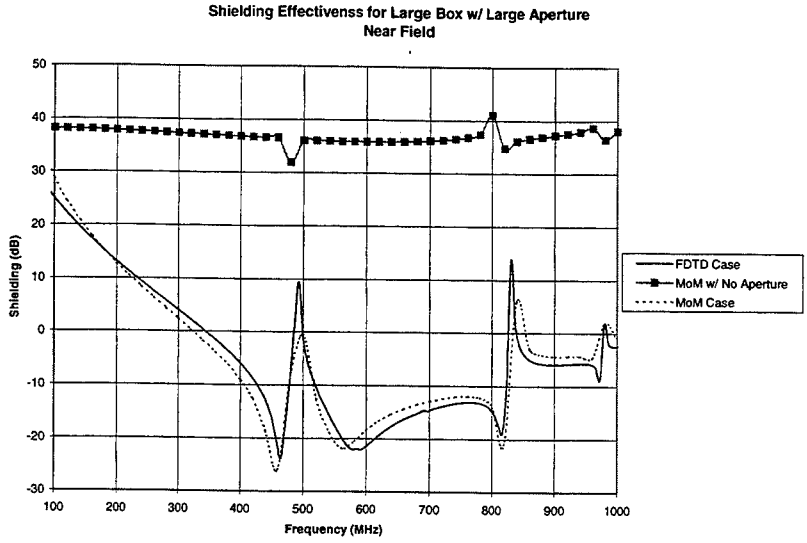


Figure 4

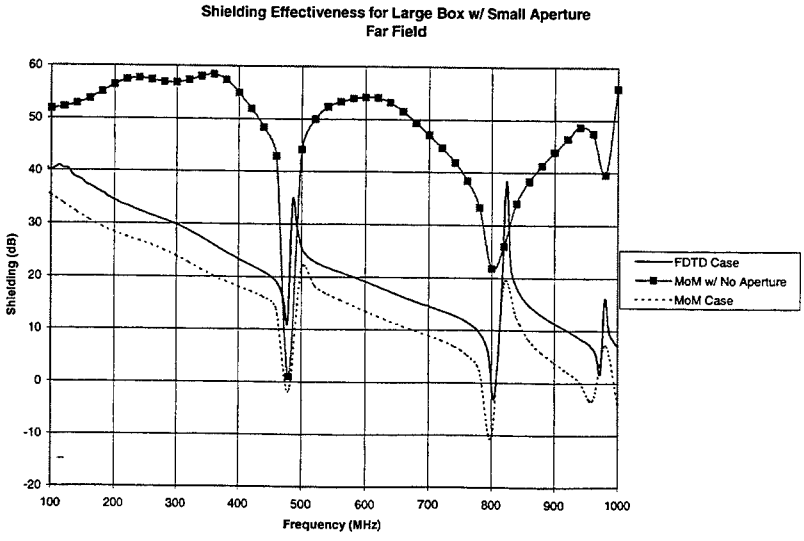


Figure 5

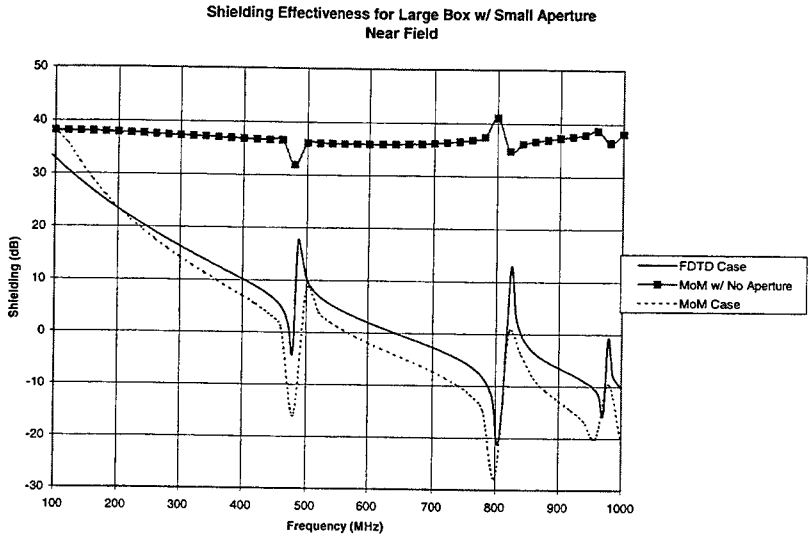


Figure 6

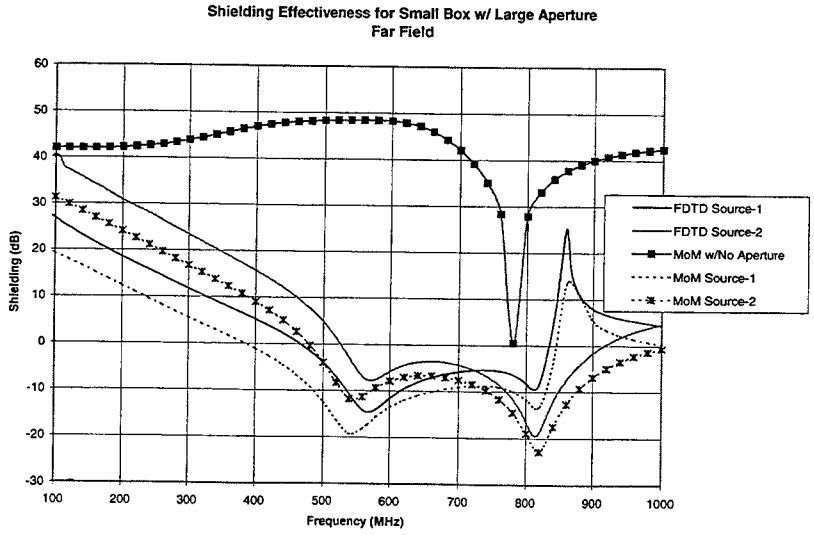
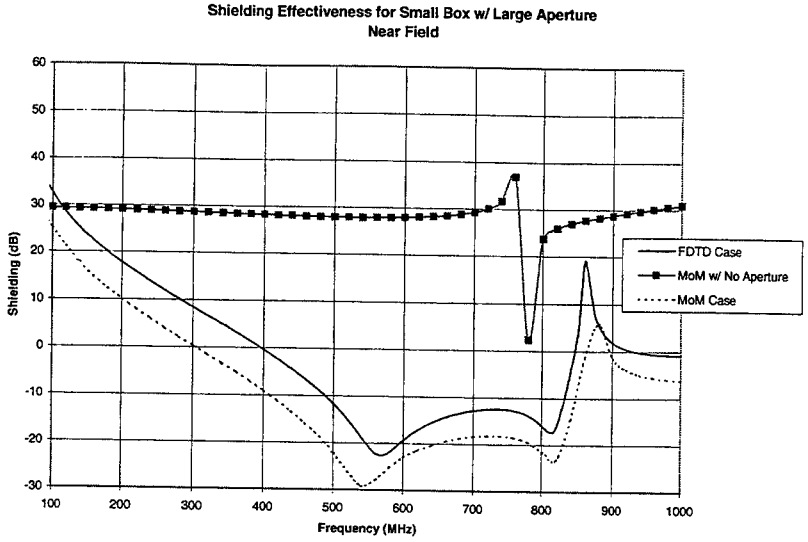


Figure 7



As an additional test, the source position was moved from 5 cm away from the aperture to 15 cm away from the aperture. Figure 6 shows the difference in apparent shielding effectiveness due to the source position, which indicates the need to maintain consistency between shielding effectiveness testing and modeling.

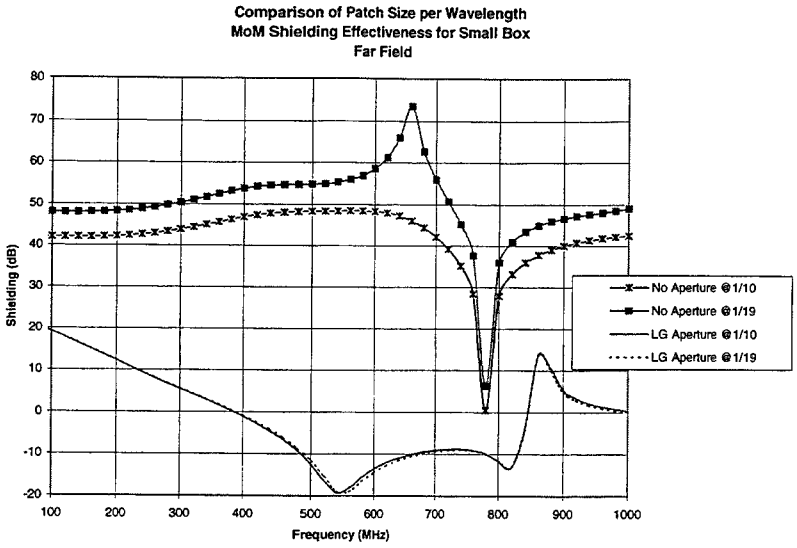
Figure 8 shows a comparison of the MoM shielding effectiveness for the small box with the large aperture for two different patch densities. The MoM patches were reduced to the point where there was at least 19 patches per wavelength at 1 GHz. The shielding effectiveness remained unchanged but the maximum dynamic range increased with the finer patch resolution.

Model Parameters

As can be seen from the previous figures, the MoM models determined the shielding results at 50 discrete frequencies between 100 MHz and 1 GHz. The large box model required 4300 patches while the small box model required 1900 patches. The large and small box models required about 500M bytes and 100M bytes (respectively) of RAM and approximately two hours to complete the large box problem and about 30 minutes to complete the small box problem on a typical UNIX workstation.

The FDTD models for the large box and the small box contained 1.1 Meg and .32 Meg cells, respectively. For resonant structures such as the boxes under study in this work, it was found that for the frequency range of interest (100 to 10,000 MHz), a modest-accuracy absorbing boundary condition such as Liao's 2nd order operator is sufficient to yield good accuracy. The large and small box models required about 42 Mbytes or 12.8 Mbytes of RAM (respectively) and a slightly longer time to run the models (3 hours and 1 hour) but provided results up to 10 GHz (not shown) as well as a wide dynamic range.

Figure 8



Conclusions

Both the MoM and the FDTD techniques can provide accurate shielding effectiveness model results as long as the results are within the dynamic range of the basic model. The MoM technique's dynamic range was limited to 40 - 50 dB unless extremely fine segmentation was used, while the FDTD technique's dynamic range is effectively unlimited. The dynamic range of the MoM models varied depending on the size of the basic enclosure as well, indicating that care must be taken to determine shielding results are truly due to the test configuration and not an artifact of the MoM technique.

When a limited number of frequencies are required for the analysis, then MoM allowed faster results. However, when a wide frequency range was required, or the resolution between frequencies must be fine to ensure all resonances are found, then FDTD was a much faster solution. The FDTD models required less RAM to run than MoM, making it a more attractive option for many applications.

Overall, either modeling technique can be used for shielding effectiveness applications as long as care is taken to understand the limitations of the modeling technique being used.

A study in the Proper Design of Grounding for SMPS Converters and the role of CEM.

Reinaldo Perez
Jet Propulsion Laboratory
California Institute of Technology

Abstract

Converters that are used in Switching Mode Power Supply (SMPS) are usually well designed internally so as to maximize efficiency and minimize output noise. However, deficiencies in the distribution of power and grounding within the PCB where the SMPS converters will be located can negate many of the advantages provided by these converters. It is shown that inductive and capacitive effects in the way SMPS converters are used in a PCB are one of the major causes of high conductive emission in PCBs. Furthermore, it is shown that changes in grounding layouts can affect these inductive and capacitive effects. Finally, it is show how CEM tools can be used in the modeling of these parasitic effects. The emphasis on this paper is on design principles.

1.0 Introduction

The switching mode power supply [1] is a class of power supply that makes use of electronic switching to process electrical power. Because ideal switches do not dissipate power, the SMPS can be designed to have a high efficiency. In the SMPS a high frequency of switching is used and the size of the transformer and filtering circuits can be minimized. Because of the great advantages, the SMPS has become the power processing unit of choice in low power circuits or in circuits where interference must be kept to a minimum.

The heart of a SMPS is a dc-to-dc converter. The converter accepts a dc input and produces a controlled dc output. The three basic types are the buck converter, the boost converter, and the buck boost converter. For each of these converters there is an electronic switch that is driven on/off a high frequency (5-500 KHz). It is the duty cycle of the electronic switch which controls the dc output voltage V_{out}. There is an output filtering capacitor C_{out} which is used to smooth out the ripple components of the output voltage resulting from the high frequency switching. By adding a feedback circuit in a converter, the output voltage of the converter can be regulated. In each converter circuit there is a energy-storage inductance L which can be chosen large enough so that the current in it is substantially smoothed [2-3].

Because SMPS dc-dc converters are very sensitive to input noise there is a need to filter out the noise from the inputs of dc-dc converters as much as possible. For that purpose an EMI filter is chosen as shown in Figure 1. The figure shows a single EMI filter (FM-461) at the inputs of several dc-dc converters which are used to supply different voltage levels for different loads. The filter EMI modules are specially designed to reduce the input line reflected ripple current of dc-dc converters.

Inside the dc-dc converter good design techniques are applied to minimize the output noise form such converters. It is well understood by designers that a clean output voltage is essential for the proper functioning of ICs and specially analog devices and circuits which are highly susceptible to bus voltages noise. Therefore, concerning noise issues in dc-dc converters is as important to control input noise as output noise. We first outline some principles of limiting input noise and then concentrate on the main subject of the paper concerning limiting output noise from PCBs where converters are found.

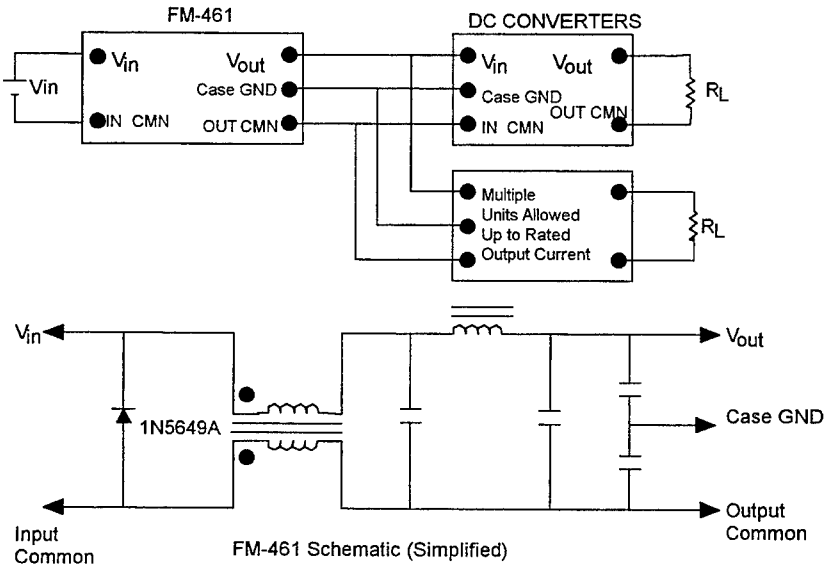


Figure 1.0 EMI Filter serving a cascaded line of dc-dc converters.

These filters are intended for use in applications of high frequency switching (100 KHz). These EMI filters are capable of reducing the input ripple current by as much as 40 dB within the frequency band of 100 KHz to up to 100 MHz.

2.0 Transients Effects in SMPS.

In a simplified circuit of a switching mode power supply an error amplifier compares output voltage V_{out} with a reference V_{ref} and controls the duty cycle, D, via a pulse width modulator as shown in Figure 2. The output capacitor C_{out} is represented by its equivalent circuit that includes the equivalent series resistance (ESR) and the equivalent series inductance (ESL). When we have a load step ΔI current through the choke inductance L can not be instantly changed. There will always be a finite time t needed for L to accommodate ΔI and is given by the expression:

$$t > \frac{L\Delta I}{(V_{in}D_{max}) - V_{out} - V_{diode}} \qquad (1.0)$$

where D_{max} is the maximum duty cycle and V_{diode} is the diode's voltage drop. The choke current I_{choke} slews to the new load current but before it does that I_{load} flows through C_{out}. This results in an output voltage deviation ΔV_{out} that may be as much as

$$\Delta V_{out} \leq \frac{ESL(dI_{load})}{dt} + ESR \bullet \Delta I \qquad (2.0)$$

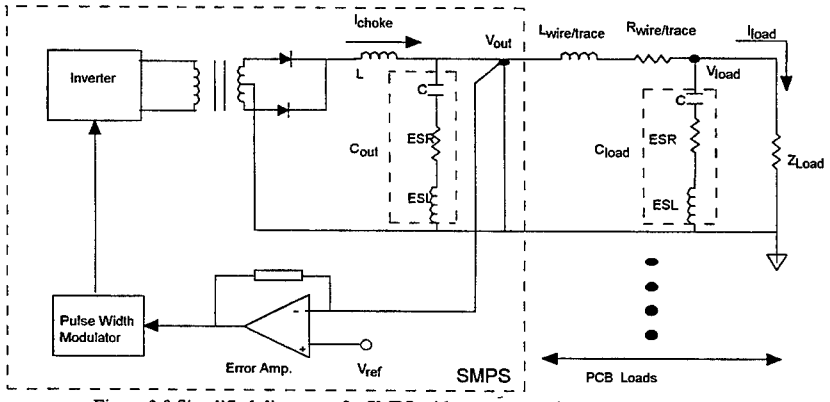


Figure 2.0 Simplified diagram of a SMPS with output capacitor model (ESR & ESL)

where dI_{load}/dt is the load's current slew rate (amps/sec).The SMPS's C_{out} acts as a reservoir for these current transients. The delay that is observed is compounded due to the wiring and possible long traces in the PCB. As can be observed in Figure 2 traces have self inductances and resistances and when I_{load} changes from Faraday's law $L_{wire/trace}$ will cause an initial voltage deviation ΔV given by

$$\Delta V = \frac{-L_{wire/traces}(dI_{load})}{dt} \qquad (3.0)$$

Furthermore, $R_{wire/trace}$ will cause an input voltage drop as I_{load} slews. The time that is needed to change a current through load wires/traces in a PCB is given by the expression

$$t = t_{delay} + t_{rise} \qquad (4.0)$$

where t_{delay} is the SMPS delay time and t_{rise} is the time needed for $I_{wire/traces}$ to catch up to the load current and given by

$$t_{rise} = t_{delay}\left(\frac{\frac{dI_{load}}{dt}}{\left(\frac{V_{max}-V_{load}}{L_{wire/trace}}\right)-\frac{dI_{load}}{dt}}\right) \qquad (5.0)$$

where V_{max} is the maximum output voltage during the transient recovery of the supply. The output load will experience a dip of as much as

$$\Delta V_{out} \approx ESR\left(\frac{t\,dI_{load}}{dt}\right)+t^2\left(\frac{dI_{load}/dt}{C_{load}}\right) \qquad (6.0)$$

A computer simulation of a circuit of the type shown in Figure 2.0 using SPICE can show the effects of load wire/traces inductances and external capacitance as shown in Figure 3.0.

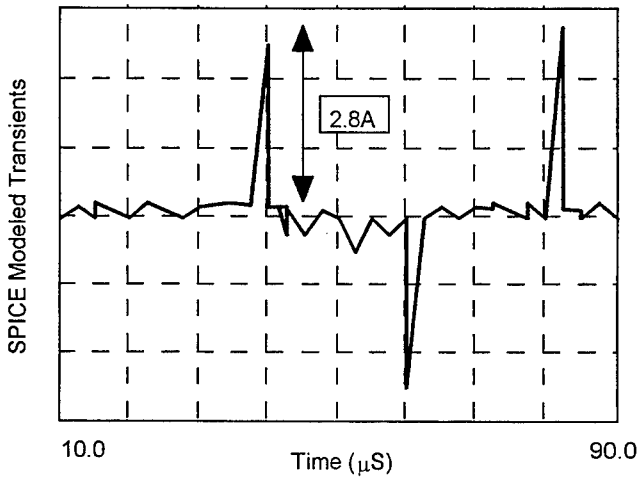


Figure 3.0 Load Transients resulting from modeling output loads of a SMPS

3.0 Proper Grounding to Suppress Transient Effects

Since parasitic inductance and resistive effects in the loads of SMPS are greatly responsible for the transients effects as shown in Figure 3.0 (the Fourier Transform of Figure 3.0 will show a series of discrete frequencies conducted emissions) the efforts to minimize such transients effects will have a great impact in reducing the conducted and radiated emissions which are so common in power supply busses.

Designers of PCB which are to accommodate SMPS converters must control the noise emanating from the on-board converter so that it does not interfere with other systems circuitry, or propagate into the main power bus. Converters are usually designed to pass CE and FCC radiated and conducted emissions. However, this is really not enough since the containment of emission at the PCB level must also be exercised. Conducted emissions containment is seldom provided within the latest high power modules. One reason is that it gives the designer greater flexibility in meeting design requirements. Second it reduces costs and real state requirements. Therefore, good board layout is essential for minimizing the amount of noise an on-board converters conducts or radiates. Good board layout is essential for maximizing power efficiency from on-board converter to other PCB loads. Ideally, the board should provide wide power paths routed closely together in parallel. In addition, all closed loop areas, which can behave as antennas should be kept to a minimum. To help shield other circuitry from the radiated noise in fast switching power train, board designers should avoid running signal lines under the converter. Common mode noise, which is coupled through the capacitance between components such as heat sinks and transformer isolation windings, appear between frame ground and the converter's input conductors. Differential noise appears across the input conductors.

Common mode noise showing high frequency content can be routed back to the on-board converter by ceramic capacitors placed between input and output conductor and the case ground. Lower frequency differential mode noise can be diminished using ceramic capacitors placed close to the converter between the input leads. All of these bypass capacitors should be placed as closed as possible to the converter to minimize loop areas. Figure 4 shows the proper placement of capacitors in dc-dc converters.

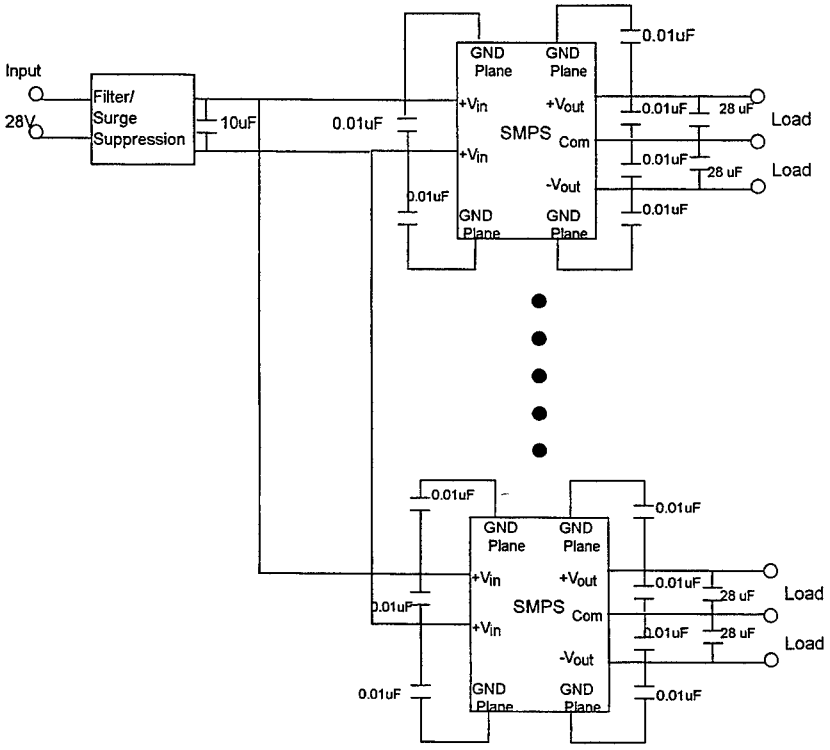


Figure 4. Proper placement of capacitors in dc-dc converters.

Once we had placed bypass capacitance in a smart configuration to reduce common mode noise from the dc-dc converters and PCB we take a look at the ground and power layouts from the CAD system to see what improvements can be made in the layout that would reduce even more the conducted and radiated emissions. In Figure 5.0 we see the layout of power and ground planes corresponding to the dc-dc converter and PCB schematics shown for Figure 4.0

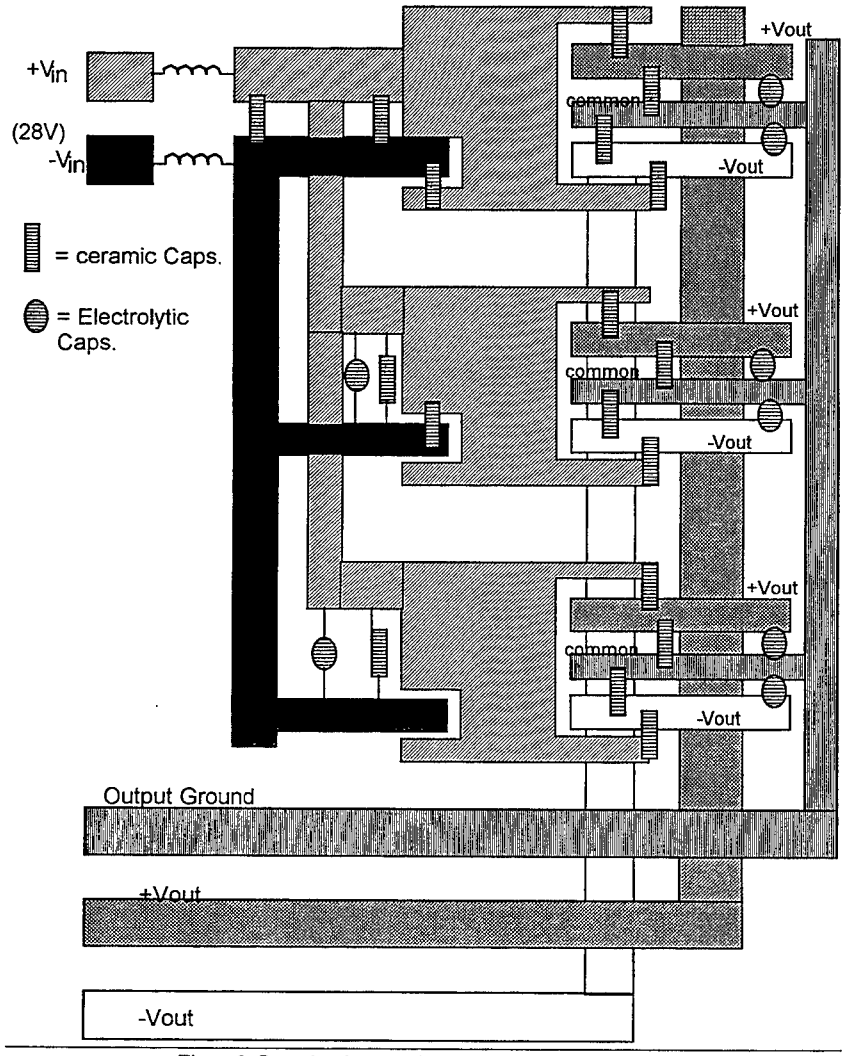


Figure 5. Ground and power planes layouts for dc-dc converter

In Figure 6 we observe some experimental data concerning conducted emissions which was obtained for a PCB based on the design of Figure 5.0 (other components of the PCB are not shown)

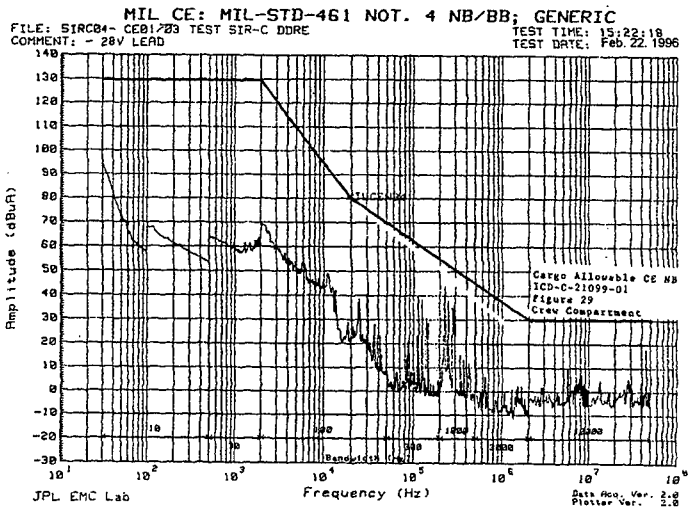
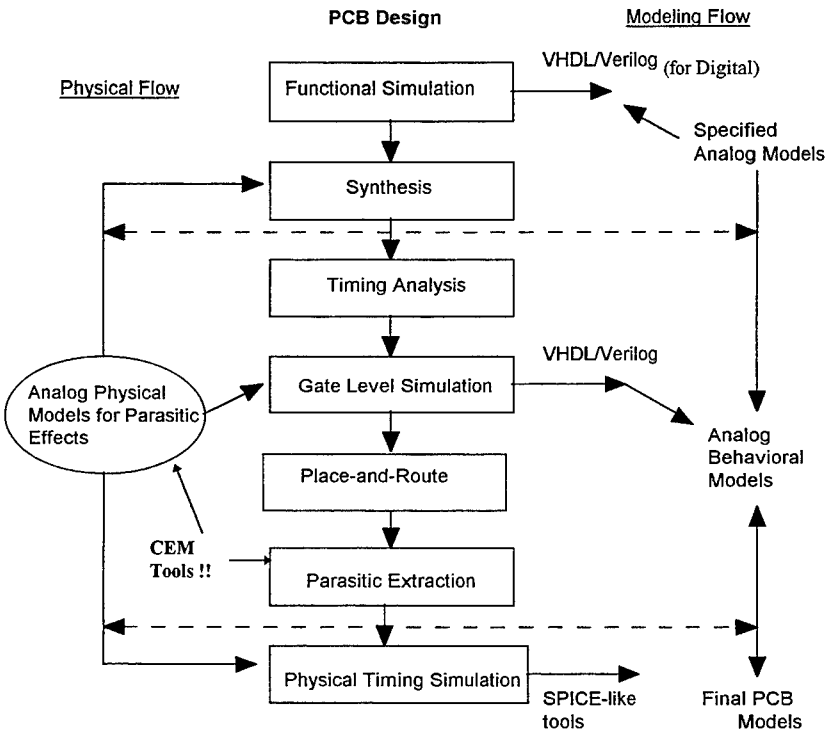


Figure 6.0 Conducted emissions for a three dc-dc converter PCB with ground and power layout for dc-dc converters as shown in figure 5.0

There are three basic rules commonly used by designers for containing the noise generated by the power module: a) return the noise current to the source using as short as possible a return loop, b) reduce the impedance of these loops by reducing inductance and increasing capacitance, and c) identify alternate routes and suppress them by adding impedance.

4.0 Using CEM Tools for Optimizing Power and Ground Layout for dc-dc converters in a PCB

The flow chart below shows a brief outline of the procedures followed for designing PCB using high level hardware description languages such as Verilog and VHDL. Notice that an integral part of this modeling process (e.g for designing a PCB with dc-dc converters) is the use of Computational Electromagnetic Tools (CEM) to perform a complex parasitic extraction process and obtain parasitic effects such as inductances, resistances and capacitances effects in the PCB. This procedure is used for obtaining power planes, ground planes and other ICs parasitics. As the flow chart shows final timing simulations are performed to estimate power & ground plane performance . The procedure can be repeated several times for each change of the power, ground planes and ICs layout until an optimum design is obtained



A new and more optimized layout of the one shown in Figure 5 is now shown in Figure 7 after the optimization procedure shown above.

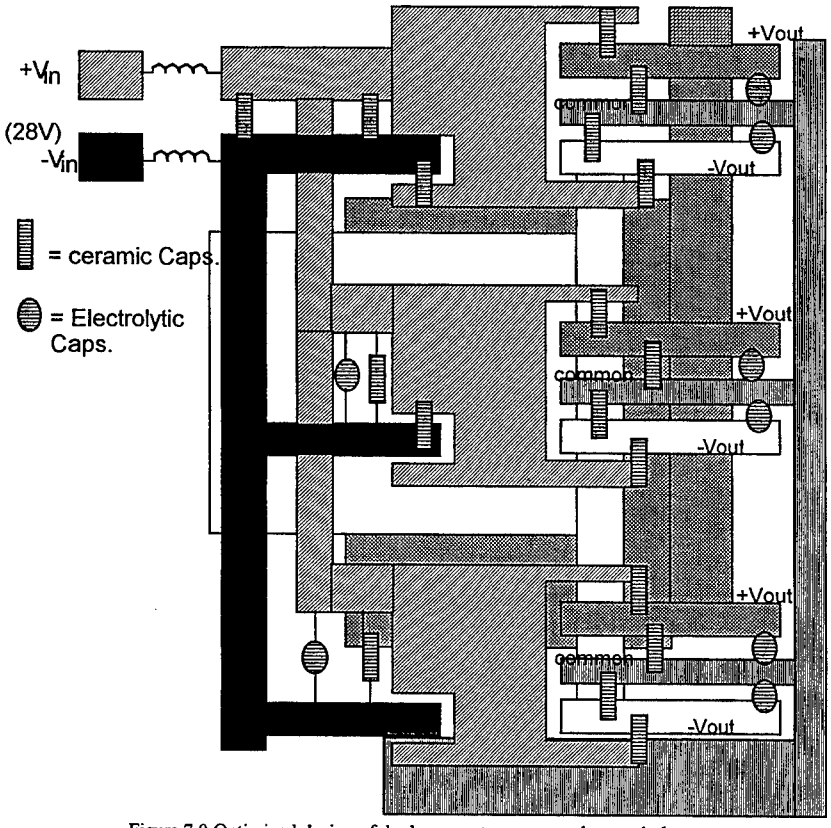


Figure 7.0 Optimized design of dc-dc converter power and ground planes.

A better ground and power planes design often translate into a less noisy board. Therefore, some conducted emission measurements are made again on this new design in Figure 7.0. The figure shows that some improvement in a couple of frequencies are made where the emissions were reduced to the spectrum analyzer's noise floor level. Further optimization can be obtain if the same techniques are applied at other parts of the PCB beyond the dc-dc converters.

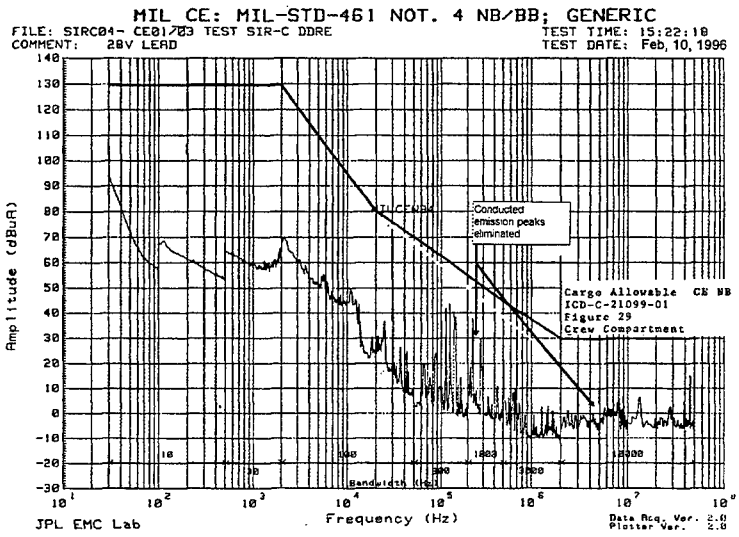


Figure 8.0. Conducted emissions for a three dc-dc converter PCB with ground and power layout for dc-dc converters as shown in figure 7.0

5.0 Conclusion

Transient effects and common mode noise current in SMPS are directly responsible for the conducted emissions often seen in the 30 kHz to 100 MHz region. This paper has shown some of the origins of these transient effects and modeling associated with such effects. It has also been shown that minimizing parasitics in power and ground planes of dc-dc converters will not only diminish possible transients effects but will diminish common mode noise as shown in conducted emissions.

References.

[1] Pressman, Abraham, L. "Switching Power Supply Design", McGraw Hill.
[2] Y.S.Lee and Y.C.Cheng, "Computer Aided Analysis of Electronic dc-dc Converters," IEEE Trans. on Aerospace ad Electronics Systems," Vol 24, No. 2, pp. 124-132, March 1988.
[3] E.T.Moore and T.G. Wilson, "Basic consideration for dc to dc conversion network," IEEE Trans. on Magnetics, Vol. 2, No. 3, September 1993.

Expert System Algorithms for EMC Analysis

T. Hubing, N. Kashyap[1], J. Drewniak, T. Van Doren, and R. DuBroff
University of Missouri-Rolla

Abstract – Expert system algorithms that analyze printed circuit board designs, anticipate EMC problems, and help designers to correct these problems are being developed by the EMI Expert System Consortium at the University of Missouri-Rolla. This paper reviews the basic structure of the EMI expert system and describes newly developed algorithms.

Introduction

In order to achieve the short development cycles that are necessary to be competitive in the electronics industry, it is becoming increasingly important to get the design correct before the first prototypes are built. This means that printed circuit board designs must be capable of meeting radiated EMI and EM susceptibility requirements the very first time they are tested in a lab. Experienced EMC engineers with a detailed knowledge of a printed circuit board design can often identify potential EMC problems in a design, evaluate the severity of these problems, and help designers to correct them before a prototype is built. Unfortunately, most companies cannot afford to have an experienced EMC engineer looking over the shoulder of the designers at every phase of the design process.

Expert system EMC software is designed to help provide EMC expertise to circuit designers and the people who do printed circuit board layouts. Expert system EMC software reads data from automated board layout files, component files and an EMC knowledge database. It then uses this information to find and evaluate potential EMC problems. Unlike numerical EM software or design rule checkers, expert system software is capable of identifying and quantifying critical EMC problems and helping the non-expert user to solve them.

The following sections describe the ongoing work of the EMI Expert System Consortium at the University of Missouri-Rolla. The consortium consists of hardware and software companies who are working with the university to develop expert system software for EMC analysis.

The EMC Expert System

Figure 1 shows the basic structure of the EMC expert system. The shaded boxes represent those algorithms that have been implemented. The expert system consists of four stages – the input stage, the evaluation stage, the estimation stage and the output stage. Each stage is made up of several modules, with each module performing a certain task. This modular structure makes it easy for a person to understand and modify the functional capability of the system.

[1] Navin Kashyap is currently a graduate student at the University of Michigan in Ann Arbor.

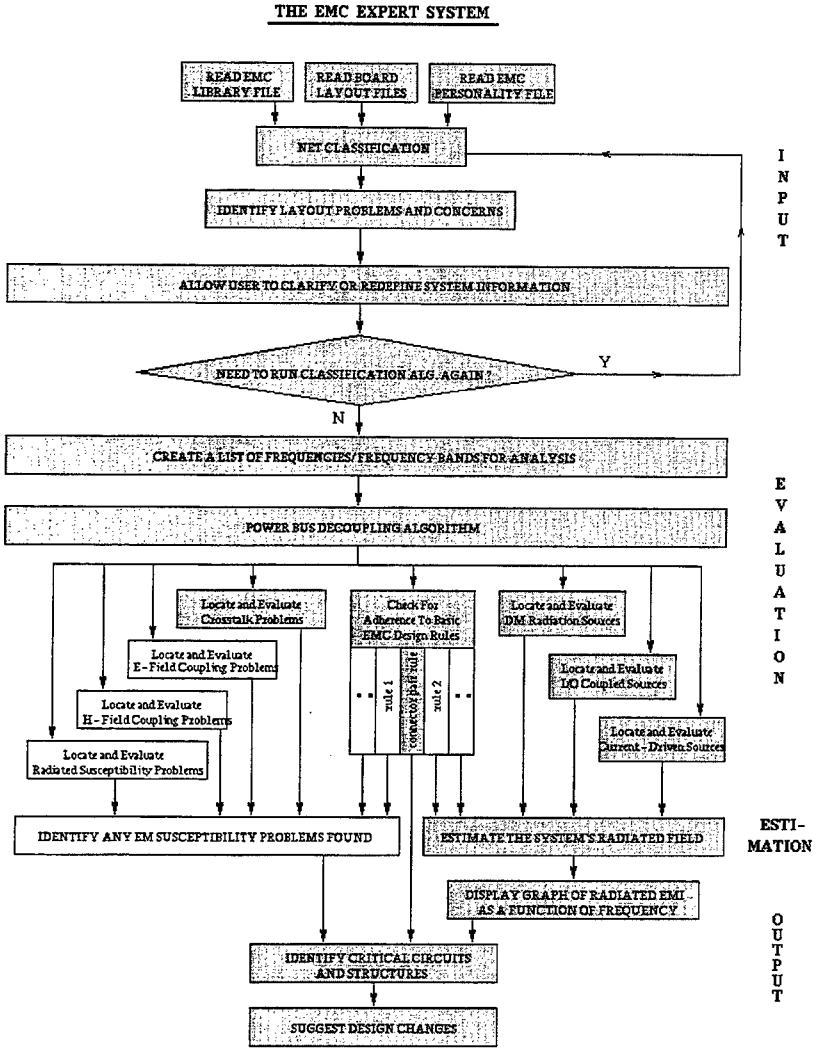


Figure 1: EMC Expert System Flow

The Input Stage

Information about the printed circuit board under analysis is collected by the input stage of the expert system. Physical information about the board, such as board geometry, names and locations of all nets and components, trace lengths and thicknesses etc., is obtained from board layout files generated by automated layout tools. The electrical properties of each net, such as signal frequencies, currents, voltages etc., are deduced by collating information from the layout files and the *component library*.

The component library is a file that contains information about components that is not present in the board layout files. It is a database of information about all components that the system may encounter when analyzing PCB's for a particular set of users. The component library contains component information at two levels – the package level and the pin level. Package level information about a component includes the component name, package size and type, pin pitch etc. Pin-level information about a component is provided for each pin of the component and varies depending on the type of component and the function of the pin. For example, each output pin of an active digital device would have an entry in the component library that specifies the risetime, maximum voltage, maximum current, clock frequency, and type of signal (e.g. data, clock, etc.).

A third source of information for the expert system algorithms is the *EMC personality file*. This file is used to tailor the expert system software to meet the needs of a particular company. The EMC personality file contains industry-specific information that controls the way the expert system algorithms execute. It also contains information that helps the expert system to recognize circuits and structures commonly used by a particular company.

The data from the layout files and the component library is used by the *net classification* algorithm to determine information about the signal properties, noise margin and function of each net on the board. It also searches for possible layout problems, such as nets being referenced to more than one power source, or nets being driven by more than one driver, and alerts the user to such problems. The algorithm identifies all power and ground nets on the board by checking each net to see if any of the pins attached to it are specified to be power or ground in the component library. Nets that are neither power nor ground are called *signal* nets.

The classification algorithm determines various signal parameters for each signal net. These parameters are determined from the component library entry for the driver for the net. The algorithm locates a driver by checking to see if any active device output pin is connected to the net either directly or through passive devices. The signal parameters determined by the classification algorithm consist of the clock frequency associated with each digital net, the range of signal frequencies on each analog net, the signal transition time for each digital net, the maximum and minimum voltages on each net, the maximum current on each net, the reference voltage for each net, and the utilization classification of each net.

Each signal net is also assigned a *noise margin*, which is the maximum voltage that may exist on the net without interfering with the normal behavior of the components. This assignment is based on the noise margins of the active device input pins on the net, as specified in the component library.

After the classification algorithm finishes its run, its results are made available to the user, who is given a chance to modify the results, or provide information that may fill in any gaps in the available information. At no point does the expert system ever require the user to provide information about the circuits or board design. If the user is satisfied with the results of the net classification, these results are passed to the evaluation stage of the EMC expert system.

The Evaluation Stage

The evaluation stage of the expert system contains the modules that perform a detailed EMC analysis of the board. These modules search for potential radiation and susceptibility problems with the board, and also test the board for compliance with basic EMC design guidelines

The expert system creates a list of all the clock frequencies on the board, and their harmonics, and all narrow-band analog signal frequencies. The narrow-band radiation from the board is calculated at these frequencies only. The frequency spectrum is also divided into blocks at which the broadband radiation is calculated. These blocks are created in such a way that each block is centered at a narrow-band frequency, and fills the space between narrow-band frequencies.

The *power bus noise* algorithm estimates the voltage induced on the power bus of printed circuit boards that utilize power and ground planes. This estimate is based on information about the currents drawn from the power bus by the active devices and the effective decoupling at each frequency of interest. A time-domain analysis is used to predict the peak voltage induced on the power bus and a frequency-domain approach is used to determine the noise on the power bus as a function of frequency. Power bus noise information is utilized by other algorithms and therefore the power bus noise algorithm must be run before the remaining algorithms in the evaluation stage.

The basic approach used by the expert system to locate and quantify radiated EMI problems is to locate all possible sources of high-frequency energy and all structures likely to radiate that energy. Different algorithms are used in the evaluation stage to locate different kinds of EMI sources.

The DM radiation source algorithm searches for signal nets that carry high-frequency currents and are long enough or large enough to serve as their own antenna. DM refers to *differential-mode* radiation sources. Differential-mode sources are rare on well-designed boards, but they are relatively easy to locate and quantify.

I/O coupled sources are fairly common, particularly on dense boards with many signal layers. An I/O-coupled source results when signal energy from one net couples to another net that carries this energy off the board. The expert system algorithms look for both magnetic and electric field coupling between nets with high-frequency signals and nets that attach to connector pins.

The most common radiated EMI problems below about 500 MHz are due to *current-driven* sources. Current driven sources result when signal return currents create a small potential difference between two points in the ground structure. This potential difference can create currents in cables or enclosures attached to ground that result in radiation. The expert system estimates the two-dimensional voltage variation across the return plane structure, due to currents returning on the power and ground planes. It then locates the antennas that may be driven by this voltage variation.

The expert system is capable of identifying antenna configurations such as a cable being driven relative to another cable or a heatsink, a cable or heatsink being driven relative to the board etc. For each such antenna, it determines the voltage difference between the two halves of the antenna, and then calculates the E-field radiated from the antenna at each narrow-band and broad-band frequency.

Algorithms are also included that identify crosstalk problems and check the design for violations of basic EMC design guidelines.

The Estimation and Output Stages

The results from all the modules in the evaluation stage are passed to the estimation stage, which combines these results to form an overall estimate of the radiated EMI from the board. The radiated EMI modules in the evaluation stage calculate the magnitudes of the electric fields due to each of the radiated EMI mechanisms, at each frequency and frequency block.

The output stage presents the expert system's evaluation of the board to the user. It displays a graph of the estimated radiated EMI as a function of frequency, and identifies the circuits and structures on the board that are mainly responsible for the board's radiated EMI problems. It also suggests design changes that will alleviate the problems reported.

The radiated EMI plot displayed by the expert system is similar to that which would be obtained from an actual EMI test. It plots the board's radiated field in dB(μV/m) as a function of frequency. An FCC or CISPR limit line is placed on the plot, so as to give the user an immediate idea of the frequencies at which the board radiation exceeds the limit, and the amount (in dB) of excess radiation at those frequencies.

Significant contributions of individual nets to the radiated E-field are recorded at each frequency by the modules of the evaluation stage. These are used to construct a list of nets causing the worst problems at any particular frequency. So, if the user would like to know which nets are causing the radiation to exceed the limit at any frequency, the expert system can list all such nets and display a diagram of the board layout that highlights these nets. Information about the mechanisms that cause these violations is also available to the user.

The expert system also offers suggestions that will help in reducing radiated EMI levels. As the chief contributors to the emissions are known to the system, it uses simple rules to come up with viable suggestions that will reduce the contributions from the worst offenders.

New Algorithms

The next prototype software will contain improvements to the existing algorithms based on evaluation of these algorithms against actual hardware. Improvements to the current-driven algorithm will reduce the probability that this algorithm will be fooled by an unusual component placement. Also voltages induced in the power and ground planes by the components themselves will be estimated in addition to the voltages induced by the currents through the traces.

Experiments using real computer hardware in the laboratory have shown that radiation at frequencies near 1 GHz is dominated by different source mechanisms than radiation below 500 MHz. At the higher frequencies, enclosure resonances play a critical role in the way that products radiate. A new algorithm has been developed to predict and analyze radiated emissions at frequencies above 500 MHz in products with metal enclosures.

The next prototype software will also use a different method to sum the contributions from the various EMI sources that are identified. The original version used a root-mean-square sum of the field strengths resulting from each individual source-antenna combination. However, the algorithms assume that the cables are oriented in the position that "tends to maximize" radiated emissions (per the FCC and CISPR test procedures). Since it is not usually possible to find a cable position that maximizes the contributions from all sources at the same time, this root-mean-square summing technique has been shown to be too harsh. The new algorithms will sum all of the sources to determine their relative contribution, but the level reported to the user will be the predicted emissions from the worst-case source-antenna pair at each frequency.

Summary

The EMC expert system described in this paper models the thinking process of a human EMC expert. It reads board layout information and information about the components on the board. It uses information stored in its knowledge base (i.e. the component library and the personality file) to deduce properties of the signals on each board trace. This information is used to identify and evaluate possible radiation sources and antennas, and provide an overall estimate of board radiation and board susceptibility.

The EMC expert system is not designed to replace human EMC experts. However, it provides a means of automating many of the tasks that human EMC experts normally perform. Also, it is capable of analyzing a design before a prototype has been built. And since the expert system does not require the user to be an expert, this analysis can be done at any point in the design process by circuit designers, board layout personnel, or anyone with access to the board layout files.

Finally, the EMC expert system is not a replacement for numerical electromagnetic modeling software. It does not do a thorough analysis of EMI sources with well-defined parameters. However, it excels at the one thing that numerical electromagnetic modeling software does not do well: locating and prioritizing potential EMC problems. Ideally, future printed circuit board designers will have a suite of tools at their disposal. They will use expert system tools to identify EMC sources, antennas and coupling paths; and numerical electromagnetic modeling tools to analyze these structures and evaluate alternatives.

The Electromagnetic Compatibility Characteristics of Buildings in Mobile Radio Waves Propagation Channel

Yasumitsu MIYAZAKI and Paul SELORMEY

Department of Information and Computer Sciences,
Faculty of Engineering, Toyohashi University of Technology
Hibarigaoka 1-1, Tenpaku-cho, Toyohashi, Aichi 441, Japan.
Tel: 0532-47-0111, FAX: 0532-47-0152
E-mail: miyazaki@emlab.tutics.tut.ac.jp

Abstract

Detailed characterization of radio propagation channel is a major requirement for successful design of mobile communication systems. In this paper, mobile radio channel characterization process based on the FDTD method is presented. The merits and demerits of the currently used methods, namely impulse-response method and ray-tracing methods are briefly considered, and the total field formulation of the FDTD method is discussed. The simulation model consists of a main street with six concrete buildings. The wave propagation patterns in the whole channel and the received signals at some line of sight and out of sight locations are presented.

Keywords : FDTD method, propagation characteristics, wave scattering, mobile radio waves, urban area

1 Introduction

A typical mobile radio environment consists of two parts; propagation loss and multipath fading, as shown in Fig. 1. This research seeks to model the multipath fading due to scattering by buildings and other outdoor structures. Multipath fading, which results from reflection, refraction and scattering of radio waves by buildings and other structures, gives rise to more than one path reaching the receiver and produces a distorted version of the transmitted signal. The multipath fading in mobile and indoor communication systems cannot be eliminated, therefore multipath channel must be well characterized in order to reduce its effect in the design of such systems [1]. Most reported mobile channel modeling process, as in [2, 3], are based on measurements which are expensive and time consuming. Until recently, the time-varying indoor and mobile radio propagation channels are usually modeled as: the channel, for each point in the 3-dimensional space, is a linear filter [4] having the impulse response:

$$h(t,\tau) = \sum_{k=0}^{N(\tau)-1} a_k(t) g[\tau - \tau_k(t)] e^{j\theta_k(t)} \qquad (1)$$

where t, τ are the observation time and time of impulse application respectively, N(t) is the number of multipath components, g(t) is a basic pulse shape, and $\{a_k(t)\}$, $\{\tau_k(t)\}$, $\{\theta_k(t)\}$ are the random time-varying amplitude, arrival-time and phase sequences respectively. This model is illustrated in Fig. 1(b). A time-invariant form, suggested by Turin [5] for the multipath channel, has been applied successfully to some mobile radio applications [6, 7]. In this case (1) reduces to:

$$h(t) = \sum_{k=0}^{N-1} a_k g[\tau - \tau_k] e^{j\theta_k} \qquad (2)$$

The output of the channel, y(t), to a transmitted signal, s(t), is given by the equation

$$y(t) = \int_{-\infty}^{\infty} s(\tau) h(t-\tau)\, d\tau + n(t) \qquad (3)$$

where n(t) is Gaussian noise. The following are some limitations of the impulse response method: the detailed structure of the scatterer is not modeled, no single statistical distribution exists to model all situations of arrival time sequences and path amplitude distributions, it is not practically intuitive, the characteristics properties of the buildings and other scatterers are not completely modeled. These limitations form the basis of our choice of the FDTD method, which has been successfully applied to many electromagnetic problems including [8], for the multipath fading channel modeling and simulations. Alternatively, the modeling based on the 3-dimensional Uniform Theory of Diffraction (UTD), used in [9, 10], is also receiving much attention. The UTD method is known to be very accurate at high operating frequencies, and requires less computer resources (memory) when compared with the FDTD method. The UTD calculation times T_r grow as

$$T_r \alpha n_{RX} \cdot n_{ob}^{n_{re}} \tag{4}$$

where n_{RX} is the number of points at which fields is to be calculated, n_{ob} is number of obstacles and n_{re} is the number of reflections.

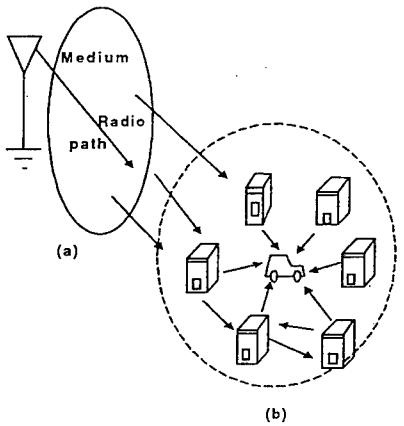


Fig. 1: A Mobile Radio environment (a) Propagation loss, (b) Multipath fading.

However, the following are some of the limitations of the UTD method: in a highly reflective environment it is very difficult to compute, it is reliable when the scatterers are many with complex geometry, it is not practical if field strength at many different locations are required, to determine the most critical receiver location for instance. Other methods used in [11, 12] are not only difficult to implement, the results are not very consistent[12].

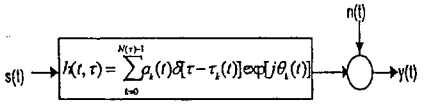


Fig. 2: Impulse response model of the multipath fading.

2 The FD-TD Algorithm for Radio Propagation

In the isotropic medium, Maxwell's equations on which Yee's FDTD algorithm is based, are given by

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} \tag{5}$$

$$\nabla \times \mathbf{H} = \sigma \mathbf{E} + \varepsilon \frac{\partial \mathbf{E}}{\partial t} \tag{6}$$

where μ, σ, ε are the magnetic permeability, electric conductivity and permittivity respectively. For the simulation, the total field formulation is used. In this case the total fields for 2-dimensional TM mode of $\mathbf{E} = E_z i_z$ are expressed as

$$\frac{\partial H_x}{\partial t} = -\frac{1}{\mu}\frac{\partial E_z}{\partial y} \tag{7}$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu}\frac{\partial E_z}{\partial x} \tag{8}$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - J_z\right) \tag{9}$$

A grid point in Yee's notation is define by the relation $(i,j) = (i\Delta x, j\Delta y)$ in 2- dimensions and any function of space and time is expressed as $F^n(i,j) = F(i\Delta x, j\Delta y, n\Delta t)$. By centered finite-difference, a space derivative can be expressed as

$$\frac{\partial F^n(i,j)}{\partial x} \approx \frac{F^n(i+1/2,j) - F^n(i-1/2,j)}{\Delta x} \tag{10}$$

The time derivative is also expressed as

$$\frac{\partial F^n(i,j)}{\partial t} \approx \frac{F^{n+1/2}(i,j) - F^{n-1/2}(i,j)}{\Delta t} \quad (11)$$

Using these notations the FDTD difference relations for the above equations are given by:

For free space,

$$E_z^{n+1}(i,j) = E_z^n(i,j) + \frac{\Delta t}{\varepsilon_0 \Delta x}[H_y^{n+1/2}(i,j) - H_y^{n+1/2}(i-1,j)] - \frac{\Delta t}{\varepsilon_0 \Delta y}[H_x^{n+1/2}(i,j) - H_x^{n+1/2}(i,j-1)] - \frac{\Delta t}{\varepsilon_0} J_z^{n+1/2} \quad (12)$$

On a perfect conductor, $E_z^n(i,j) = -E_z^n(i,j)$.

In dielectric material (buildings),

$$E_z^{n+1}(i,j) = E_z^n(i,j) + \frac{\Delta t}{\varepsilon \Delta x}[H_y^{n+1/2}(i,j) - H_y^{n+1/2}(i-1,j)] - \frac{\Delta t}{\varepsilon \Delta y}[H_x^{n+1/2}(i,j) - H_x^{n+1/2}(i,j-1)] - \frac{\Delta t}{\varepsilon} J_z^{n+1/2} \quad (13)$$

where Δx, Δy, and Δt are the increments in x, y and time, respectively.

The magnetic fields are given by the relations

$$H_y^{n+1/2}(i,j) = H_y^{n-1/2}(i,j) + \frac{\Delta t}{\mu_0 \Delta x}[E_z^n(i+1,j) - E_z^n(i,j)] \quad (14)$$

$$H_x^{n+1/2}(i,j) = H_x^{n-1/2}(i,j) + \frac{\Delta t}{\mu_0 \Delta y}[E_z^n(i,j+1) - E_z^n(i,j)] \quad (15)$$

The orientation of the electric and the magnetic fields in a cell is as shown in Fig. 3.

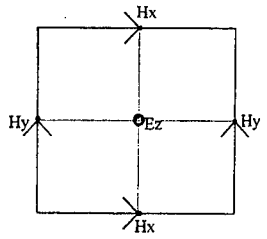


Fig. 3: The arrangement of the fields in a Yee cell

For accuracy, cells size $\delta = \min(\Delta x, \Delta y)$ must be smaller than $\lambda/10$, where λ is the smallest wavelength in problem space. For stability, the time increment Δt must satisfy the Courant inequality

$$c\Delta t \leq \frac{1}{\sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2}}} \quad (16)$$

where c is the velocity of light. For this simulation, 80% of the time given by the Courant equality is used for the time step, Δt.

The Mur's Absorption Boundary Condition (ABC) is used to limit the simulation region. The first order ABC is applied at the corners of the problem space. For example, at point x=0:

$$E_{0,j}^{n+1} = E_{1,j}^{n-1} + \frac{c\Delta t - \Delta x}{c\Delta t + \Delta x}[E_{1,j}^{n+1} - E_{0,j}^{n-1}] \quad (17)$$

At all other boundary points the second order ABC is applied. For example, along the line x=0:

$$E_{0,j}^{n+1} = -E_{1,j}^{n-1} + c_6[E_{1,j}^{n+1} + E_{0,j}^{n-1}] + c_7[E_{0,j}^n + E_{1,j}^n] + c_8[E_{0,j+1}^n - 2E_{0,j}^n + E_{0,j-1}^n + E_{1,j+1}^n - 2E_{1,j}^n + E_{1,j-1}^n] \quad (18)$$

where

$$c_1 = \frac{c\Delta t - \Delta x}{c\Delta t + \Delta x} \quad (19)$$

$$c_2 = \frac{2\Delta x}{c\Delta t + \Delta x} \quad (20)$$

$$c_3 = \frac{(c\Delta t)^2 \Delta x}{2(\Delta y)^2 (c\Delta t + \Delta x)} \quad (21)$$

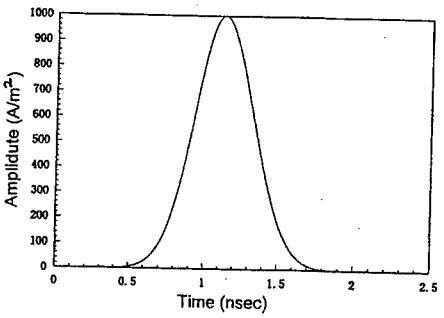


Fig. 4(a): The Gaussian line current.

For the simulation of electromagnetic propagation, we assumed that an antenna of point type at Tx in street, vertical polarization and generates a Gaussian line current of the following form,

$$J_z^{i,n} = J_{\max}^{i,n} \exp\{-\alpha(t - \zeta\Delta t)^2\}$$
$$\alpha = \left(\frac{4}{\zeta \Delta t}\right)^2 \quad (22)$$

where t is the time elapsed, $J_{max}^{i,n}$ is the amplitude, ζ is the number of time steps in Gaussian pulse from the peak value to the truncation value, and α is a constant related to ζ as given above.

The plot of Gaussian line current and the corresponding Fourier transform are shown in Fig. 4(a) and 4(b), respectively.

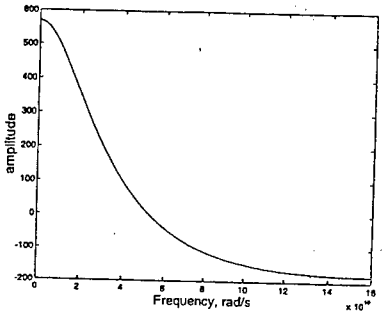


Fig. 4(b): Fourier Transform of the Gaussian line current.

3 The Simulation Model and Parameters of Radio Propagation

The problem space, as shown in Fig. 3, is 30×30 meters (or 2000×2000 in cells units). The model consists of six buildings as the scatterers of the wave, aligned symmetrically for simplicity and each is of dimensions 9×6 meters. The buildings are lineup along a main street, 9 m wide. The buildings are separated from each other by streets 4.5 meters wide, and in all the separation between the buildings and the problem space's boundaries are maintained at 1.5 meters, equivalent of 100 cells. Currently, though not very practical, the buildings are considered to be a solid of homogenous density. For the building walls, the relative permittivity is 3 and the conductivity is 0.005 mho/m [10]. The summary of the simulation parameters is given in Table I. Results are also presented for the simple case, where the buildings are considered as perfect electric conductors (PEC), for clearer propagation patterns and comparison. We used $\zeta = 32$, since the concrete buildings have relative permittivity of 3. In the simulated region, there are three line of sight (LOS) sites of receiver locations Li with coordinates as follows; L1 (4.5m, 15.0m), L2 (15.0m, 15.0m), L3 (25.5m, 15.0m), and out of sight (OOS) sites of receiver locations Li with coordinates; L4 (9.75m, 6.0m), L5 (9.75m, 24.0m), L6 (20.25m, 6.0m) and L7 (20.25m, 24.0m). The transmitter Tx is located at point (1.50m, 15.0m).

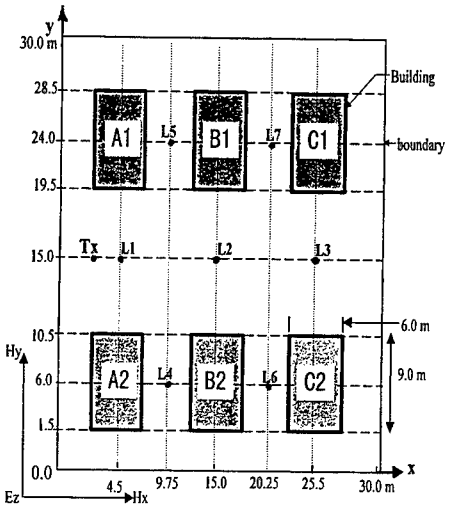


Fig. 5: The layout of problem space for simulation

Table I: Simulation Parameters

Frequency of source	=850 MHz
Cell Size, δ	=0.015 m
Time increment, Δt	=28.32 ps
Relative Permittivity(building)	=3.0
Conductivity of building, σ	=0.005 S/m
Current amplitude, $J_{max}^{i,n}$	=1000.0 A/m^2
Pulse duration	=1.81 ns

4 Numerical Results and Discussions

The total received signals at the locations L1, L2, L3, L4 and L7 for the case where the buildings are considered as having a dielectric permittivity of 3 and conductivity of 0.005 mho/m are shown in Figs. 6 -

10. Similar plots for the case where the buildings are considered to be perfect conductors are shown in Figs. 11-14. In this latter case, for the location L4, the sharp initial power fall at about 60 ns can be attributed to the shadowing effect, which is expected to be more effective in this case where the buildings are considered to be perfect conductors. In both cases, the received signal at location L4 shows much variations with time, since it is an out of sight location. For each plot, the E-values are taken starting from time t=0, therefore each plot shows an initial fast fading effects after which approximately regular patterns develop. The signal propagation patterns in the problem space are shown in Fig. 16-18 for the concrete buildings . The Fig. 16 represents the total electric field after 28.3 ns of propagation time corresponding to 1000 time steps. After the another 1000 time steps, that is a total time of 56.6 ns, the pattern in Fig. 17 is obtained and the after the next 1000 steps, total time of 84.9 ns, the propagation pattern in Fig. 18 is obtained. The propagation patterns are as expected from the scattering geometry shown. The buildings are observed to reflect back much of the transmitted signal, a lost to a receiver within the buildings. Similar, plots are shown for case where the buildings are considered as perfect electric conductor. In general, the electric field patterns show high peaks near the building corners along the line of sight. These are mainly due to diffractions and to some extent reflections at the corner points, which in these cases increase the received signal intensity. In all, the main street acts as a waveguide.

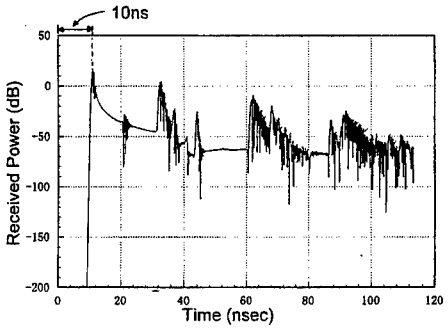


Fig. 6: Received electric signal at L1 (LOS).

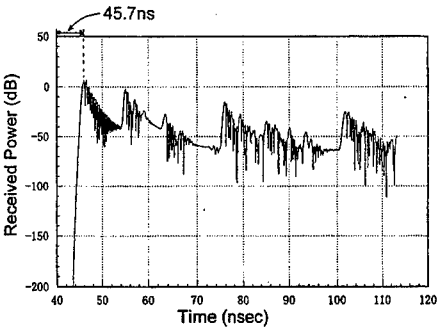


Fig. 7: Received electric signal at L2 (LOS).

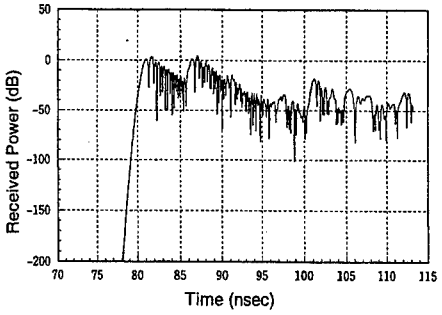


Fig. 8: Received electric signal at L3 (LOS).

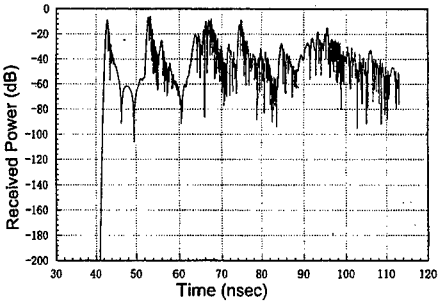


Fig. 9: Received electric signal at L4 (OOS).

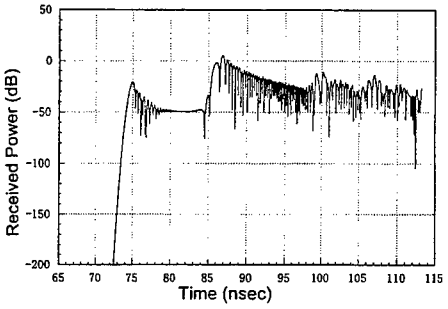


Fig. 10: Received electric signal at L7 (OOS).

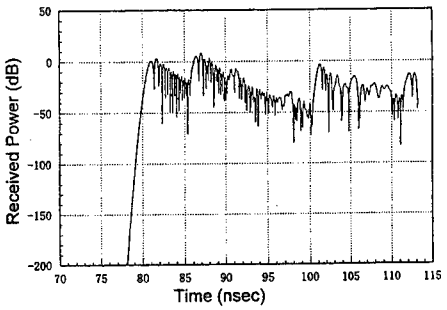


Fig. 13: Received electric signal at L3 (LOS) for PEC.

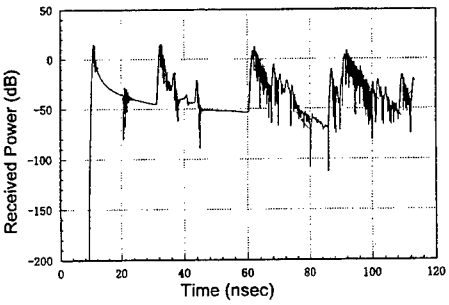


Fig. 11: Received electric signal at L1 (LOS) for PEC.

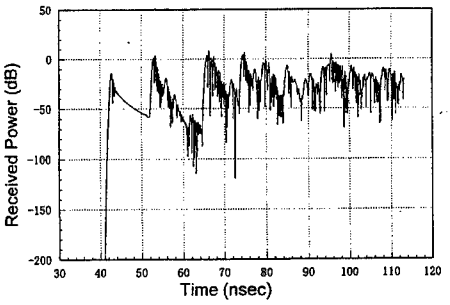


Fig. 14: Received electric signal at L4 (OOS) for PEC.

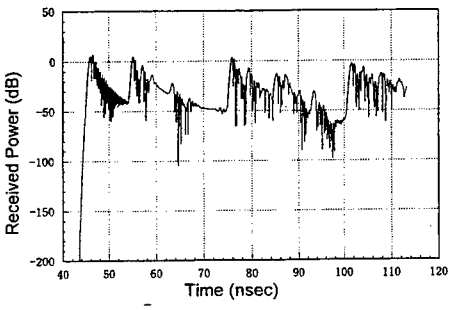


Fig. 12: Received electric signal at L2 (LOS) for PEC.

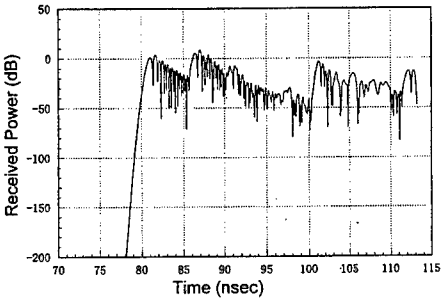


Fig. 15: Received electric signal at L7 (OOS) for PEC.

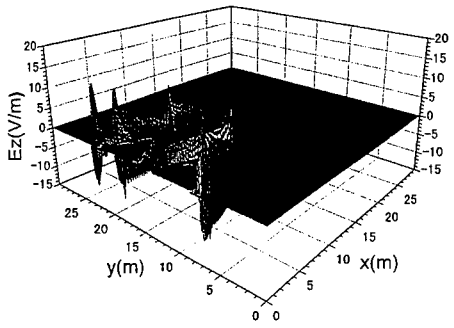


Fig. 16: Electric field patterns after 28.3 ns.

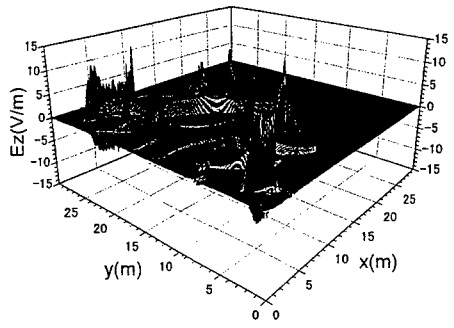


Fig. 17: Electric field patterns after 56.6 ns.

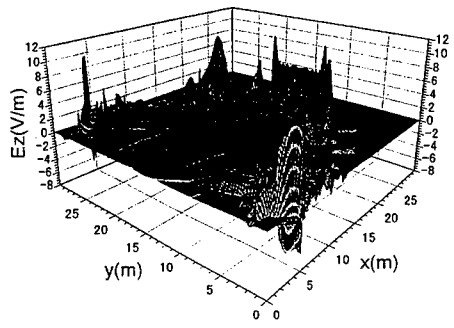


Fig. 18: Electric field patterns after 84.9 ns.

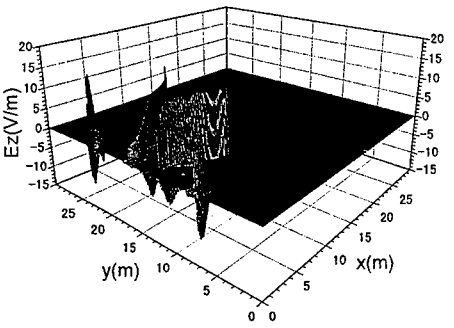


Fig. 19: Electric field patterns after 28.3 ns for PEC.

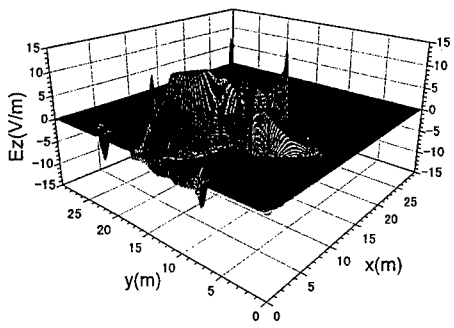


Fig. 20: Electric field patterns after 56.6 ns for PEC.

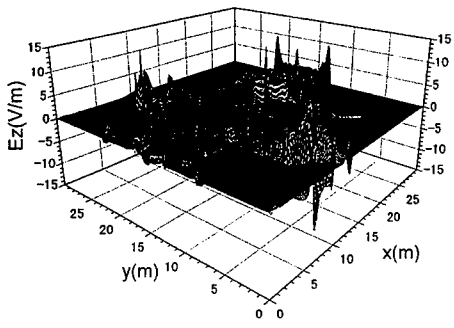


Fig. 21: Electric field patterns after 84.9 ns for PEC.

5 Conclusions and Future Plans

With the FDTD method, it is becoming increasingly possible to simulate the outdoor radio wave propagation. The results have much more intuitive meaning than the impulse response method that is currently being used. The main limitation is the two-dimensional approach as a result of the computer resources limitations. Therefore ground reflections, which are observed in practical situations, cannot be accounted for in this simulation. When used together with the UTD method, very complex mobile communication environments can be completely modeled. As a future plan, these results will be further optimized, and statistical properties of amplitude variation, path loss, mean excess delay, rms delay spread will also be determined. Finally, the results will be compared with similar models using the UTD methods in both 2-dimensions and 3-dimensions as used in [10, 14].

References

[1] H. Hashemi : "The Indoor Radio Propagation Channel", Proc. of the IEEE, vol. 81, p. 943, 1993.

[2] T. Aulin: "Characteristics of a Digital Radio Channel", IEEE Trans. on Veh. Tech., vol. 30, p. 45, 1997.

[3] P.J. Cullen, P.C. Fannin, A. Molina : "Wide-Band Measurement and Analysis Techniques for the Mobile Radio Channel", IEEE Trans. on Veh. Tech., vol. 42, p.589, 1993.

[4] Y. Miyazaki : "Optical waveguide type functional devices for signal processing in high speed mobile radio communication", 1996 Asia-Pacific MC Proc., vol. IV. p. 1333, 1996.

[5] G.L. Turin: "Communication through noisy, random-multipath channels", 1956 IRE Convention Record, part 4, p. 154.

[6] G.L. Turin : "A Statistical model for urban multipath propagation", IEEE Trans. on Veh. Tech., vol. 21, p. 1, 1972.

[7] H. Suzuki : "A Statistical model for urban radio propagation", IEEE Trans. on Comm., vol. COM-25, p. 673, 1977.

[8] Y. Miyazaki, K. Takahashi : "Wavelet Analysis of randomly reflected electromagnetic waves of subsurface radar", Proc. of ISAP, vol. 4, p. 1153, 1996.

[9] V. Erceg, A.J. Rustako, R.S. Roman : "Difraction Around Corners and its Effects on the Microcell Coverage Area in Urban and Suburban Environments at 900 MHz, 2 GHz and 6 GHz", IEEE Trans. on Vech. Tech., vol. 43, p. 762, 1994.

[10] A.G. Kanatas, I.D. Kountouris, G.B. Kostaras, P. Constantinou: "A UTD Propagation model in Urban Microcellular Environments", IEEE Trans. on Vech. Tech., vol. 45, p. 185, 1997.

[11] L.M. Hawthorne : " 800 MHz Mobile Radio Propagation Prediction Using Kalman Filtering Techniques ", IEEE Trans. on Vech. Tech., vol. 38, p. 55, 1989.

[12] H.S. Wang, N. Moayeri : "Finite-State Markov Channel - A Useful model for Radio Communication Channels", IEEE Trans. on Vech. Tech., vol. 44, p. 163, 1995.

[13] P. Selormey, Y. Miyazaki : "FDTD simulation of mobile radio waves propagation in streets", Proc. of IEICE General Conf., p. 1, Mar. 1997.

[14] S. Y. Tan, H. S. Tan : "A Theory for Propagation Path-loss Characteristics in a City-Street Grid", IEEE Trans. on EMC., vol. 37, p. 333, 1995.

SESSION 16:

HYBRID TECHNIQUES

Chairs: U. Jakobus and W.D. Wood, Jr.

Extension of the MoM/PO Hybrid Technique to Homogeneous Dielectric Bodies

U. Jakobus

Institut für Hochfrequenztechnik, University of Stuttgart,
Pfaffenwaldring 47, D-70550 Stuttgart, Germany

Abstract

A current–based technique hybridizing the method of moments (MoM) with physical optics (PO) is extended in its range of application from perfectly conducting bodies to scattering problems composed of metallic and homogeneous dielectric bodies. To treat electrically large dielectrics, the electric and magnetic surface current densities resulting from an application of the equivalence principle are approximated by PO, thus avoiding the need of solving a large system of linear equations. In an example the exact solution of a short dipole antenna radiating in front of a dielectric sphere is compared to the numerical results. MoM results are almost identical to the exact values, while PO leads to a drastic reduction of memory and CPU–time with results still accurate enough for most applications.

1 Introduction

Even though volume discretization techniques such as FDTD or FEM have gained much popularity these days due to the increased computer power available and the general range of applicability, the MoM is able to produce results with the same or an even higher degree of accuracy consuming considerably less memory and CPU–time for a certain class of scattering and radiation problems involving e.g. perfectly conducting metallic surfaces and wires or homogeneous dielectric bodies.

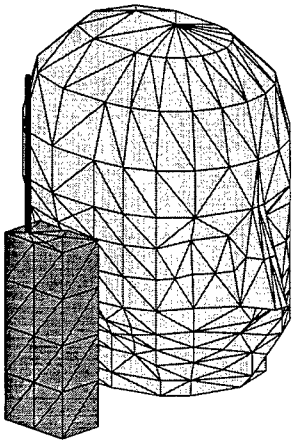


Fig. 1: Radiation of a mobile telephone close to the human head.

One common problem of all the techniques mentioned above is the strong dependency of memory and CPU–time on the frequency f [1], resulting from the need of discretizing the geometrical structure into volume or surface elements small in size as compared to the wavelength λ. The application of pure asymptotic high frequency techniques such as PO or diffraction theory (UTD) is often restricted to specific geometries.

Therefore, we will in the following concentrate on a hybrid technique combining the MoM for the resonance region and below with PO for electrically large regions. For perfectly conducting geometries, this technique has already been presented by the author, e.g. [2, 3].

The present paper extends this formulation by addressing the solution of problems where a metallic structure radiates in presence of a homogeneous dielectric body. A practical example is depicted in Fig. 1, where a mobile telephone is located in front of the human head. If, and this is our intention, we focus on optimizing the antenna by comparing different antenna concepts with respect to radiation pattern, gain, input impedance versus frequency or antenna efficiency, then a homogeneous head model with average tissue parameters is sufficient. This model is also able to closely predict the total absorbed power. Only for studies where detailed SAR (*specific absorption rate*) images are required, an inhomogeneous head model must be used e.g. in connection with the FDTD method.

In section 2 the theoretical background of the hybrid method is presented. Section 3 gives a brief review of the PO for dielectric bodies, while section 4 concentrates on some aspects of the hybridization. An example with results is considered in section 5.

2 Theoretical background of the hybrid method

Details of treating metallic problems by the MoM/PO hybrid method can be found elsewhere (e.g. [2]), therefore we assume in the following that all metallic parts are assigned to the MoM–region. According to the example depicted in Fig. 1, metallic as well as dielectric surfaces are subdivided into triangular patches. On metallic surfaces, basis functions $\vec{f}_n$ according to [4] together with unknown coefficients $\alpha_{J,n}$ are used in the superposition of the surface current density $\vec{J}$.

For determining the matrix elements or the near– and far–fields, the radiation of such a basis function $\vec{f}_n$ in the presence of the dielectric body is required as indicated in Fig. 2.

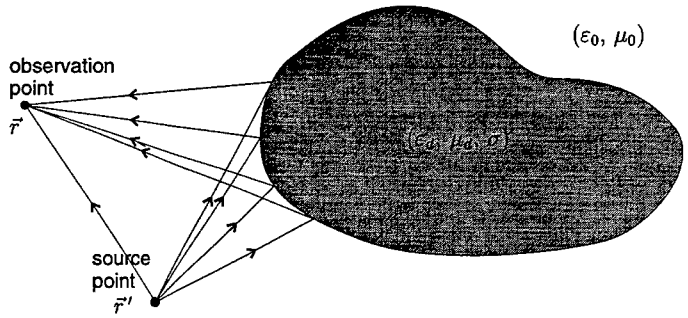


Fig. 2: Radiation of a basis function located at the source point $\vec{r}'$ in presence of the homogeneous dielectric body with material properties $(\varepsilon_d, \mu_d, \sigma)$.

The surface equivalence principle (e.g. [5]) is applied and equivalent electric $\vec{J}$ and magnetic $\vec{M}$ surface current densities radiating in free space are introduced according to Fig. 3:

$$\vec{J} = \hat{n} \times \vec{H}(S^+) \tag{1a}$$

$$\vec{M} = -\hat{n} \times \vec{E}(S^+). \tag{1b}$$

These currents are unknown. Within the MoM formulation, an integral equation is constructed and the currents are obtained from the solution of a system of linear equations. To avoid this time and memory consuming process, we will investigate the PO to determine $\vec{J}$ and $\vec{M}$ in the next section.

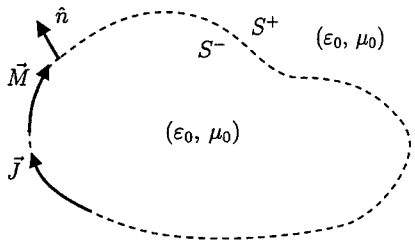


Fig. 3: Equivalent electric and magnetic surface currents $\vec{J}$ and $\vec{M}$ radiating in a homogeneous medium (ε_0, μ_0).

