

SLAC-431
UC-405
(A)

HIGH ACCURACY ELECTROMAGNETIC FIELD SOLVERS
FOR CYLINDRICAL WAVEGUIDES
AND AXISYMMETRIC STRUCTURES
USING THE FINITE ELEMENT METHOD*

Eric Michael Nelson

*Stanford Linear Accelerator Center
Stanford University
Stanford, CA 94309*

December 1993

Prepared for the Department of Energy
under contract number DE-AC03-76SF00515

Printed in the United States of America. Available from the National
Technical Information Service, U.S. Department of Commerce, 5285 Port
Royal Road, Springfield, VA, 22161.

* Ph.D. thesis

ABSTRACT

Some two-dimensional finite element electromagnetic field solvers are described and tested. For TE and TM modes in homogeneous cylindrical waveguides and monopole modes in homogeneous axisymmetric structures, the solvers find approximate solutions to a weak formulation of the wave equation. Second-order isoparametric lagrangian triangular elements represent the field.

For multipole modes in axisymmetric structures, the solver finds approximate solutions to a weak form of the curl-curl formulation of Maxwell's equations. Second-order triangular edge elements represent the radial (ρ) and axial (z) components of the field, while a second-order lagrangian basis represents the azimuthal (ϕ) component of the field weighted by the radius ρ . A reduced set of basis functions is employed for elements touching the axis.

With this basis the spurious modes of the curl-curl formulation have zero frequency, so spurious modes are easily distinguished from non-static physical modes. Tests on an annular ring, a pillbox and a sphere indicate the solutions converge rapidly as the mesh is refined. Computed eigenvalues with relative errors of less than a few parts per million are obtained.

Boundary conditions for symmetric, periodic and symmetric-periodic structures are discussed and included in the field solver. Boundary conditions for structures with inversion symmetry are also discussed. Special corner elements are described and employed to improve the accuracy of cylindrical waveguide and monopole modes with singular fields at sharp corners.

The field solver is applied to three problems: (1) cross-field amplifier slow-wave circuits, (2) a detuned disk-loaded waveguide linear accelerator structure and (3) a 90° overmoded waveguide bend. The detuned accelerator structure is a critical application of this high accuracy field solver. To maintain low long-range wakefields, tight design and manufacturing tolerances are required.

PREFACE

This dissertation is a summary of my work in the Accelerator Theory and Special Projects Group at the Stanford Linear Accelerator Center (SLAC). It concentrates on my work on finite element methods for computing electromagnetic modes in microwave cavities and waveguides.

Several applications of these methods to the design of microwave components are discussed. These are not intended to be complete discussions of such components, and it should be noted that the development of these components was a collaborative effort. My contribution, which mainly involved computing fields in various structures, was only a small part of the design effort. The discussions include some results from my collaborators.

The motivation for this work on finite element methods and its applications can be best understood by reviewing my history at SLAC. As will become apparent below, my entire graduate career has involved the numerical simulation of Maxwell's equations.

My first quarter at SLAC was spent with Prof. Elliot Bloom. He gave me the task of computing the signals from a beam position monitor given the position of the beam. With help from others in the TPC group, Gary Godfrey in particular, I employed a simple electrostatic finite-difference model to solve this problem. Not only was this an introduction to numerical solutions of partial differential equations, but it was also an introduction to the problems concerning the solution of large matrix problems. I used an iterative technique, SOR, to solve the matrix equations. Trying different size meshes and increasing the number of iterations raised the big question which would later be the focus of my work at SLAC: how well have I solved the problem? The answer at this time, unfortunately, was not well enough.

The following quarter I started working with Prof. Roger Miller, my future thesis advisor, on porting a klystron simulation code to the local IBM mainframe. SLAC has an active program working on RF power sources for the Next Linear Collider (NLC). This work got me interested in microwave power sources, and eventually lead to work on cross-field amplifiers (CFAs). Originally my interest in these devices concerned the interaction between the electron beam and the fields, but it turned out other problems would have to be solved first.

The anode circuit in the prototype SLAC CFA is a periodic structure, so much of the design and measurement effort concentrated on the dispersion diagram of this structure. To understand the modes and the dispersion diagram of the anode

circuit I wrote a finite difference eigenmode field solver for periodic two-dimensional structures. My motivation for doing so was that the popular field solvers at this time did not have quasi-periodic boundary conditions. This program, called QUAP, was my first success. With the help of my advisor Prof. Miller, Prof. Joseph Feinstein and Terry Lee, I used QUAP to understand the modes of the CFA and to interpret measurements of test structures which had previously been misunderstood.

The field calculations were so helpful in understanding the measurements of the test structure that I was soon asked to design a better CFA structure. At this point it became apparent that while QUAP was useful for understanding the qualitative behavior of the structure, it was not very good quantitatively. QUAP could compute the dispersion diagram with about 1% accuracy, but better accuracy was desired. A few attempts were made to improve the accuracy of QUAP. Some progress was made experimenting with iterative matrix solvers to handle finer meshes, but the improvement was insufficient, and QUAP was eventually abandoned.

My subsequent attempts to accurately model CFA structures were along two directions. First, with the help of Kwok Ko, I added quasi-periodic boundary conditions to the three-dimensional finite difference code ARGUS. The CFA structures were modelled on a CRAY supercomputer in order to accomodate very large meshes. The code was particularly useful as a tool to model the three-dimensional behavior of CFA structures, and it was particularly valuable for realistic simulations of the waveguide coupled structures. However, despite the use of large meshes with almost one million unknowns, the accuracy was still insufficient to obtain directly the design parameters to much better than 1%.

At the same time I was pursuing a second approach: use the finite element method instead of the finite differences to discretize the equations. I had recognized that much of my difficulty with QUAP and ARGUS was due to the inability of the finite difference mesh to closely model the true boundary of the structure. I felt the finite element method was an answer to this problem, and I proceeded to write a finite element field solver, again with quasi-periodic boundary conditions in order to model the periodic CFA circuits.

I started with a two-dimensional version as a test. Bill Hermansfeldt was at first dismayed that I was writing "yet another program" to solve Maxwell's equations, although I believe he later saw some utility in this effort. Regardless, the name YAP has been stuck to this finite element work ever since.

Within a few months I had written a mesh generator and a field solver which could solve for the modes of simple waveguides and cavities with very high accuracy

— much better than was necessary for the design of the CFA. Applications of the solver to the CFA, however, produced results which, while an improvement over the finite difference calculations, were not as fantastically accurate as the simple cavity calculations. This was due to the presence of sharp corners in the CFA model, and after another month of work the program was modified to give sharp corners special treatment. Much of the accuracy was recovered. Excellent accuracy was obtained without supercomputers.

Now it was time to consider a three-dimensional finite element field solver and related problems, such as matching the periodic CFA structure to input and output waveguides. I felt confident that with the exceptional accuracy I had achieved in two dimensions, I could get reasonable accuracy in three dimensions, and perhaps still avoid having to use supercomputers. Unfortunately, around the same time SLAC decided to stop pursuing CFAs as a power source for the NLC, and thus I looked for other thesis research.

Prof. Miller suggested I become involved in the NLC accelerator structure effort by adding to YAP the ability to compute modes in axisymmetric structures. This was fairly straightforward to do for $m = 0$ modes, and tests on simple cavities and waveguides showed impressive accuracy. Tests on the SLAC disk loaded waveguide accelerator structure also demonstrated the accuracy of the code.

With the advent of detuned accelerator structures at SLAC, I was asked to extend YAP to compute dipole ($m = 1$) modes, and subsequently design a detuned accelerator structure. The success of the detuned accelerator structure depends greatly on the accuracy with which the wakefields of the excited modes can be made to cancel by the time the next bunch arrives. Thus this problem uniquely requires the accuracy YAP can provide.

Solving for $m = 1$ modes was much more difficult than I expected due to the presence of spurious modes. However, once I solved the problem I finally felt I was part of the state of the art in field solving. Furthermore, it is an experience which will be valuable if I ever pursue three-dimensional field solvers.

In collaboration with Juwen Wang, the axisymmetric portion of the detuned accelerator structure was designed. Upon presenting the design, my department head, Prof. Ron Ruth, smiled and announced to the group that if the design was wrong, my Ph.D. would be revoked. To this day I'm not sure if he's kidding, but I have since taken solace in the fact that YAP can compute structures better than anyone can build them.

Prof. Norman Kroll encouraged me to apply YAP to the design of a 90° over-

moded waveguide bend, a problem for which YAP was once again uniquely suited, this time because of YAP's ability to handle non-integral m . In collaboration with Prof. Kroll and Chris Nantista, a few designs for overmoded waveguide bends were developed.

Now a word on the organization of this dissertation. Since the application of the finite element method to electromagnetic field solvers is relatively new to SLAC, I have chosen to include a brief introduction to the finite element method. Furthermore, since periodicity is a symmetry I've employed in most of my design calculations, the first chapter includes a discussion of the exploitation of symmetries in field solvers. I have noted with pleasure that some popular field solvers have recently added quasi-periodic boundaries to their repertoire. The reader will find additional symmetries in the first chapter.

Given the history of my work at SLAC, I hope the reader will understand that among the applications of YAP I discuss, the CFA results are only qualitative, and definite design details can be found only for the detuned accelerator structure and the 90° overmoded waveguide bends. Finally, I hope the reader will appreciate what I consider the truly original component of this work: the ability to accurately compute $m \neq 0$ modes of axisymmetric structures via the finite element method.

ACKNOWLEDGEMENTS

Many people contributed to this work, supported me in this work, or prepared me for this work. I regret I can only give a partial list of these individuals. Let me start with those that helped with my research at SLAC — many have already been mentioned in the preface.

First and foremost is my thesis advisor, Prof. Roger Miller, who provided help, direction and much encouragement for my work. Kwok Ko introduced me to the field solving community, provided lots of help with ARGUS and encouraged me to pursue better field solving algorithms. Prof. Norman Kroll provided lots of encouragement to extend the capabilities of YAP and to apply YAP to various problems. These three have been involved in almost everything I've done at SLAC, and I appreciate it very much. I also appreciate recent conversations with Bob Warnock, who encouraged me to extend and test YAP as well.

Prof. H. Alan Schwettman and Prof. Malcolm Beasley kindly reviewed this work, along with Prof. Ron Ruth, who supported and encouraged this work as head of the accelerator theory group.

My graduate student peers at SLAC, particularly Chris Nantista and Xintian "Eddie" Lin, provided camaraderie which I enjoyed very much. I also enjoyed long and varied discussions with Jim Spencer, Martin Lee and Cho Ng. These people made SLAC a friendly place.

Special thanks to Juwen Wang and the whole accelerator structures group, for including YAP in the design of the detuned accelerator structure. I also thank Harry Hoag and Hank Deruyter for providing measurements of short test structures which verified the accuracy of YAP.

Terry Lee and Prof. Joseph Feinstein, in addition to others mentioned above, taught me much about CFAs and klystrons, and were patient with me while I wrote the first versions of YAP. Chris Nantista and Prof. Kroll also deserve special thanks for including YAP in their designs of overmoded waveguide bends and to check designs of other components.

Bill Hermansfeldt, Tom Knight and Barbara Woo provided much support along with answers to many questions. Kathy Asher, Marge Bengali and Tanya Boysen also provided help whenever I asked.

I would also like to acknowledge help from and discussions with Karl Bane, Ken Eppley, Dave Farkas, Sam Heifets, John Irwin, Sam Kheifets, Walter Moshhammer, Bob Siemann, Kathy Thompson, Perry Wilson and many others in the accelerator theory and klystron groups.

I am very grateful to Alex Chao, who kindly provided me with a break from field solving for one summer. He taught me about coherent beam-beam effects during a visit to the SSC. While there I also had many enjoyable discussions with people such as Yiton Yan, Bo Chen and Xiao-Tong Yu.

My Stanford peers also deserve to be mentioned. I appreciate the camaraderie of my classmates both in and out of physics. I would especially like to thank Tatsuoki Nagaishi, David Williams, and Jim Chiang for being excellent roommates, Chris Perez for being a constant source of entertainment, and Mary Jayne Adriaans for her patient help and enduring friendship.

My preparation in Minnesota should also be acknowledged. As early as high school, I have been prepared for science and engineering by teachers like Daniel Whipple and Gerry Fisher. At the University of Minnesota I had excellent teachers, particularly Prof. Earl Peterson and Prof. W.T. Peria. Stimulating interaction with friends and fellow students, such as Cheryl Cheng, Julie Ghanbari, Marvin Goulet, Elina Hayosh, Jason Kalin and Dipender Saluja, was also memorable and very helpful. During many trips home to Minnesota, I relied on many friends there, particularly Sean McCoy, Tim McNeil and Lai Nguyen.

Last, but not least, is the unwavering support from my family, especially my parents. Mom always wanted a PhD son, and now she has at least one.

I am indebted to you all. Thank you.

TABLE OF CONTENTS

Abstract	ii
Preface	iii
Acknowledgements	vii
Table of Contents	ix
List of Tables	xi
List of Figures	xiii
1. Maxwell's Equations and Symmetry	
in Microwave Structures	1
1.1. Maxwell's Equations in Source-Free Linear Media	1
1.1.1. Eigenmode Problems	3
1.2. Structure Symmetries	4
1.2.1. Reflection Symmetry About a Plane	5
1.2.2. Inversion Symmetry	9
1.2.3. Periodic Structures	11
1.2.4. Symmetric Periodic Structures	14
2. Planar Structures and an Introduction	
to the Finite Element Method.	17
2.1. Maxwell's Equations	17
2.2. Variational Formulation	20
2.3. Finite Element Method	21
2.3.1. Elements	23
2.3.2. Linear Triangle Elements	25
2.3.3. Quadratic Triangle Elements	29
2.3.4. Tests on Parallel Plates	31
2.3.5. Tests on a Circular Waveguide	34
2.4. Sharp Corners	38
2.4.1. Special Corner Elements	39
2.4.2. Tests on a Ridged Waveguide	40
3. Axisymmetric Structures: TM_0 and TE_0 Modes	43
3.1. Wave Equation in Cylindrical Coordinates	43
3.2. Variational Formulation	46
3.3. Finite Element Formulation	47
3.4. Tests	48
3.5. Sharp Corners	51

4. Axisymmetric Structures: $m \neq 0$ Modes	52
4.1. Maxwell's Equations	52
4.2. Variational Formulation	53
4.3. Finite Element Formulation	55
4.3.1. Linear Edge Elements	56
4.3.2. Quadratic Edge Elements	58
4.3.3. Tests on an Annular Ring	60
4.4. Axis Elements	66
4.4.1. Axis Elements for $m \neq 0$	68
4.4.2. Tests on a Pillbox	71
4.4.3. Tests on a Sphere	82
4.4.4. Axis Elements for $m = 0$	84
5. Cross-Field Amplifier	92
5.1. DC Fields and the Brillouin Flow Model	93
5.2. Interaction with the RF Field	98
5.3. Simulations of the Electron-Circuit Interaction	103
5.4. Tests on a Hole and Slot Circuit	104
5.5. Other Slow-Wave Circuits	107
6. X-band Accelerator Structure	113
6.1. Accelerating Mode	113
6.2. Dipole Modes	117
6.3. Structure RF Parameters	119
7. 90° Overmoded Circular Waveguide Bend	124
7.1. Modes in Curved Guide	124
7.2. Scattering at the Interface	125
7.3. Around the Bend	126
7.4. 90° Bend Design	127
7.5. Further Work	129
8. Conclusion	130
Appendix A: Implementation of YAP	131
A.1. Planar and Monopole Formulations	131
A.1.1. Mesh Generator and Boundary Model Input Format	131
A.1.2. Finite Element Mesh File Format	140
A.1.3. Field Solver	142
A.2. Multipole Mode Formulation	143
A.2.1. Mesh Format and Refinement	144
A.2.2. Field Solver	149
References	154

LIST OF TABLES

Table 2.1: Weights and points for a 12-point scheme for numerical integration over the domain $\hat{\Omega}$. The first two weights are repeated three times for the three permutations of the corresponding triangular coordinates. The last weight is repeated six times for the six permutations of its triangular coordinates.	32
Table 4.1: Some numerical values for the cutoff wavenumbers $\gamma_{m,n}$ (for TM_m modes) and $\gamma'_{m,n}$ (for TE_m modes) of an annular ring with $a = 1/2$ and $b = 1$.	61
Table 4.2: Comparison between E and H calculations for the 10 lowest $m = 1$ modes of an annular ring with $a = 1/2$, $b = 1$ and $l = 1/2$. The mesh consists of 512 quadratic triangular elements.	64
Table 4.3: Some numerical values for the cutoff wavenumbers $\gamma_{m,n}$ (for TM_m modes) and $\gamma'_{m,n}$ (for TE_m modes) of circular waveguide with radius $a = 1$.	71
Table 4.4: Relative error of the computed eigenvalues for the 10 lowest $m = 1$ modes of a pillbox with $a = 1$ and $l = 1$ and using a fine mesh (512 quadratic elements).	74
Table 4.5: Relative error of the computed eigenvalues for the 10 lowest $m = 2$ modes of a pillbox with $a = 1$ and $l = 1$ and using a fine mesh (512 quadratic elements).	76
Table 4.6: Relative error of the computed eigenvalues for the 10 lowest $m = 10$ modes of a pillbox with $a = 1$ and $l = 1$ and using a fine mesh (512 quadratic elements).	78
Table 4.7: Some approximate numerical values for the wavenumbers $k_{l,n}$ (for TE modes) and $k'_{m,n}$ (for TM modes) of a sphere with radius $a = 1$.	83
Table 4.8: Relative error of the computed eigenvalues for the 10 lowest $m = 1$ modes of a sphere with $a = 1$ and using a fine mesh (1024 quadratic elements).	86

Table 4.9: Relative error of the computed eigenvalues for the 10 lowest $m = 0$ modes of a pillbox with $a = 1$ and $l = 1$ and using a fine mesh (512 quadratic elements).	91
Table 5.1: Selected SLAC Backward Wave CFA Parameters.	95
Table 6.1: Structure RF parameters.	123
Table 7.1: Dimensions and propagation constants for various bend designs. The cutoff frequency for the sixth mode is also listed.	128
Table A.1: Node boundary flags in <code>mg1.yap2h</code> .	141
Table A.2: Side boundary flags in <code>mg1.yap2h</code> .	141

LIST OF FIGURES

- Figure 1.1: A generalized microwave structure. The interior of the microwave structure is Ω , the boundary (metal walls) of the structure is Γ , and \hat{n} is the outward unit normal vector at the boundary. 2
- Figure 1.2: (a) A structure with symmetry plane P . The shaded region is the interior Ω , and the normal to the symmetry plane is \hat{p} . (b) The region over which Maxwell's equations must be solved is reduced to Ω_r for a symmetric structure. The symmetry boundary is Γ_{sym} . 6
- Figure 1.3: (a) A structure with two symmetry planes, P_1 and P_2 . (b) The region over which Maxwell's equations must be solved is reduced to Ω_r . The symmetry boundaries are $\Gamma_{\text{sym-1}}$ and $\Gamma_{\text{sym-2}}$. 8
- Figure 1.4: (a) A structure with reflection symmetry about a point. (b) A 2d structure which is symmetric about its center \mathbf{x}_o . The shaded region is the interior Ω . (c) The region over which Maxwell's equations must be solved is reduced to Ω_r for a symmetric structure. The symmetry boundaries are Γ_A and Γ_B . 9
- Figure 1.5: An example of a finite difference mesh exploiting symmetry about a center \mathbf{x}_o . 10
- Figure 1.6: Some examples of periodic structures. The dashed lines delimit one period, or cell, of the structure. The arrows indicate the action of the symmetry operations: (a) translation, (b) rotation and (c) glide reflection. The symmetry operation for a helical structure (d) is a combination of translation and rotation. 12
- Figure 1.7: One cell of the periodic structure of figure 1.6a. Γ_{left} and Γ_{right} are the quasi-periodic boundaries. 13
- Figure 1.8: (a) A symmetric periodic structure. The action of operator \hat{R} is indicated by the arrow. The dashed lines are symmetry planes. (b) The region over which Maxwell's equations must be solved is reduced to Ω_{1r} . The symmetry boundaries are $\Gamma_{\text{sym-left}}$ and $\Gamma_{\text{sym-right}}$. 14

Figure 2.1:	A cylindrical waveguide with arbitrary cross-section. The cross-section is shown at right.	18
Figure 2.2:	Two finite element meshes for the arbitrary cylindrical waveguide in figure 2.1. Mesh (a) is composed of linear elements. Mesh (b) is composed of quadratic elements.	23
Figure 2.3:	The element transformation \mathbf{x}_e maps the master element $\hat{\Omega}$ to the e th element Ω_e . The map \mathbf{x}_e is linear.	26
Figure 2.4:	Local basis functions \hat{N}_i for the linear triangle element. The size of the circles are proportional to the magnitude of the basis function at the center of the circles.	27
Figure 2.5:	A global basis function composed of linear elements. The basis function is associated with the field at the common corner. The size of the circles are proportional to the magnitude of the basis function at the center of the circles.	28
Figure 2.6:	The element transformation \mathbf{x}_e maps the master element $\hat{\Omega}$ to the e th element Ω_e . The map \mathbf{x}_e is quadratic.	29
Figure 2.7:	Local basis functions \hat{N}_i for the quadratic triangle element. The size of the circles are proportional to the magnitude of the basis function at the center of the circles. A cross in the circle indicates the basis function is negative.	30
Figure 2.8:	Two global basis functions composed of quadratic elements. The left basis function is associated with the field at a corner node, while the right basis function is associated with the field at a midside node. The size of the circles are proportional to the magnitude of the basis function at the center of the circles. A cross in the circle indicates the basis function is negative.	31
Figure 2.9:	Two parallel plates separated by distance d treated as a periodic structure with period L .	32
Figure 2.10:	Computed eigenvalue k^2 for the parallel plate test as the mesh is refined. Results from three different algorithms are shown. The dashed line is the exact result obtained analytically.	33

Figure 2.11:	Relative error of the computed eigenvalue k^2 for the parallel plate test as the mesh is refined. Results from three different algorithms are shown. The curves for the finite difference and linear finite element algorithms coincide.	33
Figure 2.12:	The region $\tilde{\Omega}$ (a) and a finite difference mesh (b) for a circular waveguide with radius a . The finite difference mesh was generated for the program QUAP.	35
Figure 2.13:	Computed eigenvalue k^2 for the circular waveguide test as the mesh is refined. Results from five different algorithms are shown. The dashed line is the exact result obtained analytically.	36
Figure 2.14:	Relative error of the computed eigenvalue k^2 for the circular waveguide test as the mesh is refined. Results from five different algorithms are shown.	36
Figure 2.15:	Computed and adjusted eigenvalues k^2 for the circular waveguide test as the mesh is refined. The dashed line is the exact result obtained analytically. The solid line is a fit to the adjusted eigenvalues.	38
Figure 2.16:	Polar coordinates near a sharp corner.	39
Figure 2.17:	The master element (a) for the seven node special corner element. Three lines of constant u (b) and constant v (c) on the master element are shown. Lines of constant v intersect at the sharp corner, node 3. The coordinate u corresponds to a distance from the sharp corner, while the coordinate v corresponds to an angle around the sharp corner. The basis functions for the special corner element are written in (u, v) coordinates.	40
Figure 2.18:	Ridged waveguide test problem. Symmetry reduces the problem by one-half.	40
Figure 2.19:	Relative error of the computed eigenvalue k^2 versus mesh refinement for the lowest TE mode of the ridged waveguide test problem.	41

- Figure 3.1: The region $\tilde{\Omega}$ for circular waveguide with radius a treated as a periodic structure with length l . The z - ρ (a) and x - y (b) cross-sections are shown. The dashed lines are quasi-periodic boundaries. 49
- Figure 3.2: The region $\tilde{\Omega}$ (a) for a sphere with radius a . The dashed line is a symmetry plane. A coarse finite element mesh (b) for the sphere. 49
- Figure 3.3: Relative error of the computed eigenvalues k^2 for the TE₁₁₀ mode of a sphere (solid line) and the cutoff TE₀₁ mode of circular waveguide (dashed line). 50
- Figure 3.4: Relative error of the computed eigenvalues k^2 for the TM₁₁₀ mode of a sphere (solid line) and the cutoff TM₀₁ mode of circular waveguide (dashed line). 50
- Figure 4.1: Vector fields \mathbf{V}_i (a) tangential to side i , reciprocal vector fields \mathbf{R}_i (b) normal to side i , local basis functions N_i (c) representing \tilde{E}_ϕ and local vector basis functions \mathbf{N}_i (d) representing \mathbf{E}_t . The triangles represent the vector field in the plane of the page: the size of the triangle is proportional to the magnitude of the field and the triangle points in the direction of the field. 57
- Figure 4.2: A global vector basis function comprising one local vector basis function on each of two elements. The field tangential to the element boundary is continuous while the field normal to the boundary is discontinuous. 58
- Figure 4.3: Quadratic local basis functions: N_1 through N_6 represent \tilde{E}_ϕ , \mathbf{N}_7 through \mathbf{N}_{12} represent \mathbf{E}_t . 59
- Figure 4.4: The region Ω for an annular ring with inner radius a , outer radius b and length l . The z - ρ (a) and x - y (b) cross-sections are shown. 61
- Figure 4.5: A course mesh (two elements) and four refined meshes for the annular ring problem. Each mesh covers the shaded region Ω . 62

- Figure 4.6: Field plots of the 10 lowest $m = 1$ modes of an annular ring. The ϕ -component of the fields is imaginary. The size of the circles are proportional to the ϕ -component of the field at the center of the circle. A cross in the circle indicates the ϕ -component is negative. The triangles indicate the size and direction of the z and ρ components of the field. 63
- Figure 4.7: Relative error of the computed eigenvalues for the $m = 1$ modes of an annular ring as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} . 64
- Figure 4.8: Frequency $k = \omega/c$ versus azimuthal symmetry number m for (a) four TM modes and (b) four TE modes. 66
- Figure 4.9: Relative error of the computed eigenvalue k^2 versus the azimuthal symmetry number m . The top two figures show the TM mode errors, the bottom two figures show the TE mode errors. The mesh consists of 512 quadratic elements. 67
- Figure 4.10: Local basis functions for the $m = 1$ problem for a linear element touching the axis. The bottom edge or node is on axis. The vectors (triangles) represent \mathbf{E}_t and the circles represent \tilde{E}_ϕ . 69
- Figure 4.11: Local basis functions for the $m = 1$ problem for a quadratic element touching the axis. The bottom edge or node is on axis. The vectors (triangles) represent \mathbf{E}_t and the circles represent \tilde{E}_ϕ . 70
- Figure 4.12: The region Ω for a pillbox with radius a and length l . The z - ρ (a) and x - y (b) cross-sections are shown. 71
- Figure 4.13: A course mesh (two elements) and four refined meshes for the pillbox problem. Each mesh covers the shaded region Ω . 72
- Figure 4.14: Field plots of the 10 lowest $m = 1$ modes of the pillbox. The ϕ -component of the fields is imaginary. 73
- Figure 4.15: Relative error of the computed eigenvalues for the $m = 1$ modes of a pillbox as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} . 74