3 PO for dielectric bodies

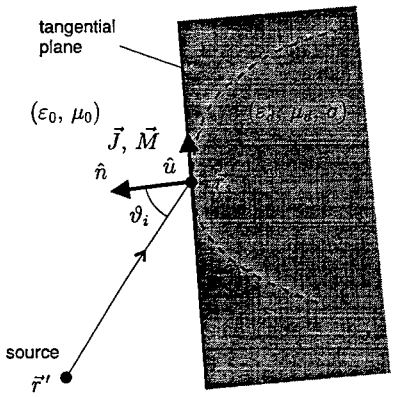


Fig. 4: Approximation of the curved surface by an infinite plane in order to determine the PO currents.

Similar to the PO approximation for metallic bodies, which is exact for an infinite plane, we locally approximate the dielectric surface at a point $\vec{r}_s$ by the tangential plane perpendicular to the normal vector $\hat{n}$. Introducing the two reflection coefficients

$$\Gamma_\perp = \frac{\xi \cos\vartheta_i - \sqrt{1 - \xi^2 (\frac{\mu_0}{\mu_d})^2 \sin^2\vartheta_i}}{\xi \cos\vartheta_i + \sqrt{1 - \xi^2 (\frac{\mu_0}{\mu_d})^2 \sin^2\vartheta_i}} \tag{2a}$$

$$\Gamma_\parallel = \frac{\xi \sqrt{1 - \xi^2 (\frac{\mu_0}{\mu_d})^2 \sin^2\vartheta_i} - \cos\vartheta_i}{\xi \sqrt{1 - \xi^2 (\frac{\mu_0}{\mu_d})^2 \sin^2\vartheta_i} + \cos\vartheta_i} \tag{2b}$$

with the ratio

$$\xi = \frac{Z_{Fd}}{Z_{F0}} \tag{3}$$

of the two wave impedances

$$Z_{Fd} = \sqrt{\frac{\mu_d}{\varepsilon_d - j\frac{\sigma}{\omega}}} \quad \text{and} \quad Z_{F0} = \sqrt{\frac{\mu_0}{\varepsilon_0}}, \tag{4}$$

the equivalent currents can be found exactly for an incident plane wave with incidence angle ϑ_i (see Fig. 4):

$$\vec{J}^{PO} = \hat{n} \times \left[\vec{\bar{I}} - \Gamma_\perp (\vec{\bar{I}} - \hat{u}\hat{u}) - \Gamma_\parallel \hat{u}\hat{u} \right] \cdot \vec{H}_{inc}(\vec{r}_s) \tag{5a}$$

$$\vec{M}^{PO} = -\hat{n} \times \left[\vec{\bar{I}} + \Gamma_\parallel (\vec{\bar{I}} - \hat{u}\hat{u}) + \Gamma_\perp \hat{u}\hat{u} \right] \cdot \vec{E}_{inc}(\vec{r}_s). \tag{5b}$$

Here $\hat{u}$ represents a unit vector perpendicular to $\hat{n}$ and $\vec{r}_s - \vec{r}'$. An alternative formulation to eqn. (5) based on an application of the equivalence principle and the local constraint of an impedance boundary condition is derived in [6]. There the result, which is exact only for perpendicular incidence $\vartheta_i = 0$, is:

$$\vec{J}^{PO} = \frac{2}{1+\xi}\,\hat{n} \times \vec{H}_{inc}(\vec{r}_s) \tag{6a}$$

$$\vec{M}^{PO} = -\frac{2\xi}{1+\xi}\,\hat{n} \times \vec{E}_{inc}(\vec{r}_s). \tag{6b}$$

Even though the two PO approximations according to eqns. (5) and (6) look rather different, the deviation for typical examples and for various incidence angles ϑ_i is usually less than one percent.

Because equations (6) are simpler to apply and since the determination of ϑ_i can be avoided, these formulations are preferred in the following. Note that in the shadowed region the equivalent currents are set to zero, the equations presented in this section are valid only in the illuminated zone.

4 Some details of the hybridization

As indicated in Fig. 1, the surface of the dielectric body is also subdivided into triangular patches. The equivalent surface current densities are expressed as linear superposition of basis functions with unknown coefficients:

$$\vec{J} = \sum_{k=1}^{N_J} \alpha_{J,k}\,\vec{f}_k \tag{7a}$$

$$\vec{M} = \sum_{k=1}^{N_M} \alpha_{M,k}\,\vec{h}_k. \tag{7b}$$

The basis functions $\vec{f}_k$ are identical to those used for metallic regions, $\vec{h}_k$ is approximately orthogonal to $\vec{f}_k$, see [7].

In principle, it might be possible to divide the surface of the dielectric body into a MoM– and a remaining PO–region. For instance, it could be useful to assign the shadowed part to the MoM– and the illuminated part to the PO–region. An example for such an allocation can be found in [3], where this principle is applied to a perfectly conducting sphere.

However, for a dielectric body we have not implemented the necessary coupling between a dielectric PO– and a dielectric MoM–region. Hence, in the following we do assume that the whole surface of the dielectric body (not metallic parts located nearby) is treated by PO. In this case, all the coefficients $\alpha_{J,k}$ and $\alpha_{M,k}$ in (7) can be determined by equating (7) with (6). After some straightforward manipulations one obtains

$$\alpha_{J,k} = \frac{1}{1+\xi}\,(\delta_k^+ + \delta_k^-)\,\frac{\vec{a}_{2,k} - \vec{a}_{1,k}}{|\vec{a}_{2,k} - \vec{a}_{1,k}|} \cdot \vec{H}_{inc}(\vec{r}_k) \tag{8a}$$

$$\alpha_{M,k} = -\frac{2\xi}{1+\xi}\,\frac{1}{l_k}\left(\frac{2A_k^+}{k_k^+}\,\delta_k^+\,\hat{t}_k^+ + \frac{2A_k^-}{k_k^-}\,\delta_k^-\,\hat{t}_k^-\right) \cdot \vec{E}_{inc}(\vec{r}_k). \tag{8b}$$

Some vectors and lengths required in this equation are defined in Fig. 5. Shadowing coefficients

$$\delta_k^{\pm} = \begin{cases} 0 & \text{triangle } T_k^{\pm} \text{ shadowed} \\ 1 & \text{triangle } T_k^{\pm} \text{ illuminated by the source} \end{cases} \tag{9}$$

have also been introduced in eqn. (8).

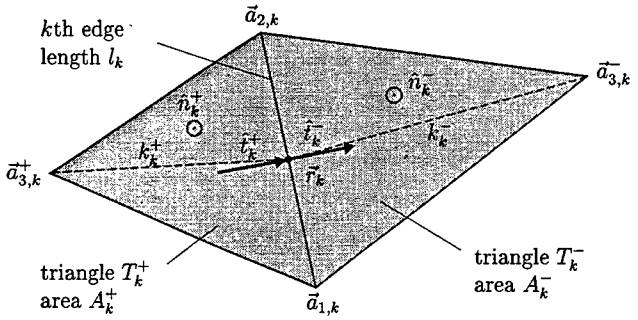
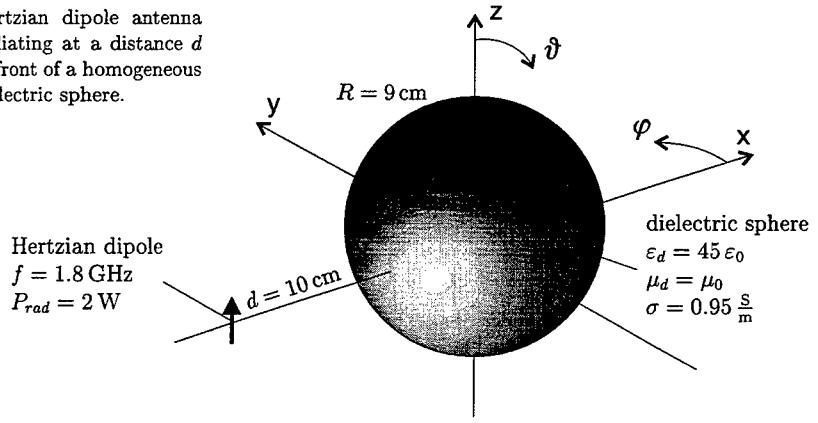


Fig. 5: Definition of some vectors and lengths in the two triangular patches $T_k^{\pm}$ adjacent to the kth edge.

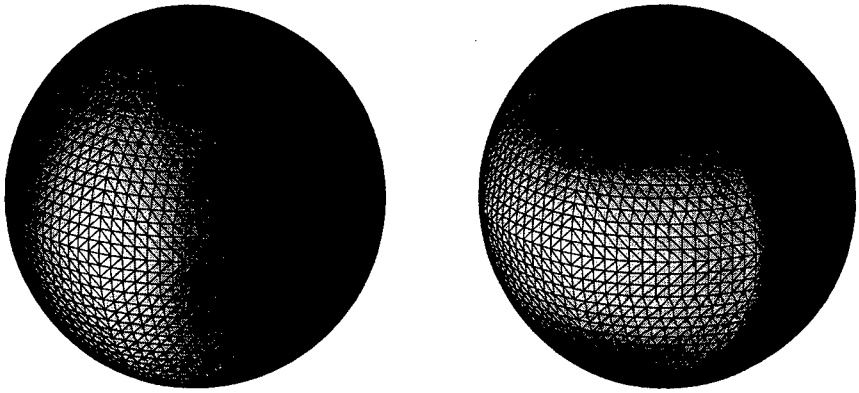
5 Example and results

Fig. 6: Hertzian dipole antenna radiating at a distance d in front of a homogeneous dielectric sphere.



The simple example of a structure consisting of a Hertzian dipole radiating in front of a dielectric sphere has been chosen here as an example to validate the formulation. The advantage of this configuration is that an exact solution is available by means of a special Green's function [8, 9]. The disadvantage is, that no metallic MoM–region is involved. However, if the field strength at an observation point $\vec{r}$ in Fig. 2 radiated by a source (here the Hertzian dipole can be interpreted

as a basis function $\vec{f}_n$ located at $\vec{r}'$) is correctly predicted by the PO approximation, which can be judged by considering this example, then the generalization to more complex geometries is straightforward.

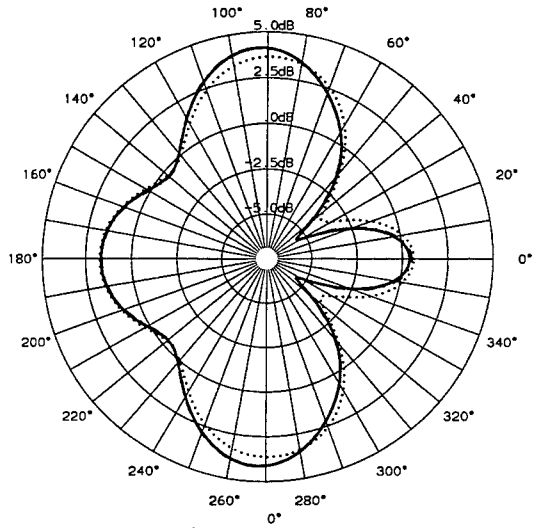


(a) equiv. electric current $\vec{J}^{PO}$ (b) equiv. magnetic current $\vec{M}^{PO}$

Fig. 7: Magnitude of the equivalent currents on the surface of the sphere based on PO according to eqn. (6).

The magnitude of the equivalent currents on the surface of the sphere based on PO is depicted in Figs. 7 (a) and (b). The shadow boundary is clearly visible in both figures: Only 26.3% of the spherical surface is illuminated, on the remaining 73.7% the current is approximated by zero. Comparing these currents to the MoM–results, which are not depicted here but which do not show the shadow boundary and have currents different from zero in the shadowed region, might lead to the conclusion that the PO solution cannot predict the scattered fields. The two radiation patterns in Figs. 8 and 9 demonstrate that the opposite is true. The solid line there represents the exact solution, the dotted line is the PO result. The two MoM solutions (dashed line: electric field integral equation EFIE, e.g. [10, 11]; dashed–dotted line: PMCHW formulation [12]) are in excellent agreement to the exact curve, while there are some differences visible in the PO solution, especially in the vertical cut in Fig. 9. However, e.g. for the optimization of mobile communication antennas, the achieved accuracy is sufficient, which is confirmed e.g. by the computed antenna efficiency, see Table 1. By hybridizing PO and MoM for the dielectric body as indicated above, a further improvement in accuracy can be expected.

In Table 1 we have also compared memory and CPU–time requirement. The surface of the sphere is $A = 3.66\lambda_0^2 = 165.5\lambda^2$ with the free space wavelength λ_0 and the wavelength λ in the dielectric material. We have used 5512 triangular patches, i.e. about 33 per square wavelength area. The memory requirement for the matrix of the MoM solution is about 4 GByte (this can be reduced to 261 MByte using two planes of symmetry). The superiority of the PO solution becomes clearly obvious from Table 1.



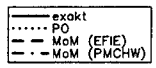


Fig. 8: Horizontal radiation pattern (directivity) in the plane $\vartheta = 90°$ as a function of the angle φ.

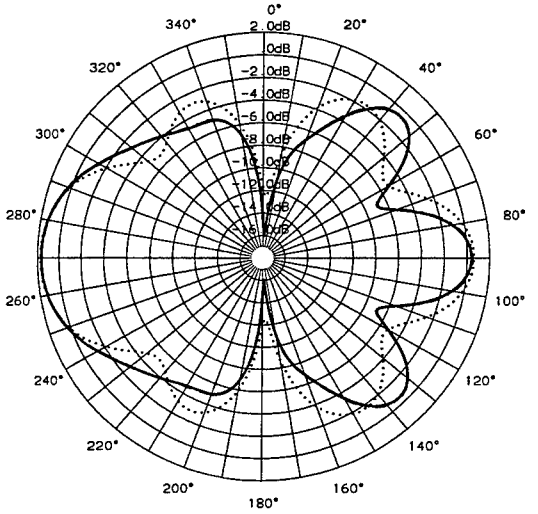




Fig. 9: Vertical radiation pattern (directivity) in the plane $\varphi = 0°$ as a function of the angle ϑ.

6 Conclusions

It has been shown that by hybridizing MoM for metallic structures with PO for dielectric bodies, a flexible and fast tool is available e.g. for the optimization of antennas on mobile telephones operating at high frequencies (e.g. PCS 1800 system at 1.8 GHz) taking the effect of the human body into account.

TABLE 1: Computed antenna efficiency, CPU–time and memory requirement for the analysis of a Hertzian dipole radiating in front of a homogeneous dielectric sphere.

	exact solution	MoM (EFIE)	MoM (PMCHW)	PO
no. of triangles	—	5512	5512	5512
no. of basis functions	—	16536	16536	16536
antenna efficiency	0.8684	0.8690	0.8695	0.8904
memory for the matrix	—	4,07 GByte	4,07 GByte	—
computer	Pentium PC 100 MHz	CRAY T3E (32 nodes)	CRAY T3E (32 nodes)	Pentium PC 100 MHz
CPU–time	11.3 sec	18.7 min	52.1 min	20.4 min

References

[1] E. K. Miller, "A selective survey of computational electromagnetics," *IEEE Transactions on Antennas and Propagation*, vol. 36, pp. 1281–1305, 1988.

[2] U. Jakobus and F. M. Landstorfer, "Improved PO–MM hybrid formulation for scattering from three–dimensional perfectly conducting bodies of arbitrary shape," *IEEE Transactions on Antennas and Propagation*, vol. 43, pp. 162–169, Feb. 1995.

[3] U. Jakobus and F. M. Landstorfer, "Current–based hybrid moment method analysis of electromagnetic radiation and scattering problems," *Applied Computational Electromagnetics Society Journal*, vol. 10, pp. 38–46, Nov. 1995.

[4] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Transactions on Antennas and Propagation*, vol. 30, pp. 409–418, May 1982.

[5] E. C. Jordan and K. G. Balmain, *Electromagnetic Waves and Radiating Systems*. Prentice–Hall, Englewood Cliffs, 1968.

[6] M. Ando, "Physical optics," in *Analysis Methods for Electromagnetic Wave Problems* (E. Yamashita, ed.), vol. 2, ch. 4, pp. 131–176, Artech House, Boston, 1990.

[7] U. Jakobus and F. M. Landstorfer, "Novel basis function for the equivalent magnetic current in the method of moments solution of dielectric scattering problems," *Electronics Letters*, vol. 29, pp. 1272–1273, July 1993.

[8] C.-T. Tai, *Dyadic Green Functions in Electromagnetic Theory*. IEEE Press, New York, 1994.

[9] H.-O. Ruoß, U. Jakobus, and F. M. Landstorfer, "Efficient EM analysis of hand–held mobile telephones close to human head using modified method of moments," *Electronics Letters*, vol. 31, pp. 947–948, June 1995.

[10] E. Arvas, A. Rahhal-Arabi, A. Sadigh, and S. M. Rao, "Scattering from multiple conducting and dielectric bodies of arbitrary shape," *IEEE Antennas and Propagation Magazine*, vol. 33, pp. 29–36, Apr. 1991.

[11] J. R. Mautz and R. F. Harrington, "Electromagnetic scattering from a homogeneous material body of revolution," *Archiv für Elektronik und Übertragungstechnik*, vol. 33, no. 2, pp. 71–80, 1979.

[12] A. J. Poggio and E. K. Miller, "Integral equation solutions of three–dimensional scattering problems," in *Computer Techniques for Electromagnetics* (R. Mittra, ed.), ch. 4, pp. 159–264, Pergamon Press, Oxford, 1973.

EMAP5: A 3D HYBRID FEM/MoM CODE

Yun Ji and Todd H. Hubing
University of Missouri-Rolla

Abstract--EMAP5 is a numerical software package designed to model electromagnetic problems. It employs the finite element method (FEM) to analyze a volume, and employs the method of moments (MoM) to analyze the current distribution on the surface of the volume. The two methods are coupled through the electric fields on the dielectric surface. The filed behavior at dielectric/metal junctions is modeled by three-way basis functions. EMAP5 can model three kinds of source: incident plane wave, voltage sources on metal patches and impressed current sources in the finite element region. Three numerical examples are provided to demonstrate the validity of the code.

I. FORMULATION

Although details of EMAP5 formulation are provided in [1][2], a brief summary is provided below. The general structure of interest is shown in Figure 1. A dielectric volume V_2 has electrical properties (ε_2, μ_2). It is enclosed by a surface S_2. A conductive volume V_3 is enclosed by a conductive surface S_c. The fields within V_3 vanish. V_1 denotes the volume outside of V_2 and V_3, and has electrical properties (ε_1, μ_1). V_1 is assumed to be free space. ($\mathbf{E}_1$, $\mathbf{H}_1$) and ($\mathbf{E}_2$, $\mathbf{H}_2$) denote the electric and magnetic fields in V_1 and V_2, respectively. The unit normal vectors for S_2 and S_c are defined pointing outward toward V_1. The structure is excited by an incident wave ($\mathbf{E}^i$, $\mathbf{H}^i$) or impressed sources ($\mathbf{J}^{int}$, $\mathbf{M}^{int}$). The scattered electric and magnetic fields are ($\mathbf{E}^s$, $\mathbf{H}^s$). The objective is to solve for the scattered fields ($\mathbf{E}^s$, $\mathbf{H}^s$) or the surface electric current density on S_c.

1. Discretization of FEM From Maxwell equations, the double curl equation in terms of $\mathbf{E}$ can written:

$$\nabla\times\left(\frac{1}{j\omega\mu_0\mu_r}\nabla\times\mathbf{E}(\mathbf{r})\right)+j\omega\varepsilon_0\varepsilon_r\mathbf{E}(\mathbf{r})=-\mathbf{J}^{int}(\mathbf{r})-\frac{1}{j\omega\mu_0\mu_r}\nabla\times\mathbf{M}^{int}(\mathbf{r}) \tag{1}$$

After multiplying Eq. (1) by a weighting function $\mathbf{w}(\mathbf{r})$ and integrating over the finite element domain V_2, one obtains the FEM weak form as follows:

$$\int_{V2}\left[\left(\frac{1}{j\omega\mu_0\mu_r}\nabla\times\mathbf{E}(\mathbf{r})\right)\bullet(\nabla\times\mathbf{w}(\mathbf{r}))+j\omega\in_0\in_r\mathbf{E}(\mathbf{r})\bullet\mathbf{w}(\mathbf{r})\right]dV=\int_{S2}(\hat{\mathbf{n}}\times\mathbf{H}(\mathbf{r}))\bullet\mathbf{w}(\mathbf{r})\,dS$$
$$-\int_{V2}\left[\mathbf{J}^{int}(\mathbf{r})+\frac{1}{j\omega\mu_0\mu_r}\nabla\times\mathbf{M}^{int}(\mathbf{r})\right]\bullet\mathbf{w}(\mathbf{r})\,dV \tag{2}$$

Tetrahedral elements are used to discretize the volume V_2. Basis and weighting functions proposed by M. L. Barton and Z. J. Cendes [3] are chosen here. Each basis function is defined within a tetrahedron and is associated with one of the six edges. The electric field $\mathbf{E}$ within volume V_2 can be expanded as:

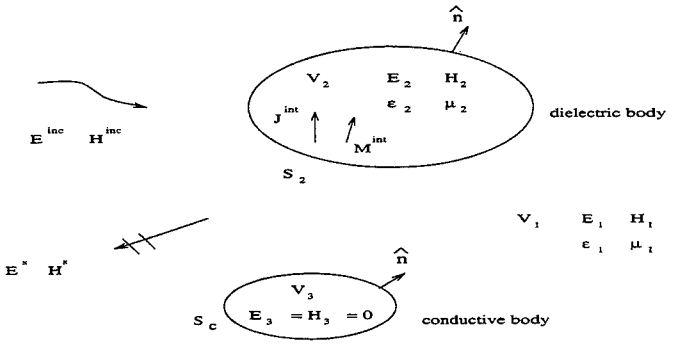


Fig. 1. A dielectric obect and a conductive object illuminated by E^i, H^i or J^{int}, M^{int}.

$$\mathbf{E}(\mathbf{r}) = \sum_{n=1}^{N_v} E_n \, \mathbf{w}_n(\mathbf{r})$$

where N_v is the total number of interior edges, and $\{E_n\}$ is a set of unknown complex scalar coefficients.

The surface integral term in Eq. (2) can be evaluated by using a surface basis function $\mathbf{f}(\mathbf{r})$, which are discussed later. Using a Galerkin's approach, a discrete form of Eq. (2) is obtained:

$$\begin{bmatrix} A_{ii} & A_{id} \\ A_{di} & A_{dd} \end{bmatrix} \begin{bmatrix} E_i \\ E_d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & B_{dd} \end{bmatrix} \begin{bmatrix} 0 \\ J_d \end{bmatrix} + \begin{bmatrix} 0 \\ g \end{bmatrix} \tag{3}$$

where $[J_d]$ is a set of unknown complex scalar coefficients for the surface electric current densities on S_d. S_d is defined as S_2 if the conductive body is not adjacent to the dielectric body; Otherwise, $S_d = S_2 - (S_2 \cap S_c)$. The unknown coefficients [E] are partitioned according to edge type. The two categories are interior edges, which are denoted by a subscript *i* in Eq. (3), and dielectric boundary edges, which are denoted by a subscript *d* in Eq. (3). $[g^{int}]$ is the forcing term. Details of how to evaluate the elements of [A], [B] and $[g^{int}]$ are provided in [1][2].

2. Discretization of MoM The MoM surface integral equation is [4]:

$$\mathbf{E}^{inc}(\mathbf{r}) = \frac{1}{2}\mathbf{E}(\mathbf{r}) + \int_S \{ \mathbf{M}(\mathbf{r}') \times \nabla' G_0(\mathbf{r},\mathbf{r}') + j k_0 \eta_0 \mathbf{J}(\mathbf{r}') G_0(\mathbf{r},\mathbf{r}') - j\frac{\eta_0}{k_0} \nabla' \bullet \mathbf{J}(\mathbf{r}') \nabla G_0(\mathbf{r},\mathbf{r}') \} dS' \tag{4}$$

where $\mathbf{r} \in S$, $S = S_c \cup S_2$, η_0 and k_0 are the intrinsic impedance and wavenumber in free space, respectively, and

$$G_0(\mathbf{r},\mathbf{r}') = \frac{e^{-jk_0|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}$$

is the Green's function in free space. The surface equivalent electric and magnetic currents are defined as:

$\mathbf{J}(\mathbf{r}') = \hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r}')$ $\mathbf{r}'$ on S; $\mathbf{M}(\mathbf{r}') = \mathbf{E}(\mathbf{r}') \times \hat{\mathbf{n}}$ $\mathbf{r}'$ on S_d.

$\mathbf{M}(\mathbf{r}')$ vanishes on S_c. $\mathbf{J}(\mathbf{r}')$ and $\mathbf{M}(\mathbf{r}')$ can be approximated by using the triangular basis function $\mathbf{f}_n(\mathbf{r})$ proposed by S. M. Rao *et al.* [5].

On surface S_d, the MoM basis function $\mathbf{f}_n(\mathbf{r})$ and the FEM basis function $w_n(\mathbf{r})$ are related by:

$$w_n(\mathbf{r}) = \hat{\mathbf{n}} \times \mathbf{f}_n(\mathbf{r}) \qquad \mathbf{r} \in S_d$$

$\mathbf{J}(\mathbf{r}')$, $\mathbf{M}(\mathbf{r}')$ can be expanded as:

$$\mathbf{J}(\mathbf{r}') = \hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r}') = \sum_{n=1}^{N_s} J_n \mathbf{f}_n(\mathbf{r}') \qquad \mathbf{M}(\mathbf{r}') = \mathbf{E}(\mathbf{r}') \times \hat{\mathbf{n}} = \sum_{n=1}^{N_d} E_n \mathbf{f}_n(\mathbf{r}')$$

where N_s is the total number of edges on the surface S, and N_d is the total number of edges on the surface S_d. $\{J_n\}$ and $\{E_n\}$ are unknown complex scalar coefficients.

The weighting functions chosen are $\mathbf{f}_n(\mathbf{r})$, n=1, ... N_s. After $\mathbf{f}_n(\mathbf{r})$ are multiplied to Eq. (4), Eq. (4) can be discretized into Eq. (5), which is a matrix equation. Edges on S_d and S_c are grouped together respectively.

$$\begin{bmatrix} C_{dd} & C_{dc} \\ C_{cd} & C_{dd} \end{bmatrix} \begin{bmatrix} J_d \\ J_c \end{bmatrix} = \begin{bmatrix} D_{dd} & 0 \\ D_{cd} & 0 \end{bmatrix} \begin{bmatrix} E_d \\ 0 \end{bmatrix} - \begin{bmatrix} F_d^i \\ F_c^i \end{bmatrix} \qquad (5)$$

$[F^i]$ is the forcing term due to the incident wave. A description of how to evaluate the elements of [C], [D] and $[F^i]$ are is provided in [1][2]. $[J_c]$,$[J_d]$, $[E_i]$, $[E_d]$ can be solved from Eq. (3) and Eq. (5).

II. COMPONENTS OF THE EMAP5 SOFTWARE PACKAGE

The EMAP5 software package includes three major components: SIFT5, EMAP5 and FAR.

1. SIFT5: The Input File Translator Standard Input File Translator Version5 (SIFT5) is designed to generate input files for the field solver EMAP5. SIFT5 reads a text file in the SIFT format [5]. Users can describe the structure of interest by using eleven keywords shown in Table I. The physical geometry, source, and the output requirements must be specified.

The input file for SIFT5 should have a .sif suffix. The output file of SIFT5 has a .in suffix. For example, if a user has composed an input file E1.hbs, the following command will generate an input file E1.in for EMAP5.

% sift5 E1.sif

2. EMAP5: The FEM/MoM Field Solver EMAP5 is the hybrid FEM/MoM field solver. It reads a file generated by SIFT5. The input file should have a .in suffix. A file with a .log suffix is generated to log running status and error messages. EMAP5 will print fields within areas specified by the keyword "output", to one or more output files. All equivalent surface currents **J** and **M** will be printed out by using the keyword "default_out". An example of how to run EMAP5 follows:

% emap5 E1.in

EMAP5 will read the mesh file E1.in as its input. In addition, E1.log will be generated immediately as the log file.

3. FAR: The Far Field Calculator FAR is a program used to calculate the far field radiation pattern. The far fields are calculated from the equivalent surface currents **J** and **M**. FAR needs two input files. One is the file generated by SIFT5, and the other is the default output file generated by

Table I. Syntax of keywords for SIFT5.

keyword	position coordinates	cell dimensions	sub attributes
#			
unit			number in m, cm or mm
boundary	x1 y1 z1 x2 y2 z2		
celldim	p1, p2	Δp	axis (x, y or z)
dielectric	x1 y1 z1 x2 y2 z2		ε' ε'' (The real and imaginary part of the complex permittivity)
conductor	x1 y1 z1 x2 y2 z2	Δx Δy Δz	
eplane			frequency,$\theta 1$, $\varphi 1$, $\theta 2$, $\varphi 2$, magnitude,
vsource	x1 y1 z1 x2 y2 z2		frequency, polarization (x, y, z), magnitude,
isource	x1 y1 z1 x2 y2 z2		frequency, polarization (x, y, z), magnitude,
output	x1 y1 z1 x2 y2 z2		axis(x, y, z) filename
default_out			filename

EMAP5. Assuming the default output file of EMAP5 is E1.out, and the input file is E1.in, the following command will run the far field calculator.

% far E1.in E1.out far.out

where far.out is the file to which the far field will be printed when the program terminates. The far.out file will contain an array of (θ, φ, E_θ, E_φ) data. E_θ and E_φ, whose units are volts/meter, are the E fields at point (R, θ, φ) in spherical coordinates.

III. NUMERICAL RESULTS

The first configuration is a flat dipole antenna in free space. Although EMAP5 is a FEM/MoM code, it can model configurations that require only one method to analyze. In this case, only the MoM portion of the code is employed. As shown in Figure 2, a center-fed flat dipole has a width of one millimeter and a length of 44 centimeters. It is fed by a 300-MHz voltage source with a magnitude of one volt. The input file for SIFT5 is as follows:

```
# example 1: a flat dipole antenna driven by a voltage source in the middle
unit 0.5 mm
conductor    0  0  2  880  2  2   10 1  1
vsource    440  0  2  440  2  2  300 x  1.0
```

output 0 0 2 840 2 2 y example1.out

The structure is divided into 268 triangles. The total number of unknown edges is 438. The output of EMAP5 is the surface electric current J along the flat dipole. The current at the source is given by:

$I = J_f * w$

where J_f is a complex number denoting the surface current across the source edge, and w is the width of the edge. In this example, two edges are used to model the source. Thus, the total current is the sum of currents across the two edges. The input impedance of the dipole is given by:

$$Z_{in} = \frac{1.0\,v}{I}$$

Figure 3 shows the input resistance, input reactance and impedance obtained by EMAP5 as the dipole length is adjusted from 38-53 cm, with a comparison of analytical results by treating the flat dipole as a cylindrical dipole with an equivalent radius[6]. Good agreement between EMAP5 and theoretical results is achieved.

The second configuration, as shown in Figure 4, is a flat dipole consisting of two quarter-wavelength traces driven by a 533-MHz voltage source with a magnitude of one volt. In this case, the source is located within the FEM region. Since the width of the traces is very small compared with the width of the FEM region, a non-uniform mesh is used to discretize the structure. Near the junction and source areas, small cells are used. Initially, the relative permittivity of the dielectric slab is set to 1.0. Thus, the configuration is a half-wave dipole in free space. The source is modeled as a current filament that coincides with two tetrahedron edges. After the E fields along these two edges are obtained, the voltage drop along the current filament can be calculated. The input file for SIFT5 is as follows,

 # example 2: a dipole driven by a current source located within the FEM region
unit 0.25 mm
boundary 0 0 0 164 82 2
celldim 0 2 2 x
celldim 2 162 16 x

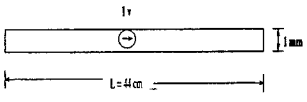


Fig. 2. A flat dipole with a width width of 1 mm.

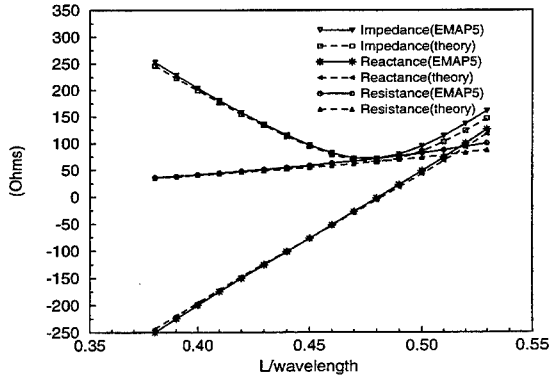


Fig. 3. Input impedance of a flat dipole length of 44 cm and a antenna with L=38~53 cm, width = 1 mm.

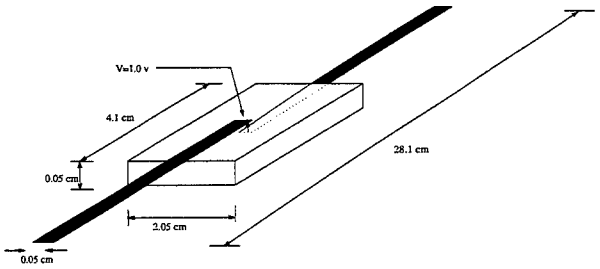


Fig. 4. Geometry of a PCB antenna.

```
celldim     162 164   2    x
celldim     0    40   10   y
celldim     40   42   1    y
celldim     42   82   10   y
celldim     0    2    1    z
dielectric  0    0    0  164  82  2   1.0  0.0
conductor  -480  40   2   82   42  2   10   1   1
conductor   82   40   0  644   42  0   10   1   1
isource     82   41   0   82   41  2  533 z   1
output     -480 40   2    0   42  2   y  example2.out
output      164  40   0  644   42  0   y  example2.out
```

The FEM region is divided into 1200 tetrahedra. The boundary is divided into 2632 edges. The number of unknown edges for the final matrix equation is 1900. The current distribution on the traces obtained using EMAP5 is plotted in Figure 5. For comparison, the results obtained using the Numerical Electromagnetics Code (NEC) and the IBM EM Simulator are also plotted. Figure 6 shows the results obtained by EMAP5 and the IBM EM Simulator when the dielectric constant is set to 10.0. In both cases, the results obtained using the different methods are similar.

The third configuration shown in Figure 7 is a dielectric cube with 0.2λ on a side, where λ is the wavelength in free space. It is illuminated by an incident wave, which travels along the +z axis. The E field is polarized along the x axis with a magnitude of one volt/meter. This example has been previously analyzed by T. K. Sarkar *et al.*[7], B. J. Rubin and S. Dajiavad [8]. First, the dielectric constant of the cube is set to 1-j1000. The input file for SIFT5 is as follows,

```
# example 3:  a dielectric cube (εr =1-j1000 ) illuminated by a plane wave
unit        1  mm
boundary    -50  -50  -50  50  50  50
celldim     -50   50   25   x
celldim     -50   50   25   y
celldim     -50   50   25   z
dielectric -50  -50  -50  50  50  50  1.0  -1000.0
eplane      600  90    0   0    0  1.0
default_out     example3.out
```

The dielectric cube is divided into 64 bricks, then 320 tetrahedra. In Figure 8, the normalized far field obtained by EMAP5 is compared to those calculated in [7][8]. Second, the dielectric constant of the cube is set to 9.0. The "dielectric" line in the input file for SIFT5 need to be changed as follows,

dielectric -50 -50 -50 50 50 50 9.0 0.0

In Figure 9, the normalized far field obtained by EMAP5 is compared to those calculated in [7] [8]. In both cases, the results obtains by EMAP5 agrees with the references.

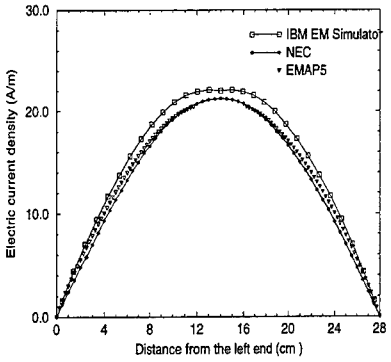


Fig. 5. Current distribution on the dipole(ε_r=1.0).

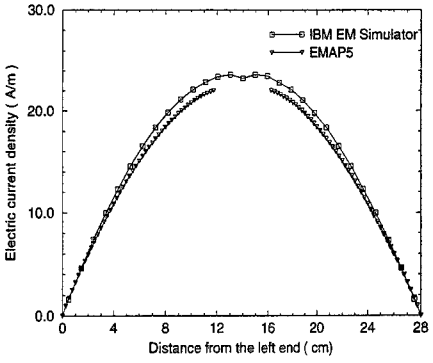


Fig. 6. Current distribution on the dipole (ε_r=10.0).

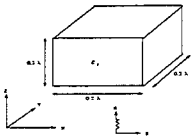


Fig. 7. A dielectric cube illuminated by a plane wave

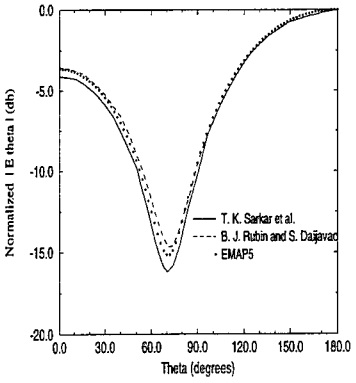


Fig. 8. Comparison of far field E_θ when ε_r = 1-j1000.

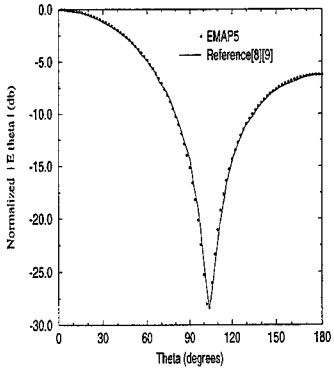


Fig. 9. Comparison of far field E_θ when ε_r =9.

Reference:

[1] M. W. Ali, T. H. Hubing and J. L. Drewniak, "A Hybrid FEM/MoM Technique for Electromagnetic Scattering and Radiation from Dielectric Objects with Attached Wires," *IEEE Trans. on Electromagnetic Compatibility*, vol. 39, pp. 304-314, November 1997.

[2] Mohammad W. Ali, "Development of a Hybrid 3D numerical Modeling Technique for Analyzing Printed Circuit Models with Attached Wires," Ph.D. dissertation, University of Missouri-Rolla, December 1996.

[3] M. L. Barton and Z. J. Cendes, "New Vector Finite Elements for Three-Dimensional Magnetic Field Computation," *Journal of Applied Physics*, vol. 61, pp. 3919-3921, April 1982.

[4] Johnson J. H. Wang, *Generalized Moment of Methods in Electromagnetics*, New York: John Wiley & Sons, 1990.

[5] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic Scattering by Surfaces of Arbitrary Shape," *IEEE Trans. on Antennas and Propagat.*, vol. 30, pp. 409-418, May 1982.

[6] T. H. Hubing, C. Hong-Him Lim and J. Drewniak, "A Geometry Description Language for 3D Electromagnetic Analysis Codes," *Proceedings of the 10th Annual Review of Progress in Applied Computational Electromagnetics*, pp. 417-422, March 1994.

[7] Constantine A. Balanis, *Antenna Theory Analysis and Design,* 2nd Edition, New York: John Wiley & Sons Inc, 1997.

[8] T. K. Sarkar, E. Arvas, and S. Ponnapalli, "Electromagnetic Scattering from Dielectric Bodies," *IEEE Trans. Antenna Propagat.*, vol 32, pp. 77-85, Jan. 1984.

[9] B. J. Rubin and S. Daijavad, "Radiation and Scattering from Structures Involving Finite-Size Dielectric Regions", *IEEE Trans. Antenna Propagat.*, vol 38, pp. 1863-1873, Nov., 1990.

Iterative Coupling of MoM and MMP for the Analysis of Metallic Structures Radiating in the Presence of Dielectric Bodies

Hans-Oliver Ruoss, Ulrich Jakobus and Friedrich M. Landstorfer

Institut für Hochfrequenztechnik, University of Stuttgart,
Pfaffenwaldring 47, D-70550 Stuttgart, Germany

Abstract

An efficient hybrid calculation method combining the Method of Moments (MoM) with a Multiple Multipole (MMP) technique is proposed. It provides an accurate modelling of complex metallic structures in regions treated by the MoM, while dielectric bodies are taken into account by means of the MMP. An iterative coupling scheme is applied, taking the scattered field of the one method as a corrective term to the other. This kind of coupling requires only small changes to the conventional MoM and MMP formulations, hence it is very attractive for the combination of already existing codes. Data exchange is done using the Message Passing Interface (MPI) allowing single processes to be executed in parallel.

1 Introduction

Three independent „classes" of numerical techniques have so far been established for computational electromagnetics: The first might be called as *method of fields*, as the electric and magnetic fields are the basis of calculation. It yields differential equations and generally necessitates a 3-dimensional discretisation of the space considered. The latter has to be limited by absorbing boundary conditions, but may contain any kind of inhomogeneities. Representatives of this class of numerical methods are e.g. the Finite Element Method (FEM) and the Finite Difference Time Domain (FDTD) technique.

In a second class, the *method of sources*, currents and charges are taken as the basis of the calculations, which leads to integral equations that are usually solved applying the Method of Moments (MoM) [1]. With this class, only a 2-dimensional discretisation is necessary for metallic surfaces and for dielectric bodies when using the surface equivalence principle [2]. The infinity of free space is easily and exactly taken into account, which makes this method particularly attractive for solving radiation problems. Bodies with inhomogeneous dielectric properties can be treated using the volume equivalence principle [3], which, however, is very time- and memory-consuming when the bodies considered are electrically large.

In the third class, the electric and magnetic fields are calculated by a weighted superposition of particular solutions of Maxwell's equations, the so-called *Multiple Multipoles* (MMP) [4]. While the

MoM is more attractive for metallic structures with sharp edges (e.g. thin wire antennas, flat plates, cubes, etc.), bodies with a smooth surface (e.g. sphere-like dielectric bodies) can often be dealt with in a more efficient manner by applying the MMP technique.

With a great number of electromagnetic radiation and scattering problems hybrid techniques combining the advantages of the single methods allow the investigation of very complex structures and yield good results with small computation time and moderate main memory requirements (e.g. a hybrid MoM / PO method as presented in [5] or the use of a specific Green's function in connection with MoM, see [6]). A combination of MoM and MMP can be carried out in different ways: In [7] rooftop basis functions (as usually applied with the MoM) have been included as a new type of basis function in the MMP, while in [8] the two methods are directly coupled. In the latter, one large system of linear equations is constructed by using the generalised point matching principle. In the following a new iterative coupling mechanism combining MoM and MMP is introduced. It avoids large system matrices and consequently offers short calculation time and small memory requirements. A similar iterative scheme but for the combination of two MMP calculations has been proposed in [9].

In *section 2 and 3* the theoretical background of the two conventional techniques MoM and MMP is briefly reviewed, while *sections 4 and 5* concentrate on the hybrid formulation and the aspects of coupling the two methods. In *section 6* the possibilities of data exchange applying the MPI are presented. An example is given in *section 7*.

2 Method of Moments (MoM)

Here the MoM is restricted in its application to metallic structures. Consider the mobile telephone depicted in Fig. 1. The surface current density $\vec{J}$ on the case and the line current I along the antenna are approximated by a linear superposition of basis functions f_n with unknown coefficients α_n:

$$\vec{J} = \sum_{n=1}^{N_J} \alpha_n^J \cdot \vec{f}_n^J \text{ and } I = \sum_{n=1}^{N_I} \alpha_n^I \cdot f_n^I. \qquad (1)$$

These currents radiate, and the scattered electromagnetic fields can be expressed as

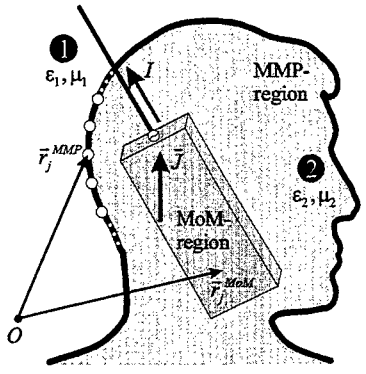


Figure 1: Calculation situation

$$\vec{E}_s^{MoM} = \vec{\mathcal{L}}_E^J\{\vec{J}\} + \vec{\mathcal{L}}_E^I\{I\} = \sum_{n=1}^{N_J} \alpha_n^J \cdot \vec{\mathcal{L}}_E^J\{\vec{f}_n^J\} + \sum_{n=1}^{N_I} \alpha_n^I \cdot \vec{\mathcal{L}}_E^I\{f_n^I\} \qquad (2a)$$

$$\vec{H}_s^{MoM} = \vec{\mathcal{L}}_H^J\{\vec{J}\} + \vec{\mathcal{L}}_H^I\{I\} = \sum_{n=1}^{N_J} \alpha_n^J \cdot \vec{\mathcal{L}}_H^J\{\vec{f}_n^J\} + \sum_{n=1}^{N_I} \alpha_n^I \cdot \vec{\mathcal{L}}_H^I\{f_n^I\}. \qquad (2b)$$

Operators $\vec{\mathcal{L}}_E^J$, $\vec{\mathcal{L}}_E^I$, $\vec{\mathcal{L}}_H^J$ and $\vec{\mathcal{L}}_H^I$ have been introduced in eqns. (2) representing the electric and magnetic field strengths, respectively, caused by $\vec{J}$ and I, e.g. $\vec{\mathcal{L}}_E^J$ is defined as follows:

$$\vec{\mathcal{L}}_H^J\{\vec{J}\} = -\frac{j}{4\pi\,\varepsilon\,\omega}\cdot\vec{\nabla}\iint_{A'}\left(\vec{\nabla}'_A\cdot\vec{J}(\vec{r}\,')\right)\cdot G(\vec{r},\vec{r}\,')\,dA' - j\omega\frac{\mu}{4\pi}\iint_{A'}\vec{J}(\vec{r}\,')\cdot G(\vec{r},\vec{r}\,')\,dA' \,. \tag{3}$$

$G(\vec{r},\vec{r}\,')$ denotes the free space Green's function, and ε_1, μ_1 are the material parameters in region ❶ according to Fig. 1.

3 Multiple Multipole Method (MMP)

When applying the MMP method, the calculation of the scattered electric and magnetic fields is based on a *direct field expansion*. From the scalar wave equation (Helmholtz equation)

$$\Delta\psi + k^2\psi = 0 \tag{4}$$

in the case of spherical coordinates $\vec{r} = (r,\theta,\phi)$ the scalar wave functions ψ_{mn} can be derived as

$$\psi_{mn}^{(c)}(k,\vec{r}) = z_n^{(c)}(kr)\cdot P_n^{|m|}(\cos\theta)\cdot e^{jm\phi} \,. \tag{5}$$

Here $z_n^{(c)}$ denotes the spherical Bessel functions of the n$^{\text{th}}$ order, while $P_n^{|m|}$ are the associated Legendre functions of the 1$^{\text{st}}$ kind with the order $|m|$ and degree n. Herewith, the vector wave functions $\vec{L}$, $\vec{M}$ and $\vec{N}$ can be constructed as follows:

$$\vec{L}_{mn}^{(c)} = \vec{\nabla}\cdot\psi_{mn}^{(c)} \quad , \quad \vec{M}_{mn}^{(c)} = \vec{\nabla}\times\left[\psi_{mn}^{(c)}\,\vec{r}\right] \quad , \quad \vec{N}_{mn}^{(c)} = \frac{1}{k}\vec{\nabla}\times\vec{\nabla}\times\left[\psi_{mn}^{(c)}\,\vec{r}\right] \tag{6}$$

For a linear, homogeneous, isotropic medium without charges, the scattered electric field can now be formulated as

$$\vec{E}_s^{MMP}(\vec{r}) = \sum_{n=1}^{\infty}\sum_{m=-n}^{n} j\,Z_F\cdot\left(a_{mn}\cdot\vec{N}_{mn}^{(4)}(k,\vec{r}) + b_{mn}\cdot\vec{N}_{mn}^{(1)}(k,\vec{r})\right) + \left(c_{mn}\cdot\vec{M}_{mn}^{(4)}(k,\vec{r}) + d_{mn}\cdot\vec{M}_{mn}^{(1)}(k,\vec{r})\right) \tag{7}$$

where Z_F denotes the field wave impedance in the medium considered, and a_{mn}, b_{mn}, c_{mn} and d_{mn} are unknown coefficients of the field expansion. When assuming $b_{mn} = d_{mn} = 0$ (for unbounded space), eqn. (7) is called a *multipole expansion*, while in the case of $a_{mn} = c_{mn} = 0$ the now formulated *normal expansion* is regulary in the origin of the coordinate system. In an arbitrary interim region all four terms of eqn. (7) have to be considered. To obtain the unknown coefficients, the boundary conditions on the surface between any two regions are fulfilled numerically (e.g. by applying a point matching algorithm). In a similar way, the scattered magnetic field is given by

$$\vec{H}_s^{MMP}(\vec{r}) = \sum_{n=1}^{\infty}\sum_{m=-n}^{n}\left(a_{mn}\cdot\vec{M}_{mn}^{(4)}(k,\vec{r}) + b_{mn}\cdot\vec{M}_{mn}^{(1)}(k,\vec{r})\right) - j/Z_F\cdot\left(c_{mn}\cdot\vec{N}_{mn}^{(4)}(k,\vec{r}) + d_{mn}\cdot\vec{N}_{mn}^{(1)}(k,\vec{r})\right) \tag{8}$$

4 Hybrid method

In Fig. 1 an exemplary calculation situation of a scattering problem is shown: A mobile telephone with monopole antenna is radiating in the close vicinity of the human tissue. The handset (metallic case and wire antenna) is located in the homogeneous region ❶ with material properties ε_1, μ_1. In this example, the region containing the handset represents the only MoM-region; the currents I on the metallic wires and the surface current density $\vec{J}$ shall be computed by means of the MoM. The antenna radiates in the presence of a homogeneous dielectric body (region ❷ with material properties ε_2, μ_2), which shall be taken into account by means of MMP. In any region, the electric and magnetic field can be expressed as

$$\vec{E}(\vec{r}) = \vec{E}_s^{\,MoM}(\vec{r}) + \vec{E}_s^{\,MMP}(\vec{r}) + \vec{E}_i(\vec{r}) \quad \text{and} \quad \vec{H}(\vec{r}) = \vec{H}_s^{\,MoM}(\vec{r}) + \vec{H}_s^{\,MMP}(\vec{r}) + \vec{H}_i(\vec{r}) \; . \tag{9a,b}$$

Index i represents the incident field ($\vec{E}_i$, $\vec{H}_i$ are the known impressed sources in the region considered), while index s denotes the scattered fields. $\vec{E}_s^{\,MoM}$ and $\vec{H}_s^{\,MoM}$ acc. to eqn. (2) denote the contribution of the MoM-region, while $\vec{E}_s^{\,MMP}$ and $\vec{H}_s^{\,MMP}$ acc. to eqns. (7) and (8) are the scattered fields calculated by means of the MMP technique.

For the MoM, the unknown coefficients $\alpha_n^{J,I}$ acc. to eqn. (1) can be obtained from the boundary condition $\vec{E}_{\tan} = 0$ on the perfectly conducting surfaces. Eqn. (9a) with (2a) leads to the integral equation

$$\sum_{n=1}^{N_J} \alpha_n^J \cdot \vec{\mathcal{L}}_E^J \left\{ \vec{f}_n^{\,J} \right\}_{\tan} + \sum_{n=1}^{N_I} \alpha_n^I \cdot \vec{\mathcal{L}}_E^I \left\{ f_n^I \right\}_{\tan} = -\vec{E}_{i,\tan} - \vec{E}_{s,\tan}^{\,MMP} \; . \tag{10}$$

As compared to a stand-alone MoM formulation, additional terms $\vec{E}_{s,\tan}^{\,MMP}$ are present in eqn. (10) representing the effect of the MMP region on the currents in the MoM region.

In the case of MMP, the unknown coefficients a_{mn}, b_{mn}, c_{mn} and d_{mn} acc. to eqn. (7) and (8) can be obtained by fulfilling the boundary conditions between any two regions. Considering the continuity of the tangential field components, this yields in our example (with only two regions ❶ and ❷)

$$\vec{E}_{s,1\tan}^{\,MMP} - \vec{E}_{s,2\tan}^{\,MMP} = -\vec{E}_{i,1} + \vec{E}_{i,2} - \vec{E}_{s,\tan}^{\,MoM} \tag{11a}$$

$$\vec{H}_{s,1\tan}^{\,MMP} - \vec{H}_{s,2\tan}^{\,MMP} = -\vec{H}_{i,1} + \vec{H}_{i,2} - \vec{H}_{s,\tan}^{\,MoM} \; . \tag{11b}$$

In eqn. (11) $\vec{E}_{s,\tan}^{\,MoM}$ and $\vec{H}_{s,\tan}^{\,MoM}$ are corrective terms acc. to eqn. (2) to the standard MMP formulation, caused by the influence of the currents radiating in the MoM-region and taking the coupling between the MMP- and the MoM-region into account.

In a more general situation, more than one MoM-region (e.g. sources outside and inside the dielectric body) can be easily taken into account.

5 Iterative coupling of MoM and MMP

The coupled linear system of equations as defined by eqn. (10) for any MoM-region and by eqn. (11) for the MMP-region (here an overdetermined system is solved by applying a least square approach) is now solved iteratively as follows:

As shown in Fig. 2, first the locations of the matching points $\vec{r}_j^{\,MMP}$ are introduced to the MoM-process, then the observation points of the MoM-region $\vec{r}_j^{\,MoM}$ are exchanged to the MMP-calculation using MPI (see section 6). Now the matrices of both methods can be formulated - this computation needs to be carried out only once (and the LU decomposition can be kept in memory leading to a fast solution by backwards substitution). At the begining a MoM-calculation acc. to eqn. (10) is carried out, assuming that the corrective terms $\vec{E}_{s,\tan}^{\,MMP}$ are zero for this „startup-calculation". This leads to the coefficients $\alpha_n^{J,I}$, and the field $\vec{E}_s^{\,MoM}$ acc. to eqn. (2a) can now be computed at the matching points $\vec{r}_j^{\,MMP}$. $\vec{H}_s^{\,MoM}$ can be computed as well acc. to eqn (2b) and with the help of eqn. (11) in a second step the expansion coefficients a_{mn}, b_{mn}, c_{mn} and d_{mn} of the MMP-algorithm can be obtained taking the corrective terms of the MoM-region(s) into account.

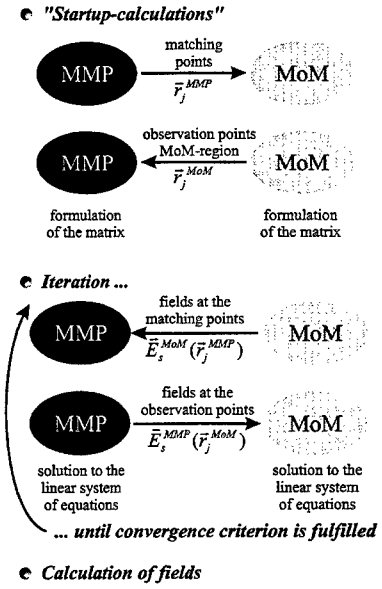


Figure 2: Iterative calculation scheme

Using eqn. (7), now $\vec{E}_s^{\,MMP}$ can be calculated at the matching points $\vec{r}_j^{\,MoM}$ of the MoM-region (note that for the MoM a Galerkin formulation is used and only the corrective terms are applied in a point matching sense), and in a new iteration the MoM integral equation (10) can be solved again, taking now the corrective terms $\vec{E}_s^{\,MMP}$ into account. Then a further MMP-calculations follows, and so on.

This iterative sequence of calculations applying the MoM- and MMP-algorithm is terminated, when the following criterion is fulfilled:

$$\left\| \vec{\alpha}_{i+1}^{\,MoM} - \vec{\alpha}_i^{\,MoM} \right\|_2 / \left\| \vec{\alpha}_i^{\,MoM} \right\|_2 < \varepsilon \,. \tag{12}$$

The vector $\vec{\alpha}_i^{\,MoM} = \left(\alpha_1^J \ldots \alpha_{N_J}^J \; \alpha_1^I \ldots \alpha_{N_I}^I \right)^T \Big|_i$ contains the MoM-coefficients of the ith iteration, and ε is a bound for the relative change of the currents in the MoM-region.

6 Communication using the Message Passing Interface (MPI)

As outlined above in section 5, we have to couple the MMP and MoM solution techniques. One might write a single code and perform the necessary data exchange (see Fig. 2) via global data structures, e.g COMMON blocks in the FORTRAN programming language. However, it is our intention to use already available MMP and MoM codes, in particular the 3D-MMP code [10] and the MoM code FEKO developed at the University of Stuttgart. These codes should run independently from one another and only a few changes should be required to the individual source codes.

One might use UNIX pipes for the communication as proposed in [11] or files on a hard disk where the different processes write and read the data. Our choice, however, has been to use MPI [12,13]. This provides a very fast and flexible means for the communication. It is also highly portable, since MPI implementations are available for a wide range of platforms and operating systems. We use the freely available MPICH package [14] on a cluster of connected PCs (running under Linux or Windows NT) and IBM as well as HP workstations (running under IBM-AIX and HP-UX), respectively. A further advantage of using MPI as compared to e.g. pipes or files on local hard disks is, that the different MoM and MMP processes can easily be executed in parallel on different workstations.

Once the initialization of MPI has been performed (only a few lines of additional code), the communication as indicated in Fig. 2 can be done by just using matching pairs of `MPI_SEND` and `MPI_RECV` commands. With these commands, single variables (such as the number of matching points corresponding to the array dimensions or a flag indicating whether the criterion (12) is fulfilled or not so that both MMP and MPI processes know whether to continue the iteration or not) can be sent, but it is also possible to send whole arrays with a single command, e.g. the field strength values $\vec{E}_s^{MMP}(\vec{r}_j^{MoM})$ during the iteration (see Fig. 2).

7 Example

As an example, Fig. 3 shows a mobile handset consisting of a cuboidal metallic case with the dimensions $2 \times 6 \times 12\,cm^3$ and a monopole antenna of length $h = 8\,cm$ and a wire radius $\rho = 0.5\,mm$. The handset operates at $f = 900\,MHz$ and the antenna radiates a power of $P_t = 2\,W$. For EMC-investigations, the human head is modeled as a homogeneous lossy dielectric sphere ($d = 2r = 18\,cm$) with the parameters $\varepsilon_r = 50$, $\mu_r = 1$ and $\sigma = 1.3\,S/m$. The position of the handset to the spherical head model is shown in Fig. 3.

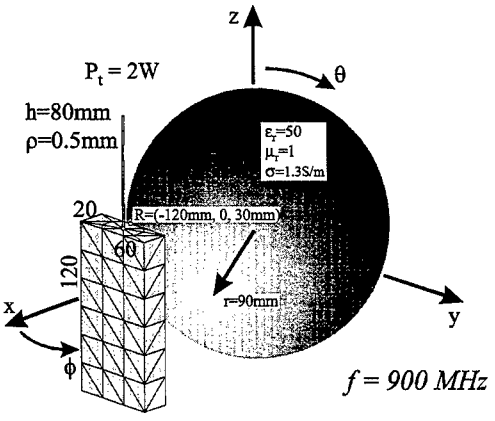


Figure 3: Example: Handset in front of a homog. spherical head model

The following Fig. 4 shows the x-component of the electric nearfield along the x-axis ($y = z = 0$). In the range $12\,cm \le x \le 14\,cm$ the observation point is inside the metal case, in the range $-9\,cm \le x \le +9\,cm$ (grey area) it is situated inside the dielectric sphere.

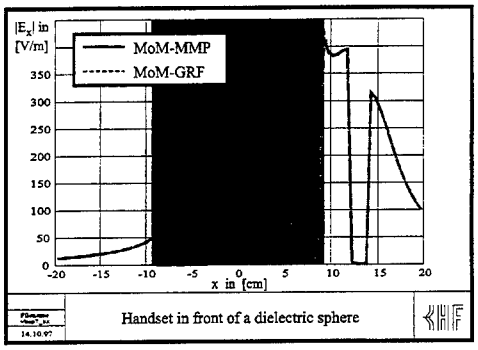


Figure 4: $|E_x|$ along the x-axis

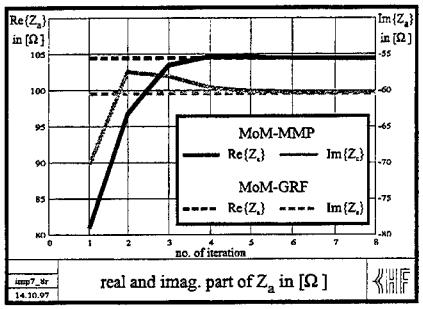


Figure 5: $\Re\{Z_a\}$ and $\Im\{Z_a\}$ as a function of iteration number

The nearfield data were calculated applying the MoM-MMP technique with one expansion each for the region inside and outside the dielectric sphere (both located in the centre of the sphere and with a maximal order $n_{max} = 8$) and 312 matching points on the surface of the sphere. In comparison a second calculation applying a MoM technique using a special Green's function (GRF) as presented in [6] was carried out. Fig. 4 shows a good agreement between the two calculations. Fig. 5 shows the real and imaginary part of the input impedance of the monopole antenna as a function of the iteration number. It can be seen from Fig. 5 that only a small number of iterations is required. The following Table 1 compares the calculation time and the main memory requirements for the two methods:

	MoM - MMP	MoM - GRF
Calculation time	6,6 min	87,7 min
Main memory	7,5 MByte	1,1 MByte

Table 1: Calculation parameters for the example

8 Conclusions

It has been shown that the analysis of metallic structures radiating in the presence of dielectric bodies can effectively be carried out by applying a hybrid MoM-MMP technique. An iterative coupling of the two methods (e.g. using Message Passing Interface (MPI) for data exchange and executing the single processes in parallel) requires only small changes to the conventional MoM and MMP formulations, hence it is very attractive for the combination of already existing codes. In an example, the radiation of a monopole antenna on a mobile handset in the close vicinity of the user's head has

been considered. Already with a small number of iterations the hybrid technique yields good results with small requirements of main memory and calculation time.

References

[1] R. F. Harrington: „*Field Computation by Moment Methods*", Macmillan Company, New York, 1968

[2] E. C. Jordan and K. G. Balmain: „*Electromagnetic Waves and Radiating Systems*", Prentice-Hall, Englewood Cliffs, 1968

[3] J. van Bladel: „*Electromagnetic Fields*", McGraw Hill Book Company, New York, 1964

[4] C. Hafner: „*The Generalized Multipole Technique for Computational Electromagnetics*", Artech House, Boston, 1990

[5] U. Jakobus and F. M. Landstorfer: „Current-based hybrid moment method analysis of electromagnetic radiation and scattering problems", *Applied Computational Electromagnetics Society Journal*, vol. 10, pp. 38-46, Nov. 1995

[6] H.-O. Ruoß, U. Jakobus, and F. M. Landstorfer: „Efficient EM analysis of hand-held mobile telephones close to human head using modified method of moments", *Electronics Letters*, vol. 31, pp. 947-948, June 1995

[7] C. Hafner, J. Waldvogel, J. Mosig, J. Zheng, and Y. Brand: „On the combination of MMP with MoM", *Applied Computational Electromagnetics Society Journal*, vol. 9, pp. 18-27, Sept. 1994

[8] F. Obelleiro, J. L. Rodriguez, and A. G. Pino: „Hybrid GMT-MoM method for solving electromagnetic scattering problems", *IEEE Transactions on Magnetics*, vol. 33, pp. 1424-1427, March 1997

[9] N. Kuster and L. Bomholt: „A Block Iterative Technique to Expand MMP's Applicability to EM Problems of Higher Complexity"; *IEEE Transactions on Microwace Theory and Techniques*, vol. 42, no. 5, pp. 875-883, May 1994

[10] C. Hafner and L. Bomholt, *„The 3D Electromagnetic Wave Simulator: 3D MMP Software and User's Guide"*, John Wiley and Sons, Chicester, 1993.

[11] P. Leuchtmann and A. Witzig: *„The Fieldinspector: A Graphic Field Representation System"*, in Proceedings of the 13th Annual Review of Progress in Applied Computational Electromagnetics, Monterey, pp 49-54, March 1997

[12] Message Passing Interface Forum: *„MPI: A message passing interface standard"*, version 1.1, Computer Science Department, University of Tennessee, Knoxville, June 1995. Available at http://www.mcs.anl.gov/mpi/index.html.

[13] W. Gropp, E. Lusk and A. Skjellum: *„Using MPI: Portable Parallel Programming with the Message Passing Interface"*, MIT Press, Cambridge, 1995.

[14] W. Gropp and E. Lusk: *„User's Guide for mpich, a Portable Implementation of MPI"*, Mathematics and Computer Science Division, Argonne National Laboratory, University of Chicago, April 1997. Available at http://www.mcs.anl.gov/mpi/mpich/index.html.

A Hybrid Algorithm for Frequency Selective Surface Analysis

Michael J. Walker
Department of Electrical Engineering, USAF Academy, CO, 80840-6236
e-mail: walkermj.dfee@usafa.af.mil; PH: 719.333.4213; Fax: 719.333.3756

November 26, 1997

Abstract

This article reviews the development of an algorithm for analyzing planar, periodic structures and predicting their specular reflection and transmission characteristics. The hybrid technique is based on the application of edge-based, finite element modelling within the structure and Moment Method radiation integrals on the exposed surfaces.

1 Introduction

The analysis of planar, periodic structures and the ability to predict their performance characteristics as Frequency Selective Surfaces (FSS) is becoming an increasingly difficult problem. When an FSS design includes exotic materials or complex shaped conducting elements, it cannot be analyzed with a technique specifically developed to compensate for any one design feature because they are often incompatible with one of the other features. There is a need for the development of a new algorithm that will allow for the accurate modeling and analysis of an FSS designed with several advanced features. It must be capable of providing specular reflection and transmission characteristics in response to an incident plane wave of arbitrary direction and polarization and still provide the accuracy and speed necessary to be a good design tool.

This paper summarizes the development of an algorithm which reduces the analysis of a planar, infinite FSS to a matrix equation using the Hybrid Finite Element Method (HFEM)[1]. It combines Finite Element Modeling (FEM) of the interior of the FSS structure with Method of Moments (MoM) radiation integrals applied to the exposed surfaces. A computer program was also created to implement the algorithm and validate its predicted results by comparison with other techniques.

2 Problem Description

The generalized model of an FSS used for the development of this algorithm is shown in Figure 1. The angle γ is called the skew angle and is used to measure the angular shift between consecutive columns in the $\hat{x}$ direction. The translational shift distance in the $\hat{z}$ direction is $D_x \cot\gamma$. The "front" face is defined as the surface upon which a plane wave is incident and occurs at $y = Y_0$, while the "back" face is the opposite surface at $y = Y_1$, where $Y_0 < Y_1$. The angles η and α describe the direction of propagation of the incident wave, $\hat{p} = -\overline{\beta}/k_o = -\sin\eta\cos\alpha\hat{x} + \cos\eta\hat{y} - \sin\eta\sin\alpha\hat{z}$. The polarization vectors are defined by first creating a vector $\hat{\alpha} = \cos\alpha\hat{x} + \sin\alpha\hat{z}$. Perpendicular polarization can then be defined as $\hat{e}_\perp = \hat{y} \times \hat{\alpha} = \sin\alpha\hat{x} - \cos\alpha\hat{y}$, and parallel polarization can be defined as $\hat{e}_\parallel = \hat{p} \times \hat{e}_\perp = -\cos\eta\cos\alpha\hat{x} - \sin\eta\hat{y} - \cos\eta\sin\alpha\hat{z}$. Finally, the polarization of the incident field is defined by the angle φ_p as $\hat{e} = \cos\varphi_p\hat{e}_\perp + \sin\varphi_p\hat{e}_\parallel$. This allows us to describe the incident plane wave using Eqn. 1.

$$\overline{E}_{in} = \hat{e}E_{in}e^{j(\omega t - k_o\hat{p}\cdot\overline{x})} = \hat{e}E_{in}e^{j(\omega t + \overline{\beta}\cdot\overline{x})} \tag{1}$$

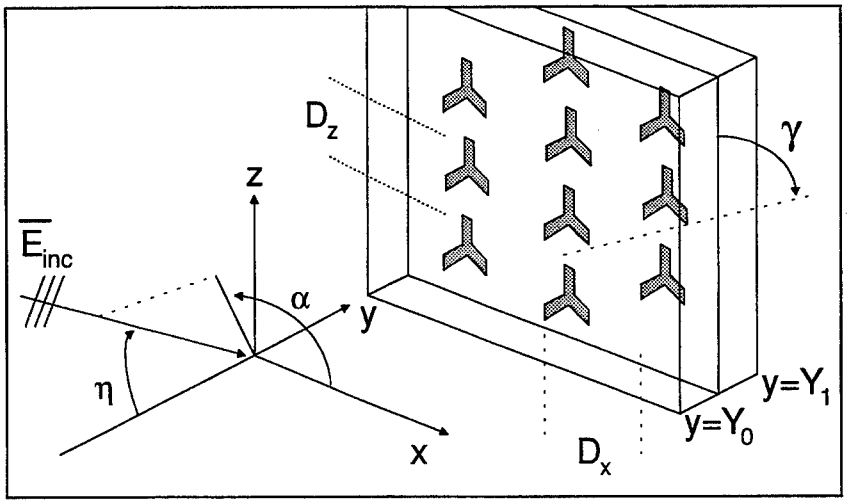


Figure 1: Generalized Geometry of a Frequency Selective Surface (FSS)

3 Analysis

Assuming the FSS is planar and infinite, the fields and currents within the structure must obey Floquet's Theorem; Eqn. 2. This limited the analysis to a single, primary cell of the FSS because it equated the fields and currents anywhere within the structure to phase-shifted copies of the fields and currents in a primary cell.

$$\overline{\Psi}(x + mD_x, z + nD_z + mD_x \cot\gamma) = \overline{\Psi}(x, z)\, e^{-jk_x mD_x} e^{-jk_z(nD_z + mD_x \cot\gamma)} \tag{2}$$

$$\text{where } k_x = k \sin\eta \cos\alpha$$

$$\text{and } k_z = k \sin\eta \sin\alpha.$$

The assumption that the FSS is planar and infinite also allowed the Equivalence Theorem to be used to divide the problem into three regions of interest: the free space region in front of the FSS ($y < Y_o$), the structural interior of the primary cell ($Y_o < y < Y_1$), and the free space region behind the FSS ($Y_1 < y$). These regions were analyzed separately, see Figure 2, but are linked by the resulting equivalent electric and magnetic surface currents defined in Eqn. 3.

$$\begin{aligned} \overline{J_1} &= -\overline{J}_{2A} = -\hat{y} \times \overline{H}(y = Y_o) \\ \overline{M_1} &= -\overline{M}_{2A} = \hat{y} \times \overline{E}(y = Y_o) \\ \overline{J_3} &= -\overline{J}_{2B} = \hat{y} \times \overline{H}(y = Y_1) \\ \overline{M_3} &= -\overline{M}_{2B} = -\hat{y} \times \overline{E}(y = Y_1) \end{aligned} \tag{3}$$

Modified field equations were used to allow the FSS design to contain exotic materials, including bianisotropic materials. These equations are

$$\begin{aligned} \overline{D} &= \varepsilon_o \overline{\overline{\varepsilon_r}} \cdot \overline{E} + \overline{\overline{\chi}}_{em} \cdot \overline{B} \\ \overline{H} &= \mu_o \overline{\overline{\mu_r^{-1}}} \cdot \overline{B} - \overline{\overline{\chi}}_{me} \cdot \overline{E} \end{aligned}$$

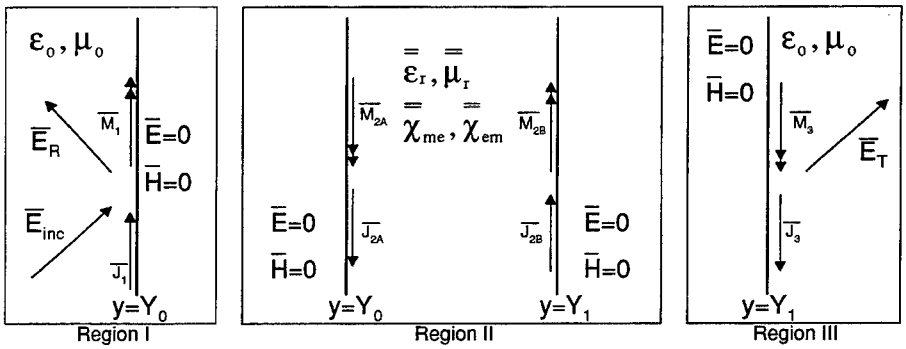


Figure 2: Three Sub-Regions of the Original FSS Problem

and they produced the reduced form of the wave equation shown in Eqn. 4.

$$L\left(\overline{E}\right) = \nabla \times \left[\overline{\overline{\mu_r^{-1}}} \cdot \left(\nabla \times \overline{E}\right)\right] + jk_o\eta_o\left[\nabla \times \left(\overline{\overline{\chi}}_{me} \cdot \overline{E}\right) - \overline{\overline{\chi}}_{em} \cdot \left(\nabla \times \overline{E}\right)\right] - k_o^2\overline{\overline{\varepsilon_r}} \cdot \overline{E} = 0. \tag{4}$$

The inner product of the wave equation with an unknown testing function, $\overline{W}$, created the weak form of the wave equation, shown in Eqn. 5.

$$\left\langle L\left(\overline{E}\right), \overline{W}\right\rangle = \iiint \left[\begin{array}{c} \left(\nabla \times \overline{W}^*\right) \cdot \overline{\overline{\mu_r^{-1}}} \cdot \left(\nabla \times \overline{E}\right) - k_o^2 \overline{W}^* \cdot \overline{\overline{\varepsilon_r}} \cdot \overline{E} \\ +jk_o\eta_o\left(\nabla \times \overline{W}^*\right) \cdot \overline{\overline{\chi}}_{me} \cdot \overline{E} \\ -jk_o\eta_o\nabla \times \left(\overline{\overline{\chi}}_{me} \cdot \overline{W}^*\right) \cdot \overline{E} \end{array} \right] dV$$

$$+jk_o\eta_o \iint \left(\overline{J}_{2A} - \overline{M}_{2A} \cdot \overline{\overline{\chi}}_{me}\right) \cdot \overline{W}^* dS$$

$$+jk_o\eta_o \iint \left(\overline{J}_{2B} - \overline{M}_{2B} \cdot \overline{\overline{\chi}}_{me}\right) \cdot \overline{W}^* dS = 0 \tag{5}$$

This weak form of the wave equation was applied to the interior region of the FSS structure's primary cell. The primary cell was divided into tetrahedra and the electric field and testing function were modelled as the weighted sum of edge-based, vector expansion functions defined over these tetrahedra.

$$\overline{E} = \sum_i c_i \overline{\Psi}_i\left(\overline{x}\right) \tag{6}$$

$$\overline{W} = \sum_j d_j \overline{\Psi}_j\left(\overline{x}\right) \tag{7}$$

The expressions in Eqns. 6 and 7 were substituted directly into the volume integral portion of Eqn. 5.

The weighted sum approximation of the electric field in the interior was also used to determine the magnetic surface currents in the two free-space regions using Eqn. 3. These currents were expanded over the surface of the entire FSS structure using Floquet's Theorem and their radiated fields were calculated. (The electric current

radiation was suppressed because the use of PEC equivalence to separate the regions created equal magnitude image currents in the opposite direction.) These radiation expressions were simplified by transforming them into the spectral domain using the Fourier Transform. This reduced the convolution integrals into multiplication, changed derivatives into dot products and accelerated the convergence of the Floquet summations. The resulting radiated fields were then inverse transformed and combined with any incident fields present (region 1 only) to get the total fields in each region. Finally, the total fields were used to derive new expressions for the surface currents which were used to evaluate the surface integrals in Eqn. 5.

A matrix equation was derived by setting the derivatives with respect to the testing function weights, d_j, equal to zero. This equation is

$$[A_{ji}]\,[c_i] = [b_j] \tag{8}$$

where

$$\begin{aligned}[A_{ji}] \;=\; & \iiint \left[\begin{array}{c} \left(\nabla \times \overline{\Psi}_j^*\right) \cdot \overline{\overline{\mu_r^{-1}}} \cdot \left(\nabla \times \overline{\Psi}_i\right) - k_o^2 \overline{\Psi}_j^* \cdot \overline{\overline{\varepsilon_r}} \cdot \overline{\Psi}_i \\ +jk_o\eta_o \overline{\Psi}_i \cdot \left[\overline{\overline{\chi}}_{em} \cdot \left(\nabla \times \overline{\Psi}_j^*\right) - \left(\nabla \times \left(\overline{\overline{\chi}}_{me} \cdot \overline{\Psi}_j^*\right)\right)\right] \end{array} \right] dV \\ & + \sum_{\substack{m,n \\ i,j\in S1 \text{ or } i,j\in S2}} \frac{j}{D_x D_z k_{ymn}} \overline{\widetilde{\Psi}}_i\left(k_{xm}, k_{zmn}\right) \cdot \overline{\overline{\widetilde{T}}}\left(k_{xm}, k_{zmn}\right) \cdot \overline{\widetilde{\Psi}}_j^*\left(k_{xm}, k_{zmn}\right) \\ & - \sum_{\substack{m,n \\ i,j\in S1 \text{ or } i,j\in S2}} \frac{jk_o\eta_o}{D_x D_z} \left(\widehat{y} \times \overline{\widetilde{\Psi}}_i\left(k_{xm}, k_{zmn}\right)\right) \cdot \overline{\overline{\chi}}_{me} \cdot \overline{\widetilde{\Psi}}_j^*\left(k_{xm}, k_{zmn}\right) \end{aligned} \tag{9}$$

$$[b_j] = \left\{ \begin{array}{ll} -2jk_o e^{j\beta_y Y_o} \left[\widehat{y} \times \left(\cos\varphi_p \widehat{e}_{\parallel} - \sin\varphi_p \widehat{e}_{\perp}\right)\right] \cdot \overline{\widetilde{\Psi}}^*_{j|_{y=Y_o}}\left(\beta_x, \beta_z\right); & j \in S1 \\ 0; & else \end{array} \right\} \tag{10}$$

and

$$\overline{\widetilde{\Psi}}\left(\overline{k}\right) = \Im\left[\overline{\Psi}\left(\overline{x}\right)\right] = \iint \Psi\left(\overline{x}\right) e^{-j\overline{k}\cdot\overline{x}} dx dz$$

$$\left[\overline{\overline{\widetilde{T}}}\right] = \left[\begin{array}{ccc} k_o^2 - k_z^2 & 0 & k_x k_z \\ 0 & 0 & 0 \\ k_x k_z & 0 & k_o^2 - k_x^2 \end{array} \right]$$

$$k_{xm} = \frac{2\pi m}{D_x} + \beta_x \qquad k_{zmn} = \frac{2\pi n}{D_z} + \frac{2\pi m \cot\gamma}{D_x} + \beta_z \qquad k_{ymn} = \sqrt{k_o^2 - k_{xm}^2 - k_{zmn}^2}$$

The solution to the matrix equation is the complex weights, c_i, of the finite element approximation of the electric field. The surface currents derived from this approximated field solution were used to determine the reflection and transmission coefficients of the FSS.

$$R_{\perp} = \left| \sum_{i\in S1} c_i \left[\frac{1}{D_x D_z \beta_y} \widehat{e}_{\perp} \cdot \overline{\overline{f}} \cdot \overline{\widetilde{\Psi}}_i\left(\beta_x, \beta_z\right) \right] - \cos\varphi_p \right| \tag{11}$$

$$T_{\perp} = \left| \sum_{i\in S2} c_i \left[\frac{1}{D_x D_z \beta_y} \widehat{e}_{\perp} \cdot \overline{\overline{f}} \cdot \overline{\widetilde{\Psi}}_i\left(\beta_x, \beta_z\right) \right] \right|$$

$$R_{\parallel} = \left| \sum_{i\in S1} c_i \left[\frac{1}{D_x D_z \beta_y} \widehat{e}_{\parallel} \cdot \overline{\overline{f}} \cdot \overline{\widetilde{\Psi}}_i\left(\beta_x, \beta_z\right) \right] - \sin\varphi_p \right|$$

$$T_{\parallel} = \left| \sum_{i\in S2} c_i \left[\frac{1}{D_x D_z \beta_y} \widehat{e}_{\parallel} \cdot \overline{\overline{f}} \cdot \overline{\widetilde{\Psi}}_i\left(\beta_x, \beta_z\right) \right] \right|$$

and

$$\overline{\overline{f}} = \left[\begin{array}{ccc} -\beta_y & 0 & 0 \\ \beta_x & 0 & \beta_z \\ 0 & 0 & -\beta_y \end{array} \right].$$

4 Results

4.1 Exotic Materials

A uniaxial dielectric has $\overline{\overline{\mu_r^{-1}}} = \overline{\overline{I}}$, $\overline{\overline{\chi}}_{me} = \overline{\overline{0}}$, and a diagonal relative permittivity tensor, $\overline{\overline{\epsilon_r}}$, with two of the three non-zero terms equal to each other. The axis with the unique permittivity component is call the optical axis. A slab of uniaxial material allows for three scenarios to be examined: Case I is when $\hat{y}$ is the optical axis, Case IIa is when $\hat{z}$ is the optical axis and the incident field is in the $\hat{y}-\hat{z}$ plane, and Case IIb is when $\hat{z}$ is the optical axis and the incident field is in the $\hat{x}-\hat{y}$ plane. The reflection coefficients of all three cases were analyzed using the HFEM algorithm and comparisons with exact calculations are shown in Figure 3. In each case, the primary cell of the FSS was a cube 1cm long on each side and divided into forty tetrahedron. The interior electric field was modelled with sixty edge-based, vector expansion functions and the frequency was 1GHz. The ordinary relative permittivity values were set to 2 and the optical axis' relative permittivity value was 4, resulting in expected Brewster angles of $\theta_{B,I} \simeq 49.1°$, $\theta_{B,IIa} \simeq 67.8°$, and $\theta_{B,IIb} \simeq 54.7°$. The ability of the HFEM algorithm to accurately locate the Brewster angles in each case verifies the algorithm's capability to analyze FSS designs which include exotic materials.