Figure 4.16: Field plots of the 10 lowest $m = 2$ modes of the pillbox. The ϕ -component of the fields is imaginary.	75
Figure 4.17: Relative error of the computed eigenvalues for the $m = 2$ modes of a pillbox as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} .	76
Figure 4.18: Field plots of the 10 lowest $m = 10$ modes of the pillbox. The ϕ -component of the fields is imaginary.	77
Figure 4.19: Relative error of the computed eigenvalues for the $m = 10$ modes of a pillbox as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} .	78
Figure 4.20: Frequency $k = \omega/c$ versus azimuthal symmetry number m for (a) four TM modes and (b) four TE modes.	80
Figure 4.21: Relative error of the computed eigenvalue k^2 versus the azimuthal symmetry number m . The top two figures show the TE mode errors, the bottom two figures show the TM mode errors.	81
Figure 4.22: The region Ω for a sphere with radius a . The dashed boundary is a symmetry plane.	83
Figure 4.23: A course mesh (one element) and five refined meshes for the sphere problem. Each mesh covers the shaded region Ω .	83
Figure 4.24: Field plots of the 10 lowest $m = 1$ modes of the sphere. The ϕ -component of the fields is imaginary.	85
Figure 4.25: Relative error of the computed eigenvalues for the $m = 1$ modes of a sphere as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} .	86
Figure 4.26: Local basis functions for the $m = 0$ problem for a linear element touching the axis. The bottom edge or node is on axis. The vectors (triangles) represent \mathbf{E}_t . There are no basis functions for \tilde{E}_ϕ .	87

- Figure 4.27: Local basis functions for the $m = 0$ problem for a quadratic element touching the axis. The bottom edge or node is on axis. The vectors (triangles) represent \mathbf{E}_t and the circles represent \tilde{E}_ϕ . 88
- Figure 4.28: Field plots of the 10 lowest $m = 0$ modes of the pillbox. The ϕ -component of the fields is imaginary. 90
- Figure 4.29: Relative error of the computed eigenvalues for the $m = 0$ modes of a pillbox as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} . 91
- Figure 5.1: The prototype SLAC backward wave cross-field amplifier. The RF wave propagates clockwise, while the electrons travel counter-clockwise. 93
- Figure 5.2: Schematic diagram of a cross-field amplifier in the x - y plane. The direction of the static electric and magnetic fields are shown. The electron drift velocity \mathbf{v} is also shown for two locations in the electron sheath. 94
- Figure 5.3: Schematic diagram of a straight cross-field amplifier. The dotted lines at the anode indicate the simplified model assumes the anode is smooth. 95
- Figure 5.4: Planar non-relativistic Brillouin flow: (a) space charge density ρ , (b) electron velocity v/c , (c) electric field E_y and (d) electric potential $V(y)$. 97
- Figure 5.5: Operating point for the SLAC CFA. The Hull cutoff voltage V_c and Buneman-Hartree threshold voltage V_{BH} delimit the three regions where the CFA is expected to (1) conduct, (2) amplify or (3) insulate. The curves are obtained from the planar non-relativistic Brillouin model. 99
- Figure 5.6: One wavelength of the electric field of a single space harmonic of the RF travelling wave: (a) RF field, (b) static field and (c) combined field. 101
- Figure 5.7: Comparison of computed and measured dispersion diagrams for a slow-wave circuit. 105

- Figure 5.8: Computed eigenvalue k^2 versus mesh refinement for the 135° mode of a hole and slot circuit for the SLAC CFA. Finite difference and various finite element methods are shown. The dashed line is the extrapolation of the quadratic FEM calculations with special corner elements to zero node spacing. 106
- Figure 5.9: Dispersion curves for (a) a slow-wave circuit and (b) a smooth waveguide. The operating point and both dispersion curves are shown in (c). The dashed line in (c) is the smooth waveguide. 108
- Figure 5.10: Waveguide coupled slow-wave circuits. The guides are coupled (a) every cell, (b) every other cell and (c) every third cell. 109
- Figure 5.11: Dispersion diagrams of multiple cell slow-wave circuits. Diagrams (a) and (b) treat the uncoupled structure as a periodic structure with a period twice and three times the length of the original cell. The dashed lines are the dispersion curves of the waveguide. Diagram (c) is the dispersion diagram of the structure coupled every third cell. 111
- Figure 6.1: Disk loaded waveguide in the z - ρ plane. 114
- Figure 6.2: Two coarse meshes for cells of the disk loaded waveguide. The meshes are composed of quadratic triangular elements. Mesh (a) is for the first cell of the structure and mesh (b) is for the last cell of the structure. 115
- Figure 6.3: The cell diameter $2b$ (in cm) which yields synchronism with the beam ($v_\phi = c$). The dots are the 206 cells of the detuned accelerator structure. The rightmost dot is the first (input) cell and the leftmost dot is the last (output) cell. 116
- Figure 6.4: The logarithmic derivative $-b/f \partial f_o / \partial b$ with $2b$ determined from the synchronism condition. The dots are the 206 cells of the detuned accelerator structure. 116
- Figure 6.5: The lowest synchronous dipole mode frequency f_1 (in GHz). The cell diameter $2b$ is determined from the synchronism condition. The open circles are the dimensions at which f_1 was computed. The dots are the 206 cells of the detuned accelerator structure. 118

- Figure 6.6: Group velocity v_g/c for the lowest synchronous dipole mode. The cell diameter $2b$ is determined from the synchronism condition. The open circles are the dimensions at which f_1 was computed. The dots are the 206 cells of the detuned accelerator structure. 118
- Figure 6.7: Dispersion diagram for the two lowest dipole modes of three different cells of the structure. The dashed line is the first (input) cell, the dotted line is the middle cell and the solid line is the last (output) cell. The dot-dash line is the velocity of light line. 120
- Figure 6.8: Dispersion diagram for the dipole modes of three different cells of the structure. The dashed line is the first (input) cell, the dotted line is the middle cell and the solid line is the last (output) cell. The dot-dash line is the velocity of light line. 121
- Figure 6.9: Accelerating gradient along the length of the structure for 100 MW input power and various structure types. 122
- Figure 6.10: Peak surface gradient along the length of the detuned accelerator structure for 100 MW input power. 122
- Figure 7.1: Outer geometry (a) and cross-section (b) of the bend. The cross-section's dashed line is a symmetry plane. 125
- Figure 7.2: Dispersion diagram of the curved guide for the first design listed in table 7.1. The dashed line is the drive frequency 11.424 GHz. The dotted line corresponds to the speed of light along the center of the guide. 125
- Figure 7.3: Transmission of the fundamental mode through a rectangular waveguide bend over angle ϕ_b . Loss is due to conversion into higher modes. 127
- Figure 7.4: Electric field patterns for the five propagating modes of the first design in table 7.1. 128
- Figure A.1: Relationship between programs, files and input parameters for the YAP implementation of the planar and $m = 0$ formulations. Programs and files are on the left, input parameters are on the right. 132

- Figure A.2: Boundary model grammar in a format similar to the format for the YACC parser generator. Uppercase words are terminal symbols. Except for `INTEGER` and `REAL`, the terminal symbols are literal. 133
- Figure A.3: An example of a boundary model. The three figures show the nodes (left), halfedges (middle), loops and regions (right). The halfedge arrows point to the next halfedge of the loop. 138
- Figure A.4: Format of the finite element mesh file `mg1.yap2h`. 140
- Figure A.5: Command line options for field solver `y2d`, from the online help file. 142
- Figure A.6: Construction of a series of successively refined quadratic meshes from a single coarse mesh of linear elements. Sample command lines are given below the flow diagram. 145
- Figure A.7: Conversion of a quadratic finite element mesh for the NeXTstation (multipole) suite of programs to a quadratic finite element mesh for the IBM/VM (monopole and planar) programs. A sample command is given below the flow diagram. 146
- Figure A.8: Mesh grammar in a format similar to the format for the YACC parser generator. Uppercase words are terminal symbols. Except for `INTEGER`, `REAL`, `PM` and `QUOTEDSTRING`, the terminal symbols are literal. 147
- Figure A.9: Two examples of linear finite element meshes. The top example is a pillbox. The bottom example is a sphere. 148
- Figure A.10: Check and plot of a mesh. A sample command is given below the flow diagram, and the command options are listed below that. 150
- Figure A.11: Relationship between programs, files and input parameters for the YAP implementation of the multipole mode formulation. Programs and files are on the left, input parameters are on the right. 151

Figure A.12: A sample command and command options for the program **axiMnz**. The first command example generates a matrix, while the second example generates a field plot from a solution vector. 152

Figure A.13: A sample command and command options for the program **ma28Ges**. The first command writes a factorization pattern **fp**, while second command reads the factorization pattern **fp**. 152

Chapter 1

Maxwell's Equations and Symmetry in Microwave Structures

This chapter presents a mathematical statement of the microwave cavity eigenmode problem. In many structures this problem cannot be solved analytically, but accurate approximate solutions can be obtained using, for example, the algorithms described in chapters 2–4. The first section reviews Maxwell's equations in a closed cavity with perfectly conducting walls. Further discussions of Maxwell's equations applied to microwave structures can be found, for example, in [1] and [2]. Electromagnetic quantities are expressed using the gaussian system of units. The second section discusses some symmetries of microwave structures. These symmetries reduce the size of the eigenmode problem by using appropriate boundary conditions.

1.1. Maxwell's Equations in Source-Free Linear Media

Let the region Ω represent the interior of a microwave structure, and let the interior be source-free, so there are no electric charges or currents in the region Ω . Furthermore, assume that the materials inside the structure behave linearly with respect to the electromagnetic fields. Then Maxwell's equations in the region Ω in the frequency domain are

$$\nabla \times \mathbf{E} = \frac{i\omega}{c} \mathbf{B} \quad \text{and} \quad \nabla \times \mathbf{H} = -\frac{i\omega}{c} \mathbf{D} \quad \text{in } \Omega, \quad (1.1a)$$

$$\nabla \cdot \mathbf{D} = 0 \quad \text{and} \quad \nabla \cdot \mathbf{B} = 0 \quad \text{in } \Omega, \quad (1.1b)$$

with linear constitutive equations

$$\mathbf{D} = \epsilon \mathbf{E} \quad \text{and} \quad \mathbf{B} = \mu \mathbf{H} \quad \text{in } \Omega. \quad (1.1c)$$

The electric permittivity ϵ and magnetic permeability μ are rank 2 tensors (3×3 matrices) which may depend on the position \mathbf{x} inside the region Ω , and they may also depend on the frequency ω . Note that for $\omega \neq 0$, equations (1.1b) follows from (1.1a). Materials are often isotropic, in which case the material properties can be replaced by scalars ϵ and μ . If the region Ω is empty (a vacuum) then $\epsilon = 1$ and $\mu = 1$.

The boundary of the region Ω is Γ . Let $\hat{\mathbf{n}}$ be the unit vector normal to the boundary and directed outward, as shown in figure 1.1. The boundary Γ may include a variety of types of boundaries. Let Γ_{metal} represent the metal walls of the structure, and for now let the entire boundary Γ be Γ_{metal} . Other types of boundaries will be introduced in the following section. Metals usually conduct well

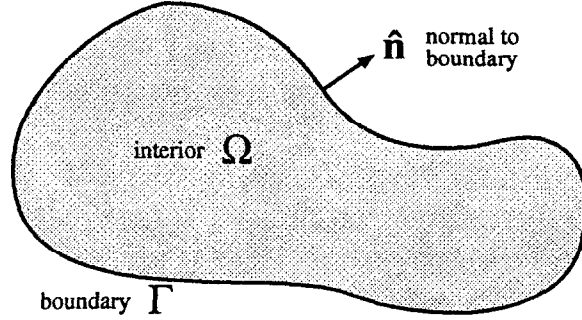


Figure 1.1: A generalized microwave structure. The interior of the microwave structure is Ω , the boundary (metal walls) of the structure is Γ , and $\hat{\mathbf{n}}$ is the outward unit normal vector at the boundary.

enough to be approximated as perfect conductors. The boundary condition at a perfectly conducting metal wall is

$$\hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}}. \quad (1.2)$$

Another boundary condition,

$$\hat{\mathbf{n}} \cdot \mathbf{B} = 0 \quad \text{on } \Gamma_{\text{metal}}, \quad (1.3)$$

can be derived from boundary condition (1.2) and Maxwell's equations (1.1), assuming there is no static magnetic field.

Equations (1.1) are in the frequency domain. The fields in the time domain are obtained from the fields in the frequency domain by using the appropriate fourier transformation. For example, consider an undriven cavity with a discrete set of normal modes. If the j th mode has frequency ω_j , and the fields of the j th mode are \mathbf{D}_j , \mathbf{E}_j , \mathbf{B}_j and \mathbf{H}_j , then the fields in the time domain are

$$\mathbf{D}(\mathbf{x}, t) = \Re e \left\{ \sum_{j=1}^{\infty} \mathbf{D}_j(\mathbf{x}) e^{-i\omega_j t} \right\}, \quad (1.4a)$$

$$\mathbf{E}(\mathbf{x}, t) = \Re e \left\{ \sum_{j=1}^{\infty} \mathbf{E}_j(\mathbf{x}) e^{-i\omega_j t} \right\}, \quad (1.4b)$$

$$\mathbf{B}(\mathbf{x}, t) = \Re e \left\{ \sum_{j=1}^{\infty} \mathbf{B}_j(\mathbf{x}) e^{-i\omega_j t} \right\} \quad (1.4c)$$

$$\text{and } \mathbf{H}(\mathbf{x}, t) = \Re e \left\{ \sum_{j=1}^{\infty} \mathbf{H}_j(\mathbf{x}) e^{-i\omega_j t} \right\}. \quad (1.4d)$$

Note that the fields in the frequency domain may be complex. Other transformations may be appropriate, especially if the mode spectrum is continuous or if the structure is driven by some source.

1.1.1. Eigenmode Problems

Consider the problem of finding the normal modes of a cavity. Let \mathbf{D} , \mathbf{E} , \mathbf{B} and \mathbf{H} represent the fields of a normal mode and let ω be the frequency of the mode. Combining Maxwell's equations (1.1) and the boundary condition (1.2) yields the following eigenmode problem: given the region Ω and its material properties ϵ and μ , find the eigenmode fields \mathbf{D} , \mathbf{E} , \mathbf{B} and \mathbf{H} , and the corresponding frequencies ω such that

$$\nabla \times \mathbf{E} = \frac{i\omega}{c} \mathbf{B} \quad \text{and} \quad \nabla \times \mathbf{H} = -\frac{i\omega}{c} \mathbf{D} \quad \text{in } \Omega, \quad (1.5a)$$

$$\nabla \cdot \mathbf{D} = 0 \quad \text{and} \quad \nabla \cdot \mathbf{B} = 0 \quad \text{in } \Omega, \quad (1.5b)$$

$$\mathbf{D} = \epsilon \mathbf{E} \quad \text{and} \quad \mathbf{B} = \mu \mathbf{H} \quad \text{in } \Omega \quad (1.5c)$$

$$\text{and} \quad \hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}}. \quad (1.5d)$$

The eigenvalue is ω and the fields \mathbf{D} , \mathbf{E} , \mathbf{B} and \mathbf{H} are time-independent field amplitudes.

An alternative formulation of the eigenmode problem combines the two curl equations (1.5a) and reduces the problem to a single vector field. The formulation using the electric field \mathbf{E} is: given the region Ω and its material properties ϵ and μ , find the eigenmode fields \mathbf{E} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}) = \frac{\omega^2}{c^2} \epsilon \mathbf{E} \quad \text{in } \Omega, \quad (1.6a)$$

$$\nabla \cdot (\epsilon \mathbf{E}) = 0 \quad \text{in } \Omega \quad (1.6b)$$

$$\text{and} \quad \hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}}. \quad (1.6c)$$

The formulation using the magnetic field \mathbf{H} is: given the region Ω and its material properties ϵ and μ , find the eigenmode fields \mathbf{H} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla \times (\epsilon^{-1} \nabla \times \mathbf{H}) = \frac{\omega^2}{c^2} \mu \mathbf{H} \quad \text{in } \Omega, \quad (1.7a)$$

$$\nabla \cdot (\mu \mathbf{H}) = 0 \quad \text{in } \Omega \quad (1.7b)$$

$$\text{and} \quad \hat{\mathbf{n}} \times (\epsilon^{-1} \nabla \times \mathbf{H}) = 0 \quad \text{on } \Gamma_{\text{metal}}. \quad (1.7c)$$

Both of these formulations are called curl-curl formulations.

If the material inside the structure is isotropic and homogeneous then the permittivity and permeability are constant scalars ϵ and μ . For such cases the vector identity $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$ can be used to obtain the wave equation

formulation of the eigenmode problem. Note that in cartesian coordinates, the i th component of $\nabla^2 \mathbf{A}$ is just $\nabla^2 A_i$. However, this is not true in other coordinate systems. Instead, the vector identity above defines $\nabla^2 \mathbf{A}$.

The wave equation formulation for the electric field \mathbf{E} is: given the region Ω and constant scalar material properties ϵ and μ , find the eigenmode fields \mathbf{E} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla^2 \mathbf{E} = -\epsilon\mu \frac{\omega^2}{c^2} \mathbf{E} \quad \text{in } \Omega, \quad (1.8a)$$

$$\nabla \cdot \mathbf{E} = 0 \quad \text{in } \Omega \quad (1.8b)$$

$$\text{and } \hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}}. \quad (1.8c)$$

Similarly, the eigenmode problem for the magnetic field \mathbf{H} is: given the region Ω and constant scalar material properties ϵ and μ , find the eigenmode fields \mathbf{H} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla^2 \mathbf{H} = -\epsilon\mu \frac{\omega^2}{c^2} \mathbf{H} \quad \text{in } \Omega, \quad (1.9a)$$

$$\nabla \cdot \mathbf{H} = 0 \quad \text{in } \Omega \quad (1.9b)$$

$$\text{and } \hat{\mathbf{n}} \times (\nabla \times \mathbf{H}) = 0 \quad \text{on } \Gamma_{\text{metal}}. \quad (1.9c)$$

A pair of boundary conditions equivalent to (1.9c) is $H_n = 0$ and $\partial \mathbf{H}_t / \partial n = 0$ on Γ_{metal} .

1.2. Structure Symmetries

Symmetries of the microwave structure can reduce the size of the problem. Reduced problems take less time and space to solve numerically. Alternatively, a finer mesh can be used with the reduced problem to obtain a more accurate solution. The symmetries described in the following sections can be exploited using appropriate boundary conditions while leaving the formulation for the interior of the structure unchanged. Other symmetries, such as cylindrical symmetry, lead to different formulations for the interior as well as different boundary conditions. Formulations for cylindrical symmetry are discussed in chapters 3 and 4. Further discussion of the use of symmetry to simplify calculations can be found, for example, in [3].

The connection between Maxwell's equations and a symmetry is that Maxwell's equations, including the boundaries and any material properties, are invariant under the symmetry operator. Then fields can be found which are simultaneously

eigenmodes of both Maxwell's equations and the symmetry operator. These eigenmodes have one eigenvalue for Maxwell's equations and a separate eigenvalue for the symmetry operator. In the language of quantum mechanics, Maxwell's equations (thought of as an operator) and the symmetry operator commute, so the operators are simultaneously diagonalizable. The size of the problem is reduced by constraining the solutions of Maxwell's equations to be eigenmodes of the symmetry operator with a particular eigenvalue.

Details of the procedure for obtaining boundary conditions are given below for reflection symmetry about a plane. Sections on other symmetries follow, and these sections focus on the aspects of a symmetry which differ from the previous cases. Example formulations which take advantage of the symmetries will be presented, but without any proof that they are equivalent to the original formulation. However, the variational formulations introduced in later chapters will indicate which boundary conditions are necessary to make the original and reduced formulations of the problem equivalent. All of the examples are based on the curl-curl formulation (1.6) for the electric field \mathbf{E} .

1.2.1. Reflection Symmetry About a Plane

Consider a structure which has reflection symmetry about a plane P , called the symmetry plane. The symmetry plane is defined by the equation $\mathbf{x}^T \hat{\mathbf{p}} = p_o$, where $\hat{\mathbf{p}}$ is a unit vector normal to the plane P and $|p_o|$ is its distance from the origin. An example is shown in figure 1.2a.

Let $\hat{\mathbf{P}}$ be the reflection operator about the symmetry plane P . The reflection operator $\hat{\mathbf{P}}$ can act on various types of objects. $\hat{\mathbf{P}}$ acting on a point \mathbf{x} yields the point

$$\hat{\mathbf{P}}\mathbf{x} = \mathbf{P}(\mathbf{x} - \hat{\mathbf{p}}p_o) + \hat{\mathbf{p}}p_o = (\mathbf{I} - 2\hat{\mathbf{p}}\hat{\mathbf{p}}^T)(\mathbf{x} - \hat{\mathbf{p}}p_o) + \hat{\mathbf{p}}p_o, \quad (1.10)$$

where $\mathbf{P} = (\mathbf{I} - 2\hat{\mathbf{p}}\hat{\mathbf{p}}^T)$ is the reflection matrix and \mathbf{I} is the identity matrix. The reflection operator acting on a vector field \mathbf{E} is the vector field $\hat{\mathbf{P}}\mathbf{E}$. Evaluating the reflected vector field $\hat{\mathbf{P}}\mathbf{E}$ at a point \mathbf{x} gives the vector

$$(\hat{\mathbf{P}}\mathbf{E})(\mathbf{x}) = \mathbf{P} \mathbf{E}(\hat{\mathbf{P}}^{-1}\mathbf{x}). \quad (1.11)$$

Finally, the reflection operator $\hat{\mathbf{P}}$ acting on a pseudovector field \mathbf{H} is $\hat{\mathbf{P}}\mathbf{H}$. The magnetic field \mathbf{H} must be transformed as a pseudovector instead of a vector in order to leave Maxwell's equations unchanged. Evaluating the reflected pseudovector field at the point \mathbf{x} gives the pseudovector

$$(\hat{\mathbf{P}}\mathbf{H})(\mathbf{x}) = -\mathbf{P} \mathbf{H}(\hat{\mathbf{P}}^{-1}\mathbf{x}). \quad (1.12)$$

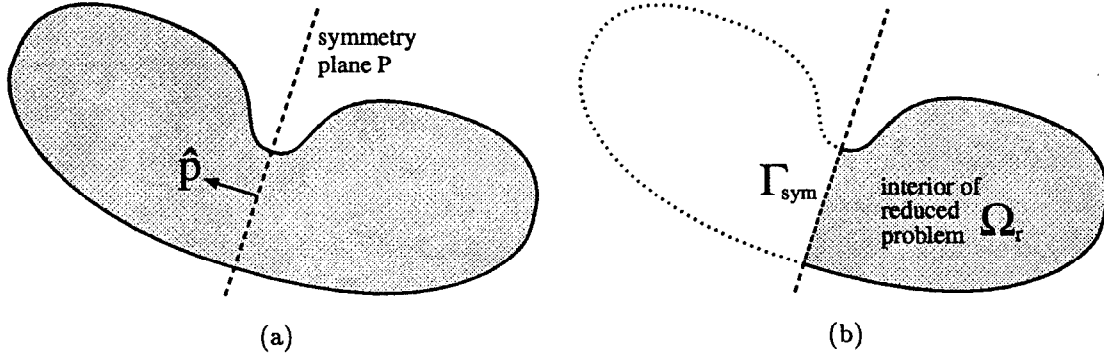


Figure 1.2: (a) A structure with symmetry plane P . The shaded region is the interior Ω , and the normal to the symmetry plane is $\hat{\mathbf{p}}$. (b) The region over which Maxwell's equations must be solved is reduced to Ω_r for a symmetric structure. The symmetry boundary is Γ_{sym} .

The minus sign is present for pseudovectors because the determinant of the reflection matrix \mathbf{P} is -1 .

Consider an eigenmode of $\hat{\mathbf{P}}$ with eigenvalue p . Denoting one of the fields (\mathbf{E} or \mathbf{H} , for example) of the eigenmode by \mathbf{A} , then $\hat{\mathbf{P}}\mathbf{A} = p\mathbf{A}$. Reflecting the mode twice gives the original mode, so

$$\hat{\mathbf{P}}(\hat{\mathbf{P}}\mathbf{A}) = \hat{\mathbf{P}}(p\mathbf{A}) = p^2\mathbf{A} = \mathbf{A}. \quad (1.13)$$

Thus $p^2 = 1$, and the eigenvalues of $\hat{\mathbf{P}}$ are $p = \pm 1$.

Let Γ_{sym} be the portion of the symmetry plane P in the structure interior Ω . The surface Γ_{sym} will be a boundary of the reduced problem. Notice that

$$\hat{\mathbf{P}}\mathbf{x} = \mathbf{x} \quad \text{for } \mathbf{x} \in \Gamma_{\text{sym}}. \quad (1.14)$$

Then

$$\begin{aligned} (\hat{\mathbf{P}}\mathbf{E})(\mathbf{x}) &= \mathbf{P}\mathbf{E}(\hat{\mathbf{P}}^{-1}\mathbf{x}) = \mathbf{P}\mathbf{E}(\mathbf{x}) = p\mathbf{E}(\mathbf{x}) \\ (\hat{\mathbf{P}}\mathbf{H})(\mathbf{x}) &= -\mathbf{P}\mathbf{H}(\hat{\mathbf{P}}^{-1}\mathbf{x}) = -\mathbf{P}\mathbf{H}(\mathbf{x}) = p\mathbf{H}(\mathbf{x}) \end{aligned} \quad \text{for } \mathbf{x} \in \Gamma_{\text{sym}} \quad (1.15)$$

and thus

$$\mathbf{P}\mathbf{E} = p\mathbf{E} \quad \text{and} \quad -\mathbf{P}\mathbf{H} = p\mathbf{H} \quad \text{on } \Gamma_{\text{sym}}. \quad (1.16)$$

Consider the action of the reflection matrix \mathbf{P} on a vector \mathbf{A} . Let $A_n = \hat{\mathbf{p}} \cdot \mathbf{A}$ be the component of \mathbf{A} normal to the symmetry plane P , and let $\mathbf{A}_t = (\mathbf{I} - \hat{\mathbf{p}}\hat{\mathbf{p}}^T)\mathbf{A}$ be the portion of \mathbf{A} tangential to the symmetry plane P . Then

$$(\mathbf{P}\mathbf{A})_t = \mathbf{A}_t \quad \text{and} \quad (\mathbf{P}\mathbf{A})_n = -A_n. \quad (1.17)$$

That is, \mathbf{P} reverses the normal component of the vector, but leaves the tangential component of the vector unchanged. Applying this to equation (1.16) yields

$$\mathbf{E}_t = p\mathbf{E}_t, \quad E_n = -pE_n, \quad \mathbf{H}_t = -p\mathbf{H}_t \quad \text{and} \quad H_n = pH_n \quad \text{on } \Gamma_{\text{sym}}. \quad (1.18)$$

The case $p = -1$ corresponds to a perfectly conducting boundary with boundary conditions

$$\text{perfectly conducting:} \quad \mathbf{E}_t = 0 \quad \text{and} \quad B_n = 0 \quad \text{on } \Gamma_{\text{sym}}, \quad (1.19)$$

and the case $p = 1$ corresponds to a perfectly insulating boundary with boundary conditions

$$\text{perfectly insulating:} \quad D_n = 0 \quad \text{and} \quad \mathbf{H}_t = 0 \quad \text{on } \Gamma_{\text{sym}}. \quad (1.20)$$

The conditions on the component of the field normal to Γ_{sym} are applied to \mathbf{D} and \mathbf{B} since D_n and B_n are continuous across any interface. The fields \mathbf{E}_t and \mathbf{H}_t are continuous across the symmetry plane, but the fields E_n and H_n are not necessarily continuous. Like the boundary condition at perfectly conducting walls, the two boundary conditions at the symmetry plane are not independent. For example, given a solution to Maxwell's equations and the boundary condition $\mathbf{H}_t = 0$, then the other boundary condition $D_n = 0$ can be derived if time-varying fields are assumed.