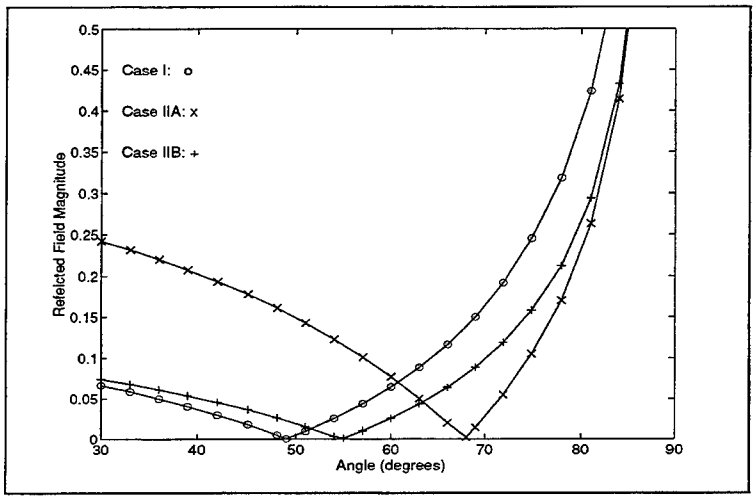


Figure 3: Uniaxial Slab Reflection Coefficients and Brewster Angles

4.2 Inductive Screen

An inductive screen is a thin layer of conductive material with an array of shaped apertures cut into it as shown in Figure 4. Inductive screens act as high pass filters with cutoff frequencies inversely proportional to aperture size. Zarillo and Aguiar calculated the transmission coefficient for an inductive screen based upon a "one-mode" approximation of the induced currents as a known function[2]. Figures 5 and 6 compare the HFEM algorithm with Zarillo and-Aguiar's results for power transmission vs frequency through a single layer, square inductive screen with $a = b = 0.8D_x = 0.8D_z$ and a plane wave incident at 30°. The electric field around the screen was modelled using 213 linear vector expansion functions. Figure 6 also includes a plot of the total power in the

directions of primary reflection and transmission. The curve accurately predicts the occurrence of a grating lobe above $D_x/\lambda = 0.67$, where the total power drops to approximately 0.707. With no lossy materials in the design, the power loss can only be accounted for by the appearance of grating lobes.

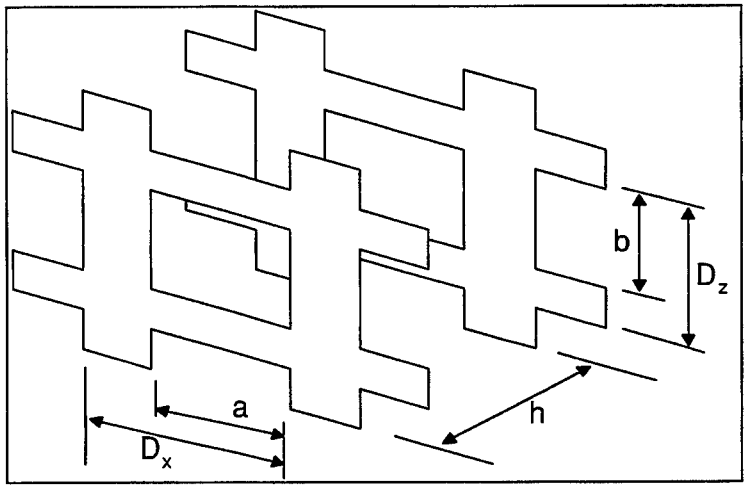


Figure 4: Geometry of a Square Inductive Screen

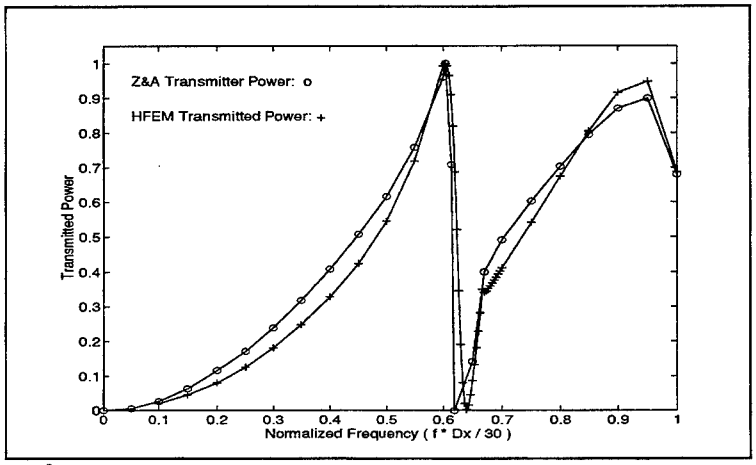


Figure 5: Square Inductive Screen, 30 Degree Incidence, Parallel Polarization

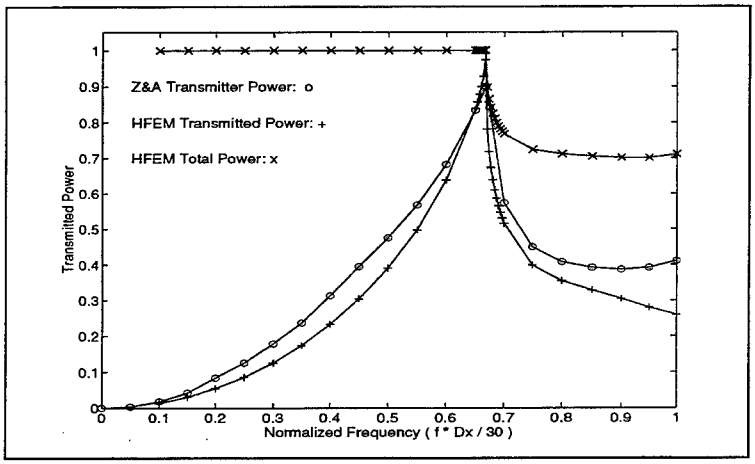


Figure 6: Square Inductive Screen, 30 Degree Incidence, Perpendicular Polarization

Finally, Lee. et. al. calculated the power transmitted through a double layer screen with $a = b = 0.7D_x = 0.7D_z$ and $h = 0.2a$[3]. They used a mode matching technique that applies the Moment Method to a frequency domain integral equation. Figure 7 compares Lee's results with the HFEM algorithm using 404 linear vector expansion functions to model the electric field between the screens. While the single layered inductive screens were positioned halfway between the front and back faces of the finite element structure, the double layered screens were positioned directly on the faces of the model.

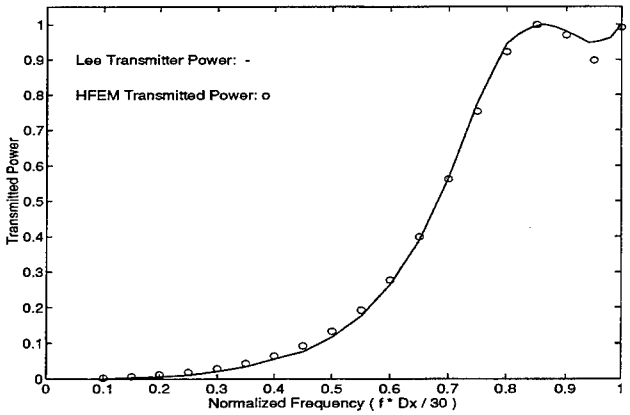


Figure 7: Square Inductive Screen, Double-Layered, Normal Incidence

5 Conclusion

The HFEM algorithm developed and reviewed here calculates the reflection and transmission coefficients of advanced Frequency Selective Surfaces. Although it is based on the assumption of infinite, planar FSS structures, it has demonstrated the ability of a single technique to accurately model complex material parameters and both single and double layered inductive screens with the same results as three other distinct analysis tools.

References

[1] Walker, Michael J. "Frequency Selective Surface Analysis Using the Hybrid Finite Element Method," Ph.D. Dissertation, Air Force Institute of Technology, December, 1995.

[2] Zarillo, G., and K. Aguiar, "Closed Form Low Frequency Solutions for Electromagnetic Waves Through a Frequency Selective Surface," *IEEE Transactions on Antennas and Propagation*, Vol. 35, No. 12, pp. 1406-1417, Dec 1987.

[3] Lee, S. W., G. Zarillo and C. L. Law, "Simple Formulas for Transmission Through Periodic Metal Grids or Plates," *IEEE Transactions on Antennas and Propagation*, Vol. 30, No. 9, pp. 904-910, Sept. 1982.

Generalized Networks for Waveguide Step Discontinuities

Mauro Mongiardo, *Member, IEEE*, Peter Russer, *Fellow, IEEE*, Marco Dionigi and Leopold B. Felsen, *Life Fellow, IEEE*

Abstract— **We have recently introduced an architecture for systematically dealing, in an efficient and rigorous manner, with electromagnetic fields representations and computations in complex structures. The approach is based on the topological partitioning of the complex structure into several subdomains joined together by interfaces. The suggested framework accommodates the use of different analytical/numerical methods (hybridization) when the latter are necessary, the choice of problem-matched alternative Green's functions and the selection of appropriate field quantities at the boundary between different regions.**
Some of these concepts are applied in this paper to the case of a waveguide step discontinuity problem: it is shown that, even for this rather well-investigated example, it is possible to select alternative Green's functions with improved convergence properties with respect to those commonly used. Moreover, a new canonical representation of the step discontinuity is derived and new original formulations of this problem are devised.

I. Introduction

Efficient electromagnetic field computations for complex waveguide components are required in various applications, especially in order to perform computer-aided optimization of the electrical response by suitably adjusting the geometrical parameters. To attack such problems systematically, it is advantageous to parameterize the overall spatial domain in terms of interactions between simpler tractable subdomains. To this end a general architecture has been proposed elsewhere [1], [2], [3], [4], [5] and it is applied here to the waveguide step discontinuity problem.

Step discontinuity problems have received considerable attention in the past (see e.g. [6, chap. 5], [7]). Due to the separability of the wave equation in the waveguide subsections [8], essentially two types of approaches have been developed: one based on mode-matching at the step discontinuity and the other based on an integral equation formulation.

The latter approach has allowed introduction of basis functions which include the edge condition [9], [10] and of the admittance matrix formulation [11], [12]. In these cases, however, the choice of the pertinent Green's function in the waveguide subregions was conventional, corresponding to an eigenfunction expansion in the transverse direction and waves propagating (and reflected) in the longitudinal direction. Accordingly, slowly convergent sums were obtained for steps with significantly different aspect ratios. In this paper we present an alternative Green's function expression which overcomes this problem and allows to use rapidly convergent sums also for fairly high aspect ratios. The theory and numerical results for this case are briefly summarized in §III.

Mode-matching has been previously considered at the step itself, i.e. in a region of zero volume; in this case mode coupling arises at the step discontinuity and one seeks a description of the step discontinuities. It has been found that, although severalalternatives are available, a description of the type employed in [13], [14] is necessary in order to obtain accurate results. In this description the independent field quantities are the electric field in the waveguide with the smaller cross-section and the magnetic field in the waveguide with the larger cross-section. Why this works has not been adequately explained so far; it is, however, readily understood by considering the canonical equivalent network introduced in this paper. Also noted is that, the only rigorous full-wave multi-mode frequency-independent equivalent circuit published for the step discontinuity [13], [14] makes use of controlled sources, while here we introduce a new canonical network based solely on transformers.

Finally, it is also illustrated that other novel approaches for the analysis of the step discontinuity are available. The latter make use of problem-matched Green's functions; in this case, the step discontinuity is partitioned into domains which can be represented by generalized networks which can be described by single term expressions. An example of this approach is also illustrated in §V and a discussion of the numerical effort for the various approaches described in this study is given in the last section.

II. Domain partitioning

We start with subdividing our geometry, i.e. the waveguide step discontinuity, into a number of subdomains (see Fig. 1) which may be of different types, and which are joined together across interfaces. It is apparent that several different topological alternatives are available. For illustration, we have selected the three different choices shown in Fig. 1, namely:

- a) subdivision into two regions of space: admittance formulation;
- b) subdivision into two regions of space with a connection network (mode-matching);
- c) subdivision into three regions of space with a connection network (subregion D).

M. Mongiardo and Marco Dionigi are with the Istituto di Elettronica, Università di Perugia, I-06100 Perugia, Italy.

P. Russer is with Lehrstuhl für Hochfrequenztechnik, Technische Universität München, Arcisstr. 21, D-80333 München 2, Germany.

L. B. Felsen is with the Department of Aerospace and Mechanical Engineering and the Department of Electrical and Computer Engineering, Boston University, 110 Cummington Street, Boston, Ma 02215, U.S.A

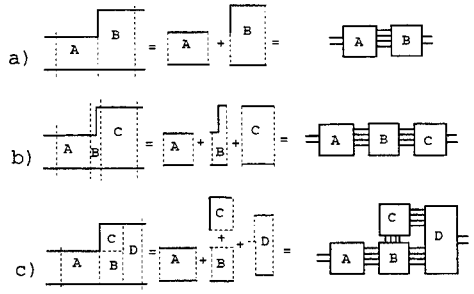


Fig. 1. Three topological alternatives for the step discontinuity segmentation: in a) is represented a subdivision in two regions of space typically used in admittance formulations; in b) is represented a subdivision into three different regions (the central one being of zero volume, i.e. a connection network) as generally used in scattering formulations; finally, in c), a subdivision is shown which provide a different network description. In this latter case region D is also of zero volume, i.e. a connection network.

We shall discuss these alternatives separately in the next three sections.

III. Use of alternative Green's functions

A. Theory

Let us consider the waveguide step discontinuity illustrated in Fig. 2. Essentially, by applying the equivalence theorem, we place on the discontinuity section a p.e.c. with equivalent magnetic currents; we then evaluate the magnetic field generated on both sides and impose the continuity of its tangential components. Typically, a Galerkin discretization procedure is adopted and the modes of the smaller waveguide are chosen as the basis function set. Consequently, most of the numerical effort is devoted to computing the elements of the admittance matrix y_{np}, representing the magnetic field tested by the n-th weighting function as generated by the p-th electric field basis function. Due to the choice of the modes of the smaller waveguide as test and basis functions, the elements pertaining to this waveguide are obtained directly. The computation is, however, less trivial for the elements relative to the larger waveguide. Usually, in this case, an eigenfunction expansion in the y direction is chosen, providing the following representation of the Green's function

$$G^y = \sum_{m=0}^{\infty} g_m(y) g_m(y') \frac{\cos(k_{zm} z_<)\cos(k_{zm}(c-z_>))}{k_{zm}\sin(k_{zm} z_<)}$$
$$g_m(y) = \sqrt{\frac{\epsilon_m}{b_2}}\cos(\frac{m\pi}{b_2} y)$$
$$k_{zm}^2 = k_0^2 - (\frac{\pi}{a})^2 - (\frac{m\pi}{l})^2 \qquad (1)$$

This choice, however, generates the problem of "relative convergence" [15], [16], i.e. the number of terms to be used in the Green's function expansion depends on the ratio b_2/b_1. The larger this aspect ratio, the larger is the number of terms to be considered for the Green's function representation.

The problem of relative convergence can be overcome by

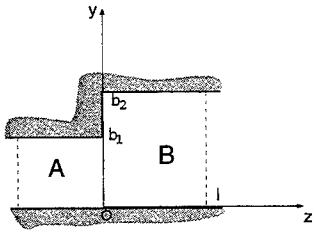


Fig. 2. Geometry of a E-plane step discontinuity between two rectangular waveguides of width a in the x direction. The computational domain lies between the two dotted vertical lines.

considering an alternative Green's function representation which emphasizes wave propagation (and reflection) in the y direction and modal expansion in the z direction. In this case the Green's function takes the form

$$G^z = \sum_{m=0}^{\infty} f_m(z) f_m(z') \frac{\cos(k_{ym} y_<)\cos(k_{ym}(c-y_>))}{k_{ym}\sin(k_{ym} y_<)}$$
$$f_m(z) = \sqrt{\frac{\epsilon_m}{l}}\cos(\frac{m\pi}{l} z)$$
$$k_{ym}^2 = k_0^2 - (\frac{\pi}{a})^2 - (\frac{m\pi}{b_2})^2 \qquad (2)$$

By using these two different Green's function representations in the evaluation of the admittance terms, we get the following expressions with different convergent properties:

$$Y_{n,p}^y = j\omega\epsilon \sum_{m=0}^{\infty} \frac{\sqrt{\epsilon_n \epsilon_p}}{b_1} \frac{\epsilon_m}{b_2} \frac{\cos(k_{zm} l)}{k_{zm}\sin(k_{zm} l)} \frac{(-1)^{n+p} k_m^2}{(k_p^2 - k_m^2)(k_n^2 - k_m^2)} \sin^2(k_m b_1) \qquad (3)$$
$$k_j = (\frac{j\pi}{b_1}) \quad j = p, n$$
$$k_m = (\frac{m\pi}{b_2})$$

$$Y_{n,p}^z = j\omega\epsilon \sum_{m=0}^{\infty} \frac{\sqrt{\epsilon_n \epsilon_p}}{b_1} \frac{\epsilon_m}{l} \frac{k_{ym}(-1)^{n+p}}{(k_n^2 - k_{ym}^2)(k_p^2 - k_{ym}^2)} \frac{\sin(k_{ym} b_1)\sin(k_{ym}(b_2 - b_1))}{\sin(k_{ym} b_2)} + \qquad (4)$$
$$j\omega\epsilon \frac{\sqrt{\epsilon_n \epsilon_p}}{b_1} \delta_{n,p} \frac{b_1}{\epsilon_n} \frac{\cot\left(l\sqrt{k_0^2 - k_p^2}\right)}{\sqrt{k_0^2 - k_p^2}}$$

A.1 Static part extraction

In both expression (3) and (4) we need to evaluate the sums in order to compute the admittance terms. It is well known that, for wide-band evaluation, a different arrangement of these sums is often convenient. In fact, denoting by S one of the above sums evaluated at a certain given frequency, we can write the generic admittance term, Y, evaluated at a different frequency, as given by $Y = S + (D - S)$; here D represents the sum evaluated a the frequency of interest. It is noted that the elements appearing in the sum

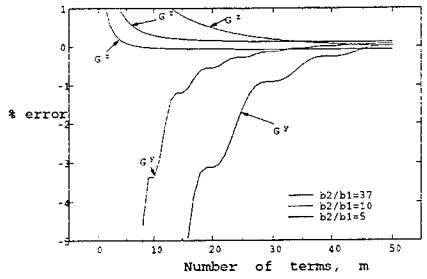


Fig. 3. Convergence behavior of the element $n = 1, p = 1$ of the admittance admittance matrix. The waveguide width is $a = 19mm$ and m is the number of terms considered in the sum in eq. (3) and (4). It is apparent that the usual Green's function, G^y, converges relatively slowly and with a strong dependence on the geometrical ratio b_2/b_1. On the contrary, the alternative Green's function G^z, which emphasize propagation and reflection in the y direction and modal expansion in the z direction, converges rapidly.

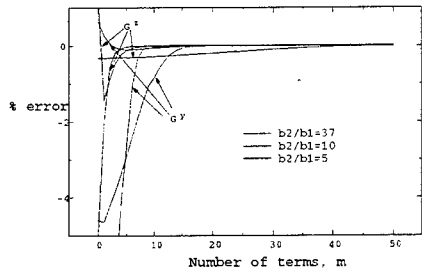


Fig. 4. As in Fig. 3 but with static-part extraction.

$(D - S)$ converge very rapidly. This is a well known technique, generally applied with S representing a static term and D the dynamic contribution. It is noted that this useful device can be applied in the evaluation of both expressions (3) and (4).

B. A numerical example

As an example, in Fig. 3 we show the convergence behavior of one element of the admittance matrix with respect to m, i.e. with respect to the number of terms used to represent the Green's functions in (1) and (2). From the figure it is apparent that a significant advantage is obtained when considering the proposed alternative representation instead of the usual Green's function expression.

Similarly, in Fig. 4 we illustrate the convergence behavior for the same case, but including the static part extraction. Clearly, this accelerates the rate of convergence. Thus, using this device and the appropriate alternative Green's function selection, convergence is achieved with just a few terms.

IV. The step discontinuity as a connection network (Mode-matching)

This approach is based on the field representation problem arising at the step discontinuity [17]. In order to investigate this problem it is convenient to refer to the bifurcation shown in Fig. 5 where three different subdomains are joined together. In particular, there is an interface which connects subdomain 1 to subdomain 3, and an interface connecting subdomain 2 to subdomain 3. In the following, for brevity, we assume that the electric (magnetic) fields at the interfaces are expanded in terms of suitable basis functions and we call by $\mathbf{V_i}$ ($\mathbf{I_i}$) the vector containing the electric (magnetic) field expansion coefficients relative to region i.

It has been shown elsewhere [18] that the connection network for this interface can be obtained by taking $\mathbf{V_1}, \mathbf{V_2}$ and $\mathbf{I_3}$ as independent variables leading to the canonical network representation in Fig. 6. The other choice of independent variables is $\mathbf{I_1}, \mathbf{I_2}$ and $\mathbf{V_3}$ which leads to the canonical network shown in Fig. 7. Both representations are equally valid in order to describe the connection network relative to a bifurcation.

However, in the case of the step discontinuity, region 1 is filled by a p.e.c., represented by a short-circuit. Thus we need to impose the condition $\mathbf{V_1} = \mathbf{0}$. The equivalent network is now the one in Fig. 6 with the ports pertaining to region 1 short-circuited.

It is useful to note that the above canonical network is frequency independent, satisfies the Tellegen theorem and admits a scattering representation with the following properties: symmetry, $\mathbf{S^T} = \mathbf{S}$, orthogonality, $\mathbf{S^T S} = \mathbf{I}$ and unitary, i.e. $\mathbf{S S^\dagger} = \mathbf{I}$, where the $\dagger$ denotes the hermitian conjugate matrix, T denotes the transposed and $\mathbf{I}$ is the identity matrix.

Also note that the above discussion is valid in general, for any choice of basis functions in regions 2 and 3. In practice, the most common choice of basis functions is the use of the modal eigenfunctions at both sides of the discontinuity; moreover it is common to place the reference planes at a certain distance from the discontinuity itself. We have therefore a certain number of modes which propagate from the discontinuity itself to the reference planes and are represented by transmission lines; by contrast the modes well below cut-off provide a localized contribution only at the discontinuity itself and can be represented by lumped, frequency-dependent reactances. It is also noted that the model proposed in this contribution, similarly to the model proposed in [13], [14] can be easily implemented in standard circuit simulators.

V. A different segmentation

The advantage of the segmentation illustrated in Fig. 1c is that when we use as basis functions the modes for each of the regions, the network terms are simple, single-term expressions [19]. With reference to the geometry of the rectangular resonator shown in Fig. 8 and the corresponding notation, the admittance matrix element y^{if}_{mn} relating the magnetic field amplitude of mode m at side i of the

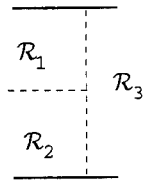


Fig. 5. The bifurcation problem: three regions of space connected at an interface.

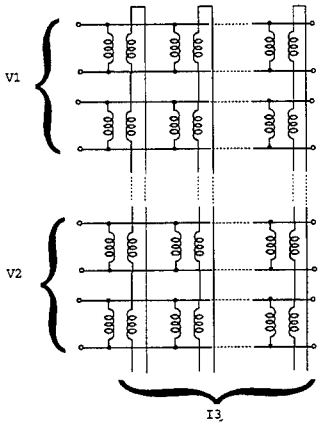


Fig. 6. A canonical network for the bifurcation: with $\mathbf{V}_1, \mathbf{V}_2$ and $\mathbf{I}_3$ chosen as independent field quantities.

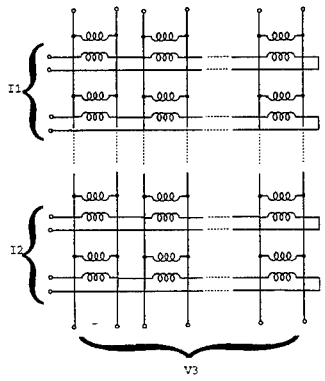


Fig. 7. A canonical network for the bifurcation: with $\mathbf{I}_1, \mathbf{I}_2$ and $\mathbf{V}_3$ chosen as independent field quantities.

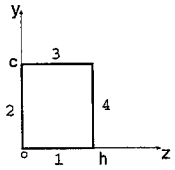


Fig. 8. Geometry of the rectangular resonator for the E-plane step segmentation.

rectangle ($i = 1, 2, \ldots, 4$) to the electric field amplitude of mode n at side ℓ ($\ell = 1, 2, \ldots, 4$) is given by the following expressions:

$$
\begin{aligned}
y_{mn}^{11} &= j\frac{\left(\frac{\pi}{a}\right)^2 - k^2}{\beta_z \omega \mu} \cot\left(\beta_z c\right) \delta_{mn} \\
y_{mn}^{12} &= j\omega\varepsilon\sqrt{\frac{\epsilon_m \epsilon_n}{ac}} \frac{\left(k^2 - \frac{\pi}{a}\right)^2}{k^2\left(k^2 - \bar{k}^2\right)} \\
y_{mn}^{13} &= -j\frac{\left(\frac{\pi}{a}\right)^2 - k^2}{\beta_z \omega \mu} \csc\left(\beta_z c\right) \delta_{mn} \\
y_{mn}^{14} &= (-1)^m y_{mn}^{12}
\end{aligned}
$$

where $\bar{k}^2 = \left(\frac{\pi}{a}\right)^2 + \left(\frac{n\pi}{h}\right)^2$ and $\beta_z = \sqrt{k^2 - \bar{k}^2}$, while δ_{mn} is the Kronecker delta.

Therefore, in this case, the network computation is trivial and the problem is completely reduced to that of the interconnection of different networks. Again, it is worthwhile to note that this problem can be advantageously solved by using standard circuit simulators. The actual advantage in using this method is closely related to the type of geometry under consideration.

VI. Results and Discussion

The topological alternatives described in the previous sections provide different generalized networks and some new physical insights for the step discontinuity problem. The techniques applied are well-suited to step discontinuity characterization for particular geometric aspect ratios. In §III-B we have already noticed that the alternative Green's function representation, which makes use of wave propagation along y, is particularly well convenient for the case of pronounced steps.

Naturally, the different approaches yield the same numerical result. As an example we have analyzed a step discontinuity via the three different topological segmentations and plotted the relative results in Fig. 9 for different aspect ratios.

The numerical expenditure associated with the various approaches is, on the other hand, quite different and a complete numerical analysis of the peculiar advantages and disadvantages is beyond the scope of this presentation. Depending on the problem geometry one approach can be more effective than the others. Moreover, although we have focused here on the example of a single step discontinuity,

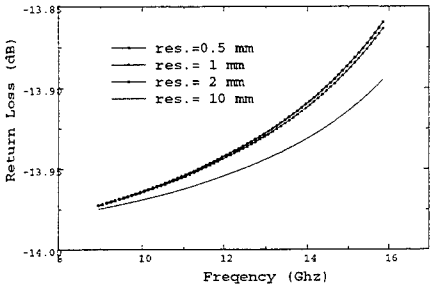


Fig. 9. A comparison of the electrical response of the step discontinuity for different values of spatial resolutions; the latter is defined as the height of the smaller waveguide divided by the number of basis functions used in order to represent the field on this domain. All three aproaches yield the same result.

the application to more complex geometries, as often required in practice, may significantly take advantage of the enlarged arsenal of options introduced.

VII. Conclusions

We have introduced elsewhere an architecture for field representation, computation and hybridization when dealing with complex problems. By applying some of the concepts developed for this general architecture we have found some new results also for fairly well-known problem such as the step discontinuity.

In particular, we have found that the systematic use of alternative Green's functions can significantly improve the convergence properties of modal sums. We have also found a canonical network representation for the step discontinuity and a different topological partitioning. Each of the proposed approaches has peculiar advantages and disadvantages that, depending on the considered geometry and the particular feature under investigation, suggest their use for various scenarios.

References

[1] L. B. Felsen, "An architecture for wave interaction with complex environments: Fields, networks, interface mappings and computations." *presented at the First Annual Symposium of Radio Science in Israel, Tel Aviv, Israel*, Dec. 1996.

[2] L. B. Felsen, M. Mongiardo, and P. Russer, "Electromagnetic field computations by a generalized network formulation," *ACES, Monterey, CA*, Mar. 1997.

[3] L. B. Felsen, M. Mongiardo, P. Russer, G. Conti, and C. Tomassoni, "Waveguide component analysis by a generalized network approach," *European Microwave Conference, Jerusalem, Israel*, Sept. 1997.

[4] M. Dionigi, M. Mongiardo, P. Russer, and L. B. Felsen, "Problem-matched Green's functions for generalized network formulation of complex waveguides," *ICEAA International Symposium, Turin*, Sept. 1997.

[5] L. B. Felsen, M. Mongiardo, and P. Russer, "Hybrid computations of electromagnetic fields via a generalized network formulation." *PIERS, Progress in Electromagnetic Research Symposium, Boston*, p. 359, July 1997.

[6] T. Rozzi and M. Mongiardo, *Open Electromagnetic Waveguides*. London: IEE, 1997.

[7] R. E. Collin, *Field Theory of Guided Waves*. New York: IEEE Press. 1991.

[8] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves*. Englewood Cliffs, NJ: Prentice Hall, 1973. Classic reissue by IEEE Press, 1994.

[9] T. Rozzi, A. Morini, F. Ragusini, and M. Mongiardo, "Direct analytical solution of iris discontinuities in waveguide by separation of variables," *accepted IEEE Trans. Microwave Theory Tech.*, Dec. 1996.

[10] T. Rozzi and M. Mongiardo, "E-plane steps in rectangular waveguide," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 1279–1288, Aug. 1991.

[11] F. Alessandri, M. Mongiardo, and R. Sorrentino, "Computer-aided design of beam forming networks for modern satellite antennas," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 1117–1127, June 1992.

[12] F. Alessandri, G. Bartolucci, and R. Sorrentino, "Admittance-matrix formulation of waveguide discontinuity problems," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 394–403, Feb. 1988.

[13] A. Weisshar, M. Mongiardo, and V. K. Tripathi, "CAD oriented equivalent circuit modeling of step discontinuities in rectangular waveguides," *IEEE Microwave and Guided Wave Letters*, vol. 6, pp. 171–173, Apr. 1996.

[14] A. Weisshar, M. Mongiardo, A. Tripathi, and V. K. Tripathi, "CAD oriented equivalent circuit models for rigorous full-wave analysis and design of waveguide components and circuits," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 2564–2570, Dec. 1996.

[15] R. Mittra, T. Itoh, and T. S. Li, "Analytical and numerical studies of the relative convergence phenomenon arising in the solution of an integral equation by the moment method," *IEEE Trans. Microwave Theory Tech.*, vol. 20, pp. 96–104, July 1972.

[16] R. Sorrentino, M. Mongiardo, F. Alessandri, and G. Schiavon, "An investigation on the numerical properties of the mode-matching technique," *Int. J. Numer. Model.*, vol. 4, pp. 19–43, 1991.

[17] R. Schmidt and P. Russer, "Modeling of cascaded coplanar waveguide discontinuities by the mode–matching approach," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 2910–2917, Dec. 1995.

[18] L. B. Felsen, M. Mongiardo, and P. Russer, "Electromagnetic field representations and computations in complex structures: the connection network," *manuscript in preparation*, Dec. 1997.

[19] F. Alessandri, M. Mongiardo, and R. Sorrentino, "Rigorous mode matching analysis of mitered E-plane bends in rectangular waveguide," *IEEE Microwave Guided Wave Lett.*, vol. 4, pp. 408–410, Dec. 1994.

SESSION 17:

SIGNAL PROCESSING TECHNIQUES IN CEM

Chairs: D.H. Werner and P.L. Werner

USING WINDOWED, ADAPTIVE SAMPLING TO MINIMIZE THE NUMBER OF FIELD VALUES NEEDED TO ESTIMATE RADIATION AND SCATTERING PATTERNS

Edmund K. Miller
3225 Calle Celestial, Santa Fe, NM 87501-9613
505-820-7371, emiller@esa.lanl.gov

1.0 INTRODUCTION

Determining radiation or scattering patterns is usually accomplished by integrating the source distribution on an object or the tangential fields these sources produce over some enclosing surface, or from summing a modal expansion on such a surface. For objects not too large relative to the wavelength, these needed far-field evaluations do not require an excessive amount of computation. However, an object's surface area in square wavelengths, A, increases with frequency, f, as f^2, and past some threshold the number of far-field observation angles needed to adequately define the radiation pattern may come to drive the total computation cost. For antenna problems modeled using a physical-optics current approximation, for example, only one current computation is needed whose cost is proportional to f^2, which is also the cost of each far-field sample. For scattering problems solved using so-called "fast" techniques whose current-computation cost ranges between f^2 and f^4 per incidence angle, the total cost to obtain the aspect-dependent RCS can be directly proportional to the number of incidence angles required. Thus, as problem size increases, the overall computation cost eventually may become proportional to the number of observation angles at which the radiation or scattering pattern is needed.

Reliably determining the far-field pattern for large objects, i.e., not missing important details of the pattern with respect to lobe maxima and null locations, can require 10 or more samples per lobe if only simple interpolation is used to approximate the pattern behavior between the samples. It would be useful were an alternate procedure available whereby the number of samples is reduced to some minimum determined by the pattern features themselves and which provides a continuous estimate of the pattern based on electromagnetic physics. One such procedure for accomplishing this objective is discussed here, based on a general technique called Model-Based Parameter Estimation (MBPE) [Miller].

MBPE involves a model, preferably physically based and called here a fitting model (FM), whose coefficients, or parameters, are estimated by matching the FM to samples of the process or data to be estimated, here called the generating model (GM). In electromagnetics, two prominent FMs having wide applicability, and related by the Laplace transform, are exponential series and pole series. The former is most obviously related to transient waveforms and the latter to frequency spectra, for which reason we refer to exponential-series as being waveform-domain FMs and pole-series as spectral-domain FMs.

Note that MBPE encompasses as special cases some classical numerical procedures. For example, Prony's Method was the first developed for handling waveform or transient responses and Pade Approximation can be recognized as being applicable to frequency or spectral responses.

2.0 BACKGROUND

The possibility of using an MBPE procedure for constructing far-field radiation patterns has been explored elsewhere [Bucci et al. (1991), Roberts and McNamara (1994)]. Bucci et al. developed a signal-processing-like procedure based on the spatial bandwidth of the field to establish the minimum number of pattern samples need to develop a radiation pattern. Roberts and McNamara applied Prony's method to angle windows to develop a radiation-pattern estimate from a sequence of discrete-source approximations (DSAs). We extend these basic ideas here by developing an adaptive procedure for far-field pattern estimation called WASPE (Windowed Adaptive Sampling Pattern Estimation),

applicable to both radiation and scattering problems, that permits the pattern to be reconstructed to a prescribed uncertainty.

WASPE begins with sparsely sampled far field values from a GM at a pre-specified set of observation angles which are used to set up an initial sequence of FMs, each of which shares two or more GM samples in their region of overlap as is illustrated conceptually in Fig. 1. Each new GM-sample angle is then selected where the maximum mismatch of all pairs of overlapping pairs of FMs occurs as determined by computing more finely sampled FM values and then increasing the order of the FMs whose angle windows include that sample. The process of comparing FM values continues until the maximum mismatch falls below some prescribed uncertainty. The approach is essentially equivalent to the adaptive method developed by Miller (1996) for estimating a frequency transfer function, but for which the FM was a pole series rather than the exponential series used here.

The performance of WASPE is dependent, among other things, on the kind of FM used for its implementation. One approach is provided by Prony's Method, where an appropriate FM is given by

$$f(\theta) = \sum_{\alpha=1}^{W} R_\alpha \exp[ikd_\alpha \cos(\theta)] \tag{1}$$

with the W point sources of strengths R_α and located at positions d_α along the x-axis from which the observation angle θ is measured. By sampling the GM far field in uniform steps of θ, Prony's Method can be used to estimate R_α and d_α and thereby obtain a FM that provides a continuous estimate of the pattern between the GM samples [Miller and Lager (1978), Miller and Goodman (1983a), Miller and Lager (1983b)].

Strictly speaking, when used in this fashion Prony's Method is only applicable to linear arrays. Furthermore, if the number of actual sources, S, and/or their effective aperture size in wavelengths, A, is too large, the matrix needing solution can become very ill-conditioned. Thus, it's logical to window the pattern so that its effective rank is maintained below some threshold. This means that in contrast to the situation where an actual source distribution can be "imaged" from its complete pattern, when using windowed data an equivalent DSA valid over only the window used for its computation will result. Since the goal for the our application is not to image the source distribution but rather to compute the pattern more efficiently, not having the actual distribution is not a disadvantage. Also note that were this done for a scattering pattern, whose source distribution depends on the angle of incidence, no single DSA can describe the backscatter pattern in any case. An alternative to Prony's Method is to instead assign the DSA source locations and use the GM samples to compute the source amplitudes. This has the advantage that the GM samples may be arbitrarily located in angle, thereby making an adaptive procedure more practical. Also, only half the samples are needed for a given value of W since the source locations are no longer unknowns.

Using this alternative requires some rationale' for assigning the source locations, or equivalently, specifying an appropriate FM. Since the highest-angular-frequency component of the pattern is proportional to kA, one possibility is to use as the FM over angle window m

$$f(\theta) = \sum_{n=S}^{N} R_{mn} \exp[ikn\cos(\theta)] + \sum_{n=-S'}^{-N} R'_{mn} \exp[ikn\cos(\theta)] \tag{2}$$

where N = int(A + 1) and 2N - S - S' = F with F the total number of exponential terms used for the m'th FM, int(X) denotes the value of X rounded off to the nearest lower integer and $|S - S'| \leq 1$. Also needed to "initialize" the FM selection is choosing a starting value for F, and the number of FMs, M. Somewhat arbitrarily, we choose a small value for F of 3 or 4, with M determined by the number of anticipated lobes in the pattern being processed and the amount of overlap chosen for the adjacent FMs.

A candidate pattern on which to test the WASPE is the back-scattered field of a finite-length circular cylinder for which and angle-dependent approximation is given by [Knott et al. (1993)]

$$E(\theta) \propto \cos(\theta_i)\frac{\sin[2\pi A\sin(\theta_i)]}{2\pi A\sin(\theta_i)} \tag{3}$$

and which is plotted in Fig. 1 for A = 10 over an incidence-angle range of 0 to $\pi/2$. If each FM is intended on average to include two lobes of the scattering pattern, then ~A = 10 angular windows would be required to cover this range of incidence angles. Using a total overlap between adjacent FMs of 2/3, then M ~ 16 FMs would be needed. As each new GM sample is added, the FMs that include it are increased in rank by alternately increasing S and S' by one, until a maximum mismatch, MM, between all pairs of overlapping FMs of

$$MM_{i,j}(\theta) = \max[|FM_i(\theta) - FM_j(\theta)|/(|FM_i(\theta)| + |FM_j(\theta)|)] \le 10^{-X} \tag{4}$$

is achieved, i.e., an X-digit match is obtained between them. It's also possible to scale X, e.g. so that it depends on the relative magnitude of the FMs being tested relative to the maximum value in the pattern, to thereby allocate the final uncertainty in the estimated pattern according to problem requirements. Note that the FM itself may represent the complex far field or its magnitude.

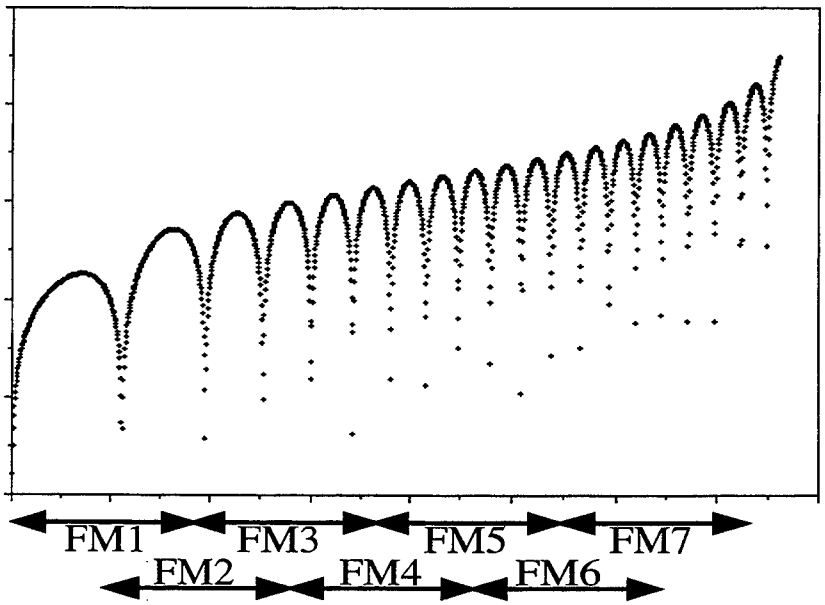


Figure 1. Conceptual illustration of one possible sequence of overlapping, windowed FMs to model a radiation or scattering pattern. The maximum mismatch between all of the pairs of FMs is used to determine where the next GM pattern sample is placed, with the process continuing until a prescribed uncertainty in the estimated pattern is achieved.

3.0 NUMERICAL EXAMPLES

An example using WASPE for a 3-wavelength sinusoidal filamentary current is shown in Fig. 2. The number of FMs

used here is 5, arranged as shown by the horizontal lines at the top of the figure. The 12 GM samples initially used and the 6 samples added during the adaptation process satisfied a maximum mismatch criterion of 0.01. Both the final FM estimates are plotted here as well as the finely sampled GM pattern, and are nearly indistinguishable.

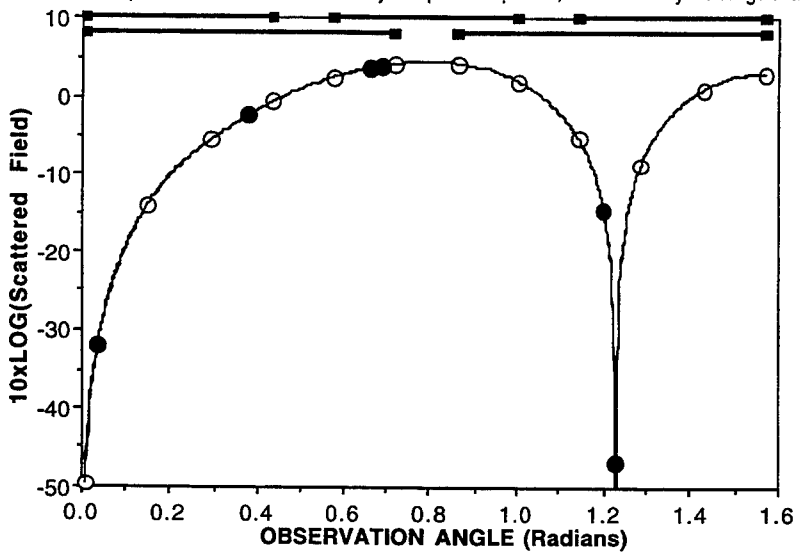


Figure 2. Use of WASPE for the radiation pattern of a 3-wavelength sinusoidal current filament is illustrated here, with the horizontal lines on top showing the angle ranges spanned by the five FMs used. The initial GM samples are shown by open circles with the solid circles showing where the adaptively added samples are located, using a maximum estimation uncertainty of 0.01.

The result of using WASPE on the pattern of Fig. 1 is shown in Fig. 3. The 16 overlapping FMs used here superimpose to graphical resolution and agree to within two digits of accuracy independent of the actual field level. The converged pattern estimate employs a total of 56 GM samples, or about 2.8 per pattern lobe. The distribution of GM samples for the cylinder problem is shown in Fig. 4. The initial 32 samples are distributed uniformly in angle, while the next 24 are added as determined by the maximum mismatch between the overlapping FMs. It can be seen that the adaptively added samples are concentrated towards broadside incidence, a result that's consistent with the closer spacing of scattered-field maxima in that region. Whereas the example presented in Fig. 2 employed FMs overlapping to both ends of the observation-angle range, that is not the case here, which explains why no additional GM samples are located at either end.

4.0 DISCUSSION

Note that two different approaches to estimating a far-field pattern using MBPE have been mentioned above, both leading to DSAs. The first attempts to model the actual source distribution, i.e., determine the source locations, using Prony's Method. The second, and the one from which the numerical results presented here were derived, instead uses angle windows and models the field, rather than the source distribution. If we wanted to solve an inverse problem, the first would be appropriate, but if needing only to estimate the pattern, the second is the better choice.

Second, radiation and scattering applications present intrinsically different problems. A radiation pattern is determined from one source distribution, whereas the backscatter radar cross section comes from a new source distribution for each incidence angle. Thus, if needing to estimate a scattering pattern, it seems preferable to model only the pattern and to avoid the need to estimate source locations as well.

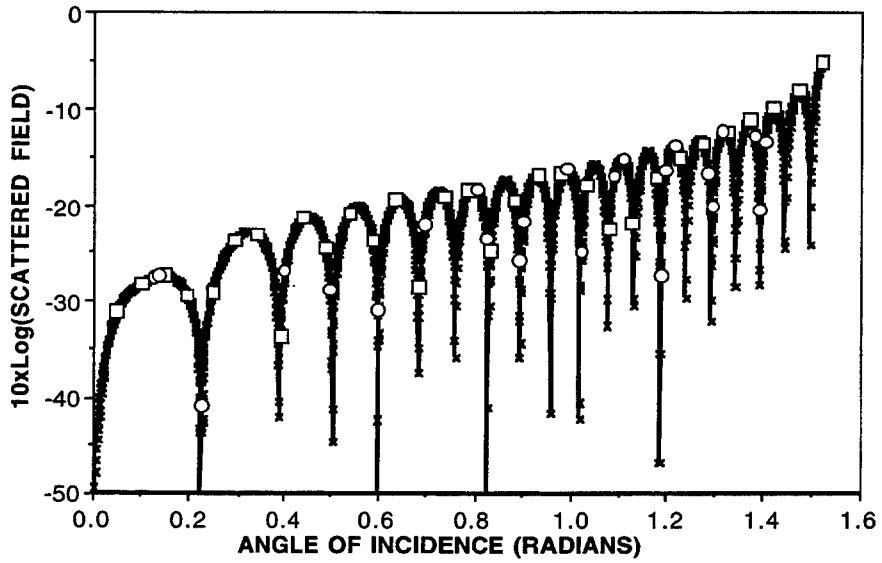


Figure 3. Example of using WASPE for estimating the back-scatter pattern of a circular cylinder 10-wavelengths long. The 32 initial GM samples are shown by the squares while the 24 adaptively added samples are shown by the circles. The prescribed estimation uncertainty used here was 0.01.

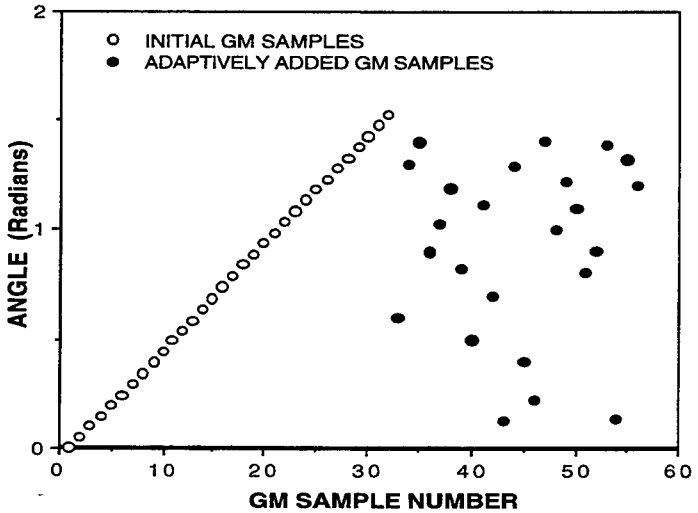


Figure 4. Location in angle of the GM pattern samples is uniform for the set used for the initial FM computation, but thereafter varies widely according to the maximum mismatch error found between overlapping FMs. It's interesting to observe that the number of added GM samples increases from end-on to broadside incidence angles, due to the finer lobe structure that occurs near broadside a can be seen in Fig. 3.

5.0 CONCLUDING COMMENTS

The feasibility of minimizing the number of samples needed to reconstruct a radiation or scattering pattern using Windowed Adaptive Sampling Pattern Estimation (WASPE) based on model-based parameter estimation has been demonstrated. While useful for radiation-pattern analysis, WASPE should offer even more to improving the efficiency of computing the scattering patterns of large objects.

Although promising, several possibilities for improving the performance of WASPE. One is to v window width with the anticipated lobe spacing or the number of terms used in the initial FMs. A maintain the rank of a FM below some level and to split that FM into two if more samples than th its window. A third would be to vary the mismatch specification with the pattern level or other that sampling is driven by problem requirements. Finally, use of other FMs should be explored.

1-34 29 30 00
Rm 417
Holiday Inn

5.0 REFERENCES

O. M. Bucci, C. Gennarelli and C. Savarese (1991), "Optimal Interpolation of Radiated Fields C *Trans. Antennas and Prop.*, **39**, pp. 1633-1643.

E. F. Knott, J. F. Shaeffer and M. T. Tuley (1993), *RADAR Cross Section*, 2nd Edition,Artech

E. K. Miller and D. L. Lager (1978), "Radiation Field Analysis and Synthesis Using Prony's Method", *Electronics Letters*, **14**, pp. 180-182.

E. K. Miller and D. M. Goodman (1983a), "A Pole-Zero Modeling Approach to Linear-Array Synthesis. Part I: The Unconstrained Solution", *Radio Science*, **18**, pp. 57-69.
5
E. K. Miller and D. L. Lager (1983b), "Imaging of Linear Source Distributions", *Electromagnetics*, **3**, pp. 21-40.

E. K. Miller (1996), "Minimizing the Number of Frequency Samples Needed to Represent a Transfer Function Using Adaptive Sampling," in *12th Annual Review of Progress in Applied Computational Electromagnetics*, Naval Postgraduate School, Monterey, CA, pp. 1132-1139.

R. Roberts and D. A. McNamara (1994), "Interpolating Radiation Patterns Using Prony's Method, *IEEE/South Africa Institute of Electrical Engineers AP/MTTS Proceedings.*

SPECTRAL DOMAIN INTERPOLATION OF ANTENNA RADIATION PATTERNS USING MODEL-BASED PARAMETER ESTIMATION AND GENETIC ALGORITHMS

R.J. Allard, D.H. Werner and J.S. Zmyslo
The Pennsylvania State University
Applied Research Laboratory
P.O. Box 30
State College, PA 16804

P.L. Werner
The Pennsylvania State University
College of Engineering
DuBois, PA 15801

Abstract: Model-Based Parameter Estimation (MBPE) techniques that use a rational function fitting model with complex coefficients have been shown to yield approximations which accurately describe the frequency dependence of antenna characteristics such as input impedance or admittance. In this paper we apply MPBE to the radiated electric field of a 0.5m dipole and use the rational function fitting model to interpolate the electric field spectrum of this antenna over the interval 150 to 950 MHz. A method is presented for using these electric field interpolations in conjunction with the associated impedance interpolations in order to reconstruct gain patterns at any frequency within this model range. Conventional Pade approximations require a matrix inversion operation in order to estimate the parameters of the rational function fitting model. In this paper we introduce a Genetic Algorithm (GA) approach for obtaining the required fitting model parameters. This new approach to MBPE has the advantages that it is very general and avoids the need for any kind of matrix inversion.

1. Introduction

The High Fidelity Analysis Module (HFAM) currently undergoing development at the Applied Research Laboratory, The Pennsylvania State University (ARL/PSU), will provide, for the first time, a highly-capable and flexible analysis tool for the modeling and analysis of composite (antenna, platform, terrain) electromagnetic radiation patterns. These models and patterns will potentially be of great value to a wide range of users with different data fidelity (space and frequency) requirements. For instance, the user involved in military Test and Evaluation (T&E) typically requires high-fidelity representations of antenna/platform/terrain interactions in order to simulate or assess the performance of a system or systems in various geometric configurations and at various frequencies. This particular military user may also be interested in evaluating different antenna/platform or antenna/terrain configurations at various frequencies.

The storage of high-fidelity antenna radiation pattern data imposes significant burdens upon storage media infrastructure. This burden is felt most heavily by the military forward-deployed tactical user/analyst who is typically seriously constrained by on-site computer storage resources. To illustrate the fundamental problem, one high-fidelity antenna radiation pattern sampled and stored (without compression) in one-degree increments of azimuth and elevation requires approximately 64,000 data points. Many radiation patterns are more often than not required to support the myriad of special cases required to fully support mission needs. Experience has shown that these data storage requirements will likely increase with time. Clearly a scheme whereby high-fidelity radiation pattern data can be stored, compressed and then regenerated at differing levels of fidelity in space and frequency (with a mechanism for explicitly describing fidelity/error tradeoffs) is highly desirable. This paper addresses a subset of these requirements by providing a means for reducing the amount of data required to represent an antenna's performance over its operational frequency range thereby realizing significant reductions in data storage requirements. Additional work relating to data compression and error metrics is ongoing and will be discussed in a future paper.

2. Theory

Model-Based Parameter Estimation (MBPE) is a form of "smart" curve fitting because it uses a fitting model which is based on the problem physics [1-3] as opposed to standard curve fitting techniques, which do not make use of the problem physics and consequently tend to be much less efficient. The "model-

based" part of MBPE involves using low-order analytical formulas as fitting models, while the "parameter estimation" part refers to the process of numerically obtaining coefficients for the fitting model by matching it or fitting it to sampled values (either calculated or measured). One form of a fitting model which is commonly employed in MBPE is represented by the following rational function [1-3]:

$$F(s)=\frac{N(s)}{D(s)}=\frac{N_0+N_1s+N_2s^2+\cdots+N_ns^n}{D_0+D_1s+D_2s^2+\cdots+D_{d-1}s^{d-1}+s^d} \tag{1}$$

The standard approach to solving for the n+d+1 unknown complex coefficients in (1) is to first sample the data set which is being interpolated at n+d+1 frequencies [1-3]. The n+d+1 equations which result may then be written in matrix form and subsequently inverted to solve for the required coefficients [2]. In this case we have a square matrix and only as many fitting points as there are unknown coefficients are required to find the solution, however this procedure can easily be extended to perform interpolation on an over-sampled data set via a least-squares approach. When applying this technique to the interpolation of antenna radiation patterns, the function $F(s)$ represents the complex far-zone radiated electric field at a particular value of θ and φ, and the argument s represents the complex frequency $j\omega$ at which the antenna is operated. Hence, this technique may be applied repeatedly over different values of θ and φ in order to obtain an approximation for the antenna radiation pattern at any frequency within the predetermined range of the fitting model.

The gain of an antenna may be expressed in the form [4]

$$G(\theta,\varphi)=\frac{4\pi U(\theta,\varphi)}{P_{in}} \tag{2}$$

where

$$P_{in}=\frac{1}{2}\mathrm{Re}\left\{V_{in}I_{in}^*\right\}=\frac{1}{2}\left|I_{in}\right|^2R_{in} \quad \text{and} \quad U(\theta,\varphi)=\frac{r^2}{2\eta_0}\left(\left|E_\theta\right|^2+\left|E_\varphi\right|^2\right) \tag{3}$$

which represent the input power accepted by the antenna and the antenna radiation intensity, respectively. A technique for interpolating the input impedance of an antenna via Pade approximations has been demonstrated previously by Miller [1-3]. A similar technique may be employed using Pade approximations to estimate the input impedance $Z_{in}=R_{in}+jX_{in}$ required in order to calculate the input power P_{in} of a particular antenna as a function of frequency.

GAs have been used for solving a wide variety of engineering electromagnetics problems [5,6]. Here we demonstrate a GA approach to MBPE which is very general and eliminates the need for any type of matrix inversion operation. A GA optimization procedure using real value encoding has been adapted for use in this application [7]. In this case, the objective function $f(P)$ to be maximized is

$$f(P)=-\sum_{i=1}^{N_f}\left|F(s_i)-\overline{F}_P(s_i)\right|^2 \tag{4}$$

where $F(s_i)$ represents the sample values for electric field versus frequency as determined by the Numerical Electromagnetics Code (NEC), $\overline{F}_P(s_i)$ is the approximated value of the field determined by the evolutionary process of the GA, N_f is equal to the total number of frequency fitting points, and P is the population index number or the member of the current population to which the fitness $f(P)$ is to be assigned. The parameters which comprise the chromosomes of the GA are the coefficients $[N_0,\cdots,N_n]$ and $[D_0,\cdots,D_{d-1}]$ of equation (1), and so $\overline{F}_P(s_i)$ is simply the rational function (1) with the approximated

chromosomes as its coefficients. When $f(P)$ is maximized (in this case that means as close to zero as possible) then we say that the GA has found an optimal fit to the data set $F(s_i)$ for $i = 1, 2, \cdots, N_f$.

3. Results

MBPE using matrix inversion is applied to a radiation pattern slice $(0^\circ \le \theta \le 180^\circ)$ of a 0.5m dipole test case over a frequency interval from 150 to 950 MHz. Figure 1 shows a sequence of frequency spectra which have been interpolated via MBPE using only six fitting frequencies for each specified value of the angle θ (i.e., θ=10°, 30°, 50°, 70° and 90°). It is clear from these plots that this method for interpolation is quite powerful and can be used to achieve significant model-order reduction, which is especially attractive for problems with large computational domains. By interpolating the frequency response at a reasonable resolution (2° in this case) and storing the six coefficients for each point, a slice of the radiation pattern at any frequency between 150 and 950 MHz can be reproduced quickly and accurately. The input impedance interpolations shown in Figure 2 can then be used to convert the interpolated electric field patterns into gain patterns using (2) and (3). Examples of these gain plots at four different frequencies are shown in Figures 3 and 4.

The GA form of MBPE is applied over a shorter frequency interval (150-450 MHz) which requires only three fitting frequencies. In this case there are two unknown numerator coefficients and one unknown denominator coefficient. Figure 5 shows the matrix-inverted interpolation with the fitting frequencies denoted by circles, while Figure 6 shows the GA interpolation which used the same three frequencies (250, 300, and 350 MHz). The two interpolations are virtually identical and the comparison demonstrates the usefulness of GAs in this type of application.

4. Conclusions

A MBPE technique has been introduced in this paper which can be used to accurately estimate not only the input impedance, but also the radiation pattern of an antenna at any frequency within the operational range of the fitting-model. This technique has the advantage of providing a method for efficiently generating and storing model parameters for problems with large computational domains. A genetic algorithm technique for performing MBPE has also been introduced which produces fits that are virtually identical to the matrix inversion technique.

Acknowledgments

The authors gratefully acknowledge the assistance provided by Raj Mittra and Ed Miller. This work was supported by the Naval Information Warfare Activity (NIWA).

References

[1] E.K. Miller, "Model-Based Parameter Estimation in Electromagnetics: I—Background and Theoretical Development," *Applied Computational Electromagnetics Society Newsletter,* Vol. 10, No. 3, pp. 40-63, 1995.

[2] E.K. Miller, "Model-Based Parameter Estimation in Electromagnetics: II—Applications to EM Observables," *Applied Computational Electromagnetics Society Newsletter,* Vol. 11, No. 1, pp. 35-56, 1996.

[3] E.K. Miller, "Model-Based Parameter Estimation in Electromagnetics: III---Applications to EM Integral Equations," *Applied Computational Electromagnetics Society Journal,* Vol. 10, No. 3, pp. 9-29, 1995.

[4] W.L. Stutzman and G.A. Thiele, *Antenna Theory and Design,* 598 pp., John Wiley & Sons, Inc., New York, 1981.

[5] J.M. Johnson and Y. Rahmat-Samii, "Genetic Algorithms in Engineering Electromagnetics," *IEEE Antennas and Propagation Magazine,* Vol. 39, No. 4, pp. 7-25, 1997.

[6] D.S. Weile and E. Michielssen, "Genetic Algorithm Optimization Applied to Electromagnetics: A Review," *IEEE Transactions on Antennas and Propagation,* Vol. 45, No. 3, pp. 343-353, 1997.

[7] M. Gulsen, A.E. Smith and D.M. Tate, "A Genetic Algorithm Approach to Curve Fitting," *Int. J. Prod. Res.,* Vol. 33, No.7, pp. 1911-1923, 1995.

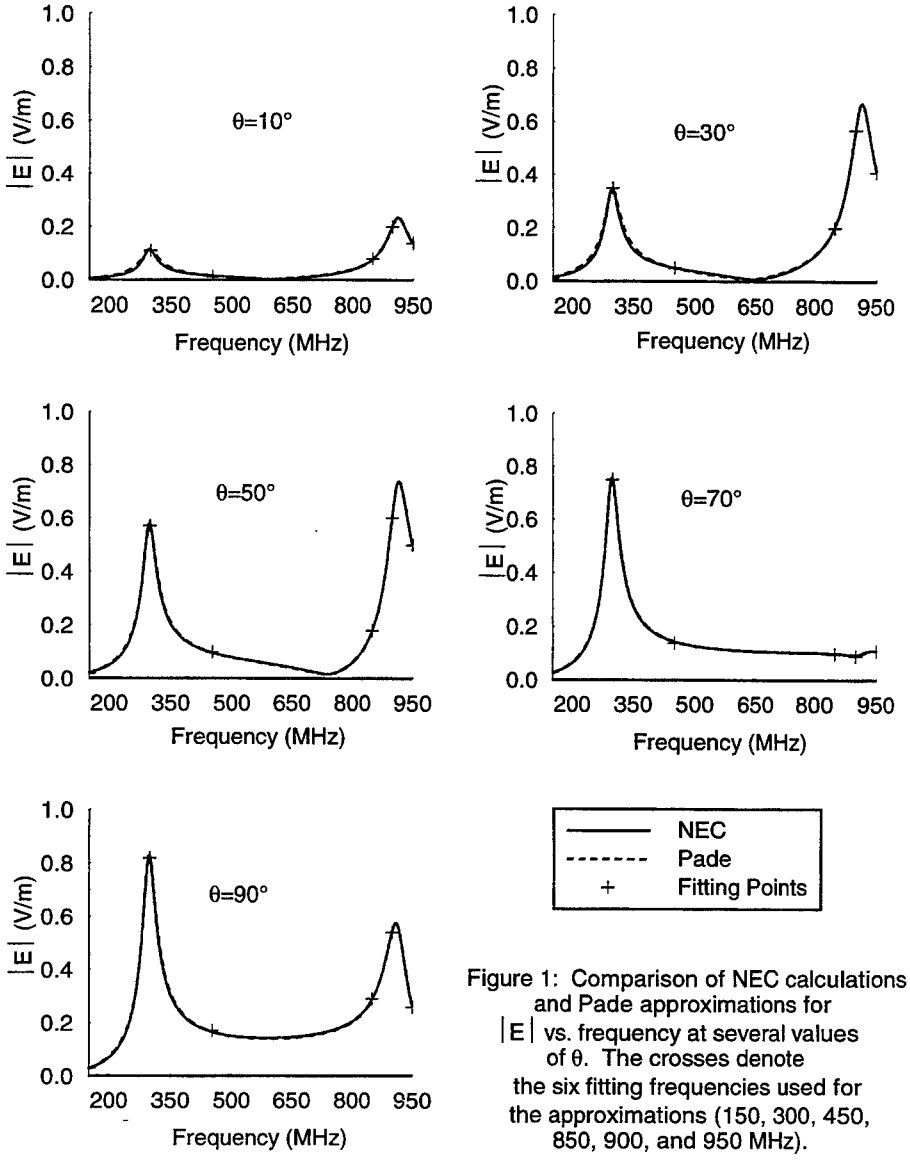


Figure 1: Comparison of NEC calculations and Pade approximations for $|E|$ vs. frequency at several values of θ. The crosses denote the six fitting frequencies used for the approximations (150, 300, 450, 850, 900, and 950 MHz).

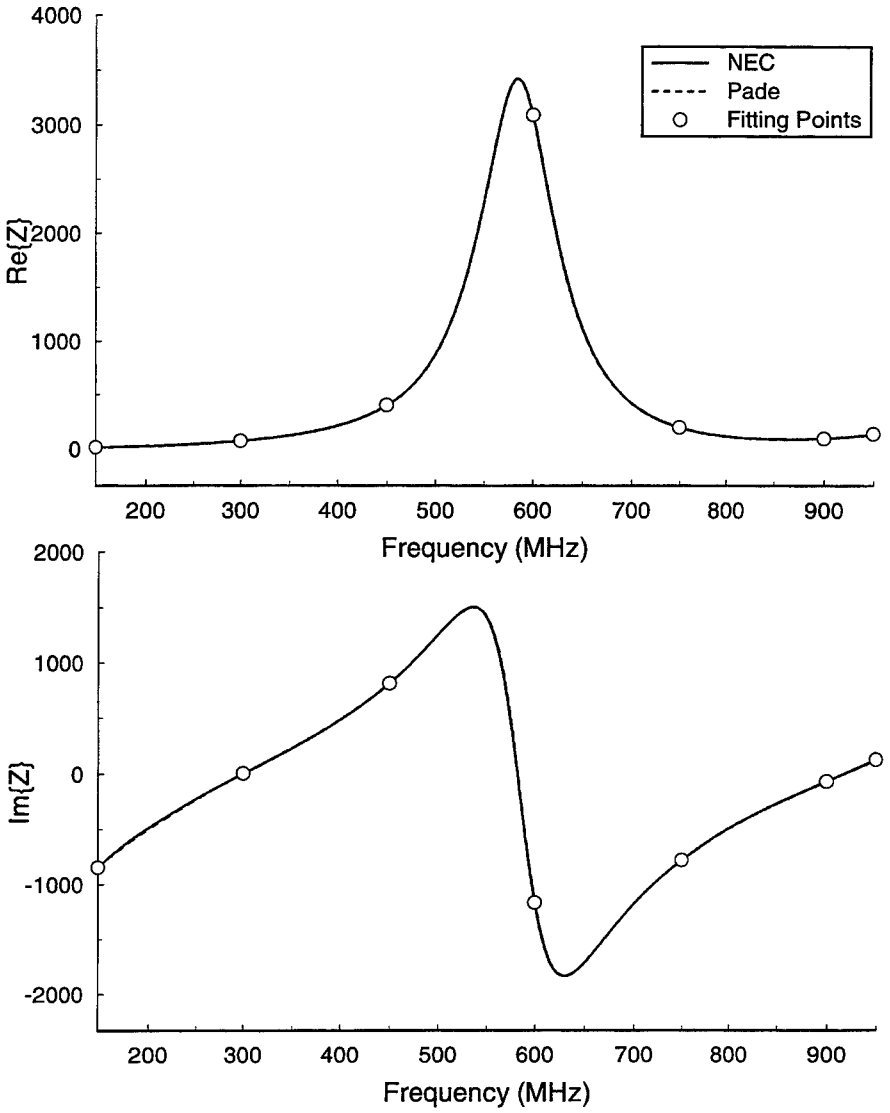
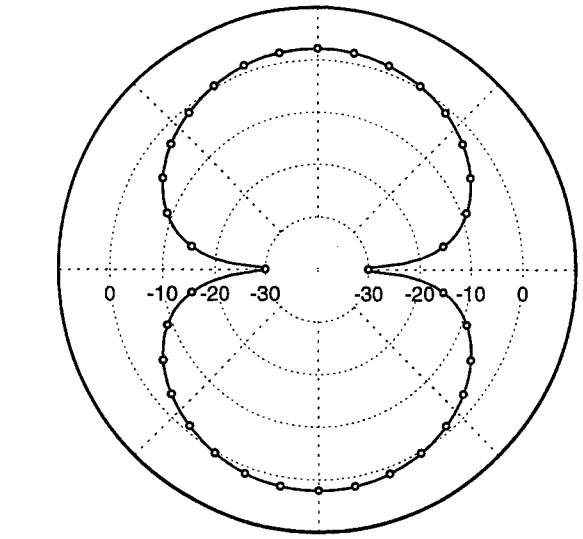
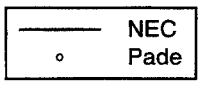
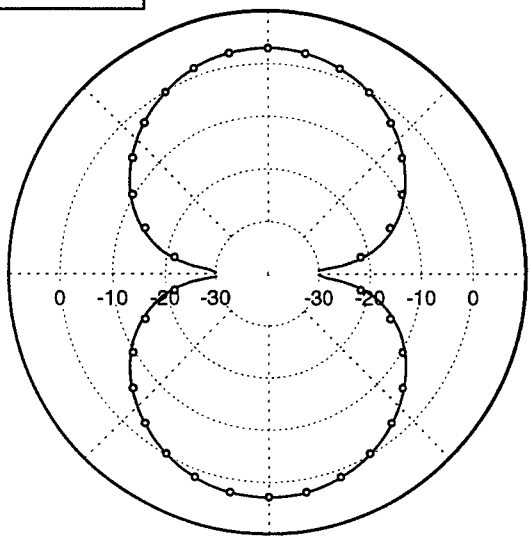


Figure 2: Re{Z} and Im{Z} showing the 7-point Pade approximation of the NEC curves vs. frequency. The fitting frequencies are shown as circles on the plot.



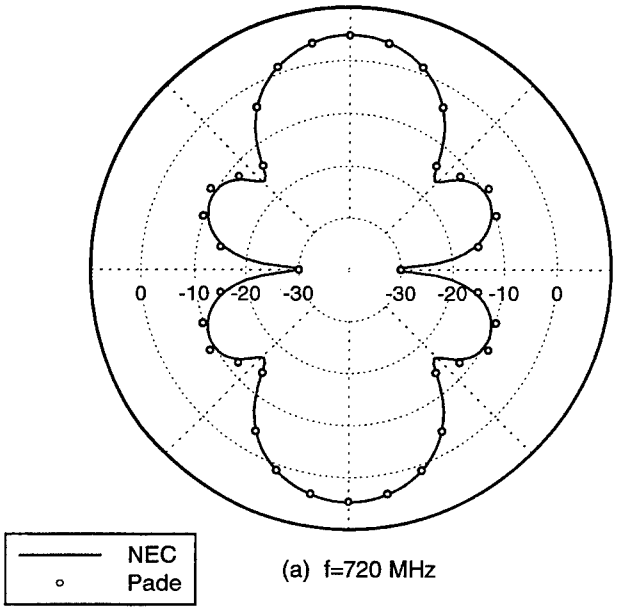


(a) f=300 MHz

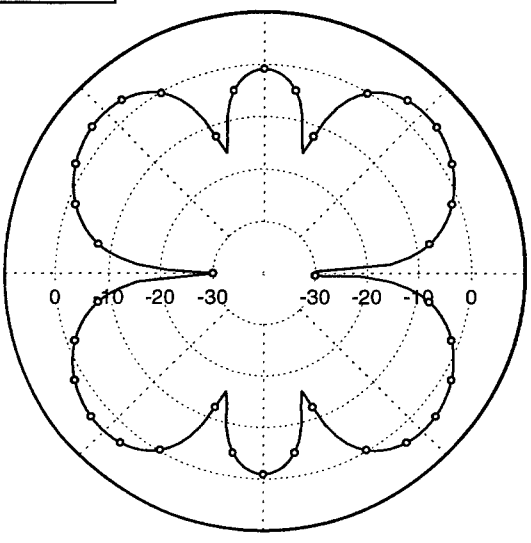


(b) f=500 MHz

Figure 3: NEC vs. Pade gain plots at (a) 300 MHz and (b) 500 MHz



(a) f=720 MHz



(b) f=933 MHz

Figure 4: NEC vs. Pade gain plots at (a) 720 MHz and (b) 933 MHz

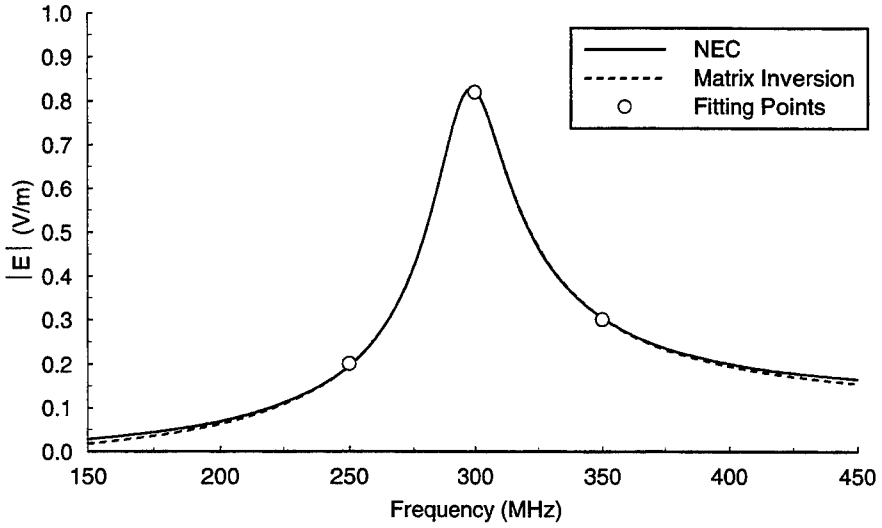


Figure 5: NEC vs. MBPE matrix inverted fit to 150-450 MHz frequency range

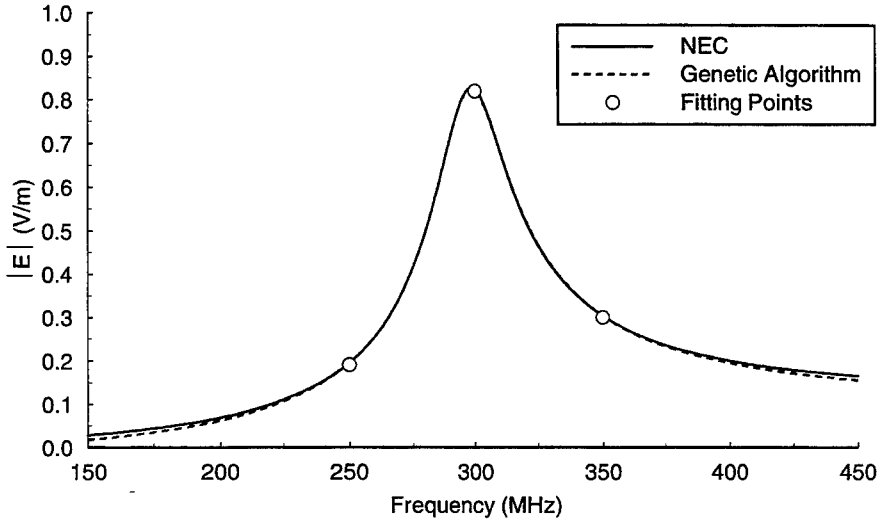


Figure 6: NEC vs. MBPE Genetic Algorithm fit to 150-450 MHz frequency range

AN ACCURATE ALGORITHM FOR NONUNIFORM FAST FOURIER TRANSFORMS (NUFFT) AND ITS APPLICATIONS

Q. H. Liu[1] and N. Nguyen[2]
[1]Klipsch School of Electrical and Computer Engineering
[2]Department of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003

Abstract

Based on the (m, N, q)-regular Fourier matrix, a new algorithm is proposed for fast Fourier transform of nonuniform (unequally spaced) data. Numerical results show that the accuracy of this algorithm is much better than previously reported results with the same computation complexity of $O(N \log_2 N)$. Numerical examples are shown for the applications in computational electromagnetics.

I. Introduction

Fast Fourier transform (FFT) has been enjoying widespread applications in numerical analysis and other areas of applied mathematics since Cooley and Tukey [1] established, in the 1960s, a powerful fast algorithm for calculating discrete Fourier transforms. The requirement for using FFT algorithms is that the input data must be equally spaced. In many practical situations, however, the input data is nonuniform (i.e., not equally spaced), and hence the regular FFT does not apply. To overcome this difficulty Dutt and Rokhlin [2] and Beylkin [3] studied the problem of FFT for nonuniform (unequally spaced) data.

We propose a new approach to achieve the fast Fourier transform for nonuniform data by using a new class of matrices, the regular Fourier matrices [4]. This algorithm, also with a complexity of $O(N \log_2 N)$ where N is the number of data points, is more accurate than that proposed in [2] because our approximation error is minimized in the least-square sense.

One of the important applications of this NUFFT algorithm is to enhance the newly developed pseudospectral time-domain (PSTD) method [5], which requires only two cells per minimum wavelength, with the capability of having a nonuniform grid.

II. Formulation

Our aim is to develop a fast algorithm to find the following summation [2]:

$$f_j = F(\alpha)_j = \sum_{k=0}^{N-1} \alpha_k e^{i\omega_k \cdot t_j} \quad \text{for} \quad j = -N/2, \cdots, N/2 - 1, \tag{1}$$

where $\omega = \{\omega_0, \cdots, \omega_{N-1}\}$ and $t = \{t_{-N/2}, \cdots, t_{N/2-1}\}$ are finite sequences of real numbers, with $\omega_k \in [-N/2, N/2 - 1]$ for $k = 0, \cdots, N-1$ and $t_j = 2\pi j/N \in [-\pi, \pi]$ for $j = -N/2, \cdots, N/2 - 1$; $\alpha = \{\alpha_0, \cdots, \alpha_{N-1}\}$ and $f = \{f_{-N/2}, \cdots, f_{N/2-1}\}$ are finite sequences of complex numbers. Note that, unlike the regular FFT, ω_k's are nonuniform (i.e., unequally spaced).

The idea of Dutt and Rokhlin [2] for solving this problem by a regular FFT is to approximate a function $F : [-\pi, \pi] \to C$ of the form:

$$F(x) = e^{-bx^2} e^{icx} \quad \text{for} \quad x \in [-\pi, \pi], \tag{2}$$

where $b > 1/2$ and c is a real number, by a small number of equally spaced points on the unit circle.

We recognize that, in applications, the function F defined by (2) takes its values on a finite set only. Therefore, instead of (2), we can consider the following finite sequence

$$F(j) = s_j e^{i2\pi cj/N} \quad \text{for} \quad j = -N/2, \cdots, N/2 - 1, \tag{3}$$

where $s_j > 0$ (called "accuracy factors") are chosen to minimize the approximation error. The novelty of this algorithm is that its approximation is optimal in the least-square sense, which leads to much more accurate results.

A. The regular Fourier matrices

For an integer $m \geq 2$ let $w = e^{i2\pi/mN}$, q be an even positive integer, s_j $(j = -N/2, \cdots, N/2 - 1)$ be positive numbers, and c be a real number. Our aim is to find $x_{k-q/2}$ $(k = 0, \cdots, q)$ to satisfy the following condition:

$$s_j w^{jmc} = \sum_{k=[mc]-q/2}^{[mc]+q/2} x_{k-[mc]}(c) w^{jk} \quad \text{for every} \quad j = -N/2, \cdots, N/2 - 1, \tag{4}$$

where $[mc]$ denotes the integer nearest to mc. Defining matrices and vectors

$$A = \begin{bmatrix} w^{-\frac{N}{2}([mc]-q/2)} & w^{-\frac{N}{2}([mc]-q/2+1)} & \cdots & w^{-\frac{N}{2}([mc]+q/2)} \\ w^{(-\frac{N}{2}+1)([mc]-q/2)} & w^{(-\frac{N}{2}+1)([mc]-q/2+1)} & \cdots & w^{(-\frac{N}{2}+1)([mc]+q/2)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ w^{(\frac{N}{2}-1)([mc]-q/2)} & w^{(\frac{N}{2}-1)([mc]-q/2+1)} & \cdots & w^{(\frac{N}{2}-1)([mc]+q/2)} \end{bmatrix}, \tag{5}$$

$$v(c) = \begin{bmatrix} s_{-N/2} w^{-\frac{N}{2}mc} \\ s_{-N/2+1} w^{(-\frac{N}{2}+1)mc} \\ \vdots \\ s_0 \\ \vdots \\ s_{N/2-1} w^{(\frac{N}{2}-1)mc} \end{bmatrix}, \tag{6}$$

we obtain the equation:

$$Ax(c) = v(c). \tag{7}$$

Observe that (7) is a system of N linear equations with $(q + 1)$ unknowns. Since in our applications $q << N$, equation (7) cannot be expected to have an exact solution. However we can find the *least squares solution* of the inconsistent system (7). That is to find $x(c)$ such that $\|Ax(c) - v(c)\|$ is smallest possible:

$$x(c) = F^{-1} a(c), \tag{8}$$

where

$$a(c) = A^{\dagger}(c) \cdot v(c), \tag{9}$$

$$F(m,N,q) = \begin{bmatrix} N & \frac{w^{-N/2}-w^{N/2}}{1-w} & \cdots & \frac{w^{-qN/2}-w^{qN/2}}{1-w^{q}} \\ \frac{w^{N/2}-w^{-N/2}}{1-w^{-1}} & N & \cdots & \frac{w^{-(q-1)N/2}-w^{(q-1)N/2}}{1-w^{q-1}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{w^{qN/2}-w^{-qN/2}}{1-w^{-q}} & \frac{w^{(q-1)N/2}-w^{-(q-1)N/2}}{1-w^{-(q-1)}} & \cdots & N \end{bmatrix}, \tag{10}$$

where $A^{\dagger}$ denotes the complex conjugate transpose of A matrix. Observe that, while A, and hence $A^{\dagger}$, depends on c, the product matrix $F(m,N,q) = A^{\dagger}A$ is independent of c and is uniquely determined by m, N and q. This remarkable property of A is of great importance because it will reduce the number of operations by our algorithms and is a crucial point of this work. The matrix $F(m,N,q)$, for $m,N,q \in \mathcal{N}$, called the (m,N,q)-*regular Fourier matrix*, is a Hermitian matrix of dimension $(q+1)\times(q+1)$. The elements of $a(c)$ are given by

$$a_k(c) = \sum_{j=-N/2}^{N/2-1} s_j e^{i\frac{2\pi}{Nm}(\{mc\}+q/2-k)j}, \tag{11}$$

where $\{mc\} = mc - [mc]$, and $k = 0, \cdots, q$.

B. The NUFFT Algorithm

We may choose two different accuracy factors, namely (i) the Gaussian $s_j = e^{-b(\frac{2\pi j}{Nm})^2}$ and (ii) the cosine $s_j = \cos\frac{\pi j}{Nm}$ accuracy factors. In particular, for the cosine accuracy factor, a closed-form solution can be found for (11):

$$a_k(c) = i\sum_{\gamma=-1,1} \frac{\sin[\frac{\pi}{2m}(2k-\gamma-q-2\{mc\})]}{1-e^{i\frac{\pi}{Nm}(2\{mc\}+q-2k+\gamma)}}. \tag{12}$$

This solution saves many arithmetic operations. Unfortunately, we are not able to find a corresponding closed-form solution for the Gaussian accuracy factors. Because of this, it is only sensible to use the Gaussian accuracy factors when many repeated NUFFTs are required for the same ω_k points, since then one can pre-compute (8) for all the subsequent NUFFTs.

In summary, our NUFFT algorithm consists of following steps:

(1) Compute $x_j(\omega_k)$ by (8) for $j = 0, \cdots, q$ and $k = 0, \cdots, N-1$.

(2) Calculate Fourier coefficients

$$\tau_l = \sum_{j,k,[m\omega_k]+j=l} \alpha_k \cdot x_j(\omega_k).$$

(3) Use uniform FFT to evaluate

$$T_j = \sum_{k=-mN/2}^{mN/2-1} \tau_k \cdot e^{2\pi ikj/mN}.$$

(4) Scale the values to arrive at the approximated NUFFT

$$\tilde{f}_j = T_j \cdot s_j^{-1}.$$

The asymptotic number of arithmetic operations of this algorithm is $O(mN \log_2 N)$, where $m \ll N$. Usually we choose $m = 2$ and $q = 8$.