Many field solvers call the perfectly conducting case a metal boundary condition and reserve the words symmetry boundary to mean only the perfectly insulating case. While this is reasonable for calculating the fields in an RF structure, some post-processing calculations, such as power loss due to the finite conductivity of the metal walls, need to distinguish between a real metal wall and the perfectly conducting case of a symmetry plane.

Here is a formulation for the eigenmode problem in a symmetric structure. The problem is reduced to a region Ω_r which is half of the original structure, as depicted in figure 1.2b. The formulation for the perfectly conducting ($p = -1$) case is: given the region Ω_r and its material properties ϵ and μ , find the eigenmode fields \mathbf{E} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}) = \frac{\omega^2}{c^2} \epsilon \mathbf{E} \quad \text{in } \Omega_r, \quad (1.21a)$$

$$\nabla \cdot (\epsilon \mathbf{E}) = 0 \quad \text{in } \Omega_r \quad (1.21b)$$

$$\text{and} \quad \hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym}}. \quad (1.21c)$$

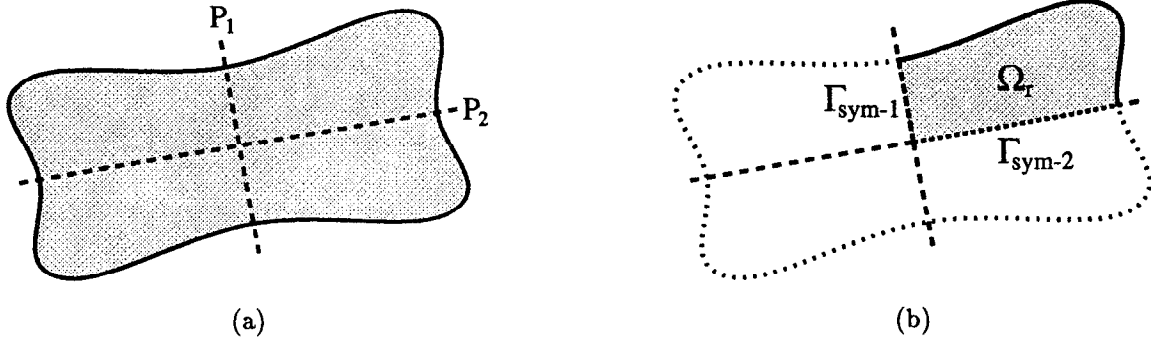


Figure 1.3: (a) A structure with two symmetry planes, P_1 and P_2 . (b) The region over which Maxwell's equations must be solved is reduced to Ω_r . The symmetry boundaries are $\Gamma_{\text{sym-1}}$ and $\Gamma_{\text{sym-2}}$.

The formulation for the perfectly insulating ($p = 1$) case is: given the region Ω_r and its material properties ϵ and μ , find the eigenmode fields \mathbf{E} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}) = \frac{\omega^2}{c^2} \epsilon \mathbf{E} \quad \text{in } \Omega_r, \quad (1.22a)$$

$$\nabla \cdot (\epsilon \mathbf{E}) = 0 \quad \text{in } \Omega_r, \quad (1.22b)$$

$$\hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}} \quad (1.22c)$$

$$\text{and } \hat{\mathbf{n}} \times (\mu^{-1} (\nabla \times \mathbf{E})) = 0 \quad \text{on } \Gamma_{\text{sym}}. \quad (1.22d)$$

For a non-static mode, the boundary condition (1.22d) is the same as $\hat{\mathbf{n}} \times \mathbf{H} = 0$ on Γ_{sym} .

Some structures have more than one symmetry plane, in which case there is a symmetry operator for each symmetry plane. If the symmetry planes are perpendicular to each other then the symmetry operators commute and modes can be found which are simultaneously eigenmodes for all of the symmetry operators and of Maxwell's equations. An example with two symmetry planes is shown in figure 1.3. Let $\Gamma_{\text{sym}(\text{cond})}$ be the symmetry planes for which the perfectly conducting ($p = 1$) case is chosen, and let $\Gamma_{\text{sym}(\text{ins})}$ be the symmetry planes for which the perfectly insulating ($p = -1$) case is chosen. Then a formulation for the eigenmode problem is: given the region Ω_r and its material properties ϵ and μ , find the eigenmode fields \mathbf{E} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}) = \frac{\omega^2}{c^2} \epsilon \mathbf{E} \quad \text{in } \Omega_r, \quad (1.23a)$$

$$\nabla \cdot (\epsilon \mathbf{E}) = 0 \quad \text{in } \Omega_r, \quad (1.23b)$$

$$\hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym}(\text{cond})} \quad (1.23c)$$

$$\text{and } \hat{\mathbf{n}} \times (\mu^{-1} (\nabla \times \mathbf{E})) = 0 \quad \text{on } \Gamma_{\text{sym}(\text{ins})}. \quad (1.23d)$$

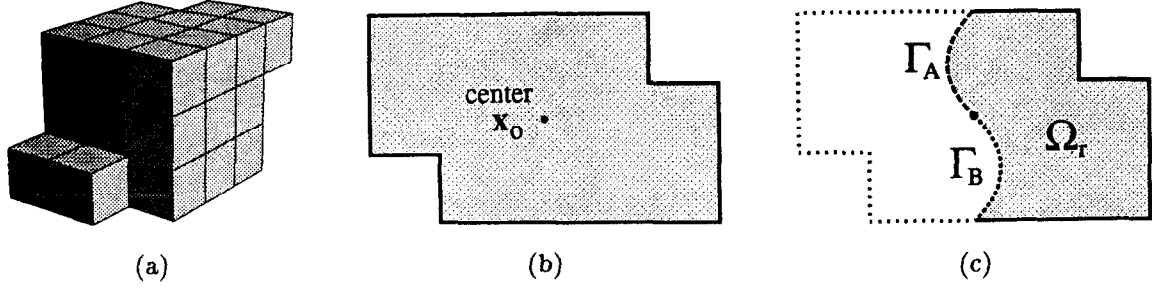


Figure 1.4: (a) A structure with reflection symmetry about a point. (b) A 2d structure which is symmetric about its center \mathbf{x}_o . The shaded region is the interior Ω . (c) The region over which Maxwell's equations must be solved is reduced to Ω_r for a symmetric structure. The symmetry boundaries are Γ_A and Γ_B .

Reflection symmetry about a plane may be combined with another type of symmetry, for example periodic symmetry. The guiding rule is that both symmetries can be used as long as the corresponding symmetry operators commute.

1.2.2. Inversion Symmetry

Consider a structure with reflection symmetry about a point \mathbf{x}_o , the center of the structure. This is called inversion symmetry. An example of such a structure is shown in figure 1.4. Let $\hat{\mathbf{P}}$ be the reflection operator about the center. The reflection operator acting on a point \mathbf{x} gives the point

$$\hat{\mathbf{P}}\mathbf{x} = -\mathbf{I}(\mathbf{x} - \mathbf{x}_o) + \mathbf{x}_o = -\mathbf{x} + 2\mathbf{x}_o. \quad (1.24)$$

The reflection operator acting on a vector field \mathbf{E} and a pseudovector field \mathbf{H} gives

$$(\hat{\mathbf{P}}\mathbf{E})(\mathbf{x}) = -\mathbf{E}(\hat{\mathbf{P}}^{-1}\mathbf{x}) \quad \text{and} \quad (\hat{\mathbf{P}}\mathbf{H})(\mathbf{x}) = \mathbf{H}(\hat{\mathbf{P}}^{-1}\mathbf{x}). \quad (1.25)$$

Let Ω_r be half of the interior Ω of the structure, and let the symmetry boundaries Γ_A , Γ_B and possibly the center \mathbf{x}_o be the portion of the boundary of Ω_r which is in Ω . The boundaries Γ_A and Γ_B are chosen such that the symmetry operator $\hat{\mathbf{P}}$ maps Γ_A to Γ_B and vice versa. An example is shown in figure 1.4c. Note that the boundaries Γ_A and Γ_B are not required to be planes.

As in the case of reflection symmetry about a plane, reflecting an eigenmode of $\hat{\mathbf{P}}$ twice gives the original mode (see equation (1.13)), so the eigenvalues of $\hat{\mathbf{P}}$ are $p = \pm 1$. The boundary conditions for the boundaries Γ_A and Γ_B are just the eigenvalue equation applied at the boundaries,

$$(\hat{\mathbf{P}}\mathbf{E})(\mathbf{x}) = -\mathbf{E}(\hat{\mathbf{P}}^{-1}\mathbf{x}) = p\mathbf{E}(\mathbf{x}) \quad \text{and} \quad (\hat{\mathbf{P}}\mathbf{H})(\mathbf{x}) = \mathbf{H}(\hat{\mathbf{P}}^{-1}\mathbf{x}) = p\mathbf{H}(\mathbf{x}). \quad (1.26)$$

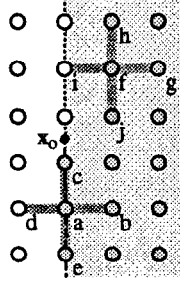


Figure 1.5: An example of a finite difference mesh exploiting symmetry about a center \mathbf{x}_o .

A special case occurs at the center \mathbf{x}_o since $\hat{\mathbf{P}}\mathbf{x}_o = \mathbf{x}_o$. At the center $-\mathbf{E}(\mathbf{x}_o) = p\mathbf{E}(\mathbf{x}_o)$ and $\mathbf{H}(\mathbf{x}_o) = p\mathbf{H}(\mathbf{x}_o)$, so

$$\mathbf{E}(\mathbf{x}_o) = 0 \quad \text{if } p = 1, \quad (1.27a)$$

$$\mathbf{H}(\mathbf{x}_o) = 0 \quad \text{if } p = -1. \quad (1.27b)$$

A formulation for the $p = \pm 1$ case is: given the region $\Omega_{\mathbf{r}}$ and its material properties ϵ and μ , find the eigenmode fields \mathbf{E} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}) = \frac{\omega^2}{c^2} \epsilon \mathbf{E} \quad \text{in } \Omega, \quad (1.28a)$$

$$\nabla \cdot (\epsilon \mathbf{E}) = 0 \quad \text{in } \Omega, \quad (1.28b)$$

$$\hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}}, \quad (1.28c)$$

$$\hat{\mathbf{n}} \times \mathbf{E} \Big|_{\mathbf{x}} = -p \hat{\mathbf{n}} \times \mathbf{E} \Big|_{\hat{\mathbf{P}}\mathbf{x}} \quad \text{for } \mathbf{x} \in \Gamma_A \quad (1.28d)$$

$$\text{and } \hat{\mathbf{n}} \times (\mu^{-1} (\nabla \times \mathbf{E})) \Big|_{\mathbf{x}} = p \hat{\mathbf{n}} \times (\mu^{-1} (\nabla \times \mathbf{E})) \Big|_{\hat{\mathbf{P}}\mathbf{x}} \quad \text{for } \mathbf{x} \in \Gamma_A. \quad (1.28e)$$

A digression on implementation will help illustrate the non-local nature of this boundary condition. To exploit this symmetry in a field solver, the fields at points on one boundary, say $\mathbf{x} \in \Gamma_A$, are constrained to be the same as the fields at the corresponding points on the other boundary, $\hat{\mathbf{P}}\mathbf{x} \in \Gamma_B$, multiplied by a factor $\pm p$. This is easily accomplished if the mesh, or discretization, of the problem has the same inversion symmetry as the structure. Numerous examples of the finite element method are presented in later chapters, so consider a finite difference example here. A portion of a finite difference mesh is shown in figure 1.5. The shaded region is in $\Omega_{\mathbf{r}}$ and the dotted and dashed lines are the boundaries Γ_A and Γ_B , respectively. The fields are computed for nodes (filled circles) in the region $\Omega_{\mathbf{r}}$ and on the boundary Γ_B . Whenever a field at a point \mathbf{x} outside of $\Omega_{\mathbf{r}}$ or Γ_B is required by the finite difference operator, the field at the point $\hat{\mathbf{P}}\mathbf{x}$ (which is in $\Omega_{\mathbf{r}}$ or Γ_B) multiplied by $\pm p$ is used

instead. For example, dark gray lines indicate the field values necessary for a 5-point finite difference operator at nodes a and f in figure 1.5. The field at node d is the field at node f times $\pm p$, while the field at point i is the field at node a times $\pm p$. Note that the symmetry of the finite difference matrix \mathbf{M} is preserved since $M_{af} = M_{fa}$.

1.2.3. Periodic Structures

A periodic structure has a symmetry operator $\hat{\mathbf{R}}$ which rigidly moves the structure by one period. The symmetry operator $\hat{\mathbf{R}}$ acting on a point \mathbf{x} can be written generally as

$$\hat{\mathbf{R}}\mathbf{x} = \mathbf{R}\mathbf{x} + \mathbf{x}_o, \quad (1.29)$$

where the matrix \mathbf{R} is an orthogonal matrix. That is, $\mathbf{R}^T\mathbf{R} = \mathbf{I}$. A common symmetry operation for periodic structures is translation by a cell length l along an axis, say $\hat{\mathbf{z}}$. In this case $\mathbf{x}_o = l\hat{\mathbf{z}}$ and $\mathbf{R} = \mathbf{I}$. Thus $\hat{\mathbf{R}}\mathbf{x} = \mathbf{x} + l\hat{\mathbf{z}}$. However, as indicated by the examples in figure 1.6, the general form of the operator $\hat{\mathbf{R}}$ allows more than translations. The symmetry operation can include a rotation as shown in figure 1.6b, and it can include a reflection as shown in figure 1.6c. A combination of translation and rotation describes helical structures, such as the example shown in figure 1.6d. These are all periodic structures and, as will be shown below, the modes of a periodic structure can be found by modelling a single period of the structure and using a boundary condition called the quasi-periodic boundary condition.

Periodicity differs from the previous symmetry operators in that $\hat{\mathbf{R}}^2$ is not the identity operator, so the eigenvalues of the symmetry operator are not simply ± 1 . A generalization of Floquet's theorem gives the eigenvalues for the symmetry operator $\hat{\mathbf{R}}$ of a periodic structure.

Let \mathbf{A} represent one of the fields, perhaps \mathbf{E} or \mathbf{H} , of a mode which is a solution to the eigenmode problem in the periodic structure. The field \mathbf{A} is complex and represents a wave propagating along the structure. The real field $\mathbf{A}(\mathbf{x}, t)$ can be obtained from the complex field $\mathbf{A}(\mathbf{x})$ using

$$\mathbf{A}(\mathbf{x}, t) = \Re \{ \mathbf{A}(\mathbf{x}) e^{-i\omega t} \}. \quad (1.30)$$

The eigenvalues of the symmetry operator $\hat{\mathbf{R}}$ are complex numbers α . The eigenvalue relates the complex field in one cell with the complex field in the following cell. If the eigenvalue has $|\alpha| > 1$ or $|\alpha| < 1$ then the mode is geometrically growing or damping along the structure. Formulations for these evanescent modes are typically

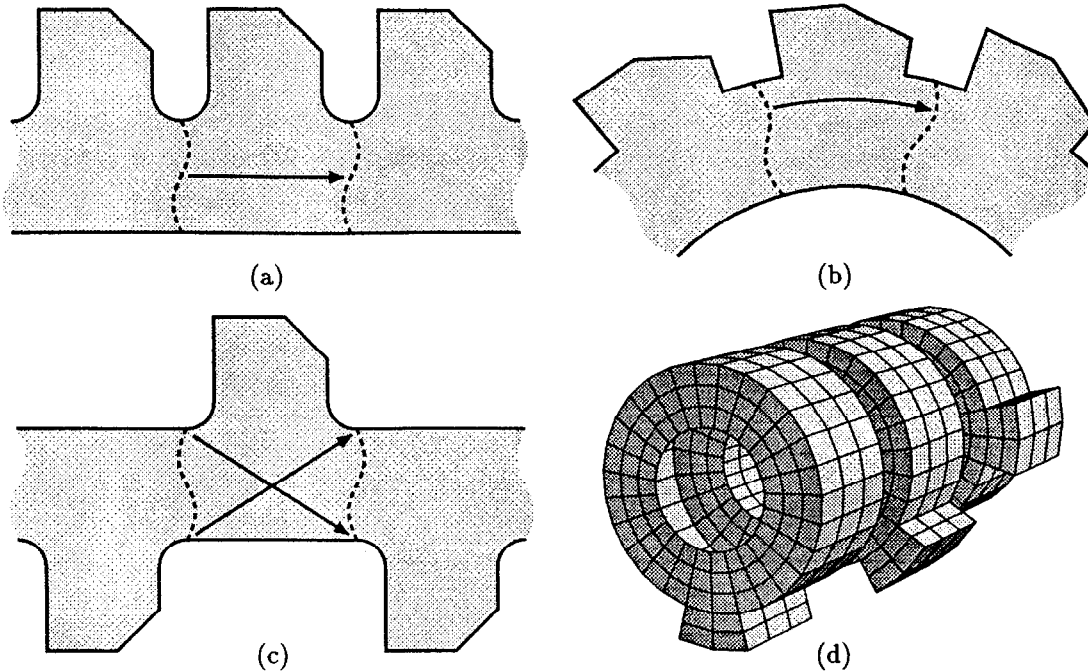


Figure 1.6: Some examples of periodic structures. The dashed lines delimit one period, or cell, of the structure. The arrows indicate the action of the symmetry operations: (a) translation, (b) rotation and (c) glide reflection. The symmetry operation for a helical structure (d) is a combination of translation and rotation.

non-symmetric or not positive semi-definite. This work focuses on symmetric formulations with real eigenvalues, so evanescent modes will not be considered further here.

If $|\alpha| = 1$ then the eigenvalue is $\alpha = e^{-i\psi}$ for some real phase advance ψ . These modes are travelling waves in the periodic structure, and are the modes of interest in this work.

To exploit this symmetry the solutions are restricted to be eigenmodes of the symmetry operator $\hat{\mathbf{R}}$ with a particular eigenvalue $e^{-i\psi}$. In other words, a phase advance ψ is selected and the vector field \mathbf{E} must satisfy

$$\mathbf{E}(\hat{\mathbf{R}}\mathbf{x}) = \mathbf{R}\mathbf{E}(\mathbf{x})e^{i\psi} \quad (1.31)$$

and the pseudovector field \mathbf{H} must satisfy

$$\mathbf{H}(\hat{\mathbf{R}}\mathbf{x}) = \pm\mathbf{R}\mathbf{H}(\mathbf{x})e^{i\psi}, \quad (1.32)$$

where the sign is the determinant of the matrix \mathbf{R} . The sign is positive for all of the examples of figure 1.6 except for the glide reflection example in figure 1.6c. For glide reflection the sign is negative.

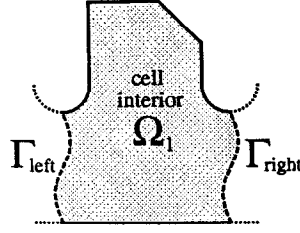


Figure 1.7: One cell of the periodic structure of figure 1.6a. Γ_{left} and Γ_{right} are the quasi-periodic boundaries.

Let the region Ω_1 be the interior of one period, or cell, of the periodic structure. The portion of the cell boundary in the interior Ω comprises the quasi-periodic boundaries Γ_{left} and Γ_{right} . The symmetry operator $\hat{\mathbf{R}}$ acting on the boundary Γ_{left} is the boundary Γ_{right} . An example is shown in figure 1.7. Notice that the normal vectors at the quasi-periodic boundaries are related by

$$\hat{\mathbf{n}}|_{\hat{\mathbf{R}}\mathbf{x}} = -\mathbf{R}\hat{\mathbf{n}}|_{\mathbf{x}} \quad \forall \mathbf{x} \in \Gamma_{\text{left}}. \quad (1.33)$$

There is no unique choice for the cell and its boundaries. The boundaries Γ_{left} and Γ_{right} are usually chosen to be planes, but they are allowed to be curved surfaces as shown in the example.

Here is a formulation for the eigenmode problem reduced to one cell Ω_1 of the periodic structure. Given the region Ω_1 , its material properties ϵ and μ and a phase advance ψ , find the complex eigenmode fields \mathbf{E} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}) = \frac{\omega^2}{c^2} \epsilon \mathbf{E} \quad \text{in } \Omega_1, \quad (1.34a)$$

$$\nabla \cdot (\epsilon \mathbf{E}) = 0 \quad \text{in } \Omega_1, \quad (1.34b)$$

$$\hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}}, \quad (1.34c)$$

$$\hat{\mathbf{n}} \times \mathbf{E}|_{\hat{\mathbf{R}}\mathbf{x}} = \mp \mathbf{R} \hat{\mathbf{n}} \times \mathbf{E}|_{\mathbf{x}} e^{i\psi} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left}}, \quad (1.34d)$$

$$\text{and } \hat{\mathbf{n}} \times (\mu^{-1} (\nabla \times \mathbf{E}))|_{\hat{\mathbf{R}}\mathbf{x}} = -\mathbf{R} \hat{\mathbf{n}} \times (\mu^{-1} (\nabla \times \mathbf{E}))|_{\mathbf{x}} e^{i\psi} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left}}. \quad (1.34e)$$

The sign in (1.34d) is $-|\mathbf{R}|$. The last two conditions are the quasi-periodic boundary conditions for this formulation. In words, the portion of the fields \mathbf{E} and \mathbf{H} tangential to Γ_{right} are the same as the corresponding tangential fields on Γ_{left} multiplied by the phase factor $e^{i\psi}$. Note that condition (1.34e) combined with the curl-curl equation (1.34a) implies

$$\hat{\mathbf{n}} \cdot \mathbf{D}|_{\hat{\mathbf{R}}\mathbf{x}} = -\hat{\mathbf{n}} \cdot \mathbf{D}|_{\mathbf{x}} e^{i\psi} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left}}. \quad (1.35)$$

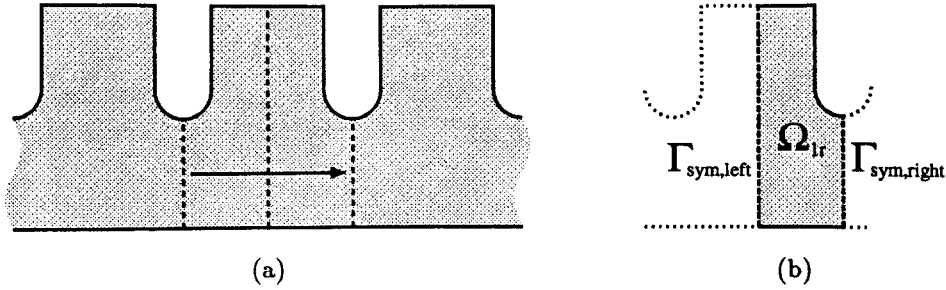


Figure 1.8: (a) A symmetric periodic structure. The action of operator $\hat{\mathbf{R}}$ is indicated by the arrow. The dashed lines are symmetry planes. (b) The region over which Maxwell's equations must be solved is reduced to Ω_{1r} . The symmetry boundaries are $\Gamma_{\text{sym-left}}$ and $\Gamma_{\text{sym-right}}$.

A lattice-like structure with two or three directions of periodicity can be handled similarly. There will be two or three independent rigid motions $\hat{\mathbf{R}}_i$. For each operator $\hat{\mathbf{R}}_i$ there is a pair of quasi-periodic boundaries and a phase advance. The quasi-periodic boundaries are $\Gamma_{\text{left-}i}$ and $\Gamma_{\text{right-}i}$ and the phase advance is ψ_i .

1.2.4. Symmetric Periodic Structures

Now consider a structure with two non-commuting symmetry planes. In other words, the structure is both periodic and symmetric. An example is shown in figure 1.8. Let $\hat{\mathbf{P}}_L$ be the reflection operator about a symmetry plane $\Gamma_{\text{sym-left}}$ (see equation (1.10)), and let $\hat{\mathbf{P}}_R$ be the reflection operator about a symmetry plane $\Gamma_{\text{sym-right}}$. The subscripts refer to the left and right sides of the half-cell Ω_{1r} . The rigid motion operator $\hat{\mathbf{R}} = \hat{\mathbf{P}}_R \hat{\mathbf{P}}_L$ moves the structure or field right one period (see equation (1.29)).

If the two symmetry planes are parallel, then the rigid motion $\hat{\mathbf{R}}$ is a translation. Otherwise, the two planes must intersect, and $\hat{\mathbf{R}}$ is a rotation about the line of intersection. Note that the matrix \mathbf{R} has determinant $|\mathbf{R}| = 1$, so this rigid motion is never a reflection.