III. Numerical Results

We first compare (i) Dutt-Rokhlin's algorithm, and our algorithm with (ii) Gaussian and (iii) cosine accuracy factors for different values of N. For this comparison, we use $m = 2$ and $q = 8$. The computation was done with Matlab on a SUN Ultra 1 Workstation. We compute the errors for equation (1) by the formulae defined in [2]:

$$E_\infty = \max_{0 \le j \le N} |\tilde{f}_j - f_j| \Big/ \sum_{j=0}^{N-1} |\alpha_j|$$

and

$$E_2 = \sqrt{\sum_{j=0}^{N-1} |\tilde{f}_j - f_j|^2 \Big/ \sum_{j=0}^{N-1} |f_j|^2}.$$

Figure 1 shows the error E_2 and E_∞ of the three algorithms for $N = 64$, 128, 256, 512, 1024, 2048, and 4096. Both ω_j and α_j are given by pseudo-random number generators. It is observed that overall the errors of the our algorithm with Gaussian and cosine accuracy factors are respectively about 8 and 12 times smaller than Dutt-Rokhlin's algorithm for E_∞, and 16 and 19 times smaller for E_2.

Figure 2 displays the errors E_2 and E_∞ as functions of q for $N = 64$. It is seen that for $2 \le q \le 12$, on average, our algorithm is 45 times (for cosine accuracy factors) or 24 times (for Gaussian accuracy factor) more accurate than the algorithm in [2]. This conclusion is independent of N.

We apply this NUFFT algorithm to perform spectral analysis of electromagnetic waves near sharp medium discontinuities. Shown in Figure 3(a) is the transverse electric field due to a transient plane wave normally incident to a thin conductive dielectric slab of 15 cm thick. The slab has $\epsilon_r = 4$ and $\sigma = 1$ S/m, and the background is vacuum. The center frequency of the transient incident wave is 166.7 MHz (a Blackman-Harris window time function). In terms of the center frequency, the slab is only 1/12 of the wavelength in vacuum. The fast spatial variation of the field (obtained by the FDTD method with a very fine grid) is depicted in Figure 3(a) near the slab. As shown in the figure, in order to effectively describe the field variation, a fine sampling is used near the slab, while a much coarser sampling is used away from the slab where the field has a slow variation. Figures 3(b) and 3(c) show the excellent agreement of the real and imaginary parts of the (spatial) spectrum obtained by the NUFFT (with cosine accuracy factors and $q = 8, m = 2$) and direct evaluation. Figure 3(d) displays the absolute error from our NUFFT algorithm and that from [2]. Quantitatively, the L_2 and L_∞ errors defined in [2] are $E_2 = 2.731 \times 10^{-6}$ and $E_\infty = 2.956 \times 10^{-6}$ for our algorithm, and $E_2 = 3.849 \times 10^{-5}$ and $E_\infty = 3.694 \times 10^{-5}$ for the algorithm in [2]. Our NUFFT algorithm is more than one order of magnitude more accurate. This algorithm will benefit the development of a nonuniform pseudospectral time-domain (PSTD) method for Maxwell's equations [5].

Figure 4 shows the CPU time as a function of N in the NUFFT algorithm. Both the input data α_k and its locations ω_k ($k = 0, \cdots, N-1$) are obtained by a pseudo-random number generator with large variations. It clearly verifies that the algorithm is of complexity $O(N \log_2 N)$.

IV. Conclusions

Based on a class of regular Fourier matrices, a new nonuniform fast Fourier transform (NUFFT) algorithm is developed for unequally spaced data. With a comparable complexity of $O(N \log_2 N)$, this algorithm is much more accurate than previously reported results since it is optimal in the least squares sense. The algorithm is useful for computational electromagnetics and other fields of applied mathematics.

Acknowledgment

This work was supported by Environmental Protection Agency under a PECASE grant CR-825-225-010, by a SURP grant from Sandia National Laboratories, and by National Science Foundation under a CAREER grant ECS-9702195.

References

[1] J. W. Cooley, and J. W. Tukey, "Algorithm for the machine computation of complex Fourier series," *Math. Comp.*, vol. 19, pp. 297–301, 1965.

[2] A. Dutt, and V. Rokhlin, "Fast Fourier transforms for nonequispaced data," *SIAM J. Sci. Comp.*, vol. 14, pp. 1368–1393, 1993.

[3] G. Beylkin, "On the fast Fourier transform of functions with singularities," *Applied and Computational Harmonic Analysis*, vol. 2, pp. 363–382, 1995.

[4] N. Nguyen, and Q. H. Liu, "The regular Fourier matrices and nonuniform fast Fourier transforms," *New Mexico State University Technical Report*, July 1997.

[5] Q. H. Liu, "The PSTD algorithm: a time-domain method requiring only two cells per wavelength," *Microwave Opt. Tech. Lett.*, vol. 15, no. 3, pp. 158–165, June 1997.

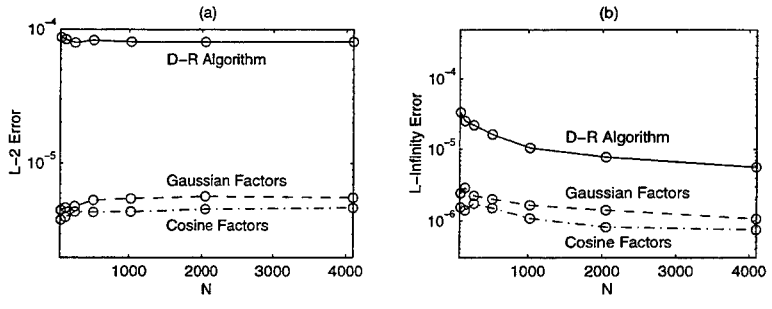


Figure 1. Comparison of E_2 and E_∞ from Dutt-Rokhlin algorithm, and from this algorithm with Gaussian and cosine accuracy factors. (a) E_2 as a function of N. (b) E_∞ as a function of N.

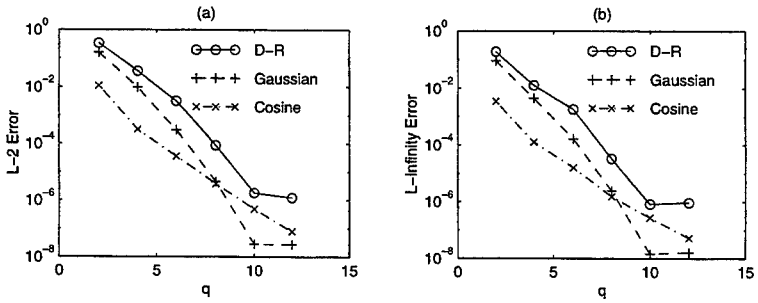


Figure 2. (a) E_2 and (b) E_∞ as functions of q for Dutt-Rokhlin's algorithm (D-R), and for this algorithm with Gaussian and cosine accuracy factors. The size of the FFT array is $N = 64$. Similar results are obtained for other values of $64 \leq N \leq 4096$.

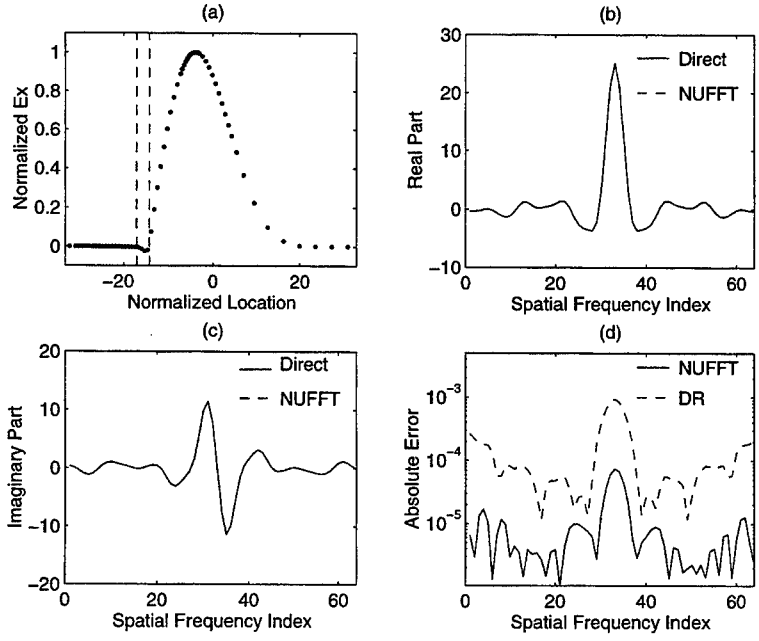


Figure 3. (a) Spatial distribution of transient electromagnetic field near a conductive dielectric slab. Dashed lines show interfaces of the slab. Nonuniform sampling is used to increase the resolution close to the slab. (b) Real and (c) imaginary parts of the (spatial) spectrum of the field obtained by direct evaluation and by the NUFFT algorithm. (d) Absolute errors from this algorithm and that by Dutt-Rokhlin [2].

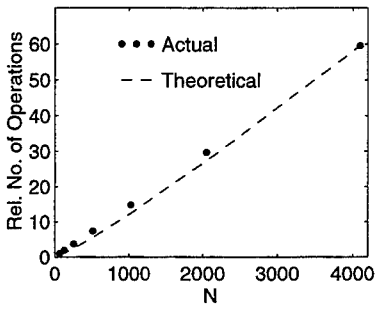


Figure 4. Relative number of operations as a function of N. Both input data and the locations of the sampling points are random. The dashed curve is the theoretically predicted curve $O(N \log_2 N)$ passing through the last point.

A SUBSPACE APPROACH TO FAST MOMENT METHOD SCATTERING PREDICTIONS OVER LIMITED SECTORS

John Stach <stach@erim-int.com>
ERIM International, http://erim-int.com
P.O. Box 134008
Ann Arbor, MI 48113-4008

ABSTRACT

The number of basis functions needed for sufficient accuracy and convergence of a moment method scattering solution is often much more than the number of degrees of freedom required to reconstruct the scattering pattern, especially when the desired pattern is limited to a small angular sector. In this paper, we describe a technique by which the computational complexity of the scattering solution can be significantly reduced by projecting the impedance matrix onto a subspace that spans a limited prediction angle sector. In addition to the derivation of the technique, we also include some comments on the application of this technique, especially to iterative problems, and note the relationship between this technique and the MEI approach to solving finite element problems.

1.0 INTRODUCTION

Typical applications of moment method (MM) solutions for scattering require basis function sampling densities of about ten samples per wavelength to achieve adequate accuracy and convergence. However, the degrees of freedom required to accurately reconstruct the scattered field over limited angles may be much smaller than the number of basis functions. In this paper, we will describe a technique by which the computations required for an accurate MM prediction can be significantly reduced by limiting the size of the matrix inverse to be the size of the subspace spanned by the sector of predictions rather than the number of basis functions. Furthermore, we will show how this technique is closely related to the measured equation of invariance (MEI) approach [1] which has been applied to finite element problems.

In Section 2, we will describe the subspace technique for MM scattering predictions. In Section 3, we will discuss some practical issues and the relationship between the subspace approach and MEI. In Section 4, we will describe the computational advantages of the subspace approach, and in Section 5, we will summarize the approach.

Lower case characters will be used to represent scalars (normal) or vectors (bold), and upper case bold characters will be used to represent matrices.

2.0 SUBSPACE APPROACH

A scattering prediction based on a moment method solution can be formulated as a bilinear equation [3]

$$s(r,t) = \sum_{i}^{N}\sum_{j}^{N} a_i(t) b_j(r) y_{ij}, \qquad \text{(Eq 1)}$$

where s is the measured scattering value, t and r represent the transmitter and receive geometries (e.g., polarization, angle) for the radar respectively, a_i describes the coupling of the transmitter to the ith basis function, b_j describes the coupling of the receiver to the jth basis function, and the admittance elements, y_{ij}, describe the

coupling between the ith and jth basis function due to the target geometry. The elements of the admittance matrix are computed by taking the inverse of the impedance matrix.

If we separate the sums, the first sum represents the computation of currents, j, due to the radar transmitter,

$$j_i(t) = \sum_j^N a_j(t) y_{ij}, \qquad \text{(Eq 2)}$$

and the second sum represents the propagation of those currents to the radar receiver, i.e.,

$$s(r,t) = \sum_i^N b_i(r) j_i(t). \qquad \text{(Eq 3)}$$

Note that if we know the currents, there is no need to invert the impedance matrix. Therefore, applications where the technique is useful are (1) where currents must be computed many times for iterative computations, and (2) where we can get away with pseudo-currents that are much simpler to compute yet span the prediction subspace. These applications will be explored further in Section 3.

The sum in Eq 2 can be written as a matrix-vector product, and furthermore, if there are several scattering geometries of interest, then the corresponding currents can be represented by the matrix product

$$\boldsymbol{J} = \boldsymbol{YA}, \qquad \text{(Eq 4)}$$

where the columns of $\boldsymbol{J}$ represent the currents, and the columns of $\boldsymbol{A}$ represent the corresponding excitations, i.e., transmitter geometries. We assume that $\boldsymbol{Y}$ is full rank, and $\boldsymbol{J}$ and $\boldsymbol{A}$ are full column rank but not full row rank, i.e., there are a limited sector of angles for which we will compute scattering predictions. Since $\boldsymbol{Y} = \boldsymbol{Z}^{-1}$, we can multiply both sides of Eq 4 by $\boldsymbol{Z}$ to get

$$\boldsymbol{ZJ} = \boldsymbol{A}. \qquad \text{(Eq 5)}$$

This equation represents a projection of $\boldsymbol{Z}$ onto a subspace spanned by the columns of $\boldsymbol{J}$. Therefore, we can find another matrix, $\boldsymbol{Z}_\alpha$, such that.

$$\boldsymbol{Z}_\alpha \boldsymbol{J}_\alpha = \boldsymbol{A}, \qquad \text{(Eq 6)}$$

where $\boldsymbol{Z}_\alpha$ is full column rank and $\boldsymbol{J}_\alpha$ is derived from $\boldsymbol{J}$. The matrix, $\boldsymbol{J}_\alpha$, can simply be a subset of the rows of $\boldsymbol{J}$ or linear combinations of the rows of $\boldsymbol{J}$ based on averages, wavelet decompositions, geometric partitioning, or other nonlinear combinations that span the prediction subspace. We will discuss the choice of rows for $\boldsymbol{J}_\alpha$ in Section 3.

If we treat $\boldsymbol{J}_\alpha$ as known quantities based on $\boldsymbol{J}$ as described above, then we can treat each row of $\boldsymbol{Z}_\alpha$ as unknowns which are to be solved for each corresponding row of $\boldsymbol{A}$. Since $\boldsymbol{J}_\alpha$ is full column rank, the solution for $\boldsymbol{Z}_\alpha$ that satisfies Eq 6 is given by

$$\boldsymbol{Z}_\alpha = \boldsymbol{ZJ}(\boldsymbol{J}_\alpha^+), \qquad \text{(Eq 7)}$$

where superscript + represents the pseudo-inverse [2, p139], i.e., $(\boldsymbol{J}_\alpha^+)\boldsymbol{J}_\alpha = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix. Given $\boldsymbol{Z}_\alpha$ which is full column rank, we can now solve for $\boldsymbol{J}_\alpha$ given the excitation, $\boldsymbol{A}$, from

$$Z_\alpha^+ A = J_\alpha . \qquad \text{(Eq 8)}$$

The final step is to find a way to compute the scattered field from J_α. If we begin by assuming that we must be able to solve for the scattering at all receiver geometries from all currents, then we can start with

$$B^T J = S , \qquad \text{(Eq 9)}$$

where superscript T denotes transpose, B is full column rank, J is full column rank as before, and S is the bistatic scattering matrix whose rows represent receiver geometries and whose columns represent transmitter geometries. We need to find a matrix, B_α that when multiplied by J_α, results in the same S. By the same approach as before, we find that

$$B_\alpha^T = B^T J(J_\alpha^+) . \qquad \text{(Eq 10)}$$

Now we can put the equation back together for a single scattered field computation.

$$s = b^T J(J_\alpha^+)[ZJ(J_\alpha^+)]^+ a . \qquad \text{(Eq 11)}$$

We note that we have essentially defined a projection operator,

$$I_\alpha = J(J_\alpha^+) , \qquad \text{(Eq 12)}$$

based on the currents, which reduces the rank of the impedance matrix to the size of the subspace spanned by the limited angle sector of the scattered field.

3.0 APPLICATIONS

As noted before, if we need the currents in order to construct the projection operator in Eq 12, then why bother? In the first place, there are many optimization problems for which the scattered field is computed many times for small perturbations in the target or radar geometries. For these problems, the computations associated with a single computation of currents may be insignificant compared to the many iterations that follow. In such cases, we are guaranteed a matrix of currents that perfectly spans the appropriate subspace. We have used this approach for interpolation/extrapolation problems [4] and have found excellent agreement between the final scattering predictions from the subspace approach and the true scattering predictions from direct inversion of the impedance matrix.

For other problems where the scattering predictions are needed over a limited sector only once, we must find a matrix, J, that is (1) efficient to compute, and (2) spans the same subspace as the true currents. In the measured equation of invariance (MEI) approach to finite element solutions, these false currents are called metrons. The metrons are used in MEI to derive coefficients for local finite element mesh points that can be used to reliably compute nearby fields which would otherwise require an integral over the target surface. There are many considerations in the derivation of adequate metrons, and generally useful metrons have not yet been derived for this problem. However, it seems likely that some variation on the physical optics currents would be a good starting point.

Another issue is the number of rows needed in J_α. We need to use enough rows to satisfy the equality in Eq 6 to an acceptable level. If the prediction angles are oversampled, then it is possible that fewer rows will be required to satisfy Eq 6 than there are columns of A. Note that we can also regularize the generalized inverse of J_α to account for anticipated errors in the model or currents.

4.0 COMPUTATIONS

Table 1 shows some comparisons between subspace and direct inverse computations for scattering predic-

Table 1: Computational comparisons between subspace and direct inversion

Computation	Complexity	Method	Comments
$b'I_\alpha(ZI_\alpha)^\dagger$ and $I_\alpha(ZI_\alpha)^\dagger a$	$O[n_2^2 n_3 + 2n_2 n_3^2 + n_3^3 + 4n_1 n_2 n_3]$	subspace	Takes advantage of common terms
$b'Z^{-1}$ and $Z^{-1}a$	$O[n_2^3 + 2n_1 n_2^2]$	direct	
$s = b'I_\alpha(ZI_\alpha)^\dagger a$	$O[2n_1 n_2 n_3 + n_1 n_3]$	subspace	Includes scattering over entire sector (not just scalar)
$s = b'Ya$	$O[n_1 n_2^2 + n_1 n_2]$	direct	
$I_\alpha = J(J_\alpha^+)$ n_1 - Number of measurements, n_2 - Number of model elements n_3 - Number of subspace coefficients (typically slightly greater than n1)			

tions. The number of predictions and subspace coefficients, n_1 and n_3 respectively, are typically comparable and much smaller than the number of model basis functions, n_2,. If we let n_1 and n_3 be a fraction of n_2, i.e. $n_1 = n_3 = \alpha n_2$, then the subspace method reduces the total number of computations per iteration by about a factor of about α. If we perturb only the diagonal impedances (complex loads) for both methods, the subspace approach results in a total per iteration cost reduction factor of α^2. In general we would expect α to be about an order of magnitude. Note that there is a relatively fixed set up cost associated with the subspace approach to compute currents and a projection for each angular sector.

5.0 SUMMARY

We have presented a method by which the computational complexity of MM scattering predictions for limited angle sectors can be significantly reduced. We develop a subspace projection for the impedance matrix using either currents or metrons which span the same prediction subspace as the currents. The approach is analogous to that used in the MEI approach to finite element solutions.

The computational savings is highly dependent on the model fidelity, sector size, and sampling density of the scattered field. However, savings between one and two orders of magnitude should be typical.

6.0 REFERENCES

[1] .K.K. Mei, R. Pous and Z.Q. Chen, "The Measured Equation of Invariance: A New Concept in Field Computation," *IEEE Trans. Antennas Propagation*, March 1994.

[2] G.H. Golub and C.F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Maryland, 1985.

[3] R.F. Harrington, *Field computations by moment methods*, Robert E. Krieger Publishing Co, Malabar, FL, 1968

[4] J. Stach and L. Heck, "RCS Extrapolation Using Network Models," *Proceedings of the URSI Radio Science Meeting*, Ann Arbor, MI., June 1993.

APPLICATION OF BIORTHOGONAL B-SPLINE-WAVELETS TO TELEGRAPHER'S EQUATIONS

MARTIN AIDAM AND PETER RUSSER

LEHRSTUHL FÜR HOCHFREQUENZTECHNIK AN DER TECHNISCHEN UNIVERSITÄT MÜNCHEN, GERMANY

ABSTRACT. Biorthogonal B-spline-wavelets are used to represent voltage and current of telegrapher's equations with respect to space. A Petrov-Galerkin procedure using B-splines as trial functions (more precisely the primal scaling function which is the B-spline and the derived primal wavelet) and the dual scaling functions and wavelets as test functions, is applied to obtain a set of ordinary differential equations (method of lines) which are integrated by a simple explicit two-step scheme, namely Nyström's method and the modified Euler scheme to start the procedure. Results of numerical experiments including thresholding effects are presented.

1. INTRODUCTION

The interest in investigating wavelets in the scope of computational electromagnetics originates mainly from three properties of wavelets: the capability to construct adaptive schemes easily without evaluating costly error estimates, the uniform boundedness of the matrices of some operators in a wavelet basis, and their efficiency in data compression, i.e. a function can be represented by much fewer coefficients. For overviews see e.g. [2, 3].

So far, the application of wavelets to electromagnetic problems —formulated as partial differential equations— has found limited use, where the advantages of wavelets have not been fully exploited. In some cases only scaling functions [9, 12], and in others only one additional wavelet level [5, 11] have been used. In contrast, the algorithm, we implemented, allows —at least theoretically— an arbitrary number of additional wavelet levels. The more wavelet levels one uses, the more efficient will be the lossy data compression via thresholding, i.e. in the adaptive procedure one has less unknowns.

In this paper, we investigate the capability of biorthogonal B-spline-wavelets for the application to telegrapher's equations.

2. B-SPLINE-WAVELETS

To easily incorporate ideal electric and magnetic boundaries, one likes to use the mirror principle. This means, that the scaling and wavelet functions we use have to be symmetric or antisymmetric. Moreover, we favour compactly supported wavelets, since difference operators should only have a finite extent. Unfortunately, compactly supported, symmetric, real wavelets which are orthogonal do not exist, so one is forced to drop orthogonality or to use wavelets with an infinite support and approximate the infinite difference operator by a finite one [6]. In this paper, the first option is made, and the biorthogonal, symmetric, compactly supported and real B-spline-wavelets constructed by Cohen et. al. [1] are chosen.

As scaling function, the cardinal B-splines

$$\varphi_d(x) := d\,[0,1,\ldots,d]\left(\cdot - x - \left\lfloor \frac{d}{2} \right\rfloor\right)_+^{d-1} \tag{1}$$

are used. $[t_i,\ldots,t_{i+d}]f$ is the divided difference of $f \in C^d(\mathbb{R})$ for the knot sequence $t_i \leqslant \ldots \leqslant t_{i+d}$, $x_+^d := (\max\{0,x\})^d$, $\lfloor x \rfloor := \max\{z \in \mathbb{Z} : z \leqslant x\}$, $x \in \mathbb{R}$ $\lceil x \rceil := \min\{z \in \mathbb{Z} : z \geqslant x\}$, $x \in \mathbb{R}$. These functions are symmetric around $l(d)/2$, with

$$\varphi_d(x + l(d)) = \varphi_d(-x) \quad \text{for} \quad x \in \mathbb{R}\,, \quad l(d) := d \bmod 2 = \begin{cases} 0 & \text{for } d \text{ even}, \\ 1 & \text{for } d \text{ odd}, \end{cases} \tag{2}$$

and have compact support $\operatorname{supp}\varphi_d = [l_1, l_2]$ with $l_1 := -\lfloor \frac{d}{2} \rfloor$ and $l_2 := \lceil \frac{d}{2} \rceil$. They are refinable

$$\varphi_d(x) = \sqrt{2} \sum_{k=l_1}^{l_2} h_k \varphi_d(2x-k) \quad \text{with} \quad h_k := 2^{1-d} \binom{d}{j + \lfloor \frac{d}{2} \rfloor}, \; j \in \{l_1, \dots, l_1 + d\} \tag{3}$$

and of order d, i.e. all polynomials at most of degree $d-1$ can be represented exactly as linear combinations of the translates $\varphi_d(\cdot - k)$, $k \in \mathbb{Z}$

The dual functions $\tilde{\varphi}_{d,\tilde{d}}$, $\tilde{d} \geq d$, $\tilde{d} \in \mathbb{N}$, $l(\tilde{d}+d) = 0$, also have compact support $\operatorname{supp}\tilde{\varphi}_{d,\tilde{d}} = [l_1 - \tilde{d} + 1, l_2 + \tilde{d} - 1]$ and the same symmetry property

$$\tilde{\varphi}_{d,\tilde{d}}(x + l(d)) = \tilde{\varphi}_{d,\tilde{d}}(-x) \tag{4}$$

for $x \in \mathbb{R}$. They are biorthogonal to φ_d with respect to the inner product of $L_2(\mathbb{R})$

$$(\varphi_d, \tilde{\varphi}_{d,\tilde{d}}(\cdot - k))_{L_2(\mathbb{R})} = \delta_{0,k} \quad \text{for} \quad k \in \mathbb{Z}, \tag{5}$$

and of order $\tilde{d}$. They are also refinable

$$\tilde{\varphi}_{d,\tilde{d}}(x) = \sqrt{2} \sum_{k=\tilde{l}_1}^{\tilde{l}_2} \tilde{h}_k \tilde{\varphi}_{d,\tilde{d}}(2x-k) \tag{6a}$$

where

$$\tilde{l}_1 := l_1 - \tilde{d} + 1 \qquad \tilde{l}_2 := l_2 + \tilde{d} - 1 \tag{6b}$$

$$\tilde{h}_k := \frac{1}{\sqrt{2}} \sum_{j \in \mathbb{Z}} \tilde{r}_j \tilde{p}_{k-j} \qquad k \in \left\{ -\left\lfloor \frac{\tilde{d}}{2} \right\rfloor - \frac{d+\tilde{d}}{2} + 1, \dots, \left\lceil \frac{\tilde{d}}{2} \right\rceil + \frac{d+\tilde{d}}{2} - 1 \right\} \tag{6c}$$

$$\tilde{r}_j := 2^{1-\tilde{d}} \binom{\tilde{d}}{j + \lfloor \frac{\tilde{d}}{2} \rfloor} \qquad j \in \{\tilde{l}_1, \dots, \tilde{l}_1 + \tilde{d}\} \tag{6d}$$

$$\tilde{p}_j = (-1)^j \sum_{n=|j|}^{\frac{d+\tilde{d}}{2}-1} \frac{1}{4^n} \binom{\frac{d+\tilde{d}}{2} - 1 + n}{n} \binom{2n}{n+j} \qquad j \in \left\{ -\frac{d+\tilde{d}}{2} + 1, \dots, \frac{d+\tilde{d}}{2} - 1 \right\}. \tag{6e}$$

As usual, the wavelets are defined as

$$\psi_{d,\tilde{d}}(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} g_k \varphi_d(2x-k) \qquad g_k = (-1)^k \tilde{h}_{1-k} \tag{7}$$

$$\tilde{\psi}_{d,\tilde{d}}(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} \tilde{g}_k \tilde{\varphi}_{d,\tilde{d}}(2x-k) \qquad \tilde{g}_k = (-1)^k h_{1-k}. \tag{8}$$

They are symmetric or antisymmetric around $1/2$, depending on $l(d)$, i.e.

$$\psi_{d,\tilde{d}}(x+1) = (-1)^{l(d)} \psi_{d,\tilde{d}}(-x), \quad \tilde{\psi}_{d,\tilde{d}}(x+1) = (-1)^{l(d)} \tilde{\psi}_{d,\tilde{d}}(-x). \tag{9}$$

FIGURE 1. Example (3,3)-B-spline wavelets: scaling function $\varphi_3(x-1)$, wavelet $\psi_{3,3}(x-2)$, dual scaling function $\tilde{\varphi}_{3,3}(x-3)$ and dual wavelet $\tilde{\psi}_{3,3}(x-2)$.

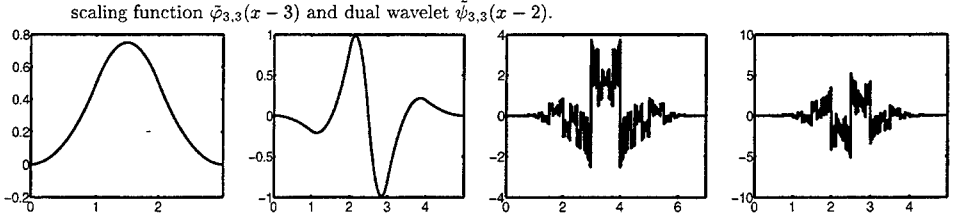


Fig. 1 shows as an example the (3,3)-B-spline scaling function, the wavelet, the dual scaling function and the dual wavelet. Although the dual functions look a little odd, they are able to represent quadratic polynomials exactly.

3. Functions in Wavelet Spaces

A function $f \in L_2(\mathbb{R})$ can be represented by the sum

$$f(x) = \sum_{k\in\mathbb{Z}} c_k^C \varphi_{C,k}(x) + \sum_{m=C}^{\infty}\sum_{k\in\mathbb{Z}} d_k^m \psi_{m,k}(x)\,, \tag{10}$$

where the superscript C indicates an arbitrary coarsest resolution level (we number higher resolution levels by higher superscripts) and

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)\,, \quad \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)\,. \tag{11}$$

For convenience, we drop the subscripts denoting the primal and dual orders of the expansion (there should be no possibility for confusion).

Practically, we truncate the infinite sum at some finest wavelet level M. With the help of the fast wavelet transform, our approximation can be represented equivalently by

$$f(x) = \sum_{k\in\mathbb{Z}} c_k^{M+1}\varphi_{M+1,k}\,. \tag{12}$$

For the solution of initial value problems using wavelet algorithms one needs to be able to express more or less arbitrary functions in terms of wavelet coefficients. The easiest way to do this is to expand the function only in terms of scaling coefficients at some finest level $M+1$ where the approximation error is acceptable and then to apply the fast wavelet transform to calculate all the wavelet coefficients at coarser levels. To calculate the scaling coefficients, we apply the oversampling technique proposed by Ware [14]

$$c_{j,k} = \theta_{j,k}(f) = 2^{-j/2}\sum_i w_i f(2^{-j}(k+\tfrac{1}{2}))\,. \tag{13}$$

The weights w_i are chosen such that

$$\theta_{j,k}(\varphi_{j,k'}) = \sum_i w_i \varphi(\tfrac{i}{2} + k - k') = \delta_{k,k'}\,. \tag{14}$$

Ware also discusses how to calculate directly the wavelet coefficients and how to calculate them in an adapted lacunary basis. This will be implemented soon, but was not used in our numerical experiments, yet.

To incorporate ideal boundary conditions, say at $x = a$ and $x = b$, it is possible to extend this interval by using the mirror principle or by periodizing. Straight forward calculations give the properties of the coefficients of symmetric, antisymmetric and periodic functions in a B-spline-wavelet basis for the problem at hand.

4. Discretizing Differential Operators

Petrov-Galerkin discretizations of linear differential equations with constant coefficients lead to integrals of the type

$$\int_{-\infty}^{\infty} f^{(n)}(x)g(x)dx\,,$$

where $f^{(n)}$ is the trial function's nth derivative and g is the test function. Apparently, when using scaling functions and wavelets as trial functions and their duals as test functions, we have four different types

$$A(j,i,l,k) = \int_{-\infty}^{\infty} \varphi_{j,l}^{(n)}(x)\tilde{\varphi}_{i,k}(x)dx \qquad B(j,i,l,k) = \int_{-\infty}^{\infty} \psi_{j,l}^{(n)}(x)\tilde{\varphi}_{i,k}(x)dx \tag{15a}$$

$$C(j,i,l,k) = \int_{-\infty}^{\infty} \varphi_{j,l}^{(n)}(x)\tilde{\psi}_{i,k}(x)dx \qquad D(j,i,l,k) = \int_{-\infty}^{\infty} \psi_{j,l}^{(n)}(x)\tilde{\psi}_{i,k}(x)dx\,. \tag{15b}$$

Fortunately, all these integrals can be calculated from the knowledge of $A(0,0,l,0)$. To calculate this integral, one has to solve an eigenvalue problem and add appropriate conditions to uniquely define the eigenvector [7]. We use the program written by Angela Kunoth [8], which implements the method presented in [7] and the references therein.

Let $K(i) = A(0,0,i,0)\,, i \in \mathbb{Z}$ denote the results given by Kunoth's program. A straight forward calculation yields

1. $A(j,i,l,k)$
   (a) $i = j$

(16a) $$A(j,j,l,k) = 2^{jn} K(l-k)$$

   (b) $i < j$

(16b) $$q = j - i \qquad A_{1,0}(p) = 2^n \sum_\nu \tilde{h}_\nu K(-\nu + p)$$

(16c) $$A(j,i,l,k) = 2^{in} A_{q,0}(l - 2^q k) \qquad A_{q,0}(p) = 2^n \sum_\nu \tilde{h}_\nu A_{q-1,0}(p - 2^{q-1}\nu)$$

   (c) $i > j$

(16d) $$q = i - j \qquad A_{0,1}(p) = 2^n \sum_\nu h_\nu K(\nu - p)$$

(16e) $$A(j,i,l,k) = 2^{jn} A_{0,q}(k - 2^q l) \qquad A_{0,q}(p) = 2^n \sum_\nu h_\nu A_{0,q-1}(p - 2^{q-1}\nu)$$

2. $B(j,i,l,k)$ (in this case $j \geqslant i$)

(17a) $$q = j - i \qquad B_q(p) = \sum_\nu g_\nu A_{q+1,0}(\nu + 2p)$$

(17b) $$B(j,i,l,k) = 2^{in} B_q(l - 2^q k)$$

3. $C(j,i,l,k)$ (in this case $i \geqslant j$)

(18a) $$q = i - j \qquad C_q(p) = \sum_\nu \tilde{g}_\nu A_{0,q+1}(\nu + 2p)$$

(18b) $$C(j,i,l,k) = 2^{jn} C_q(k - 2^q l)$$

4. $D(j,i,l,k)$
   (a) $i = j$

(19a) $$D(j,j,l,k) = 2^{(j+1)n} D_{0,0}(l-k) \qquad D_{0,0}(p) = \sum_\nu g_\nu \sum_\mu \tilde{g}_\mu K(\nu - \mu + 2p)$$

   (b) $i < j$

(19b) $$q = j - i \qquad D_{q,0}(p) = \sum_\nu g_\nu \sum_\mu \tilde{g}_\mu A_{q,0}(\nu - 2^q \mu + 2p)$$

(19c) $$D(j,i,l,k) = 2^{(i+1)n} D_{q,0}(l - 2^q k)$$

   (c) $i > j$

(19d) $$q = i - j \qquad D_{0,q}(p) = \sum_\nu g_\nu \sum_\mu \tilde{g}_\mu A_{0,q}(\mu - 2^q \nu + 2p)$$

(19e) $$D(j,i,l,k) = 2^{(j+1)n} D_{0,q}(k - 2^q l)$$

5. Application to Telegrapher's Equations

The simplest wave equations of electrodynamics, telegrapher's equations, write in normalized variables

$$\frac{\partial u(t,x)}{\partial t} = -\beta \frac{\partial i(t,x)}{\partial x} \tag{20a}$$

$$\frac{\partial i(t,x)}{\partial t} = -\beta \frac{\partial u(t,x)}{\partial x}\,, \tag{20b}$$

where $\beta \in \mathbb{R}, \beta > 0$ is the normalized speed of the travelling waves on the line, $u : \mathbb{R}_0^+ \times [0,1] \to \mathbb{R}$ the normalized voltage and $i : \mathbb{R}_0^+ \times [0,1] \to \mathbb{R}$ the normalized current. The initial boundary value problem $x \in \mathbb{R}, 0 \leqslant x \leqslant 1$, $t \in \mathbb{R}_0^+ = \{x \in \mathbb{R} : x \geqslant 0\}$ with the initial conditions $u_0(x) = u(0,x)$, $i_0(x) = i(0,x)$ and ideal electric, magnetic and periodic boundary conditions will be investigated.

We assume, that the solutions $u(t,x)$ and $i(t,x)$ can be written as a linear combination of our scaling function (and its translates) and our wavelet (and its translates and dilates) with time dependent coefficients, e.g.

$$u(t,x) = \sum_{k \in \mathcal{K}_{C-1}} u_k^{C-1}(t)\varphi_{C,k}(x) + \sum_{m=C}^{M} \sum_{k \in \mathcal{K}_m} u_k^m(t)\psi_{m,k}(x)\,. \tag{21}$$

$\mathcal{K}_m$ denote the appropriate sets of indices for each level and the superscript $C-1$ now indicates the scaling coefficients.

These expansions are inserted into telegrapher's equations, and then the resulting equations are tested with the biorthogonal scaling and wavelet functions. This leads to a system of ordinary differential equations

$$\frac{du_k^m(t)}{dt} = -\beta D_k^m I(t) \quad C-1 \leqslant m \leqslant M \tag{22a}$$

$$\frac{di_k^m(t)}{dt} = -\beta D_k^m U(t) \quad k \in \mathcal{K}_m\,, \tag{22b}$$

which has to be integrated (method of lines). $U(t)$, $I(t)$ denote the vectors of all coefficients $u_k^m(t)$, $u_k^m(t)$ of the expansions of $u(t,x)$ and $i(t,x)$, e.g.

$$I(t) = \left(i_{\min}^{C-1}(t), \ldots, i_{\max}^{C-1}(t), i_{\min}^{C}(t), \ldots, i_{\max}^{C}(t), \ldots, i_{\min}^{M}(t), \ldots, i_{\max}^{M}(t)\right)^T\,. \tag{23}$$

The subscripts min and max denote the minimum resp. the maximum of each set of indices (denoted by the superscript). The difference operators are given by the row vectors

$$\mathcal{D}_k^{C-1} = (A(C,C,l_{\min},k), \ldots, A(C,C,l_{\max},k), B(C,C,l_{\min},k), \ldots, B(M,C,l_{\max},k)) \tag{24a}$$

$$\mathcal{D}_k^{m} = (C(C,m,l_{\min},k), \ldots, C(C,m,l_{\max},k), D(C,m,l_{\min},k), \ldots, D(M,m,l_{\max},k)) \tag{24b}$$

where $C \leqslant m \leqslant M$ and k runs through the appropriate index sets.

6. Time Integration

For time integration, we use Nyström's method with $q = 0$ [10]. With this, equations (22) write

$$U(l+1) = U(l-1) - 2\Delta t \beta \mathcal{D} I(l) \tag{25a}$$

$$I(l+1) = I(l-1) - 2\Delta t \beta \mathcal{D} U(l) \tag{25b}$$

with

$$\mathcal{D} = (\mathcal{D}_{\min}^{C-1}, \ldots, \mathcal{D}_{\max}^{C-1}, \mathcal{D}_{\min}^{C}, \ldots, \mathcal{D}_{\max}^{M})^T\,. \tag{25c}$$

Of course, this is a simple multi-step scheme, which has to be started with a one-step method. We use the modified Euler scheme [10] which is also second order accurate

$$U(1) = U(0) - \Delta t \beta \mathcal{D}\left(I(0) - \frac{\Delta t}{2}\beta \mathcal{D} U(0)\right) \qquad C-1 \leqslant m \leqslant M \qquad k \in \mathcal{K}_m \tag{26a}$$

$$I(1) = I(0) - \Delta t \beta \mathcal{D}\left(U(0) - \frac{\Delta t}{2}\beta \mathcal{D} I(0)\right) \qquad l \in \mathbb{N} \qquad l \geqslant 1 \tag{26b}$$

TABLE 1. Bounds on the time step (three digits).

$(d,\tilde{d})$	$\gamma_{\max}$
(2,4),(3,3)	1.27
(2,6),(3,5),(4,4)	1.16
(2,8),(3,7),...,(5,5)	1.10
(2,10),(3,9),...,(6,6)	1.05
(2,12),(3,11),...,(7,7)	1.01
(2,14),(3,13),...,(8,8)	0.989
(2,16),(3,15),...,(9,9)	0.964
(2,18),(3,17),...,(10,10)	0.945

Apparently, (25) is the famous leap-frog scheme. $U(2l)$ depends neither on $U(2l+1)$ nor on $I(2l)$, only on $I(2l+1)$. For special choices of initial conditions, the calculation of only one part of the scheme suffices (either $U(2l+1)$ and $I(2l)$ or $U(2l)$ and $I(2l+1)$), but not for general initial conditions.

Equations (25) have to fulfill Neumann's condition to be stable [13]. The evaluation of the eigenvalues of the amplification matrix of (25) gives upper bounds for the time step (of course depending on the orders of the expansion functions used). All eigenvalues, i.e. their absolute value, must not be greater than one to avoid exponentially growing solutions. This evaluation, which has been carried out numerically —since the mask coefficients also have been calculated numerically— results in Tab. 1, where

$$\gamma := \Delta t \beta 2^{M+2} \text{ and } \gamma_{\max} := \max\{\gamma : \text{absolute values of all eigenvalues} \leqslant 1\}. \tag{27}$$

7. Numerical Experiments

As examples we present our results for periodic boundary conditions and the initial conditions

$$u_0(x) = \begin{cases} (x-0.4)^2(x-0.6)^2 & \text{for } 0.4 \leqslant x \leqslant 0.6, \\ 0 & \text{else} \end{cases} \tag{28a}$$

$$i_0(x) = \begin{cases} (x-0.4)^2(x-0.6)^2 & \text{for } 0.4 \leqslant x \leqslant 0.6, \\ 0 & \text{else.} \end{cases} \tag{28b}$$

For different combinations of electric and magnetic boundaries we did not find any difference to the periodic case. And as long as the initial conditions are smooth functions, the algorithm works very well. But if one tries e.g. a box function, the results deteriorate with a lot of oscillations, just the same situation like e.g. with finite elements (apart from some extraordinary cases with magic time steps).

First, let's have a look at the behaviour of the error with respect to time and as a function of the time step. For easy evaluation of the error, we choose $\beta = 1$ and $\Delta t = 1/K$, $K \in \mathbb{N}$, so that after K time steps, the solution should equal the initial conditions exactly. As voltage and current are the same, it suffices to look at the current. We define the error as

$$\varepsilon(nK\Delta t) = \frac{\|i(nK\Delta t, x) - i_0(x)\|_{M+1}}{\|i_0(x)\|_{M+1}}, n \in \mathbb{N} \text{ with } \|f(x)\|_n = \sqrt{2^{-n} \sum_{\nu=0}^{2^n-1} |f(2^{-n}\nu)|}. \tag{29}$$

To see the dispersion behaviour, we set $M = 7$, $C = 3$ and used the (3,3)-B-spline-wavelets. As time step, we used $\Delta t_{\max} = \gamma_{\max}/2^{M+2}$,$\Delta t_{\max}/2$,$\Delta t_{\max}/4$, $\Delta t_{\max}/8$,$\Delta t_{\max}/16$ and $\Delta t_{\max}/32$. Results can be seen in Fig. 2, left diagram. The curves correspond from top to bottom to the time step in descending order. To enhance readability, the curve for $\Delta t_{\max}/32$ was dashed.

Obviously, as one decreases the time step, the error becomes smaller and the error stays relatively small for a longer time. Note that e.g. for the biggest time step $t = 100$ means around 40,000 time steps and for the smallest around 1,200,000. The second interesting thing is, that decreasing the time step below a certain bound (here approx. $\Delta t_{\max}/16$) improves the solution only for a short time. We presume that from this property it might be possible to estimate the at least necessary highest resolution level to obtain a desired accuracy and the allowed maximum time step to achieve this.

FIGURE 2. Left: Time behaviour of the error for six different time steps (top to bottom: $\Delta t = \Delta t_{\max}, ..., \Delta t_{\max}/16$ (solid) and $\Delta t_{\max}/32$ (dashed)). Right: Current at $t = 0$, $t = 50$ and $t = 100$ (solid, dashed, dashdotted).

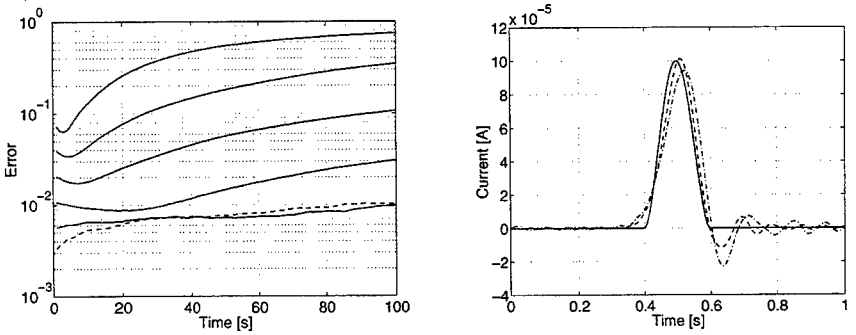


TABLE 2. Influence of thresholding and increasing the dual order: Scaling level 5 and 3 additional wavelet levels.

τ	0.0		10^{-6}		10^{-5}		10^{-4}		10^{-3}		10^{-2}	
$\tilde{d}$	ε	η	ε	η	ε	η	ε	η	ε	η	ε	η
3	0.0343	1.0	0.0343	0.996	0.0343	0.574	0.0326	0.391	0.0661	0.211	1.26	0.0469
5	0.0316	1.0	0.0316	0.984	0.0316	0.602	0.0313	0.355	0.0445	0.23	1.1	0.0508
7	0.0301	1.0	0.0301	0.98	0.0301	0.629	0.0299	0.32	0.0416	0.246	0.803	0.176
9	0.0289	1.0	0.0289	0.969	0.0288	0.633	0.029	0.391	0.038	0.262	0.604	0.227
11	0.0279	1.0	0.0279	0.973	0.0279	0.594	0.0282	0.348	0.0401	0.277	0.711	0.223
13	0.0273	1.0	0.0273	0.965	0.0273	0.664	0.0276	0.398	0.0307	0.211	0.753	0.234
15	0.0267	1.0	0.0267	0.949	0.0267	0.676	0.0269	0.355	0.0395	0.285	0.77	0.188

To visualize what the error means, the right hand side of Fig. 2 shows the initial current (solid), the current after 50 (dashed) and after 100 (dashdotted) seconds for $\Delta t = \Delta t_{\max}/2$, i.e. an error about 18% resp. 35%. The pulse travels to the right, so in contrast to FDTD, it has no tail behind but in front of it. So the high frequency components are faster than they should be. A rigorous characterization of the dispersion error is on its way.

To investigate the influence of thresholding, we define a sparsity coefficient

$$\eta = \frac{\text{number of nonzero coefficients}}{\text{total number of coefficients}}. \tag{30}$$

A simple thresholding procedure which sets every coefficient with absolute value smaller than $\tau \cdot I_{\max}$, $I_{\max} = \max\{|i_k^{C-1}| : k \in \mathcal{K}_{C-1}\}$ to zero, is applied.

To see the influence of the dual order —increasing the dual order should improve the compression, i.e. decrease our sparsity coefficient— we use again as primal order 3 and as scaling level 5, i.e. 32 scaling coefficients. We looked at two cases. First, three additional wavelet levels (i.e. a total number of 256 unknowns) and second, five additional wavelet levels (resulting in 1024 unknown coefficients). Results are presented in Tab. 2 and Tab. 3.

Increasing the dual order improves the error slightly, but not significantly. In both cases, thresholding up to a certain threshold only reduces the sparsity coefficient, but does not affect the error. In the first case a sparsity of approximately 25% seems to be obtainable, in the second case about 15%. Note that if one is interested only in a rough estimation of global parameters (such as S-parameters) and doesn't need high accuracy in the field values (like one would for calculating impedances of e.g. antennas), with five additional wavelet levels less than 2% of the expansion coefficients suffice to obtain an accuracy of 10% in the fields. Increasing τ above a certain number, increases the error dramatically. A precalculation of this bound $\tau_{\max}$ would be extremely valuable. Moreover, using more wavelet levels improves the sparsity coefficient. But unfortunately, $\tau_{\max}$ seems to depend on the number of wavelet levels. We presume that this is due to the simple thresholding procedure we applied, and conclude that it

TABLE 3. Influence of thresholding and increasing the dual order: Scaling level 5 and 5 additional wavelet levels.

τ	0.0		10^{-6}		10^{-5}		10^{-4}		10^{-3}		10^{-2}	
$\tilde{d}$	ε	η	ε	η	ε	η	ε	η	ε	η	ε	η
3	0.0105	1.0	0.0105	0.266	0.0106	0.206	0.0163	0.146	0.188	0.0176	1.29	0.00879
5	0.00962	1.0	0.00962	0.293	0.00966	0.211	0.0135	0.143	0.112	0.0225	1.26	0.00781
7	0.00912	1.0	0.00912	0.290	0.00913	0.192	0.0123	0.159	0.0957	0.0176	1.26	0.00781
9	0.00872	1.0	0.00871	0.291	0.00886	0.199	0.0116	0.179	0.103	0.0146	1.22	0.00781
11	0.00839	1.0	0.0084	0.300	0.0084	0.196	0.0109	0.195	0.107	0.0127	1.19	0.00879
13	0.00821	1.0	0.00821	0.299	0.0083	0.215	0.0112	0.181	0.106	0.0137	1.23	0.009772
15	0.00801	1.0	0.00801	0.301	0.00809	0.201	0.0116	0.158	0.107	0.0137	1.22	0.00781

is not good enough, especially if one wants to construct adaptive algorithms where the number of wavelet levels is not fixed.

8. Conclusion

We presented how to solve telegrapher's equations using biorthogonal B-spline-wavelets with a Petrov-Galerkin method using Nyström's method for time integration. The results shown validate our approach and point out, what has to be done to create more efficient algorithms. First of all, we conclude that it is unavoidable to implement fully adaptive algorithms to exploit wavelets as much as possible. Don't forget that using wavelets increases the length of the difference operators involved. This means that sparsity is the only key to reduce again the calculation time. Results by Göttelmann [4] indicate that one needs five to six wavelet levels to be computationally as efficient as using only scaling functions. Further we conclude that an appropriate thresholding procedure has to be invented to become independent of the number of wavelet levels used.

Acknowledgement

We would like to thank Siemens AG for making this work possible by generously granting an Ernst-von-Siemens scholarship.

References

1. A. Cohen, I. Daubechies, and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math. **45** (1992), 485–560.
2. W. Dahmen, *Wavelet and multiscale methods for operator equations*, Acta Numerica (A. Iserles, ed.), vol. 6, Cambridge University Press, 1997, pp. 55–228.
3. G. Erlebacher, M. Y. Hussaini, and L. M. Jameson (eds.), *Wavelets: Theory and application*, New York, Oxford University Press, 1996.
4. J. Göttelmann, *Adaptive multiscale algorithms for the spherical advection equation*, Internet, 1997, part of PhD thesis.
5. K. Goverdhanam and L. P. B. Katehi, *Applications of multiresolution based FDTD multigrid*, IEEE MTT-S Int. Microwave Symp. Digest, June 8–13, Denver, June 1997, pp. 333–336.
6. M. Krumpholz and L. P. B. Katehi, *MRTD: New time domain schemes based on multiresolution analysis*, IEEE Transactions on Microwave Theory and Techniques **44** (1996), no. 4, 555–571.
7. A. Kunoth, *On the fast evaluation of integrals of refinable functions*, Wavelets, Images and Surface Fitting (P. J. Laurent, A. Le Méhauté, and L. L. Schumaker, eds.), AKPeters, Boston, 1994, pp. 327–334.
8. ———, *Computing refinable integrals — documentation of the program — version 1.1*, Tech. Report ISC-95-02-MATH, Institute for Scientific Computation, Texas A&M University, 1995.
9. R. Robertson, E. Tentzeris, M. Krumpholz, and L. P. B. Katehi, *MRTD analysis of dielectric cavity structures*, IEEE MTT-S Int. Microwave Symp. Digest, June 17–21, San Francisco, 1996, pp. 1861–1864.
10. J. Stoer and R. Bulirsch, *Numerische Mathematik 2*, 3. ed., Springer-Verlag, Berlin, 1990.
11. E. Tentzeris, A. Cangellaris, and L. P. B. Katehi, *Space/time adaptive meshing and multiresolution time domain method (MRTD)*, Proc. ACES, March 1997, pp. 1509–1514.
12. E. Tentzeris, M. Krumpholz, and L. P. B. Katehi, *Application of MRTD to printed transmission lines*, IEEE MTT-S Int. Microwave Symp. Digest, June 17–21, San Francisco, 1996, pp. 573–576.
13. J. W. Thomas, *Numerical partial differential equations: Finite difference methods*, Springer-Verlag, New York, 1995.
14. A. Ware, *Discrete projections onto wavelet subspaces*, Numerical Analysis Report NA-97/04, Durham University, August 1997.

E-mail: aidam@hft.e-technik.tu-muenchen.de, russer@hft.e-technik.tu-muenchen.de

NEC Acceleration by the Wavelet Matrix Transform

Yee Hui Lee and Yilong Lu†
School of Electrical and Electronic Engineering, Nanyang Technological University
Singapore 639798, Republic of Singapore
† Fax: (+65) 792 0415, E-mail: eylu@ntu.edu.sg

Abstract— **The Numerical Electromagnetic Code (NEC) is one of the most popular tools used for electromagnetic simulation of wireframe structures. The application of NEC is often limited to small to medium sized problems due to its dense matrix nature. In this paper, an approach by using the wavelet transform to increase the efficiency and capability of NEC is presented. In the approach, a sparse moment matrix equation can be produced and solved by efficient sparse solver instead of solving full matrix equation in the original NEC. Under close examination, structures with less singularities are found to have much better accuracy and higher compression rates.**

I. Introduction

Numerical Electromagnetic Code (NEC) [1] is probably still the most popular tool for modelling and analysis of electromagnetic (EM) response of complex wire-frame metallic structures. NEC uses the method of moment (MoM) together with the Lower and Upper triangular factorization method (LU) to solve the full matrix equation. This makes NEC extremely memory and computational time intensive when the problem is large. It is formidable for NEC to treat problems with more than 3000 unknowns even on high-end workstations and low-end supercomputers. In engineering applications, many problems are large and complicated with unknowns much more than this 3000. To solve large problems by NEC has always been a challenging task for computational electromagnetic researchers.

In recent years, there has been growing interest in applying wavelets to EM problems. The wavelet transform method (WTM) [2] has been developed using the translating and dilating of a suitable basis function, known as the mother wavelet. This mother wavelet then undergoes the decomposition and reconstruction algorithms producing the wavelet transform matrix. In this matrix, each row stands for a wavelet basis in N-dimensional wavelet vector space. The translation of the highest resolution wavelet makes up half of the basis set and the next highest resolution makes up a quarter of it. This goes on down the hierarchy.

In this paper, the WTM is applied to NEC so as to transform the full impedance matrix into a highly sparse matrix and then solved using a sparse solver. This would largely reduce the memory required for the storage of the impedance matrix and also cut down computational time by using the sparse solver instead of the LU factorization technique.

The modified version of NEC (NEC-WTM) is tested by a number of examples. The compression rate and accuracy is then compared with the results obtained from the original version 2 of NEC (NEC-2). Some limitations of the NEC-WTM has also been examined.

II. The Wavelet Matrix Transformation

In the NEC, like many other MoM based codes, the most time consuming part in the computation is to solve the moment matrix equation

$$[Z] \cdot [I] = [V], \tag{1}$$

where $[Z]$ is a full moment matrix, $[I]$ is current intensity related unknown column vector to be solved, and $[V]$ is the known source related column vector. In NEC, the LU decomposition method which forms the bulk of the CPU time required to solve the full matrix equation is $O(N^3)$ for large matrix order N. For large problem, due to the memory required for the storage of all the elements in the full matrix, the out-of-core operation has to be performed where the $[Z]$ matrix had to be stored in 4 sequential access files each of the size of the full matrix. This manipulation of the full matrix is therefore an extremely storage and CPU demanding process.

In order to increase the efficiency and capability for larger problems, we have applied the recently developed wavelet transform method (WTM) [2] into the NEC. The wavelet matrix transformation will produce a sparse moment matrix similar to that obtained by using basis expansion in MoM. By using the sparse wavelet transform matrix $[\tilde{W}]$, the wavelet matrix transformation can be carried out on (1) as follows:

$$[\tilde{W}][Z][\tilde{W}]^T \cdot ([\tilde{W}]^T)^{-1}[I] = [\tilde{W}][V], \tag{2}$$

where $[\tilde{W}]^T$ is the transpose of the wavelet transform matrix $[\tilde{W}]$. After this transformation, we now obtain a new matrix equation.

$$[Z'] \cdot [I'] = [V'], \tag{3}$$

where , $[Z'] = [\tilde{W}][Z][\tilde{W}]^T$, $[I'] = ([\tilde{W}]^T)^{-1}[I]$, and $[V'] = [\tilde{W}][V]$. Next, a suitable threshold value τ, has to be chosen. We can then discard the elements in the matrix $[Z']$ whose magnitudes are smaller than $\tau \cdot m$, where m is the largest magnitude of the matrix elements. The threshold value τ need to be well chosen so as to balance the computational efficiency and accuracy of the approximate solutions.

The resulting matrix from this process is now a sparse one. The sparse matrix can then be solved much more efficiently by a sparse solver. Solving a sparse matrix requires $O(NlogN)$ operations, where N is the number of unknowns on the structure. This is much more efficient as compared to $O(N^3)$ for a full matrix solution in the original version of NEC-2. Once the $[I']$ is solved, matrix $[I]$ can then be reconstructed by

$$[I] = [\tilde{W}]^T[I']. \tag{4}$$

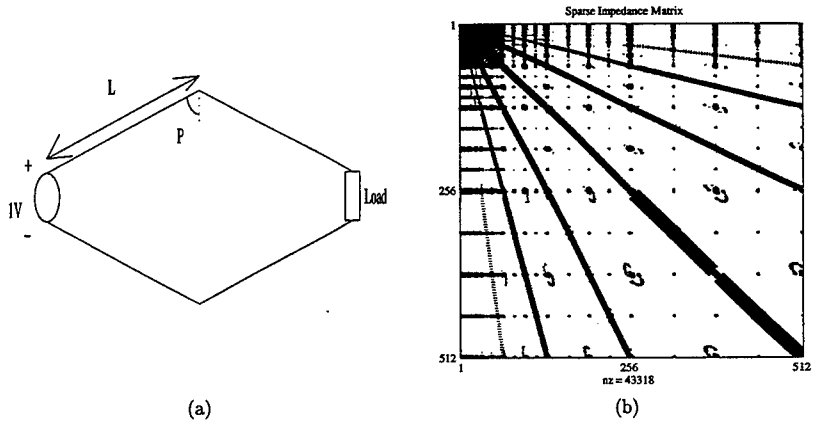
This process would therefore be an efficient method of solution which may considerably reduces the computational time and storage requirements.

III. Implementation and Experiment

The wavelet transform method was implemented into the NEC code and then the NEC-WTM code was used to test various examples and to study its effectiveness. The results obtained by the NEC-WTM code were compared with the solutions obtained by the original NEC-2.

A. Example: Rhombic Antenna

As a numerical example, the structure of a rhombic antenna [6] above a perfectly conducting ground plane was simulated. The geometry of the considered problem is shown in Fig. 1(a). The structure shows a terminated rhombic antenna with each leg (L) of 6 wavelength, the angle $P = 70^o$ (indicated in Fig. 1(a)), a height of 1.1 wavelength above a perfectly conducting ground. The frequency used in the simulation is 300 MHz and the excitation is a 1 V voltage source. The terminating resistance (load) is 800Ω.



(a) (b)

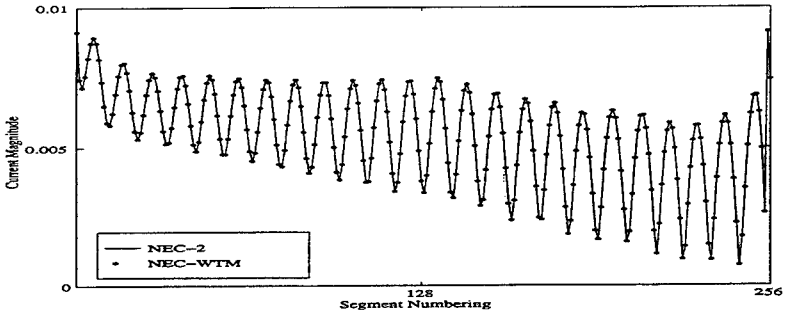
Fig. 1. (a) Terminated rhombic antenna (b) The sparsity pattern of the matrix $[Z']$ after the wavelet transform with a threshold value of $\tau = 10^{-8}$.

The sparsity pattern of the $[Z']$ matrix with a threshold value $\tau = 10^{-8}$ is shown in Fig. 1(b). The black dots show the remaining nonzero elements. In this particular case, only 43318 elements are left out of the total 262144 elements. This is 16.52% of the original full 512 by 512 matrix.

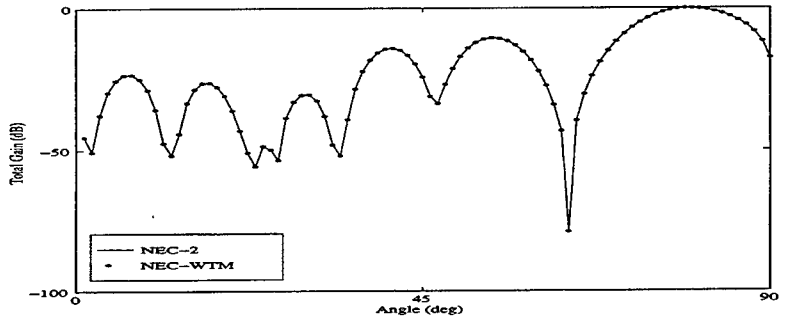
Fig. 2 compares the full matrix solution from the original NEC-2 and the sparse matrix solution from the NEC-WTM with a threshold of $\tau = 10^{-8}$.

From the comparison, it can be seen that the approximate solutions by the NEC-WTM are very close to those by the original NEC-2, showing a high accuracy even when the compression rate is very high. In this case, only 16.52% of the elements in the $[Z']$ are left. From this numerical example, it can be shown that with a very high compression rate, the accuracy of the approximate solution can be very high. With this high compression rate, computational time would be drastically reduced, showing that the matrix equation can be solved more efficiently while maintaining good accuracy. It should be noted that, the larger the size of EM problem, the more effective the wavelet matrix transform method.

However, there are some limitations to the method. It has been observed that this method is sensitive to singularity. The more the number of singularity, the less effective the method. Therefore, this application is more effective for large and smooth problems. Another problem faced is the fact that, the NEC-WTM is only applicable to problems with an impedance



(a)



(b)

Fig. 2. Comparison of the sparse matrix solution using the NEC-WTM ($\tau = 10^{-8}$, compression rate 16.52%) and the full matrix solution by the original NEC-2. (a) Current distribution on two of the four legs (b) Total gain pattern.

matrix size of 2^N where N is an integer. This problem will be solved most efficiently by using the adaptive segmentation approach dealt with in another paper [7].

B. Example: Elliptical Scatterer

The next numerical example was designed to examine the relation between the compression rate and the accuracy of both the current distribution and scattering power pattern.

This example involves an elliptical scatterer. The geometry of the considered structure is shown in Fig. 3(a). The figure shows an elliptical scatterer under the incidence of a plane wave coming down at 30^o from the z-direction in the x-y plane. The frequency used in this simulation is 3 GHz.

The sparsity pattern of the $[Z']$ matrix with a threshold value $\tau = 10^{-6}$ is shown in Fig. 3(b). For this example, 76642 elements out of a total of 1048576 elements were left. This is 7.31% of the original full 1024 by 1024 matrix.

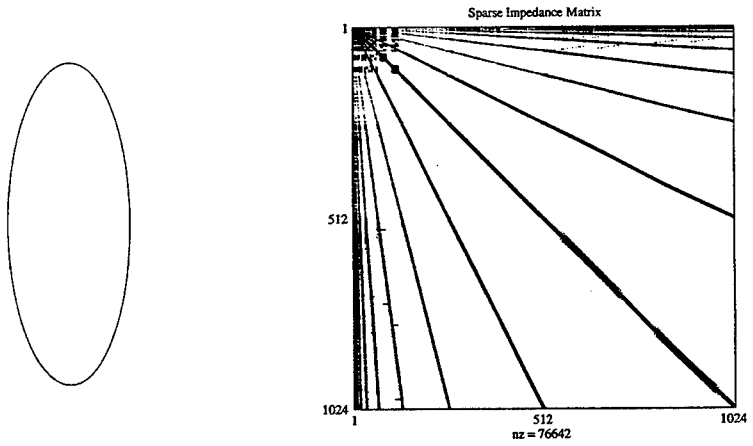


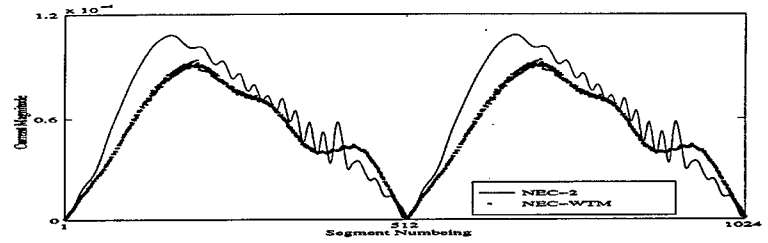
Fig. 3. (a) Elliptical Scatterer, (b) The sparsity pattern of the matrix $[Z']$ after the wavelet transform with a threshold value of $\tau = 10^{-6}$.

A comparison was done for four different threshold values ($\tau = 10^{-4}$, 10^{-5}, 10^{-6}, 10^{-7}). The compression rates corresponding to the four cases are 1.37%, 3.36%, 7.31% and 12.45%, respectively.

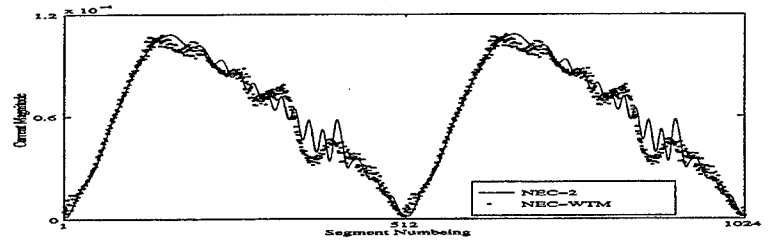
Fig. 4 shows the comparisons between the full matrix current intensity solution and the four sparse matrix solutions. With a populated rate of 1.37%, the current intensity deviates from the true value. As the populated rate increases to 7.31%, it can be noted that the result gradually approaches the true value with many points oscillating around the actual solution. At the populated rate of 12.45%, the error between the NEC-2 and the NEC-WTM solutions is negligible. This again justifies that the NEC-WTM can produce results of high accuracy with a high compression rate.

Fig. 5 shows the scattering power in the x-z plane. A review of Fig. 4 and 5 shows that, for a higher compression rate of 1.37%, although the NEC-WTM solution of the current intensity is not very accurate, the scattering power pattern obtained seems more acceptable. From Fig. 5(c), the results shows that when the compression rate is 7.31%, the accuracy of the NEC-WTM solution is sufficiently accurate. This shows that for users who are interested only in the scattering power pattern, a higher compression rate may be used to achieve results with sufficient accuracy. The use of a Z matrix of less populated rate will then translate into less computation time and memory space required.

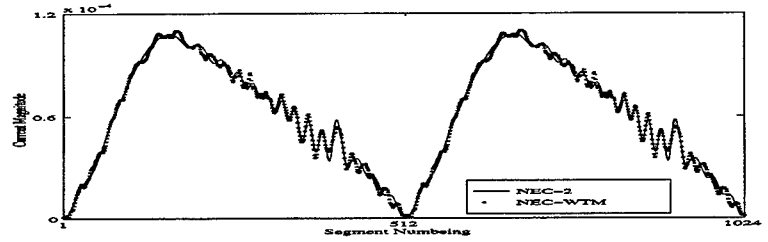
Table I shows a comparison between the CPU time required to solve the sparse matrix using a sparse solver as compared to the LU factorization method used in the original NEC-2. This was done on the matrices from this example. As seen from Table I, for the populated rate of 3.36%, the CPU time required to solve this 1024 by 1024 matrix is approximately 29 times faster than that of the original time required. This shows a great improvement in CPU time required for the NEC-WTM.



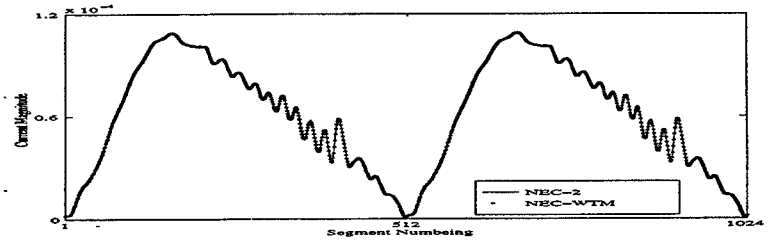
(a)



(b)

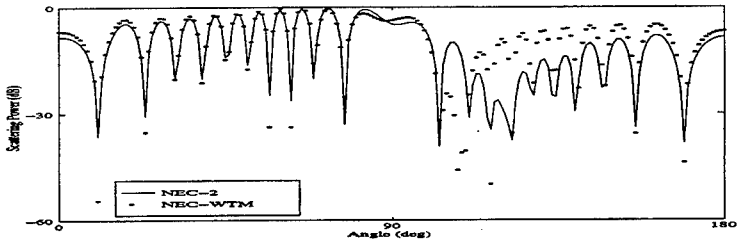


(c)

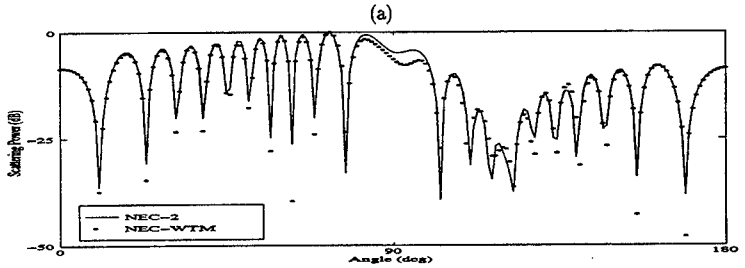


(d)

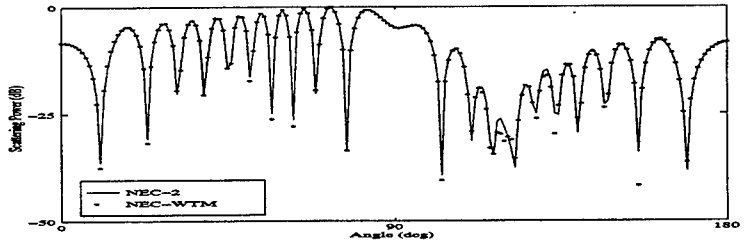
Fig. 4. Comparison of current intensity plots with compression rates of (a) 1.37% (b) 3.36%, (c) 7.31% and (d) 12.45% with the full matrix, respectively.



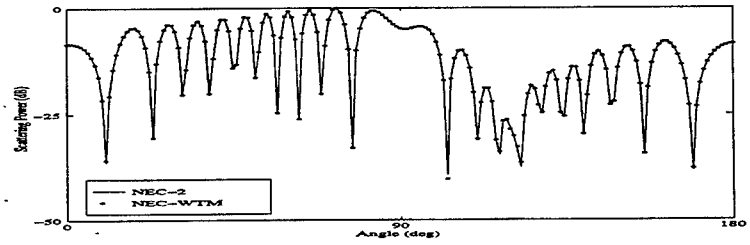
(a)



(b)



(c)



(d)

Fig. 5. Comparison of scattering power plots with compression rates of (e) 1.37% (f) 3.36%, (g) 7.31% and (h) 12.45% with full matrix for x-z plane, respectively.

TABLE I
SPARSITY AND CPU TIME

Populated Rate	CPU Time (sec)	No. Times Faster
100%	374.7	1.00
12.5%	73.62	5.09
7.31%	68.78	5.45
3.36%	12.98	28.86
1.37%	6.75	55.51

In this second example, the effect of threshold value on the accuracy of the results has been examined. The effects of threshold values to the accuracy of both the current intensity and scattering power pattern were investigated. It can be concluded that, depending on the required results, a suitable threshold should be chosen so as to balance the computational cost and the accuracy required.

IV. CONCLUSIONS

In this paper, acceleration of NEC computation by using the wavelet transform method is proposed and studied. By using the wavelet transform method, instead of solving a full matrix equation in the original NEC-2 at a computation cost of $O(N^3)$, a sparse matrix can be obtained and solved efficiently by sparse solvers with an operational cost of $O(NlogN)$,

It is shown that with the NEC-WTM approach, one can obtain accurate enough approximate solutions with a very sparse matrix equation. The effectiveness and accuracy of the method are shown by numerical examples. Some of the limitations to this method has been examined.

REFERENCES

[1] G. J. Burke and A. J. Poggio, *Numerical Electromagnetic Code (NEC) - Method of Moments,* Lawrence Livermore National Laboratory Rept. UCID-18834, January 1981.

[2] Z. Xiang and Y. Lu, "An Effective Wavelet Matrix Transform Approach for Efficient Solution of Electromagnetic Integral Equations," *IEEE Trans. Antennas Propag.*, vol.45, pp.1205-1213, Aug. 1997.

[3] C. K. Chui, *An Introduction to Wavelets,* Boston MA: Academic Press, 1992.

[4] G. Kaiser, *A Friendly Guide to Wavelets,* Boston: Birkhauser, 1994.

[5] M. N. O. Sadiku,, *Numerical Techniques in Electromagnetics,* Boca Raton: CRC Press, 1992.

[6] J. D. Kraus, *Antennas,* McGraw-Hill, 1988.

[7] Y.H. Lee and Y. Lu, "Adaptive Segmentation Algorithms for Optimal NEC Modelling of Wire-Grid Structures," submitted to *The 14th Annual Review of Progress in Applied Computational Electromagnetics,* Monterey CA, March 1998.

Adaptive Segmentation Algorithms for Optimal NEC Modelling of Wire-Grid Structures

Yee Hui Lee and Yilong Lu†
School of Electrical and Electronic Engineering, Nanyang Technological University
Singapore 639798, Republic of Singapore
† Fax: (+65) 792 0415, E-mail: eylu@ntu.edu.sg

Abstract— **The realistic modelling and simulation of wire-frame structures is an important part of computational electromagnetics. The Numerical Electromagnetic Code (NEC) is one of the most popular tools for electromagnetic simulation of wire-grid structures. The application of NEC is often limited to small to medium sized problems due to its dense matrix nature. In order to perform a simulation using NEC, it is important to model the structure accurately so as to obtain realistic simulation results. Adaptive segmentation algorithms have been developed with the aim of generating optimal NEC models so as to reduce redundancy and computational cost. Some numerical examples are done to show its validity.**

I. Introduction

Numerical Electromagnetic Code (NEC) [1] is a useful and popular tool for the modelling and analysis of electromagnetic (EM) response of complex wire-frame metallic structures. NEC is based on, the method of moment (MoM) and full matrix equation solver, and thus, it is extremely hungry for memory and computational time for large-size problems. In engineering application, many problems are large and complicated. To reduce the computational cost while maintaining the accuracy has always been a challenging problem.

In order to get results of high accuracy, the modelling of a structure using the optimal number of segmentation is critical. Therefore, the accurate modelling of a complex structure for the use of NEC itself has proven to be a laborious task. Much attention has been given to the accurate modelling of a wire-grid structure.

In this article, adaptive segmentation algorithms are presented to allow optimal segmentation for NEC models. This reduces redundancy yet maintains good accuracy. An optimized NEC model will therefore be translated into the reduction of computational time and memory space required. There are a number of ways in which adaptive segmentation can be performed. From previous experiences in the simulation of complex structures, adaptive distance and adaptive current segmentations are found to be more effective. Therefore, these algorithms have been developed so as to allow optimized segmentation for complex models.

To further enhance the usefulness of these algorithms, they have been modified such that an optimized model with an impedance matrix of 2^N can be produced so as to be used with the newly developed NEC-WTM [4].

The adaptive segmentation algorithms have aided in the optimal segmentation and flexible matrix size adjustment for NEC modelling.

II. Adaptive Segmentation

Before a simulation can be done using NEC, the NEC wire-grid model of the structure has to be carefully and accurately modelled. The accurate modelling of a structure for the use of NEC may be a frustrating job. It has been found that, very often, for a large complex structure, wires which are situated far away from the active element in the structure have little or negligible effect on the overall radiation pattern. Therefore, it is very useful for users to be able to perform some kind of adaptive segmentation. We have developed two algorithms that will allow users to deal with the above mentioned problem.

A. Adaptive Distance Segmentation

In NEC, the electric field integral equation (EFIE) and the magnetic field integral equation (MFIE) are used for the evaluation of electromagnetic responses of thin wires and surfaces respectively. The form of EFIE used in NEC for thin wires follows from an integral representation for the electric field of a volume current distribution $\mathbf{J}$,

$$\mathbf{E}(\mathbf{r}) = -jk_0 Z_0 \int_V \mathbf{J}(\mathbf{r}') \cdot \overline{\overline{G}}(\mathbf{r}, \mathbf{r}') dV' \tag{1}$$

where $\mathbf{E}$ is the electric field intensity vector, $k_0 = \omega\sqrt{\mu_0 \epsilon_0}$ the wave number in free space, $Z_0 = \sqrt{\mu_0/\epsilon_0}$ the wave impedance in free space, and $\overline{\overline{I}}$ the unit dyadic, $\overline{\overline{G}}$ the free space dyadic Green's function defined by

$$\overline{\overline{G}}(\mathbf{r}, \mathbf{r}') = \left(\overline{\overline{I}} + \frac{1}{k_0^2}\nabla\nabla'\right) \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|)}}{4\pi|\mathbf{r}-\mathbf{r}'|} \tag{2}$$

From (1), it can be seen that, if the observation point $\mathbf{r}$ is at a large distance away from the source point $\mathbf{r}'$, the effect of this point on the overall electric field will be reduced significantly. However, the effect of these points can not be totally ignored. With reference to the active source point, three field regions are first identified with R_1 being the outer boundary to the first region and R_2 the outer boundary of the second region ($R_1 < R_2$). For each region, a particular number of segments per wavelength is specified.

Range of Region	*Seg./λ*
$\lvert\mathbf{r}-\mathbf{r}_0\rvert < R_1$	N_1
$R_1 < \lvert\mathbf{r}-\mathbf{r}_0\rvert < R_2$	N_2
$\lvert\mathbf{r}-\mathbf{r}_0\rvert > R_2$	N_3

In the above, $N_1 > N_2 > N_3$ and $\mathbf{r}_0$ is the active element. This results in an adaptive segmentation algorithm, where the number of segments which are within a particular range to the active antenna would be given more segments than those located far away from the active antenna. By doing so, a more accurate model of the structure can be generated. If more then one active element is present in the structure, every active element will be taken into consideration before the adaptive distance segmentation is to be performed. With this algorithm, unwanted segmentation can be avoided thus eliminating redundancy. Therefore, the size of the matrix is reduced to produce a better model. This is especially useful for the modeling of large complex structure where the controlling of segmentation can be quite

tedious and the matrix size generated large. By reducing the size of the matrix, the memory and computational time required automatically reduce.

B. Adaptive Current Segmentation

For some cases, due to the complexity of the structure and the coupling between wires in the overall structure, it is difficult to estimate the current intensity on different parts of the structure. For such structures where higher accuracy results are required, a rough model with minimum number of segments can be used to simulate an estimation of the final results. By using this rough estimated result, an accurate model can then be generated using the adaptive current segmentation algorithm. This algorithm reads in the estimated results and then refines the segmentation of the model according to this estimated results. At a region with high current intensity, segmentation is increased. Whereas for regions with low current intensity, segmentation will be decreased. This method is slightly more time consuming than the previous, however, the resultant model can give results of higher accuracy. Similarly, this method can avoid redundancy in the total number of unknown used, therefore reducing memory space and computational time.

C. Reshaping Matrix Size Using Adaptive Segmentation

Besides the above mentioned advantages, the algorithms can allow users to reshape the matrix into the desired size. The above two algorithms have also been developed so as to overcome the bottle-neck which exists in applications of the wavelet transform method (WTM) to NEC. The NEC-WTM is only able to handle matrix of size 2^N. Therefore, these algorithms has been developed such that the segmentation would be adjusted to have exactly 2^N segments and used by the NEC-WTM.

III. Implementation and Experiment

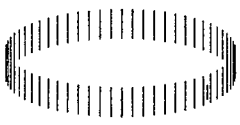
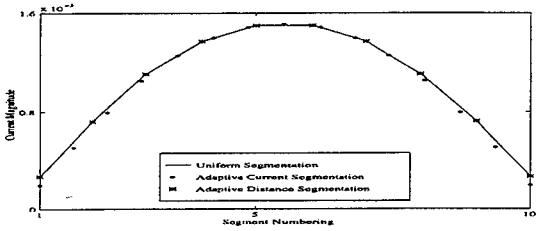


Fig. 1. A Dipole by a Large Cylinder

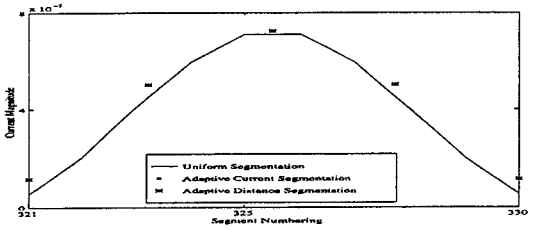
Using the newly developed adaptive segmentation algorithm, results from various simulations were compared so as to verify the usefulness of the algorithms.

A. Example: Dipole by Cylinder

Fig. 1 shows the structure used for this example. The structure shows a simple half-wavelength dipole located beside a conducting cylinder (modelled using verticle wire-grids). Due to the large radius of the cylinder, wire-grids at the far end of the cylinder (located furthest away from the active element) has little effect on the overall simulation results.



(a)



(b)

Fig. 2. Comparison of current intensity of 10 segments per wavelength for whole structure and the two cases of adaptive segmentation for (a) nearest wire (b) furthest wire from active antenna.

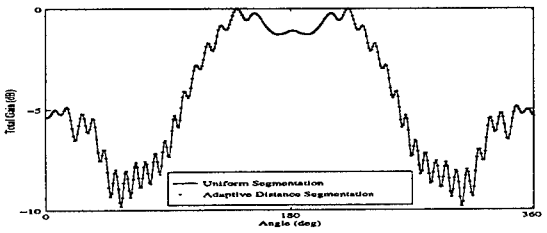
Simulation was done on the same structure with different number of segmentation. The first simulation was done on a structure with a standard of 10 segments per wavelength throughout the entire structure. The total number of segmentation for this structure is 655. This will act as a reference case. Next, using the adaptive segmentation on the same structure, two other cases of the same structure were produced. One according to distance (Adaptive Distance Segmentation), and another according to current distribution from previously simulated results (Adaptive Current Segmentation). The first case has a total of 412 segments. Adaptive segmentation according to the distance of each element from the active element, has been done on the structure. The next case has a total of 411 segments. Adaptive current segmentation was performed.