According to Floquet's theorem, the fields can be decomposed into modes with phase advance ψ . Following a technique described in [4], consider the electric field \mathbf{E} of a mode with phase advance ψ satisfying $\mathbf{E}(\hat{\mathbf{R}}\mathbf{x}) = \mathbf{R}\mathbf{E}(\mathbf{x})e^{i\psi}$. The complex conjugate of the mode has the opposite phase advance,

$$\mathbf{E}^*(\hat{\mathbf{R}}\mathbf{x}) = (\mathbf{R}\mathbf{E}(\mathbf{x})e^{i\psi})^* = \mathbf{R}\mathbf{E}^*(\mathbf{x})e^{-i\psi}. \quad (1.36)$$

$\hat{\mathbf{P}}_L\mathbf{E}$ is the mode reflected about the symmetry plane $\Gamma_{\text{sym-left}}$, and it also has the

opposite phase advance,

$$(\hat{\mathbf{P}}_L \mathbf{E})(\hat{\mathbf{R}}\mathbf{x}) = \mathbf{R}(\hat{\mathbf{P}}_L \mathbf{E})(\mathbf{x})e^{-i\psi}. \quad (1.37)$$

Assuming \mathbf{E} is a non-degenerate mode, equations (1.36) and (1.37) indicate that \mathbf{E}^* and $\hat{\mathbf{P}}_L \mathbf{E}$ are the same mode. Let

$$\mathbf{E}^* = \alpha \hat{\mathbf{P}}_L \mathbf{E} \quad (1.38)$$

for some complex number α . Conjugating equation (1.38) and substituting for \mathbf{E}^* with equation (1.38) gives

$$\mathbf{E} = \alpha^* \hat{\mathbf{P}}_L \mathbf{E}^* = \alpha^* \hat{\mathbf{P}}_L (\alpha \hat{\mathbf{P}}_L \mathbf{E}) = \alpha^* \alpha \hat{\mathbf{P}}_L^2 \mathbf{E} = |\alpha|^2 \mathbf{E} \quad (1.39)$$

which implies $|\alpha| = 1$. Without loss of generality choose $\alpha = 1$. A different α would just multiply the mode by an overall phase factor.

At the symmetry plane $\Gamma_{\text{sym-left}}$, $\hat{\mathbf{P}}_L \mathbf{x} = \mathbf{x}$ so

$$\mathbf{E}^*(\mathbf{x}) = (\hat{\mathbf{P}}_L \mathbf{E})(\mathbf{x}) = \mathbf{P}_L \mathbf{E}(\hat{\mathbf{P}}_L \mathbf{x}) = \mathbf{P}_L \mathbf{E}(\mathbf{x}) \quad \forall \mathbf{x} \in \Gamma_{\text{sym-left}}. \quad (1.40)$$

Let E_n be the component of \mathbf{E} normal to $\Gamma_{\text{sym-left}}$ and let \mathbf{E}_t be the portion of \mathbf{E} tangential to $\Gamma_{\text{sym-left}}$. Then the above conditions are $E_n^* = -E_n$ and $\mathbf{E}_t^* = \mathbf{E}_t$, or

$$\Re E_n = 0 \quad \text{and} \quad \Im \mathbf{E}_t = 0 \quad \forall \mathbf{x} \in \Gamma_{\text{sym-left}}. \quad (1.41)$$

In other words, E_n is imaginary and \mathbf{E}_t is real on $\Gamma_{\text{sym-left}}$.

At the other symmetry plane $\Gamma_{\text{sym-right}}$, $\mathbf{x} = \hat{\mathbf{P}}_R \mathbf{x} = \hat{\mathbf{R}} \hat{\mathbf{P}}_L \mathbf{x}$. Replacing \mathbf{x} with $\hat{\mathbf{P}}_L \mathbf{x}$ in equation (1.36) and using equation (1.38) gives

$$\begin{aligned} \mathbf{E}^*(\mathbf{x}) &= \mathbf{E}^*(\hat{\mathbf{R}} \hat{\mathbf{P}}_L \mathbf{x}) = \mathbf{R} \mathbf{E}^*(\hat{\mathbf{P}}_L \mathbf{x}) e^{-i\psi} = \mathbf{R}(\hat{\mathbf{P}}_L \mathbf{E})(\hat{\mathbf{P}}_L \mathbf{x}) e^{-i\psi} \\ &= \mathbf{R} \mathbf{P}_L \mathbf{E}(\hat{\mathbf{P}}_L^2 \mathbf{x}) e^{-i\psi} = \mathbf{R} \mathbf{P}_L \mathbf{E}(\mathbf{x}) e^{-i\psi} \quad \forall \mathbf{x} \in \Gamma_{\text{sym-right}}. \end{aligned} \quad (1.42)$$

$\mathbf{R} \mathbf{P}_L$ is the reflection matrix \mathbf{P}_R about the symmetry plane $\Gamma_{\text{sym-right}}$, so let E_n be the component of \mathbf{E} normal to $\Gamma_{\text{sym-right}}$ and \mathbf{E}_t be the vector tangential to $\Gamma_{\text{sym-right}}$. Then the above conditions are $E_n^* = -E_n e^{-i\psi}$ and $\mathbf{E}_t^* = \mathbf{E}_t e^{-i\psi}$, or

$$\Re \{E_n e^{-i\psi/2}\} = 0 \quad \text{and} \quad \Im \{\mathbf{E}_t e^{-i\psi/2}\} = 0 \quad \forall \mathbf{x} \in \Gamma_{\text{sym-right}}. \quad (1.43)$$

In other words, $E_n \propto i e^{i\psi/2}$ and $\mathbf{E}_t \propto e^{i\psi/2}$ on $\Gamma_{\text{sym-right}}$.

The eigenmode problem reduced to one half cell Ω_{1r} is: given the region Ω_{1r} , material properties ϵ and μ and the phase advance ψ , find the eigenmode fields \mathbf{E} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}) = \frac{\omega^2}{c^2} \epsilon \mathbf{E} \quad \text{in } \Omega_{1r}, \quad (1.44a)$$

$$\nabla \cdot (\epsilon \mathbf{E}) = 0 \quad \text{in } \Omega_{1r}, \quad (1.44b)$$

$$\hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}}, \quad (1.44c)$$

$$\Im m \{ \hat{\mathbf{n}} \times \mathbf{E} \} = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (1.44d)$$

$$\Re e \{ \hat{\mathbf{n}} \times (\mu^{-1} \nabla \times \mathbf{E}) \} = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (1.44e)$$

$$\Im m \{ \hat{\mathbf{n}} \times \mathbf{E} e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}} \quad (1.44f)$$

$$\text{and } \Re e \{ \hat{\mathbf{n}} \times (\mu^{-1} \nabla \times \mathbf{E}) e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}}. \quad (1.44g)$$

Chapter 2

Planar Structures and an Introduction to the Finite Element Method

A simple microwave structure to solve is a cylindrical waveguide containing uniform isotropic media. The cross-section $\tilde{\Omega}$ of the waveguide may be an arbitrary shape as shown in figure 2.1. The cross-section must be constant over the length of the waveguide and the material properties ϵ and μ are assumed to be scalar constants in the waveguide.

In the first section below Maxwell's equations for the TE and TM modes are reduced to a two dimensional scalar eigenmode problem. The second section introduces an equivalent variational formulation of the eigenmode problem, and in the third section the problem is further reduced to a finite element formulation which can be solved numerically on a computer. The final section discusses the deleterious effects of sharp corners and a technique to improve the calculations.

More details on the finite element method can be found, for example, in [5]–[7].

2.1. Maxwell's Equations

Maxwell's equations in a cylindrical waveguide containing uniform isotropic media are discussed, for example, in [1] (ch. 8). Separation of variables leads to fields with an assumed z dependence $e^{ik_z z}$, where z is the coordinate along the length of the waveguide. The modes of a cylindrical waveguide split naturally into three types: TM, TE and TEM modes. The TEM modes will not be considered here. The fields of a TM mode can be derived from the single field component E_z , while the fields of a TE mode can be obtained from H_z alone. For TM modes, $H_z = 0$, while for TE modes, $E_z = 0$. The z -component of the wave equation (1.8) can be written as

$$\nabla_t^2 E_z = (k_z^2 - \epsilon\mu \frac{\omega^2}{c^2}) E_z = -\gamma^2 E_z \quad \text{in } \tilde{\Omega} \quad (2.1)$$

where $\nabla_t^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$. The z dependence $e^{ik_z z}$ has been factored out, so the field $E_z(x, y)$ is independent of z . The eigenvalue γ^2 corresponds to the mode's cutoff frequency. Equation (2.1) along with appropriate boundary conditions constitutes an eigenmode problem for the TM modes. A similar equation for the magnetic field H_z is the foundation for the corresponding TE eigenmode problem.

The boundary condition $\hat{\mathbf{n}} \times \mathbf{E} = 0$ on a metal wall becomes $E_z = 0$. The corresponding boundary condition (1.9c) for the magnetic field becomes $\hat{\mathbf{n}} \cdot \nabla H_z = \partial H_z / \partial n = 0$. Boundary conditions for the symmetries described in section 1.2 can

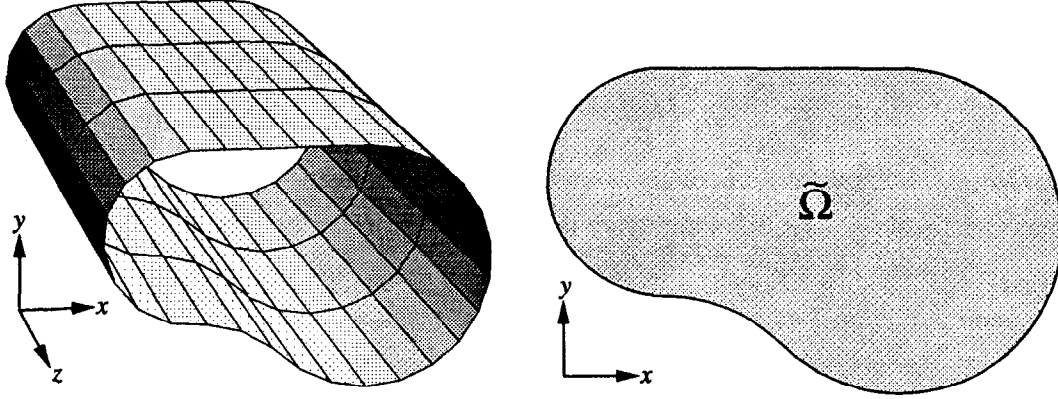


Figure 2.1: A cylindrical waveguide with arbitrary cross-section. The cross-section is shown at right.

be deduced from the corresponding 3d problem or they can be derived ab initio for the 2d problem. The reduction of the quasi-periodic boundary conditions from 3d to 2d is illustrated here. The component of $\hat{\mathbf{n}} \times \mathbf{E}$ in the x-y plane,

$$(\hat{\mathbf{n}} \times \mathbf{E})_t = \hat{x} n_y E_z - \hat{y} n_x E_z = (\hat{\mathbf{n}} \times \hat{\mathbf{z}}) E_z, \quad (2.2)$$

is useful since $\hat{\mathbf{n}} \times \hat{\mathbf{z}}$ can be factored out of the vector boundary conditions on \mathbf{E} to obtain scalar boundary conditions on E_z . Recall relation (1.33) between the normal vectors on the quasi-periodic boundaries, $\hat{\mathbf{n}}|_{\hat{\mathbf{R}}\mathbf{x}} = -\mathbf{R} \hat{\mathbf{n}}|_{\mathbf{x}}$ for $\mathbf{x} \in \Gamma_{\text{left}}$. The symmetry operation does not involve z so $(\mathbf{R}\mathbf{E})_z = E_z$. Using the identity

$$\mathbf{R}^T((\mathbf{R}\hat{\mathbf{n}}) \times \mathbf{E}) = |\mathbf{R}| \hat{\mathbf{n}} \times (\mathbf{R}\mathbf{E}), \quad (2.3)$$

the quasi-periodic boundary condition (1.34d)

$$\hat{\mathbf{n}} \times \mathbf{E}|_{\hat{\mathbf{R}}\mathbf{x}} = (-\mathbf{R} \hat{\mathbf{n}}|_{\mathbf{x}}) \times \mathbf{E}|_{\hat{\mathbf{R}}\mathbf{x}} = \mp \mathbf{R} \hat{\mathbf{n}} \times \mathbf{E}|_{\mathbf{x}} e^{i\psi} \quad (2.4)$$

multiplied by \mathbf{R}^T becomes

$$\hat{\mathbf{n}}|_{\mathbf{x}} \times (\mathbf{R}\mathbf{E})|_{\hat{\mathbf{R}}\mathbf{x}} = \hat{\mathbf{n}} \times \mathbf{E}|_{\mathbf{x}} e^{i\psi} \Rightarrow E_z|_{\hat{\mathbf{R}}\mathbf{x}} = E_z|_{\mathbf{x}} e^{i\psi}. \quad (2.5)$$

For the other quasi-periodic boundary condition use the z -component of the relation $\hat{\mathbf{n}} \times (\nabla \times \mathbf{E}) = \nabla(\hat{\mathbf{n}} \cdot \mathbf{E}) - (\hat{\mathbf{n}} \cdot \nabla)\mathbf{E}$, where the differentiation acts only on \mathbf{E} and not $\hat{\mathbf{n}}$. The quasi-periodic boundary condition (1.34e)

$$\hat{\mathbf{n}} \times (\nabla \times \mathbf{E})|_{\hat{\mathbf{R}}\mathbf{x}} = -\mathbf{R} \hat{\mathbf{n}} \times (\nabla \times \mathbf{E})|_{\mathbf{x}} e^{i\psi} \quad (2.6)$$

becomes

$$\left(\frac{\partial}{\partial z} (\hat{\mathbf{n}} \cdot \mathbf{E}) - (\hat{\mathbf{n}} \cdot \nabla) E_z \right)|_{\hat{\mathbf{R}}\mathbf{x}} = - \left(\frac{\partial}{\partial z} (\hat{\mathbf{n}} \cdot \mathbf{E}) - (\hat{\mathbf{n}} \cdot \nabla) E_z \right)|_{\mathbf{x}} e^{i\psi}. \quad (2.7)$$

Relation (1.35) between the fields normal to the two quasi-periodic boundaries eliminates the $\hat{\mathbf{n}} \cdot \mathbf{E}$ terms, so the quasi-periodic boundary condition becomes

$$(\hat{\mathbf{n}} \cdot \nabla) E_z \Big|_{\hat{\mathbf{R}}_x} = -(\hat{\mathbf{n}} \cdot \nabla) E_z \Big|_x e^{i\psi}. \quad (2.8)$$

The formulation for TM modes is: given the region $\tilde{\Omega}$ and possibly phase advances ψ or ψ_j (if periodicity is exploited), find the eigenmode fields E_z and the corresponding eigenvalues γ^2 such that

$$\nabla_t^2 E_z + \gamma^2 E_z = 0 \quad \text{in } \tilde{\Omega}, \quad (2.9a)$$

$$E_z = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym}(\text{cond})}, \quad (2.9b)$$

$$\hat{\mathbf{n}} \cdot \nabla E_z = 0 \quad \text{on } \Gamma_{\text{sym}(\text{ins})}, \quad (2.9c)$$

$$E_z \Big|_{\hat{\mathbf{R}}_j \mathbf{x}} = E_z \Big|_x e^{i\psi_j} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left-}j}, \quad (2.9d)$$

$$\hat{\mathbf{n}} \cdot \nabla E_z \Big|_{\hat{\mathbf{R}}_j \mathbf{x}} = -\hat{\mathbf{n}} \cdot \nabla E_z \Big|_x e^{i\psi_j} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left-}j}, \quad (2.9e)$$

$$\Im E_z = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (2.9f)$$

$$\Re \{(\hat{\mathbf{n}} \cdot \nabla) E_z\} = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (2.9g)$$

$$\Im \{E_z e^{-i\psi/2}\} = 0 \quad \text{on } \Gamma_{\text{sym-right}} \quad (2.9h)$$

$$\text{and } \Re \{(\hat{\mathbf{n}} \cdot \nabla) E_z e^{-i\psi/2}\} = 0 \quad \text{on } \Gamma_{\text{sym-right}}. \quad (2.9i)$$

This formulation does not include inversion symmetry since an equivalent 180° rotation may be used instead. The 2d problem may have two independent periodicities $\hat{\mathbf{R}}_1$ and $\hat{\mathbf{R}}_2$, hence there may be two pairs of quasi-periodic boundaries, $\hat{\mathbf{R}}_1: \Gamma_{\text{left-}1} \rightarrow \Gamma_{\text{right-}1}$ and $\hat{\mathbf{R}}_2: \Gamma_{\text{left-}2} \rightarrow \Gamma_{\text{right-}2}$, with phase advances ψ_1 and ψ_2 respectively. Many symmetry boundary conditions are included in the formulation but real structures will not use them all at once. For example, a problem which exploits two periodicities and a symmetry plane simultaneously would represent an unphysical problem.

A formulation for the TE modes is: given the region $\tilde{\Omega}$ and possibly phase advances ψ or ψ_j (if periodicity is exploited), find the eigenmode fields H_z and the corresponding eigenvalues γ^2 such that

$$\nabla_t^2 H_z + \gamma^2 H_z = 0 \quad \text{in } \tilde{\Omega}, \quad (2.10a)$$

$$\hat{\mathbf{n}} \cdot \nabla H_z = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym}(\text{cond})}, \quad (2.10b)$$

$$H_z = 0 \quad \text{on } \Gamma_{\text{sym}(\text{ins})}, \quad (2.10c)$$

$$H_z \Big|_{\hat{\mathbf{R}}_j \mathbf{x}} = H_z \Big|_x e^{i\psi_j} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left-}j}, \quad (2.10d)$$

$$\hat{\mathbf{n}} \cdot \nabla H_z \Big|_{\hat{\mathbf{R}}_j \mathbf{x}} = -\hat{\mathbf{n}} \cdot \nabla H_z \Big|_{\mathbf{x}} e^{i\psi_j} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left-}j}, \quad (2.10e)$$

$$\Im H_z = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (2.10f)$$

$$\Re \{(\hat{\mathbf{n}} \cdot \nabla) H_z\} = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (2.10g)$$

$$\Im \{H_z e^{-i\psi/2}\} = 0 \quad \text{on } \Gamma_{\text{sym-right}} \quad (2.10h)$$

$$\text{and } \Re \{(\hat{\mathbf{n}} \cdot \nabla) H_z e^{-i\psi/2}\} = 0 \quad \text{on } \Gamma_{\text{sym-right}}. \quad (2.10i)$$

This formulation treats H_z as a scalar field similar to E_z . This formulation does not treat H_z as a pseudoscalar field, which means the electric field computed from H_z will transform under symmetry operations like a pseudovector instead of a vector. This reversal of transformation properties is benign since the formulation does not include any source charges or currents. A field solver which exploits symmetries and includes sources will have to transform fields appropriately to match the transformation properties of the (scalar) charge and the (vector) current.

2.2. Variational Formulation

A variational formulation can be obtained from the above formulations by multiplying the wave equation by a test function and then integrating over the domain $\tilde{\Omega}$. Integration by parts using the vector identity $\nabla \cdot (v \nabla u) = v \nabla^2 u + \nabla v \cdot \nabla u$ reduces the integral to a form which is symmetric. A variational formulation for TM modes which is equivalent to (2.9) is: given the region $\tilde{\Omega}$ and possibly phase advances ψ or ψ_j (if periodicity is exploited), find the eigenmode fields $E_z \in \mathcal{U}_E$ and the corresponding eigenvalues γ^2 such that for all test functions $F_z \in \mathcal{U}_E$

$$\Re \int_{\tilde{\Omega}} \nabla F_z^* \cdot \nabla E_z - \gamma^2 F_z^* E_z d\Omega = 0 \quad (2.11a)$$

where the space \mathcal{U}_E of test and trial functions is

$$\begin{aligned} \mathcal{U}_E = \{ & E_z \in \mathcal{H}_1(\tilde{\Omega}) : \\ & E_z = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym(cond)}}, \\ & E_z \Big|_{\hat{\mathbf{R}}_j \mathbf{x}} = E_z \Big|_{\mathbf{x}} e^{i\psi_j} \quad \forall \mathbf{x} \in \Gamma_{\text{left}}, \\ & \Im E_z = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \\ & \Im \{E_z e^{-i\psi/2}\} = 0 \quad \text{on } \Gamma_{\text{sym-right}} \} \end{aligned} \quad (2.11b)$$

and

$$\begin{aligned} \mathcal{H}_1(\tilde{\Omega}) = \{ & E_z: \tilde{\Omega} \rightarrow \mathbb{C} : \\ & \int_{\tilde{\Omega}} \nabla E_z^* \cdot \nabla E_z + E_z^* E_z d\Omega \text{ exists and is finite} \}. \end{aligned} \quad (2.11c)$$

A solution of (2.9) is also a solution of this variational formulation. Conversely, a solution of this variational formulation is a solution of (2.9) provided that the fields

are sufficiently differentiable. In particular, formulation (2.9) requires the existence of the second derivative $\nabla^2 E_z$. The variational formulation only requires that the first derivative ∇E_z be integrable in the sense indicated by (2.11c). The formulation (2.9) is called a strong formulation and the variational formulation (2.11) is called a weak formulation because the conditions on the differentiability of the fields are stronger in (2.9) than in (2.11).

Only some of the boundary conditions which appeared in the strong formulation are listed in the definition (2.11b) of the function space \mathcal{U}_E . The boundary conditions listed in (2.11b) are called essential boundary conditions. The boundary conditions not listed in (2.11b) are called natural boundary conditions because any solution to the variational formulation naturally satisfies these boundary conditions without any additional constraints. The natural boundary conditions are typically the conditions which involve field derivatives at the boundaries. The Neumann boundary condition $(\hat{\mathbf{n}} \cdot \nabla)E_z = 0$ is a typical example. The Dirichlet boundary condition $E_z = 0$ is a typical example of an essential boundary condition.

The corresponding variational formulation for the TE modes is: given the region $\tilde{\Omega}$ and possibly phase advances ψ or ψ_j (if periodicity is exploited), find the eigenmode fields $H_z \in \mathcal{U}_H$ and the corresponding eigenvalues γ^2 such that for all test functions $G_z \in \mathcal{U}_H$

$$\Re \int_{\tilde{\Omega}} \nabla G_z^* \cdot \nabla H_z - \gamma^2 G_z^* H_z d\Omega = 0 \quad (2.12a)$$

where the space \mathcal{U}_H of test and trial functions is

$$\begin{aligned} \mathcal{U}_H = \{ & H_z \in \mathcal{H}_1(\tilde{\Omega}) : \\ & H_z = 0 \quad \text{on } \Gamma_{\text{sym}(\text{ins})}, \\ & H_z \Big|_{\hat{\mathbf{R}}_j \mathbf{x}} = H_z \Big|_{\mathbf{x}} e^{i\psi_j} \quad \forall \mathbf{x} \in \Gamma_{\text{left}}, \\ & \Im m H_z = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \\ & \Im m \{ H_z e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}} \}. \end{aligned} \quad (2.12b)$$

The Hilbert space $\mathcal{H}_1(\tilde{\Omega})$ is the same as (2.11c). This variational formulation is equivalent to (2.10).

2.3. Finite Element Method

The function space \mathcal{U}_E is an infinite dimensional space. Restricting the problem to a finite dimensional subspace $\mathcal{U}_E^h \subset \mathcal{U}_E$ gives a formulation of the eigenmode problem called the Galerkin formulation. For the TM modes the Galerkin formulation is: given the region $\tilde{\Omega}$ and possibly phase advances ψ or ψ_j (if periodicity is

exploited), find the eigenmode fields $E_z^h \in \mathcal{U}_E^h$ and the corresponding eigenvalues γ_h^2 such that for all test functions $F_z^h \in \mathcal{U}_E^h$

$$\Re \int_{\tilde{\Omega}} \nabla F_z^{h*} \cdot \nabla E_z^h - \gamma_h^2 F_z^{h*} E_z^h d\Omega = 0. \quad (2.13)$$

There are many possibilities for the subspace \mathcal{U}_E^h . Choosing a good space \mathcal{U}_E^h means the approximate eigenfields E_z^h obtained from the Galerkin formulation are close to the true eigenfields E_z obtained from the variational formulation. Likewise, the approximate eigenvalues γ_h^2 will be close to the true eigenvalues γ^2 if the space \mathcal{U}_E^h is chosen well.

The Galerkin formulation reduces to a problem involving matrices of finite order. Let the linearly independent set of functions $\{N_1, \dots, N_M\}$ be a basis for the space \mathcal{U}_E^h . Then any function $E_z^h \in \mathcal{U}_E^h$ is a unique linear combination of the basis functions,

$$E_z^h(x, y) = \sum_{i=1}^M u_i N_i(x, y), \quad (2.14)$$

where the coefficients u_i are real. A basis function N_i may be complex, particularly if the problem includes quasi-periodic or symmetric quasi-periodic boundaries. Decomposing the fields E_z^h and F_z^h of (2.13) into linear combinations of the basis functions gives

$$\begin{aligned} \Re \int_{\tilde{\Omega}} \nabla \left(\sum_{i=1}^M v_i N_i \right)^* \cdot \nabla \left(\sum_{j=1}^M u_j N_j \right) - \gamma_h^2 \left(\sum_{i=1}^M v_i N_i \right)^* \left(\sum_{j=1}^M u_j N_j \right) d\Omega \\ = \sum_{i=1}^M \sum_{j=1}^M v_i u_j \Re \left\{ \int_{\tilde{\Omega}} \nabla N_i^* \cdot \nabla N_j - \gamma_h^2 N_i^* N_j d\Omega \right\} = 0. \end{aligned} \quad (2.15)$$

Equation (2.13) must be true for any test field $F_z^h \in \mathcal{U}_E^h$, so (2.15) must be true for any v_i . The variables v_i are independent, so

$$\begin{aligned} \sum_{j=1}^M \Re \left\{ \int_{\tilde{\Omega}} \nabla N_i^* \cdot \nabla N_j - \gamma_h^2 N_i^* N_j d\Omega \right\} u_j \\ = \sum_{j=1}^M (A_{ij} - \gamma_h^2 B_{ij}) u_j = 0 \quad \text{for all } i = 1, \dots, M. \end{aligned} \quad (2.16)$$

The notation

$$A_{ij} = \Re \int_{\tilde{\Omega}} \nabla N_i^* \cdot \nabla N_j d\Omega \quad \text{and} \quad B_{ij} = \Re \int_{\tilde{\Omega}} N_i^* N_j d\Omega \quad (2.17)$$

has been introduced. Equation (2.16) is written concisely in matrix notation as

$$\mathbf{A}\mathbf{u} - \gamma_h^2 \mathbf{B}\mathbf{u} = 0 \quad (2.18)$$

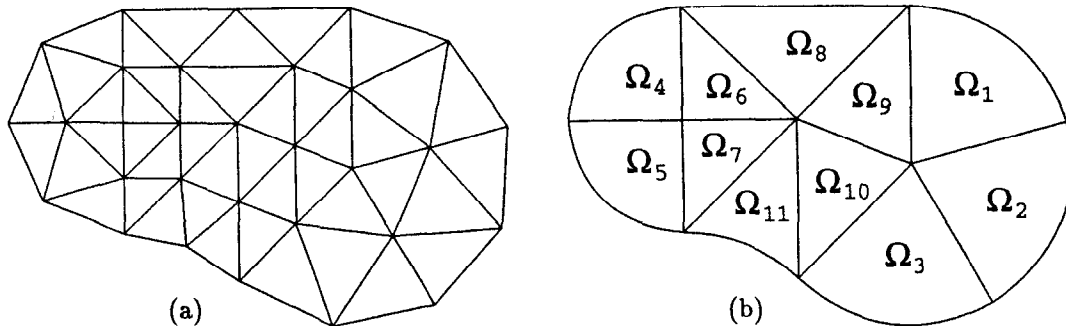


Figure 2.2: Two finite element meshes for the arbitrary cylindrical waveguide in figure 2.1. Mesh (a) is composed of linear elements. Mesh (b) is composed of quadratic elements.

where the components of matrices \mathbf{A} and \mathbf{B} are given by (2.17) and the vector of field coefficients is $\mathbf{u} = (u_1, \dots, u_M)^T$. Thus the Galerkin formulation of the eigenmode problem reduces to the generalized algebraic eigenvalue problem (2.18) for the eigenvectors \mathbf{u} and the corresponding eigenvalues γ_h^2 .