As can be seen from Fig. 2, the current distribution on the wire nearest and furthest from the active dipole for all three cases are very close. Fig. 3 shows the radiation pattern of each adaptive segmentation case as compared to the reference case. Fig. 3(a) shows adaptive distance segmentation while Fig. 3(b) shows adaptive current segmentation.

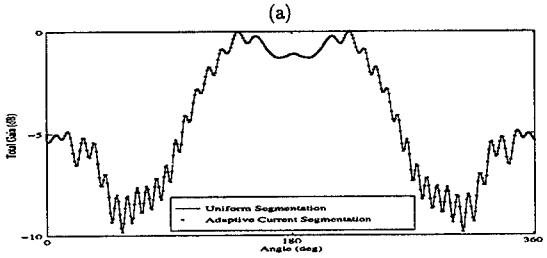
From the above, it can be concluded that, the number of segmentation required for this structure can be reduced from 655 to about 411 keeping the same accuracy. This shows that about 37% of the total segments used can become redundant if appropriate segmentation has been done.

B. Example: Monopole on Car

The next numerical example (Fig. 4) shows a monopole antenna mounted on a structure of a car [3]. The frequency of simulation is 300 MHz. Similarly, both adaptive distance and adaptive current segmentation were done on this model. The resulting comparison of the



(a)



(b)

Fig. 3. Comparison of total gain of 10 segments per wavelength for whole structure and (a) adaptive distance segmentation (b) adaptive current segmentation.

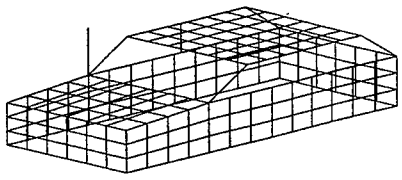
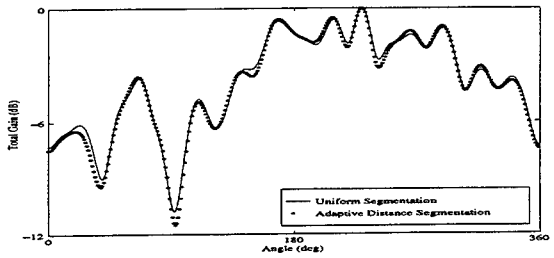


Fig. 4. A Monopole mounted on a Car

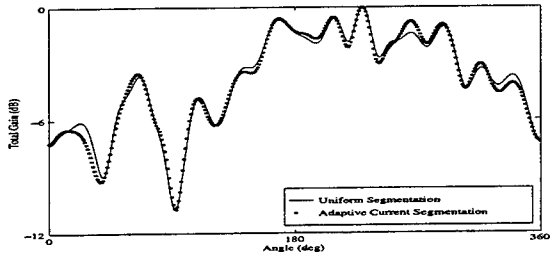
total gain for both cases are as shown in Fig. 5.

For this structure, the reference case is a structure with uniform segmentation of 15 segments per wavelength. This gives a total of 1677 segments for the entire structure. When adaptive distance segmentation was performed on the structure, the total number of segments was reduced to 1241. This reduction in segmentation is obtained with minimal difference in the total gain result obtained before and after the reduction as shown in Fig. 5(a). Similarly, when adaptive current segmentation is performed, the total number of segments was reduced to 1253 and yet maintaining total gain results of high accuracy. This is a 25% reduction in segmentation. This example further shows that adaptive segmentation is effective in avoiding redundancy during segmentation of a wire-frame structure.

From these examples, we can see the usefulness of adaptive segmentation algorithms. By avoiding redundancy and thus reducing matrix size, memory space and computational time can be reduced. These two algorithms have also been modified so that any structure can easily be segmented such that the total number of segment becomes 2^N. This would allow



(a)



(b)

Fig. 5. Comparison of total gain of 15 segments per wavelength for whole structure and (a) adaptive distance segmentation (b) adaptive current segmentation.

the use of NEC-WTM on any structure instead of limiting its use to structures with total segment of 2^N.

IV. Conclusions

In this paper, adaptive segmentation algorithms have been developed. One does adaptive segmentation according to the distance of each segment from the active element. Another requires a rough simulation result so as to refine the final optimized segmentation of the model. Algorithms have also been developed to reshape the impedance matrix for other applications. These adaptive segmentation algorithms have been developed for two purposes. Firstly, with these algorithms, an optimized model of a complex structure can be generated and thus avoiding the use of redundant segmentation. Secondly, the bottle-neck which exist in the NEC-WTM can be eliminated so as to improve the efficiency of the numerical electromagnetic code.

References

[1] G. J. Burke and A. J. Poggio, *Numerical Electromagnetic Code (NEC) - Method of Moments,* Lawrence Livermore National Laboratory Rept. UCID-18834, January 1981.

[2] M. N. O. Sadiku, *Numerical Techniques in Electromagnetics,* Boca Raton: CRC Press, 1992.

[3] *Wire Grid Modelling for NEC* by University of Stellenbosch.

[4] Y.H. Lee and Y. Lu, "NEC Acceleration by the Wavelet Matrix Transform," submitted to *The 14th Annual Review of Progress in Applied Computational Electromagnetics,* Monterey CA, March 1998.

SESSION 18:

HF-UHF PRACTICAL ANTENNA TOPICS

Chairs: W.P. Wheless, Jr. and N. Cohen

Comparison of Shipboard HF Transmit Fan Characteristics NEC versus Scale-Model Measurements

Keith Lysiak
Signal Exploitation
and Geolocation Division
Southwest Research Institute
San Antonio, TX 78238

LCdr Perry Dombowsky
Directorate Maritime Ship Support
National Defense Headquarters
Hull, Quebec, Canada K1A 0K2

Abstract

In support of the Canadian IROQUOIS Class HF Replacement Project (HFRP), HF antenna impedances and patterns were calculated using the Numerical Electromagnetics Code Version 2 (NEC2) [ref.1]. These results were compared to brass scale-model data collected at the SwRI rotary test facility. In this paper, the HF transmit fan antenna is discussed as it is an ideal candidate for NEC modeling applications. The results of the NEC modeling and measured impedance and pattern data for the HF transmit fan antenna are presented herein. The computed and measured results compare very favourably.

1. Introduction

The Canadian Department of National Defense (DND) is currently identifying new HF communication antenna types and locations on the DDH280 IROQUOIS class ships' as part of its ongoing HF Replacement Project (HFRP). The purpose of this project is to provide these ships with additional HF communication capabilities. The major task of the Signal and Geolocation Division at SwRI was to make scale-model range measurements to predict antenna patterns at proposed locations on the ship's topside. Since the locations of several antennas could be varied, SwRI proposed using NEC to guide the choice of such locations. The current HF transmit fan antenna was to remain fixed throughout the reconfiguration. It was therefore selected for NEC modeling and used as a benchmark to provide confidence in the NEC results. This paper describes the computed and measured results for the HF transmit fan antenna.

2. HF Transmit Fan Antenna

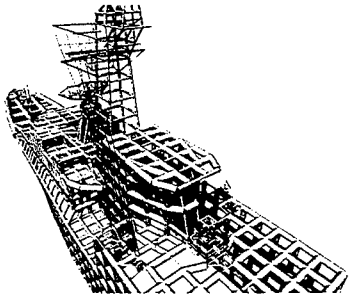
The HF transmit antenna is a large wire rope antenna extending from the ship's hanger top to the main mast. The antenna has two symmetrical halves with a single feed point at the top-center near the mast. Its frequency of operation is 2-9 MHz. The antenna, in its present configuration, does not have a matching network. Antennas similar to this are found on most large navy ships. They are relatively well matched to 50 ohms and have a fairly omni-directional radiation pattern.

3. NEC Model

A wire grid model of the full IROQUOIS class ship was developed for analysis on a Pentium Pro 200 with 256 Mbytes of memory. To balance the processing time and model resolution given this computing resource, the model was gridded for 15 MHz. This yielded a segment spacing of typically 2 meters and a segment radius of approximately 0.32 meters. The bow and stern were built with a coarser gridding to reduce the number of wire segments. The full model contains approximately 3000



Full NEC wire grid model



Detail of HF transmit fan antenna

segments and requires 150 Mbytes of memory. Each frequency run requires about 1.6 hours. NEC-Win Pro software and NEC version 2 were used.

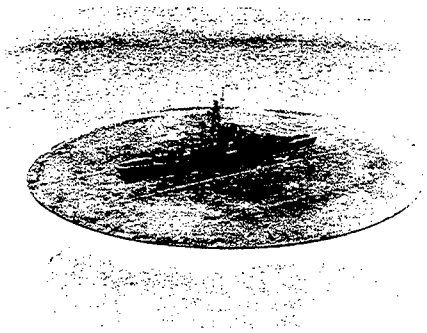
The main hull sections were gridded using the Structure Interpolation and Gridding (SIG) program [ref.2]. This program automatically grid surfaces based on a set of cross section contours. About 50 percent of the model are symmetrical and were therefore produced with the GX (reflection in coordinate planes) card in NEC [ref. 3]. Although this makes building the model easier, it does not reduce run time or memory because the entire model is not symmetrical.

Three-dimensional visualization of the NEC model was accomplished with POV-Ray for Windows. The ability to visualize the wire segments as three-dimensional cylinders greatly enhance the ability to refine the wire grid model. It provided the opportunity to visually check wire radii to ensure that they met the NEC modeling guidelines.

The HF transmit fan was relatively easy to model. It is secured to the deck at the lower ends by standoff poles and to the mast through insulators at the upper end. The feed point is simple both in practice and in the model.

4. Scale-Model Measurements

A 1/48th brass scale model of the DDH 280 IROQUOIS class ship was used to make antenna impedance and pattern measurements. The measurements were conducted at the SwRI rotary test facility. The test facility consists of a rotating copper platform surrounded by a 400-foot diameter radial ground screen. RF signals can be generated at elevation angles of 0 to 60 degrees while the rotating platform provides 360 degrees of azimuthal positioning. All RF and control cables run underground to an equipment building approximately 200 feet from the rotator's center. While measuring the outputs of scale model antennas as the ship is rotated, amplitude (and, if desired, phase) data is recorded relative to a fixed reference antenna. Eight RF channels are available for simultaneous data collection of up to eight scale model antennas at a time.

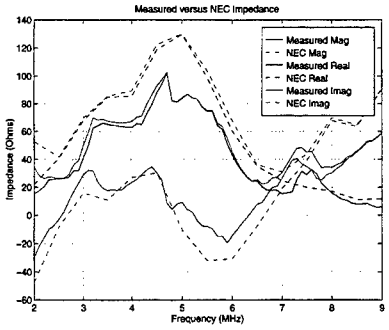


SwRI scale-model rotary test facility

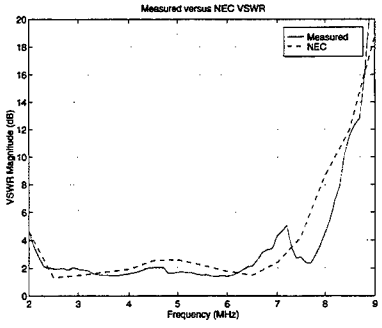
Measurements were accomplished at the scaled equivalent of fourteen frequencies (2.0, 2.5, 3.0, 3.75, 4.5, 5.5, 7.0, 8.5, 10.5, 13.0, 16.0, 19.5, 24.0, 30.0 MHz), five elevations (0°, 10°, 20°, 30°, 40°) and two polarizations (vertical and horizontal). It should also be noted that the measured antenna patterns represent relative measurements only.

5. Results and Comparisons

As shown in the figures below the computed and measured impedance and VSWR plots match well over the fan operating frequency range of 2 – 9 MHz. It should be noted that the NEC modeling was completed before the scale-model measurements and the NEC model was not "tweaked" to match the measured data. The small peak in the measured VSWR around 7 MHz is suspected to be noise or interfering signals in the measurements.

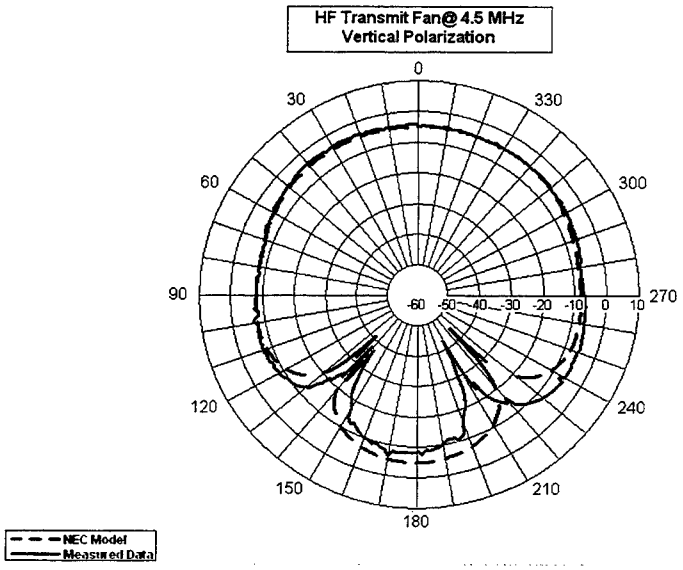


Impedance Results

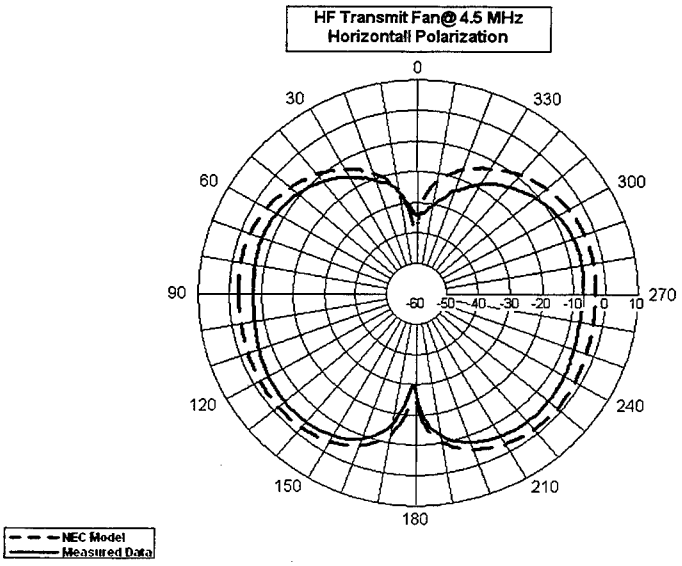


VSWR Results

The following figures show the NEC calculated radiation patterns at 4.5 MHz compared to the measured data for both vertical and horizontal polarization at 10 degrees elevation. The NEC data is plotted in dBi whereas the range data is relative to a reference antenna. The impedance, VSWR and pattern plots were made using the NEC-Win Pro software package. A FORTRAN routine was written to format the range data into a standard NEC2 output format so that it could be plotted with the NEC-Win Pro polar plotting routine.



Comparison of computed and measured data



Comparison of calculated and measured data

6. Conclusions

NEC2 did an excellent job of predicting the HF transmit fan antenna performance, as indicated by both the impedance plots and radiation patterns. These results instilled confidence that subsequent NEC models were valid and range measurements were consistent and repeatable. NEC-Win Pro served as a valuable tool for evaluating candidate antenna locations onboard naval platforms and for computing representative antenna parameters.

7. Acknowledgments

This work was sponsored by the Canadian Department of National Defence. Special thanks go to Mr. D. Robinson of DND and also Lt (N) D. Winters of the Royal Military College of Canada for his support in the 3-D visualization of the NEC models.

8. References

[1] K. Roberson, K. Lysiak, R. Landreth, "Trump HF System Upgrade Antenna Brass Modeling Final Report," SwRI Project 16-8484, Southwest Research Institute, 15 August 1997

[2] A.P.C. Fourie, D.C. Nitch, and O. Givati, "A Complex Body Structure Interpolation and Gridding Program (SIG) for NEC" (manual), EM Simulations Ltd., South Africa, 11 February 1994

[3] G.J. Burke and A.J. Poggio, "Numerical Electromagnetics Code (NEC) - Method of Moments," Technical Document 116, Volume 2, Lawrence Livermore Laboratory, Livermore, CA, January 1981

NEC Model Results for Shipboard Shielded Crossed Loop Antennas with Scale-Model Range Data

Keith Lysiak
Signal Exploitation and Geolocation Division
Southwest Research Institute
San Antonio, TX 78238

Abstract

The Signal Exploitation and Geolocation Division of the Southwest Research Institute conducted an internal research project to determine the feasibility of numerical modeling for shipboard High Frequency Direction Finding (HFDF) array design. The Numerical Electromagnetic Code (NEC4) calculated antenna responses that were compared to results measured at the SwRI scale-model rotary test facility. Both amplitude and phase results are compared. Although the shielded loops were modeled as simple unshielded loops, the results are good. These results indicate that numerical modeling for shipboard array design is feasible.

1. Introduction

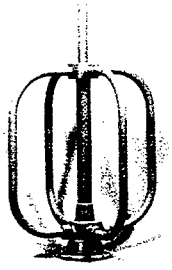
An HFDF system requires an array of distributed sensors. These sensors can be a mix of electric and magnetic elements. The response from these sensors is processed with a DF algorithm to provide angle of arrival and possibly elevation for a target emitter. The number of sensors and their placement greatly affects the performance of the DF system. Their placement is critical not only from a geometrical standpoint but also in terms of the presence of near-field scattering objects. Therefore the design of an HFDF array for an electromagnetically cluttered environment such as a Navy ship is a formidable task.

A database of antenna responses has been compiled for an HFDF array installed on the U.S. Navy CG-47 Ticonderoga class of ships. The sources of this database include two Navy ships, a $1/48^{th}$ brass scale-model, and a NEC model. The NEC model was developed to determine the feasibility of using numerical modeling to design shipboard HFDF arrays prior to installation on the scale-model. A detailed

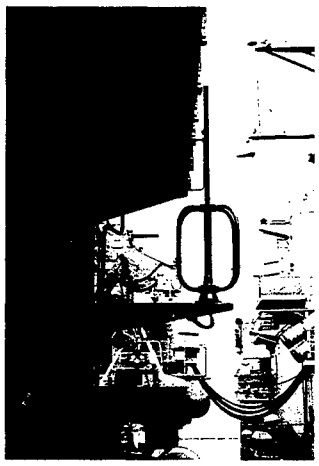
examination of the individual antenna responses indicates that numerical modeling is indeed a feasible design tool and that the CG-47 wire grid model performed well.

2. Shipboard HFDF Crossed-Loop Antenna

For the application in this paper, a unique antenna was used consisting of two orthogonal electrostatically shielded loops and a monopole. The loops are intended to sense two components of the magnetic field while the monopole senses the electric field. The loop elements are connected to matching networks installed in the base of the antenna. They are fed at the bottom and the electrostatic shield gap is at the top. The loops are approximately square and 0.6 meters on each side. The monopole is a simple wire element fed at the bottom of the loops without a matching network. A fiberglass tube running through the center of the loops supports the monopole. The monopole is approximately 1.5 meters long. The antenna has three outputs, one for each element. There are typically six antennas installed on each ship and 18 sensor inputs for the DF system. The operating frequency of the antenna is 0.5 - 30 MHz.



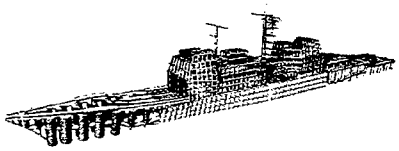
HF/DF Antenna



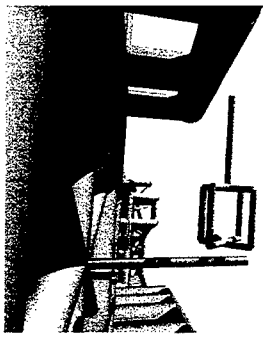
Typical shipboard installation

3. NEC Model

A wire grid model was developed of the full Ticonderoga class ship. The model was developed to run on a Pentium Pro 200 with 256 Mbytes of memory and was gridded for 15 MHz. This meant the segment spacing was typically 2 meters and the segment radius was approximately 0.32 meters. These parameters meet NEC's general guidelines for equal area. The bow and stern are built with a coarser gridding to reduce the total number of segments. The full model contains approximately 3400 segments and requires 190 Mbytes of memory. Each frequency run required 2.5 hours. A frequency run provides responses for a full 360-degree azimuth sweep at one frequency. GNEC software and NEC version 4 were used.



Full NEC wire grid model of CG-47



NEC model of HFDF antenna

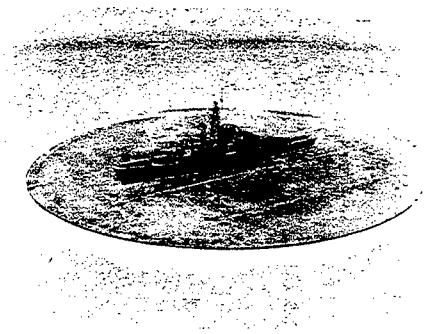
The main hull section was gridded using the Structure Interpolation and Gridding (SIG) program [ref. 1]. This program automatically grid surfaces based on a set of cross section contours. About 50 percent of the model are symmetrical and therefore produced with the GX (reflection in coordinate planes) card in NEC [ref. 2].

Three-dimensional visualization of the NEC model was accomplished with POV-Ray for Windows. The ability to visualize the wire segments in three dimensions greatly enhances the ability to refine the wire grid model. It provided the opportunity to visually check wire radii to make sure that they met NEC modeling guidelines.

The loop antennas are modeled as simple loops. They are square with five segments on each side. The center top segment is used as a current sensor. The loops are slightly vertically offset so that their segments do not touch. The monopole is modeled as a dipole above the loops. It contains nine segments and the center segment is used as a current sensor. A plane wave excitation is used with 360 azimuths at 1-degree increments. The NEC model is oriented similar to the scale-model with the main mast located at the center of rotation. In this manner, the NEC calculation is performed in very much the same way as the range measurement.

4. Scale-Model Measurements

A 1/48th scale brass model of the CG-47 Ticonderoga class ship was used to make antenna response measurements. The measurements were conducted at the SwRI scale-model rotary test facility. The test facility consists of a rotating copper platform surrounded by a radial ground screen with a diameter of 400 feet. RF signals can be generated at elevation angles of 0 to 60 degrees while the rotating platform provides 360 degrees of azimuthal positioning. All RF and control cables run underground to an equipment building approximately 200 feet from the rotator's center. While measuring the outputs of scale-model antennas as the ship is rotated, amplitude and phase data is recorded relative to a fixed reference antenna. Eight RF channels are available for simultaneous data collection of up to eight antenna outputs at a time.



SwRI scale-model rotary test facility

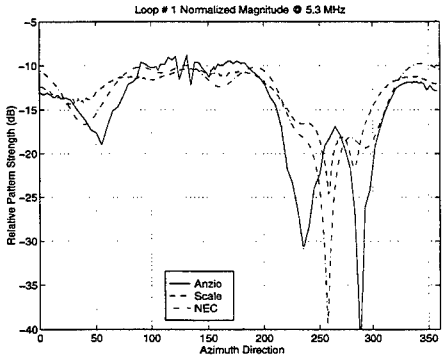
5. Ship Calibration Measurements

The shipboard HFDF system requires an antenna response calibration table to perform accurately. This table is simply the measured response for the antennas as a function of azimuth and frequency. These measurements are made using a land-based transmitter. The ship is sailed around a buoy at sea to acquire azimuth-dependent antenna response measurements. This method of performing measurements is similar to the range measurements but there is an additional phase term introduced by the variation in the distance between the ship and transmitter.

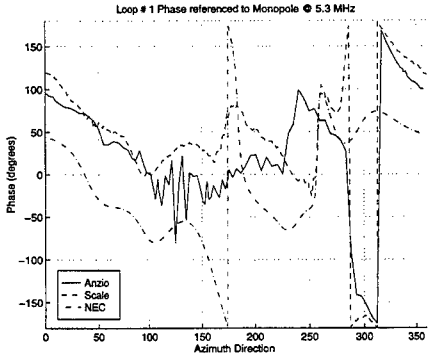
6. Results and Comparisons

A comparison for one of the six antennas will be discussed in this paper. The other antennas provided similar results. The element response magnitudes are normalized for simplicity. This allows the ship, range, and NEC data to be plotted on the same scale. For DF applications this is adequate because only the relative magnitude and phases are of interest. Both the range and NEC phase data contain a predictable oscillation due to the rotation of the antennas about the origin. The ship data has an additional phase oscillation due to the change in the ship location. Therefore, the loop phase response is referenced to the monopole phase response for these plots. Essentially, phase shifts due to the antenna movement relative to the origin have been removed. This does introduce a problem if the monopole responses do not agree. The ship data is shown as solid line, the range data is shown as dashed lines and the NEC calculations are shown as dot-dashed lines.

The Loop # 1 magnitude responses for all three data sets match well except near the null area around 250 degrees of azimuth. The NEC result appears to slightly favor the range magnitude rather than the ship magnitude. Notice that the loop no longer has a typical sine response but rather the pattern is dominated by its placement on the ship structure. The phase responses also track well. The NEC has an apparent offset of 50-100 degrees. Keep in mind that the ship and scale-model antennas have matching networks on the loop antennas that result in a phase offset between the monopole and the loops other than the theoretical 90 degrees. Given this fact, the NEC phase tracks well with the measured data. In general the two orthogonal loops will have similar amplitudes and phases once they are placed on the ship. This indicates that the antenna is sensing primarily the currents on the hull rather than from the incident field.

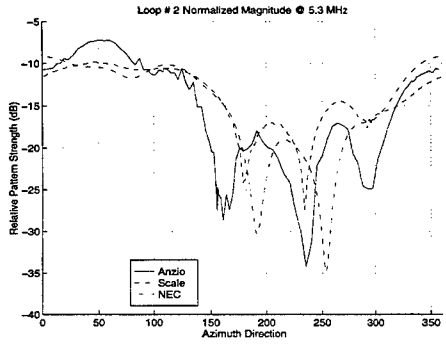


Normalized Magnitude Response

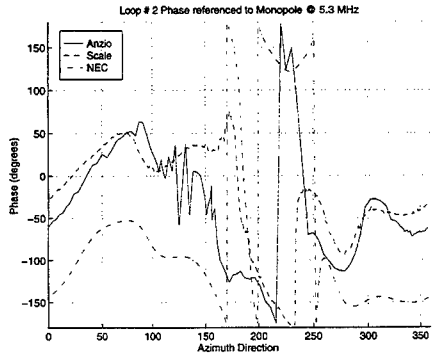


Normalized Phase Response

The magnitude responses for Loop # 2 also match well. Note that using the monopole as the phase reference raises the question as to which antenna is causing the difference in the phase results. Once again, the NEC magnitude response appears to favor the scale-model magnitude response.

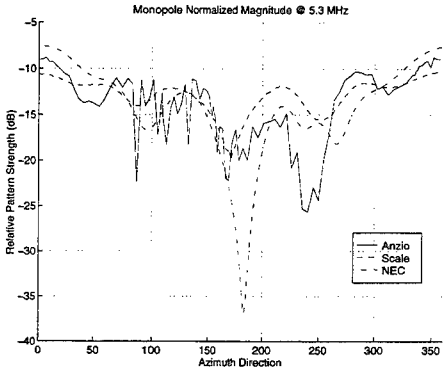


Normalized Magnitude Response



Normalized Phase Response

The magnitude response of the NEC modeled dipole also compared well to the measured monopole responses. Since the monopole is used as the phase reference, the phase plots result in a constant phase of zero.



Normalized Magnitude Response

7. Conclusions

The NEC magnitude and phase responses for the 18 elements of the HFDF array compared favorably to the ship and scale-model measured data. For DF applications, only the relative magnitudes and phases are required. Therefore, it is not necessary to accurately model the impedance or absolute gain. Related research has shown that the results presented here provide very good DF performance estimations. Numerical modeling has proven to be a useful tool for use in predicting DF performance for various HFDF array installations.

8. References

[1] A.P.C. Fourie, D.C. Nitch, and O. Givati, "A Complex Body Structure Interpolation and Gridding Program (SIG) for NEC" (manual), EM Simulations Ltd., South Africa, 11 February 1994

[2] Gerald J. Burke, "Numerical Electromagnetics Code - NEC-4, Method of Moments, Part I: User's manual" UCRL-MA-109338, Lawrence Livermore National Laboratory, Livermore, CA, January 1992

A Near-Earth and Buried HF Antenna Computer Modeling Program

W. Perry Wheless, Jr. and Larry T. Wurtz
Department of Electrical Engineering
University of Alabama
Tuscaloosa, AL 35487
e-mail wwheless@ua1vm.ua.edu

Abstract

Near-earth and buried antennas for HF (1.8 - 30 MHz) communication applications may be very accurately analyzed by computer implementation of an analytic model, independent of sophisticated NEC-3 or NEC-4 kernels. This paper describes a specialized MATLAB computer program, named SNAKE1, which models the feedpoint impedance, VSWR, pattern shape and power gain for single-element near-earth and buried wire antennas. The program requires the user to provide the characterization of the soil conductivity σ and dielectric constant ϵ_r where the antenna is deployed. The governing analytical equations for the model are believed to be useful between approximately 1 kHz to 100 MHz, so other applications in addition to HF antennas are possible. Because it has comparable accuracy, executes quickly, and may be distributed freely without restriction, SNAKE1 is an attractive alternative to NEC for this particular class of antennas.

1 Introduction

The amateur radio service and other practical radio communicators using the HF spectrum (nominally extended to mean 1.8 - 30 MHz here) have multiple interests in near-earth and buried single-element wire antennas. Near-earth dipole and traveling-wave antennas are very portable and quickly deployed. They can be particularly effective for NVIS (near vertical incidence skywave) communications over short links. It also turns out that they are effective radiators of end-fire vertically polarized fields at low takeoff angles.

Further, the potential utility of these so-called *snake* antennas, a term which encompasses both near-earth and buried deployments, for the selective reduction of objectionable static interference at frequencies up through the 40-meter band (7 MHz) was discussed in [1], where most of the basic components for an analytical model are archived for convenient reference. The remainder of the analytical model was documented in [2]. The governing equations in [1] and [2] are based on the excellent work results reported in references [3] and [4].

It is assumed that the near-earth or buried wire is straight, and aligned along the x-axis as illustrated in Figure 1. ϕ denotes azimuth angle, measured CCW from the +x axis. θ denotes elevation angle, with $\theta = 0$ representing the air-earth interface, referred to as the horizontal plane.

Because MATLAB [6] has become the premier software package for interactive numeric computation, data analysis, and graphics at numerous academic institutions, and is also gaining widespread acceptance in industry, a computer implementation for the *snake* antenna model was carried out in MATLAB.

The program SNAKE1 was developed to promote academic pursuits of computer-based modeling and experimentation by amateur radio enthusiasts and other practical radio communicators actively using the HF spectrum. Both end-fed and center-fed configurations, as shown in Figure 2, are handled by the program. SNAKE1 calculates feedpoint impedance and VSWR characteristics and plots antenna patterns; power gain in dBi, as well as pattern shape, are accurately described by the model over the entire HF band. The user may specify an elevation or azimuth plot, with a choice of either vertical ($E\theta$) or horizontal ($E\phi$) component, for each program execution.

Earth (real ground) permittivity is frequency dependent, but often it is the case that the dielectric constant and conductivity are known at just one frequency. Formulas which allow a reasonable approximation of the frequency dependence based on data at a single frequency are included in SNAKE1.

The program has been validated by comparisons to NEC-3 modeling results and to patterns contained in [5]. As noted in [1], the Eyring Communications Division was developing advanced buried antenna modeling and hardware products for several years before they ceased operation in the early 1990s. The SNAKE1 program was developed independently, from governing mathematics available in the scientific literature, and is intended for individual academic pursuits only. Comparisons of several SNAKE1 results to illustrative plots published in [5] show acceptable and consistently close, but not exact, agreement. The theoretical equations may be manipulated into different forms for computer implementation, and it is believed that the observed small discrepancies follow from the implementation of slightly different equations.

2 Modeling Geometry

There are four basic antenna element configurations, as shown in Figure 2. In all four cases, it is assumed that the wire axis is aligned with the x-axis. Two variations of the basic antenna element are end-fed, one of which is open-terminated and the other ideally match-terminated. The other two variations are center-fed, similarly with one version open-terminated and the other match-terminated.

For the power gain calculations to be realistic, feedpoint mismatch losses are taken into account. It is assumed that a balanced transmission line is attached to the two feedpoint terminals, for all the center- and end-fed variations. The source impedance specified to the program should be that of the transmission line. If a source impedance of 450 Ω is specified to the program while the actual transmitter output impedance is 50 Ω, for example, it is tacitly assumed by the program that the user will employ a 9:1 balun on the transmitter output to achieve the conversion from a 50 Ω unbalanced to a 450 Ω balanced feed.

Figure 1 identifies some parameters relevant to computation and plotting of radiation patterns. Note that the elevation angle θ is measured relative to the horizontal plane, that the $E\theta$ field component is taken by definition to be the vertical polarization component, and that the $E\phi$ field component is similarly taken to be the horizontal polarization component. The program, in its present form, can plot both azimuth and elevation radiation patterns for either vertical or horizontal polarization.

3 Program Operation

The SNAKE1 program interactively prompts the user to input the necessary parameters for program execution. Brief remarks on each of the major program setup steps follow:

1. FEED TYPE: center or end feed.
2. TERMINATION: matched or open termination of the antenna element.
3. FIELD OF INTEREST: skywave or groundwave.
4. COMPONENT OF INTEREST: vertical or horizontal.
5. FREQUENCY RANGE OF INTEREST (MHZ): either a single frequency (scalar entry) or a vector spanning the range of frequencies of interest.
6. GROUND PARAMETERS: Default values of $\varepsilon_r = 10$ and $\sigma = 5$ mS/m may be accepted or changed interactively in response to user prompts.
7. VARY GROUND PARAMETERS WITH FREQUENCY? Allows the user to specify *yes* or *no*. If *yes* is specified, the user is prompted to input the single REFERENCE FREQUENCY in MHz.
8. WIRE RADIUS AND INSULATION RADIUS: Default values of 0.001 and 0.005 m, respectively may be accepted or changed interactively in response to user prompts.
9. ANTENNA LENGTH L IN FT OR M? Allows the selection of meters or feet for input of antenna element length.
10. INPUT LENGTH OF ANTENNA ELEMENT: Note from Figure 2 that input length L is the full length of the end-fed elements, but is half the overall length of the center-fed variations. That is, the overall length of the center-fed variations is 2L.
11. ANTENNA HEIGHT IN M: Heights above ground for near-earth elements are positive; for buried antennas, this entry is a negative number.
12. FEED LINE IMPEDANCE IN OHMS: A balanced transmission line feed is assumed, as discussed earlier. Futher, if this value is specified as 0, the program does not compute and take into account mismatch loss.

After these entries, the program will compute and display the following summary information for each frequency of interest: frequency, ground dielectric constant, gamma ($\beta - j\alpha$) per [1], antenna feedpoint impedance, line characteristic impedance, reflection coefficient magnitude, VSWR, and mismatch loss (where selected). At the conclusion of this tabulation, the user is offered an opportunity to (13) DISPLAY A FEED-POINT SUMMARY TABLE and, after that, the opportunity to (14) PLOT VSWR VERSUS FREQUENCY.

Finally, the pattern plotting options are selected:

15. 0 = AZIMUTH PLOT, 1 = ELEVATION PLOT, 2 = NO PLOT.
16. If 'azimuth plot' is selected, the user is prompted to input the desired elevation angle (in degrees). If 'elevation plot' is selected, the user is prompted to input the desired azimuth angle.

4 Illustrative Results

The impedance and mismatch results of two test cases are summarized in Table I below. These typical values both illustrate the impedance levels to be expected in applications, and provide numerical values that independent programmers can use for comparison. The *snake* of Example 1 is deployed above, but near ground, at a height of 1.0 m. For Example 2, the antenna element is buried at a depth of 0.5 m.

Table I. Summary of impedance and mismatch examples.

Description	Example 1	Example 2
Input Parameters:		
Frequency (MHz)	10.0	10.0
Soil dielectric constant	10	4
Soil σ in mS/m	5	5
Wire insulation dielectric constant	5	2.25
Wire insulation σ in mS/m	0	0
Length of antenna element in m	20	13.4
Wire radius a in m	0.001	0.003175
Insulation radius b in m	0.005	0.0127
Transmission line impedance Ω	600	300
Feed (c=center, e=end)	c	c
Termination (o=open, m=matched)	o	o
Field (s=sky, g=ground)	s	s
Component (v=E_Θ, h=E_Φ)	v	v
Vary ground parameters/ref. MHz?	no	no
Wire height z (m); -z $\rightarrow$ buried	1.0	-0.5
Outputs:		
$\gamma = \beta - j\alpha$	0.2370-j0.0180	0.5252-j0.1867
Line char. impedance Z_{AC}	515.79-j39.27	159.11+j13.58
Ant. feedpoint impedance Z_{in}	359.03-j2.496	318.82+j22.90
Reflection coefficient magnitude $\|\Gamma\|$	0.2513	0.0479
VSWR	1.6712	1.1005
Return loss (dB)	-11.917	-26.400
Mismatch loss (dB)	-0.2832	-0.0100

Example 1 was subsequently re-run with the frequency entered as a vector running from 2 to 32 MHz in steps of 0.66 MHz. The program was instructed to vary the ground parameters with frequency, taking the values above as true at reference frequency 10.0 MHz. Figure 3 shows the graphical result, and clearly indicates the potential for broadband operation. The small glitch in the VSWR curve near 14 MHz is because $|k_2 z|$ becomes equal to one in that vicinity, and the two above-ground element propagation constant approximations do not meet seamlessly at the transition point.

In addition, radiation pattern plots for another four test cases are reported here. Table II summarizes the important details of these illustrative cases. Two computer runs were made

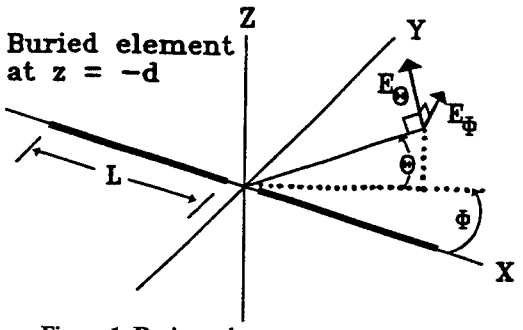


Figure 1. Basic *snake* antenna geometry.

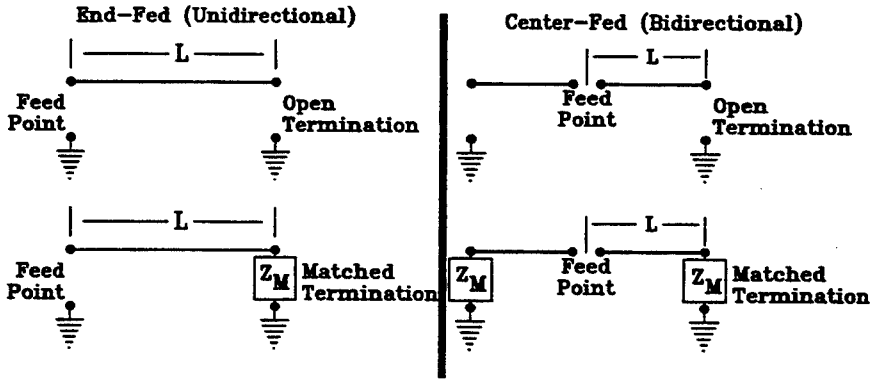


Figure 2. Unidirectional and bidirectional variations.

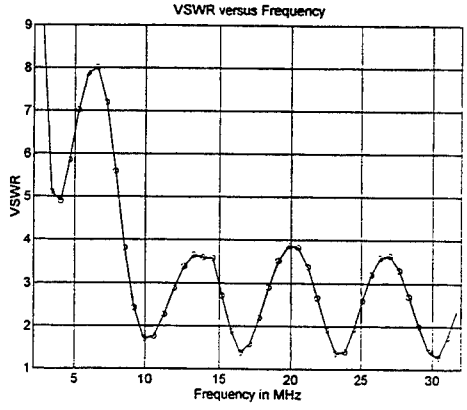


Figure 3. VSWR 3-32 MHz for the antenna of Example 1.

for each case - first using the positive z (wire height) value indicated in Table II, and then with z = -0.5 m, also indicated on the same line of the table.

Table II. Summary of four radiation pattern examples.

Description	Example 3	Example 4	Example 5	Example 6
Frequency (MHz)	10	8.015	8.015	8.015
Soil dielectric constant	10	9.5	9.5	9.5
Soil σ in mS/m	5	6.7	6.7	6.7
Wire insulation dielectric constant	5	2.7	2.7	2.7
Wire insulation σ in mS/m	0	0	0	0
Length of antenna element in m	20	22.86	22.86	22.86
Wire radius a in m	0.001	0.0008	0.0008	0.0008
Insulation radius b in m	0.005	0.0012	0.0012	0.0012
Transmission line impedance Ω	450	450	450	450
Feed (c=center, e=end)	c	c	c	c
Termination (o=open, m=matched)	o	o	o	o
Field (s=sky, g=ground)	s	s	s	s
Component (v=E_Θ, h=E_Φ)	v	v	h	v
Vary ground parameters/ref. MHz?	no	no	no	no
Wire height z (m); -z $\rightarrow$ buried	1.0/-0.5	0.66/-0.5	0.66/-0.5	0.66/-0.5
Plot	v elev.	v elev.	h elev.	v az.
Θ/Φ conditions on plot	$\Phi = 0°$	$\Phi = 0°$	$\Phi = 90°$	$\Theta = 30°$

The resultant radiation patterns are in Figures 4 through 7. The reader is reminded that the computer program includes mismatch loss into the computed gain patterns, unless a transmission line impedance of zero is specified, in which case mismatch loss is ignored. For brevity, a detailed commentary on the figures is omitted; clearly, however, single-element near-earth and buried antennas are often in the operational regime -15 dBi to -25 dBi for practical HF deployments.

5 Conclusions and Future Research

A computer-based capability, using MATLAB, for the sinusoidal steady-state analysis of near earth and buried single (insulated) wire elements has been developed. It is intended that the code developed for this application will be freely distributed to support academic research into this interesting class of low-frequency antennas.

The ability to predict impedance conditions and radiation patterns for *snake* antennas is of considerable interest to practical radio communicators. The relatively low power gains, compared to isotropic, are not necessarily objectionable for 'receive only' and certain tranmit applications. Incorporation of signal angle-of-arrival considerations is planned for a future program upgrade.

Future work will also incorporate provision for arrays of near-earth and buried elements. The additional gain provided by arraying these basic antenna elements may broaden the scope of practical transmit applications. Also, an experimental measurement program to quantify the selective rejection effectiveness of these antennas for both local man-made noise and distant, naturally occurring static sources is planned for the future.

Solid line Z = 1.0 m **Vertical Polarization** **Maximum = 0 dBi**
Dashed line Z = -0.5 m **10 dB/division**

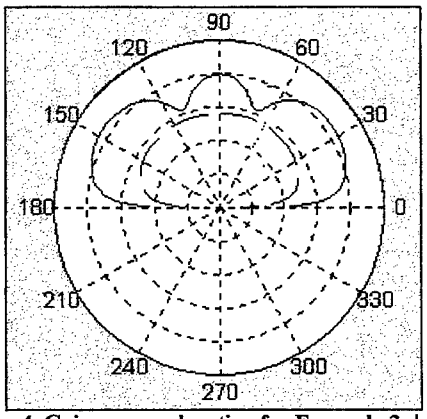


Figure 4. Gain versus elevation for Example 3, $\phi = 0°$

Solid line Z= 0.66 m **Vertical Polarization** **Maximum = 0 dBi**
Dashed line Z = -0.5 m **10 dB/division**

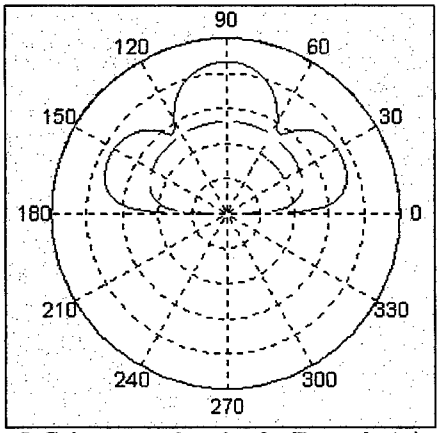


Figure 5. Gain versus elevation for Example 4, $\phi = 0°$

Solid line Z = 0.66 m Horizontal Polarization Maximum = 0 dBi
Dashed line Z = -0.5 m 10 dB/division

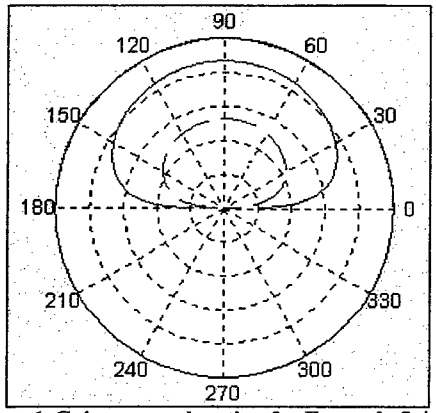


Figure 6. Gain versus elevation for Example 5, $\phi = 90°$

Solid line Z = 0.66 m Vertical Polarization Maximum = 0 dBi
Dashed line Z = -0.5 m 10 dB/division

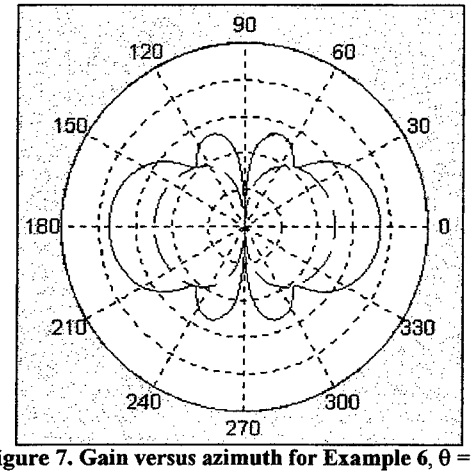


Figure 7. Gain versus azimuth for Example 6, $\theta = 30°$

6 Acknowledgment

The authors thank David Faust and his former colleagues at the now-defunct Eyring Communications Division, Tom McLees, and other experimenters for their contributions of both motivation for this study and factual information reflected in this paper.

References

[1] W.P. Wheless, Jr. and L.T. Wurtz, "Components of an analytical model for *snake* antennas," *ACES Newsletter*, vol. 12, no. 2, p. 35-44.

[2] W.P. Wheless, Jr. and L.T. Wurtz, "Computer implementation of a *snake* antenna analytical model," *ACES Newsletter*, vol. 12, no. 3, pp. 31-38.

[3] G.A. Lavrov and A.S. Knyazev, *Near Earth and Buried Antennas*, Sovetskoye Radio, Moscow, 1965 (translated by the Joint Publications Research Service, U.S. Department of Commerce, Washington, D.C.).

[4] R.W.P. King and G.S. Smith, *Antennas in Matter*, MIT Press, Cambridge, MA, 1982.

[5] M.B. King, R.B. Gilchrist, and D.L. Faust, *Eyring Low Profile and Buried Antenna Modeling Program*, Eyring (preliminary) Document 300-0085A, June 1991.

[6] MATLAB is a registered trademark of The MathWorks, Inc., Natick, Massachusetts.

Empirical and Numerical Treatment of Electromagnetic Pulse Induced Currents

Malcolm J. Packer
Harris Corporation, RF Communications Division
1680 University Avenue, Rochester, NY 14610

ABSTRACT

Short duration electromagnetic pulses created by a high-altitude nuclear event can cover an immense geographical area with peak field strengths in excess of 50 kV/m. The induced voltages and currents from this short duration pulse are capable of disabling communications by destroying susceptible semiconductors within radio equipment. It is important to be able to predict the induced antenna currents to determine the potential threat and protection mechanism for communication systems.

A crossed dipole antenna was subjected to non-ionizing electromagnetic transient energy that simulated a high-altitude nuclear event. The crossed dipole antenna was connected to a microprocessor controlled antenna coupler and radio. This paper compares these empirical results to numerically predicted data.

Predicted data based on a previously published numerical technique was first considered. This technique was enhanced by adding a correction factor to retain the initial conditions at time equals zero. A comparison of measured and predicted input impedance is presented. The empirical technique follows, starting with the test set-up and pulser antenna details. The electric and magnetic field data collected, both spectral and temporal, with and without the test antenna is presented and discussed. Next the currents induced onto the crossed dipole antenna and delivered to the antenna coupler are discussed. Finally, the currents that passed through the antenna coupler are discussed. Direct comparisons of the predicted data to measured results are complicated, due to the difficulty providing a short-duration pulse waveform with correct characteristics. However, the empirical created waveform does provide useful insight into induced currents and potential threat to a communication system.

NUMERICAL TREATMENT

The numerical approach[1] involves calculating an antenna's plane wave response using the Numerical Electromagnetic Code[2] (NEC). A one Volt/meter plane wave excites the wire-grid antenna model every 250 kHz (corresponds to the desired sampling rate in the time domain) across the frequency band from 250 kHz to 100 MHz. The calculated plane wave response is multiplied with the Discrete Fourier Transform (DFT) of the generalized high-altitude EMP double exponential transient waveform. The result of this spectral multiplication is the predicted short-circuit current response of the antenna when it is subjected to an incident EMP transient. The short-circuit transient response current is obtained by complex Inverse DFT (IDFT).

The short-circuit current when combined with the complex antenna input impedance in the frequency domain can be converted to an open-circuit Thevenin voltage. With the antenna represented as a Thevenin voltage source and source impedance, the voltage transfer to any arbitrary load can be calculated. Applying a complex IDFT provides the transient response across the load.

The numerical technique[1] presented last year was applied to five antennas: a simple dipole, a fan dipole, a crossed dipole, a whip with ground radials, and a sloping VEE. Of the five antennas, not all transient responses started at time equals zero, but tens of nanoseconds earlier. This error is introduced

EXPERIMENTAL TREATMENT

A radio system that includes a crossed dipole antenna, antenna coupler, and transceiver was exposed to electromagnetic pulse simulation. The antenna height (8.7 meters) required a simulator with the minimum pulser height of 27 meters. This factor of three is typical to limit coupling between pulser antenna and the test antenna. The Electromagnetic Transients Branch, NAWCAD, Patuxent River, Maryland Electromagnetic Pulse Test Facility had a suitable site[4].

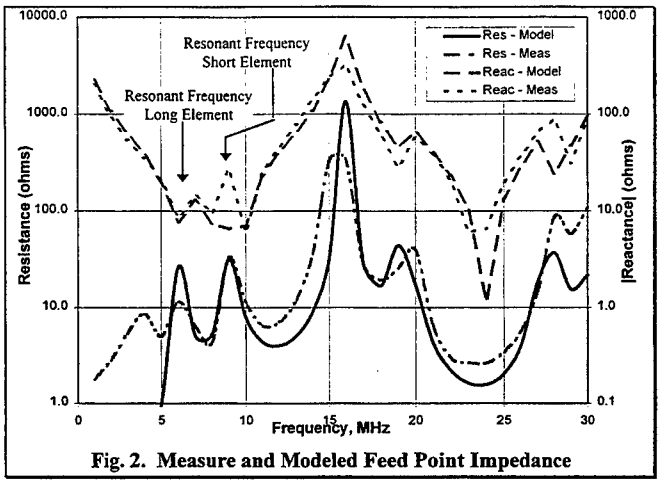


Fig. 2. Measure and Modeled Feed Point Impedance

This facility has a hybrid simulator[5, 6, 7] which combines various features of radiation simulators and static simulators. This hybrid simulator provides both the early-time (high frequency) portion of the waveform radiated from the source region and the late-time (low-frequency) portion of the waveform radiated over the entire pulser antenna. The hybrid simulator at NAWCAD included a ground plane as part of the pulser antenna. This includes the ground reflected pulse, an important component of the simulation in respect to testing antennas.

The test antenna was deployed directly below the pulser antenna as shown in Figure 3. Field strength measurements were gathered at 5.5 meters above the ground plane with and without the antenna. The x-directed magnetic field, H_x, and the z-directed electric field, E_z, were measured. The z-directed electric field is not the primary field. The primary field is E_y aligned with the pulser antenna. Data for E_y was taken, but a fiber optic transmitter failure prevented reliable data. With the crossed dipole antenna under the pulser antenna, H_x and E_x were measured. Current probes were used to measure the current induced on the crossed dipole antenna and delivered to the antenna coupler and the current delivered to the radio.

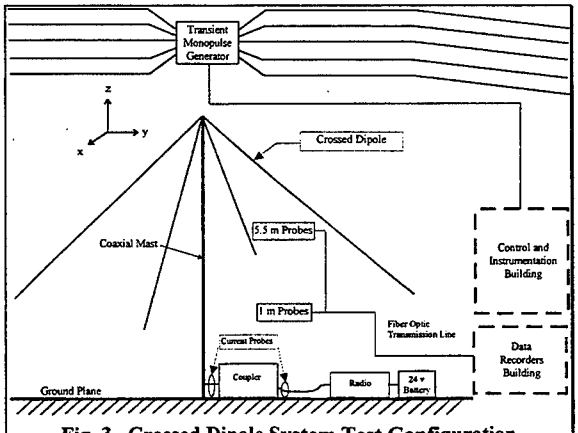


Fig. 3. Crossed Dipole System Test Configuration

in the frequency domain when the Numerical Electromagnetic Codes (NEC) computes the complex short circuit response to a unit plane wave. If the phase center of the antenna was aligned with the Cartesian coordinate center no error was introduced. As the phase center of the antenna moved away from the coordinate center, a phase error was introduced. This translates to a time error equal to the time for energy to travel the distance between the coordinate center and the phase center of the antenna. This results in the transient response appearing to start in negative time. This problem of phase reference to the coordinate center is inherent to the NEC algorithm. A correction factor versus frequency was added in MATHCAD[3] to subtract the phase difference between coordinate center and phase center of the antenna. With the phase corrected plane wave response, successive DFT and IDFT still have transient starting at time equals zero.

The antenna that underwent extensive measurements was the extended crossed dipole. This antenna is an extended version of the AS-2259/U that includes an 8.7 meter coaxial mast that supports two unequal sets of dipole elements mounted at 90°.

The predicted current transient response for the crossed dipole antenna is a damped exponential sinusoidal as displayed in Figure 1. One curve is when the antenna is grounded and the other is when the antenna is connected to an antenna coupler tuned to 8.4 MHz. The waveform reaches 900 Amperes peak current when connected to the tuned antenna coupler and 1350 Amps when connected to a short circuit. Sinusoidal oscillations occur every 170 nanosecond (ns.) for a resonant frequency of 5.9 MHz. The EMP 10 ns. rise-time has significantly decreased to 60 ns., due to the bandpass nature of the antenna, which attenuates high frequency components. Both waveforms have oscillatory behavior at about 6 MHz, but decay at different rates. The short circuit termination has the longer decay time.

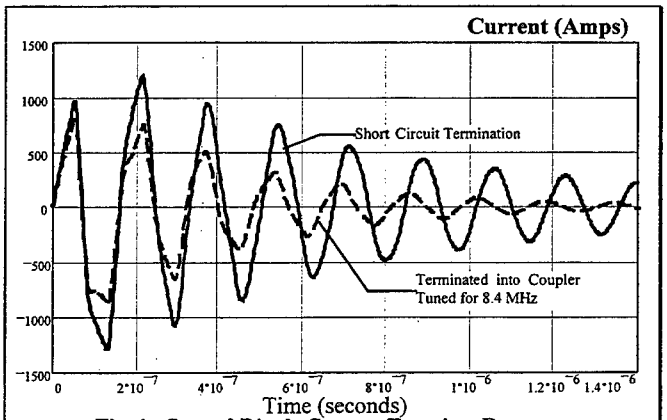


Fig. 1. Crossed Dipole Current Transient Response

Figure 2 displays the measured and modeled crossed dipole antenna's feed point impedance. The two lower curves represent the resistance and the two upper curves are the absolute value of the reactance. This figure exemplifies the ability to accurately model antenna impedances with NEC. The reactance zero crossing indicates the long elements are resonant at 6 and 18 MHz, while the short elements are resonant at 9 and 27 MHz. The transient response frequency agrees with the longest element resonant frequency. The analysis and testing were both performed with the longest element aligned with the polarization of the incident EMP wavefront.

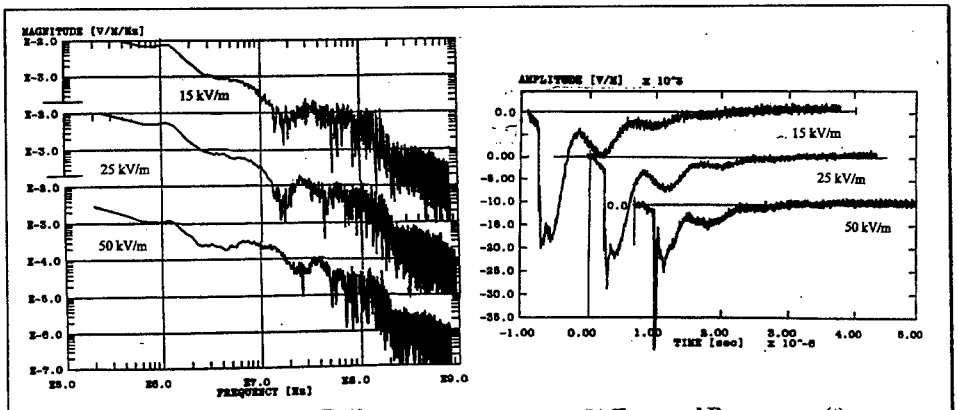
The pulser must always be triggered at 50 kV/m peak due to difficult timing issues creating the fast rise-times required. To ramp up to the 50 kV/m threat level, the crossed dipole antenna was placed at predetermined distances from the pulser's center. Resulting in exposure levels of 15 kV/m, 25 kV/m, and 50 kV/m.

ELECTRIC FIELD

At these three locations, the electric field $E_z(f)$ at 5.5 meters above the ground plane without the crossed dipole antenna is shown in Figure 4a. The logarithmic frequency scale covers four decades from 100 kHz to 1GHz. The amplitude is relatively flat below 1 MHz at -40 dBV/m/Hz (0.01) which agrees very well with the numerical results. The relative amplitude on a per Hertz basis decreases as the distance to the pulser's center decreases, but the spectrum shape is almost identical. All the spectral plots are increasingly noisier at higher frequencies. The starting frequencies of noise correlates with inverse distance. Directly below the pulser's center a noise free spectrum up to 20 MHz is obtained. At the furthest distance, 33m, noise starts at 13 MHz. The noise level is 40 dB below the signal peak. The noise is introduced from limited dynamic range of the measurement system in the time domain.

The time domain plots, $e_z(t)$, for the spectrums are displayed in Figure 4b. The curves are shifted in time to plot the them closer together. All three curves approach 30 kV/m with a quick rise-time after a slow 200 ns. ramp. A 80 ns. delay was expected, the additional 120 ns. is due to time delays associated with the measurement system. Each curve shows a second, third, and a trace of a forth peak, then a slow decay that approaches zero. The second peak appears merged with the first peak at the furthest distance. As the probe approaches the pulser's center the peaks separate and the second and third peaks decrease. These extra peaks are due to the reflection from the pulser antenna's ends. The time span between successive peaks is 800 ns. or 1.25 MHz. This frequency component can be seen as a slight peak on the spectrum plots.

At the same three measurement locations, the crossed dipole antenna was placed for free field and current measurements. The x-directed electric field, $E_x(f)$, at 5.5 meters above the ground plane is shown in Figure 5a.

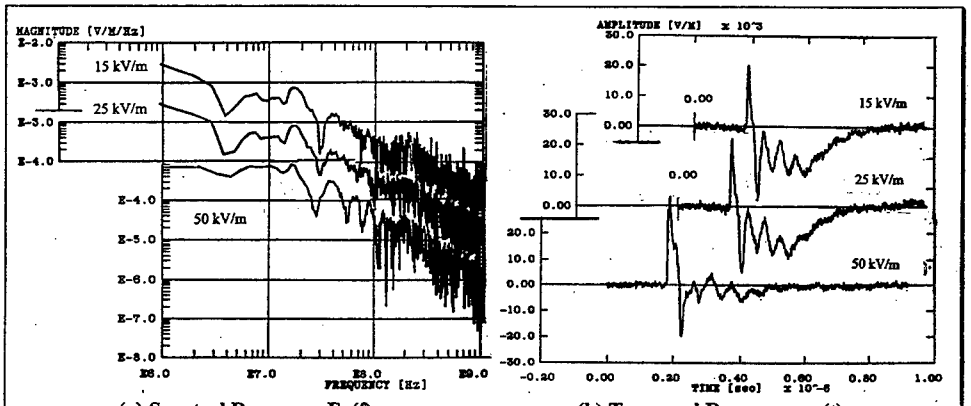


(a) Spectral Response, $E_Z(f)$, **(b) Temporal Response, $e_Z(t)$**

Fig. 4. Electric Field at Three Location under the Pulser's Antenna

The amplitude peak at 1 MHz does not achieve -40 dBV/m/Hz with this component of the E-field or the field is perturbed by the crossed dipole antenna. Clearly, the amplitude of each spectral component varies across the spectrum with low amplitudes at 4 and 30 MHz and high amplitudes at 1 and 17 MHz. The same features are seen at the location that delivered 25 kV/m. Directly under the pulser, many low and high amplitude spectra appear: lows at 4, 25, 51, and 79 MHz, highs at 10, 17, 36, 62, and 90 MHz. This indicates strong interaction or coupling between antennas. Noise is still present, the signal to noise level is only 30 dB but occurs above 50 MHz.

The time domain plots, $e_x(t)$, for the spectrums are displayed in Figure 5b. The time scale covers 1 μs. Each curve shows a 180 ns. delay, then a fast rise-time to a peak. Then a fast rising negative peak which is the ground reflected pulse. Three to four oscillations occur before the electric field approaches zero. The oscillation frequency is 16.7 MHz.



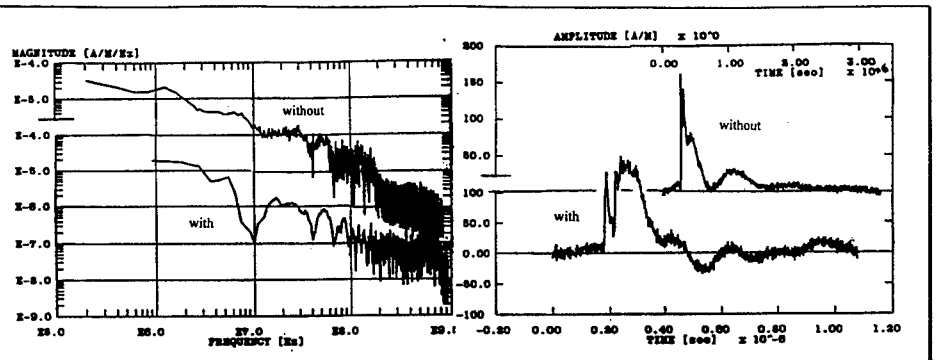
(a) Spectral Response, $E_X(f)$ **(b) Temporal Response, $e_X(t)$**

Fig. 5. Electric Field at Three Location with Crossed Dipole under the Pulser's Antenna

There is a significant difference in the electric field response when the crossed dipole antenna was included in the measurements. We have compared E_z to E_x, where the normal field, E_z, does not have a ground reflected pulse. However, the tangential field, E_x, provides a ground reflected pulse with 180° phase reversal. It is unfortunate that the same component was not measured for each case. The same component was measured for the magnetic field.

MAGNETIC FIELD

The tangential magnetic field, H_x, was examined for its response when the crossed dipole was introduced. Figure 6 displays H_x for the 50 kV/m level with and without the crossed dipole antenna. Both time response curves do not have a ground reflected pulse, as expected for a tangential magnetic field. The curves are different, particularly the amplitude and duration of the second peak. In the spectral domain, the curves display the same typical effects of frequency selectivity when the crossed dipole antenna was introduced. The exact locations of the low and high amplitudes have shifted; lows at 10, 40, and 65 MHz, and highs at 1, 15, 52, and 90 MHz.



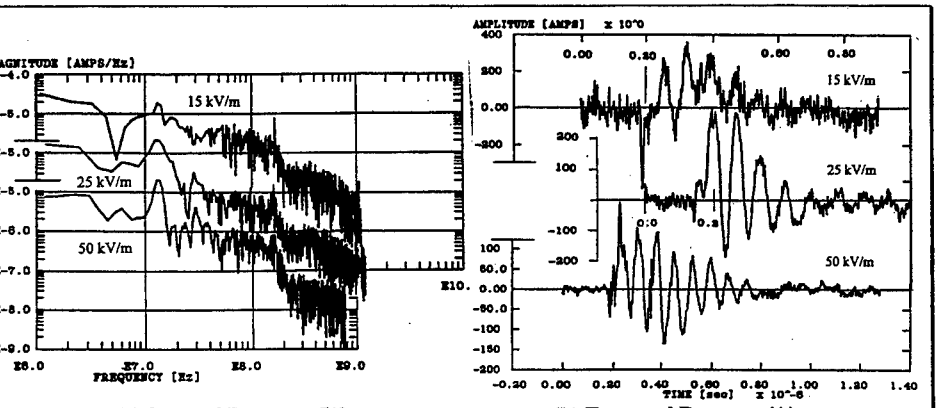
(a) Spectral Response, $H_X(f)$ (b) Temporal Response, $h_X(t)$

Fig. 6. Magnetic Field With and Without the Crossed Dipole Antenna

ANTENNA CURRENT

Current delivered to the antenna coupler by the crossed dipole antenna was measured with a current probe and the results are shown in Figure 7. At the furthest location, equivalent to a 15 kV/m EMP level, the peak current approached 400 amps. The temporal trace is rather noisy, believed to be caused by voltage arcing. The spectrum has a minimum at 5.5 MHz, but the maximum current was expected at the resonant frequency of 5.8 MHz. The maximum current was at 12 MHz, this frequency correlates to two way travel time the length of the long element.

When exposed to 25 kV/m EMP peak level, a peak of 275 amps was delivered to the coupler that decayed in four cycles. At 50 kV/m exposure, the coupler saw a 250 amp peak decaying over 8 cycles. The spectral curves reveal that as the peak exposure level increases, or as the crossed dipole approaches the pulser's center, all frequency components decrease except for 12 MHz.



(a) Spectral Response, I(f) (b) Temporal Response i(t)

Fig. 7. Current Induced on the Crossed Dipole and Delivered to the Antenna Coupler

The coupler has a gas-filled spark-gap for transient protection. The spark-gap shunts all received energy to ground if the voltage increases beyond a preset threshold, 7 kV. At 50 kV/m exposure level the current decay time increases. This increase in decay compares well to theory (Figure 1) when the antenna is shorted, indicating that the transient protection device triggered to protect the coupler's sensitive semiconductors.

Comparison of predicted results with empirical data shows that the predicted peaks are higher by a factor of two to three. The reason for this is the exciting waveshape is different, the theoretical pulse is a smooth double exponential with pure polarization aligned with the long dipole elements. Whereas, the empirical pulse creates an electric field distributed among all three electric field components. Therefore, the numerical model represents the worst case scenario. In the actual scenario, a high altitude electromagnetic pulse would approach a plane wave. Therefore, the numerical simulation is more representative, on the other hand the polarization tilt is unknown.

Comparing the primary measured frequency content was 12 MHz and 6 MHz predicted. Whereas the coupler's tuned frequency component at 8.4 MHz was not evident. A numerical model was used to examine the response of linear polarized waves aligned with the long and short elements independently. Figure 8 shows the relative current magnitude received on the antenna. Both polarizations received current peaks at 5, 8, 12, and 24 MHz; the long elements also received 20 MHz and rejected 30 MHz, while the short elements rejected 20 MHz and received 30 MHz. The greatest predicted peak is 5 MHz with the peaks at 8 and 12 MHz being 3 dB down.

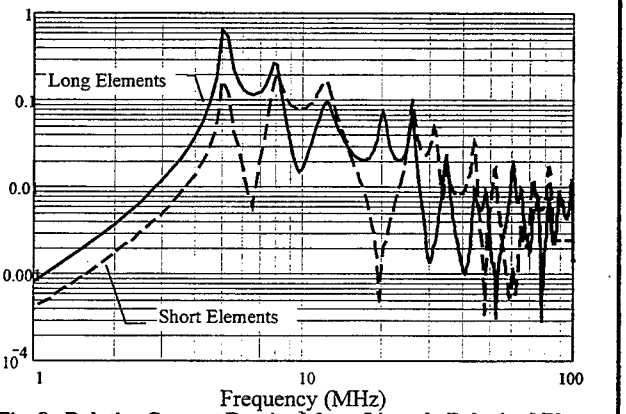
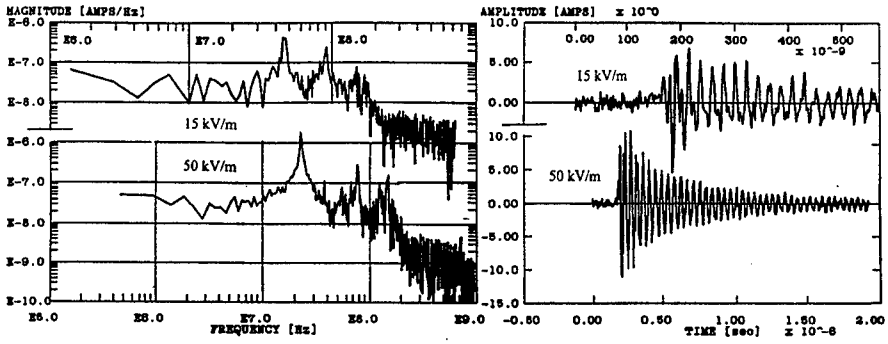


Fig. 8. Relative Current Received from Linearly Polarized Plane Waves Aligned with the Long or Short Elements

CURRENT OUT OF THE COUPLER

A current probe was placed on the center conductor of the coaxial cable which connected the antenna coupler to the transceiver. Shown in Figure 9, at 15 kV/m level, the current appears to be a summation of several frequency components with the primary one being 45 MHz. The spectrum plot reveals the main frequency content is centered on 45 MHz and 90 MHz. These components are harmonically related and pass through the coupler. The coupler must therefore have a bandpass response at 45 and 90 MHz when tuned for 8.4 MHz operation.

At the 50 kV/m level, directly under the pulser antenna, the current had a peak of 12 Amps with a fundamental frequency of 22 MHz, harmonically related to 45 and 90 MHz. This current plot appears to be created by a single excitation and rings while it decays. The single excitation is the energy that passed into the coupler before the transient spark-gap triggers.



(a) Spectral Response, I(f) **(b) Temporal Response i(t)**

Fig. 9. Current That Passed Through the Antenna Coupler

CONCLUSIONS

An enhancement was provided to the numerical approach that reestablished correct initial conditions. An input impedance comparison between numerical and measured data validated the wire-grid model. Clearly, the crossed dipole antenna introduction perturbed the electromagnetic fields created by the hybrid simulator. This caused difficulty in the comparison with the predicted data, the antenna source waveforms are different. The predicted induced current peak values are double the empirical values, which represents the worst case. The most interesting feature was the current's spectrum delivered to the antenna coupler. Along with the expected antenna resonant components there was also the anti-resonant component with a higher peak value. The coupler's tuned nature did not have any visible effect on frequency preference. Comparisons of the current's damped oscillatory decay time were good. The increased decay time at 50 kV/m level implies that the transient protection device triggered and protected sensitive semiconductors. The current that passed through the coupler was always limited to 12 amps, this also provided indication that the transient protection device triggered.

Overall, the combination of NEC and MATHCAD provides a reliable tool to predict the worst case voltages and current induced on complex antennas when subjected to short duration electromagnetic pulses.

REFERENCES

[1] M. J. Packer, Energy Transfer from Free Space Transient Waveforms Through HF Antenna to Arbitrary Loads" Applied Computational Electromagnetics Society 13th Annual Review of Progress in Applied Computational Electromagnetics, Monterey, CA, March 1997.

[2] G. J. Burke, "Numerical Electromagnetics Code - NEC-4, Method of Moments," Part I: User's Manual, Lawrence Livermore National Laboratory, Jan. 1992.

[3] MATHCAD, MathSoft Inc., 101 Main St. Cambridge, Massachusetts, 1995.

[4] I would like to acknowledge Rick Wilder and his crew at NAWCAD, Code 5.1.7.2 for outstanding effort running the pulser, collecting, and reducing the empirical data.

[5] C.E. Baum, "EMP Simulators for various Types of Nuclear EMP Environments: An Interim Categorization", IEEE Trans on Antenna and Propagation, Vol. AP-26, No. 1, January 1978.

[6] C.E. Baum, W.D. Prather, D.P. McLemore, "Topology for Transmitting Low-Level Signals from Ground Level to Antenna Position in Hybrid EMP Simulators", Sensor and Simulation Notes, Note 333, 30 September 1991.

[7] I.D. Smith, H. Aslin, " Pulsed Power for EMP Simulators", IEEE Trans. On Antennas and Propagation, Vol. AP-26, January 1978.

ADVANTAGES OF AN ALTERNATE VIEWPOINT WHEN DESIGNING HF VERTICALS FOR 80 AND 160 m

Rudy Severns
N6LF
PO Box 589
Cottage Grove, OR 97424
rudys@digiforest.com

Abstract

The traditional view of a vertical considers ground to be an integral part of the antenna. This paper suggests the use of an alternate viewpoint in which the vertical is a loaded, asymmetrical dipole in proximity to ground. The purpose for doing this is to visualize a much wider range of possibilities for a given situation and perhaps arrive at solutions which are simpler and less expensive than the traditional λ/4 wave vertical with 120 long radials, but are competitive in performance.

Introduction

The grounded vertical is one of the earliest radio antennas and is widely used today by amateurs, particularly for 80 and 160 meters. VHF verticals with "ground planes" are also widely used. The traditional way to visualize this antenna is to include ground as in integral part of the antenna – in effect supplying the "missing" part of the antenna since, at low frequencies at least, the vertical portion of the antenna is usually less than λ/2. Even when the antenna is not grounded but raised above ground we still use the terms "elevated ground system", "counterpoise ground", "ground plane", etc, etc. In this view we retain the concept that ground is an integral part of the antenna and that an ungrounded vertical must have some structure which replaces the "real" ground. While this conceptual framework has served us well for over 100 years it tends to limit our thinking to more traditional solutions. A change in viewpoint might expose useful variations better suited for particular applications.

The traditional view is that a λ/4 vertical with a ground system of 100 or more, long radials is the ideal and that anything else is an inferior compromise. Recent work[1,2] using primarily NEC modeling, has indicated that elevated ground systems with only 4 to 8 λ/4 radials are very competitive with the more traditional 120 buried radial antenna. However, elevated radial systems have their own drawbacks such as non-uniform radial currents[3], which lead to asymmetrical patterns and perhaps increased loss, and the need for an isolation choke at the feedpoint. A network of wires, λ/2 in diameter, suspended above ground may be even more trouble that simply burying the wires. There is good reason to believe that the traditional λ/4 radials used in elevated ground systems are a poor choice[3,4] and other arrangements may be superior.

Most amateurs are severely limited by available space and the cost of towers and extensive ground systems. The traditional buried radial or even the elevated λ4 radial systems are frequently not possible. What is needed is a wide range of choices for the antenna structure from which to chose those best suited for a given situation. In as far as possible the final design should sacrifice as little performance as possible.

An alternate way to look at verticals was suggested by Moxon[4]:

1)The antenna is a shortened (<½λ) vertical dipole with loading. The loading may symmetrical or asymmetrical. The loading may be inductive, capacitive or a combination of both. Usually the

loading contributes little to the radiation although some loading structures may radiate.

2) Ground is not part of the antenna but the interaction between ground and the antenna must certainly be taken into account. This includes both near and far field.

This seems a trivial conceptual change but looking at a vertical as a short, loaded dipole rather than a grounded monopole, opens up possibilities not usually considered with the more traditional point of view.

Loaded Dipoles in Free Space

One of the simplest ways to load a shortened dipole is to add capacitive elements or "hats" at the ends as shown in figure 1. As indicated in figure 1, the feed point may be anywhere along the radiating portion of the antenna. Figure 1 shows symmetric end loading. Figure 2 gives an example of extreme asymmetric loading where only one capacitive loading structure is used. This is of course the familiar ground-plane antenna being viewed as an asymmetric dipole. Actual antennas can vary between these two extremes, adjusting the size of the loading hats to suit a particular application.

When the vertical portion of the antenna (h) is < $\lambda/4$, top loading is quite commonly employed. However, top loading is usually not considered when $h \geq \lambda/4$ or more is used. This may be due to our past view that we need an extensive set of buried radials or equivalently an elevated system of $\lambda/4$ radials which "complete the antenna". In fact there are compelling reasons for adding some form of top loading or inductive loading even if the vertical section is a full $\lambda/4$. For a $\lambda/4$ vertical the diameter of the radial system will be $\approx \lambda/2$, changing only slowly as the number of radials is varied. On the other hand, if we lengthen the vertical section beyond $\lambda/4$ or add some top loading or even some inductive loading, the diameter of the radial structure drops rapidly, seemingly out of proportion to the added loading.

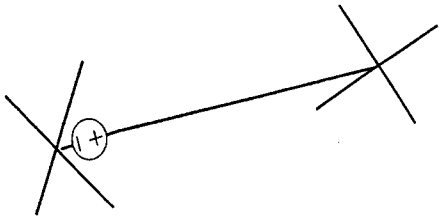


Figure 1, Short loaded dipole

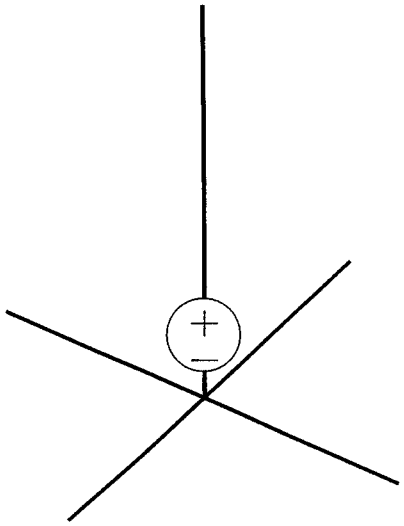


Figure 2, Asymmetrical dipole

A simple example that illustrates this point is given in figures 3 and 4. Figure 3 shows an asymmetrical $\lambda/4$ dipole with two radials (L_1 and L_2) at each end. L_2 is varied from zero to 22.3' and L_1 readjusted to resonate the antenna at 3.790 MHz. Clearly the addition of even a small amount of top loading (L_2) greatly reduces the length of the bottom radials (L_1) and consequently the land area required for installation. This is a matter of considerable practical importance to those with

restricted space in which to erect an antenna. With somewhat more complex loading elements the footprint can be reduced even further.

In addition to greatly reducing the length of the radials a number of other things happen during the above exercise:

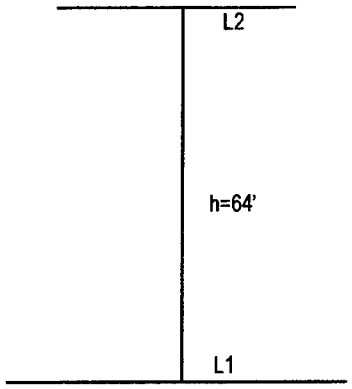


Figure 3, Asymmetric two radial dipole. Fr = 3.790 MHz

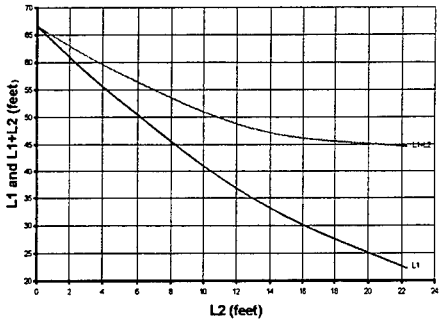


Figure 4, Effect on radial length of top loading

1) With only two radials and no top loading, the radiation pattern will vary with azimuth by about .7 dB, making the pattern slightly oval. This pattern asymmetry essentially disappears as the radials are shortened.
2) When placed over ground, the current in individual $\lambda/4$ radials will rarely be equal. This can lead to asymmetric patterns and increased loss. The current asymmetry rapidly decreases as the radials are shortened.
3) The peak gain and the angle at which it occurs, changes relatively little as top loading is added and the radials shortened.
4) Small amounts of inductive loading could also be used to supplement or even replace the top loading. As long as the vertical section is close to $\lambda/4$, the radials lengths can be reduced to $\lambda/8$ without seriously increasing losses.

Modeling Issues

The realization that everything, from the length of the radiator to the type and distribution of loading, is a potential variable which may be adjusted to achieve specific ends is a very liberating idea but it brings it's own set of problems. Which variations are best for a given application? A multitude of questions arise when judging any particular variation.

The possibilities and questions cannot be dealt with analytically, at least beyond an elementary level. The only practical way to deal with the variables is to systematically explore the possibilities with NEC, MININEC or other CAD modeling software. But even that is not a simple matter. Each modeling program has particular strengths and weaknesses that affect its use for this problem. The bottom portion of a vertical for 80 or 160 m will usually be very close to ground (<.05 λ). The modeling software should implement the Norton-Sommerfeld ground and properly model the current distribution in the lower part of the antenna as modified by induced ground currents. The loading structure may consist of a web of wires with multiple wires at each junction, perhaps of different diameters, and with small angles (<90°) between adjacent wires attached to the same node. MININEC based software can model

multiple acute angles if segment tapering is used but if many wires are used in the structure the number of segments becomes quite large. MININEC Broadcast Professional, using a different segment current distribution, does an even better job without the need for tapering. However, both of these programs do not model the interaction properly for very low antennas over real ground. NEC2 can model the ground effects correctly but may not handle the multiple small angles properly, especially if different diameter conductors are connected together. NEC4 is much better in this respect but is not widely used by amateurs because of the expense.

Real grounds are frequently stratified beginning only a few feet down. On 160 m the skin depth is of the order of 15-20' and it is not uncommon to have several different layers with different electrical properties in this distance. Even in homogeneous ground the effect of rain and subsequent drying will create a non uniform conductivity profile. None of the presently available software addresses this problem. The validity of NEC2/4 modeling for ground has been questioned because of differences between experimental measurements and predictions made by modeling. This is a critical issue. If NEC is fundamentally deficient with regard to ground modeling then the comparisons to date between buried radial and elevated radial systems are invalid. That includes the work reported in this paper! On the other hand, NEC modeling may be fine but the problem lies with the highly non uniform nature of real ground, particularly down to depths of 15-20', which cannot be simulated with NEC but which could greatly modify experimental results. Support for this view comes from experimental work at higher frequencies, where the skin depth is much less, where the modeling predictions are in much better agreement with experiment.

The presently available software, while a remarkable achievement, is not totally satisfactory to fully exploit the possibilities which the suggested point of view brings out and a great deal of care must be used when modeling a vertical with a complex loading system near ground.

A Design Example

The advantages of employing this concept can be illustrated by the 160 m vertical used at N6LF where an effective antenna was built on a very difficult site at low cost.

The site available was on a narrow ridge ($\approx$ 60' wide at the top) in a forest. There was no possibility of installing an extensive buried radial system due to the dense forest, with heavy underbrush, steep slopes and very large old growth stumps. Even an elevated system of normal size was not practical.

A support for the antenna was constructed from three trees, bolted together in the form of an A-frame. This resulted in a support 135' high. Allowing 8' spacing above ground for the bottom of the antenna and a few feet of slack at the top to allow for sway in high winds, the final vertical length was 120', very close to $\lambda/4$. The antenna was designed for a 75 Ω feedpoint impedance.

The final antenna is shown in figure 5. Four radials connected at the ends with a skirt wire were used at the bottom. The diameter of the bottom loading structure is only 40', compared to 260' for normal $\lambda/4$ radials. Two sloping wires were used for loading at the top. The use of sloping wires for loading may not be optimum but is very simple and has the advantage of allowing the antenna to be tuned by changing the angle of the wires. This can be done from the ground by shifting the attachment points for the lines connected to the ends of the sloping wires.

Christman's[1] comparison between a 120 buried radial vertical and an elevated 4 radial vertical, both with h= $\lambda/4$, indicates that the gain and radiation pattern differences between the antennas are quite small: .35 dB for peak gain, 1° for peak gain angle. Because the difference is so small I have chosen to use the 4 radial elevated antenna as the reference antenna because it is much easier to model than a complete 120 buried radial antenna.

Using NEC4D for modeling, a radiation pattern comparison between a four

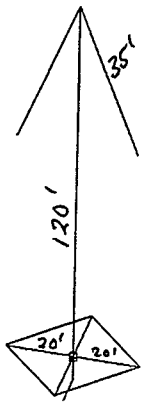


Figure 5, Antenna configuration

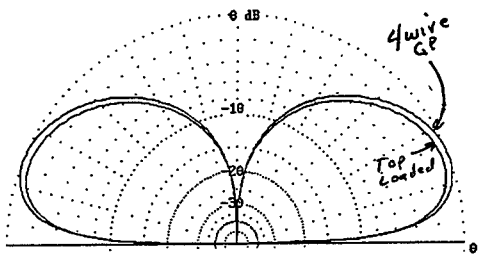


Figure 6, Comparative radiation pattern

radial ground-plane antenna and this antenna is presented in figure 6. Average ground was assumed (σ=.005S/m, ε=13). The wire used was #13 copper and loss was included in the modeling. The price paid for drastically reducing the diameter of the bottom loading structure is a peak gain reduction of 0.5 dB. This is a fair trade for dramatically easing the installation of the lower loading element because 0.5 dB will probably not be detectable in actual operation. In the real world where full size radials will very likely have non-uniform currents[3], the reduced size antenna may in fact not be inferior at all.

Any antenna with an elevated radial system needs an isolation choke (balun) on the transmission line near the feedpoint. One of the effects of moving the loading from the bottom to the top of the antenna is to increase the potential between the bottom and ground. This requires more inductance in the isolation choke to properly decouple the transmission line. For this application I happened to have a roll of ½" hardline. The roll was about 2' in diameter so I expanded it into a coil 3' long and 2' in diameter with a simple wood framework to hold it in place. The result was a choke with 350 uH of inductance (4 kΩ at 1.840 MHz). When this value of inductance was placed in the model there was still some interaction, resonance was displaced downward. On the actual antenna this was also found to be true. This illustrates one of the drawbacks of very small bottom loading structures, it may not be practical to have enough inductance in the choke to avoid some interaction, at least on 160 m. The Q of the choke must be high to limit losses.

More Modeling

In the process of developing this antenna a great deal of additional modeling was performed to explore the effect on performance of different loading arrangements. One of the more interesting variations was a symmetrically loaded, two radial antenna called a Lazy-H vertical[5]. This antenna was intended to be supported between two trees. The antenna is identical to that shown in figure 3 with L1=L2. Table 1 shows gives a comparison between a λ/2, λ/4 with 2 and 4 radials and the lazy-H with different values of h varying from 120' down to 30'. Note that the λ/4 lazy-H is within 0.3 dB of the 4 radial λ/4 vertical and has greater bandwidth. If two supports are available the lazy-H would be much easier to fabricate. In the design example shown earlier the top loading structure was simply a pair of drooping wires lead to anchor points near ground. The question arises as to the comparison between flat configuration, like that shown for the lazy-H, and the drooping wire alternative. This question can be quickly answered

Table 1. Antenna Comparison at 3.510 MHz

ant	h	L1 = L2	Z_{middle} Ω	Z_{end} Ω	peak gain, dB	peak angle °	wire loss -dB	2:1 SWR Bw kHz
λ/2	137'	0	91	>5000	+.30	16	.08	270
lazy-H	120'	4.4'	96	1096	+.28	17	.02	280
"	100'	10.4'	94	384	+.12	19	.07	280
"	80'	17.4'	81.3	180	-.06	20	.08	260
"	69.8'	21.6'	71.2	127	-.07	21	.09	240
"	60'	26.3'	59.7	90.9	-.15	22	.10	200
"	40'	38.3'	33.7	40.8	-.38	24	.16	140
"	30'	45.6'	21.5	23.8	-.59	25	.23	100
λ/4 2 radials	69.8'	———	———	38.8	.11/-.39	22	.15	200
λ/4, 4 radials	69.8'	———	———	35.7	+.21	22	.13	175

by modeling an end-loaded dipole in free space with two different configurations as shown in figure 7. The results of modeling show that the drooping wires must be made longer to achieve resonance, the radiation resistance is significantly lower with drooping wires and the far-field pattern is essentially the same. From a practical point of view the use of drooping wires greatly simplifies the structure and has very little effect on the far-field pattern but may reduce the efficiency of the antenna if the radiation resistance is lowered too much. This is the kind of trade-off information which critical to a new design.

In general the modeling of this class of antennas shows that the primary determinators of peak gain and peak gain angle are ground characteristics and the height of the vertical radiator (h). The loading means has only a second order effect on the radiation pattern and a wide variety of loading arrangements can be used to satisfy a particular situation with little loss of performance as long as attention is paid to keeping the radiation resistance high enough to control losses.

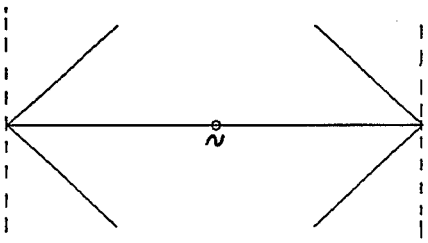


Figure 7, Flat versus drooping loading wires

Conclusions

This paper has advocated the adoption of a different conceptual view of vertical antennas: that they be viewed as loaded dipoles close to ground. The object of changing the point of view to make it easier to recognize the wide range of options available for configuring a high performance vertical to meet the needs of particular site and set of limitations. Properly assessing the many possibilities requires the use of modeling software. Unfortunately none of the available software packages can provide the computational capabilities desired at a cost attractive to amateurs. Users of MININEC and NEC2 based software must be very careful in modeling and interpreting results.

Acknowledgement

In addition to the referenced papers other workers in this field have pointed out the

advantages of the point of view presented here. In particular I am indebted to Dr. L. B. Cebik, W4RNL and Dr. Grant Bingeman, KM5KG for their comments and support.

References

[1] Christman, A., Elevated Vertical Antennas for the Low Bands: Varying the Height and Number of Radials, ARRL Antenna Compendium, Volume 5, 1966, pp. 11-18
[2] Christman, A., Elevated Vertical Antenna Systems, QST, August 1988, pp. 35-42
[3] Weber, R., Optimal Elevated Radial Vertical Antennas, Communications Quarterly, Vol. 7, Number 2, Spring 1997, pp. 9-27
[4] Moxon, L., Ground Planes, Radial Systems and Asymmetric Dipoles, ARRL Antenna Compendium, Vol. 3, 1992, pp. 19-27
[5] Severns, R., The Lazy-H Vertical, Communications Quarterly, Antennas, Communications Quarterly, Vol. 7, Number 2, Spring 1997, pp. 31-40

Tower Equivalent Radius

William F. Cummins
K6MYH

Abstract

The equivalent radius of a lattice tower is calculated by comparing the scattered power from such a tower section with the power scattered from a series of cylindrical wires of various radii. The MOM code NEC is employed at two different wavelengths which are greater than the tower dimensions to produce a result that differs from various estimates seen in the literature.

I. Introduction

The question arises in various journals and news groups as to what one should use for the equivalent radius of a typical amateur tower when it is to be part of a HF radiating structure. Although, one, in principle, can detail the tower structure in the antenna calculation; it becomes quite complex and requires a very large number of segments in an MOM code.

In most real world cases this level of detail really is not required; since the total EM environment being modeled is not specified that well. While most amateur antenna structures, proper, can be specified accurately; their surroundings cannot. Ground conductance and permitivity generally are estimates, and all of the wires in structures and power lines within a few wavelengths of the antenna are not specified. Thus, treating a tower as one or more cylindrical wires is sufficient to the problem at hand. Furthermore, HF radiation patterns are weak functions of the wire radius. The primary effect of conductor radius is on feed point impedance generally varying somewhat as $\log_{10}(h/a)$ where h is the antenna height and a the radius.[1]

In this paper the equivalent radius of a 44 ft. section of Rohn 25G[2] tower is calculated by comparing the power scattered in free space from a vertically polarized plane wave at 100 KHz and 1 MHz with the power scattered from a series of cylinders of the same length. The wavelengths are chosen to be a good deal greater than the lengths involved, and the length, in turn, is much greater than the transverse dimensions of the tower and cylinders. The Rohn tower (Fig. 1) is triangular with a side dimension of 12 in. and individual section lengths of 16 in.

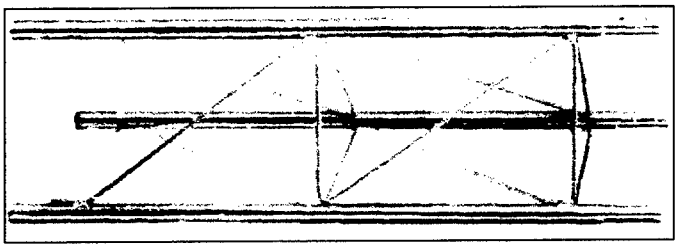


Fig. 1

There are a number of estimates for equivalent area that have been used. In Jasik[3] the equivalent radius of a triangle is given by a_{eq}= 0.4214a, where a is radius of the outscribed radius of the triangle. For the case in hand this gives a_{eq} = 2.92 in. , which is probably too small. The other two common estimates are equivalent area and equivalent perimeter. For an equivalent area we get a_{eq} = 4.7 in. and for equivlent perimeter a_{eq} = 5.73 in. As we will see later, the answer lies between these two values.

II. Calculations

In this exercise a 16 in. section of tower is generated and replicated to form the complete model. A vertically polarized plane wave illuminated this section at both 1 MHz and 100 KHz.. Also, the scattered power was calculated for the wave incident both upon a tower face and edge-on to one leg. Two widely separated frequencies were used as reliability check of the method and model. Fig. 2 is a sample input file.

```
CM      TOWER  SCATTERING X-SECTIONS     5/20/97    20:31:38
CM      INCHES SCALED TO MTRS.
CE
GW1,1,0.0,0.0,0.0,0.0,0.0,0.0,16.0,0.625        ! VERTICAL POST
GW2,1,0.0,0.0,16.0,12.0,0.0,16.0,0.1563         ! X-BRACE
GW3,1,0.0,0.0,16.0,12.0,0.0,0.0,0.1563          ! DIAG. BRACE
GM3,1,0.0,0.0,120.0,12.0,0.0,0.0                ! ROTATE & SHIFT FOR NEXT ELEMENT
GW7,1,6.0,10.3923,0.0,6.0,10.3923,16.0,0.625    ! 3RD POST
GW8,1,0.0,0.0,16.0,6.0,10.3923,16.0,0.1563      ! X-BRACE
GW9,1,0.0,0.0,16.0,6.0,10.3923,0.0,0.1563       ! DIAG. BRACE
GM10,32,0.0,0.0,0.0,0.0,0.0,16.0                ! REPEAT  VERTICALLY
GW400,1,0.0,0.0,0.0,12.0,0.0,0.0,0.1563         ! BOTTOM X-BRACE
GW401,1,12.0,0.0,0.0,6.0,10.3923,0.0,0.1563     !       "
GW402,1,6.0,10.3923,0.0,0.0,0.0,0.0,0.1563      !       "
GS0,0,0.0254                                    ! SCALE FROM INCHES TO METERS
GE                                              ! FREE SPACE
PT-1
EX1,1,1,0,90.0,0.0,0.0,0.0,0.0,0.0,1.0          ! VERT. POLARIZED PLANE WAVE
FR0,1,0,0,0.1                                   ! 100 KHZ
RP0,10,2,0001,0.0,0.0,10.0,90.0                 ! SCATTERING X-SECTIONS
EN
```

Fig. 2

Cylinder models of 4.0, 4.5, 5.0, and 5.5 in. of the same length as the tower and composed of 16 in. segments were used, and the scattered power was plotted at both frequencies. This data was used in conjunction with the power from the tower runs to interpolate a value for equivalent radius. All calculations were made at double precision.

III. Results

The tower calculations showed very little difference between broadside and edge-on illumination being 1.037 X 10^{-3} W. and 1.0421 X 10^{-3} W., respectively, at 1 MHz.; and 1.0177 X 10^{-7} W. and 1.0277 X 10^{-7} W. at 100 KHz. Averaging these values gave 1.0396 X 10^{-3} W. at 1 MHz and 1.0227 X 10^{-7} W. at 100 KHz. Fig. 3 is a plot of the data for both frequencies showing the Rohn 25G equivalent radius.

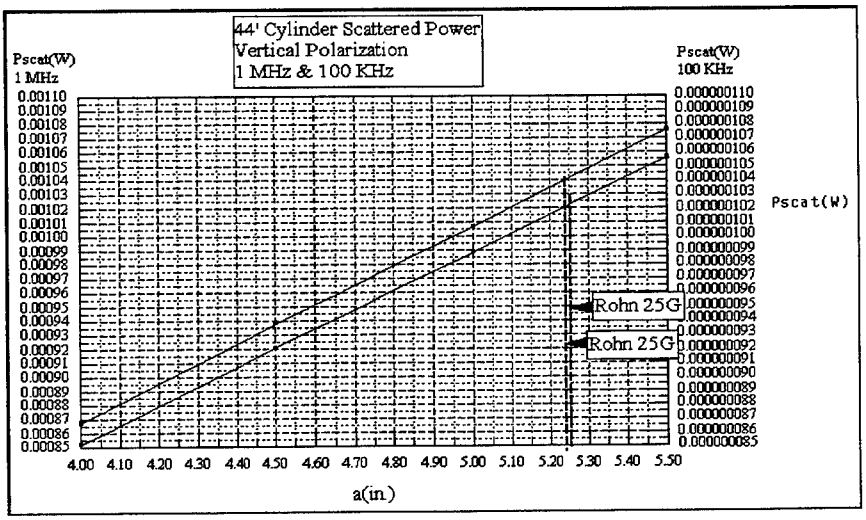


Fig. 3

The resulting equivalent radius for the Rohn 25G tower is 5.24 - 5.25 in.

IV. Discussion

The agreement between the equivalent radius obtained at two wavelengths differing by a factor of 10 is a measure of confidence in the method. In any case, comparison based on the scattering cross sections has a firmer physical foundation than some of the other guesses.

These results were used in the design of a multi-wire folded unipole antenna system for one of the local amateurs. A VHF tower and antennas near to the radiating structure and the use of elevated radials made for a sufficiently complex geometry that a detailed tower geometry was precluded. It will be interesting to see how close the input impedance predictions are on 75 and 160 meters when this design is constructed.

Calculations will be made for a variety of other tower designs to see if there is much difference between the mechanical details.

References

(1) Terman, Radio Engineering Handbook, p. 846, McGraw-Hill, New York, NY (1943)

(2) Unarco-Rohn, Division of Unarco Industries, Inc., P.O. Box 2000, Peoria, IL 61656

(3) Jasik, ed., Antenna Engineering Handbook, Sec. 3-6, McGraw-Hill, New York, NY (1961)

William F. Cummins, K6MYH
16298 Pearson Ln.
Fort Bragg, CA 95437
wfc@mcn.org
707-964-1256

Simple CP Fractal Loop Array With Parasitic

Nathan Cohen
Fractal Antenna Systems, Inc.,
5100 North Ocean Blvd Suite 1218
Ft Lauderdale FL 33308
fractenna@aol.com

I. Introduction

Circular polarization in wire antennas is acheived in a number of ways, including crossed elements with 90 degrees phase lag, or helical arrays. Here I present an alternative with a NEC4 model for a simple, circularly polarized, planar, fractal loop with a parasitic quad reflector. It may prove useful as a moderate gain system or be used as a front end element for CP in a dish or elsewhere.

II. Description

A simple, square Minkowski fractal loop of second iteration **(Cohen,1995)** was modeled with NEC4 in free space. However, in distinction from this previous work, the loop was loaded at the feedpoint to produce two current maxima which were out of phase for the feedpoint and it's nearby-spaced antipode. Such an arrangement naturally produces a small, circularly polarized fractal loop which is planar and end-fire. A Yagi-Uda type reflecting parasitic was then placed in parallel to the fractal loop and also loaded on the opposite side, for maximum end-fire gain. The array is shown in **Figure 1**, and relevant parameters are shown in **Table 1**.

Table 1

Fractal Loop Width: **0.154** waves
Quad width: **0.29** waves
Element Spacing: **0.2** waves

In the modeling, segment density was kept uniform with respect to the fractalized loop. A total of 100 wires comprised the loop while 4 wires made up the quad parasitic. Using EZNEC Pro with a NEC4 engine, the patterns and impedances were then ascertained in free space. Conservative wire limits were utilized and copper wire losses were incorporated into the field strength (gain) estimates. Loads were also loss-included and a QF of 200 was assumed. Load 1 (quad) was -260 ohms while load 2 (fractal loop) was -430 ohms. A wire width of 0.00025 waves was also assumed.

III. Results

Figure 2 reveals the (amplitude)current distributions of the elements, with the circle corresponding to the feedpoint and squares as load locations. The current distribution clearly phase shifts as the symmetry of the fractalized loop motif iterations provides both the orthogonal polarization components: the fractal structure acting as a phasing line. The quad acts purely as a reflecting parasitic and enhances the gain. **Figure 3** gives the end-fire pattern, gain **(a)**, and 3D power pattern **(b)**. The impedance gives an excellent match to a 50 ohm feed and the VSWR to 55 ohms is shown in **Figure 4**. This represents a 3 dB bandwidth of slightly less than 5%.

IV. Discussion and Conclusion

As a stand-alone element, this CP fractal loop provides a small planar antenna which could be incorporated at the front end of a dish or other system. With the parasitic below it, the array gives reasonable end fire gain for such a simple arrangement with low height.

The fractal elegantly provides a small antenna with a built in phasing arrangement for production of CP, suggesting that this may prove a powerful option in other applications where phase lags are needed for gain and/or polarization control.

Particular use of this array may arise at VHF and UHF satellite links, where broad bandwidths may prove a hindrance and the simplicity of the arrangement is obvious. LEO telecommunication may benefit from the broad elevation coverage for reasonable CP gain.

V. Acknowledgements

Aspects of this work are patent pending. Fractal Antenna Systems, Inc., is a licensed user of NEC4.

VI. Reference

Cohen,N., 1995, "Fractal Antennas Part 1", **Communications Quarterly**, Summer, 7.

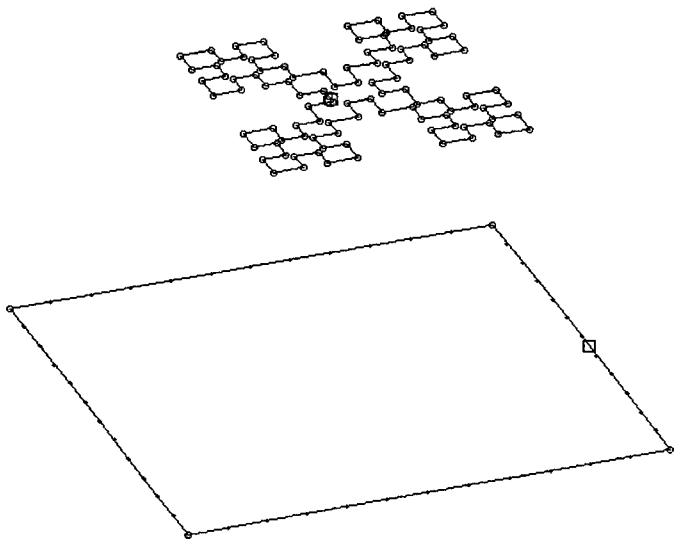


Figure 1

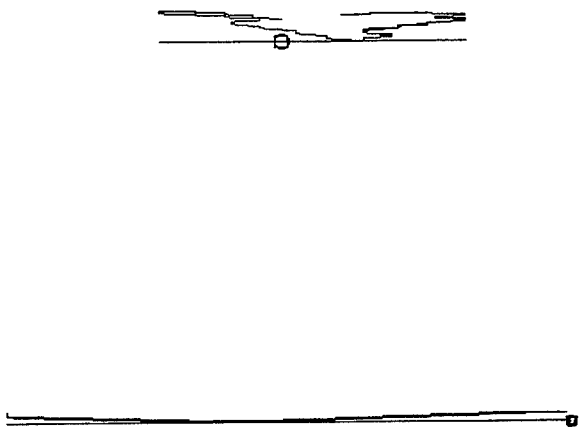


Figure 2

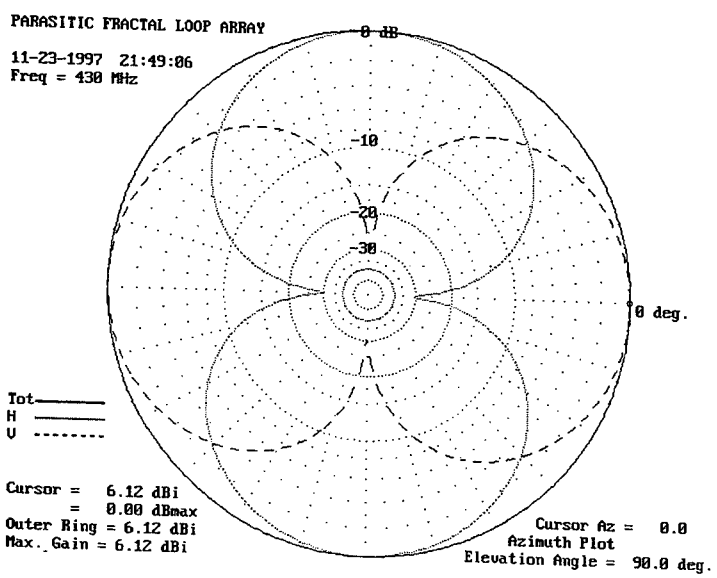


Figure 3(a)

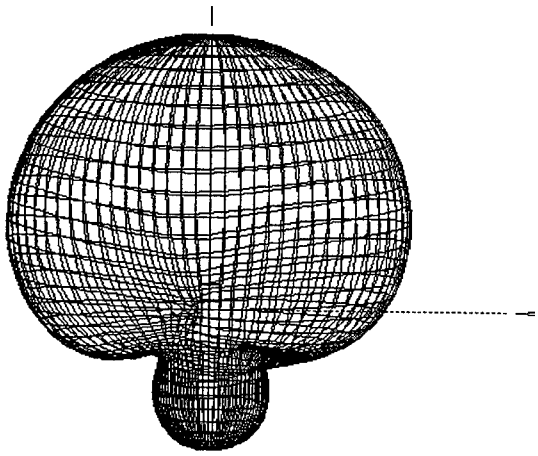


Figure 3(b)

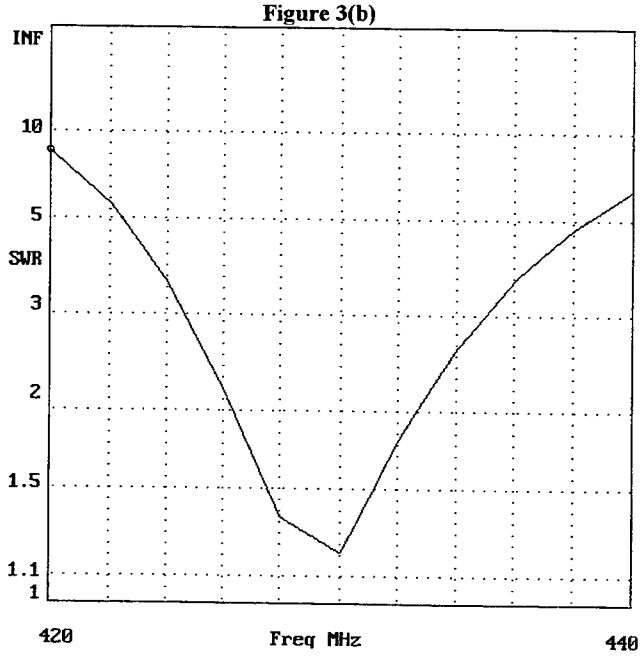


Figure 4

NEC4 Analysis of a Fractalized Monofilar Helix in an Axial Mode

Nathan Cohen
Fractal Antenna Systems, Inc.,
5100 North Ocean Blvd Suite 1218
Ft Lauderdale FL 33308
fractenna@aol.com

I. Introduction

The monofilar helix has proven to be a versatile antenna, with three main modes: normal; axial; and conical. In particular, the axial (Kraus) mode has been widely used due to its high end-fire gain, simple method of producing circular polarization, and broad bandwidth. Here we report the results of NEC4 simulations of such a fractalized helix **(Cohen, 1995a)** in an axial mode. The effect of the fractalization is to substantially shorten the length for a desired end-fire gain in the axial mode.

II. NEC4 Modeling

A multi-turn monofilar helix was fractalized with a Minkowski motif to a second iteration by generating an eight-fold symmetry as an approximation of a circle. The effect is similar to the fractalization of a loop **(Cohen,1995b)** and is related to the approach of meander line loading of a loop **(Pfeiffer ,1994)**; and axial mode helix **(Barts and Stutzman, 1997)**. A single turn's geometry is shown in **Figure 1**. The model was then extended to 4 turns and attached to a small ground plane, approximated by a wire mesh, and shown in **Figure 2**.

Segment and mesh density were enhanced near the feedpoint on the ground plane but otherwise kept uniform with respect to the fractalized helix. A total of 789 wire segments comprised the helix while 36 wire segments made up the ground plane. Using EZNEC Pro with a NEC4 engine, the patterns and impedances were then ascertained in free space. Conservative wire limits were utilized and copper wire losses were incorporated into the field strength (gain) estimates.

III. Results

In the axial mode for CP, a helix should maintain virtually a 1:1 axial ratio of orthogonal polarization components. The current falls to zero at the end of the helix and the power pattern is end-fire with little backfire component. These criteria were indeed met by the fractalized helix over a range of model frequencies. For convenience of comparison to an actual fractal helix, the resonance was modeled at and near 675 MHz. At this frequency, the length, diameter, and perimeter of the fractalized helix are shown in **Table 1**.

Table 1

Length: **0.345** waves
Diameter: **0.178** waves
Spacing Between Turns: **0.086** waves
Turn perimeter: **1.88** waves
Ground Plane diameter: **1/2** wave
Wire width: **0.001** waves

In **Figure 3** is shown the circularly polarized power pattern at 90 degrees elevation (end fire) and **Figure 4** shows the corresponding pattern from the side, at zero degrees azimuth. The axial mode criteria are met (at least) over 30 % bandwidth.

Figure 5 reveals the VSWR to a 280 ohm feed across the axial mode frequencies. The axial ratio remains very close to one in the end-fire orientation far beyond this range. The 2:1 VSWR bandwidth is approximately 20%.

IV. Comparisons

Figure 6 reveals the constructed helix , with an attached plastic radome, over a 1/2 wave ground plane. With radome removed, return loss measurements (S11) were done from 0-2300 MHz. A very large number of moderate/high Q resonances were found, which are believed to be normal or conical mode in nature; the details of which will be discussed elsewhere. However, this axial mode is readily apparent in **Figure 7**. The modest return loss (dB) and bandwidth are closely matched to that one would expect with the NEC4 results, which indicate typical real impedances of 180-350 ohms.

At 675 MHz, a circular helix was constructed with 4 turns and mounted above a 3/4 wave ground plane as described by **Kraus (1985)**. Length was 0.95 waves with turns approximately 1 wave in circumference. S12 measurements reveal the gain of these two are within 1 dB (0.5 dB RMS) of each other, with the fractalized helix being favored in the measurement over the circular helix. Both measurements were corrected for impedance mismatch to 50 ohms.

3D NEC4 Modeling of these two helices reveals very similar (within 0.3 dB) gains. However, as shown in **Figure 8**, there is a major difference in sidelobe structure between the two helices, favoring the fractalized helix, at least with this small number of turns.

V. Discussion

The fractalized helix provides for some surprising attributes in comparison to a (conventional)circular helix. In particular, the shortening of length--*almost a factor of* 3 *for a desired gain*-- provides for substantial practical advantages and opens up a variety of new applications often restricted to other antenna designs. Both the turn width and ground plane size are also substantially shrunk, allowing for more constraining form factors to be considered. Finally, the higher drive impedance, 280 ohms for the fractal versus 150 ohms for a circular helix, provides for greater efficiency to an already high efficiency antenna, suggesting that compromises in design dielectrics for support will have only minor impact on the antennas gain.

It should be noted that the smaller ground plane (1/2 wave) of the fractalized helix is not only a practical advantage, but an electrical necessity. The antenna will not function in the axial mode with a 3/4 wave diameter ground plane.

Clearly the fractalized helix is being *fractal loaded* and it is interesting to note two aspects of this. First, the fractal patterns are hardly defined by full radial symmetry. X and Y symmetry are apparent but radial is not. This is in contrast to assumptions made by **Barts and Stuztman (1997)** in a meander load or 'stub load' helix, and radial symmetry manifested by an opening angle, as seen in self-similar

arrays (log periodics). Fractal antenna work without this constraint of radial symmetry. Second, the perimeter of the fractal turn is very much larger than that for a circular helix (about 1.0 wave) or stub loaded helix (less than 1.5 waves). This suggests that if there is some fundamental restriction, it has not been acheived and further fractal loading, through different motives and/or further iterations, should enhance shrinkage and gain performance before succumbing to the restriction.

As with all antennas, including fractal ones, the tradeoff of size versus field strength versus bandwidth is evident. And here, unlike many fractal antenna designs, the tradeoff has favored size over bandwidth for the same gain. Hence the 'price to be paid' for these attributes is the smaller (20%) bandwidth. Yet the loss of bandwidth should not be too surprising given that the antenna is at a supergain condition, interpreted as meaning the gain is higher than that expected by the equation

$$G = 4\pi L \qquad (1)$$

with L in waves. At 0.345 waves length the 6.3 dB gain is virtually the same as that found from NEC4 modeling. It will be illustrative to compare these results with those from other fractal turn geometries and additional iterations. It should, in principle, be possible to use the fractal geometry to get one to a few dB more gain from a very short helix in the axial mode, and tradeoff bandwidth in turn. That effort may be additionally advantageous in many applications of point to point and satellite uplink telecommunications.

VI. Conclusions

In comparison to an axial mode monofilar circular-turn helix, this fractalized helix provides for substantially smaller size and better sidelobe response. Furthermore it requires a smaller ground plane. The 20% bandwidth represents the practical tradeoff for the shrunken size. Additional fractal motifs and/or iterations provide for yet higher gain in a small form factor.

VII. Acknowledgements

Aspects of this work are patent pending. Fractalized helix antennas are marketed under the trademark FracPole(TM). Thanks to Chuck Counselman for suggestions on ground plane mesh sampling. Fractal Antenna Systems, Inc. is a licensed user of NEC4.

VII. References

Barts, R., and Stutzman,W., 1997,"A Reduced Size Helical Antenna", **Proceedings of the IEEE AP-S Symposium,** p.1588

Cohen, N., 1995a, **patent pending**.

Cohen,N., 1995b, "Fractal Antennas Part 1", **Communications Quarterly**, Summer, 7.

Kraus, J.,1985, **Antennas**, McGraw Hill, New York.

Pfeiffer, A.,1994, "The Pfeiffer Quad Antenna System", **QST**, March, 28.

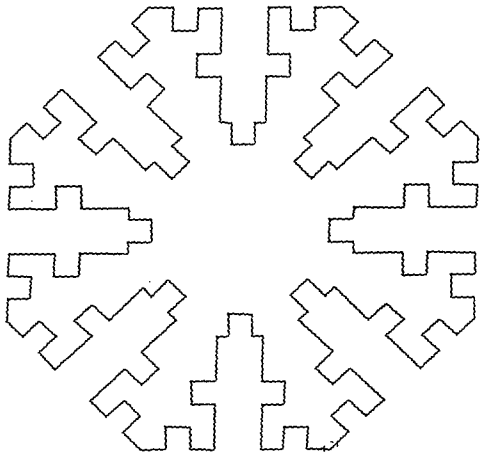


Figure 1

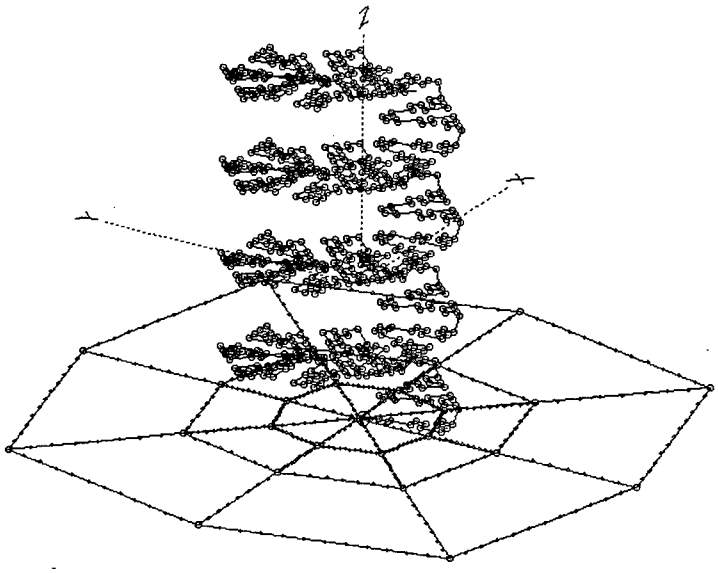


Figure 2

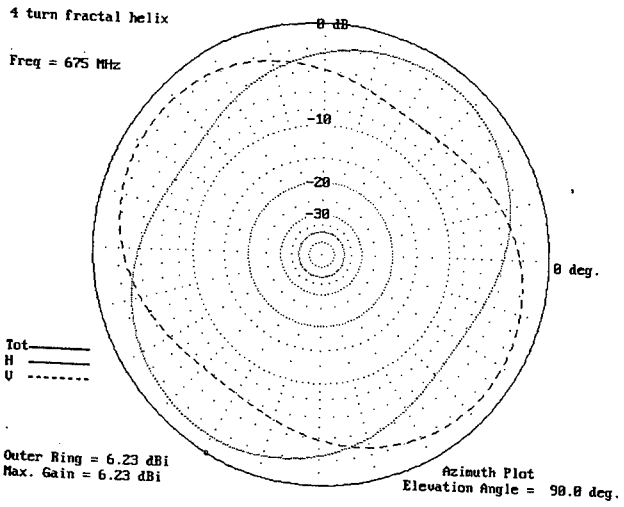


Figure 3

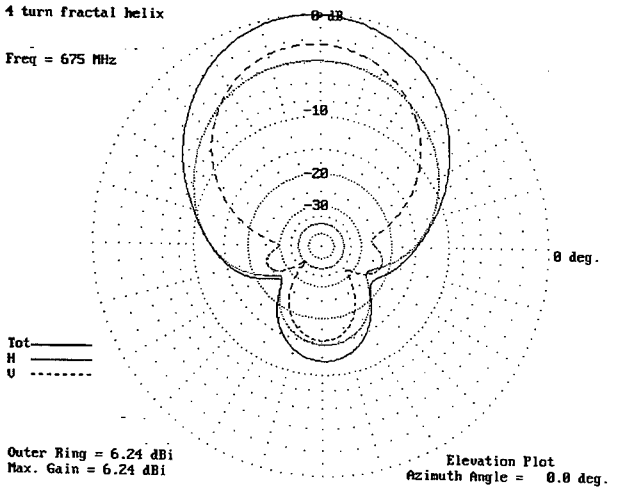


Figure 4

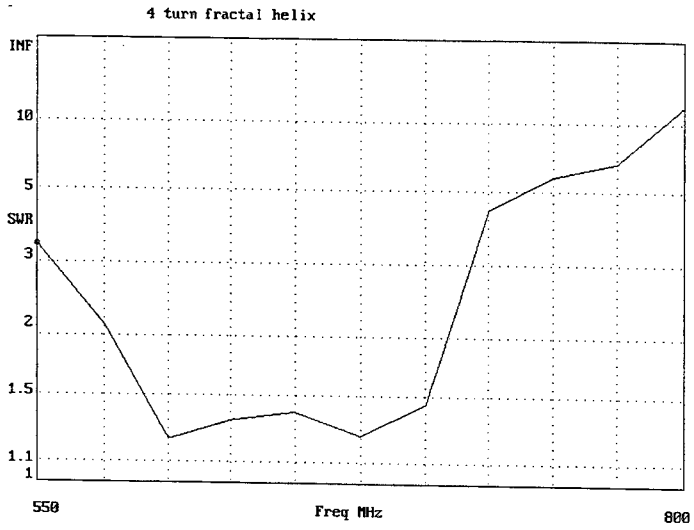


Figure 5

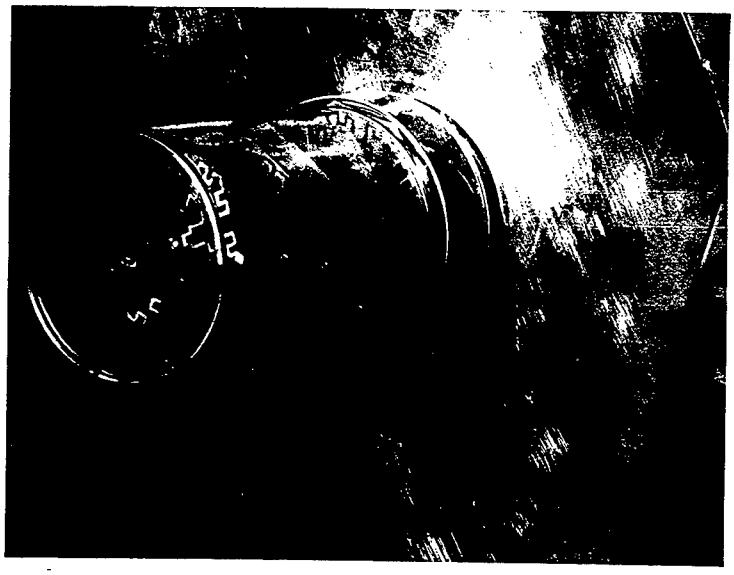


Figure 6

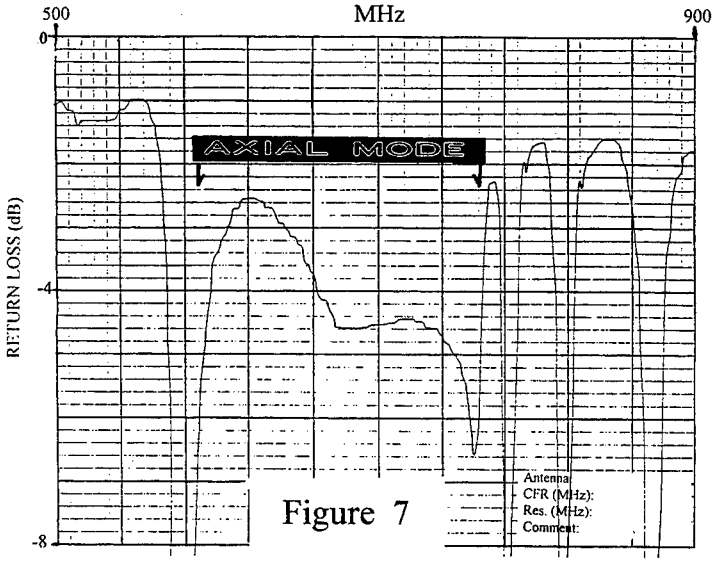
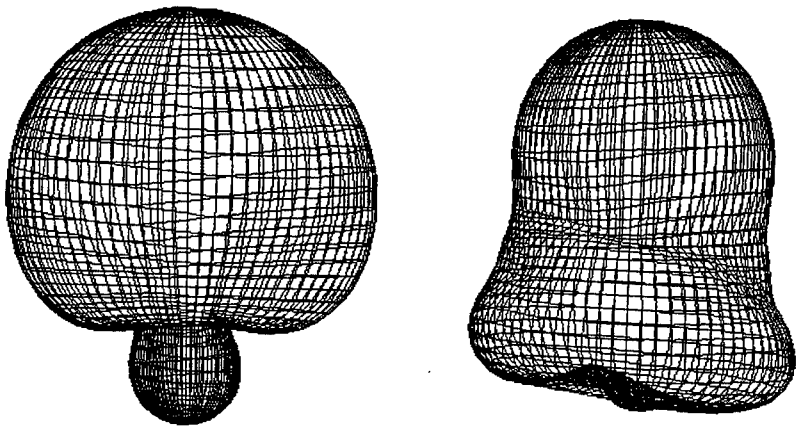


Figure 7



3D Fractal

3D Circular

Figure 8

Design of Low Sidelobe Antennas

Robert W. Hecht
California State University, Fresno
Fresno, California

Abstract

The design of a monopulse radar antenna, which would minimize the corrupting properties of jammers in the sidelobes of the antenna while tracking a skin (non-jamming) target in the antenna's main lobe, is described herein. Specifically, the antenna should have low sidelobes, particularly in the difference channels. The low sidelobes are achieved by selectively attenuating the gain in the slot elements of the antenna. Therefore, a min-max optimization algorithm was developed to find an optimal set of slot element attenuation factors between zero and one, which minimize the difference sidelobe gains.

The difference channel sidelobe levels were reduced by 14.15 dB beyond 18 degrees off boresight. MATLAB and the MATLAB Optimization Toolbox were used to design the antenna.

Introduction

The two main functions of the radar on a radar missile are to detect the target and to provide angles to the angle tracker. In order to maximize the probability of detecting a target in the presence of a sidelobe jammer, it is desirable to minimize the gain of the antenna's sum channel side lobes relative to the gain of the main channel. However, if a jammer is in the sidelobes, there would normally be considerable angle noise which is a significant contributor to miss distance. The reason for this stems from the fact that, typically, the gain in the difference channel is approximately 10 dB greater than in the sum channel (Figure 1), and jamming noise in the difference channel contributes directly to measured angle noise, while noise in the sum channel is only a minor contributor to measured angle noise. Figure 1 is the nominal antenna pattern of the antenna used in this study. That is, it is the original pattern with all slot element weighting factors set equal to unity.

If an antenna could be developed with low sum and difference channel sidelobes, then in the case of a jammer many miles to the rear of the screened target (stand off jammer), the missile could maneuver in order to put the jammer into the far sidelobes (angle off boresight greater than 18 degrees) where the sum and difference channel gains have been reduced.

How a Monopulse Radar Measures an Angle

A monopulse radar is one that can find the angle of a target off boresight (the normal to the antenna) with a single pulse. To accomplish this, the antenna of a monopulse radar is divided into four quadrants. The electromagnetic energy passes in and out of wave guide slots (or antenna elements) that are found in each quadrant. On a missile, the same antenna is used to both transmit and receive the energy.

It is assumed that the target is far from the antenna so that the energy from the target is a plane wave. It is also assumed that there is no cross-coupling between channels. Therefore, it is assumed that the antenna pattern can be calculated by multiplying the array factor by the element factor. The array factor is simply the fact that the wave front hits some slots before others when the target is not on boresight. Therefore, the array factor gives the relative phase for each slot (i) for target locations

at different roll angles (ϕ), and angles off boresight (θ). The element factor gives the pattern for a single slot element which, basically, is a function of the polarization. The element factor is a function of the roll and angle off boresight. It is assumed that the polarization is in one direction only; therefore, for zero roll, the element factor is unity for all angles off boresight.
The array factor is :

$$element = a_i\left[e^{\frac{j2\pi}{\lambda}\sin\theta(x_i\cos\phi + y_i\sin\phi)}\right]E$$

The a_i's are the slot element attenuation factors. Note that the antenna chosen for this research is a typical antenna that could be used on a missile and has 35 slot elements per quadrant, or 140 elements total.
The element factor is:

$$E = \sqrt{\frac{(\cos^2\theta\sin^2\phi + \cos^2\phi)\cos^2\left(\frac{\pi}{2}\sin\theta\sin\phi\right)}{(1 - \sin^2\theta\sin^2\phi)^2}}$$

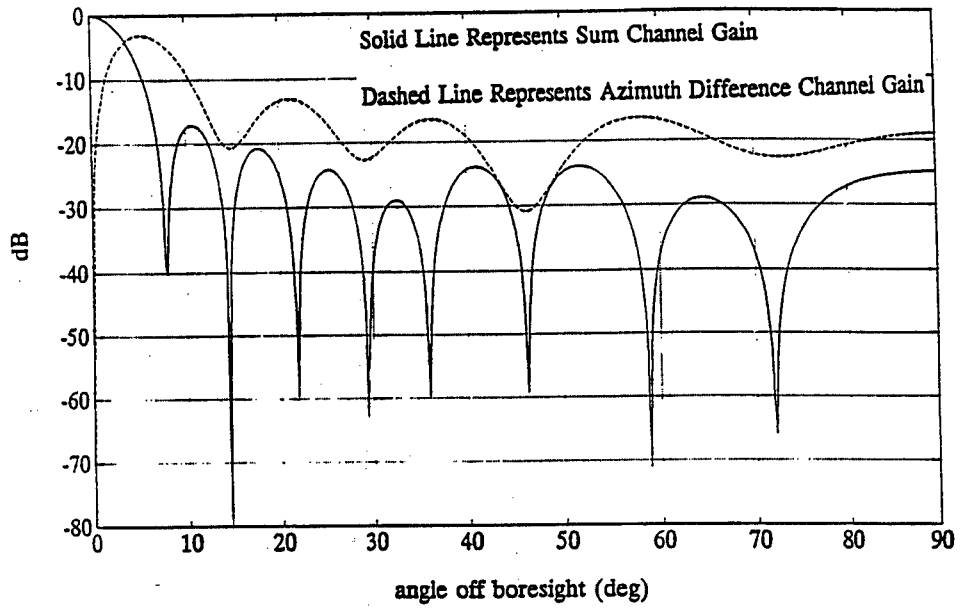


Figure 1. Nominal Sum and Difference Patterns

In order to form the sum, and azimuth and elevation difference signals, it is necessary to add and subtract the slots in the four quadrants; namely, the sum signal (S) is the sum of the four quadrants. The horizontal difference signal (D_H) consists of the difference between the left and right halves. The vertical difference signal (D_V) consists of the difference between the top and bottom halves. The real part of the resultant phasor gives the sum signal. To find the magnitude of the difference signals, one must take the imaginary part of the resultant signals.

In this algorithm, when the sum and difference signals are calculated, they are divided by the sum of the slot attenuation factors. The reason for this is twofold. First, when the antenna pattern is graphed, the peak of the sum channel is equal to unity. The second, and more important, reason is to prevent the min-max algorithm from setting all slot values to zero. That is, if the goal is to minimize the sidelobe levels, then zero for all slot locations would yield zero sidelobe levels. However, zero attenuation factors would also eliminate the sum and difference main lobe gains as well. Therefore, if the sum and difference signals are divided by the sum of the attenuation factors, then if the slots were set to zero, the sidelobe gains would be divided by zero which would not yield low values. Consequently, the min-max algorithm would not attempt to force the values to become zero.

The radar energy coming from a distant target is an electromagnetic plane wave that will strike the antenna at theta (θ), the angle off boresight. Unless theta is zero, the wave will arrive at some slots before others, and there will be a phase change from one slot to the next. This phase change is the basis for the angle measurement process. If one moves a source of radiation from a position normal to the antenna (boresight) to some large angle off boresight, one could plot the magnitudes of the output of the sum and difference signals to get a plot similar to Figure 1.

Since the sum signal near boresight resembles a cosine wave ($S = \cos(\theta)$) and the difference signal resembles a sine wave ($D = \sin(\theta)$), the measured angle (θ_m) is (essentially) equal to a constant K (described below) times the $\mathrm{Tan}^{-1}(D/S)$.

$$\theta_m = K * \mathrm{Tan}^{-1}(D/S) \qquad \text{(Equation 1)}$$

Since S and D are phasors, one can not divide the difference by the sum signal as indicated in Equation 1. There are various ways to achieve the measured angle without dividing phasors, and one of these methods is as follows:

If the target is in the main lobe of the antenna, it can be shown that the following angle discriminant gives the angle off boresight of the target:

$$\theta_m = K * [\angle (S+jD) - \angle (S-jD)] \qquad \text{(Equation 2)}$$

When the sum and difference signals are formed, it can be shown that they are ninety degrees out of phase; therefore, the signal S+jD can be obtained by simply adding the difference signal to the sum signal and S-jD can be obtained by subtracting the difference signal from the sum signal. The symbol $\angle$ refers to taking the phase of the signal. The constant K is necessary because the first zero cross of the sum pattern is not the spacial angle of 90 degrees, but is an angle less than that. Note that the first zero cross is sometimes called an "electrical 90 degrees". If the target is not in the main lobe, the angle discriminant will yield some random angle having no relation to the actual angle off boresight.

Let us assume that a skin target is in the main lobe and a jammer is in the sidelobe. Generally, when detection occurs, the radar is much closer to the target than to the jammer, even though the jammer is in the sidelobes, because the transmitted power of the jammer is usually greater than that of the radar power reflected off the target. Naturally, if the sum channel sidelobes are very small, then detection is facilitated. One can think of the effect of the jammer as an increase of receiver noise, thus reducing the signal to noise ratio (SNR) level. Note, however, that when a target is tracked, it

is on boresight, and the sum signal is at its maximum value while the difference signal is zero (or close to zero). The jammer corrupts both the sum signal and the difference signal, but noise in the difference channel is more deleterious than noise in the sum channel because the difference signal noise (from the jammer) is added to the very small difference signal from the tracked target. (Recall, the difference channel gain of the target on boresight is zero). Note also, that usually the difference channel gain in the sidelobes is much larger than the sum channel gain as shown in Figure 1; therefore, there is more angle noise than one would expect from a calculation based on signal to noise ratio if it was assumed that the noise in the signal to noise calculation was receiver noise. The reason for this is that the signal to noise ratio is based on the sum channel calculations only. Therefore, one of the problems of large difference channel sidelobe gains is a mismatch of predicted and actual angle noise, and this is a serious problem if the angle tracker employs a Kalman filter.

Consequently, the objective of this research is to find a method of designing a monopulse antenna with low difference channel gains in the sidelobes without raising the sum channel gains excessively. The method used to minimize the sidelobe gains is a min-max optimization scheme.

Min-Max Optimization Applied to Antenna Design

Parameter min-max optimization involves finding the minimum of a maximum function. That is, one has a function of two sets of variables **x** and **y** where **x** is the minimizing vector and **y** is the maximizing vector. For the example we are considering, the vector **x** is the set of attenuation factors $\{a_i\}$. For our example, the maximizing vector is equal to the roll angle phi, and the angle off boresight theta ($\mathbf{y} = \{\phi,\theta\}$). The min-max solution, denoted with asterisks, of $f(\mathbf{x},\mathbf{y})$ is the set of values $x1^*$, $x2^*$..., etc. which minimizes the max function $g(\mathbf{x}) = \max_{\mathbf{Y}} f(\mathbf{x},\mathbf{y})$. That is, in order to find the function $g(\mathbf{x})$, the function $f(\mathbf{x},\mathbf{y})$ is maximized over the multidimensional vector space y1, y2, etc. For example, to find the min-max solution of $f(\mathbf{x},\mathbf{y})$ one picks a set of values for the **x** variables and then maximizes $f(\mathbf{x},\mathbf{y})$ over the **y** variables. Then having this max function $g(\mathbf{x})$, one will calculate first and/or second order gradients (this research employed only first order gradients) at the various peaks of the max function. Then one finds a combination vector which is a combination of the above gradients.

The mathematical reason for calculating the gradient at not only the global maximum, but at all peaks close to the global maximum, can be found in Reference 1; however, an intuitive reason will be given at this point. Let us assume that the gradient is calculated only at the global maximum. Then the next step is to calculate a new set of slot attenuation factors. The new set is equal to the previous set (set initially equal to unity) and then a modification vector is added which is equal to some small stepsize times the negative gradient which was calculated at the global maximum. If one recalculates the peaks in the sidelobes, one would find that the peak that previously had been the global maximum has been reduced; however, one would also find that a peak that previously had been near in height to the global maximum is now a global maximum and its height may be actually higher than the previous global maximum.

Therefore, to prevent other peaks from rising up unexpectedly, it is necessary to calculate a direction to travel which will minimize not only the global maximum, but will minimize all other peaks close to the global maximum. This is accomplished by finding the gradients at the global maximum and at the peaks close in height to the global maximum and then finding a direction which is a combination of all of the above gradients. The derivation to the direction to move is found in Reference 1; however, the combination gradient direction is given as follows:

For a given set of attenuation factors **x**, let the global difference channel sidelobe peak, beyond

18 degrees off boresight, be denoted P_1. That is, $P_1 = \max_y\{D_H(x,y),D_V(x,y)\}$. Let all other difference channel sidelobe peaks within some small value of P_1 be denoted P_2, P_3, etc. Find the gradient of the horizontal and vertical difference channel antenna patterns with respect to the attenuation factors **x**, at the current set of attenuation factors, and at the phi and theta that gives the local peak of attenuation factors. Let these gradients be denoted $del(P_i)$ for the i^{th} gradient respectively. Let each gradient be a row in the matrix G, which consists of all gradients of local peaks within a small distance of the global maximum.

The direction to move is the result of the quadratic minimization problem, where the column vector $\mathbf{x}^*$ which is the minimum of $\mathbf{x}^T\mathbf{x}$, with the linear constraint $G\mathbf{x} <= -1$.

The new set of attenuation factors can now be computed from the previous set of attenuation factors augmented by a vector which is equal to a small stepsize times this combination direction (the output of the quadratic programming algorithm).

Having a new direction to move in the minimization space, one can pick a new value of **x**, find the max function and repeat. Note that even if the function f(x,y) is continuous and smooth over the full set of variables **x** and **y**, the max function g(x) is generally not smooth. Therefore, the minimum is typically at the base of a V. Consequently, the gradient is generally not equal to zero at the minimum.

The iterative procedure to find a low sidelobe antenna is to choose an initial set of weights. For example, an initial set of weights could be the nominal design where all of the weights are set to unity (equal weighting). Then the algorithm, after approximately 200 iterations, will terminate and give a set of weighting factors which produce the desired low sidelobe antenna. Various other sets of initial weights were chosen at random, and for most of these other starting conditions, the algorithm terminated at approximately the same set of final weighting factors as in the case where all initial weights were set to unity. Some sets of initial conditions did not lead to the same final set of attenuation factors; however, the final design for those cases had higher sidelobes than the design where the starting set of weights were all set to unity. The final design presented here had unity weights for the initial values.

Find The Maximum Peak(s)

The initial design criterion was to reduce the peaks of the difference sidelobes as much as possible beyond some number of degrees off boresight. The threshold number of degrees off boresight examined were 18, 16, 14, and 11 degrees, but only the cases where the sidelobes were reduced beyond 18 degrees will be discussed here. For angles less than this set number of degrees off boresight (e.g. 18), the sum and difference peaks were allowed to rise without limit. Therefore, the first step after choosing the initial attenuation factors was to find the maximum peaks in the sum and azimuth and elevation difference channels. The MATLAB Optimization Toolbox has a subprogram, called CONSTR for constrained minimization, which was used for this purpose. Since we are concerned with finding the maximum peaks of the sidelobes, and since CONSTR finds the minimum of a function, the sidelobes were expressed as the negative dB of the absolute value of the antenna pattern. Then, by finding the minimum of the negative pattern, we are actually finding the maximum peaks. An algorithm employing a constrained minimization is necessary since we are minimizing the sidelobes for some angle off boresight beyond some value (e.g., 18 degrees).

The subprogram CONSTR will only find a local maximum peak. That is, the subprogram CONSTR starts with an initial starting location for the roll angle, phi, and angle off boresight, theta, and then it is an iterative algorithm which moves in the gradient direction of the theta and phi space

until the maximum peak closest to the starting position is found. Since this peak may or may not be the global maximum (the absolute maximum peak), it is necessary to use a grid of starting points. It was determined by observing a large set of two and three dimensional graphs that the starting grid of points should consist of 12 points in the phi direction and 9 in the theta direction. Thus, the subprogram CONSTR is initiated from a total of 108 starting points in each of the three channels. Note that the antenna chosen for this study did not have 108 separate peaks in each channel; therefore, many duplicate peaks were found. Consequently, an algorithm was developed which would eliminate the duplicate peaks. The reason that the number of elements in the initial grid of starting points was greater than the total number of peaks was to guarantee that every peak would be found at least once.

The subprogram CONSTR uses the gradients of the antenna pattern with respect to the theta and phi maximizing variables. The gradient can be calculated either numerically or analytically. The antenna pattern must be calculated three times for each iteration step if the gradient is calculated numerically, and only once if the gradient is calculated analytically. Therefore, in order to have the algorithm operate at maximum speed, the gradient was calculated analytically. This entailed taking the derivative of the antenna pattern (the array factor times the element factor) with respect to theta and phi.

The initial design consisted of minimizing the azimuth and elevation difference channel peaks for angles off boresight beyond 18 degrees. There was no attempt to minimize either the sum or difference channel peaks for angles off boresight less than the threshold number of degrees (e.g. 18) nor were they prevented from rising above the original levels

Gradient of Antenna Pattern With Respect to Attenuation Factors

After finding the peaks in the region where they are to be minimized, one must find the overall peak (global maximum) and then all other peaks that are close in height to the global maximum. The next step is to calculate the gradient of the antenna pattern with respect to the parameters in the minimization space, and to calculate the gradient at the global peak and all peaks close to the global peak. The parameters in the minimization space are the slot element attenuation weighting factors. Each weighting factor is considered a dimension in the minimization space. Therefore, for an antenna with 140 slot elements, with 35 elements per quadrant, the minimization space has 35 dimensions. In other words, the minimization space is 35-dimensional. It is 35-dimensional instead of 140-dimensional because it is assumed that the four quadrants are symmetric. Consequently, it is necessary to minimize with respect to the elements in one quadrant; the other three will be the same.

Results

The maximum difference channel peak for angles off boresight beyond 18 degrees for the original nominal antenna (all weighting factors set equal to unity) is 13.06 dB below the sum channel main lobe peak. The comparable maximum difference channel peaks using the attenuation factors found as a result of the above algorithm are 27.21 dB below the sum channel main lobe peak. Thus, the above algorithm reduced the sidelobe peaks in the design space by 14.15 dB. The horizontal difference channel antenna pattern with sidelobes lowered 14.15 dB beyond 18 degrees off boresight is shown in Figure 2.

An added benefit of reducing the difference channel sidelobes is a reduction on the sum channel beamwidth. The sum channel of the main beam was reduced from 3.36 degrees to 3.16 degrees. A narrow beam is extremely desirable for tracking multiple skin or multiple jamming targets in the main

beam.

One of the less desirable consequences of the approach taken to reduce the sidelobe levels is a reduction of the sum channel main lobe gain. This gain reduction is unavoidable since, in order to shape the sidelobes, all but one of the antenna slot elements have been attenuated by various amounts. As a result, the total power entering the radar was reduced. For the design given, the total power is reduced by 5.5 dB.

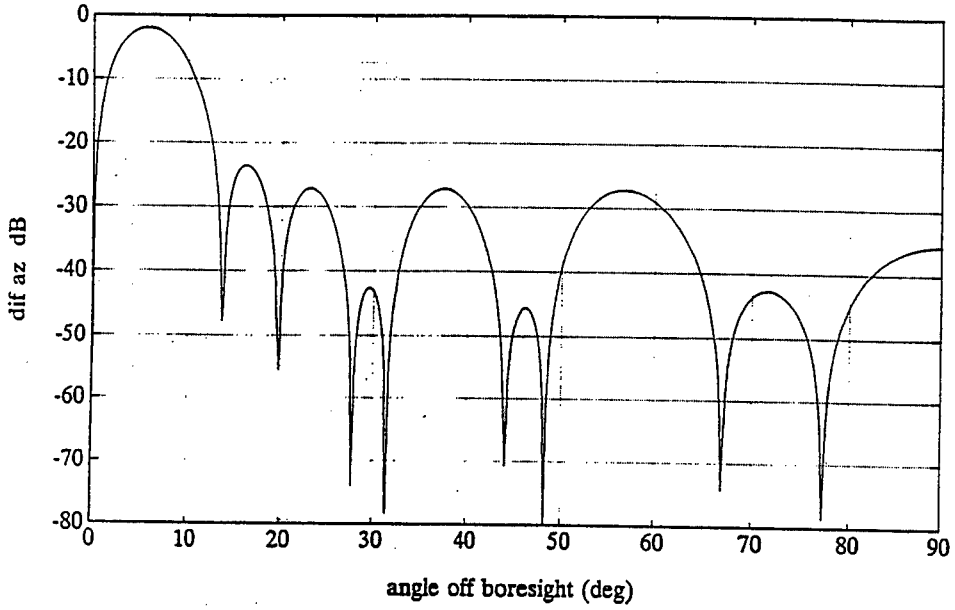


Figure 2. Horizontal Difference Pattern With Reduced Sidelobes

Support:

This work was performed during FY 1991 and FY 1992, was authorized and funded by Mr. Tom Loftus under Navy contract number N60530-91-C-0337, and the work was supervised by Mr. Sam Ghaleb, both of Naval Air Warfare Center, Weapons Division, China Lake, CA 93555

References:

[1] R. W. Hecht, "Algorithms For Min-Max Problems In Hilbert Spaces", Report R-548, Coordinated Science Laboratory, University of Illinois, Urbana, Illinois, January, 1972.

[2] J. E. Heller, "A Gradient Algorithm for Minimax Design", Report R-406, Coordinated Science Laboratory, University of Illinois, Urbana, Illinois, January 1969.

[3] J. Medanic, "On Some Theoretical and Computational Aspects of the Minimax Problem", Report R-423, Coordinated Science Laboratory, University of Illinois, Urbana, Illinois, July 1969.

[4] J. M. Danskin, The Theory of Max-Min, Springler-Verlag, New York, 1967.

[5] V. F. Demjanov, "Algorithms for Some Minimax Problems", Journal of Computer and System Sciences, Vol. 2, pp. 342-380, 1968.

[6] Fan Ky, "Minimax Theorems", Proc. Natl. Acad. Science, Vol. 39, 1953.

[7] H. Kunzi and W. Krelle, Nonlinear Programming, Blaisdell Publishing Company, Massachusetts, 1966.

[8] C. Wood, "Review of Design Optimization Techniques", IEEE Trans. on Systems Science and Sybernetics, Vol. SSC-1, No. 1, pp. 14-20, November 1965.

[9] W. Stutzman and G. Thiele, Antenna Theory and Design, John Wiley and Sons, New York, 1981.

[10] R. T. Compton, Jr. Adaptive Antennas, Prentice Hall, New Jersey, 1988.

[11] R. A. Monzingo and T. W. Miller, Introduction to Adaptive Arrays, John Wiley and Sons, New York, 1980.

[12] R. W. Hecht, "Comparison of Discriminants For Purpose of Tracking in Multiple Target Environments", Hughes Aircraft Company IDC, Confidential, 23 December 1981.

[13] R. W. Hecht, "Update to 6-DOF Receiver Model", Hughes Aircraft Company IDC, Unclassified, 17 February 1988.

[14] R. W. Hecht, "Addition of Clutter, Stand-off Jamming, Images, and Detection Logic Models to the AMRAAM 3-DOF Trajectory Simulation", Hughes Aircraft Company IDC, Confidential, 23 January 1985.

[15] R. W. Hecht, "Verification of Detailed Receiver Model for 3-DOF Trajectory Simulation", Hughes Aircraft Company IDC, Confidential, 6 February 1985.

EMP simulation of near field enhancement of wire antenna

M.EL HACHEMI, C.TOSSER-ROUSSEY and A.TOSSER-ROOSEY

L.P.L.I, I.P.E.M . Université de Metz

1.Bld ARAGO 57078 Metz cedex 03 FRANCE

Tel: 0387315866 Fax: 0387315801

Email: hachemi@ipc.univ-metz.fr

Abstract

EMP simulation is used for testing the capability of dielectrics windows (groved within the insulating layer of linear antenna) for enhancing the near electric field.

EMP calculation is based on the FDTD code, introducing absorbing limit conditions in the way of Majda et al [1], in order to obtain convenient procedures of calculation.

Good agreement by using these new limit condition is obtained with previous simulation of Maloney et al [2]. The effects of the geometrical parameters of dielectrics windows is then simulated for the sake of optimizing the enhancement of near field. Qualitative interpretations are presented.

INTRODUCTION

Prvious theoretical approximate equation [3] have given insight in the capability of dielectrics windows (around a wire antenna) for enhancing the near field. In order to obtain an exhaustive view on this subject a general simulation of the near electric field is practiced.

Previous works related to modelling electric field [4,5,6] have been published but only a few papers were related to the radiation of antennas [2,7]. For a complete analysis of radiation procedure in the FDTD method, arbitrary limits with total absorption capability is required. The usual absorbing conditions lead to heavy calculation and it is attempted to use analogous conditions previously used by Majda et al [1] for a quite different purpose.

PRACTICAL SIMULATIONS

EMP simulation is based on the following gaussian excitation [2] applied to a wire antenna. As shown in figure (5) to (7), the simulation results obtained from this new mixed procedure are in good agreement with those derived from the EMP procedure initially used by [2] . Varations in the near electric field due to the number of windows are shown in figures (2) and (3). Variations due to height of windows can be shown in figure (4).

DISCUSSION

It clearly appers that five windows, at least, are required for a marked enhancement of the near electric field in the vicinity of the wire antenna, practically in the top half of the wire. No significant further enhancement is obtained for more than five windows.

the effect of window height is not marked in the experimented range.

Radiation patterns due to the emission from gaussian pulse clearly show that the extent of the high near field domain is more important as the number of windows increasess.

All these results are in qualitative agreement with the assumption of multiple interferences due to windows.

REFERENCES

[1] B. Engquist and A. majda, "Absorbing boundary conditions for the numerical simulation of waves," Math. comp.,vol. 31, pp. 629-651, july 1977.

[2] J. G. Maloney, 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 Propagation. vol.AP-38,7, pp. 1059-1068, 1990.

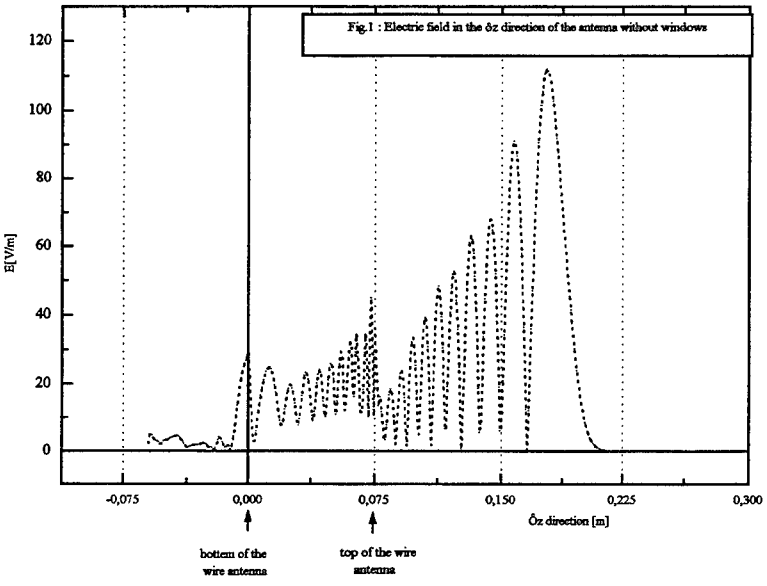
[3] Raymonde Mouecoucou "Effect of the heterogeneous dielectric upon metallic wire antenna at 915MHz" Thèse Univ.Metz.1995.

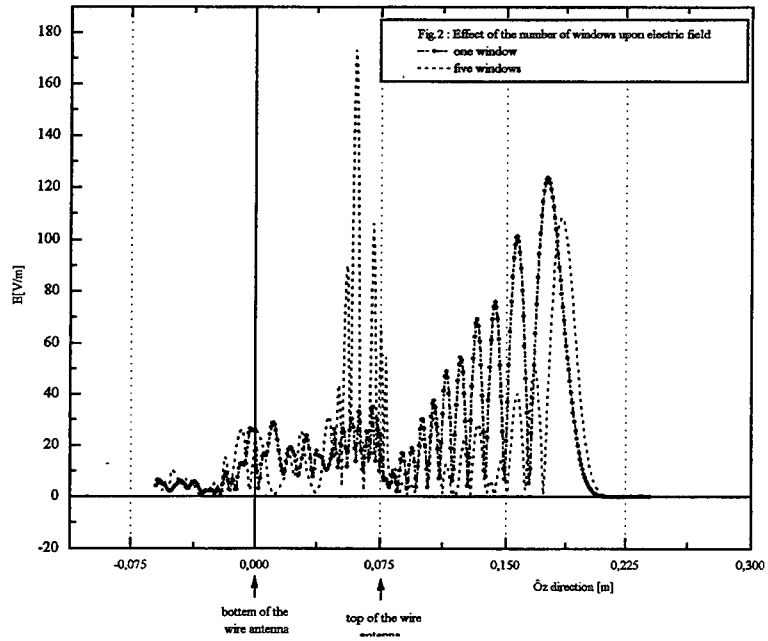
[4] R.Holand, L. Simpson, and K. s. Kunnz, "finite difference analysis of EMP coupling to lossy dielectric structures," IEEE Trans. Electromagn. Compat., Vol. EMC-22, pp. 203-209, Aug. 1980.

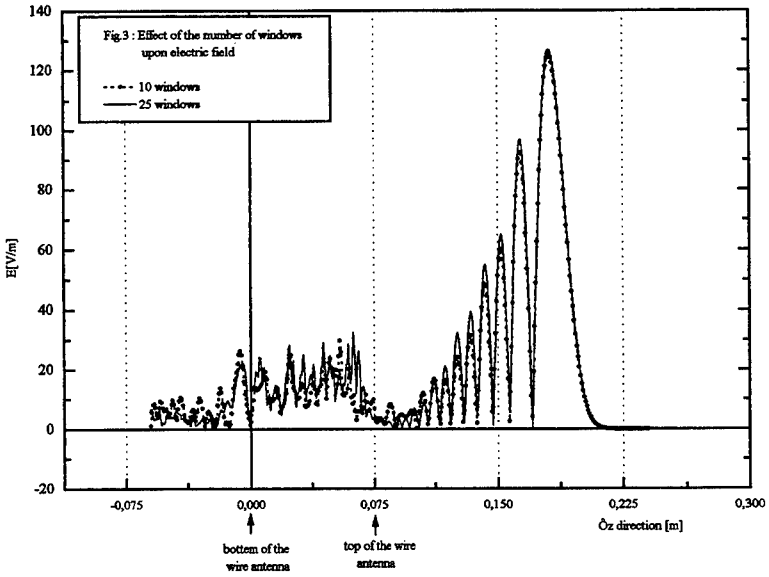
[5] A. Taflove and K. Umashankar, "A hybrid moment mrthod/finite-difference time-domain approach to electromagnetic coupling and aperture penetration into complex geometries," IEEE Trans.Antennas Propagation. vol.AP-30, pp. 617-627, July 1982.

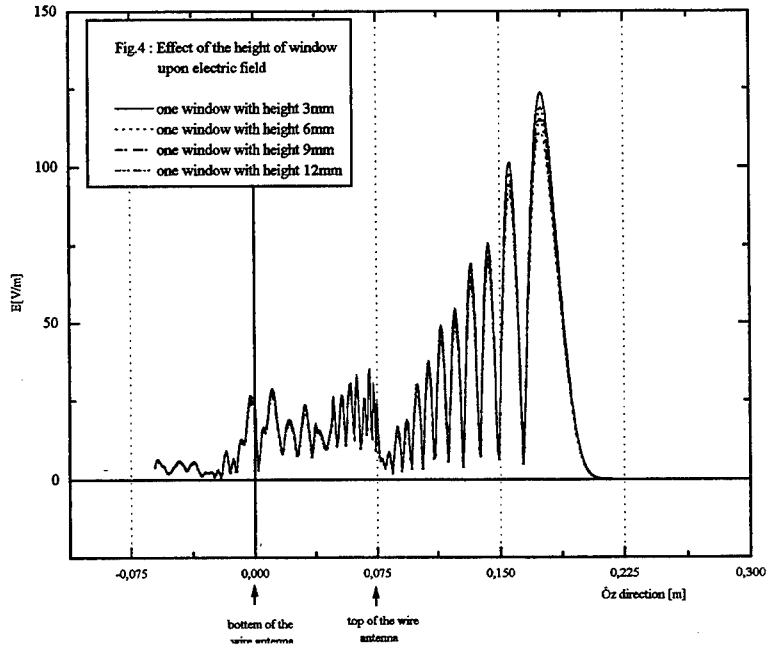
[6] A. Taflove and K. Umashankar, B. Beker, F. Harfoush, and K. Yee, "Detailed FD-TD analysis of electromagnetic fields penetrating narrow slots and lapped joints in thick conducting screens, " IEEE Trans.Antennas Propagation. vol. 36, pp. 247-257, feb 1988.

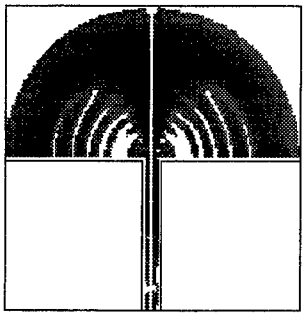
[7] P. A Tirkas and C. A Balanis, "finite-difference time-domain method for antenna radiation," IEEE Trans.Antennas Propagation. vol.AP-40, pp. 334-340, 1992


Fig.1 : Electric field in the ôz direction of the antenna without windows
E[V/m]
120
100
80
60
40
20
0
-0,075
0,000
0,075
0,150
0,225
0,300
Ôz direction [m]
bottom of the wire antenna
top of the wire antenna

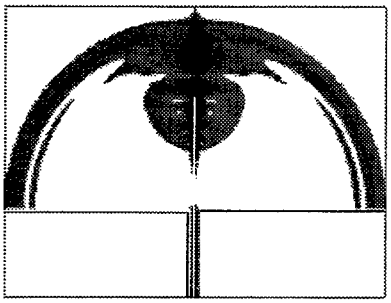

Fig.2 : Effect of the number of windows upon electric field
one window
five windows
E[V/m]
180
160
140
120
100
80
60
40
20
0
-20
-0,075
0,000
0,075
0,150
0,225
0,300
Ôz direction [m]
bottom of the wire antenna
top of the wire


Fig.3 : Effect of the number of windows upon electric field
10 windows
25 windows
E[V/m]
140
120
100
80
60
40
20
0
-20
-0,075
0,000
0,075
0,150
0,225
0,300
Öz direction [m]
bottem of the wire antenna
top of the wire antenna

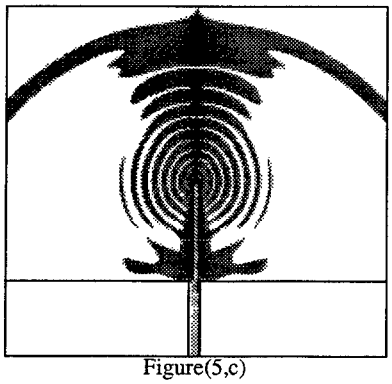

Fig.4 : Effect of the height of window upon electric field
one window with height 3mm
one window with height 6mm
one window with height 9mm
one window with height 12mm
E[V/m]
150
100
50
0
-0,075
0,000
0,075
0,150
0,225
0,300
Öz direction [m]
bottem of the
top of the wire antenna



Figure(5,a)

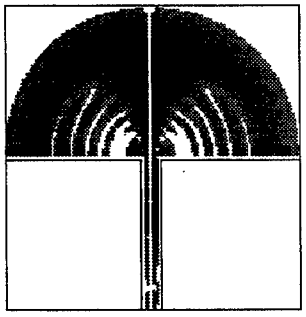


Figure(5,b)

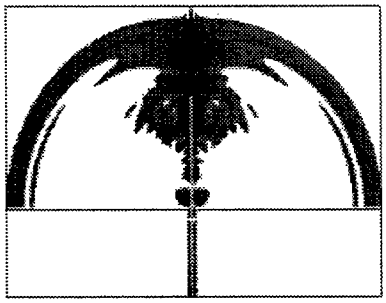


Figure(5,c)

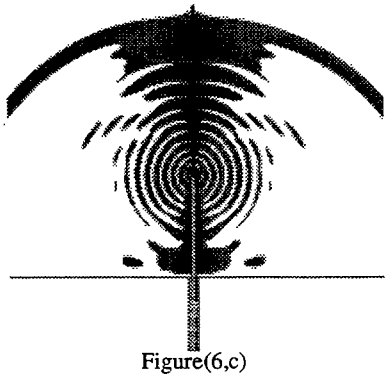
Figures(5,a,b,c)Radition of gaussian pulse of wire antenna without windows for tree different times. Gray scale show magnitude of electric field



Figure(6,a)

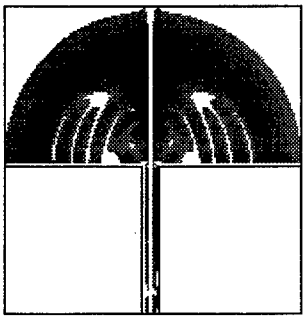


Figure(6,b)

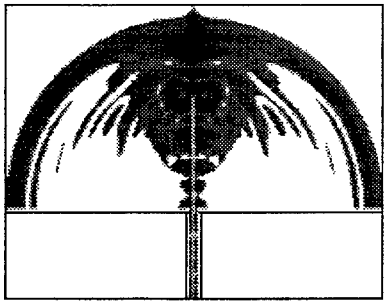


Figure(6,c)

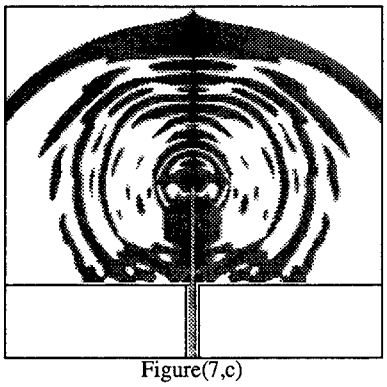
Figures(6,a,b,c) : Radition of gaussian pulse of wire antenna with one window for tree different times. Gray scale show magnitude of electric field



Figure(7,a)



Figure(7,b)



Figure(7,c)

Figures(7,a,b,c) : Radition of gaussian pulse of wire antenna with five windows for tree different times. Gray scale show magnitude of electric field

SESSION 19:

OPTIMIZATION TECHNIQUES FOR ELECTROMAGNETICS

Chairs:E. Michielssen and R.L. Haupt

Genetic Algorithm Design of the Conical Interdigitated Log-Periodic Antenna[1]

†P. D. Mannikko, ‡P. J. O'Brien and ††K. W. Ommodt

†Software Technology, Inc., 5904 Richmond Highway, Suite 610, Alexandria, VA 22303.
e-mail: paul_mannikko@alx.sticomet.com

‡GeoCom, Inc., 11800 Sunrise Valley Drive, Sixth Floor, Reston, VA 20191.
e-mail: pjo@geo-com.com

††Raytheon-TI Systems, 2501 West University, MS 8019, McKinney, TX 75070.
e-mail: ommodt@ti.com

1. Introduction

Genetic algorithms have recently been embraced in the computational electromagnetics community as a robust means to optimize designs over a multi-dimensional non-linear search space. Search algorithms offer the potential to enhance the design and understanding of complex antenna systems which, heretofore, have had limited engineering understanding. In this context, novel wideband log-periodic direction-finding antennas are being studied. Successful application of any search algorithm for wideband log-periodic antenna design requires 1) efficient numerical electromagnetics, 2) parametric geometry and grid generation capability and 3) parallel implementations for reasonable engineering design times. In this paper, we address these issues in the development of an efficient, network-parallel, genetic-algorithm optimization of the conical interdigitated log-periodic antenna (IDLPA)[2]. Results to date have yielded families of compliant designs and have shown correlations between antenna performance and antenna design parameters that were not previously known.

2. Conical Interdigitated Log-Periodic Array

In Figure 1(a), we represent a planar interdigitated log-periodic structure. It comprises N identical log-periodic antennas arranged radially and distributed uniformly in a rotationally symmetric array. Any single array element, depicted in Figure 1(b), is characterized by the usual log-periodic angle α and the scale factor τ defined as the ratio of sequential radiating element radii (ρ_n/ρ_{n+1}). Additionally, the radial transmission line which excites the antenna has subtended angle β and the width of each radiating element may be determined from a constant central arclength-to-width ratio r. We form a conical interdigitated log-periodic antenna, represented in Figure 1(c), by projecting or rotating the planar antenna onto a cone of half-cone angle γ. The parameter set $(\alpha, \tau, \beta, \gamma, r)$ forms a natural search space over which a genetic algorithm may be employed to optimize a design.

A typical application of this antenna is in the front-end of a wideband two-channel monopulse direction finding system. Here, two basic signal modes, the sum and difference modes, are received from which azimuth and elevation of the signal source may be estimated. Substantial calibration

[1]This work was partially supported by the Naval Research Laboratory under contract number N00014-95-C-2044.

[2]U.S. Patent No. 5,212,494 awarded 18 May 1993.

and measurement reductions are achieved if the antenna can be designed to be a circularly polarized receiver over the main lobe of the antenna. This forms the primary design constraint on the antenna. Additional constraints might include gain, input impedance, sidelobe levels and variations of each of these over a log-frequency period. Physical size limitations impose yet another constraint on the search. Since no published set of design curves exists which allow the engineer to design directly this antenna, genetic algorithms form a logical approach to automate and study the design and operation of this unique log-periodic structure.