Notice that the matrices \mathbf{A} and \mathbf{B} are symmetric. The symmetry of the matrices reflects the self-adjoint nature of the eigenmode problem.

Matrix \mathbf{A} is non-negative definite and matrix \mathbf{B} is positive definite, so the eigenvalues γ_h^2 are real and non-negative. The TE mode formulation often has a zero frequency mode which corresponds to a static constant magnetic field in the structure. The TM mode formulation is positive definite unless a rare combination of boundary conditions is chosen (e.g., all symmetry boundaries) to make the problem look like a TE mode problem. In either case there are at most only a few zero eigenvalues. The remaining eigenvalues are positive.

The finite element method provides, among other things, a scheme for choosing subspaces \mathcal{U}_E^h for the Galerkin formulation of the problem.

2.3.1. Elements

In the finite element method the region $\tilde{\Omega}$ is partitioned into a mesh of non-overlapping elements Ω_e as shown in figure 2.2. The elements Ω_e are simple shapes, typically triangles or quadrilaterals, which cover a region $\tilde{\Omega}^h$. The work described in this thesis uses only triangular elements. Note that $\tilde{\Omega}^h$ is often only an approximation to the region $\tilde{\Omega}$. For example, the approximate nature of the mesh in figure 2.2a is obvious because the mesh looks very different from $\tilde{\Omega}$ in figure 2.1b. Careful scrutiny will reveal that the mesh in figure 2.2b is also only an approximation to $\tilde{\Omega}$.

The finite element method prescribes basis functions for the Galerkin formulation which are simple functions on each element. These simple functions are called

local basis functions and are typically low order polynomials. The domain of a local basis function N_i^e is just the e th element.

The local basis functions are assembled into the global basis functions $N_i \in \mathcal{U}_E$ which are zero everywhere except on a single element (Ω_8 in figure 2.2b, for example) or a pair of elements which share a side (Ω_1 and Ω_2 , for example) or a set of elements which share a corner ($\Omega_4, \Omega_5, \Omega_6$ and Ω_7 , for example). The support (the domain over which the function is non-zero) of a global basis function N_i consists of at most a few elements. Elsewhere $N_i = 0$. Examples of global basis functions are shown in figures 2.5 and 2.8 in the following sections.

The finite element method also suggests a scheme for computing the matrices **A** and **B**. The integration for the matrix components (2.17) can be split into a sum of integrals over each element,

$$\begin{aligned} A_{ij} &= \sum_e \Re e \int_{\Omega_e} \nabla N_i^* \cdot \nabla N_j d\Omega \\ \text{and } B_{ij} &= \sum_e \Re e \int_{\Omega_e} N_i^* N_j d\Omega. \end{aligned} \tag{2.19}$$

A global basis function N_i on Ω_e is either zero or it is one of the local basis functions for element e , say N_i^e . The (global) index i for the global basis function is not the same as the (local) index \hat{i} for the corresponding local basis function on element e . This is demonstrated in an example on the following page, where element matrices are accumulated to global matrices.

Consider a pair of global basis functions, N_i and N_j . Equation (2.19) indicates that the components A_{ij} and B_{ij} can be non-zero only if the two global basis functions N_i and N_j have an element in common. Since the support of N_i consists of at most a few elements, there are a limited number of basis functions N_j which share an element with N_i . This implies that there are few non-zero entries in the i th row of the matrices **A** and **B**, so the matrices **A** and **B** are sparse matrices.

For a given i and j few elements, if any, will have a non-zero contribution to the sum in (2.19), so computing A_{ij} and B_{ij} by summing the contributions to (2.19) over every element in the mesh is not the most practical method to compute **A** and **B**. It is more efficient to first compute the non-zero terms of (2.19) in the form of element matrices. The element matrices \mathbf{A}^e and \mathbf{B}^e are

$$\begin{aligned} A_{\hat{i}\hat{j}}^e &= \Re e \int_{\Omega_e} \nabla N_i^{e*} \cdot \nabla N_j^e d\Omega \\ \text{and } B_{\hat{i}\hat{j}}^e &= \Re e \int_{\Omega_e} N_i^{e*} N_j^e d\Omega. \end{aligned} \tag{2.20}$$

There are separate element matrices for each element of the mesh. The order of an element matrix is the number of local basis functions in the element, so the element matrices are small. The element matrices are also dense, so little computational effort is spent computing zero. The element matrices are also symmetric.

The element matrices \mathbf{A}^e and \mathbf{B}^e are easily assembled into the matrices \mathbf{A} and \mathbf{B} . The components of the small dense matrices \mathbf{A}^e and \mathbf{B}^e are scattered to the appropriate components of the large sparse matrices \mathbf{A} and \mathbf{B} . This is done for each element of the mesh.

As an example, suppose the 8th element (i.e., $e = 8$) of a mesh has 3 local basis functions. The element matrices \mathbf{A}^8 and \mathbf{B}^8 are dense symmetric 3×3 matrices. Suppose further that global basis $i = 23$ is local basis $\hat{i} = 1$ on the 8th element, and that global basis $j = 27$ is local basis $\hat{j} = 2$ on the 8th element. Then one step of the assembly procedure for the 8th element would be the accumulation of the element matrix component A_{12}^8 to the global matrix component $A_{23,27}$. Similar steps accumulate the remaining components of \mathbf{A}^8 to \mathbf{A} , and then the components of \mathbf{B}^8 to \mathbf{B} .

One further optimization should be mentioned. When the fields are complex there are typically pairs of basis functions which are identical except for a complex phase factor. One basis function of the pair will be purely real, while the other basis function of the pair will be purely imaginary. The presence of these pairs arose from the desire to factor the coefficients v_i and u_j outside of the \Re operator in (2.15), and thus simplify the form of the matrices and their components. Instead of a pair of basis functions with one real coefficient each, a single real basis function with one complex coefficient could be used instead.

The field solver should compute the element matrices (2.20) for real basis functions only. The components of the element matrices which involve imaginary or complex basis functions are easily computed from the components involving only real basis functions. Only multiplication by a complex factor is involved. For this reason, the basis functions described in the remainder of this thesis are real, and complex basis functions will not be mentioned further.

2.3.2. Linear Triangle Elements

Figure 2.2a is a mesh composed of linear triangle elements. Each element is a linear map of the master element $\hat{\Omega}$ as shown in figure 2.3. The master element $\hat{\Omega}$ provides a common domain for all triangular elements. The utility of having a single master element for all triangular elements will become apparent later, particularly

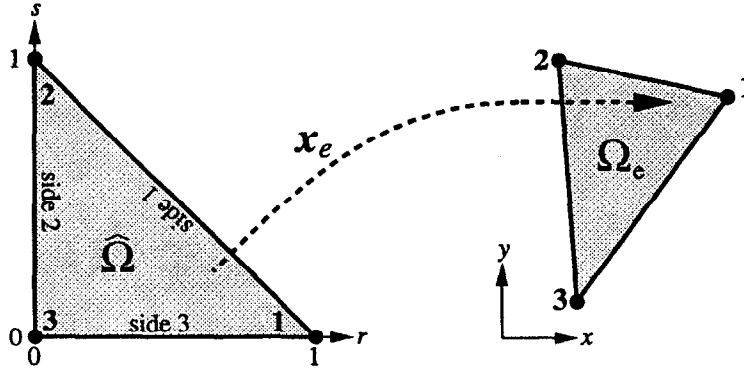


Figure 2.3: The element transformation \mathbf{x}_e maps the master element $\hat{\Omega}$ to the eth element Ω_e . The map \mathbf{x}_e is linear.

when quadratic triangle elements are discussed.

Let $\mathbf{x}_e: \hat{\Omega} \rightarrow \Omega_e$ be the linear map for the eth element. In terms of components the map is,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{x}_e(r, s) = \begin{pmatrix} x_e(r, s) \\ y_e(r, s) \end{pmatrix}. \quad (2.21)$$

While $\hat{\Omega}$ is the same for all triangle elements, different elements have different domains Ω_e and thus different maps \mathbf{x}_e . Coordinates x and y are global coordinates representing a point in $\tilde{\Omega}$, while coordinates r and s are local coordinates representing a point in the master element $\hat{\Omega}$. The local coordinates are defined in figure 2.3.

Local basis functions \hat{N}_i are defined on $\hat{\Omega}$ and mapped (using \mathbf{x}_e) to Ω_e . The three local basis functions are

$$\hat{N}_1 = r, \quad \hat{N}_2 = s \quad \text{and} \quad \hat{N}_3 = t, \quad (2.22)$$

where $t = 1 - r - s$. Each of the local basis functions is a linear polynomial in r and s . The local basis function \hat{N}_i is 1 at corner \hat{i} and 0 at the other two corners, as shown in figure 2.4.

In figure 2.4, the size of the circles is proportional to the magnitude of the basis function at the center of the circle. They can be considered arrows directed out of the page, since the basis functions are a basis for the field E_z , which is normal to the page. These basis functions are all positive, but later figures will have both negative and positive fields. A cross through the circle indicates the field at the center of the circle is negative.

The local basis functions on Ω_e are simply

$$N_i^e(x, y) = \hat{N}_i(\mathbf{x}_e^{-1}(x, y)). \quad (2.23)$$

Notice that $N_i^e(x, y)$ is a function of the global coordinates while $\hat{N}_i(r, s)$ is a function of the local coordinates.

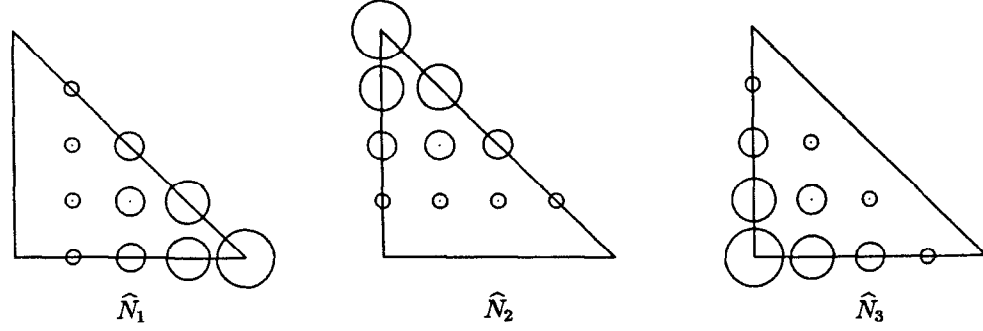


Figure 2.4: Local basis functions \hat{N}_i for the linear triangle element. The size of the circles are proportional to the magnitude of the basis function at the center of the circles.

It is convenient to integrate over $\hat{\Omega}$ instead of Ω_e when computing the element matrices. This is easily accomplished by including the Jacobian J_e of the element transformation \mathbf{x}_e . Hence

$$A_{ij}^e = \Re \int_{\Omega_e} \nabla N_i^{e*} \cdot \nabla N_j^e d\Omega = \Re \int_{\hat{\Omega}} \nabla \hat{N}_i^* \cdot \nabla \hat{N}_j J_e dr ds \quad (2.24)$$

and $B_{ij}^e = \Re \int_{\Omega_e} N_i^{e*} N_j^e d\Omega = \Re \int_{\hat{\Omega}} \hat{N}_i^* \hat{N}_j J_e dr ds,$

where

$$J_e = \left| \frac{\partial(x_e, y_e)}{\partial(r, s)} \right| = \det \begin{pmatrix} \partial x_e / \partial r & \partial x_e / \partial s \\ \partial y_e / \partial r & \partial y_e / \partial s \end{pmatrix}. \quad (2.25)$$

The gradients $\nabla \hat{N}_i^*$ and $\nabla \hat{N}_j$ above are still with respect to the global coordinates x and y . A simple application of the chain rule yields the gradient in terms of the local coordinates r and s on the master element $\hat{\Omega}$,

$$\frac{\partial \hat{N}_i}{\partial x} = \frac{\partial \hat{N}_i}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial \hat{N}_i}{\partial s} \frac{\partial s}{\partial x}, \quad (2.26)$$

and $\frac{\partial \hat{N}_i}{\partial y} = \frac{\partial \hat{N}_i}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial \hat{N}_i}{\partial s} \frac{\partial s}{\partial y}.$

Derivatives like $\partial r / \partial x$ are derivatives of \mathbf{x}_e^{-1} , which can be computed on $\hat{\Omega}$ from derivatives of \mathbf{x}_e using

$$\begin{pmatrix} \partial r / \partial x & \partial r / \partial y \\ \partial s / \partial x & \partial s / \partial y \end{pmatrix} = \frac{1}{J_e} \begin{pmatrix} \partial y_e / \partial s & -\partial x_e / \partial s \\ -\partial y_e / \partial r & \partial x_e / \partial r \end{pmatrix}. \quad (2.27)$$

The map \mathbf{x}_e can be written in terms of the local basis functions \hat{N}_i and the corners of the element \mathbf{x}_i . The map is

$$\mathbf{x}_e(r, s) = \sum_{i=1}^3 \mathbf{x}_i \hat{N}_i(r, s). \quad (2.28)$$

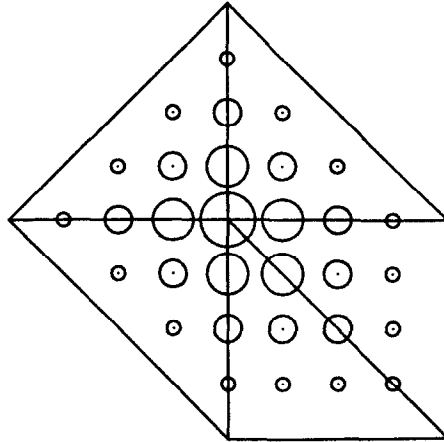


Figure 2.5: A global basis function composed of linear elements. The basis function is associated with the field at the common corner. The size of the circles are proportional to the magnitude of the basis function at the center of the circles.

This element is called an isoparametric element because the element transformation \mathbf{x}_e can be written in terms of the local basis functions N_i . In other words, the element transformation and the field have the same parametrization.

The local basis functions have the property that $\hat{N}_j(r_i, s_i) = \delta_{ij}$, where (r_i, s_i) are the local coordinates of node \hat{i} . A polynomial basis with this property is called a lagrangian basis.

Figure 2.5 shows a global basis function constructed from one local basis function on each element surrounding a corner. The global basis function is 1 at the common corner and zero at all other corners. Hence the global basis function is associated with the field at the common corner: the coefficient for this global basis function is just the field at the common corner. The global basis function is zero on any other element not shown in figure 2.5. That is, the global basis function is zero on any element not touching the common corner.

Notice that the global basis function is continuous across the element boundaries. This occurs for two reasons. First, a common element boundary is defined solely by the two corners shared by the two adjacent elements. The location of other corners do not affect the common boundary or its parametrization. Second, the local basis functions are linear on each element, and thus they are also linear on the shared boundary. The local basis functions agree at the two shared corners, thus they agree along the entire shared boundary. This may seem like a trivial point for linear elements, but a similar argument applies to the quadratic elements, where continuity of the global basis functions is not as obvious.

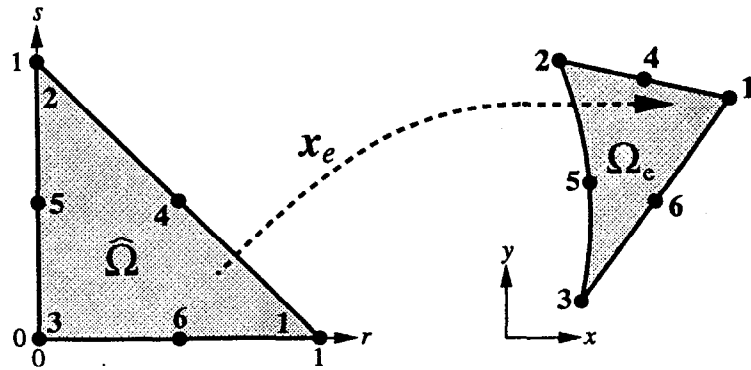


Figure 2.6: The element transformation \mathbf{x}_e maps the master element $\hat{\Omega}$ to the eth element Ω_e . The map \mathbf{x}_e is quadratic.

The corners of the linear triangle element are also called the nodes of the element because global basis functions are associated with the field at the corners of the element. That is, if \mathbf{x}_i is the i th node (corner) of a mesh composed of linear triangle elements, and \mathbf{u} is the coefficient vector for a field $E_z^h \in \mathcal{U}_E^h$ then $E_z^h(\mathbf{x}_i) = u_i$. For the quadratic triangle element discussed later there are basis functions associated with the field at the midside points of the triangle. Hence the quadratic element has midside nodes in addition to corner nodes. In chapter 4 some basis functions are associated with the field tangential to the sides of the elements. These can also be considered nodes of the element.

2.3.3. Quadratic Triangle Elements

Quadratic triangle elements are similar to linear triangle elements in that a quadratic triangle element is also based on a map \mathbf{x}_e from a master element $\hat{\Omega}$ to the element domain Ω_e . The master element is the same, but the map may be quadratic instead of just linear. This allows an element Ω_e to have curved sides as shown in figure 2.6.

There are six local basis functions for this element. Each of the basis functions is associated with the field at one of six nodes. In this work the three corner nodes are labelled first in counter-clockwise order, just like the linear element. Then the three midside nodes are labelled starting with side 1, again in counter-clockwise order. The basis functions are

$$\begin{aligned} \hat{N}_1 &= r(2r - 1), & \hat{N}_2 &= s(2s - 1), & \hat{N}_3 &= t(2t - 1), \\ \hat{N}_4 &= 4rs, & \hat{N}_5 &= 4st & \text{and} & \hat{N}_6 = 4rt. \end{aligned} \quad (2.29)$$

The basis functions are quadratic in r and s . The set of basis functions forms a complete basis for the space of quadratic polynomials. That is, any function $u(r, s)$

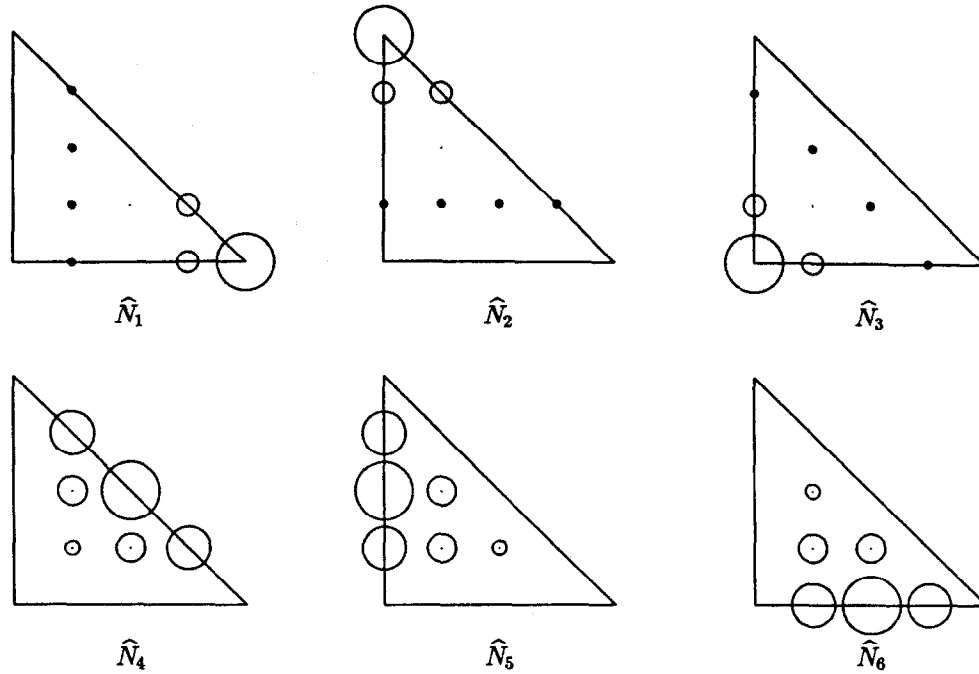


Figure 2.7: Local basis functions \hat{N}_i for the quadratic triangle element. The size of the circles are proportional to the magnitude of the basis function at the center of the circles. A cross in the circle indicates the basis function is negative.

of the form

$$u(r, s) = \alpha_1 + \alpha_2 r + \alpha_3 s + \alpha_4 r^2 + \alpha_5 r s + \alpha_6 s^2 \quad (2.30)$$

can be written as a unique linear combination of the basis functions

$$u(r, s) = \sum_{i=1}^6 u_i \hat{N}_i(r, s). \quad (2.31)$$

Local basis \hat{N}_i is 1 at node i and 0 at all other nodes. Thus this is a lagrangian basis. Plots of the local basis functions are shown in figure 2.7.

This element is isoparametric just like the linear element. That is, the map \mathbf{x}_e can be written in terms of the local basis functions \hat{N}_i and the nodes \mathbf{x}_i of the element,

$$\mathbf{x}_e(r, s) = \sum_{i=1}^6 \mathbf{x}_i \hat{N}_i(r, s). \quad (2.32)$$

Two examples of global basis functions are shown in figure 2.8. As in the linear element case, the global basis functions are continuous across element boundaries. Note, however, that the component of the derivative normal to the boundary is not continuous. This is the case for both linear and quadratic global basis functions.

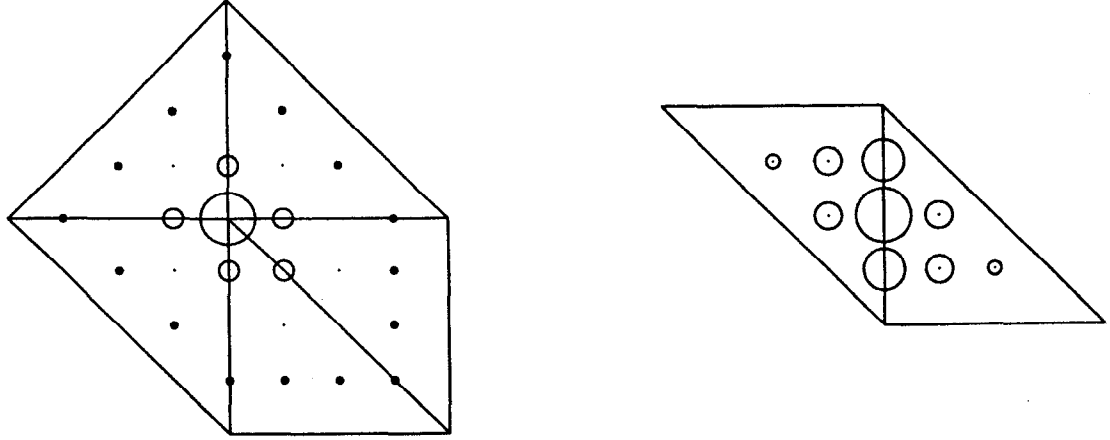


Figure 2.8: Two global basis functions composed of quadratic elements. The left basis function is associated with the field at a corner node, while the right basis function is associated with the field at a midside node. The size of the circles are proportional to the magnitude of the basis function at the center of the circles. A cross in the circle indicates the basis function is negative.

The condition (2.11c) demands that the field be continuous across any boundary, but it does not demand that the derivatives be continuous across any boundary.

The integrals (2.20) for the element matrices \mathbf{A}^e and \mathbf{B}^e are transformed to integrals (2.24) over the master element $\hat{\Omega}$ in order to simplify the domain of integration. While this transformation may seem to be only an added complication when the map \mathbf{x}_e is linear, the simplification of the integration domain is very helpful when the element domain Ω_e has curved boundaries (and thus \mathbf{x}_e is non-linear). At this point the integration could still be carried out analytically for \mathbf{B}^e , but the integrand for \mathbf{A}^e is not simply a polynomial in r and s , so \mathbf{A}^e is difficult to obtain analytically. To circumvent this difficulty numerical integration is used instead.

A list of integration rules for the triangular domain $\hat{\Omega}$ is given in [6] (ch. 3). The weights and points for the 12-point quadrature formula used in this work are listed in table 2.1. The approximate integral is

$$\int_{\hat{\Omega}} f d\Omega \cong \frac{1}{2} \sum_{i=1}^{12} f(r_i, s_i, t_i) W_i. \quad (2.33)$$

Additional accuracy was obtained by subdividing the domain $\hat{\Omega}$ into four similar triangles. A number of test cases showed that this numerical integration scheme is more than sufficient for the problems solved here. Numerical integration was used for all element matrices for both the linear and quadratic elements.

i	W_i	r_i	s_i	t_i
1	0.050844906370207	0.873821971016996	0.063089014491502	0.063089014491502
2	0.050844906370207	0.063089014491502	0.873821971016996	0.063089014491502
3	0.050844906370207	0.063089014491502	0.063089014491502	0.873821971016996
4	0.116786275726379	0.501426509658179	0.249286745170910	0.249286745170911
5	0.116786275726379	0.249286745170911	0.501426509658179	0.249286745170910
6	0.116786275726379	0.249286745170910	0.249286745170911	0.501426509658179
7	0.082851075618374	0.636502499121399	0.310352451033785	0.053145049844816
8	0.082851075618374	0.636502499121399	0.053145049844816	0.310352451033785
9	0.082851075618374	0.053145049844816	0.636502499121399	0.310352451033785
10	0.082851075618374	0.053145049844816	0.310352451033785	0.636502499121399
11	0.082851075618374	0.310352451033785	0.053145049844816	0.636502499121399
12	0.082851075618374	0.310352451033785	0.636502499121399	0.053145049844816

Table 2.1: Weights and points for a 12-point scheme for numerical integration over the domain $\hat{\Omega}$. The first two weights are repeated three times for the three permutations of the corresponding triangular coordinates. The last weight is repeated six times for the six permutations of its triangular coordinates.