3. DBOR Moment Method Solution

The interdigitated log-periodic antenna, along with most other direction-of-arrival antennas of its class, is a discrete-body-of-revolution (DBOR) of order M. A sequence of M $360/M$ degree rotations of a single "generating" arm about the axis-of-symmetry with subsequent duplication generates the entire structure. Assuming a moment method surface patch model of this antenna which is also discretely rotationally symmetric, there will be a total of MN unknown current coefficients which must be determined where N is the number of basis functions per generating arm. Since MN is typically tens of thousands for IDLPAs, substantial computational savings will necessarily be enjoyed if we employ the discrete rotational symmetry to advantage. To this end, we express the familiar conducting patch moment method matrix equation in a less familiar double index-set form

$$V_i[l] = \sum_{j=1}^{N} \left\{ \sum_{k=0}^{M-1} Z_{ij}[l-k] I_j[k] \right\} \qquad i = 1, \ldots, N, \quad l = 0, \ldots, M-1 \tag{1}$$

where an index in brackets refers to an arm and a subscripted index refers to a basis function on a given arm. The arm indices are ordered circularly such that the arm k geometry is obtained from arm l by a $2\pi(k-l)/M$ rotation about the axis of symmetry. The basis function indices are ordered identically from arm to arm. Thus, $V_i[l]$ is the forcing function tested across basis function i of arm l, $I_j[k]$ is the unknown current amplitude of the j^{th} Rau-Wilton-Glisson basis function on arm k, while $Z_{ij}[l-k]$ is the impedance matrix element representing the coupling from the j^{th} unknown current on arm k to the i^{th} basis function on arm l. The discrete rotational symmetry is manifested in the arm index difference $[l-k]$ in the coupling matrix. Recognizing the discrete circular convolution over indices k, l within the braces in (1), we discrete Fourier transform both sides of (1), yielding the reduced matrix equation

$$\tilde{V}_i[m] = \sum_{j=1}^{N} \tilde{Z}_{ij}[m] \tilde{I}_j[m] \qquad i = 1, \ldots, N. \tag{2}$$

valid for the m^{th} ($m = 1, \ldots, M$) orthogonal discrete Fourier mode. Tildes denote a transformed variables. Solutions of (2) are orthogonal mode solution of (1). General excitations require the superposition of all M modal solutions to synthesize the complete solution.

The rank of (2) is reduced by a factor of M from that of (1) resulting in an M^2 savings in memory requirements and an order M^3 savings in matrix factor time per mode assuming direct LU decomposition. Further reduction of numerical effort is obtained by noting that two-channel monopulse direction finding applications employ two orthogonal circular modes, the sum mode ($m = 1$) and the difference mode ($m = 2$). Equation (2) need only be solved twice in a complete analysis of IDLPA radiation.

4. Parametric Grid Generation

Engineering optimization of this antenna using computational electromagnetics involves three distinct processes. Geometry representation and gridding, numerical solution, and mensuration and assessment. While the second processes is inately computational and the third is readily automated, the process of geometry representation and subsequent grid generation suitable for computational electromagnetic codes is not well developed for automatic solution. While several successful efforts to marry CAD packages, grid generation and computational electromagnetic software have been developed, this approach is not conducive to automatic implementation in a design cycle due to the presence of the human interface to the CAD package. Even in non-automated approaches, the human-CAD interface may severely slow an iterative design process.

To remedy these defficiencies, we developed a suite of software libraries which enables automatic generation of an antenna grid given only the antenna design parameters. This library is implemented in a three level hierarchy. At the highest level are the algorithms which mathematically describe the antenna and sweep regions of space which define the physical surfaces or volumes of the antenna. The central level comprises packaged convenience routines with common primitive objects, *e.g.,* spiral arms or conic sections, and algorithms for their subsequent structured decomposition into common elements, *e.g.,* triangle, quadrilateral, tetrahedral or hexahedral elements. At the lowest level in the hierarchy, a global database of nodes, edges, faces, *etc.,* is formed which describes the decomposed structure. At this level, data are also output as complete, properly formatted input files to commonly available electromagnetic analysis codes including the PATCH Code, FERM and EIGER.

A small amount of invested time building a parametric representation of the antenna of interest at the top level of the hierarchy results in "immeasurable" savings in the design cycle time and allows automation of the engineering design process. Since the IDLPA geometry is completely defined by $(\alpha, \tau, \beta, \gamma, r)$ and its structural bandwidth, a grid generator has been devised which requires only these basic geometrical parameters to construct a grid suitable for computational electromagnetics and creates an input file to the conducting surface PATCH Code of Sandia National Laboratory modified for DBOR symmetry as described in the previous section.

5. Genetic Algorithm Implementation

We employ a steady-state replacement genetic algorithm with binary encoding of the search parameters, single-point crossover with mutation and power-law fitness scaling. Cost-functions σ are chosen to achieve desired $\hat{\theta}$ and $\hat{\phi}$-component sum-mode pattern beamwidths B_θ and B_ϕ, respectively, over a log-frequency period. A useful and simple cost-function for minimization is

$$\sigma = \sqrt{\frac{1}{N_f}\sum_{n=1}^{N_f}[BW_\theta(f_n) - B_\theta]^2} + \sqrt{\frac{1}{N_f}\sum_{n=1}^{N_f}[BW_\phi(f_n) - B_\phi]^2} \qquad (3)$$

where f_n is the n^{th} frequency in a log-frequency period sampled N_f times and BW_θ and BW_ϕ are the beamwidths of the $\hat{\theta}$ and $\hat{\phi}$ component sum-mode patterns, respectively, taken from defined azimuthal cuts. We choose the design goal beamwidths B_θ and B_ϕ to control the desired receive polarization properties as well as the component beamwidths.

The computational cost of evaluating the genetic algorithm is driven by three factors: 1) the size of the moment method problem, 2) the wide operational bandwidth over which the antenna

must be evaluated and 3) the large "antenna populations" which are evaluated in the genetic algorithm. Use of DBOR symmetries addresses the first factor. We address the second by evaluating the cost function over a single log-frequency period. Since the structure is log-periodic, one log-frequency period characterizes the entire operating bandwidth assuming structure end-effects may be neglected. The final factor is addressed, in part, through parallelization of the genetic algorithm. Evaluation of the genetic algorithm cost function is "embarrassingly parallel" on a per generation basis. Since genetic algorithm execution time is dwarfed by the time required to evaluate the cost function, a network-parallel approach is employed for cost function evaluation. Load-balancing and control over the network are achieved through the use of freely available network queuing software. This implementation offers many advantages including real-time control over when a given computer on the network will be available to the genetic algorithm and the execution priority of the cost function.

6. Results

We optimized the sum-mode beam-shape of a six-arm IDLPA where $\tau = 0.921$. We searched over a $24^o < 2\gamma < 50^o$ and $20^o < \alpha < 55^o$ space to minimize the cost-function in (3) with -10 dB design beamwidths $B_\theta = B_\phi = 80^o$ and $N_f = 5$ samples in a log-frequency period. The initial random population, represented in Figure 2(a), is 50 members followed by 8 member generational replacement. In Figure 2(b), we depict the results after 9 generations. Symbols on the plot represent parameter values which meet the sum-mode design specification that both E_θ and E_ϕ have a -10 dB beamwidth with an RMS deviation from 80 degrees of less than 4^o. Using the genetic algorithm, we not only found a design which met beamwidth specifications, we identified trends not easily determined, *i.e.*, the high dependence of the beam design on α with an almost lack thereof on cone-angle. In Figure (3), we demonstrate with pattern cuts that disparate values of cone-angle do yield nearly identical main beam patterns. The primary difference between the designs is the expected sidelobe level increase with increasing cone-angle. Given this information, a designer can make tradeoffs such as reduced antenna size (length) accompanying larger cone-angle with increased front-to-back ratio accompanying smaller cone-angles. Further results with higher dimensional searches will be presented.

7. Conclusions

A genetic algorithm approach to wideband rotationally symmetric log-periodic array design is presented. The approach is made computationally feasible through the introduction of symmetries, through automatic, parametric grid generation, through evaluation over a single log-frequency period and through a network-parallel implementation using freely available network queuing software. Not only did use of the genetic algorithm identify engineering solutions, but also trends in the behavior of the conical interdigitated log-periodic antenna.

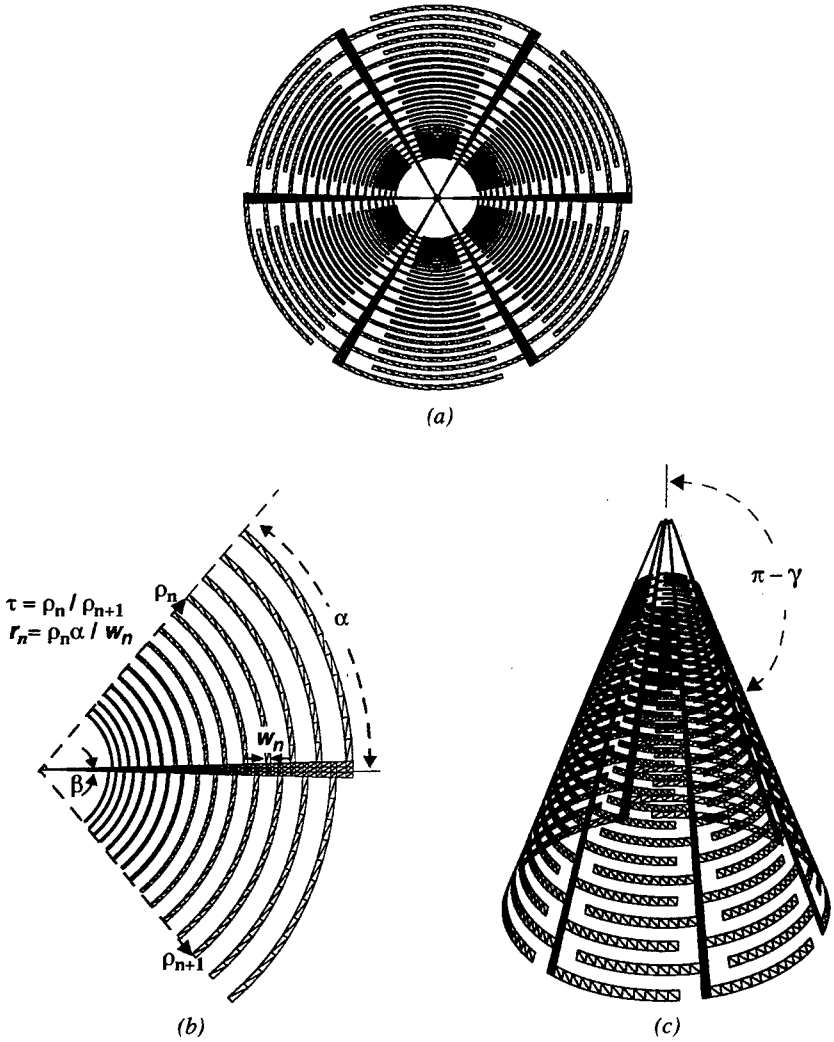


Figure 1. (a) A planar six-arm interdigitated log-periodic antenna geometry, (b) a single generating arm with log-periodic parameter definitions and (c) a conical six-arm interdigitated log-periodic antenna with half-cone angle γ.

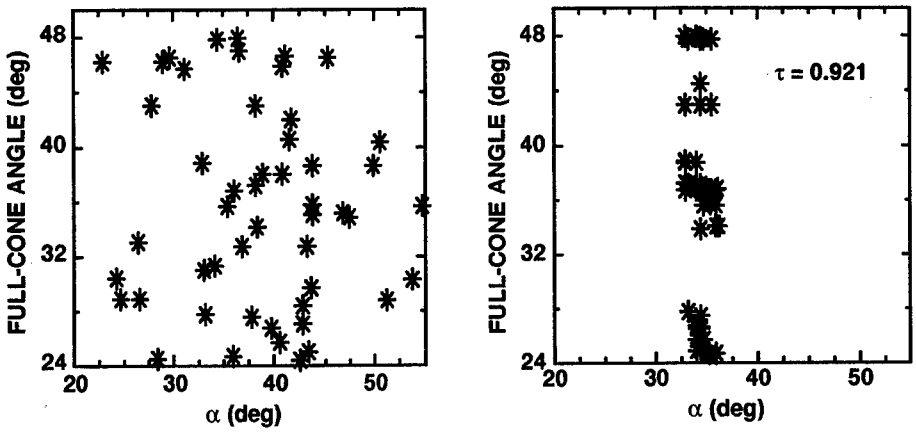


Figure 2. (a) Genetic algorithm initial population distribution and (b) genetic algorithm two-parameter search results after 9 generations. Asterisks in (b) indicate parameters which meet the designed sum-mode beamshape goals.

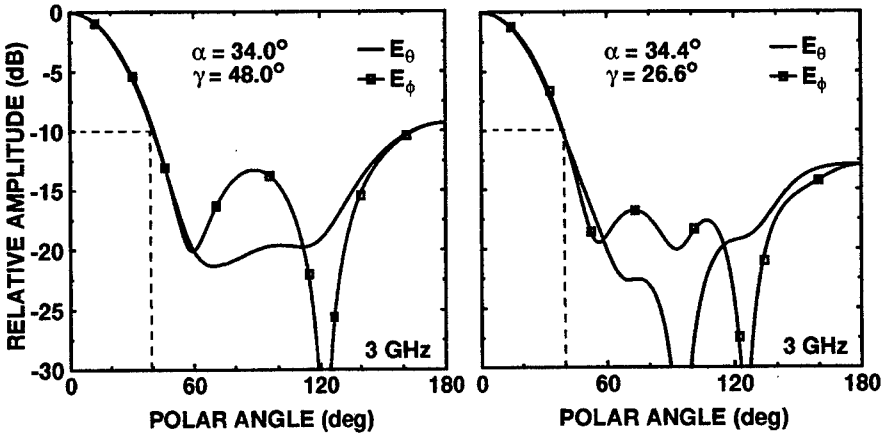


Figure 3. Sum-mode, 0°azimuth pattern cuts from two different designs resulting from genetic algorithm search above.

A COMPARISON OF SIMPLE AND COMPLEX GENETIC ALGORITHMS IN WIRE ANTENNA DESIGN

Brian S. Sandlin
Andrew J. Terzuoli
Air Force Institute of Technology
P.O. Box 3402
Dayton, OH 45401

Abstract

A simple genetic algorithm and the GENOCOP III software package are each integrated with Numerical Electromagnetics Code Version 4.1 (NEC4.1) for the purpose of determining the geometry of a wire antenna to be used as a ground antenna. After ten unique trials of each integrated routine, the resulting fitness values are compared. Also, a direct comparison is made between the antenna designs achieved in terms of power gain, azimuthal symmetry and input impedance.

1 Introduction

The impact of the earth upon the fields and power radiated by near-ground antennas has been extensively studied in both emperical and theoretical domains. Understood to a lesser extent is the impact of antenna geometry upon the power radiated at low elevation angles. No effort has been made to optimize the antenna geometry given the real-earth consideration.

A wire-antenna design is clearly desirable for the type of context associated with a remote intrusion monitoring system, but because of the problem complexity combined with the design constraints, a classical design approach is impractical. A stochastic search method, the genetic algorithm (GA), not only makes a solution attainable, it finds a solution that performs better than thought possible.

Thus, the research problem is simple: to use a GA to optimize a wire antenna geometry in the presence of a less-than-perfectly-conducting half-space for the objectives of power gain, symmetry of radiated power in azimuth, and matched input impedence. The reason for the first two objectives is obvious given the design context and the performance of existing designs. Meeting the last objective, that of matching the impedence will allow maximum power transfer, a topic reminiscent of a basic circuits course.

2 Approach

The first step in the approach to solving the research problem was to develop GAs which interface with the moment method code (MoM), Numerical Electromagnetics Code Version 4.1 (NEC4.1), to develop a wire antenna geometry. The wire endpoints become the features that the GAs search to find the optimal design. The fitness is determined by a weighted sum of multiple objectives. The next step is to compare the resulting antenna design found by the simple version of the integrated GA with one found by the integration of NEC4.1 and the more sophisticated GA software package, GENOCOP III, using a simple geometry definition. The third and final step is to develop the method by which the resulting genetically-designed antennas are to be evaluated. Not only will the gain be investigated, but symmetry and input impedance are investigated.

The first of the integrated codes is titled the *simple genetic algorithm* (SGA) because it is based upon the fundamental principles behind the genetic methodology. The second of the two codes takes advantage of a GA software package developed by Zbigniew Michalewicz called GENOCOP III, a highly sophisticated program developed over the course of seven years, and its associated integrated code will be known as the GENOCOP III-GA (GGA). Additionally, some issues in interfacing with the NEC4.1 code are investigated.

3 Results

For the runs incorporating the basic series geometry also investigated by Altshuler and Linden in [1], four wires were investigated to allow some complexity to enter the design while keeping the number of wires to a reasonable level to avoid a messy conglomeration of wires. For this simple geometry, the antennas obtained by both the SGA and GGA are directly compared. In order to make a comparison, the number of fitness evaluations was limited for each algorithm. For the SGA, a population of 50 strings was used and it was allowed to iterate for 100 generations, resulting in a total of

	Best Fitness	
Run #	SGA	GGA
1	70.1	75.5
2	72.4	73.0
3	69.8	69.8
4	70.1	78.8
5	70.0	70.2
6	70.1	73.2
7	60.1	75.5
8	70.0	74.2
9	70.1	78.7
10	70.0	70.4
Mean (μ)	69.3	73.9
Variance (σ^2)	9.9	9.6

Table 1: **Comparison of SGA and GGA Results for the Four-Wire Series-Connected Geometry**

3020 fitness evaluations. Similarly, the GGA was limited to a maximum of 3020 fitness evaluations. The results of this experiment involved 10 unique runs for both GAs. The best fitness obtained by each of the GAs for each run is shown in Table 1. The superiority of the GGA is evident by a significantly higher mean and a lower variance in the trials. This result proves that the variety of crossover and mutation operators used by GENOCOP III performs a more adequate search of the landscape.

3.1 Optimized Geometries

From the second run of the SGA, the antenna with the highest fitness is displayed in Figure 1. The antenna with the highest fitness found by the GGA came from the fourth run and has a geometry shown in Figure 2. By simply looking at the geometry, it is not perfectly clear why this antenna performs better than any other arbitrary conglomeration of wires. However, the GGA design exhibits some definite characteristics that would be expected for this application. A mostly vertical element rises from the source and is augmented by some sort of top-loading structure. The SGA design is a little more peculiar but still exhibits the height characteristic.

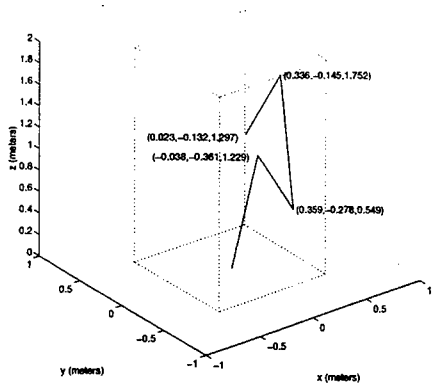


Figure 1: **Four-Wire Geometry Found by SGA**

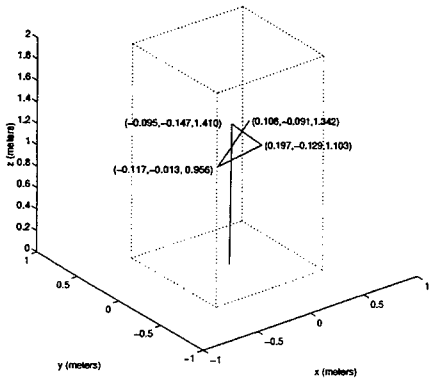


Figure 2: **Four-Wire Geometry Found by GGA**

3.2 Power Gain

Elevation cuts of the power gain are shown in Figures 3 and 4. On each plot, the gain of the $\lambda/4$ monopole is also given to provide the opportunity for a direct visual comparison with the GA designs.

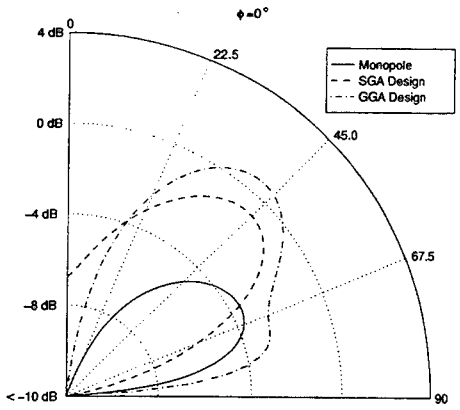


Figure 3: **Four-Wire Geometry: Elevation Cut of Power Gain at an Azimuth of $\phi = 0°$**

The power gain plots in Figures 3 and 4 are particularly disturbing from a user point-of-view because they show a major deficiency in the SGA design. For angles of $\theta > 67.5$, the monopole gain is at least 4 dB greater than the SGA design. For the other azimuth cuts given, this deficiency is not present, but it is this lack of symmetry in the SGA design that differentiates it from the GGA design by a lower fitness score.

The symmetry of the gain for the GGA design is very good considering its asymmetrical geometry. This result is attributable to its long, mostly-vertical wire which serves a monopole function. The height of the vertical wire in the GGA design is greater than that of the RIMS monopole by nearly a meter, which explains its superiority in gain at the lower elevations.

For the GGA design at most all of the azimuth positions, the GGA design offers a 1 dB improvement in power gain over the monopole at $\theta = 67.5°$. This improvement increases to approximately 4 dB at $\theta = 82°$.

At several azimuth positions, the SGA design exhibits better gain (up to 3 dB for $\theta > 67.5$), particularly at $\phi = 90°, 135°, 270°$, and $315°$. However, this does not make

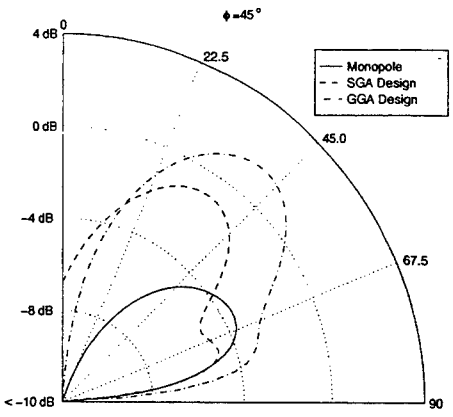


Figure 4: **Four-Wire Geometry: Elevation Cut of Power Gain at an Azimuth of $\phi = 45°$**

the SGA design a better antenna for the RIMS application because of the previously discussed lack of symmetry in its gain patterns.

3.3 Input Impedance

It is clear from the Smith Chart of Figure 5 that the four-wire antenna designed by GGA has a far superior impedance match at the center frequency than the monopole. The SGA design is less well matched than the GGA design but is still better than the monopole at the center frequency. When moving away from the center frequency, however, the reactance of both designs becomes very large.

4 Conclusions

Many in the electromagnetics community are using simple genetic algorithms for optimization. The direct comparison in this research effort hopes to show that the elementary GA might not be the most extensive search tool given the availability of sophisticated GA codes from the computer science community.

The more complex GGA algorithm proved to be more capable in both domains. In the GA domain, its variety of crossover and mutation operators made it possible for

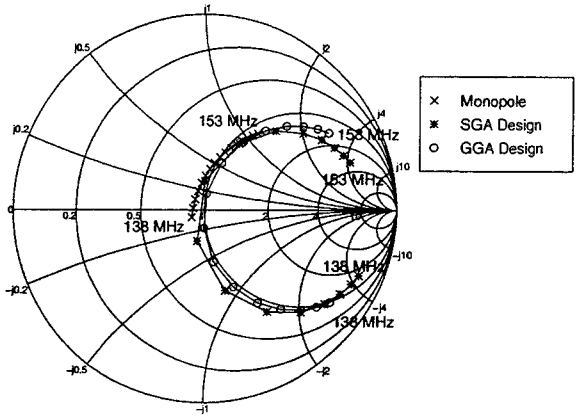


Figure 5: **Input Impedance of the GA Designs vs. a Typical RIMS Monopole**

the GGA to find a wire-geometry with a high fitness not achievable using the SGA. In the antenna domain, the resulting antenna produced a much more desirable antenna because the radiation from the SGA design was not very symmetrical.

5 Acknowledgments

The authors gratefully acknowledge the following people for their help and advise: Dave Van Veldhvizen, Robert Marmelstein, Gary Thiele, Gary Lamont, Peter Collins, Jeff Berrie, Errol English, Leo Kempel, Eric Walton, Ben Munk, Gerald Burke, and Ed Altshuler.

References

[1] Derek S. Linden and Edward E. Altshuler, "Automating Wire Antenna Design Using Genetic Algorithms", *Microwave Journal*, pp. 74–86, March 1996.

Array Failure Correction with a Genetic Algorithm

Brian Beng Kiong Yeo and Yilong Lu†
School of Electrical and Electronic Engineering, Nanyang Technological University
Singapore 639798, Republic of Singapore
† Fax: (+65) 792 0415, E-mail: eylu@ntu.edu.sg

Abstract— **Array failure correction has been achieved using a genetic algorithm (GA). One is employed in this paper to re-calculate the new attenuation ratios and phase differences among the remaining functional elements of a digitally beamformed linear array. The same method may be applied for different failure conditions. A double element failure correction can make use of the chromosome obtained from correcting the failure of a single element, if the latter is one of the two damaged elements in the former scenario, and so forth. Though the nature of GA makes real-time computational results unavailable, the normalised weightage of the remaining elements may be pre-calculated and stored in the memory of the beamforming computer.**

I. Introduction

Instead of replacing the defective array elements of a phased-array antenna, the attenuation or power level and the phase differences of the remaining elements may be re-calculated to produce an array pattern close to the original. The idea is not new. More recent applications in satellite or extra-terrestial communications, where antenna damage caused by radiation or age cannot be rectified by element replacements, have renewed the interest in this area of research.

In addition, the mutual coupling effect among the antenna elements may result in an array pattern output which is different from that desired. Digital beamforming using an array of analogue-to-digital convertors may resolve the above problem. Consequently, the motivation to study the re-distribution of elemental weightage or power, in face of mutual coupling or less than optimum performance of one or more elements. So that an array pattern that is much closer to the original specification can be produced.

Currently, there are two algorithms aimed at reducing the sidelobe levels of arrays with defective elements. The first [1] reconfigured the amplitude and phase distribution of the remaining elements by minimizing the ratio of the average peak sidelobe power level to the power in the main beam, via a conjugate gradient method. The second algorithm [2] was shown to replace the signals from failed elements in a digitally beamformed receive array. Experimental data further confirm its operation for the case of one signal and one interfering source.

On the other hand, GA is a stochastic search and optimisation technique. It searches from a population of points, not a single point. It works with the coding of the parameters and not the parameters themselves. It uses objective function information instead of derivatives or other auxiliary knowledge. In addition, it relies on probabilistic transition rules and not

deterministic rules. However, due to its slow convergence, it is not suitable for real-time applications.

Nevertheless, in [3] Haupt applied GA to determine which element should be on, in large thinned linear and planar arrays to obtain low sidelobes. Yan and Lu [4] then used a GA to restrict the phase and magnitude to certain discretized values for easy implementation by commercially available digital phase shifters and attenuators, thereby greatly reducing the complexity and cost of array antennas.

In this article, a GA based on [4] is applied to array failure corrections. Since the above is a much more difficult task than simple sidelobe reduction of a uniformly spaced linear array, considerable improvement and new additional features have to be introduced. The approach using the GA is elaborated in the following text, with the assumption that the reader has sufficient background knowledge on the relevant antenna theories.

II. The Genetic Algorithm

Natural evolution is a search for the fittest individual in species-space. The success of life on earth demonstrates the computational power of this search process.

Based after natural evolution [6]-[7], genetic algorithms (GAs) capitalize on tools that work well in nature. GAs often succeed where other algorithms fail. It is considered a sophisticated search algorithm for complex, poorly understood mathematical search spaces.

GAs mimic biological evolution to solve computational problems. Living things are encoded by chromosomes, with GAs one encodes the problems in the form of data structures. Thus, GAs are capable of arriving at an optimal solution without the benefit of explicit knowledge about the problem area.

A. Chromosome Structure

Until [8], most GAs use binary coding and binary genetic operations. The proposed approach however, applies floating-point genetic operations on complex array weighting vectors.

Hence, each chromosome is a vector which has a length equivalent to the number of array elements. It represents the normalised weighting coefficients, $\mathbf{w}_n$, as follows:

$$\mathbf{w} = \{w_1, w_2, \ldots, w_N\}, w_n \in C^n, \tag{1}$$

where C^n is the set or subset of all complex numbers and

$$AF = \sum_{n=1}^{N} w_n * e^{jknd\cos\theta} \tag{2}$$

is the array factor of a linear array.

$$size(cPop3') = P * 2 - N \tag{3}$$

The best P children out of $cPop3'$ are stored into $cPop3$.

EMP usually yields the best sample among the three methods. It is the only method that allows the fittest individual to procreate freely with the rest of the population. However, it requires nearly twice the computation time as compared to the other two methods. Eventually, four populations are available for comparison, namely the original Pop, $cPop1$ from BMW, $cPop2$ from AFP, and $cPop3$ from EMP. Subsequently, a ranking exercise sorts out the best P individuals to produce $tPop1$. Meanwhile, a multi-modal non-uniform mutation operator is applied to a side population $mPop1$, comprising of P copies of the best individual prior to the mating operation above. The Gaussian distribution shape parameter S, for the amount of mutational change, is adaptively reduced once stagnated growth is detected. To ensure intrinsic parallelism, the same mutation operator is performed across a copy of the original population, Pop, giving $mPop2$. The best P individuals from $mPop1$ and $mPop2$ are selected to produce $tPop2$. Finally, the new generation of P individuals to form Pop are those from the best of $tPop1$ and $tPop2$.

D. Fitness Evaluation

A template, formed by the shape of the main lobe and the specified sidelobe level (SLL), is cast over the array pattern, produced by each candidate, to compute their cumulative difference as a form of fitness measure, in dB. Thus, the ideal array pattern must conform to the original main beam shape with the specified SLL.

E. Termination Criteria

The maximum number of generations must be defined together with the desired fitness level. By satisfying either of the above, the GA will terminate. A log file of the GA progress in terms of the increasing fitness per generation, and the matrix containing the chromosomes of the current population are saved onto the hard disk. By reviewing the above data, a better control of the GA convergence through fine-tuning the shape S of Gaussian distribution or introducing new heuristik marriage routines can be achieved.

F. Convergence Observation

The best sample of each generation may be produced through linear crossover, after one of the selection methods, or from a mutated individual. Usually, the offsprings of fit individuals from the previous generation show greater fitness, in the beginning of a GA run. However, when approaching convergence, the mutation operation may tend to produce better individuals.

A lower shape value will result in rapid convergence in the beginning of a GA run, but ends up with premature stagnation, far from the desired fitness. Too high a S value results in a

B. Initial Population

An initial population of at least 100 random samples is generated. The weighting vector or chromosome, $\mathbf{w}_d$, of the damaged array pattern and that of a Taylor (one-parameter) synthesized array, $\mathbf{w}_t$, with an identical beamwidth are then added, to replace two of the weakest individuals among the initial population.

Their insertion helps to improve the rate of convergence. In fact, It is observed that the best individual grown for m^{th} element failure correction, should be inserted into the initial population of a double element failure, if one of the failed elements is in the m^{th} position. In so doing, the rate of convergence is observed to be increased depending on the position of the failed elements.

C. Reproduction

Ranked-based fitness assignment sorts the individuals in a descending order of fitness for a population *Pop* of P individuals. Mating is performed by three different selection methods using duplicate populations: *Pop*1, *Pop*2 and *Pop*3 respectively.

N-point linear crossover is performed [4], where $N = 2$, thus two parents produce two children.

1. Best-Mate-Worst (BMW)

Adapted from [4] and [5], BMW effectively spreads the superior genetic material in *Pop*1 to give *cPop*1. It is maximally disruptive, but weaker individuals with any desirable traits do get a chance to produce offsprings with stronger partners.

In BMW, the best gets to mate with the worst, and second best with the second worst individual. Thus, the difference in fitnesses between the best and the worst individuals is reduced. In addition, the bias for an elitist group is low.

2. Adjacent-Fitness-Pairing (AFP)

AFP marries two individuals with adjacent fitnesses. Such that the best marries the second best, the third best marries the fourth best and so forth, resulting in *cPop*2. It is highly conservative of genetic information but may result in premature convergence. However, AFP ensures the union of strong individuals whose offsprings may prove to be fitter than their parents.

In [5], a similar method known as fit-fit selection, steps through the ordered list of individuals of a population that does not remain static for an entire cycle. Unlike [5], AFP does not allow any individual to breed twice. Moreover, the population *Pop*2 that it works on stays static throughout the mating process.

3. Emperor (EMP)

The best individual in *Pop*3 gets to mate with every other sample in the population. The population of children, *cPop*3′, so generated has a size

slower convergence rate with a much steady and continuous improvement in the fitness of future generations.

III. Simulation Results

A Dolph-Tschebyscheff linear array design with a SLL of –35dB is used as a reference. The array consists of 32 identical dipoles, each spaced at half a wavelength apart.

A. Single Element Failure Correction

Fig. 1 depicts the three fitness progress curves for the different main beam directions. Notice that in the broadside case, the cumulative error after 200 generations is the highest. More importantly, convergence is observed for all the above cases at around 200 generations.

Shown in Fig. 2(a), (b) and (c) are the corrected array patterns for a 5^{th}-element failure, with the main beam pointing at broadside, 52 degrees and 138 degrees respectively. All corrected patterns have a SLL of at most –35.5dB, and their main beams retain practically the same shape and half-power beamwidth as the original.

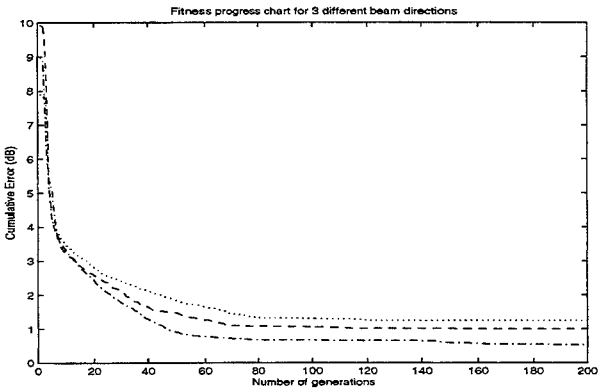
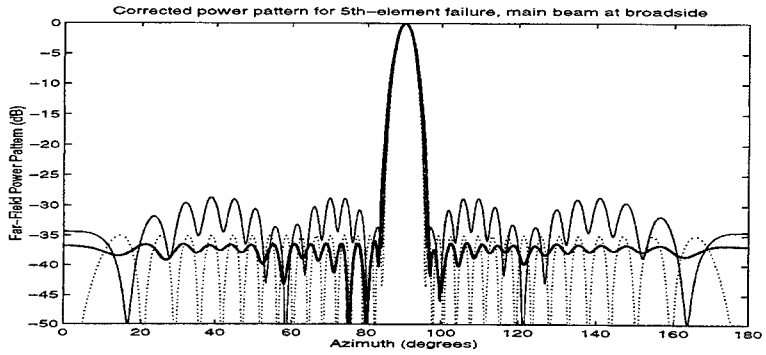


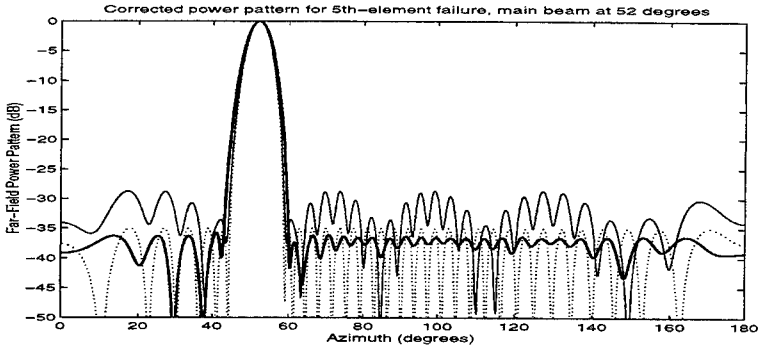
Fig. 1. Fitness progress curves, obtained from an average of 20 runs, with main beam directed at(i) broadside - dotted ,(ii) 52 degrees - dot-dashed and (iii) 138 degrees - dashed.

B. Double Element Failure Correction

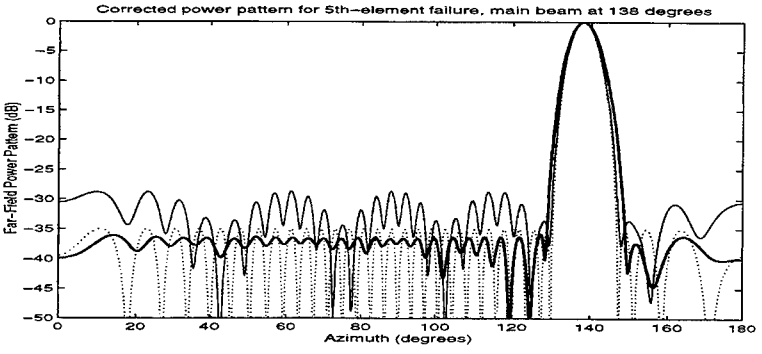
Now, if a 2^{nd}-element failure follows, the fitness progress curve, obtained from an average of 20 runs, is illustrated in Fig. 3. Similarly, convergence is observed at around 200 generations, even though two elements have failed. This is made possible by the insertion of the solution for a 5^{th}-element failure correction. Or else, more generations will be required before the GA reaches a satisfactory fitness level.



(a)



(b)



(c)

Fig. 2. (a) Corrected field pattern with main beam at broadside, (b) at 52 degrees (c) 138 degrees.

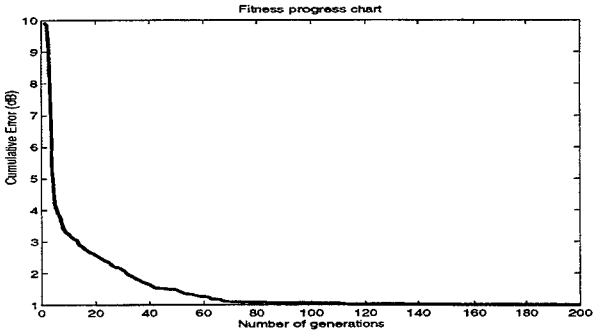


Fig. 3. Fitness progress curve with main beam directed broadside.

The corrected far-field pattern for 2^{nd}- and 5^{th}-element failure is shown in Fig. 4. The highest SLL is -35dB. Corrected patterns for other main beam directions are not shown, since they are essentially similar.

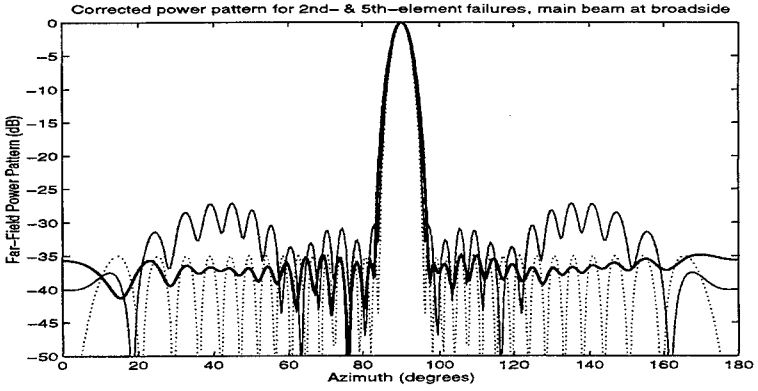


Fig. 4. Corrected beam pattern for 2^{nd}- and 5^{th}-element failure.

Usually, the number of generations required to obtain a satisfactory fitness value does increase with the number of failed elements. However, the increase in the number of generations is largely dependent on the position of the failed element(s). This applies even if the solution for a single element failure correction is planted in the initial population for a double element failure correction, and so forth.

IV. Conclusions

A genetic algorithm is proposed for the (32-element) array failure correction of single and double element failures, which translate to 3.125% and 6.25% array failure percentage respectively. For a triple element failure (or 9.372% failure), the solution for a double element

failure can be included in the initial population for the correction of the former, if two out of the three failed elements are identical to those involved in the latter, and so forth.

The genetic algorithm demonstrates the possibility of its application for non-linear array synthesis, since damaged linear arrays are essentially non-linear in nature. Though the rate of convergence may be too slow for real-time applications, the results for different combinations of element failure may be stored in the memory of a digital beamformer. It can then dynamically set the weight of each element, if an array failure scenario of a similar nature arises.

References

[1] T. J. Peters, "A Conjugate Gradient-Based Algorithm to Minimize the Sidelobe Level of Planar Arrays with Element Failures," *IEEE Trans. Antennas Propag.*, vol.39, pp.1497-1504, Oct 1991.

[2] R. J. Mailloux, "Array Failure Correction with a Digitally Beamformed Array," *IEEE Trans. Antennas Propag.*, vol.44, pp.1542-1550, Dec 1996.

[3] R. L. Haupt, "Thin arrays using Genetic Algorithms," *IEEE Trans. Antennas Propag.*, vol.42, pp.993-999, Jul 1994.

[4] K. K. Yan and Y. Lu, "Sidelobe Reduction in Array Pattern Synthesis Using Genetic Algorithm," *IEEE Trans. Antennas Propag.*, vol.45, pp.1117-1121, Jul 1997.

[5] Lance Chambers, "Practical Handbook of Genetic Algorithms : Applications," *CRC Press*, vol.1, pp45-46, 1995.

[6] J. H. Holland, "Adaptation in Natural and Artificial Systems," *Ann Arbor: The University of Michigan Press*, 1975.

[7] G. J. E. Rawlins, (Ed). *Foundations of Genetic Algorithm*, San Mateo, CA: Morgan Kaufmann Publishers, 1991, pp.211-213.

[8] L. Davis, (Ed), *Handbook of Genetic Algorithms*, New York: Van Nostrand Reinhold, 1991.

BACKSCATTERING SYNTHESIS FROM TAPERED RESISTIVE GRIDS

Randy L. Haupt
University of Nevada
Electrical Engineering Dept./260
Reno, NV 89557-0153
702-784-6927
fax: 702-784-6627
email: haupt@ee.unr.edu

Abstract

This paper describes how to synthesize a tapered resistive grid that produces a desired backscattering pattern. The grid consists of equally spaced, equal width strips. Each strip has a resistivity that is found using a genetic algorithm. Physical optics is used to calculate the backscattering. Results are compared with Method of Moments calculations.

1. INTRODUCTION.

The control of backscattering patterns using resistive surfaces has received considerable attention in the literature. Resistive tapers for strips have been synthesized to produce bistatic scattering and backscattering patterns similar to those of antenna arrays [1]. Physical optics (PO) proved useful in the resistivity synthesis, because the resistivity significantly dampened current interactions on the strip. Closely spaced grids have backscattering patterns similar to those of a continuous strip of the same size. The problem with physical optics backscattering calculations is that any perfectly conducting strips in the grid make the calculations inaccurate. Figure 1 shows the maximum return from a finite grid of 8 resistive strips spaced 0.5λ with η=2. The physical optics calculations are compared with more accurate method of moments calculations. Interactions between the strips are strong when the strips are perfectly conducting; however, the interactions are considerably smaller when the strips are resistive [2]. Consequently, PO is a viable technique for calculating the induced currents on the strips. Note that the MOM and PO results converge as the strip width approaches 0.5λ or the eight separate strips become one.

This paper introduces a method of optimizing the resistivity of the strips in the grid using a genetic algorithm. The objective is to reduce the maximum relative sidelobe level of the backscattering pattern. Genetic algorithms mimic natural selection, reproduction, and mutation to arrive at an optimum solution. The genetic algorithms are slow but can find an optimum solution to a problem with a large number of parameters. Excellent results are obtained from reasonable resistive tapers.

2. APPROACH.

Figure 2 shows a diagram of a finite grid of periodic resistive strips that lies along the x-axis. The incident electric field is parallel to the edges of the strips. A single strip with a constant resistivity has an induced current density given by [1]

$$J_z(x) = \frac{\sin\phi_o}{.5 + \eta \sin\phi_o} e^{jkx\cos\phi_o} \tag{1}$$

where

ϕ_0 = angle of incidence as measured from the x-axis
$k = 2\pi/\lambda$
λ = wavelength
η = normalized resistivity of strip

The backscattering RCS from a single strip is calculated from

$$\sigma(\phi) = \left| \int_{-w/2}^{w/2} \frac{\sin\phi}{0.5 + \eta \sin\phi} e^{jkx'\cos\phi} dx' \right|^2 \tag{2}$$

where ϕ is the angle of observation. Placing many strips side-by-side to form a grid results in a composite RCS of

$$\sigma(\phi) = \left| \sum_{m=1}^{M} \int_{x_m - w/2}^{x_m + w/2} \frac{\sin\phi}{0.5 + \eta_m \sin\phi} e^{jkx'\cos\phi} dx' \right|^2 \tag{3}$$

where

M = total number of strips
x_m = distance to center of strip m
w = width of strip
η_m = resistivity of strip m

If we assume the grid is symmetric about its center then

$$\sigma(\phi) = k \left| \sum_{m=1}^{M/2} w R_m \cos(2kx_m \cos\phi) Sa(kw\cos\phi) \right|^2 \tag{4}$$

where

$$R_m = \frac{\sin\phi}{0.5 + \eta_m \sin\phi} \tag{5}$$

$$Sa(x) = \frac{\sin x}{x} \tag{6}$$

A genetic algorithm encodes the resistivity parameters into a binary sequence that undergoes numerical evolution to arrive at the optimum resistivities that yields the lowest possible maximum relative sidelobe level.

3. RESULTS.

A grid having 40 uniform resistive strips 0.1λ wide and spaced 0.1λ apart has a maximum relative sidelobe level of approximately -13 dB. Keeping the same spacing and strip widths, but optimizing (with 5 bit accuracy for resistivity values $0 \le \eta \le 1.9375$) for the resistive taper that produces the lowest maximum relative sidelobe level results in the optimized backscattering pattern shown in Figure 3. The genetic algorithm optimization used physical optics (PO) to evaluate the sidelobe level. In this case the goal is to reduce the maximum relative backscattering. The algorithm arrived at a maximum relative sidelobe level of -27 dB. A plot of the optimum resistivity is shown in Figure 4. Figure 3 also shows the method of moments calculation of the backscattering pattern for the optimum resistivity shown in Figure 4. The method of moments plot shows a maximum relative sidelobe level of about -24.3 dB. It is interesting to note that the PO and MOM patterns have approximately the same relative characteristics. The MOM pattern has a much higher peak than the PO pattern, because PO is directly proportional to the surface area, while the MOM contributions include interactions between the grid elements.

4. CONCLUSIONS.

Resistive tapers can be approximately synthesized to control the backscattering patterns from a grid of strips using physical optics and genetic algorithms. Agreement between PO and MOM improves as the resistivity of the grid elements increase.

5. REFERENCES

[1] R. L. Haupt and V. V. Liepa, "Synthesis of tapered resistive strips," IEEE AP-S Trans., Vol. 35, No. 11, Nov 1987, pp. 1217-1225.

[2] R. L. Haupt, "Backscattering from aperiodic resistive grids using physical optics," 1995 National Radio Science Meeting Program and Abstracts, Boulder, CO, Jan 1995, p. 292.

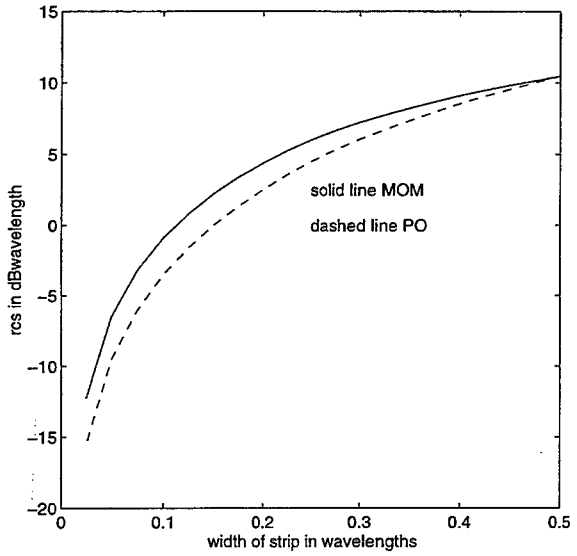


Figure 1. Comparison of the maximum RCS values for MOM an PO for a grid of 8 strips with η=2 and spacing = 0.5λ as the element widths vary.

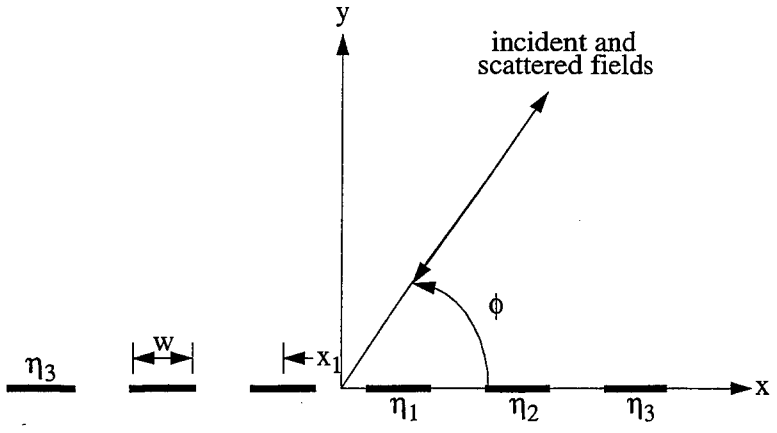


Figure 2. Model of a grid of strips.

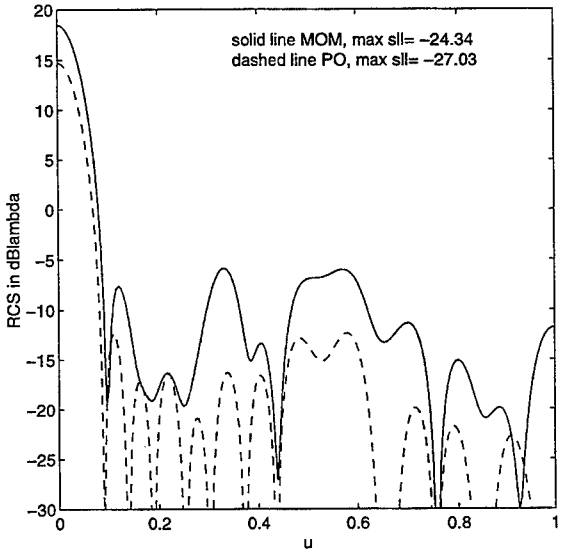


Figure 3. RCS backscattering of the grid in Figure 3.

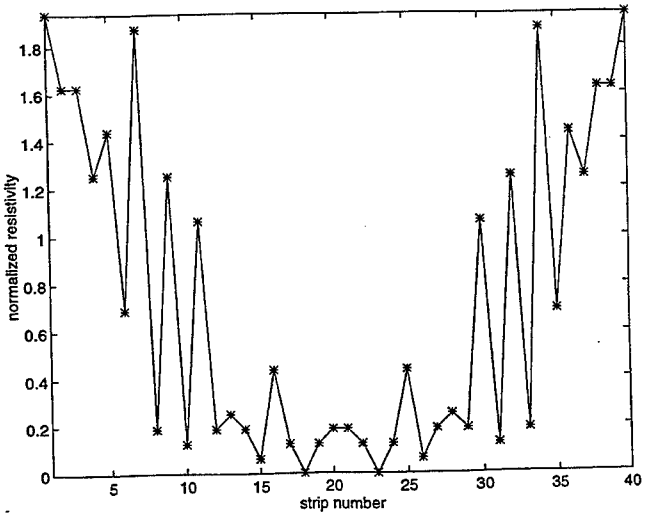


Figure 4. Optimized resistivities for a grid of strips.

Obtaining linear and circular apertures with smooth amplitude distribution and high efficiency

J. A. Rodriguez, F. Ares
Grupo de Sistemas Radiantes,
Dpto. de Física Aplicada, Facultad de Física,
Universidad de Santiago de Compostela
15706 Santiago de Compostela, SPAIN.
E-mail: faares@usc.es

Abstract

This paper presents a method of synthesising sum patterns yielding aperture distributions without edge brightening and high efficiency. The method uses the simulated annealing technique to introduce small perturbations to the root positions of linear and circular Taylor patterns. In the process, a cost function, which takes into account some design parameters such as the sidelobe level, the smoothness of the amplitude distribution, and the efficiency, is minimised.

1. Introduction

A common feature of equal sidelobe level sum patterns possessing deep nulls (linear Taylor [1] and circular Taylor [2]), is to have aperture distributions with large excitation peaks at the ends (non-monotonic). These peaks, called edge brightening, are indicative of an increase in the tolerance sensitivity [3]. Also, this rapid variation in the current (severe inverse tapering near the edge of the array with maximum efficiency) is difficult to approximate with a discrete array and may be unrealisable in a practical size [4].

Recently, a technique, based on filling the nulls of the patterns, allowed to alleviate the edge brightening in the sum patterns distributions [5]. However, this improvement requires the use of a complex aperture distribution and yields a loss in the efficiency when comparing to the conventional pattern with deep nulls.

In this work, the simulated annealing technique [6] is used to calculate small perturbations to the roots of linear and circular Taylor patterns and synthesise those associated to a high efficient amplitude distribution without edge brightening. These parameters as well as the sidelobe level of the radiation patterns are controlled by means of a given cost function, which is minimised.

2. Description of the method

a) Linear apertures:

The linear Taylor pattern [1] is given by the expression:

$$F(u)=\frac{\sin(\pi u)}{\pi u}\frac{\prod_{n=1}^{\bar{n}-1}\left[1-\frac{u^2}{u_n^2}\right]}{\prod_{n=1}^{\bar{n}-1}\left[1-\frac{u^2}{n^2}\right]} \tag{1}$$

The corresponding aperture distribution can be calculated by:

$$g(\zeta)=\frac{1}{2a}\left[F(0)+2\sum_{m=1}^{\bar{n}-1}F(m)\ \cos\frac{m\pi\zeta}{a}\right] \tag{2}$$

where $u=(2a/\lambda)\ cos\ \theta$, with $2a$ the total length, in terms of wavelength λ of the continuous linear distribution, and the aperture range $-a\le\zeta\le a$.

The method begins with the positions of the roots u_i calculated by the Taylor technique [1] and introduces small perturbations to them δu_i using the simulated annealing method [6]. These perturbations are calculated by minimising a given cost function which is defined in the following way:

$$C(\delta u_i)=c_1\cdot|I_{max}/I_{min}|+c_2\cdot V+c_3\cdot\eta+c_4\cdot f_i,\quad i=1,2,...,\bar{n}-1 \tag{3}$$

where $|I_{max}/I_{min}|$ is the dynamic range ratio, V measures the smoothness of the amplitude distribution (it allows to avoid the edge brightening), c_i are the weights of each term which depend on the design parameters, and η is the aperture efficiency defined as the ratio between the directivity peak of the obtained distribution to the directivity peak of the uniform distribution, both of the same length. Finally, f_i allows us to take each sidelobe peak under control and is defined as

$$f_i=\begin{cases}0 & if\ SLL_{i,o}\le SLL_{i,d}\\ (SLL_{i,o}-SLL_{i,d})^2 & if\ SLL_{i,o}\ge SLL_{i,d}\end{cases}\quad i=1,2,...,\bar{n}-1 \tag{4}$$

in which $SLL_{i,o}$ and $SLL_{i,d}$ are the obtained and desired peaks in dB of the *ith*-sidelobe. Note that f_i in (4) does not specify a strict level for each sidelobe peak, but a maximum allowed level. This gives to the algorithm the possibility of finding that optimal topography of the pattern that best verifies the design specifications.

b) Circular apertures:

For this case, the procedure is the same as the used for linear apertures. The circular Taylor pattern and its corresponding aperture distribution are now given by [2]:

$$F(u)=\frac{J_1(\pi u)}{\pi u}\,\frac{\prod_{n=1}^{\bar{n}-1}\left[1-\frac{u^2}{u_n^2}\right]}{\prod_{n=1}^{\bar{n}-1}\left[1-\frac{u^2}{\gamma_{1n}^2}\right]} \tag{5}$$

$$g(\rho)=\frac{2}{\pi^2}\sum_{m=1}^{\bar{n}-1}\frac{F(\gamma_{1m})}{J_0(\gamma_{1m}\pi)}\,J_0(\gamma_{1m}\pi(\rho/a)) \tag{6}$$

where $u=(2a/\lambda)\cdot sin\theta$, with $2a$ the diameter of the circular boundary of the aperture, ρ the radial coordinate at the aperture, and γ_{1m} is the root of the Bessel function J_1, defined by $J_1(\pi\gamma_{1m})=0$, with m=0,1,2,...

As before, the simulated annealing technique is used to minimise the cost function (3) introducing small perturbations to the roots of the circular Taylor patterns.

3. Results

a) Linear apertures:

Let it be desired to synthetise a pattern with a sidelobe level requirement of -20 dB. The initial roots were obtained from a linear Taylor pattern with $a=5\lambda$, SLL=-20 dB, and $\bar{n}=6$ (this value of $\bar{n}$ corresponds to a Taylor pattern of -20 dB with maximum efficiency [4]).The positions of the roots, the efficiency of the pattern as well as the dynamic range ratio of the aperture distribution $|I_{max}/I_{min}|$ are shown in the table I. In the figures 1 and 2, the pattern and its corresponding aperture distribution are shown (dashed lines).

Using our method, it was possible to obtain a pattern with a smoother amplitude distribution keeping a high efficiency. The final pattern and the aperture distribution are shown in the figures 1 and 2 respectively (solid lines), whereas Table I shows the final roots. As it can be seen, the method achieved to eliminate the edge brightening and reduce the dynamic range ratio of the aperture distribution without a significant loss in the efficiency compared to the Taylor pattern (of only about 0.2%).

TABLE I: Positions of the roots, dynamic range ratio and efficiency of the linear Taylor pattern (a=5λ, SLL=-20 dB, $\bar{n}$=6) and that obtained after the optimisation.

	Roots positions (u_i)					$\lvert I_{max}/I_{min}\rvert$	η
Linear Taylor	1.156	1.910	2.876	3.899	4.944	1.75	0.966
Final pattern	1.162	1.928	2.936	4.029	5.018	1.54	0.964

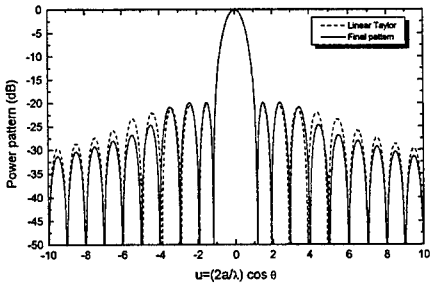


Fig. 1. Power patterns corresponding to a linear Taylor (a=5λ, SLL=-20 dB, $\bar{n}$=6) and that obtained after the optimisation.

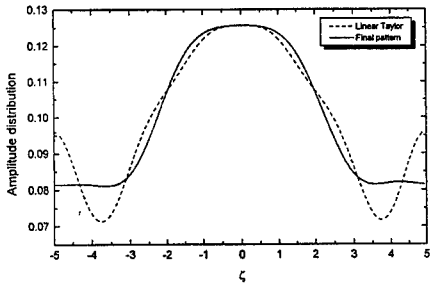


Fig. 2. Amplitude distributions corresponding to a linear Taylor pattern (a=5λ, SLL=-20 dB, $\bar{n}$=6) and that obtained after the optimisation.

b) Circular apertures:

In this case, the design sidelobe level was -25 dB. The initial roots were obtained from a circular Taylor pattern with a=5λ, SLL=-25 dB, and $\bar{n}$=5 (optimal in terms of the efficiency) [7]. Table II shows these roots, the dynamic range and the efficiency of the Taylor distribution. The power pattern and its aperture distribution are plotted in the figures 3 and 4 respectively (dashed lines).

Using our method, it was possible to obtain a pattern without edge brightening and even achieving a slight improvement in the efficiency (since the circular Taylor pattern has a SLL of about -25.8 dB and not strictly -25 dB). Table II presents the results, whereas the resulting pattern and the aperture distribution are shown in the figures 3 and 4 (solid lines).

Table II. Positions of the roots, dynamic range ratio and efficiency of the circular Taylor pattern (a=5λ, SLL=-25 dB, $\bar{n}$=5) and that obtained after the optimisation.

	Roots positions (u_i)				$\lvert I_{max}/I_{min}\rvert$	η
Circular Taylor	1.403	2.126	3.102	4.157	2.25	0.940
Final pattern	1.398	2.135	3.195	4.266	1.93	0.947

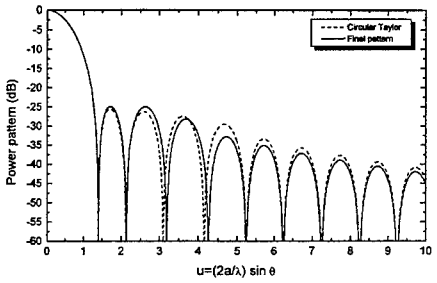


Fig. 3. Power patterns corresponding to a circular Taylor (a=5λ, SLL=-25 dB, $\bar{n}$=5) and that obtained after the optimisation.

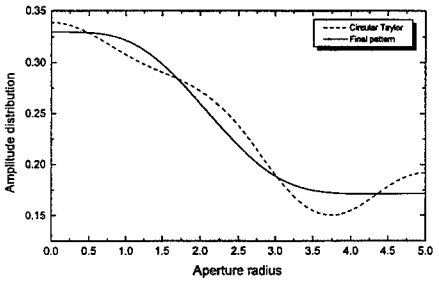


Fig. 4. Amplitude distributions corresponding to a circular Taylor pattern (a=5λ, SLL=-25 dB, $\bar{n}$=5) and that obtained after the optimisation.

4. Conclusions

A new technique of synthesising sum patterns yielding linear and circular aperture distributions without edge brightening, and high efficiency has been described. The method perturbates the roots of the linear and circular Taylor patterns through minimisation of a given cost function by means of the simulated annealing technique. The process is fast, taking about two minutes on a 200 Mhz Pentium processor for the examples shown, in which the optimal perturbations were found to be very small. This method is extensible to difference as well as to shaped beam patterns for both linear and circular distributions. The obtained apertures may be useful in reflector antenna applications because peculiar illuminations such as those with a spike near the edge may be difficult to realize in practice.

Acknowledgements

The authors wish to acknowledge financial support from Xunta de Galicia, XUGA20601A96.

References

[1] T. T. Taylor, "Design of Line Source Antennas for Narrow Beamwidth and Low Side Lobes", *Trans. I. R. E.*, 1955, vol. AP-3, pp. 16-28.

[2] T. T. Taylor, "Design of Circular Apertures for Narrow Beamwidth and Low Side Lobes", *Trans. I. R. E.*,1960, vol. AP-8, pp. 17-22.

[3] R. C. Hansen, "Array Pattern Control and Synthesis", *Proc. IEEE*, 1992, vol. 80, pp. 141-151.

[4] R. J. Mailloux, Phased Array Antenna Handbook. Norwood, MA: Artech House, 1994, pp. 124-134.

[5] F. Ares, S. R. Rengarajan, and E. Moreno, "Optimization of Aperture Distributions for sum patterns", *Electromagnetics*, 1996, vol. 16, 2, pp. 129-141.

[6] W. H. Press, W. T. Vetterling, S. A. Teukolsky, and B. P. Flannery, Numerical Recipes in C, Second Edition, Cambridge, 1992, pp. 444-455.

[7] R. S. Elliott, Antenna Theory and Design. Englewood Cliffs, N. J.: Prentice-Hall Inc, 1981, pp.261-265.

SESSION 20:

INTEGRAL EQUATION METHODS AND ERROR CONTROL

Chairs: G. Eriksson and C.Y. Shen

Force Calculations and Error Estimates with Boundary Element Methods

Kent Davey
American Electromechanics, New Smyrna Beach, FL 23168-5941

Dalian Zheng
Boulder Microwave Technologies, 4840 Meridith Way, Boulder, CO 80303

Abstract - Most boundary element techniques have been formulated directly using either the vector potential or the B and E fields. Static and time harmonic, linear and non linear problems (but not transient) can be formulated using surface currents which are placed on the skin of every ferrous or conducting region. In static problems, this surface current completely replaces the ferrous medium. Time harmonic problems require a double layer of surface current. Estimates of errors both in force and field calculations can be obtained by the secondary calculation of an intermediate set of surface currents which account for any discontinuity in the tangential component of the H field. Confirmation of this approach is obtained by comparing forces and fields with those obtained in two international TEAM workshop problems.

Index Terms - Eddy Currents, Forces, Error Prediction

INTRODUCTION

Boundary element approaches have classically employed the vector potential A or the field quantities B and E [1], [2]. These quantities are chosen in an integral equation format in such a way as to insure the continuity of normal B and tangential H. An alternative is to lace the material interfaces with surface currents, and to chose these unknown surface currents to insure the continuity conditions on B and H. This approach has the advantage that the computation of fields and forces can be performed using Biot-Savart type integrals rather than derivatives.

Boundary element methods (BEM) lack convenient quantities such as energy norms having a clear physical interpretation in predicting errors [3]. A second objective in this research in addition to outlining how this technique is used accurately for predicting these forces is to offer another technique for computing an error norm. In a previous paper by the authors, such an error norm was outlined and defined primarily in a static field arena [4]. In this paper such a technique is developed and explicitly defined in an eddy current context.

FORMULATION

In a static problem, the directive is to replace all magnetizable media with a skin of surface current as depicted in Fig. 1. This skin of surface current once determined represents entirely both the external and internal field effects properly of the original medium. In post processing, the entire field of space is represented as being occupied by air. All field calculations

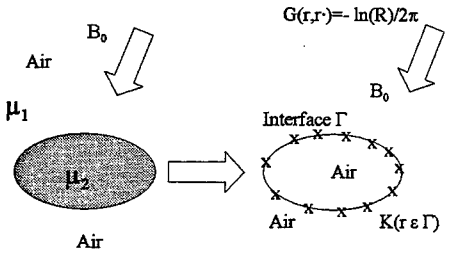


Fig. 1 A static formulation conceptual approach, replacing the magnetizable media with a skin of surface current enveloping a "sack of air".

are represented as a superposition of the field sources, the effect of the magnetizable medium being to simply add an additional set of sources which are now represented as a skin of surface current. Assume that $\vec{B}_0$ represents the effect of all impressed fields upon the problem. The magnetic field anywhere in space would then be represented as

$$\vec{B}(r) = \vec{B}_0(r) + \mu_1 \int_{s'} \nabla x \, \vec{K}_1(r') \, G_1(r,r') ds'. \tag{1}$$

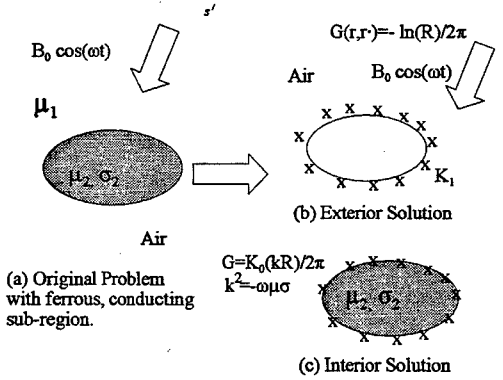


Fig. 2 Representation of an eddy current problem separating it into a laced-perimeter of surface current for the exterior solution and a separate surface current for the interior solution.

Note that the Green's function in this problem is that of the external medium, which is normally air. In a two dimensional problem the Green's function would simply be $\frac{-\ln(R)}{2\pi}$. The reader should convince himself that indeed the normal component of magnetic field density is assured to be continuous from this starting point. The same Green's function applies to all of space. The problem condenses to insuring the continuity of tangential H across the interface,

$$\hat{n} \times || \vec{H} || = \hat{n} \times (\vec{H}_2 - \vec{H}_1) = 0. \tag{2}$$

It is noted that the evaluation of the intergrand in (1) as the point R moves to the boundary is determined by noting the principle part of this integral as

$$\int_{s'} \nabla x \, \vec{K}_1(r') \, G_1(r,r') \, ds' \, |_{tangential} = \int \vec{K}_1 \frac{\partial G_1(r,r')}{\partial n} ds' \pm \frac{\vec{K}_1}{2}; \begin{cases} inside \\ outside \end{cases} \tag{3}$$

Combining (1) - (3) yields the result

$$\oint \vec{K}(r') \frac{\partial G(r,r')}{\partial n} ds' + \frac{\vec{K}}{2}\left(\frac{\mu_1 + \mu_2}{\mu_1 - \mu_2}\right) = - \frac{\hat{n} \times \vec{B}_0}{\mu_0}. \tag{4}$$

Linear basis functions are assumed for the unknown surface current K and a standard Galerkin approach used to determine the unknowns. The local error in the tangential H field is easily computed in terms of the surface currents. A typical source of the error incidentally is the failure to adequately discretize the interface; additional error accrues due to numerical solution accuracy. The discontinuity can be thought of as a new surface current, an error current defined as

$$\vec{K}_e = \hat{n} \times || \vec{H} ||. \tag{5}$$

The error in the magnetic field can now be approximated using

$$B_{error}(r) = \int_{\Gamma} \nabla x \ \vec{K}_e(r') G(r,r') ds'. \tag{6}$$

Indeed with this surface current determined, it is straight forward to predict forces, torques, and a number of other global quantities. For example, the force on a given object on which both the primary and error surface current is known becomes

$$\vec{F} = \int_{\Gamma} \vec{K} \times \vec{B} \ ds \pm \int_{\Gamma} \vec{K}_e \times \vec{B} \ dS. \tag{7}$$

EDDY CURRENTS

In a time harmonic problem, the formulation becomes slightly more complicated in that a common Green's function cannot be used throughout space. The original problem as shown in Fig. 2 must now be broken into two isolated problems. The medium having in general some permeability μ and conductivity σ must be represented as far as the exterior solution is concerned with an external equivalent skin of surface current. A secondary skin of surface current just inside the interface must also be incorporated to represent only the field interior to that region. The exterior and interior field is represented respectively as

$$\vec{B}_1 = \vec{B}_0 + \mu_1 \int \nabla x \ \vec{K}_1(r') G_1(r,r') \ ds' \tag{8}$$

$$\vec{B}_2(r) = \int_{\Gamma} \nabla \times \left[\vec{K}_2(r) \, G_2(r,r') \right] dS'. \tag{9}$$

As the point R moves to the interface, the integrands on the right-hand side of (8) and (9) become in that limit

$$\int_{r \to \Gamma} \nabla x \ \ K_1(r') \, G_1(r,r') \, ds' \, |_{tangential} = \int K_1 \frac{\partial G_1}{\partial n} \, ds + \frac{K_1}{2} \tag{10}$$

$$\int_{r \to \Gamma} \nabla x \ \ K_2(r') \, G_2(r,r') \, ds' \, |_{tangential} = \int K_2 \frac{\partial G_2}{\partial n} \, ds - \frac{K_2}{2}. \tag{11}$$

is again necessary to enforce the boundary condition in (2); this yields the condition

$$\frac{\hat{n} \times \vec{B}_0}{\mu_1} + \int K_1 \frac{\partial G_1}{\partial n} - \int K_2 \frac{\partial G_2}{\partial n} + \frac{K_1 + K_2}{2} = 0. \tag{12}$$

What remains is to enforce the continuity condition on tangential E, that $\hat{n} \times \| \vec{E} \| = 0$. Since E is related to the vector potential A as

$$\vec{E} = - \nabla \varphi - j\omega \vec{A}, \tag{13}$$

This boundary condition can now be imposed in terms of the vector potential A. Note that in two dimensions the scalar potential φ can be set equal to 0. Sometimes this is referred to as the modified vector potential solution. This inherently assumes a coulomb gauge in which $\nabla \cdot \vec{A} = 0$. The vector potentials in both regions 1 and 2 can be represented as

$$\vec{A}_1 = \vec{A}_0 + \mu_1 \int_{s'} \vec{K}_1(r') \, G_1 \, (r,r') \, ds' \tag{14}$$

$$\vec{A}_2 = \mu_2 \int_{s'} \vec{K}_2(r') \, G_2 \, (r,r') \, ds'. \tag{15}$$

Incorporating these equations into (13) yields the result

$$j\omega\left[\vec{A}_o + \mu_1 \int_\Gamma \vec{K}_1 G_1(r,r')\,ds - \mu_2 \int_\Gamma \vec{K}_2(r') G_2(r,r')ds'\right] = 0. \tag{16}$$

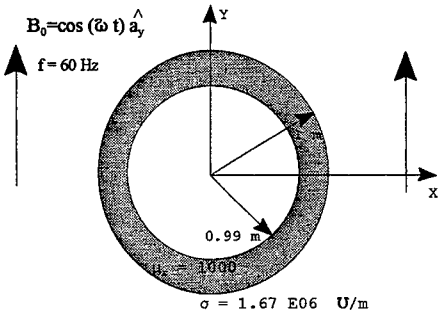


Fig. 3 Thin shell conducting sphere problem.

In three dimensions it is also necessary to impose the constraint of the gauge, i.e., that $\nabla\cdot\vec{A} = 0$. In a double-layer formulation such as this, the continuity of normal B is no longer inherently enforced. Thus errors can accumulate both in jumps in tangential H and normal B. The two types of error sources associated with these discontinuities are

$$\vec{K}_e = \hat{n}\times\|\vec{H}\|,\ \sigma_m^e = \hat{n}\cdot\|\vec{B}\|. \tag{17}$$

Here it is clear that the magnetic surface charge σ_m^e must be employed as an additional error source due to the discontinuity of normal B. The error term describing the inaccuracy of the force prediction is now composed of two contributions due to both of these sources as

$$\vec{F}_e = \int_\Gamma (\vec{K}_e\times\vec{B} + \sigma_m^e \vec{B})\,ds. \tag{18}$$

Equation (17) correctly infers that the $\vec{H}$ is equivalent to a new kind of current on the interface and the discontinuity in B is equivalent to a new kind of surface charge. However this would force a new formulation inside the code, and therefore would not be an efficient means of calculating the error quantities. In practice with an eddy current problem, it has been found that this discontinuity in normal B is very small. The primary contribution to the error comes from a discontinuity in the tangential component of $\vec{H}$. An alternative therefore to (17) would be to seek an error current just external and just internal to the conducting body's interface as in the original problem. The governing equations for these two contributions to the error current would be

$$\int \vec{K}_1^e(r')\frac{\partial G_1(r,r')}{\partial n}\,ds' - \int \vec{K}_2^e(r')\frac{\partial G_2(r,r')}{\partial n}\,ds' + \frac{\vec{K}_1^e + \vec{K}_2^e}{2} = -\hat{n}\times\|\vec{H}\|;\ r\varepsilon s \tag{19}$$

$$j\omega\left[\mu_1 \int \vec{K}_1^e\, G_1(r,r')\,ds' - \mu_2 \int \vec{K}_2^e\, G_2(r,r')\,ds'\right] = 0. \tag{20}$$

Rather than employ (18) with this modified formulation to determine the error in the force computation, it can be computed directely with the external error current K_1^e as

$$\vec{F}_e = \int_\Gamma \vec{K}_1^e \times \vec{B}_0 \, ds. \tag{21}$$

A quantification of the accuracy of the error force prediction guaranteed by (21) will now be examined.

NUMERICAL RESULTS FOR A TIME HARMONIC PROBLEM

Fig. 3 shows the first problem examined in the prediction of these error fields. A conducting shell sphere having a permeability of 1000 and a conductivity $\sigma = 1.67 \times 10^6$ Ʊ/m is stimulated by a vertical 1 T field oscillating at 60 Hz. Because the Galerkin formulation is employed, the jump in tangential H must be determined at three points over any given element, with a Gaussian integration performed over that element to solve equations (19) and (20). Because these equations are identical in format to (12) and (16), no new matrix need be defined. This translates into an immense savings if a direct solver is used, and a reasonable one if an iterative solver is employed since the set up and preconditioning need not be repeated. Because the surface currents are computed using linear basis functions there are in fact at least three options for actually formulating the surface currents to bound this error.

1. Use a linear basis function as directly implied by $\hat{n} \times \| \vec{H} \|$.
2. Use the largest value over any element unless the sign changes.
3. Use the absolute value over any element and assume the sign is always positive.

The reader might appreciate the fact that these error currents typically oscillate due to numerical inaccuracies implied both by the basis functions chosen and the numerical handling within the computer. Option 1 yields an approximation to the error, but does not bound it. Option 2, wherein the largest value over any one element is employed unless a sign change occurs, does give a much better approximation to the error and usually bounds that error. What is implied is the following: If the surface current for the left edge is -10 and that for the right edge is -15, -15 is chosen as the equivalent service current over the whole area. By contrast Option 3 for the same example would dictate a choice of +15 for the surface current over the entire element. This choice does give a better bound for the error but usually over-estimates it. Option 3 is a better indicator for the actual bounding of the error than is option 2 for time harmonic problems.

Table 1 shows the results of the error field calculations for a thin shell conducting sphere placed in a 60Hz, 1 Tesla field.

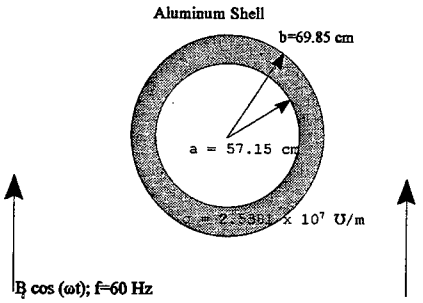


Fig. 4 Felix shell cylinder.

The errors commensurate with an option 1 assignment for the jump in tangential H over any element as well as an option 3 assignment follow in columns 5 and 6 respectively. Column 4 represents what is close to an analytic approximation to the solution, one which was obtained by using an extremely large number of unknowns. It is clear that an option 1 assignment of the jump in tangential H over any element yields too low an approximation to the error.

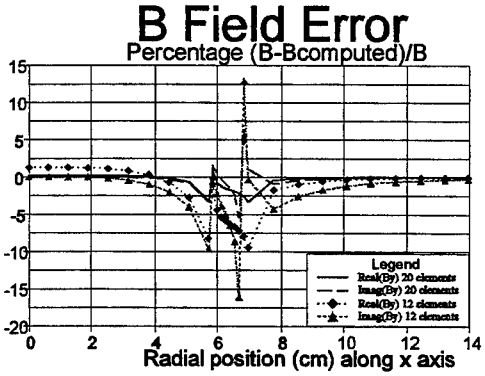


Fig. 5 Percentage field error computed along the x axis for different element densities.

Fig. 4 shows the second problem considered, the Felix shell cylinder. Fig. 5 shows a prediction in the field along some line and some error as indicated. The finer density demonstrates a much lower field prediction error as expected. The 20 element distribution predicts a force on a quarter of the cylinder of Fx = -333.34 N and Fy = -474.99 N. The analytic force predictions were Fx = -333.433 N, and Fy = -466.7166 N.

Table 1 THIN SHELL CONDUCTING SPHERE INNER RADIUS = 0.99m, OUTER RADIUS = 1m, μ_r = 1000, $\sigma = 1.67 \times 10^6$ Mho/m, f = 60 Hz IN A 1 TESLA VERTICAL FIELD.

* elements	x	calculated	(close to analytical)	error option 1	error option 3
20	0.5	0.90E-03	0.24E-02	0.00138	0.00861
20	0.10	35.78	34.91	0.0827	0.1084
20	1.1	0.8205	0.815	0.000963	0.00504
40	0.5	0.24E-02	0.24E-02	0.000153	0.003688
40	0.10	34.6	34.91	0.01081	0.04604

CONCLUSIONS

A technique for approximating not only the size of the field error, but also that of force and torque errors has been suggested. Although the technique is not rigorous, and violations to the bounding can be demonstrated, the technique does offer a reasonably close approximation to the field uncertainty at any point in space. The examples here are two dimensional; the approach has equal validity to 3 dimensions. Perhaps one of the greater benefits of the technique is that no new matrix problem need be defined and solved; the same solution matrix is used both for this primary field and secondary error field calculations. Furthermore, the technique offers the advantage that when global quantities such as force and torque are required the technique has a higher accuracy in bounding the error even as the number of elements is radically altered.

REFERENCES

[1] K. Davey and Y. Kanai, "Use of Null Field Equations in Magnetoquasistatic Problems," *IEEE Transactions on Magnetics*, vol. MAG-22, no. 4, pp. 292-298, July 1986.

[2] K. Davey, H. Han, and L. Turner, "The Calculation of Transient Eddy Current Fields Using Null Field Techniques," *IEEE Transactions on Magnetics*, vol. MAG-23, no. 2, pp. 1811-1818, March 1987.

[3] G. C. Hsiao and W.L. Wendland, "A finite element method for some integral equations of the first kind," *J. Appl. Math. Anal.*, vol. 58, pp. 449-481, 1977.

Use of Residual Error Bounds to Obtain Stable Numerical Solutions of a Fredholm Integral Equation of the First Kind

Thomas Schwengler
Dept. of Electrical & Computer Engineering
University of Colorado
Boulder, CO 80309-0425
303-541-6052
tschwen@Colorado.edu

Edward F. Kuester
Dept. of Electrical & Computer Engineering
University of Colorado
Boulder, CO 80309-0425
303-492-5173
kuester@boulder.colorado.edu

Abstract—Residual error bounds are derived for solutions of a Fredholm integral equation of the first kind. A simple two-dimensional scattering problem is used to illustrate and compute the error bounds. The problem is solved by the method of moments using the usual Galerkin method and a least square method based on the error bound. Two sets of basis functions are considered: rooftop and step functions. In the case of rooftop functions, the least square method leads to better and more stable numerical results when compared to the Galerkin method.