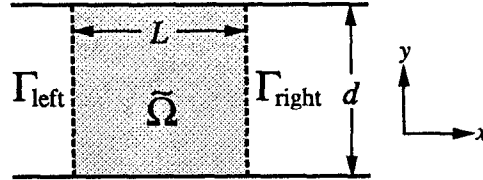


Figure 2.9: Two parallel plates separated by distance d treated as a periodic structure with period L .

2.3.4. Tests on Parallel Plates

The first test of these finite element formulations is on a pair of parallel plates spaced 1 cm apart. The plates are treated as a periodic structure with period $L = 1$ cm in order to test the implementation of quasi-periodic boundary conditions as well as the implementation of the finite element method. The test problem is illustrated in figure 2.9.

The finite element program is called YAP. Comparisons will be made with a simple finite difference program called QUAP. QUAP uses the usual five-point difference operator on a uniform grid to discretize and subsequently solve the wave equation. Later tests will compare YAP with the established field solvers SUPERFISH[8], URMEL[9] and URMEL-T[10].

Figures 2.10 and 2.11 show the results of calculations for various mesh sizes and various field solvers. The computed mode is the lowest frequency mode with phase advance $\psi = 135^\circ$. The field H_z is uniform along y . Analytically the solution is

$$H_z(x, y) = H_0 e^{ik_x x} = H_0 e^{i\psi x/L} \quad (2.34)$$

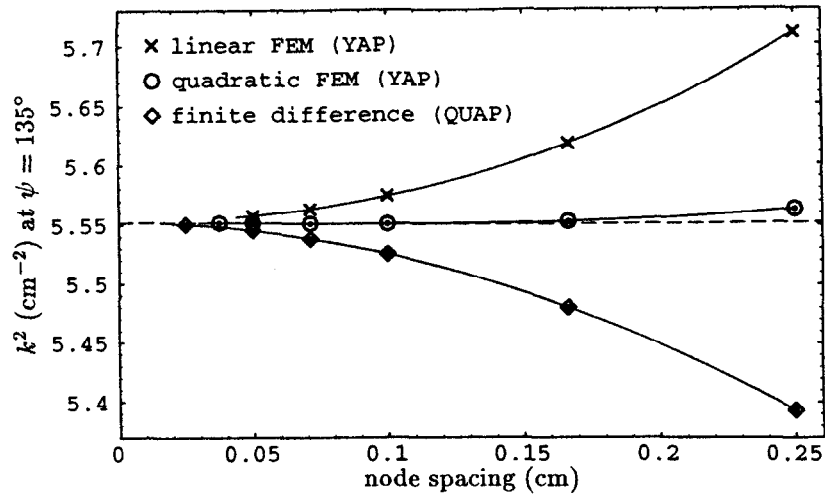


Figure 2.10: Computed eigenvalue k^2 for the parallel plate test as the mesh is refined. Results from three different algorithms are shown. The dashed line is the exact result obtained analytically.

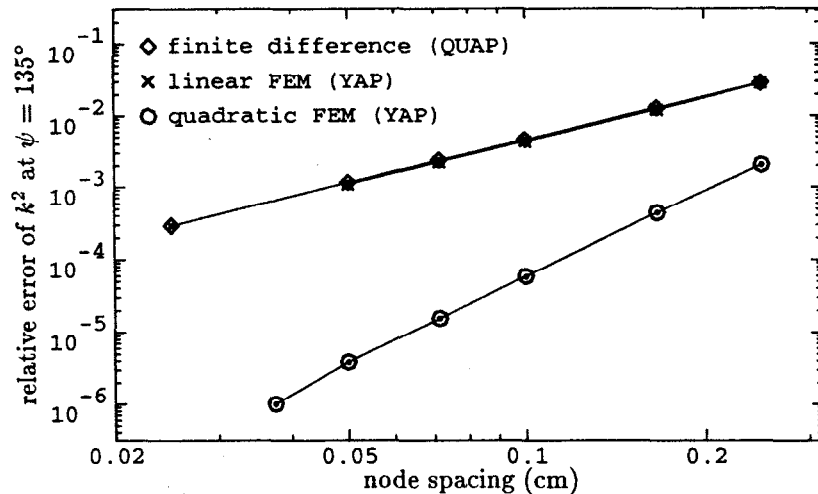


Figure 2.11: Relative error of the computed eigenvalue k^2 for the parallel plate test as the mesh is refined. Results from three different algorithms are shown. The curves for the finite difference and linear finite element algorithms coincide.

so the eigenvalue is $k^2 = \omega^2/c^2 = 5.5516525 \text{ cm}^{-2}$. The dashed line in figure 2.10 corresponds to this analytically computed eigenvalue.

Figure 2.10 shows the computed eigenvalues converging smoothly toward the correct eigenvalue as the mesh is refined. The coarse meshes have widely separated nodes and thus appear at the right side of the graph. The coarsest meshes yield the largest errors due to the coarse discretization of the problem.

The node spacing decreases as the mesh is refined, so fine meshes appear at the left side of the graph. The eigenvalues obtained from a fine mesh are more accurate

than the eigenvalues obtained from a coarse mesh. Ideally the exact eigenvalue is obtained when the node spacing becomes zero. Zero node spacing corresponds to an infinitely refined mesh. Of course, zero node spacing is never achieved in practice. In this work the limiting factor was storage space (computer memory). If storage space was not limited, one would eventually encounter numerical round-off problems while solving the generalized eigenvalue problem. This occurs when the matrices become sufficiently large and ill-conditioned. This work used double precision arithmetic to avoid such problems, and no indication of round-off problems has been found.

The smooth convergence toward the correct eigenvalue shown in figure 2.10 is not surprising since the problem has a simple rectangular boundary which a uniform mesh can model well. The two finite element algorithms overestimate the eigenvalue while the finite difference method underestimates the eigenvalue. In other words, the finite element methods approach the correct eigenvalue from above, while the finite difference method approaches the eigenvalue from below. It can be shown that the finite element methods will always overestimate the eigenvalue of a mode.

The magnitude of the discretization error for the linear finite element method and the 5-point finite difference method is nearly the same. This can be seen clearly in figure 2.11 where the relative error, $(k^2 - k_{\text{exact}}^2)/k_{\text{exact}}^2$, is plotted on a log-log scale. The relative error for the finite element method using quadratic elements is much smaller than the other two algorithms. Furthermore, the rate at which the quadratic finite element method converges is superior to the rate at which the linear finite element and the 5-point finite difference methods converge. The eigenvalue error for the linear finite element and 5-point finite difference methods is $\mathcal{O}(h^2)$, where h is the size of the elements or the spacing between nodes. The eigenvalue error for the quadratic finite element method is $\mathcal{O}(h^4)$.

For quadratic elements the node spacing h corresponds to half of the element size due to the presence of the midside nodes. Hence, for a given node spacing h in figures 2.10 and 2.11, the number of unknowns in the finite difference and the two finite element algorithms is identical.

2.3.5. Tests on a Circular Waveguide

A more stringent test of a field solver is a circular waveguide. Figure 2.12a shows the cross-section for a circular waveguide with radius a . In the tests below, the radius of the circular waveguide was chosen to be $a = 0.9$ cm.

Like many real-life problems, the cross-section of the circular waveguide is not rectangular. This will present a problem for finite difference algorithms which insist

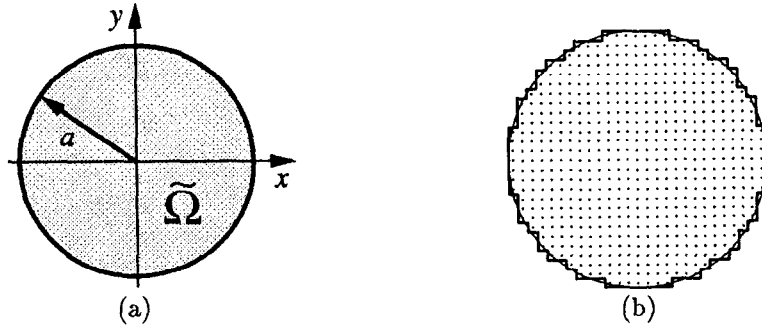


Figure 2.12: The region $\tilde{\Omega}$ (a) and a finite difference mesh (b) for a circular waveguide with radius a . The finite difference mesh was generated for the program QUAP.

on using a regular grid. An example of such a grid for QUAP is shown in figure 2.12b. The grid boundary is the thick line while the desired circular boundary is the thin line. The grid only coarsely approximates the circular domain of the problem.

URMEL is another finite difference field solver. For problems with a rectangular domain URMEL and QUAP give identical results. For this problem, however, URMEL generates a better grid than QUAP by including an occasional right triangular element next to the boundary. URMEL also allows a non-uniform grid, but this feature was not employed in the tests below.

URMEL-T is also considered a finite difference field solver. It uses a grid based on triangles and a dual grid composed of hexagons. These grids can approximate a circular domain well.

Figures 2.13 and 2.14 show calculations of the TE_{11} mode of the circular waveguide. The analytically computed eigenvalue is $k = j'_{1,1}/a$ where $j'_{1,1}$ is the location of the first maximum of the bessel function $J_1(x)$. Numerically the analytic eigenvalue is $k^2 = 4.185132984 \text{ cm}^{-2}$. This exact eigenvalue is represented by the dashed line in figure 2.13.

In figure 2.13 notice the erratic behavior exhibited by the two finite difference codes QUAP and URMEL. As the mesh is refined the computed eigenvalue does not necessarily become more accurate. In many cases the computed eigenvalue becomes less accurate for a refined mesh. This test demonstrates the difficulty these codes have modelling problems with curved boundaries. The replacement of the desired circular boundary with an approximate stair-step boundary dominates the discretization error of the problem. This boundary-induced error fluctuates unpredictably as the node spacing is reduced. This not only makes it difficult to decide which result (or mesh) is better, it also makes extrapolation of the calculations to zero node spacing difficult.

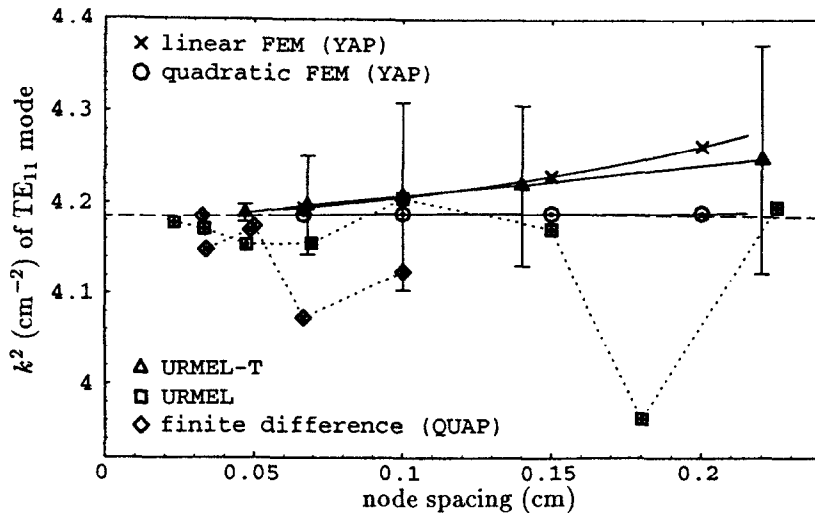


Figure 2.13: Computed eigenvalue k^2 for the circular waveguide test as the mesh is refined. Results from five different algorithms are shown. The dashed line is the exact result obtained analytically.

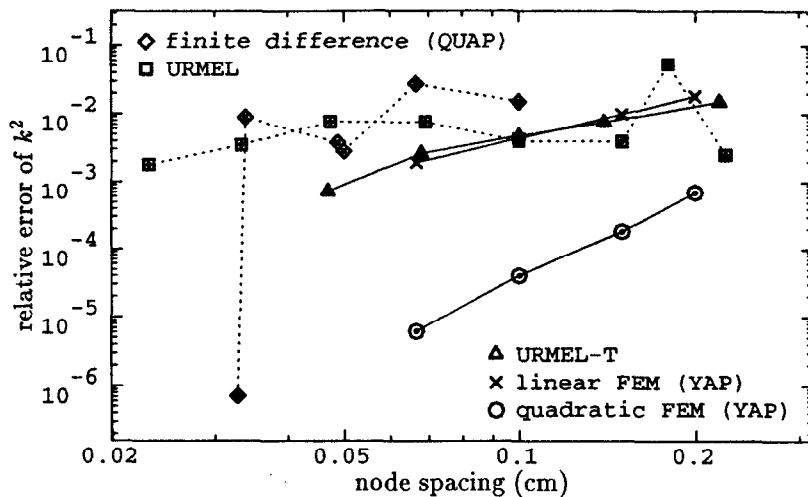


Figure 2.14: Relative error of the computed eigenvalue k^2 for the circular waveguide test as the mesh is refined. Results from five different algorithms are shown.

This example also indicates that caution must be exercised when verifying the accuracy of a finite difference code on this type of problem. If a test calculation is performed on just one mesh then the results could easily be misleading. For example, if the only eigenvalue computed with QUAP was the eigenvalue computed using the finest mesh in figures 2.13 and 2.14 then one might believe that QUAP was the best of the five codes tested. But when a variety of meshes are used then it becomes apparent that this is not true. QUAP should have the worst accuracy of the five codes tested.

Likewise, URMEL on the coarsest mesh tested looks good. Further tests using

different meshes indicate the impressive accuracy obtained using the coarse mesh was merely coincidence.

Similarly, when verifying measurement data the engineer or scientist using a field solver may feel an urge to adjust the finite difference mesh in order to make it “better” correspond to the actual structure. This usually means the mesh is adjusted until the measured frequency and the calculated frequency agree, although the user may conjure other criteria to rationalize his or her adjustments. In fact, one such rationalization will be used below. Figure 2.13 indicates that the eigenvalues can easily be adjusted by a few percent. The code’s mythical accuracy is perpetuated when the user reports the “nearly perfect” results to his or her colleagues.

For this problem the error of the QUAP calculations is correlated with the area enclosed by the finite difference mesh. If the calculations are scaled such that the area of the mesh is correct then the adjusted eigenvalues k_{adj}^2 are

$$k_{\text{adj}}^2 = k^2 \frac{A}{\pi a^2} \quad (2.35)$$

where k^2 is the eigenvalue computed by QUAP using a mesh with area A . The adjusted eigenvalues are relatively predictable as the mesh is refined. This is shown in figure 2.15. The dashed line is the analytically computed eigenvalue and the solid line is a fit to the area-adjusted eigenvalues. The dependence of the discretization error on node spacing h looks like $\mathcal{O}(h)$, which is worse than the $\mathcal{O}(h^2)$ error observed in the parallel plate test. Extrapolation of the adjusted eigenvalues to zero node spacing gives an improved estimate of the eigenvalue. While area normalization is useful for this test problem it is probably not useful for any problems more complex than this.

The eigenvalues computed by URMEL-T are unusual in that the eigenvalues for the two degenerate modes are separated. The two eigenvalues are at the ends of the error bars in figure 2.13. In figures 2.13 and 2.14 the arithmetic average of the two eigenvalues computed by URMEL-T is plotted. While the individual modes are a little unpredictable as the average node spacing is reduced, the average eigenvalue seems to converge smoothly toward the correct eigenvalue as the mesh is refined.

The two finite element methods (YAP) converge smoothly to the desired solution. Furthermore, the eigenvalues for the two degenerate modes are nearly identical. The accuracy obtained using linear elements is similar to URMEL-T, while the accuracy obtained using quadratic elements is superior. Figure 2.14 illustrates that, as in the parallel plate test, the error using linear elements is $\mathcal{O}(h^2)$ and the error using quadratic elements is $\mathcal{O}(h^4)$. In addition, extrapolation of the computed

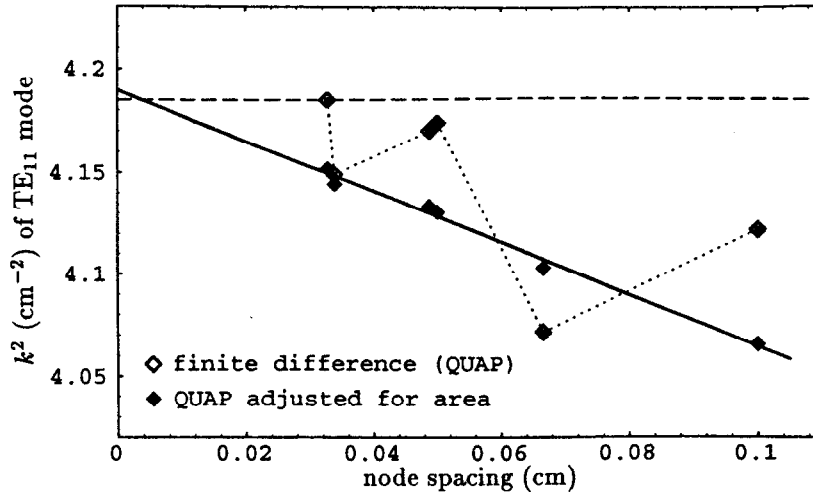


Figure 2.15: Computed and adjusted eigenvalues k^2 for the circular waveguide test as the mesh is refined. The dashed line is the exact result obtained analytically. The solid line is a fit to the adjusted eigenvalues.

eigenvalues to zero node spacing may be used to obtain an improved estimate of the eigenvalue.

2.4. Sharp Corners

Some microwave structures have sharp (convex) corners. Structures may also be modelled with sharp corners when the radius of curvature of the corner is much smaller than the element size or node spacing.

Consider the field H_z of a TE mode near the sharp corner depicted in figure 2.16. The field near the corner can be expanded in a power series in polar coordinates,

$$H_z(\rho, \phi) = a_0 + a_1 \cos(\pi\phi/\beta)\rho^{\pi/\beta} + a_2 \cos(\pi\phi/\beta)\rho^{2\pi/\beta} + \dots, \quad (2.36)$$

where ρ is the distance from the corner, β is the corner angle and ϕ is the angle from one wall, $0 \leq \phi \leq \beta$. The field E_z for a TM mode has a similar expansion near the sharp corner,

$$E_z(\rho, \phi) = a_1 \sin(\pi\phi/\beta)\rho^{\pi/\beta} + a_2 \sin(\pi\phi/\beta)\rho^{2\pi/\beta} + \dots. \quad (2.37)$$

Notice in both cases that the power series satisfies the boundary conditions at the metal walls near the corner.

The derivatives of the field are singular at the corner. Singular derivatives are not approximated well by linear or quadratic elements, which means problems containing sharp corners are more difficult to solve than problems with smooth boundaries. This difficulty is illustrated in a ridged waveguide example following

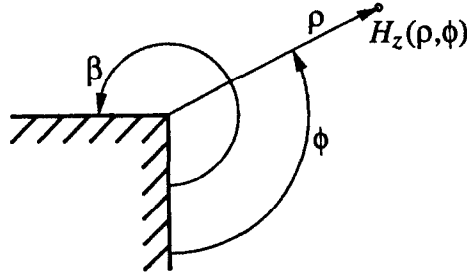


Figure 2.16: Polar coordinates near a sharp corner.

the introduction of a special corner element which improves the accuracy of the solution.

2.4.1. Special Corner Elements

Figure 2.17a shows the master element for a seven node special corner element for use with quadratic elements. Node 3 is the corner node. The map from the master element $\hat{\Omega}$ to the corner element Ω_e is quadratic and is completely specified by the coordinates of the first six nodes of the element using equation (2.32). The seven basis functions for the special corner element are

$$\begin{aligned}
 N_1 &= u^\alpha(2^\alpha u^\alpha - 1)(1 - v)(1 - 2v)/(2^\alpha - 1) \\
 N_2 &= u^\alpha(2^\alpha u^\alpha - 1)(2v - 1)v/(2^\alpha - 1) \\
 N_3 &= (1 - 2^\alpha u^\alpha)(1 - u^\alpha) \\
 N_4 &= u^\alpha(2^\alpha u^\alpha - 1)4v(1 - v)/(2^\alpha - 1) \\
 N_5 &= 4^\alpha u^\alpha(1 - u^\alpha)v(2v - 1)/(2^\alpha - 1) \\
 N_6 &= 4^\alpha u^\alpha(1 - u^\alpha)(1 - v)(1 - 2v)/(2^\alpha - 1) \\
 N_7 &= 4^\alpha u^\alpha(1 - u^\alpha)4v(1 - v)/(2^\alpha - 1)
 \end{aligned} \tag{2.38}$$

where $\alpha = \pi/\beta$ and the coordinates u and v are

$$\begin{aligned}
 u &= r + s & r &= u(1 - v) \\
 v &= s/(r + s) & s &= uv.
 \end{aligned} \tag{2.39}$$

The coordinates u and v are illustrated in figure 2.17b. The coordinate u corresponds to the distance from the corner and the coordinate v corresponds to the angle relative to one wall. In this work the mesh generator found all of the sharp corners in the mesh and computed the corner angle β for each sharp corner. This information was made available to the field solver so that each element touching a corner could be a special corner element with the appropriate $\alpha = \pi/\beta$.

Note that only elements adjacent to a corner are affected. All other elements are still the usual quadratic element described earlier in this chapter. The special

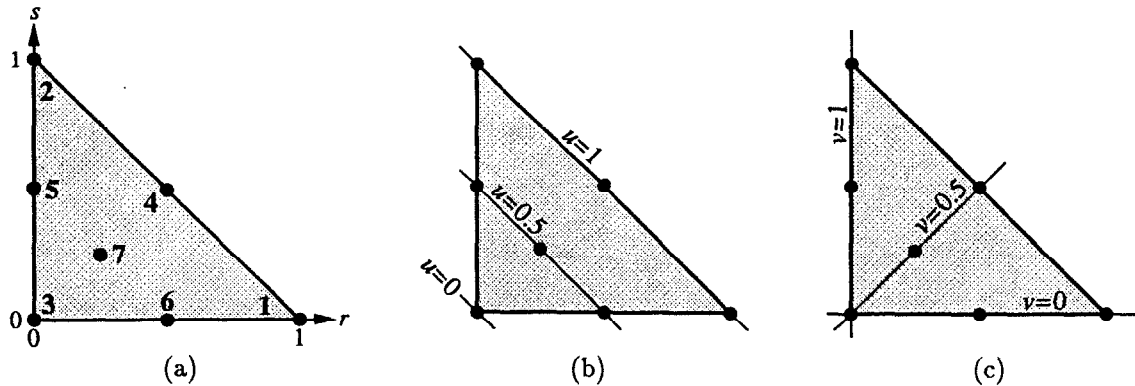


Figure 2.17: The master element (a) for the seven node special corner element. Three lines of constant u (b) and constant v (c) on the master element are shown. Lines of constant v intersect at the sharp corner, node 3. The coordinate u corresponds to a distance from the sharp corner, while the coordinate v corresponds to an angle around the sharp corner. The basis functions for the special corner element are written in (u, v) coordinates.

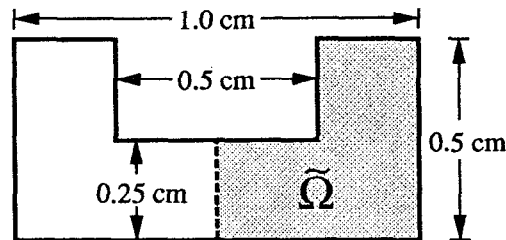


Figure 2.18: Ridged waveguide test problem. Symmetry reduces the problem by one-half.

corner elements are constructed to match with normal quadratic elements on the edge opposite the corner. The special corner element should be adjacent to only one sharp corner, thus the corner angle β is uniquely defined for the element.

The 12-point integration scheme described earlier has poor numerical accuracy when used for special corner elements. This is due to the singular derivative at the corner. An adaptive numerical integration scheme is used instead to obtain the element matrices A^e and B^e . While this adaptive scheme takes much more computer time, only the few corner elements in the mesh require this scheme. The remaining elements can be integrated using the 12-point scheme.

2.4.2. Tests on a Ridged Waveguide

As a test case consider the ridged waveguide shown in figure 2.18. The waveguide is 1.0 cm wide and 0.5 cm high. The ridge is 0.5 cm wide and 0.25 cm high. There are two sharp corners in this problem. However, exploiting the symmetry of this structure reduces the problem so that it contains a single sharp corner.

Figure 2.19 shows the results from three different algorithms for the lowest

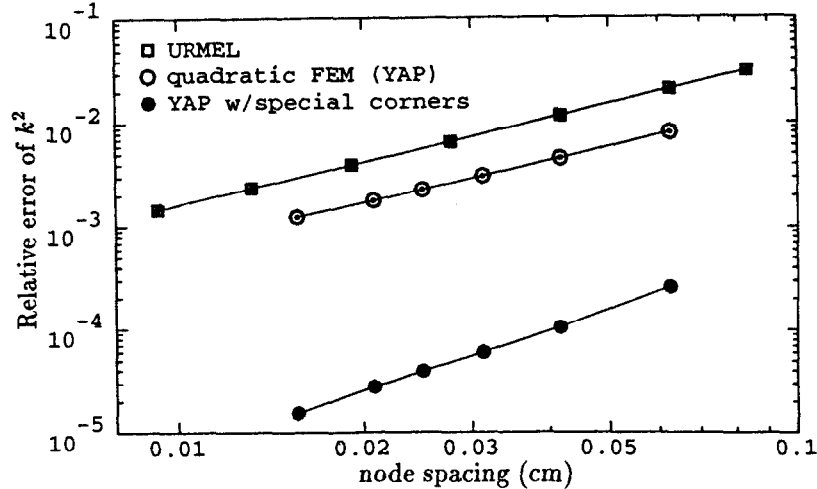


Figure 2.19: Relative error of the computed eigenvalue k^2 versus mesh refinement for the lowest TE mode of the ridged waveguide test problem.

TE mode of the ridged waveguide. Extrapolation of the YAP results which employed special corner elements to an infinitely refined mesh yields the eigenvalue $k^2 = 5.06016 \text{ cm}^{-2}$. This is the “exact” eigenvalue used to compute the relative error in figure 2.19.

For this structure the behavior of URMEL and YAP without special corner elements is similar. YAP is only a factor of two more accurate than URMEL. For both of these cases the dependence of the error on mesh size is $\mathcal{O}(h^{4/3})$. This is worse than the $\mathcal{O}(h^2)$ error exhibited by URMEL for the parallel plate test, and it is much worse than the $\mathcal{O}(h^4)$ error exhibited by YAP in the previous two tests.

Figure 2.19 agrees with the theoretical prediction for structures with sharp corners. The error is expected to be $\mathcal{O}(h^{2\pi/\beta})$ where $\beta > \pi$ is the sharpest corner in the structure. For the ridged waveguide $\beta = 3\pi/2$. Notice that the sharpness of the corner affects the convergence. An infinitely thin plate has corner angle $\beta = 2\pi$ and thus the convergence rate of the field solver would be only $\mathcal{O}(h)$.

The special corner elements improve the accuracy of YAP dramatically. However, while the dependence of the error on mesh size is improved to $\mathcal{O}(h^2)$, it is still short of the $\mathcal{O}(h^4)$ error observed earlier. An improved corner element may be used to recover the $\mathcal{O}(h^4)$ error if desired.

Special treatment for sharp corners is discussed in [7]. They include special basis functions for the singular fields at the corner which cover multiple elements and extend to some fixed distance r_0 from the corner. The special basis functions are not defined in terms of the elements, so there is no assembly process for these basis functions. The treatment of the special corner basis functions is different in

this work. The special corner basis functions are defined locally on the elements and are assembled into global basis functions surrounding the corner. Related work is described in [11]. This technique conforms nicely with the finite element method.

An alternative to special corner elements is a technique called adaptive mesh refinement. This technique uses only ordinary elements, but the mesh is locally refined near the sharp corner since most of the error occurs there. The field solution is significantly improved using a modest number of additional elements near the corner, and algorithms are available to indicate the degree of refinement required near a sharp corner. This technique also conforms nicely with the finite element method.

Chapter 3

Axisymmetric Structures: TM_0 and TE_0 Modes

Many interesting microwave structures are axisymmetric. In particular, accelerator structures made from disk loaded waveguide are axisymmetric. The gain cavities and drift tube of a klystron are also typically axisymmetric.

The eigenmode equations can be written in the cylindrical coordinates ρ , ϕ and z . The azimuthal symmetry of the structure implies that the fields of a mode have azimuthal dependence $e^{im\phi}$ for some integer m . In this chapter an algorithm for computing the monopole ($m = 0$) modes is presented. The following chapter considers modes with azimuthal index $m \neq 0$.

3.1. Wave Equation in Cylindrical Coordinates

Consider the wave equation formulation (1.8) written in cylindrical coordinates. Using the definition $\nabla^2 \mathbf{E} = \nabla(\nabla \cdot \mathbf{E}) - \nabla \times (\nabla \times \mathbf{E})$ and the representation of div, grad and curl in cylindrical coordinates, the wave equations for the three field components E_ρ , E_ϕ and E_z are

$$\nabla^2 E_\rho - \frac{E_\rho}{\rho^2} - \frac{2}{\rho^2} \frac{\partial E_\phi}{\partial \phi} + \epsilon\mu \frac{\omega^2}{c^2} E_\rho = 0, \quad (3.1a)$$

$$\nabla^2 E_\phi - \frac{E_\phi}{\rho^2} + \frac{2}{\rho^2} \frac{\partial E_\rho}{\partial \phi} + \epsilon\mu \frac{\omega^2}{c^2} E_\phi = 0 \quad (3.1b)$$

$$\text{and } \nabla^2 E_z + \epsilon\mu \frac{\omega^2}{c^2} E_z = 0, \quad (3.1c)$$

where the laplacian ∇^2 acting on a scalar field E in cylindrical coordinates is

$$\nabla^2 E = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial E}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 E}{\partial \phi^2} + \frac{\partial^2 E}{\partial z^2}. \quad (3.2)$$

Notice that the field components E_ρ and E_ϕ are coupled in the first two wave equations.

Now assume the azimuthal dependence of the fields is $e^{im\phi}$ for some chosen m . This reduces the problem from a three dimensional domain to a two dimensional domain in the z - ρ plane. Then the wave equations for the three field components are

$$\nabla_t^2 E_\rho - \frac{m^2 + 1}{\rho^2} E_\rho - \frac{2im}{\rho^2} E_\phi + \epsilon\mu \frac{\omega^2}{c^2} E_\rho = 0, \quad (3.3a)$$

$$\nabla_t^2 E_\phi - \frac{m^2 + 1}{\rho^2} E_\phi + \frac{2im}{\rho^2} E_\rho + \epsilon\mu \frac{\omega^2}{c^2} E_\phi = 0 \quad (3.3b)$$

$$\text{and } \nabla_t^2 E_z - \frac{m^2}{\rho^2} E_z + \epsilon\mu \frac{\omega^2}{c^2} E_z = 0, \quad (3.3c)$$

where ∇_t^2 denotes the z - ρ portion of the laplacian,

$$\nabla_t^2 E = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial E}{\partial \rho} \right) + \frac{\partial^2 E}{\partial z^2}. \quad (3.4)$$

In the three wave equations of (3.3) the azimuthal dependence $e^{im\phi}$ has been factored out, so the three field components E_ρ , E_ϕ and E_z depend on z and ρ only.

Cylindrical symmetry implies that the unit vectors $\hat{\mathbf{n}}$ normal to the boundaries are of the form $\hat{\mathbf{n}} = \hat{z}n_z + \hat{\rho}n_\rho$. The normal vectors have no azimuthal component. Then the boundary condition $\hat{\mathbf{n}} \times \mathbf{E} = 0$ at perfectly conditions walls becomes

$$E_\phi = 0 \quad \text{and} \quad E_z n_\rho - E_\rho n_z = 0 \quad \text{on } \Gamma_{\text{metal}}. \quad (3.5)$$

Notice that the boundary condition connects the field components E_z and E_ρ . Also, for $m \neq 0$, the wave equations of (3.3) for E_ρ and E_ϕ are coupled. All three components of the field are coupled. The problem is not easily reduced to a single field component.

When $m = 0$ the three wave equations are uncoupled. While the boundary condition still couples E_z and E_ρ , the azimuthal electric field E_ϕ can be found independently of E_z and E_ρ . Solutions can be obtained for the field E_ϕ with $E_z = E_\rho = 0$. These solutions are the TE₀ modes, where the subscript 0 denotes $m = 0$. According to Maxwell's curl equations, the TE₀ modes will have magnetic field \mathbf{H} in the z - ρ plane and the azimuthal magnetic field will be $H_\phi = 0$.

Let $\tilde{\Omega}$ be the cross-section of the structure in the z - ρ plane. That is, spinning the region $\tilde{\Omega}$ about the axis sweeps out the interior Ω of the microwave structure. The strong formulation for the TE₀ modes is: given the region $\tilde{\Omega}$, constant scalar material properties ϵ and μ , and possibly a phase advance ψ , find the eigenmode fields E_ϕ and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla_t^2 E_\phi - \frac{E_\phi}{\rho^2} + \epsilon\mu \frac{\omega^2}{c^2} E_\phi = 0 \quad \text{in } \tilde{\Omega}, \quad (3.6a)$$

$$E_\phi = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym(cond)}}, \quad (3.6b)$$

$$\hat{\mathbf{n}} \cdot \nabla_t E_\phi = 0 \quad \text{on } \Gamma_{\text{sym(ins)}}, \quad (3.6c)$$

$$E_\phi|_{\hat{\mathbf{R}}\mathbf{x}} = E_\phi|_{\mathbf{x}} e^{i\psi} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left}}, \quad (3.6d)$$

$$\hat{\mathbf{n}} \cdot \nabla_t E_\phi|_{\hat{\mathbf{R}}\mathbf{x}} = -\hat{\mathbf{n}} \cdot \nabla_t E_\phi|_{\mathbf{x}} e^{i\psi} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left}}, \quad (3.6e)$$

$$\Im m E_\phi = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (3.6f)$$

$$\Re e \{ (\hat{\mathbf{n}} \cdot \nabla_t) E_\phi \} = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (3.6g)$$

$$\Im m \{ E_\phi e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}}, \quad (3.6h)$$

$$\text{and } \Re e \{ (\hat{\mathbf{n}} \cdot \nabla_t) E_\phi e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}}, \quad (3.6i)$$

where $\nabla_t E_\phi$ denotes the z - ρ components of the gradient of the scalar field E_ϕ ,

$$\nabla_t E_\phi = \hat{z} \frac{\partial E_\phi}{\partial z} + \hat{\rho} \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho E_\phi). \quad (3.7)$$

Boundary conditions for exploiting various symmetries are included. Note that the symmetry planes are necessarily planes of constant z , so $\hat{\mathbf{n}} = \pm \hat{z}$ for Γ_{sym} , $\Gamma_{\text{sym-left}}$ and $\Gamma_{\text{sym-right}}$. The corresponding boundary conditions become

$$\frac{\partial E_\phi}{\partial z} = 0 \quad \text{on } \Gamma_{\text{sym(ins)}}, \quad (3.8a)$$

$$\Re e \frac{\partial E_\phi}{\partial z} = 0 \quad \text{on } \Gamma_{\text{sym-left}} \quad (3.8b)$$

$$\text{and } \Re e \left\{ \frac{\partial E_\phi}{\partial z} e^{-i\psi/2} \right\} = 0 \quad \text{on } \Gamma_{\text{sym-right}}. \quad (3.8c)$$

The divergence of the electric field is

$$\nabla \cdot \mathbf{E} = \frac{1}{\rho} \frac{\partial E_\phi}{\partial \phi} = 0 \quad (3.9)$$

so the field automatically satisfies the zero-divergence condition in formulation (1.8).

The wave equation formulation (1.9) for the magnetic field \mathbf{H} can be treated similarly. Assuming the fields are constant with respect to ϕ , the wave equation and boundary conditions for the azimuthal field H_ϕ are independent of the fields H_z and H_ρ . Solutions for H_ϕ with $H_z = H_\rho = 0$ yield the TM_0 modes. These modes have electric field in the z - ρ plane and no azimuthal electric field E_ϕ .

The strong formulation for the TM_0 modes is: given the region $\tilde{\Omega}$, constant scalar material properties ϵ and μ , and possibly a phase advance ψ , find the eigenmode fields H_ϕ and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla_t^2 H_\phi - \frac{H_\phi}{\rho^2} + \epsilon \mu \frac{\omega^2}{c^2} H_\phi = 0 \quad \text{in } \tilde{\Omega}, \quad (3.10a)$$

$$\hat{\mathbf{n}} \cdot \nabla_t H_\phi = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym(cond)}}, \quad (3.10b)$$

$$H_\phi = 0 \quad \text{on } \Gamma_{\text{sym(ins)}}, \quad (3.10c)$$

$$H_\phi \Big|_{\hat{\mathbf{R}}\mathbf{x}} = H_\phi \Big|_{\mathbf{x}} e^{i\psi} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left}}, \quad (3.10d)$$

$$\hat{\mathbf{n}} \cdot \nabla_t H_\phi \Big|_{\hat{\mathbf{R}}\mathbf{x}} = -\hat{\mathbf{n}} \cdot \nabla_t H_\phi \Big|_{\mathbf{x}} e^{i\psi} \quad \text{for } \mathbf{x} \in \Gamma_{\text{left}}, \quad (3.10e)$$

$$\Im m H_\phi = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (3.10f)$$

$$\Re e \{ (\hat{\mathbf{n}} \cdot \nabla_t) H_\phi \} = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (3.10g)$$

$$\Im m \{ H_\phi e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}} \quad (3.10h)$$

$$\text{and } \Re e \{ (\hat{\mathbf{n}} \cdot \nabla_t) H_\phi e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}}. \quad (3.10i)$$

Concerning field transformation properties, the magnetic field H_ϕ is treated as a scalar field in this formulation, not as a pseudoscalar field. This is just like the electric field E_ϕ in the TE₀ mode formulation (3.6) above. The only difference between the two formulations is that the boundary conditions for perfectly conducting and perfectly insulating surfaces are exchanged.

3.2. Variational Formulation

Multiplying (3.6a) by a test function F_ϕ^* and a weight ρ and then integrating over the domain $\tilde{\Omega}$ leads to a variational formulation. Integrating by parts yields the following symmetric variational formulation for the TE₀ modes: given the region $\tilde{\Omega}$, constant scalar material properties ϵ and μ , and possibly a phase advance ψ , find the eigenmode fields $E_\phi \in \mathcal{U}_E$ and the corresponding eigenvalues ω^2/c^2 such that for all test functions $F_\phi \in \mathcal{U}_E$

$$\Re \int_{\tilde{\Omega}} \rho \nabla_\perp F_\phi^* \cdot \nabla_\perp E_\phi + \frac{1}{\rho} F_\phi^* E_\phi - \epsilon \mu \frac{\omega^2}{c^2} \rho F_\phi^* E_\phi d\Omega = 0 \quad (3.11a)$$

where the space \mathcal{U}_E of test and trial functions is

$$\begin{aligned} \mathcal{U}_E = \{ & E_\phi \in \mathcal{H}_1(\tilde{\Omega}) : \\ & E_\phi = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym}(\text{cond})}, \\ & E_\phi|_{\hat{\mathbf{R}}\mathbf{x}} = E_\phi|_{\mathbf{x}} e^{i\psi} \quad \forall \mathbf{x} \in \Gamma_{\text{left}}, \\ & \Im m E_\phi = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \\ & \Im m \{ E_\phi e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}} \} \end{aligned} \quad (3.11b)$$

and

$$\mathcal{H}_1(\tilde{\Omega}) = \{ E_\phi: \tilde{\Omega} \rightarrow \mathbb{C} : \int_{\tilde{\Omega}} \rho \nabla_\perp E_\phi^* \cdot \nabla_\perp E_\phi + \frac{1}{\rho} E_\phi^* E_\phi d\Omega \text{ exists and is finite} \}. \quad (3.11c)$$

The notation $\nabla_\perp E_\phi$ denotes the gradient of the scalar field E_ϕ treating z and ρ as ordinary cartesian coordinates instead of cylindrical coordinates,

$$\nabla_\perp E_\phi = \hat{\mathbf{z}} \frac{\partial E_\phi}{\partial z} + \hat{\rho} \frac{\partial E_\phi}{\partial \rho}. \quad (3.12)$$

This gradient is different from $\nabla_t E_\phi$ in (3.7). This variational formulation for the TE₀ modes is equivalent to the strong formulation (3.6) provided the fields are sufficiently smooth. Condition (3.11c) implies $E_\phi = 0$ at $\rho = 0$. This can be considered a boundary condition for the axis.

The corresponding variational formulation for TM₀ modes is: given the region $\tilde{\Omega}$, constant scalar material properties ϵ and μ , and possibly a phase advance ψ , find

the eigenmode fields $H_\phi \in \mathcal{U}_H$ and the corresponding eigenvalues ω^2/c^2 such that for all test functions $G_\phi \in \mathcal{U}_E$

$$\begin{aligned} \Re \left\{ \int_{\Gamma_{\text{metal}}} n_\rho G_\phi^* H_\phi d\Gamma + \int_{\tilde{\Omega}} \rho \nabla_\perp G_\phi^* \cdot \nabla_\perp H_\phi \right. \\ \left. + \frac{1}{\rho} G_\phi^* H_\phi - \epsilon \mu \frac{\omega^2}{c^2} \rho G_\phi^* H_\phi d\Omega \right\} = 0 \end{aligned} \quad (3.13a)$$

where the space of test and trial functions is

$$\begin{aligned} \mathcal{U}_H = \{ H_\phi \in \mathcal{H}_1(\tilde{\Omega}) : \\ H_\phi = 0 \quad \text{on } \Gamma_{\text{sym(ins)}}, \\ H_\phi|_{\hat{\mathbf{R}}\mathbf{x}} = H_\phi|_{\mathbf{x}} e^{i\psi} \quad \forall \mathbf{x} \in \Gamma_{\text{left}}, \\ \Im m H_\phi = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \\ \Im m \{ H_\phi e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}} \} \end{aligned} \quad (3.13b)$$

and

$$\begin{aligned} \mathcal{H}_1(\tilde{\Omega}) = \{ H_\phi : \tilde{\Omega} \rightarrow \mathbb{C} : \int_{\Gamma_{\text{metal}}} n_\rho H_\phi^* H_\phi d\Gamma \\ + \int_{\tilde{\Omega}} \rho \nabla_\perp H_\phi^* \cdot \nabla_\perp H_\phi + \frac{1}{\rho} H_\phi^* H_\phi d\Omega \text{ exists and is finite} \}. \end{aligned} \quad (3.13c)$$

Notice the presence of two integrals in (3.13a). The first integral is a line integral over the perfectly conducting boundaries Γ_{metal} of the problem. This integral is the boundary term from the integration by parts. In the TE_0 mode formulation this boundary term vanished, but in the TM_0 mode formulation the boundary term remains. The integrand includes the factor n_ρ , the radial component of the unit vector normal to the metal wall. The second integral is the same as integral (3.11a) for the TE_0 mode variational formulation.

3.3. Finite Element Formulation

The finite element formulation of the eigenmode problem is constructed in a manner similar to the planar eigenmode problem of the previous chapter. A Galerkin formulation of the TE_0 mode problem is constructed by restricting the variational formulation to a finite dimensional subspace $\mathcal{U}_E^h \subset \mathcal{U}_E$ spanned by a linearly independent set of basis functions N_1, \dots, N_M . The field $E_\phi^h \in \mathcal{U}_E^h$ is expressed as a linear combination of the basis functions,

$$E_\phi^h(z, \rho) = \sum_{i=1}^M u_i N_i(z, \rho). \quad (3.14)$$

The problem reduces to a generalized algebraic eigenvalue problem

$$\mathbf{A}\mathbf{u} - \epsilon\mu\frac{\omega_h^2}{c^2}\mathbf{B}\mathbf{u} = 0 \quad (3.15)$$

with eigenvectors \mathbf{u} and eigenvalues $\lambda_h = \epsilon\mu\omega_h^2/c^2$. The components of the matrices \mathbf{A} and \mathbf{B} are

$$A_{ij} = \Re e \int_{\tilde{\Omega}} \rho \nabla_{\perp} N_i^* \cdot \nabla_{\perp} N_j + \frac{1}{\rho} N_i^* N_j d\Omega \quad (3.16a)$$

$$\text{and } B_{ij} = \Re e \int_{\tilde{\Omega}} \rho N_i^* N_j d\Omega. \quad (3.16b)$$

The matrices \mathbf{A} and \mathbf{B} are symmetric and positive definite, so the eigenvalues λ_h are real and positive. A good set of basis functions will provide approximate eigenvalues λ_h and approximate eigenmode fields E_{ϕ}^h which are close to the true eigenvalues $\lambda = \epsilon\mu\omega^2/c^2$ and eigenmodes E_{ϕ} of the problem.

The Galerkin formulation for the TM_0 mode problem is the same as the TE_0 mode formulation except the basis functions span a finite dimensional subspace $\mathcal{U}_H^h \subset \mathcal{U}_H$ and the components of the matrix \mathbf{A} are

$$A_{ij} = \Re e \left\{ \int_{\Gamma_{\text{metal}}} n_{\rho} N_i^* N_j d\Gamma + \int_{\tilde{\Omega}} \rho \nabla_{\perp} N_i^* \cdot \nabla_{\perp} N_j + \frac{1}{\rho} N_i^* N_j d\Omega \right\}. \quad (3.17)$$

The matrix \mathbf{A} is still symmetric and positive definite.

The triangle elements described in section 2.3 can be used for this problem. This work used mainly the quadratic elements of section 2.3.3. As in the planar problem, the large sparse matrices \mathbf{A} and \mathbf{B} were constructed by computing small dense element matrices \mathbf{A}^e and \mathbf{B}^e and then assembling the element matrices into the global matrices \mathbf{A} and \mathbf{B} .

The $m = 0$ calculations are only slightly more complicated than the cylindrical waveguide calculations. The $m = 0$ calculations have additional terms and factors to integrate when constructing the element matrices, and in TM_0 mode calculations the elements next to metal walls need to include the line integral in \mathbf{A}^e .

3.4. Tests

Consider two test problems. The first test problem is the cutoff mode of circular waveguide. The waveguide has radius a and is treated as a periodic structure with period l as shown in figure 3.1. The cutoff eigenvalue for the TE_{01} mode is $k^2 = 14.68197064/a^2$, and the cutoff eigenvalue for the TM_{01} mode is $k^2 = 5.783185964/a^2$. In this work the cutoff modes are easily obtained by choosing the phase advance to be $\psi = 0$. That is, the fields are required to have period l just

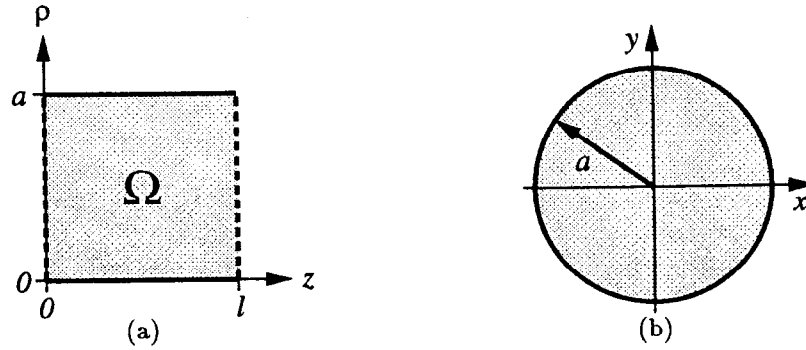


Figure 3.1: The region $\tilde{\Omega}$ for circular waveguide with radius a treated as a periodic structure with length l . The z - ρ (a) and x - y (b) cross-sections are shown. The dashed lines are quasi-periodic boundaries.

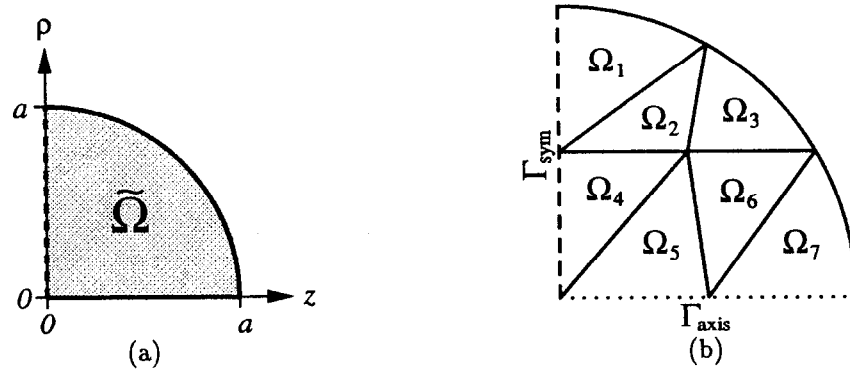


Figure 3.2: The region $\tilde{\Omega}$ (a) for a sphere with radius a . The dashed line is a symmetry plane. A coarse finite element mesh (b) for the sphere.

like the model of the structure. The lowest mode will have fields which are uniform along z . Notice that the domain $\tilde{\Omega}$ for the circular waveguide is rectangular in the z - ρ plane.

The second test problem is a sphere with radius a as shown in figure 3.2a. The domain $\tilde{\Omega}$ is not rectangular, so this is expected to be a more stringent test of the field solvers. The frequency of the TE_{110} mode is $k^2 = 20.19072856/a^2$, and the frequency of the TM_{110} mode is $k^2 = 7.527929583/a^2$. The modes of a sphere are discussed further in section 4.4.3. In particular, note that the third index of a spherical mode is the azimuthal index m . Figure 3.2b shows a coarse finite element mesh for a sphere composed of quadratic triangle elements.

The lowest TE modes for both test problems with $a = 1$ cm were computed using YAP, which employs the finite element method with quadratic elements. The relative error of the eigenvalues k^2 are shown in figure 3.3. In both cases the program converges smoothly toward the correct eigenvalue as the mesh is refined. The error

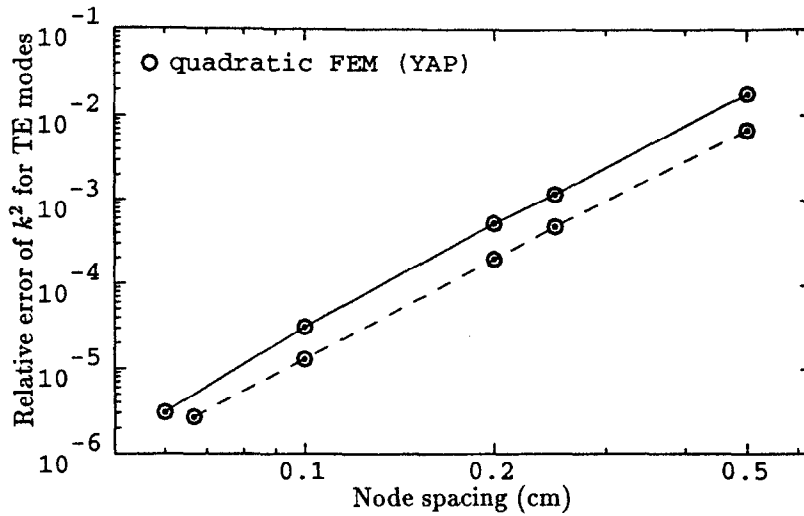


Figure 3.3: Relative error of the computed eigenvalues k^2 for the TE_{110} mode of a sphere (solid line) and the cutoff TE_{01} mode of circular waveguide (dashed line).

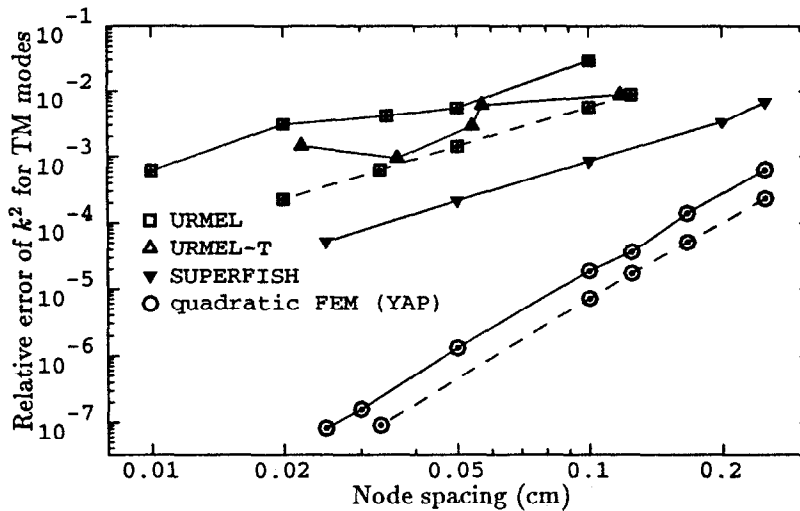


Figure 3.4: Relative error of the computed eigenvalues k^2 for the TM_{110} mode of a sphere (solid line) and the cutoff TM_{01} mode of circular waveguide (dashed line).

is $\mathcal{O}(h^4)$, where h corresponds to the size of the elements. This is the same error observed with YAP on planar waveguide problems.

The lowest TM modes were also computed using a number of field solvers. The relative error of the eigenvalues k^2 are shown in figure 3.4. The circular cylindrical waveguide TM_{01} cutoff mode problem is easily solved since the domain is rectangular. Calculations performed by YAP and URMEL for this problem converge smoothly toward the correct eigenvalue. The relative eigenvalue error is $\mathcal{O}(h^2)$ for URMEL and $\mathcal{O}(h^4)$ for YAP.