1. Introduction

1.1. Generalities

We consider the following linear operator equation

$$\mathcal{L}f = g \tag{1}$$

where f and g are complex functions belonging respectively to the domain D and range R of $\mathcal{L}$. The function g is known (by an exact expression or by measurements), and the function f is the unknown. $\mathcal{L}$ is a linear operator with D and R in a vector space X:

$$\begin{aligned} \mathcal{L}: \quad & D \to R \\ & v \mapsto \mathcal{L}v \end{aligned} \tag{2}$$

We also consider spaces of functions, X and their duals X', with the following duality pairing:

$$\begin{aligned} \langle \cdot | \cdot \rangle: \quad & X \times X' \to \mathbb{C} \\ & (u, v) \mapsto \langle u | v \rangle \end{aligned} \tag{3}$$

such that (whenever the expressions make sense):

$$\langle u|v\rangle = \overline{\langle v|u\rangle},\ \langle a_1u_1 + a_2u_2|v\rangle = a_1\langle u_1|v\rangle + a_2\langle u_2|v\rangle,\ \|u\|^2 = \langle u|u\rangle \geq 0 \text{ and } (u \overset{a.e.}{=} 0) \Leftrightarrow \|u\|^2 = 0 \tag{4}$$

where u, u_1, u_2 are functions of the space X, v is a function of X', and a_1 and a_2 are complex scalars*. A bar over an

* The letters *a.e.* above the equal sign denote equality almost everywhere, that is that u differs from 0 at most on a set of measure zero.

expression denotes the complex conjugate. If both u and v are in X, the duality pairing coincides with a hermitian inner product on X^2. In our concrete example below, we shall use the well-known product used to confer to L_2 [†] its usual structure of Hilbert space:

$$\langle\cdot|\cdot\rangle_{\mathbf{R},0}: \qquad X\times X'\to\mathbb{C}$$
$$(u,v)\mapsto\langle u|v\rangle_{\mathbf{R},0}=\int_{-\infty}^{\infty}u(x)\overline{v(x)}dx \tag{5}$$

And finally we need the adjoint operator $\mathcal{L}^\dagger$

$$\mathcal{L}^\dagger: \qquad R'\to D'$$
$$v\mapsto\mathcal{L}^\dagger v \tag{6}$$

such that

$$\forall(u,v)\in X\times X',\langle\mathcal{L}u|v\rangle_{\mathbf{R},0}=\langle u|\mathcal{L}^\dagger v\rangle_{\mathbf{R},0} \tag{7}$$

In our following simple example of a scattering by a strip, as in many other cases, the interest is more in the calculation of some other physical parameter, rather than in purely knowing f. Let us call F the parameter to calculate and express it as the following functional of f:

$$F=\langle f|r\rangle \tag{8}$$

Since $f\in D$, $r\in D'$, and we also seek q in R' such that

$$\mathcal{L}^\dagger q=r \tag{9}$$

This last equation constitutes the adjoint problem – to the direct problem (1). Ultimately, the purpose of the present paper is to derive an upper bound of the error made in the evaluation of F defined in (8). And, as we shall see further, obtaining an error bound on the parameter involves both problems.

1.2. Residual

In general, we approximate f and q by their projections, f^{ap} and q^{ap} to a subspace of basis and testing functions respectively:

$$f^{ap}=\sum_{i=1}^{i=N}\alpha_i\varphi_i \quad\text{and}\quad q^{ap}=\sum_{j=1}^{j=N}\beta_j\psi_j \tag{10}$$

where the complex coefficients α_i, and β_j are to be determined by some method in order to approximate the solutions f and g in some sense. We denote by e_f and e_q the respective errors of the approximation:

$$e_f=f^{ap}-f \quad\text{and}\quad e_q=q^{ap}-q \tag{11}$$

and we refer to the following functions as residuals of the operator equation and its adjoint:

$$\mathcal{L}e_f=\mathcal{L}f^{ap}-\mathcal{L}f=\mathcal{L}f^{ap}-g \quad\text{and}\quad \mathcal{L}^\dagger e_q=\mathcal{L}^\dagger q^{ap}-\mathcal{L}^\dagger q=\mathcal{L}^\dagger q^{ap}-r \tag{12}$$

These residuals are useful in the approximation of the parameter F, since its stationary (or variational) approximate form is given by:

$$F^{st}=\langle f^{ap}|r\rangle+\langle g|q^{ap}\rangle-\langle f^{ap}|\mathcal{L}^\dagger q^{ap}\rangle=F-\langle\mathcal{L}e_f|e_q\rangle \tag{13}$$

In this paper we shall consider two separate methods of computing the aforementioned α_i, and β_j coefficients for a given illustrative problem; we shall use the Petrov-Galerkin method (method of moments with identical sets of basis and testing function) and a least square method minimizing a norm of the residuals. In both cases we shall look at two different sets for the basis and testing functions, piecewise constant (step functions) and piecewise linear (rooftop functions).

[†] L_2 is the space of square Lebesgue-integrable functions over the real line.

2. Simple One-dimensional Problem

2.1. Scattering by a Strip

In order to evaluate error bounds involved with a specific operator equation, we now consider the simple problem of scattering of a time-harmonic electromagnetic plane incident wave by an infinitely long, perfectly conducting, strip (of width *2a*) in free space.

The incident wave is TM polarized, whose direction of propagation is at an angle φ_i with the strip, as shown on Figure 1 below. For the physical parameter of interest in this case, that is the expression as defined above in (8), we choose the far field of the resulting scattered wave, at an observation direction φ_s with respect to the strip, also represented in Figure 1 below.

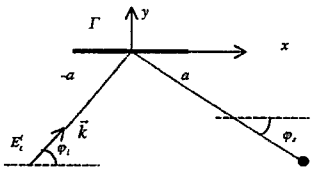


Figure 1: Simple scattering problem.

This example is the same as that considered in [1]. The incident wave (omitting the time harmonic factor *exp(jωt)*) is $\overrightarrow{E^i} = E_z^i \overrightarrow{u_z} = \overrightarrow{u_z} E_0 e^{-jk(x\cos\varphi_i + y\sin\varphi_i)}$. A surface current density is induced on the strip: $\overrightarrow{J_s} = \overrightarrow{u_z} J_{s_z}$, and a scattered field results:

$$\overrightarrow{E^s} = E_z^s \overrightarrow{u_z} = \overrightarrow{u_z}\left(-\frac{\omega\mu}{4} \int_\Omega J_{s_z}(x') H_0^{(2)}\left(k\sqrt{(x-x')^2 - y^2}\right) dx' \right) \tag{14}$$

where Ω denotes the strip (which occupies the interval [*-a, a*] in the *xy* plane) and $H_0^{(2)}$ is the Hankel function of the second kind of order zero. By enforcing the boundary condition on the strip (total *E*-field $E_z^i + E_z^s$ vanishes), we obtain the integral equation:

$$\int_\Omega f(x') H_0^{(2)}\left(k|x-x'|\right) dx' = g(x) \tag{15}$$

where

$$g(x) = e^{-jkx\cos\varphi_i} \quad \text{and} \quad f(x) = \begin{cases} \dfrac{\omega\mu}{4E_0} J_{s_z}(x) & \text{for } x \in \Omega \\ 0 & \text{for } x \notin \Omega \end{cases} \tag{16}$$

Equation (15) is in the form of (1) with the following integral operator:

$$\mathcal{L}h = \int_{-\infty}^{\infty} h(x') H_0^{(2)}\left(k|x-x'|\right) dx' \tag{17}$$

The adjoint operator with respect to (5) is then:

$$\mathcal{L}^\dagger h = \int_{-\infty}^{\infty} h(x') H_0^{(1)}\left(k|x-x'|\right) dx' \tag{18}$$

where $H_0^{(1)}$ is the Hankel function of the first kind, and is equal to the complex conjugate of $H_0^{(2)}$ for real arguments.

The far-field expression for the scattered field is:

$$E_z^s = -E_0 \langle f | r \rangle \sqrt{\frac{2}{\pi k \rho}} e^{-jk\rho + j\pi/4} \tag{19}$$

where r is the function:

$$r(x) = e^{-jkx\cos\varphi_s} \tag{20}$$

This completes the necessary settings for our problem.

2.2. Sobolev Spaces

It has been shown (e.g. in [2], and also in [3], [4] & [1]) that Sobolev spaces of fractional order *(n+½)* are the appropriate setting for studying the functions involved. The reader is referred to these publications, as well as [5], [6] & [7] for a detailed treatment of Sobolev spaces. We shall make use in the following of the Sobolev space of order ½ and its dual, of order -½. In particular, [1] shows that the error between F and F^{st} is bound by the following:

$$\left|F^{st} - F\right| = \left|\left\langle \mathcal{L}e_f \,\middle|\, e_q \right\rangle\right| \le \left\|\chi \mathcal{L}e_f\right\|_{\mathbb{R},½} \left\|\chi \mathcal{L}^\dagger e_q\right\|_{\mathbb{R},½} \tag{21}$$

where χ is an extension operator extending functions defined on Ω to the entire real line. The extension must be done in a smooth enough manner that the resulting functions belong to $H^{½}(\mathbb{R})$, the Sobolev space of order ½. A norm and inner product of $H^{½}(\mathbb{R})$ are given by (22) and (23) below. They vary from the usual norm given to Sobolev spaces by the factors k & $1/2\pi$; this difference allows us to make the expression consistent with its physical dimensions; it also leads to a tighter bound. These norms, however, are equivalent to the usual ones (obtained by replacing the factors by unity), and their treatment is similar to that found in the aforementioned literature.

Now, one further step is useful; the functions we deal with are mostly in Sobolev spaces $H^s(\Omega)$, rather than in $H^s(\mathbb{R})$. And, typically, the functions that we deal with that belong to $H^s(\mathbb{R})$ are extensions of functions defined on Ω. In particular, the functions of our error bound are $H^{½}(\mathbb{R})$ extensions of $\mathcal{L}e_f$ and $\mathcal{L}^\dagger e_q$, and both functions are only known on the domain Ω. The intuitive extension operator would be to truncate the function outside of Ω, unfortunately, this produces a discontinuity and the resulting function no longer belongs to $H^{½}(\mathbb{R})$. Another straightforward operator is to transition linearly to zero, over an interval T, as detailed by (24) below and shown on Figure 2.

The latter extension was used in [1] and led to good results. It is however a little cumbersome to evaluate, and leads to further difficulty in a multi-dimensional case. In principle, the best approach would be to minimize $\left\|\chi \mathcal{L}e_f\right\|_{\mathbb{R},½}$ and $\left\|\chi \mathcal{L}^\dagger e_q\right\|_{\mathbb{R},½}$, over all possible extension operators χ, thus obtaining a norm – as impractical as it may be to compute – on $H^{½}(\Omega)$. We may denote that norm: $\inf_\chi \left\|\chi u\right\|_{\mathbb{R},½}$. We now wish to find a norm that involves only integration over Ω, like $\inf_\chi \left\|\chi u\right\|_{\mathbb{R},½}$, but that is practically computable; the extension defined in (25), and represented in Figure 3, will lead us to such a norm.

The inner product and norm in $H^{½}(\mathbb{R})$ are referred to with the subscript $\mathbb{R},½$:

$$\langle u | v \rangle_{\mathbb{R},½} = k \langle u | v \rangle_{\mathbb{R},0} + \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\left(u(x) - u(x_0)\right)\overline{\left(v(x) - v(x_0)\right)}}{\left|x - x_0\right|^2} \cdot dx \cdot dx_0 \tag{22}$$

$$\left\|u\right\|_{\mathbb{R},½}^2 = k \left\|u\right\|_{\mathbb{R},0}^2 + \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\left|u(x) - u(x_0)\right|^2}{\left|x - x_0\right|^2} \cdot dx \cdot dx_0 \tag{23}$$

$$\chi_{lin}u = \begin{cases} u(x) & \forall x \in [-a,a] \\ \dfrac{u(a)}{T}(a+T-x) & \forall x \in]a,T+a[\\ \dfrac{u(-a)}{T}(a+T+x) & \forall x \in]-T-a,a[\\ 0 & \text{elsewhere} \end{cases} \quad (24)$$

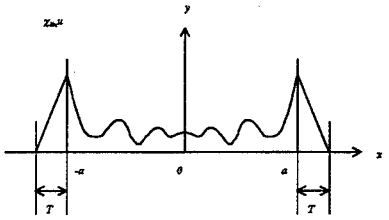


Figure 2: $\chi_{lin}u$, linear tapering on a function u defined on $\Omega=[-a,a]$.

$$\chi_{sym}u = \begin{cases} u(x) & \forall x \in [-a,a] \\ \dfrac{(2a-x)}{a}u(2a-x) & \forall x \in]a,2a[\\ \dfrac{(2a+x)}{a}u(-2a-x) & \forall x \in]-2a,-a[\\ 0 & \text{elsewhere} \end{cases} \quad (25)$$

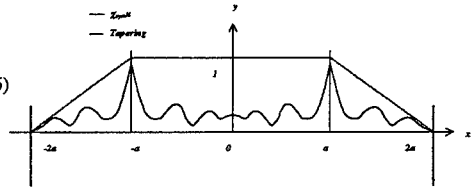


Figure 3: $\chi_{sym}u$, symmetrical extension with linear tapering on a function u defined on $\Omega=[-a,a]$.

Estimations of the extended portions of the norm lead to the bound [8]:

$$\|u\|^2_{\mathbb{R},\gamma_2} \le \|u\|^2_{\Omega,\gamma_2} = \left(2k + \frac{0.521}{a}\right)\|u\|^2_{\Omega,0} + \frac{2}{\pi}\int_{\Omega}\int_{\Omega}\frac{|u(x)-u(x_0)|^2}{|x-x_0|^2}\cdot dx \cdot dx_0 \quad (26)$$

This last norm is certainly more convenient, since it does not require any type of extension outside of Ω, however, it is still cumbersome to compute. The trouble comes from the fact that we need to evaluate a double integral and that the integrand has some sort of singularity $\left(1/|x-x_0|^2\right)$. A remedy is to try to bound the norms above by H^1 norms. In particular, [1] showed that when the function u has an integrable derivative:

$$\|u\|^2_{\mathbb{R},\gamma_2} \le \|u\|^2_{\mathbb{R},1^{bis}} = k\|u\|^2_{\mathbb{R},0} + 2\left\|\frac{du}{dx}\right\|_{\mathbb{R},0}\|u\|_{\mathbb{R},0} \quad (27)$$

Note that we denote that norm with the subscript $\mathbb{R},1^{bis}$, because that norm uses the L_2 norm of the function and of its derivative, thus resembling a norm in the Sobolev space of order 1. Still the product of the two does not appear in the usual norm of $H^1(\mathbb{R})$, which is the reason for the use of the notation *bis*. Similar bounding techniques to those which led to (26) give us [8]:

$$\|\chi u\|^2_{\mathbb{R},1^{bis}} \le \|u\|^2_{\Omega,1} = \left(2k + \frac{2\sqrt{2}}{a}\right)\|u\|^2_{\Omega,0} + \left(2\sqrt{2}a\right)\left\|\frac{du}{dx}\right\|^2_{\Omega,0} \quad (28)$$

which is a norm on the Sobolev space $H^1(\Omega)$. The latter norm also has the precious advantage of having a straightforward associated scalar product (unlike $\|u\|_{R,1^{bis}}$).

We now have an assortment of norms, $\|u\|_{\mathbb{R},\gamma_2}$, $\|u\|_{\Omega,\gamma_2}$, $\|u\|_{\mathbb{R},1^{bis}}$, $\|u\|_{\Omega,1}$ that are increasingly more convenient to compute, unfortunately, they also lead to decreasingly tight bounds – we trade numerical ease of calculation for tightness of the bound.

3. COMPUTATION OF THE BOUND

3.1. Petrov-Galerkin method

We now consider more particularly the Petrov-Galerkin method to find an approximation to our problem.

The well-known method leads to the following equation[‡]:

$$[[\langle \varphi_i | \mathcal{L}\varphi_j \rangle_{\mathbf{R},0}]]_{N\times N} \cdot [\overline{\alpha_j}]_N = [\langle \varphi_j | g \rangle_{\mathbf{R},0}]_N \tag{29}$$

Which is resolved by matrix inversion.

We take in particular the following two examples of basis functions: piecewise constant functions and piecewise linear functions as represented below, so that the coefficients of each matrix can be evaluated analytically[§].

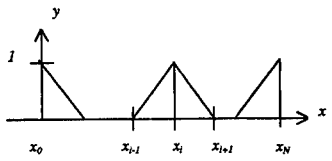


Figure 4: Piecewise linear basis functions (rooftop functions).

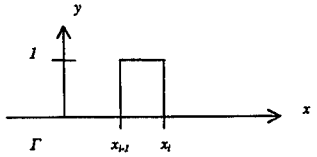


Figure 5: Piecewise constant basis functions (step functions).

The advantage of step functions is double, simplicity of calculation in the integration of the Hankel function, and convenience of the Toeplitz matrix in the evaluation of the inner products. Nevertheless, we still consider the rooftop functions for two reasons. Firstly, it seems that, intuitively, they might lead to a better (slightly smoother) approximation of the surface current on the strip. Secondly, and most importantly, certain more general problems will require continuous functions (such as rooftops) if there are any transverse currents in the x-direction, like for instance for an incident wave in TE mode. We should also recall here that, in the Galerkin method, the three terms used to evaluate parameter F, in the right hand-side of (13) are equal, therefore it is sufficient to evaluate only one of these terms.

Results are plotted as a function of the number of substrips, N, in the following Figures, for step functions, and for rooftop functions. Three of the four norms defined above are plotted (for ka=0.1).

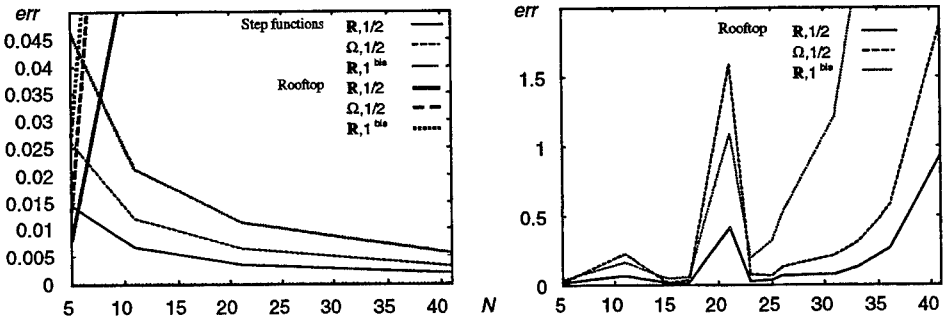


Figure 6 & 7: Error bound *err* for Galerkin method with step functions and rooftop.

[‡] We use the notation $[[m_{i,j}]]_{N\times N}$ for an N by N matrix of elements $m_{i,j}, i \in \{1,..,N\}, j \in \{1,..,N\}$ and $[v_i]_N$ for an N-dimensional vector of elements $v_i, i \in \{1,..,N\}$ – that is a 1 by N matrix.

[§] Step functions lead to a double integration of the Hankel functions, and rooftop functions lead to a quadruple integration (after integration by part). We express these in terms of Hankel functions and of their first primitive which make use of Struve functions. Formulae for computation of all these functions are found in [9] & [10]. Matrix inversions are computed by Toeplitz inversion or by LUD decomposition, both found in [10].

We cannot help but notice the poor result of rooftop functions compared to step functions – our intuitive feeling expressed above was wrong. Several papers that make use of the method of moments briefly allude to this puzzling result, without further comments. A few papers, however, such as [3] and [11], [12] comment on the phenomenon, even though only [12] attempts an explanation. In [13], Druchinin also attempts an explanation, arguing that the longitudinal direction of the current needs one more degree of smoothness than the transversal in order to obtain the best numerical results. Another interesting fact is that [3] also shows that when the singularities are taken into account at the edges of the strip, then rooftop lead to better results than step functions. All in all, it is difficult to find a thoroughly convincing explanation for that phenomenon, and it seems that we simply have to accept it.

We also note that rooftop calculations seem more unstable than those for step functions, perhaps due to the aforementioned phenomenon, or perhaps due to the added complexity of the numerical calculations.

In conclusion, although the Petrov-Galerkin method leads to acceptable results, it can exhibit instabilities in the case of rooftop functions, and we will require a better solution for the TE case. This is the motivation to explore the other method of this paper, the least square (LS) minimization of the error bound.

3.2. LS method

We now consider another method of calculation of the α_i and β_j coefficients, by writing a variational equation on the error bound. We should emphasize here that even though results of the error bounds are expected to be better (since, by definition, we determine the coefficients to minimize the bound), the actual error may not be. Minimizing the bound of the error in one of the norms used above sometimes causes an increase in the bound computed with another norm and may have a similar effect on the actual error.

We now choose a norm from among the ones we detailed earlier [8], say for the purpose of the following derivation $\mathbb{R},½$, and we write the variational principle on $\left\|\chi\mathcal{L}e_f\right\|_{\mathbb{R},½}$ and $\left\|\chi\mathcal{L}^\dagger e_q\right\|_{\mathbb{R},½}$ as follows.

$$\forall p\in\{1,..,N\},\ \frac{\partial\left\|\chi\mathcal{L}e_f\right\|^2_{\mathbb{R},½}}{\partial x_p}=0\ \text{ and }\ \frac{\partial\left\|\chi\mathcal{L}e_f\right\|^2_{\mathbb{R},½}}{\partial y_p}=0 \tag{30}$$

This leads to $2N$ real equations, or N complex equations for the $\alpha_i=x_i+jy_i$ coefficients, and similar conditions on $\left\|\chi\mathcal{L}^\dagger e_q\right\|_{\mathbb{R},½}$ will give us the β_j coefficients. Finally, in a similar manner to earlier, we can express these conditions by matrix equations:

$$[[\langle\chi\mathcal{L}\varphi_i\,|\,\chi\mathcal{L}\varphi_j\rangle_{\mathbb{R},½}]]_{N\times N}\cdot[\overline{\alpha_j}]_N=[\langle\chi\mathcal{L}\varphi_i\,|\,\chi g\rangle_{\mathbb{R},½}]_N \tag{31}$$

$$[[\langle\chi\mathcal{L}^\dagger\varphi_i\,|\,\chi\mathcal{L}^\dagger\varphi_j\rangle_{\mathbb{R},½}]]_{N\times N}\cdot[\overline{\beta_j}]_N=[\langle\chi\mathcal{L}^\dagger\varphi_i\,|\,\chi r\rangle_{\mathbb{R},½}]_N \tag{32}$$

Similarly to Galerkin method, the linear system is solved[**] and the resulting coefficients allow us to calculate the parameter and the error bound. The same equations can be used with our $\Omega,½$ duality pairing or with $\Omega,1$. Results are plotted on Figure 8, for rooftop functions; comparison to their corresponding Galerkin case is immediate, the LS method leads to better and stable results. Interestingly, [14] favors the Galerkin method for stability, but that may be because the spaces considered for that LS method are not $H^{\pm½}$, nonetheless, in the case of rooftop functions, our computations clearly show the superiority of the LS method.

4. Conclusion

By example of a simple scattering problem on a strip, we evaluate the error bound on the far-field pattern of a scattered plane wave. The error bound is expressed in four different norms, in the Sobolev spaces of order one half or one. Firstly the Petrov-Galerkin method is used with step functions or with rooftop functions as basis functions. The latter turns out to

[**] Again, step functions lead to a double integration of the Hankel functions, and rooftop functions lead to a quadruple integration. Matrices are Toeplitz for step functions, and Hermitian in both cases.

be numerically unstable and leads to poor results of the error bound. Secondly, a least square method is used on one of the norms derived. Step functions lead to similar (yet still slightly better) results than earlier, rooftop functions, however, lead to an impressive improvement, in stability and in the value of the bound, leading to our best results.

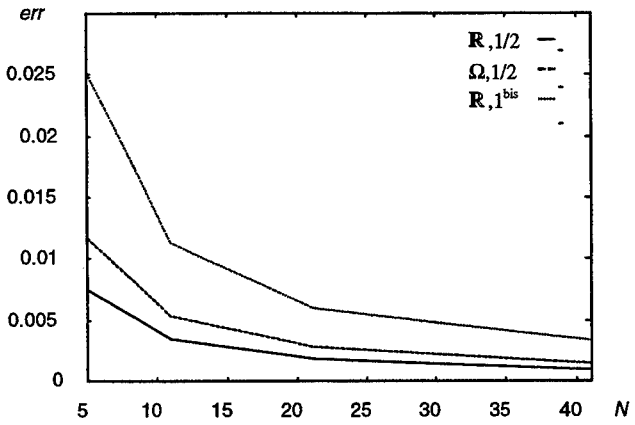


Figure 8: Error bound *err* for LS method with rooftop

REFERENCES

[1] E. F. Kuester, "Computable Error Bounds for Variational Functionals of solutions of a Convolution Integral Equation of the First Kind", *Wave Motion* **22**, pp 171-185, 1995.

[2] J. L. Lions & E. Magenes, *Non-Homogeneous Boundary Value Problem, vol. 1*, (Berlin: Springer Verlag, 1972), ch. 1.

[3] G. C. Hsiao, R. E. Kleinman, R.-X. Li, & P. M. van den Berg, "Residual Error - A Simple and Sufficient Estimate of Actual Error in Solutions of Boundary Value Integral Equations", *Computational Engineering with Boundary Elements*, (S. Grilli & al, Eds. Southamton UK: Computational Mechanics Publications), **Vol. 1**, pp. 73-83, 1990.

[4] P. M. van den Berg, A. P. M. Zwamborn, G. C. Hsiao & R. E. Kleinman, "Iterative Solutions of First Kind Integral Equations", *Direct and Inverse Boundary Value Problems*, (R. Kleinman, R. Kress, & E. Martensen, Eds. Frankfurt: Peter Lang), pp. 213-232, 1991.

[5] R. A. Adams, *Sobolev Spaces*, (N. Y.: Academic Press, 1975), ch. 7.

[6] G. I. Eskin, *Boundary Value Problem for Elliptic Pseudodifferential Operator* (Providence, RI: American Mathematical Society, 1981), ch. 1.

[7] J. T. Oden & J. N. Reddy, *An Introduction to the Mathematical Theory of Finite Elements*, (N. Y.: Wiley, 1972), ch. 4.

[8] T. Schwengler & E. F. Kuester, "Residual Error Bounds in Various Norms for Solutions of a Fredholm Integral Equation of the First Kind", submitted to *SIAM Fourth International Conference on Mathematical and Numerical Aspects of Wave propagation*, June 1-5, 1998, Colorado School of Mines, Golden, CO.

[9] M. Abramowitz & I. Stegun, *Handbook of Mathematical Functions* (Washington, D. C.: National Bureau of Standards, Dept. of Commerce, 1964). Reprinted by Dover Publ., N. Y., 1965, chaps. 9 & 10.

[10] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes in C, The Art of Scientific Computing*, New York: Cambridge University Press.

[11] D. R. Wilton, S. Govind, "Incorporation of Edge Conditions in Moment Method Solutions", *IEEE Transactions on Antennas and Propagation*, **vol. AP-25**, pp. 845-850, 1977.

[12] A. W. Glisson, D. R. Wilton, "Simple and Efficient Numerical Methods for Problems of Electromagnetic Radiation and Scattering from Surfaces", *IEEE Transactions on Antennas and Propagation*, **vol. AP-28**, pp. 583-603, September 1980.

[13] S. V. A. Druchinin, "A Criterion for Distortion-Free Solutions to the Electric-Field Integral Equation when it is Solved by the Method of Moments", *Journal of Communications Technology and Electronics*, **40**, pp. 94-104,1995.

[14] G. R. Richter, "Numerical Solutions of Integral Equations of the First Kind with Nonsmooth Kernels", *SIAM Journal Numerical Analysis*, **vol. 15**, pp. 511-521, June 1978.

Eigenvalue Studies of Matrices Resulting from EFIE Simulations for Planar Structures

John M. Dunn and Hugh MacMillan
University of Colorado
Department of Electrical and Computer Engineering
Campus Box 425
Boulder, CO, 80309–0425

phone: (303)–492–5926
fax: (303)-492-5323
e-mail: John.Dunn@Colorado.EDU

Abstract:

In this paper, we examine the eigenvalues and eigenvectors for the matrices resulting from simulation of planar structures using the electric field integral equation (EFIE). A number of interesting features peculiar to this equation are observed. The eigenvalues fall into two distinct groups at low enough frequencies. These two groups of eigenvalues are associated with the zero curl and divergence spaces of the basis functions, as has been observed by other researchers. We give a number of examples showing this separation, and discuss the poor conditioning that results. We also observe how the low frequency solution is dominated by a few modes. Finally, we discuss why sparsification of the matrix sometimes works, even though the condition number can be quite poor.

Motivation for the Study

The moment method based on the electric field integral equation (EFIE) is on the most popular methods in use today for the numerical solution of planar electromagnetic problems. In fact, it is the numerical basis for most of the commercial electromagnetic moment method products today. Despite the method's popularity and demonstrated usefulness, there are a number of well known difficulties inherent to its use. These include poorly conditioned matrices resulting from the discretization of the equation, difficulties in maintaining accuracy of solution at low frequency, and relatively poor performance of iterative techniques in solving the matrix equation.

Conventional wisdom for these problems has been that the matrices are poorly conditioned because the EFIE is an integral equation of the first kind. Such equations have notoriously poor condition numbers. Therefore a poor condition number is to be expected. Furthermore, any iterative technique converges at a rate proportional to the size of the condition number. Therefore, systems with large condition numbers, as is the case here, should be expected to perform poorly when iterative methods are used. However, this does not explain the poor performance at low frequencies. Naively, we would expect the condition number of the problem to get worse as the frequency goes up, not down. Therefore, we decided to examine the structure of the integral equation and the resulting matrices more closely.

We also wished to better understand why some researchers have been able to make the matrices sparser by converting them to different bases, for example wavelets, and then thresholding the matrix to make it sparse. This results in a system can be solved much faster, as there are

fewer elements in the matrix. The problem with this procedure is that the wavelet matrix is an orthogonal basis transformation which leaves the eigenvalue, eigenvector structure of the matrix intact; the condition number of the matrix is not changed. Normally, we are not allowed to discard values from a poorly conditioned matrix, even though some of the values are much smaller than the largest ones.

Therefore, we decided to examine a number of canonical microstrip structures to see if we could understand the issues mentioned above. In so doing, we have followed the closely the work of Vecchi et al. [1] In their paper, they demonstrate how the integral equation's basis can be subdivided into two subspaces, and how this can be used to advantage in preconditioning the system. Essentially, this paper presents a number of simple canonical structures and shows how they fit into the ideas presented in Vecchi.

Results

We examined a number of canonical structures: through sections of microstrip transmission lines, microstrip bends, and similar simple discontinuities. For example, a microstrip line is shown in figure 1. The line has been divided into a number of cells. The current is approximated by the well known rooftop basis functions. Each one of these rooftops has one unknown. Essentially, each interior edge can be associated with a rooftop. In addition, the two ports have unknowns, corresponding to half rooftops. The matrices for the resulting structures were then examined. For example, the magnitude of a matrix corresponding to 123 unknowns is shown in figure 2. (This corresponds to a strip discretization of 3 cells across and 24 cells long.) Notice that the matrix appears to be sparse. The diagonal elements, the largest elements in the matrix have maximum values of 1. The smallest elements in the matrix have values of 10^{-6}. This sparsity occurs because of the rapid falloff of the values in the matrix.

Unfortunately the condition number is quite large for these matrices. The matrix shown in figure 2 has a condition number of $3x10^5$. It is therefore not possible to discard small elements. This is shown in figure 4, where we have set small matrix elements to zero. We see that the solution for the current rapidly becomes corrupted.

The magnitude of the eigenvalues is shown in figure 3. The condition number is the ratio of the largest to smallest eigenvalue. Notice that the eigenvalues fall into two groups. At first this is a surprising development. Normally, the eigenvalues would decay toward zero at a uniform rate. The separation of the eigenvalues into two groups is explained in the work by Vecchi et al. [1] The current can be broken into two groups, a zero curl group and a zero divergence group. The most common realization of these two groups is in a "loop" and "star" separation. The zero divergence space gives very small eigenvalues, in that the eigenvectors correspond to magnetostatic solutions. Such solutions can exist in principle for any conductor. The zero divergence eigenvalues are the second group of eigenvalues in figure 4. A second thing to notice is that the first group of eigenvalues, the zero curl eigenvalues have some small entries too. These values correspond to the dominant electrostatic modes. We will give further examples of this grouping into two sets of modes.

References

[1] G. Vecchi, L. Matekovits, P. Pirinoli, M. Orefice, "A numerical regularization of the EFIE for three-dimensional planar structures in layered media", Int. Journal of Microwave and Millimeter Wave CAE, November, 1997.

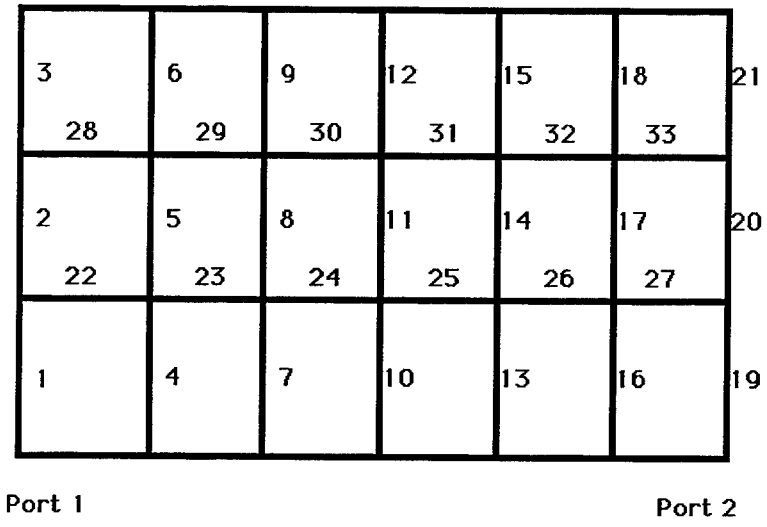


Figure 1: A two port circuit is discretized. There are 33 unknowns. Each interior edge corresponds to a rooftop current with unknown amplitude.

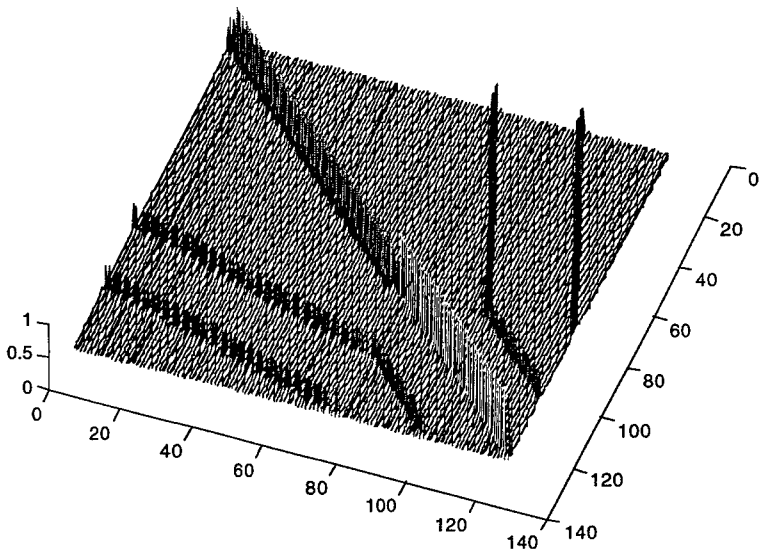


Figure 2: Absolute value of a matrix elements resulting from a structure similar to that shown in figure 1, except with 133 unknowns. The matrix looks sparse, but the condition number is very poor.

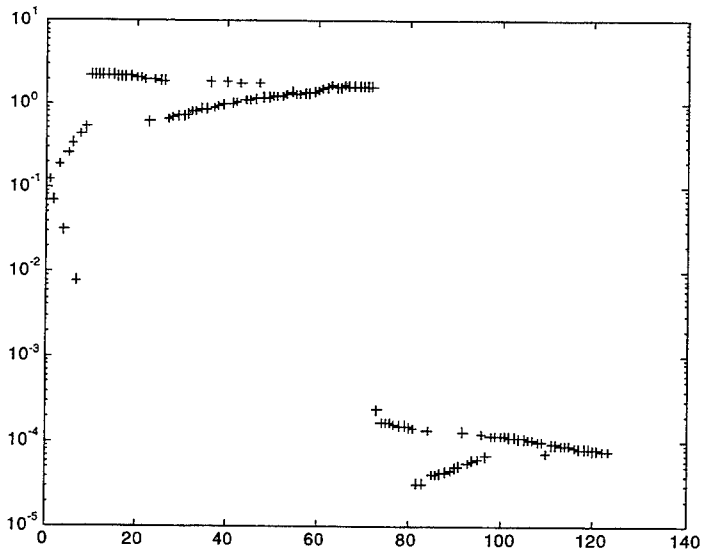


Figure 3: Magnitude of the eigenvectors of the previous matrix. The eigenvalues are clearly broken up into two groups. These groups can be associated with curl zero and div zero subspaces of the eigenvectors.

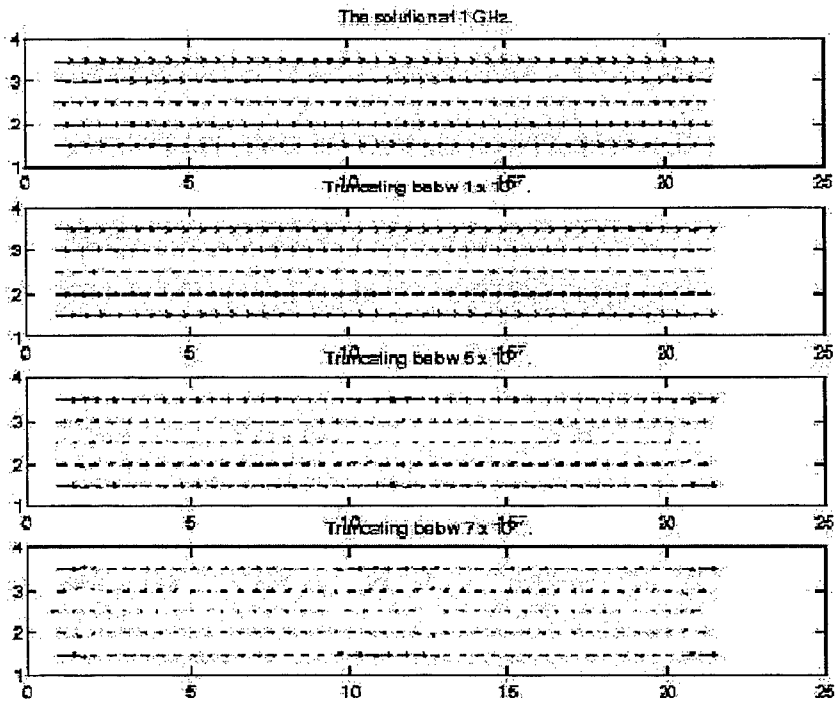


Figure 4: The matrix is made more sparse, and the corresponding current solution plotted. The currents are shown for the complete solution in the upper picture. The cutoff level in the matrix is 1x10^(-7) in the second picture; 5x10^(-7) in the third picture; and 7x10^(-7) in the fourth picture. Clear degradation in the solution can be seen.

ITERATIVE SOLUTION OF DENSE LINEAR SYSTEMS IN ELECTROMAGNETIC SCATTERING CALCULATIONS

Jussi Rahola
E-mail: Jussi.Rahola@cerfacs.fr

CERFACS, 42 av G. Coriolis, 31057 Toulouse Cedex, France

and

Center for Scientific Computing, P.O. Box 405, FIN-02101 Espoo, Finland

Abstract

Dense systems of linear equations can be efficiently solved with modern iterative methods. In many cases, the linear system needs to be preconditioned to guarantee fast convergence. This article presents two case studies, a volume and a surface integral equation for electromagnetic scattering. Convergence of the iterative methods is analyzed together with the eigenvalues of the coefficient matrices. For the surface integral equation, a sparse approximate inverse preconditioner is implemented. Finally, the possibilities of computing the matrix-vector product with fast multipole techniques is discussed.

1 Introduction

In computational electromagnetics, approaches based on differential equations have been in many cases preferred over the use of integral equations because the latter lead to very expensive calculations involving the solution of dense systems of linear equations. However, recent developments in iterative methods and in fast multipole techniques have much increased the range of applicability of integral equations in electromagnetics research and industry.

The computational time required by the direct methods for solving dense linear systems (with, *e.g.*, LAPACK) increases rapidly with the size of the linear system. Moreover, the memory required to store a dense matrix quickly becomes a bottleneck, restricting the size of the system to maybe 30000 unknowns for present supercomputers.

Iterative solvers can reduce both the computational complexity and memory requirements of the solution of dense linear systems. With a good preconditioner, the iterative solver can converge with much fewer iterations than the number of unknowns. If this is the case, the iterative solution becomes superior to the direct one.

Iterative Krylov-subspace methods only access the matrix by a series of matrix-vector products. Thus, the coefficient matrix need not be formed explicitly, and only its effect on a given vector needs to be computed. This offers some interesting possibilities for further reduction of the solution time. With modern methods for computing the matrix-vector product, such as the fast multipole method or the FFT, dense linear systems of millions of unknowns can be solved efficiently. The memory requirements of such schemes only grow linearly with the size of the system.

This article gives an example of a volume and a surface integral equation for electromagnetic scattering. The use of various iterative solvers for these equations is then discussed. Especially the relation of the eigenvalues of the coefficient matrix and the convergence of iterative solvers is considered. Finally, some experiences on the use of the fast Fourier transform and the fast multipole method are reported.

2 Integral equations of electromagnetic scattering

In this section we will describe a volume integral equation formalism and a surface integral equation formalism for electromagnetic scattering.

The volume integral equation is of course more costly than a surface integral equation, because of the greater number of unknowns. However, it allows a simple description of the scatterer in terms of cubic computational cells, and it offers the possibility of computing scattering by inhomogeneous and anisotropic scatterers. The volume integral equation of electromagnetic scattering is given by [11]

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_{\text{inc}}(\mathbf{r}) + k^3 \int_V (m(\mathbf{r}')^2 - 1)\mathbf{G}(\mathbf{r},\mathbf{r}') \cdot \mathbf{E}(\mathbf{r}')\, d^3\mathbf{r}', \tag{1}$$

where $\mathbf{E}(\mathbf{r})$ is the electric field inside the particle, $\mathbf{E}_{\text{inc}}(\mathbf{r})$ is the incident field, k is the wave number, m is the complex refractive index, $\mathbf{G}$ is the dyadic Green's function

$$\mathbf{G}(\mathbf{r},\mathbf{r}') = \left(\mathbf{1} + \frac{\nabla\nabla}{k^2}\right) g(|\mathbf{r} - \mathbf{r}'|) \tag{2}$$

and

$$g(r) = \frac{e^{ikr}}{4\pi kr}. \tag{3}$$

In the above equations, we assume that the electric field has harmonic time-dependence $\exp -i\omega t$.

The scattering integral equation can be discretized in various ways. The simplest discretization uses cubic cells and assumes that the electric field is constant inside each cube. In our calculations we require that the integral equation (1) is satisfied at the centers $\mathbf{r}_i$ of the N cubes (point-matching or collocation technique) and we use simple one-point integration.

The simple discretization of the volume integral equation naturally leads to some errors in the solutions. These errors were analyzed by Hoekstra and Rahola [9], who showed that the magnitude of the error is greatest near the surface of the scatterer.

We have also studied the solution of dense linear systems for surface integral equations. We have used a new formulation of Bendali *et al.* [2] who use a boundary flux finite-element method. Their scheme is well-posed at all frequencies and it also avoids the use of a free parameter that is used in the combined-field integral equation. The formulation uses the impedance boundary condition. In the present article we only give examples for a perfect conductor.

3 Iterative solvers

Both the volume and surface integral equations lead to a dense system of linear equation with complex symmetric (*i.e.* non-Hermitian) coefficient matrices. Thus they cannot be solved using the conjugate gradient method. Instead, non-Hermitian Krylov subspace methods are used.

We have tested various iterative solvers for the volume integral equation and the complex symmetric version of the quasi-minimal residual method (QMR) [5] turned out to be the best [16, 13, 15]. It uses only one matrix-vector product per iteration. We found out that if the same particle is discretized with increasing resolution, the number of iterations remains constant. Note that if multiple right-hand sides are available simultaneously, a block version of QMR [3] can further speed up the convergence.

The convergence of iterative solvers is determined by the eigenvalues of the coefficient matrix. In Figure 1 we show the eigenvalues of the coefficient matrix for the volume integral equation in the case of a dielectric sphere. Note that almost all the eigenvalues cluster on the line with only a few outliers. We can explain and analytically calculate the eigenvalues off the line and have a proposition for the eigenvalues on the line [15].

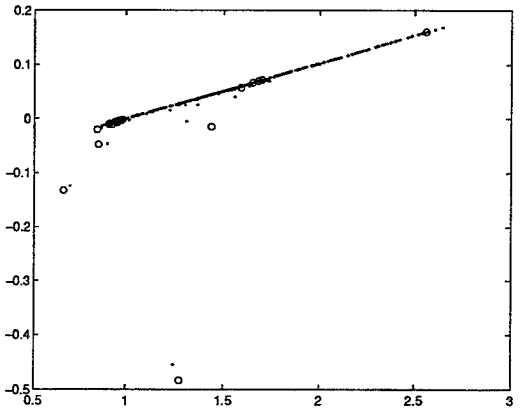


Figure 1: Eigenvalues (small black dots) of the coefficient matrix for the volume integral equation. The circles represent analytical eigenvalues of the volume integral operator.

For the volume integral equation, the QMR method converges nicely even without preconditioning. However, near some resonance frequencies the convergence can slow down [9]. For the surface integral equation, the situation is rather different and there is clearly need a good preconditioner.

We propose the use of a sparse approximate inverse (SPAI) preconditioner. The SPAI concept in a factorized from was given in [10]. However, the factorized SPAI only works for positive definite matrices, which is not the case here. The version of SPAI adopted in the present article is related to the one described in [1].

The aim of preconditioning is to transform the linear system $Ax = b$ into a preconditioned system

$$M_2AM_1y = M_2b,\ x = M_1y, \tag{4}$$

so that the new system is easier to solve with iterative methods. In many iterative solvers, the left and right parts M_2 and M_1 of the preconditioner are only accessed through the multiplication by $M := M_1M_2$. The eigenvalues of the preconditioned matrix M_2AM_1 are the same as those of AM. Ideally, the preconditioner M should be close to the inverse of the matrix A and should not have too many nonzero elements, so that the products of M with a vector can be computed cheaply.

Because we use a complex symmetric version of QMR, we are also looking for a symmetric M. We are using a version of preconditioned QMR that does not require M to have a symmetric factorization (i.e. $M_1^T = M_2$ is not required). In the sequel we only consider preconditioners with $M_2 = I$ and thus $M_1 = M$.

Let a symmetric sparsity pattern for the right preconditioning matrix M be given. Now we would like to determine the values of M by solving the problem

$$\min_M \|I - AM\|_F^2, \tag{5}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Note that this leads to independent least-squares problems for each column of M, *i.e.*

$$\min_{m_j} \|e_j - Am_j\|_F^2,\ j = 1, \ldots, N, \tag{6}$$

where m_j and e_j are the jth columns of M and I, respectively.

As the coefficient matrix is full, the solution of the above least-squares problems would require the computation of full columns or the matrix, which is a rather heavy task for really large matrices. Instead, for each column of M, we only pick the rows of A corresponding to a new sparsity set that is somewhat larger than the sparsity set for M.

In the surface integral equation employed, each unknown corresponds to a midpoint in the triangular mesh of the surface. The sparsity set for each row of M is prescribed by including only the elements corresponding to the kth neighbors of a given node. In the least-squares problems, we use the matrix rows corresponding to the $(k+1)$th neighbors of the node.

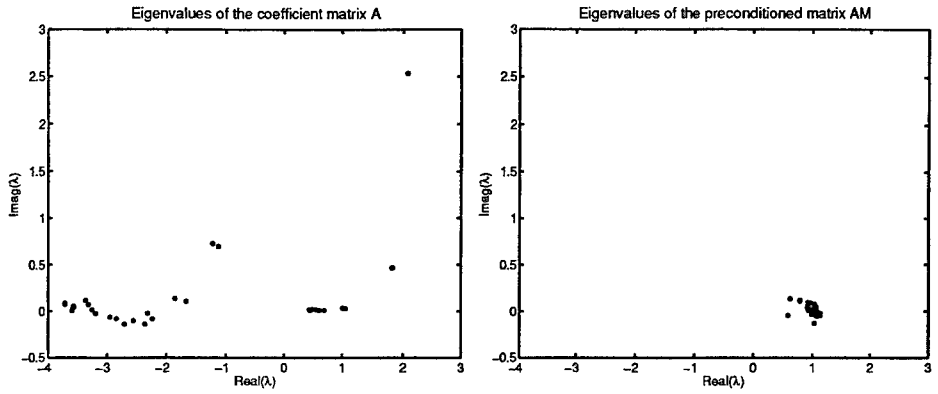


Figure 2: The left figure shows the eigenvalues of the coefficient matrix for the surface integral equation. The right figure shows the eigenvalues of the preconditioned matrix.

In the left-hand part of Figure 2 we show the eigenvalues of the coefficient matrix for the surface integral equation in the case of a perfectly conducting sphere. Note that the eigenvalues are located on both sides of the origin, which can cause problems for some iterative solvers. The right-hand part of Figure 2 show the eigenvalues of the preconditioned matrix AM. Figure 3 gives an example on how the convergence of QMR depends on the neighbor level k for the SPAI preconditioner.

In practical calculations we set k to three, which means that a row typically has about 25 nonzeros and we need to solve a least-squares problem of size 40×25 to determine the values for the nonzeros. Because M has to be symmetric, we construct the final preconditioner as $M' = (M + M^T)/2$.

In Table 1 we report the CPU times that are needed to build the preconditioner and to solve the linear system with direct and iterative solvers together with the number of iterations for the iterative method QMR. It can be seen that the number of iterations for QMR grows rather slowly when bigger scattering problems are solved. The iterative method quickly becomes faster than the direct method.

4 Matrix-vector product

When iterative Krylov-subspace methods are applied to dense linear systems, by far the most expensive operation is the matrix-vector product. Both the time needed to compute the matrix-vector product and the memory to store the matrix grow as $\mathcal{O}(N^2)$ if the matrix is formed explicitly.

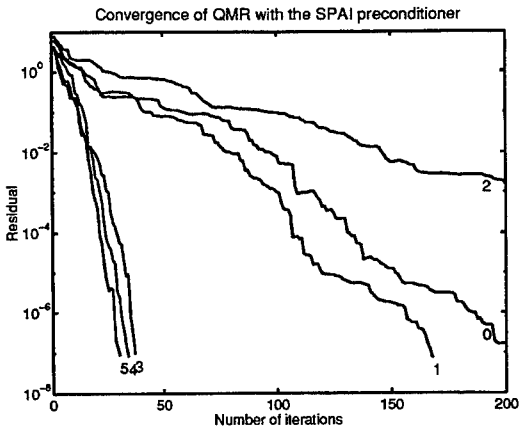


Figure 3: Convergence of QMR with the SPAI preconditioner. The numbers below the curves indicate the level of preconditioner used, with 0 denoting unpreconditioned QMR.

N	n_{iter}	T_{SPAI}	T_{QMR}	T_{direct}
1080	24	5.7	0.5	2
1920	31	10	1.7	11
3000	38	16	4	42
6750	61	36	31	477
9720	73	52	71	1410

Table 1: Number of iterations n_{iter} for QMR and CPU times (in seconds) on a Cray C90 for construction of the SPAI preconditioner (T_{SPAI}), the iterative solution (T_{QMR}) and for the direct solve (T_{direct}). N denotes the size of the linear system. QMR was stopped when the 2-norm of the initial residual was reduced by a factor of 10^{-8}.

For the volume integral equation the matrix-vector product can in some cases be computed by a 3D FFT [6, 12]. If the computational cells sit on a regular lattice, the lattice can be enlarged to a cube and the matrix-vector product with a cubic lattice reduces to a 3D-convolution, which can be computed efficiently with a 3D fast Fourier transform (FFT). The computational complexity of the FFT depends on the number of lattice points in the enlarged lattice. The FFT has also been used in volume integral calculations to compute the scattering by more than seven million computational points [8]

For the surface integral equation method the FFT is not practical because the surface elements do not sit on a regular lattice.

Another method to compute the matrix-vector product is the fast multipole method (FMM) by Greengard and Rokhlin [7]. The FMM is based on truncated potential expansions which are

used to combine the field of many far-away computational cubes to a single set of expansion coefficients. The use of the FMM in scattering calculations is described, e.g., in [4, 14]. The FMM has recently been applied to surface integral equations involving over 2 million unknowns [17].

The author has analyzed the errors in the fast multipole method when the so-called diagonal translation operators are used. This formulation must be used with care if small distances and large-order terms of the series expansion are used [14]. In this case the accuracy of the FMM can be destroyed by the finite-precision arithmetic of the computers.

5 Conclusion

This article has presented two integral equation formulations of electromagnetic scattering where the dense system of linear equations has been solved with iterative solvers. For the volume integral equation case, the iterative solver converges quickly even without preconditioning. The linear system arising from a surface integral equation is preconditioned by a sparse approximate inverse preconditioner, which is relatively simple to apply and nicely reduces the number of iterations.

The sparse approximate inverse preconditioner is constructed by solving a least-squares problem for each column of the preconditioning matrix. After this, the application of the preconditioner is relatively cheap, because it involves the multiplication of a vector by a matrix with roughly 25 nonzeros in each row.

The author has used the FFT to compute matrix-vector products by a matrix of almost half a million rows and also has some experience on the use of the FMM in scattering problems. Future research will consider the application of the FMM to large-scale surface-integral equation problems. Also, the work on sparse approximate inverse preconditioners will be continued.

Acknowledgments: I would like to thank Abderrahmane Bendali and M'Barek Fares for providing me with the surface integral code and Luc Giraud for discussions on the sparse approximate inverse preconditioners.

References

[1] G. Alléon, M. Benzi, and L. Giraud. Sparse approximate inverse preconditioning for dense linear systems arising in computational electromagnetics. CERFACS TR/PA/97/05, CERFACS, 1997. To appear in Numerical Algorithms.

[2] A. Bendali, M.G. Fares, and J. Gay. Finite element solution to impedance boundary integral equation in electromagnetic scattering. CERFACS TR/EMC/97/35, CERFACS, 1997. Submitted to IEEE Trans. Antennas Propagat.

[3] W. E. Boyse and A. A. Seidl. A block QMR method for computing multiple simultaneous solutions to complex symmetric systems. *SIAM J. Sci. Comput.* **17**, 263–274, 1996.

[4] R. Coifman, V. Rokhlin, and S. Wandzura. The fast multipole method for the wave equation: a pedestrian prescription. *IEEE Antennas and Propagation Magazine* **35**, No. 3, 7–12, June 1993.

[5] R. W. Freund. Conjugate gradient-type methods for linear systems with complex symmetric coefficient matrices. *SIAM J. Sci. Stat. Comput.* **13**, 425–448, 1992.

[6] J. J. Goodman, B. T. Draine, and P. J. Flatau. Application of fast-Fourier-transform techniques to the discrete-dipole approximation. *Optics Letters* **16**, 1198–1200, 1991.

[7] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comp. Phys.* **73**, 325–348, 1987.

[8] A.G. Hoekstra, M. D. Grimminck, and P.M.A. Sloot. A parallel fast discrete dipole approximation for simulations of elastic light scattering from micron-sized particles. *Int. J. Mod. Phys. C.* (submitted).

[9] A.G. Hoekstra and J. Rahola. Accuracy of internal fields in VIEF simulations of light scattering. *Appl. Opt.* (submitted).

[10] L. Yu. Kolotilina and A. Yu. Yeremin. Factorized sparse approximate inverse preconditionings I. Theory. *SIAM J. Matrix Anal.* **14**, 45–58, 1993.

[11] A. Lakhtakia and G. W. Mulholland. On two numerical techniques for light scattering by dielectric agglomerated structures. *J. Res. Natl. Inst. Stand. Technol.* **98**, No. 6, 699–716, 1993.

[12] K. Lumme and J. Rahola. Light scattering by porous dust particles in the discrete-dipole approximation. *Astrophys. J.* **425**, 653–667, 1994.

[13] J. Rahola. Solution of dense systems of linear equations in electromagnetic scattering calculations. Licenciate's thesis, Helsinki University of Technology, 1994.

[14] J. Rahola. Diagonal forms of the translation operators in the fast multipole algorithm for scattering problems. *BIT* **36**, 333–358, 1996.

[15] J. Rahola. *Efficient solution of dense systems of linear equations in electromagnetic scattering calculations.* PhD thesis, Helsinki University of Technology, 1996. CSC Research Reports R06/96, Center for Scientific Computing, 1996.

[16] J. Rahola. Solution of dense systems of linear equations in the discrete-dipole approximation. *SIAM J. Sci. Comput.* **17**, 78–89, 1996.

[17] J.M. Song, C.C. Lu, W.C. Chew, and S.W. Lee. Fast Illinois solver code (FISC) solves problems of unprecedented size at the Center for Computational Electromagnetics, University of Illinois. Research Report No. CCEM-23-97, University of Illinois, 1997. Submitted to IEEE Antennas and Propagation Magazine.

EMCP2 A PARALLEL BOUNDARY ELEMENT SOFTWARE PACKAGE USING A NOVEL PARAMETERISATION TECHNIQUE.

Göran Eriksson and Ulf Thibblin

Ericsson Saab Avionics AB
Electromagnetic Technology Division
SE-581 88 Linköping
Sweden

ABSTRACT

The development of a new software package based on a frequency domain boundary element solver and novel parameterization techniques is presented. The parameterization technique is introduced to overcome the need to assemble and invert the resulting impedance matrix for each parameter configuration. The technique in essence makes it possible to obtain the solution as a function of parameters like frequency and geometry with only one matrix inversion and matrix-matrix and matrix vector multiplication involving high order derivatives of the matrix. Comparison with measurements are presented. The software package is developed and validated within the European ESPRIT-HPCN project EMCP2 (Electro-Magnetic Compatibility using Parallel Parameterization) where the partners are: Aerospatiale CCR (F)), Alenia Aerospazio (I), CADOE (F), Centro Ricerche FIAT (I), COREP (I), Ericsson Saab Avionics AB (S), Eurocopter (F), KTH/PSCI (S) and MIP (F).

1. INTRODUCTION

Computational Electromagnetics (CEM) has during the last years become a tool that is used more and more to solve complex problems in areas like EMC, antennas and radar cross section to mention a few examples. This is mainly due to the rapidly increasing performance of computers at a lower cost. Moreover this development has also inspired the development of new methods [1]. In particular there is a great need for simulation tools that can be used at an early stage of a project for virtual prototyping. At that stage of a project no objects exists. Therefore measurements, except on smaller parts, are out of the question. In particular so since perhaps several hundreds of different prototypes have to be studied with hundreds of different parameter configurations. However to carry out such extensive parameter studies efficiently we need efficient computational tools.

The methods used either solves Maxwell's equations in time or frequency domain. Both approaches have their advantages and disadvantages. The methods can be approximate like asymptotic high frequency methods [2] or uses a numerical method starting from an exact formulation of Maxwell's equations. Most of the high frequency methods executes quite fast and can be further improved by straightforward parallelisation so therefore we limit ourselves in this paper to direct numerical solutions by discretization

of space and time if it is a time domain method. The direct methods can either be boundary element methods (BEM) where only the boundaries of the object is discretized or volume element methods where a truncated space surrounding the object is discretized [3,4]. Although more and more time domain codes emerges that either is based on unstructured boundary element meshes or unstructured volume meshes [1] the classical FDTD method [5] is still one of the most used. Hybrid methods on hybrid meshes where for example FDTD is hybridized with a time domain FEM solver on an unstructured mesh are also being studied [1]. However many of these methods are still on a research level and the computational burden is large. The advantages with FDTD is its robustness, the many sub cell models that have been developed, the high performance that can be achieved on parallel computers and the fact that it is very easy to use. The main disadvantage is the resulting stairstepped approximation of the geometry and the difficulties with local mesh refinement. This leads to small possibilities to accurately model for example the substructure on an aircraft and curved surfaces. Another disadvantage is that the computation has to be repeated for each excitation.

While a frequency BE or FE method models the true geometry very accurately, in comparison with FDTD, the advantage with FDTD is that a pulse can be used as the excitation. Thereby we can obtain the frequency response in a frequency band by Fourier transform. The FE and BE methods requires a matrix inversion, so if the response is desired for many frequencies the inversion has to be repeated many times. This can make parameter studies where the response over a frequency band is required very time consuming with BEM or FEM in the frequency domain. In particular so since changes in other parameters like geometry and impedance, also requires a new inversion of the matrix. It is therefore very desirable to find a method by which the solution in a parameter interval can be obtained with only one matrix inversion to be able to carry out comprehensive parameter studies and optimization.

In this paper we present the development of a new software package based on a BEM solver provided by Aerospatiale CCR and novel parameterization techniques provided by CADOE that overcomes the disadvantage with having to invert a matrix many times to obtain the response in a parameter interval.

The software is being developed and validated by a concortium in the European ESPRIT-HPCN project EMCP2 (ElectroMagnetic Compatibility using Parallel Parameterization) [6]. The concortium consists of Aerospatiale CCR (F)), Alenia Aerospazio (I), CADOE (F), Centro Ricerche FIAT (I), COREP (I), Ericsson Saab Avionics AB (S), Eurocopter (F), KTH/PSCI (S) and MIP (F).

In section 2 the objectives of the project and an overview of the software package is presented. In section 3 the technical aspects are discussed. Section 5 presents the project status and results. Conclusions are presented in section 6.

2. OBJECTIVES AND TECHNICAL OVERVIEW

The objective of the project is to develop a software suitable for EMC and antenna applications and in particular well suited for parameter studies and optimization. For this purpose a frequency domain BEM solver and novel parameterization tools were chosen. The BEM solver has been used and validated for many EMC and antenna applications. The parameterization techniques for integral equation methods have been used by CADOE sucessfully in other areas like structural analyses, see for example [7].

The parameterization technique uses Taylor and Pade polynomials to approximate the solution in a parameter interval. The computation of the Taylor polynomials includes the inversion of the resulting matrix for a given parameter value and high order automatic differentiation of the same matrix. Since the solutions are given as polynomials it is very easy to carry out parameter studies within the obtained parameter interval and to find optimal solutions. The parameters that have been chosen to demonstrate the technology are frequency, impedance and geometry. The frequency is an obvious parameter to choose for EMC studies. The impedance can for example be the terminating impedance for a cable or the input impedance for an antenna. As regards geometry the connectivity between the nodes has to be maintained during the geometrical change which of course gives rise to limitations. The changes that can be carried out are however sufficiently large to be of importance. A typical application would be to change the cable routing so as to obtain as low currents as possible when a system is excited by an electromagnetic disturbance. Another would be antenna positioning and optimization of the shape of antenna elements.

Since the computation of the polynomials approximating the solution, in a parameter interval, is very complex the use of supercomputers is necessary for this part. However since this computation is carried out once or relatively few times it is motivated. This gives the most accurate solution and allows the user to tackle complex and realistic applications. On the other hand the preprocessing can be carried out on a workstation. Moreover the postprocessing including searching for an optimal solution can also be carried out on a workstation or even a PC since the parametrized solutions are represented by simple polynomials. This division between a supercomputer for the computation of the polynomials and a workstation environment for pre and postprocessing is very practical for the EMC or antenna specialist who is not necessarily a specialist on large scale computations.

The preprocessing is based on I-DEAS from SDRC as regards geometry and meshing. Either a CAD model is imported to, and if necessary repaired in, I-DEAS or the geometry is created within I-DEAS. The standard boundary element meshing tools in I-DEAS are used. I-DEAS was chosen for the project since it is a well known tool used by many companies, the BEM solver was already before the project started linked to I-DEAS and CADOE has an extensive knowledge of I-DEAS and have integrated their tools within I-DEAS. After the project the software could in a straightforward way be ported to similar tools like I-DEAS. The preprocessor also includes a graphical user interface (GUI) where the relevant parameters are set. In the GUI, which is linked to I-DEAS, general electromagnetic parameters like angle of incidence for an exciting wave, parameters linked to nodes in the mesh and parameters and limits for the parameterization is set.

The datavisualisation is mainly based on I-DEAS but other postprocessing tools will also be included. Furthermore the optimization software for the parametrized solutions is also linked to I-DEAS.

3. SOLVER AND PARAMETERISATION TECHNIQUE

The BE solver is based on the Stratton-Chu formulation of the EFIE [8]. The geometry can be modeled with triangular or line (beam) elements generated with I-DEAS where the line elements often are used to model cables or wire antennas. A large linear complex equation system, $Z\vec{J} = \vec{V}$, where Z is a matrix depending on the parameters of the system like geometry, impedance and frequency, $\vec{J}$ is a vector containing the unknowns and $\vec{V}$ is a vector containing the source terms, has to be solved. By inverting the so called impedance matrix Z, the unknowns $\vec{J}$ can be obtained. Normaly the inversion of this matrix

has to be repeated a large number of times when carrying out parameter studies. To obtain the solution in a parameter interval $\vec{J}$ is approximated by an asymptotic polynomial expansion including Taylor polynomials and Pade approximations. The coefficients of the Pade approximation, a rationale function of two polynomials, are computed from the coefficients of the Taylor polynomials. Each coefficient of the polynomial is obtained by computing the successive derivatives of $\vec{J}$, with respect to a parameter p, at a given value p_0. The higher order derivatives of Z are obtained by using an automatic differentiation tool, ADOC, developed by CADOE. The major interest of this approach comes from the ease of use of the Taylor polynomials. Even though the computation of the Taylor polynomials can be very complex it is only done one or a few times. The Taylor polynomials then allows the computation of $\vec{J}$ for a large set of parameters:

$$\vec{J}_p \approx \sum_{n=0}^{N} \frac{\vec{J}_{p_0}^{(n)}}{n!}(p - p_0)^n$$

High degree polynomial interpolation provide oscillating solutions but, for a large class of industrial problems, the solution depends analytically on the parameters of the studied structure. It means that the Taylor expansion converges to that solution. This is true for the continous problem and the discrete problem processed by the computer.

Moreover, the automatic differentiation tool ADOC, provides exact derivatives of the latter one by differentiating its algorithm. There is no truncation error like for the finite difference technique. To illustrate the analycity of the solution with respect to its parameters, the same binary data representation is obtained by a direct computation and by a Taylor's expansion (computed by automatic differentiation) [9]. To improve the range of validity for the parameterization different methods will be studied where the joint use of variable change, where a variable is replaced by a function, and the introduction of complex wave numbers is one example.

To obtain maximum efficiency the computation of the polynomials will be implemented on parallel machines. The most demanding part is the computation of the (inverse of the) matrix and its derivatives. This is a computation that is well suited for parallelization by using a message passing system like PVM or MPI. Furthermore, the storage of intermediary matrices during the computations requires large resources in terms of memory and disk. Only a parallel system can handle it easily and can deal with the required computing power, local storage and high I/O bandwidth.

4. DEMONSTRATOR APPLICATIONS

To validate and demonstrate the performance of the software several demonstration cases have been defined with applications regarding EMC and antenna problems for cars, aircraft and helicopters. To start with all industrial partners will study a simple object, i. e. a rectangular metallic box, on which measurements also are carried out. Secondly a complex and realistic test case will be studied where measurement results already exists. Finally a new complex test case will be studied to evaluate the full functionality of the software.

5. PROJECT STATUS AND RESULTS

Presently a first serial version of the software with parameterization in frequency have been tested and validated for the simple object mentioned in section 4. A parallel version of the code have just been installed on a 113 node IBM SP2 at KTH and the tests using parameterization in frequency will start for the complex applications described in section 4. First simple tests of the version including parameterization with respect to geometry and impedance are presently being carried out with promising results.

The simple object developed by CRFIAT is a metallic box with several apertures that can be closed or open and with a wire inside. The wire is terminated at the interior walls of the box with variable termination impedances. Measurements have been carried out by CRFIAT in an anechoic chamber at CRFIAT. The measured quantities are electric fields inside the box and wire currents in the frequency range 200 - 1000 MHz. The configuration that we have studied is shown in figure 1.

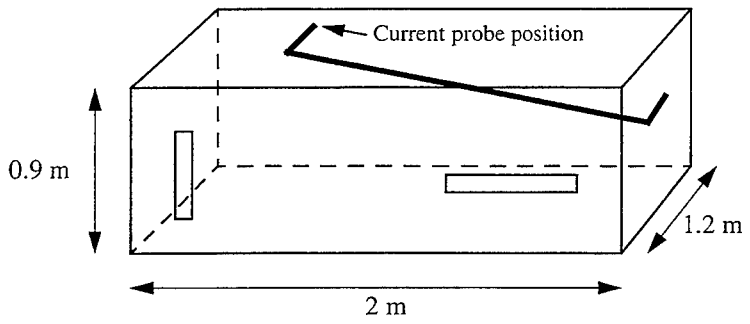


Figure 1. Metallic box.

In our case the wire termination impedances are 50 ohm at both ends. The apertures have dimensions $50mm \times 500mm$. The illumination is broadside to the apertures. In the computations a plane wave excitation is used. The computed quantity is the wire current at the position shown in figure 1. Two different meshes are used, one with a constant mesh size of 10 cm and the other one with a locally refined mesh at the apertures. The locally refined mesh is shown in figure 2. The number of unknowns for the 10 cm mesh is 3464 and 7383 for the finer mesh. Normally for a general purpose method like this an aperture has to be resolved by at least three nodes in each direction to obtain a good result which is also seen from the comparison with measurements.

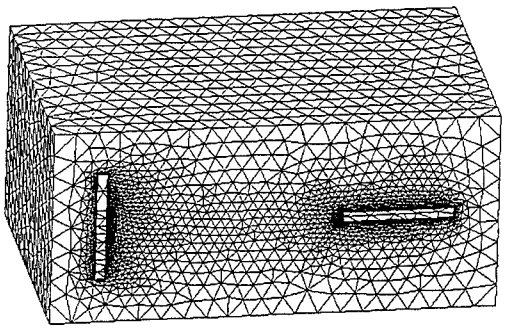


Figure 2. Mesh of the metallic box with 1 cm mesh size at the apertures.

In figure 3 the computed wire current using the non parametrized solver and the measurement results for vertical polarization is shown. As can be seen the agreement is very good. In particular if the finer mesh is used.

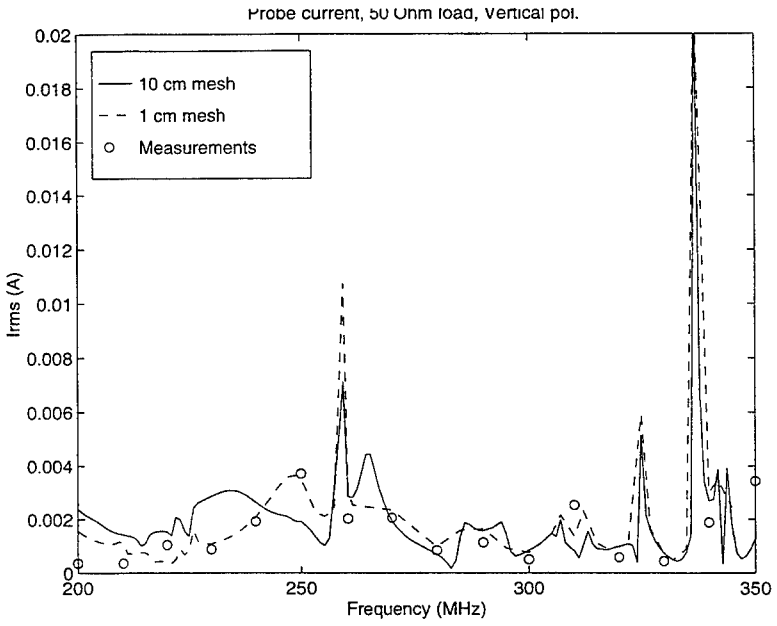


Figure 3. Wire RMS current as a function of frequency computed for the 10 cm mesh, solid line, 1 cm mesh, dashed. The results from measurements are shown by circles.

The parametrized version in frequency is still being improved as regards the range of validity so no final conclusions can be drawn at this stage. As an example figure 4 shows the results computed with the direct version compared with the results computed with the parametrized version for the same case as described above.

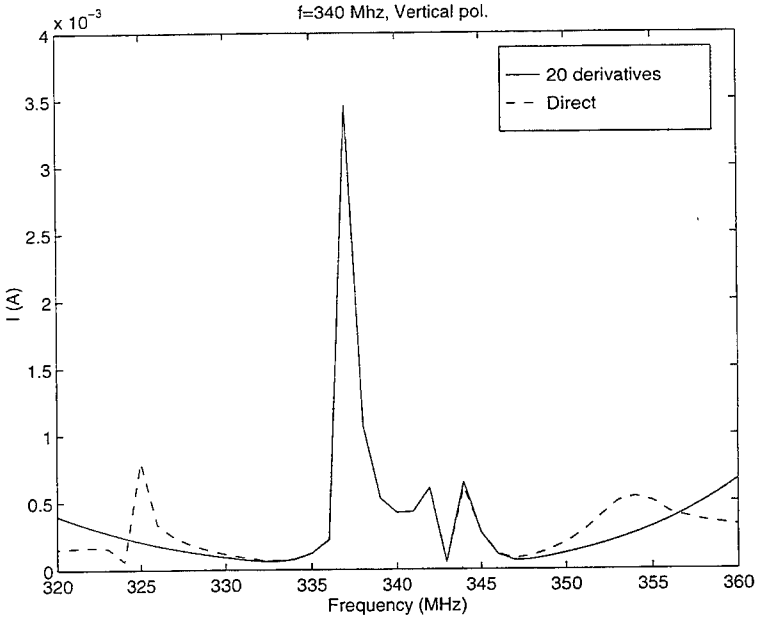


Figure 4. Wire current computed with parametrized version, solid, and with direct version, dashed.

The order of derivation is 20 and a good agreement is found between 330 and 350 MHz which yields a validity range of approximately 20 MHz in this case. Note that several sharp features are captured which demands many computations with the direct solver. It is important to notice that the parameterized version can accurately catch very narrow resonances by one computation which with a direct solver would require many computations.

6. CONCLUSIONS

Initial tests show promising results for using a parameterization technique for a frequency BEM v gives the solution in a parameter interval, for example a frequency interval. The method will be f improved to increase the range of validity and impedance and geometry will also be considered rameters.

AUTHORS INDEX

Abrams, R. I-20
Aidam, M. II-983
Alexopoulos, N.G. I-151, 179
Allard, R.J. II-964
Allen. J.M. I-36
Alpert, B.K. II-600
Archambeault, B. II-880, 888
Ares, F. II-1100
Arkko, A. I-121
Axe, M.R. I-554
Aygun, K. II-838, 866, 873
Azar, T. I-216
Baghai-Wadji, A.R. I-297, 305
Beng, B. II-1087
Bernstein, D. II-750
Beyer, A. I-344
Bindiganavale, S. I-405
Biswas, R. I-144
Blake , D.C. II-808
Bleszynski, E. I-441
Bleszynski, M. I-441
Blocher, T.W. I-36
Bogotyubov, A.N.
Bomholdt, J.M. I-554
Boots, B. II-687
Botros, Y.Y. II-639
Brauer, J.R. II-763
Brennan, C. I-390
Bridges, G. I-326
Burkholder, R. I-243
Bush, G.G. I-167
Byers, A. II-687
Cangellaris , A.C. II-721
Cai, M. I-73
Canino, L.S. I-413
Canning, F.X. I-449
Carroll, Jr., C.E. I-36
Caputa, K. I-476
Cascio, L. I-367
Castro, F. I-173
Chamberlin, K. II-880
Chang, Y. I-20
Cherkashin, Y.N. I-278
Chew, W.C. I-397, 420, 434,
Chew, W.C. I-500, 507, II-712
Christopoulos, C. I-490
Chung, Y-C. II-795
Cohen, E. I-576
Cohen, N. II-1047, 1051
Coren, R. I-216
Cown, B.J. I-270
Crosta, G.F. I-210
Crouch, D.D. I-144
Cuhaci, M. I-326
Cown, B.J. I-270
Crosta, G.F. I-210
Crouch, D.D. I-144
Cuhaci, M. I-326
Culien, P. I-390
Cummins, W.F. II-1043
Davey, K. I-188, 1106
Dawson, T.W. I-258, 476
De Cogan, D. I-331
DeLyser, R.R. I-105 359, 374
Demkowicz, L. II-799
Diaz, R.E. I-159, 179
Dinesen, P.G. II-858
Dionigi, M. II-952
Dombowsky, P. II-1006
Donepudi, , K. II-712
Drewniak, J.L. II-672, 905
Drozd, A.L.S. I-36
DuBroff, R. II-905
Dunn, J.M. II-1120
Ehmann, R. II-825
El Hachemi, M. II-1066
Elsherbeni, A.Z. I-99
Eriksson, G. I-87, II-1134
Eremenko, V.A. I-278
Ergin, A.A. II-838, 866, 873
Estrada, J.P. I-270
Evans, J. I-9, 15
Fan , G-X. II-655
Felsen, L.B. II-952
Feng, M. I-204
Filipovic, D.S I-405
Fisher, S.E. II-838
Fujiwara, K. II-779
Furse, C.M. I-482
Fuzi , J. I-283, 291
Gallivan, K. II-732
Gamer, D. I-581
Gaudine, D. I-546
Gisin, F. I-313
Gong, J. II-750
Gottlieb, D. II-607
Goverdhanam, K. I-127
Gravrok, R. II-687
Greengard, L.F. II-600
Greenwood , A.D. II-756
Grimes, C.A. I-194
Grimes, D.M. I-194
Gutschling, S. II-832
Hagstrom, T. II-600
Hallatt, D.M. I-15
Hantman, B. I-20
Hariharan, S.I. II-600
Haupt, R. II-1095
Haussmann, G. I-113, 531
He , J-Q. II-631, 845
Hecht, R.W. II-1058
Hesthaven, J.S. II-858
Hill, K.C. I-50
Ho, K.M. I-144
Hoefer, W.J.R. I-367
Horn, K. I-538
Hori, T. I-66
Hubing, T. II-905, 928
Ingersoll, G. I-28
Ivanyi, A. I-283
Jakobus, U. II-920, 936
Jandhyala, V. I-204, 434
Jaroszewicz, T. I-441
Jensen, M.A. I-428, 470
Ji, Y. II-928
Jin, J-M. I-250, 397, II-712, 756
Joseph, B. I-576
Kagoshima, K I-66
Kang, T-W. II-795
Kantartzis, N.V. II-623
Kashyap, N. II-905
Katehi, L.P.B. I-127, II-647
Kettunen, L. I-121
Kim, B-W. II-795
Kim, K.W. I-462
Kopylov, Y.V. I-381
Kyriazidou, S.A. I-151, 179
Kruger, H. II-832
Kubina, S.J. I-93, 546
Kuester, E.F II-1112
Kun, C. I-331
Kurokhtin, A.N. I-381
LaBelle, J. I-20
Landstorfer, F.M. II-936
Lauber, W.R. I-93
Lebaric, J. I-581
Lee, J-F. II-787
Lee , Y.H. II-991, 999
Leskiw, D.M. I-28
Levy, M.F. I-455
Ljepojevic, N. I-73
Lindenmeier, S. I-320
Ling, F. I-250
Liu , C. II-787
Liu , F. II-695
Liu, G. I-194
Liu, Q-H. II-631, 655, 845, 852, 972
LoVetri, J. I-326
Lu, C.C. I-397
Lu, Y. II-991, 999, 1087
Luu, Q. I-546
Lynov, J.P II-858

Lysiak. K. II-1006, 1012
MacMillan, H. II-1120
Mangold, T. I-351
Mannikko, P.D. II-1074
Marhefka, R.J. I-568
Martin, T. I-79
Maruyama, T. I-66
McFarlane, R. I-367
Merrill, M. I-151
Michielssen, E. I-204, 434, II-732, 838
Michielssen, E. II-866, 873
Miller, E.K. I-583, II-958
Minaev, D.V.
Mittra, R. I-515
Miyazaki, Y. II-911
Mongiardo, M. II-952
Muller, U. I-344
Nabet, B. I-173
Nguyen, N. II-972
Nott , A. 560
O'Brien, P.J. II-1074
Ommodt, K.W. II-1074
Ong, K.G. I-194
Ottusch, J.J. I-413
Paboojian, A. I-576
Packer, M.J. II-1028
Pantic-Tanner, F. I-313
Paul, J. I-490
Paynter, G.F. I-568
Pearlman, R.A. I-554
Perez, R. II-895
Pertz, O, I-344
Peterson, A.F. II-771
Petropoulos, P.G. II-615, 623
Pflug, D.R. I-42
Pierantoni, L. I-320
Piket-May, M. I-113, 531, II-687
Poltz, J. II-680
Popov, A.V. I-381
Rachowicz, W. II-799
Rahmat-Samii, Y. I-462
Rahola, J. II-1126
Ramahi, O.M. II-592, 880
Rebel, J. I-336
Redmond, R. I-28
Regarajan, S. I-264
Remis , R.F. II-704
Righi, M. I-367
Riley , D.J. II-816
Robertson, R. II-647
Rodriguez, J.A. II-1100
Rodriguez-Pereyra, V. I-99
Roedder, J.M. 554
Rogovin, K. I-449
Routier, R. I-243, 270
Roy, J.E. I-93
Ruoss, H-O. II-936
Russer, P. I-320, 336, 351, II-952, 983
Sandlin , B.S. II-1080
Savage, J.S. II-742
Schutt-Aine, J.E. II-695
Schwengler, T. II-1112
Selormey, P. II-911
Sengupta, D. I-204
Sertel, K. I-405
Severns, R. II-1036
Shaeffer, J. I-538
Shang, J.S. II-808
Shanker, B. I-204, 434, II-838, 866,873
Shen, C.Y. I-57
Sheng, X.Q. I-397
Shi, H. II-672
Shlepnev, Y.O. II-664
Siarkiewicz, K.R. I-2
Sigalas, M. I-144
Simons, N. I-326
Singh, D. I-560
Siushansian, R. I-326
Skochinski, E. I-264
Smith, C.E. I-99
Song, J.M. I-397, 420
Soukoulis, C.M. I-144
Speciale, R.A. I-224
Stach,J. II-979
Stalzer, M.A. I-413
Stillman, G. I-204
Stuchly, M.A. I-476
Sugiyama, K. II-779
Suzuki, T. I-136
Takahashi, N. II-779
Takehara. K. II-779
Tardioli, G. I-367
Tarhasaari, T. I-121
Teixeira, F.L. I-500, 507
Tentzeris, E. I-127, II-647
Terzuoli, A.J. II-1080
Thibblin, U. I-87, II-1134
Thomas, D.W.P. I-490
Thomas, K. I-113
Tosser-Roosey, A. II-1066
Tosser-Roussey, C. II-1066
Trueman, C.W. I-93, 523, 546
Tsiboukis, T.D. II-623
Tumer, C.D. II-816
Uehara, K. I-66
Ulander, L. I-79
Vall-llossera, M. I-93, 523
van den Berg, P.M. II-704
Vardapetyan, L. II-799
Van Doren, T. II-905
Vidoni. T. I-28
Vinogradov, A.V. I-381
Visher, J.L. I-413
Volakis, J.L. I-405, II-639
Voltmer, D. I-581
Walker, M.J. II-944
Wandzura, S.M. I-413
Wang, C-F. I-250
Wassef, K.N. II-771
Wedge, S. II-750
Weiland, T. II-825, 832
Weile, D.S. II-732
Werner, D.H. I-515, II-964
Werner, P.L II-964
Wharton, T. I-9
Wheless, Jr. W.P. II-1019
Wood, Jr., W.D. I-50
Woolf, S. I-576
Wright, H. I-20
Wurtz, L.T. II-1019
Yang, B. II-607
Yeo, K. II-1087
Yu, P.K.L. I-136
Zaporozhets, A.A. I-455
Zhao, L. II-721
Zheng, D. II-1106
Zmyslo, J.S. II-964
Zunoubi, M. II-712