Both URMEL and URMEL-T have difficulty computing the TM_{110} mode of a sphere. This is understood for URMEL because the domain was required to fit on a regular grid, and thus the boundary of the grid does not approximate the circle well. The difficulty URMEL-T has with a sphere is less well understood. The domain of the grid looks very much like a circle, but the code does not converge smoothly as the mesh is refined. This could be due to the presence of bad triangles in the mesh constructed by URMEL-T. Triangles which contain a large interior angle (approaching 180°) tend to reduce the accuracy of the solutions. One might have expected that URMEL-T should behave like SUPERFISH since both programs are considered finite difference codes based a regular mesh composed of triangles. Unfortunately, figure 3.4 shows this is not the case.

For the sphere problem both of the field solvers SUPERFISH and YAP converge smoothly toward the correct eigenvalue as the mesh is refined. The eigenvalue error is $\mathcal{O}(h^2)$ for SUPERFISH and $\mathcal{O}(h^4)$ for YAP. In both cases extrapolation to zero element size h can be employed to obtain a better estimate for the eigenvalue. It also gives the user extra confidence in the result he or she has obtained.

The finite element method using quadratic elements (YAP) is very accurate. The coarsest mesh used for the sphere problem is the seven element mesh shown in figure 3.2b. The relative accuracy of the eigenvalue computed using this coarse mesh is already better than 0.1%. For the most refined meshes the error is less than 10^{-7} , or 0.1 ppm.

3.5. Sharp Corners

Sharp corners in the domain $\tilde{\Omega}$ of the problem are detrimental to the accuracy of the field solver, for the same reasons sharp corners deteriorated the accuracy of the waveguide solutions. The field E_ϕ or H_ϕ will have a singular derivative at the sharp corner. The field near the sharp corner is similar to the planar case. The same field expansions (2.36) and (2.37) apply to H_ϕ and E_ϕ , respectively, so the special corner elements developed for the planar waveguide problem can be used to enhance the accuracy of the finite element solution for axisymmetric $m = 0$ problems.

Axisymmetric Structures: $m \neq 0$ Modes

In axisymmetric structures the modes with $m \neq 0$ are often needed. For example, the dipole modes are useful for calculating the transverse wakefield in accelerator structures. A related example is a beam position monitor based on measurement of the beam-induced dipole fields of a cavity. This chapter describes a finite element formulation for calculating the $m \neq 0$ modes of an axisymmetric structure.

4.1. Maxwell's Equations

Unlike the $m = 0$ case, when $m \neq 0$ there is no separation of the fields between TE and TM modes. Furthermore, the problem requires at least two field components, so it is not a scalar problem. There are many possible formulations for this problem. For example, URMEL uses a finite difference formulation of the wave equation reduced to two field components: either (E_ρ, E_z) or (B_ρ, B_z) .

The finite element formulations of this chapter are based on the curl-curl formulations (1.6) and (1.7) of the eigenmode problem. For now consider an arbitrary structure in three dimensions and allow all of the boundary conditions described in section 1.2. The reduction of axisymmetric problems to two dimensions will occur in the next section. The formulation for the electric field \mathbf{E} is: given the region Ω , its material properties ϵ and μ , and possibly phase advances ψ or ψ_j (if periodicity is exploited) or an inversion eigenvalue p (if inversion symmetry is exploited), find the eigenmode fields \mathbf{E} and the corresponding eigenvalues ω^2/c^2 such that

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}) = \frac{\omega^2}{c^2} \epsilon \mathbf{E} \quad \text{in } \Omega, \quad (4.1a)$$

$$\hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym}(\text{cond})}, \quad (4.1b)$$

$$\hat{\mathbf{n}} \times (\mu^{-1} \nabla \times \mathbf{E}) = 0 \quad \text{on } \Gamma_{\text{sym}(\text{ins})}, \quad (4.1c)$$

$$\hat{\mathbf{n}} \times \mathbf{E} \Big|_{\hat{\mathbf{P}}_{\mathbf{x}}} = -p \hat{\mathbf{n}} \times \mathbf{E} \Big|_{\mathbf{x}} \quad \forall \mathbf{x} \in \Gamma_A, \quad (4.1d)$$

$$\hat{\mathbf{n}} \times (\mu^{-1} \nabla \times \mathbf{E}) \Big|_{\hat{\mathbf{P}}_{\mathbf{x}}} = p \hat{\mathbf{n}} \times (\mu^{-1} \nabla \times \mathbf{E}) \Big|_{\mathbf{x}} \quad \forall \mathbf{x} \in \Gamma_A, \quad (4.1e)$$

$$\hat{\mathbf{n}} \times \mathbf{E} \Big|_{\hat{\mathbf{R}}_{j\mathbf{x}}} = \mp \mathbf{R}_j \hat{\mathbf{n}} \times \mathbf{E} \Big|_{\mathbf{x}} e^{i\psi_j} \quad \forall \mathbf{x} \in \Gamma_{\text{left-}j}, \quad (4.1f)$$

$$\hat{\mathbf{n}} \times (\mu^{-1} \nabla \times \mathbf{E}) \Big|_{\hat{\mathbf{R}}_{j\mathbf{x}}} = -\mathbf{R}_j \hat{\mathbf{n}} \times (\mu^{-1} \nabla \times \mathbf{E}) \Big|_{\mathbf{x}} \quad \forall \mathbf{x} \in \Gamma_{\text{left-}j}, \quad (4.1g)$$

$$\Im m \{ \hat{\mathbf{n}} \times \mathbf{E} \} = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (4.1h)$$

$$\Re e \{ \hat{\mathbf{n}} \times (\mu^{-1} \nabla \times \mathbf{E}) \} = 0 \quad \text{on } \Gamma_{\text{sym-left}}, \quad (4.1i)$$

$$\Im m \{ \hat{\mathbf{n}} \times \mathbf{E} e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}}, \quad (4.1j)$$

$$\text{and } \Re e \{ \hat{\mathbf{n}} \times (\mu^{-1} \nabla \times \mathbf{E}) e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}}. \quad (4.1k)$$

The formulation for the magnetic field \mathbf{H} is similar. This formulation is called a strong formulation because it requires that some second derivatives of the field exist. In particular, $\nabla \times (\mu^{-1} \nabla \times \mathbf{E})$ is required. The variational formulation will have less demanding requirements on the differentiability of the fields, so the variational formulation is called a weak formulation.

Notice that the condition $\nabla \cdot (\epsilon \mathbf{E}) = 0$ of formulation (1.6) is not present in this formulation. As a consequence of this omission, this formulation has two sets of solutions: physical modes with $\nabla \cdot (\epsilon \mathbf{E}) = 0$ and non-physical (spurious) modes with $\nabla \cdot (\epsilon \mathbf{E}) \neq 0$. If the eigenvalue ω^2/c^2 of a mode is non-zero then the mode is physical since (4.1a) and the vector identity $\nabla \cdot (\nabla \times \mathbf{A}) = 0$ give

$$\nabla \cdot (\epsilon \mathbf{E}) = \nabla \cdot \left(\frac{c^2}{\omega^2} \nabla \times (\mu^{-1} \nabla \times \mathbf{E}) \right) = 0. \quad (4.2)$$

Therefore, all of the spurious modes have eigenvalue $\omega^2/c^2 = 0$. This is unfortunate if static physical modes need to be calculated since they will be mixed with all of the non-physical modes, but in many cases the problem has no static physical modes or the static physical modes are not the modes of interest. Then the desired physical modes are easily identified by their non-zero eigenvalues.

4.2. Variational Formulation

The variational formulation is obtained from the strong formulation multiplied by a test function \mathbf{F} and then integrating over the domain Ω . Integrating by parts using the vector identity $\nabla \cdot ((\nabla \times \mathbf{E}) \times \mathbf{F}) = \mathbf{F} \cdot (\nabla \times (\nabla \times \mathbf{E})) - (\nabla \times \mathbf{E}) \cdot (\nabla \times \mathbf{F})$ yields a symmetric variational formulation. The variational formulation equivalent to (4.1) is: given the region Ω , its material properties ϵ and μ , and possibly phase advances ψ or ψ_j or an inversion eigenvalue p , find the eigenmode fields $\mathbf{E} \in \mathcal{U}_E$ and the corresponding eigenvalues ω^2/c^2 such that for all test functions $\mathbf{F} \in \mathcal{U}_E$,

$$\Re \int_{\Omega} (\nabla \times \mathbf{F}^*) \cdot \mu^{-1} (\nabla \times \mathbf{E}) - \frac{\omega^2}{c^2} \mathbf{F}^* \cdot \epsilon \mathbf{E} \, d\Omega = 0 \quad (4.3a)$$

where the space \mathcal{U}_E of test and trial functions is

$$\begin{aligned} \mathcal{U}_E = \{ & \mathbf{E} \in \mathcal{H}_{\text{curl}}(\Omega) : \\ & \hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym}(\text{cond})}, \\ & \hat{\mathbf{n}} \times \mathbf{E} \Big|_{\hat{\mathbf{p}}_{\mathbf{x}}} = -p \hat{\mathbf{n}} \times \mathbf{E} \Big|_{\mathbf{x}} \quad \forall \mathbf{x} \in \Gamma_A, \\ & \hat{\mathbf{n}} \times \mathbf{E} \Big|_{\hat{\mathbf{R}}_j \mathbf{x}} = \mp \mathbf{R}_j \hat{\mathbf{n}} \times \mathbf{E} \Big|_{\mathbf{x}} e^{i\psi_j} \quad \forall \mathbf{x} \in \Gamma_{\text{left-}j}, \\ & \Im m \{ \hat{\mathbf{n}} \times \mathbf{E} \} = 0 \quad \text{on } \Gamma_{\text{sym-left}} \text{ and} \\ & \Im m \{ \hat{\mathbf{n}} \times \mathbf{E} e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}} \} \end{aligned} \quad (4.3b)$$

and

$$\mathcal{H}_{\text{curl}}(\Omega) = \left\{ \mathbf{E}: \Omega \rightarrow \mathbb{C}^3 : \int_{\Omega} (\nabla \times \mathbf{E}^*) \cdot (\nabla \times \mathbf{E}) + \mathbf{E}^* \cdot \mathbf{E} \, d\Omega \text{ exists and is finite} \right\}. \quad (4.3c)$$

The space $\mathcal{H}_{\text{curl}}(\Omega)$ is a Hilbert space.

Now consider an axisymmetric problem. The problem can be written in cylindrical coordinates ρ , ϕ and z with field components E_ρ , E_ϕ and E_z . The solutions in cylindrical coordinates can be decomposed into modes with azimuthal dependence $e^{im\phi}$. Choose an m and restrict the solutions to have azimuthal dependence $e^{im\phi}$. Factor the azimuthal dependence out of the field components so E_ρ , E_ϕ and E_z depend on z and ρ only, and let $\tilde{\Omega}$ be the z - ρ cross-section of the problem. Then the reduced formulation for the electric field \mathbf{E} is: given the region $\tilde{\Omega}$, its material properties ϵ and μ , and possibly a phase advance ψ , find the eigenmode fields $\mathbf{E} \in \tilde{\mathcal{U}}_E$ and the corresponding eigenvalues ω^2/c^2 such that for all test functions $\mathbf{F} \in \tilde{\mathcal{U}}_E$,

$$\Re \int_{\tilde{\Omega}} \left((\nabla \times \mathbf{F}^*) \cdot \mu^{-1} (\nabla \times \mathbf{E}) - \frac{\omega^2}{c^2} \mathbf{F}^* \cdot \epsilon \mathbf{E} \right) 2\pi \rho \, d\rho \, dz = 0 \quad (4.4a)$$

where the space $\tilde{\mathcal{U}}_E$ of test and trial functions is

$$\begin{aligned} \tilde{\mathcal{U}}_E = \{ & \mathbf{E} \in \mathcal{H}_{\text{curl}}(\tilde{\Omega}) : \\ & \hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on } \Gamma_{\text{metal}} \text{ and } \Gamma_{\text{sym}(\text{cond})}, \\ & \hat{\mathbf{n}} \times \mathbf{E} \Big|_{\hat{\mathbf{R}}\mathbf{x}} = \mp \mathbf{R} \hat{\mathbf{n}} \times \mathbf{E} \Big|_{\mathbf{x}} e^{i\psi} \quad \forall \mathbf{x} \in \Gamma_{\text{left}}, \\ & \Im m \{ \hat{\mathbf{n}} \times \mathbf{E} \} = 0 \quad \text{on } \Gamma_{\text{sym-left}} \text{ and} \\ & \left. \Im m \{ \hat{\mathbf{n}} \times \mathbf{E} e^{-i\psi/2} \} = 0 \quad \text{on } \Gamma_{\text{sym-right}} \right\} \end{aligned} \quad (4.4b)$$

and

$$\mathcal{H}_{\text{curl}}(\tilde{\Omega}) = \left\{ \mathbf{E}: \tilde{\Omega} \rightarrow \mathbb{C}^3 : \int_{\tilde{\Omega}} \left((\nabla \times \mathbf{E}^*) \cdot (\nabla \times \mathbf{E}) + \mathbf{E}^* \cdot \mathbf{E} \right) 2\pi \rho \, d\rho \, dz \text{ exists and is finite} \right\}. \quad (4.4c)$$

Notice that multiple quasi-periodic boundary conditions are excluded from this formulation. An axisymmetric problem can have at most one periodicity, and the corresponding symmetry operator must be a translation along the z -axis. Also note that inversion symmetry is not included in this formulation since there is always a plane reflection symmetry which can be used instead. Finally, boundary conditions for fields on the axis are given implicitly in (4.4c). The axis boundary conditions are discussed further in section 4.4.

In terms of the field components,

$$\begin{aligned}
\nabla \times \mathbf{E} &= \hat{\rho} \left(\frac{im}{\rho} E_z - \frac{\partial E_\phi}{\partial z} \right) + \hat{\phi} \left(\frac{\partial E_\rho}{\partial z} - \frac{\partial E_z}{\partial \rho} \right) + \hat{z} \left(\frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho E_\phi) - \frac{im}{\rho} E_\rho \right) \\
&= \frac{i\hat{\rho}}{\rho} \left(m E_z - \frac{\partial \tilde{E}_\phi}{\partial z} \right) + \hat{\phi} \left(\frac{\partial E_\rho}{\partial z} - \frac{\partial E_z}{\partial \rho} \right) + \frac{i\hat{z}}{\rho} \left(\frac{\partial \tilde{E}_\phi}{\partial \rho} - m E_\rho \right) \\
&= \nabla_\perp \times \mathbf{E}_t + \frac{i\hat{\phi}}{\rho} \times (m \mathbf{E}_t - \nabla_\perp \tilde{E}_\phi),
\end{aligned} \tag{4.5}$$

where $\tilde{E}_\phi = -i\rho E_\phi$, $\mathbf{E}_t = \hat{z}E_z + \hat{\rho}E_\rho$ and $\nabla_\perp = \hat{z}\frac{\partial}{\partial z} + \hat{\rho}\frac{\partial}{\partial \rho}$. Some insight can be gained by considering a problem with no media (vacuum). Then equation (4.4a) becomes

$$\begin{aligned}
\Re e \int_{\tilde{\Omega}} \rho (\nabla_\perp \times \mathbf{F}_t)^* \cdot (\nabla_\perp \times \mathbf{E}_t) + \frac{1}{\rho} (m \mathbf{F}_t - \nabla_\perp \tilde{F}_\phi)^* \cdot (m \mathbf{E}_t - \nabla_\perp \tilde{E}_\phi) \\
- \frac{\omega^2}{c^2} (\rho \mathbf{F}_t^* \cdot \mathbf{E}_t + \frac{1}{\rho} \tilde{F}_\phi^* \cdot \tilde{E}_\phi) d\rho dz = 0.
\end{aligned} \tag{4.6}$$

4.3. Finite Element Formulation

The finite element formulation is the variational formulation restricted to a finite dimensional subspace $\tilde{U}_E^h \subset \tilde{U}_E$ with basis functions which are relatively simple functions on each element covering $\tilde{\Omega}$. The choice of basis functions is particularly important for this formulation because of the presence of spurious modes[12]. A poor choice of basis functions will allow the spurious modes to have eigenvalue $\omega^2/c^2 \neq 0$. Obtaining distinct physical and spurious solutions becomes very difficult if not impossible.

For local basis functions to be useful in the finite element method they must be easily assembled into global basis functions which are in the space \tilde{U}_E of trial and test functions. In particular, condition (4.4c) requires that the component of \mathbf{E} tangential to any interface is continuous across the interface. Notice that there is no corresponding condition on the continuity of the component of \mathbf{E} normal to an interface. For the interface at the boundary between two elements, the continuity of E_ϕ across element boundaries is accomplished as in the scalar finite elements of the previous chapters. The adjacent elements share the nodes on the common boundary and thus ensure that the global basis function is continuous. The continuity of the portion of \mathbf{E}_t which is tangential to the element boundary requires more work. It will be necessary to construct vector basis functions which are normal to all but one side of the element. There will be one or more "nodes" which specify the component

of \mathbf{E}_t which is tangential to the element boundary. Adjacent elements will share this “node” in order to construct a global basis function which satisfies condition (4.4c).

The local basis functions described below were inspired by the covariant projection elements[13]–[17] of Crowley and the work of Nedelec[18]. The vector basis functions for \mathbf{E}_t are a basis for the polynomial space \mathcal{R}^k described by Nedelec and are also similar to the elements described in [19]. Other edge elements are described in [20]–[22]. This and the modelling of \tilde{E}_ϕ instead of E_ϕ by lagrange-type basis functions leads to a finite element formulation without spurious mode problems.

A set of linear basis functions is described first. A quadratic basis follows. The elements are triangular elements. Let $\hat{\Omega}$ be the master element and $\mathbf{x}_e : \hat{\Omega} \rightarrow \Omega_e$ be the map from the master element to the e th element Ω_e . The elements are typical linear or quadratic elements, so the map \mathbf{x}_e is simply a linear or quadratic coordinate transformation. These elements were illustrated earlier in figures 2.3 and 2.6.

4.3.1. Linear Edge Elements

The basis functions representing the field \tilde{E}_ϕ are the usual lagrange-type linear basis functions,

$$N_1 = r, \quad N_2 = s \quad \text{and} \quad N_3 = t, \quad (4.7)$$

where $t = 1 - r - s$. The basis functions representing the field \mathbf{E}_t are more complex. Use the derivatives of the element transformation \mathbf{x}_e to obtain vector fields \mathbf{V}_i which have the property that \mathbf{V}_i is tangential to side i ,

$$\mathbf{V}_1 = -(\mathbf{V}_2 + \mathbf{V}_3), \quad \mathbf{V}_2 = -\frac{\partial \mathbf{x}_e}{\partial s} \quad \text{and} \quad \mathbf{V}_3 = \frac{\partial \mathbf{x}_e}{\partial r}. \quad (4.8)$$

The vector fields \mathbf{V}_i are directed in the counterclockwise direction along side i as shown in figure 4.1a. The reciprocal vector fields \mathbf{R}_i are normal to side i ,

$$\mathbf{R}_1 = -(\mathbf{R}_2 + \mathbf{R}_3), \quad \mathbf{R}_2 = \frac{\hat{\phi} \times \mathbf{V}_2}{\hat{\phi} \cdot (\mathbf{V}_2 \times \mathbf{V}_3)} \quad \text{and} \quad \mathbf{R}_3 = \frac{\hat{\phi} \times \mathbf{V}_3}{\hat{\phi} \cdot (\mathbf{V}_2 \times \mathbf{V}_3)}. \quad (4.9)$$

The reciprocal vector field \mathbf{R}_i is directed into the element along side i as shown in figure 4.1b. The reciprocal vectors can also be written as

$$\mathbf{R}_1 = \left(\frac{\partial t}{\partial z}, \frac{\partial t}{\partial \rho} \right), \quad \mathbf{R}_2 = \left(\frac{\partial r}{\partial z}, \frac{\partial r}{\partial \rho} \right) \quad \text{and} \quad \mathbf{R}_3 = \left(\frac{\partial s}{\partial z}, \frac{\partial s}{\partial \rho} \right), \quad (4.10)$$

which states that the reciprocal vectors are derivatives of \mathbf{x}_e^{-1} , the inverse of the element transformation. Now a set of local vector basis functions representing the

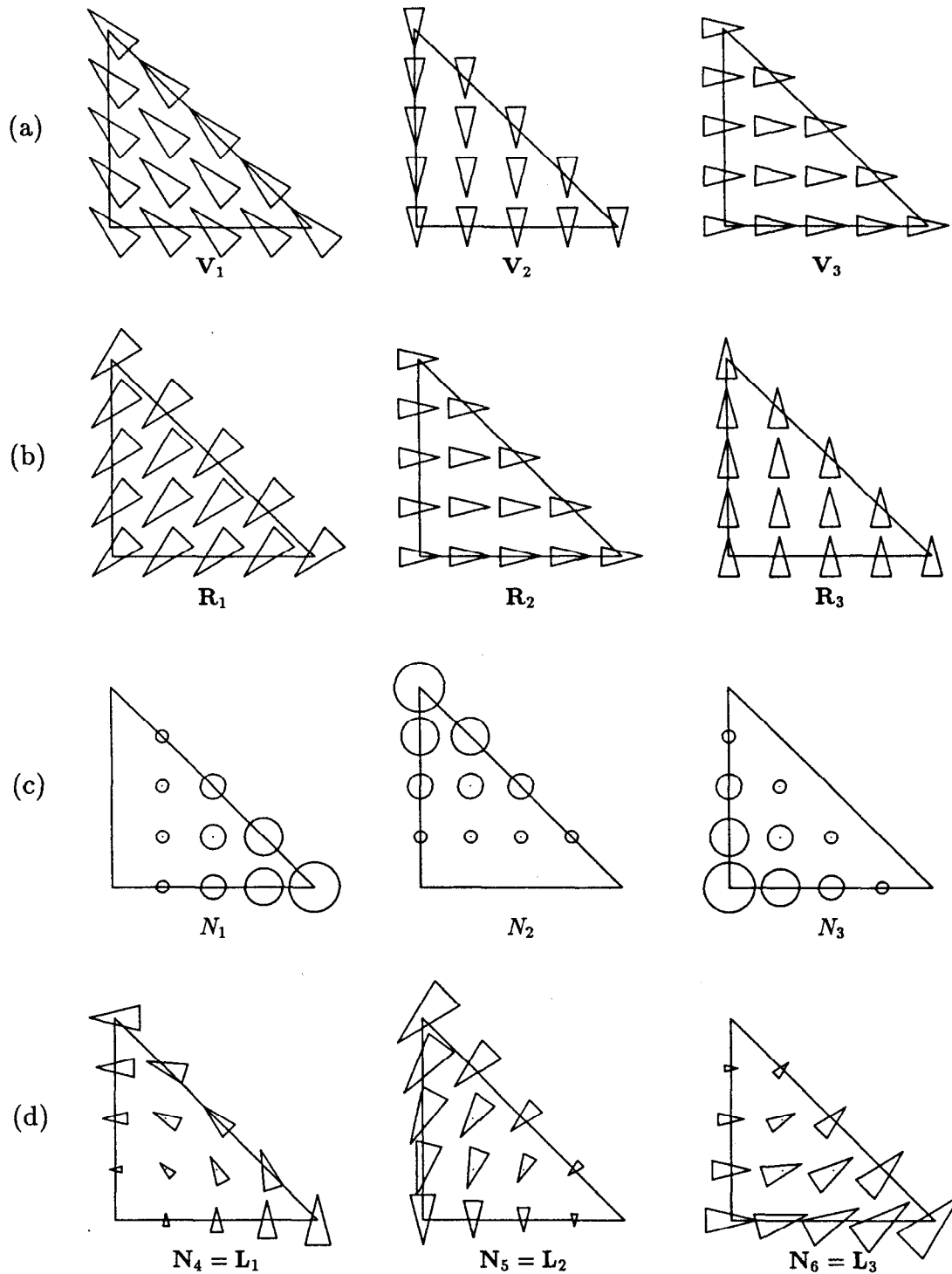


Figure 4.1: Vector fields V_i (a) tangential to side i , reciprocal vector fields R_i (b) normal to side i , local basis functions N_i (c) representing \tilde{E}_ϕ and local vector basis functions N_i (d) representing E_i . The triangles represent the vector field in the plane of the page: the size of the triangle is proportional to the magnitude of the field and the triangle points in the direction of the field.

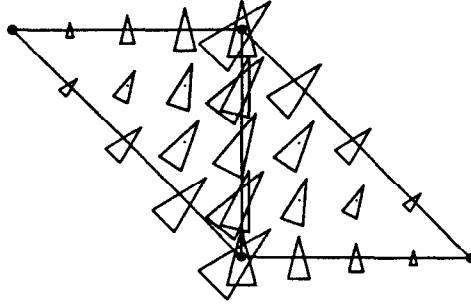


Figure 4.2: A global vector basis function comprising one local vector basis function on each of two elements. The field tangential to the element boundary is continuous while the field normal to the boundary is discontinuous.

field \mathbf{E}_t are constructed,

$$\begin{aligned}
 \mathbf{N}_4 &= \mathbf{L}_1 = r\mathbf{R}_3 - s\mathbf{R}_2, \\
 \mathbf{N}_5 &= \mathbf{L}_2 = s\mathbf{R}_1 - t\mathbf{R}_3 \\
 \text{and } \mathbf{N}_6 &= \mathbf{L}_3 = t\mathbf{R}_2 - r\mathbf{R}_1.
 \end{aligned} \tag{4.11}$$

The local vector basis functions are easily assembled into global basis functions because they are normal to all but one side of the element. In particular, the basis functions satisfy $\mathbf{L}_i \cdot \mathbf{V}_j = \delta_{ij}$ on side j of the element. Consider two elements which are adjacent to each other. The tangents \mathbf{V}_j on the common side differ only by a minus sign when computed in one element or the other. The local vector basis functions are directed in the counterclockwise direction as shown in figure 4.1d. However, to construct a global basis function for \mathbf{E}_t spanning two elements, the local basis function must be reversed in one of the elements. This will ensure that the component of \mathbf{E}_t tangential to the element boundary is continuous. An example is shown in figure 4.2.

4.3.2. Quadratic Edge Elements

The local basis functions for the quadratic element are similarly split between \tilde{E}_ϕ and \mathbf{E}_t . The local basis functions for \tilde{E}_ϕ are the quadratic lagrange-type basis functions:

$$\begin{aligned}
 N_1 &= r(2r - 1) & N_2 &= s(2s - 1) & N_3 &= t(2t - 1) \\
 N_4 &= 4rs & N_5 &= 4st & N_6 &= 4rt.
 \end{aligned} \tag{4.12}$$

There are 8 local vector basis functions representing \mathbf{E}_t . They are based on the linear local vector basis functions \mathbf{L}_1 , \mathbf{L}_2 and \mathbf{L}_3 . The local vector basis functions

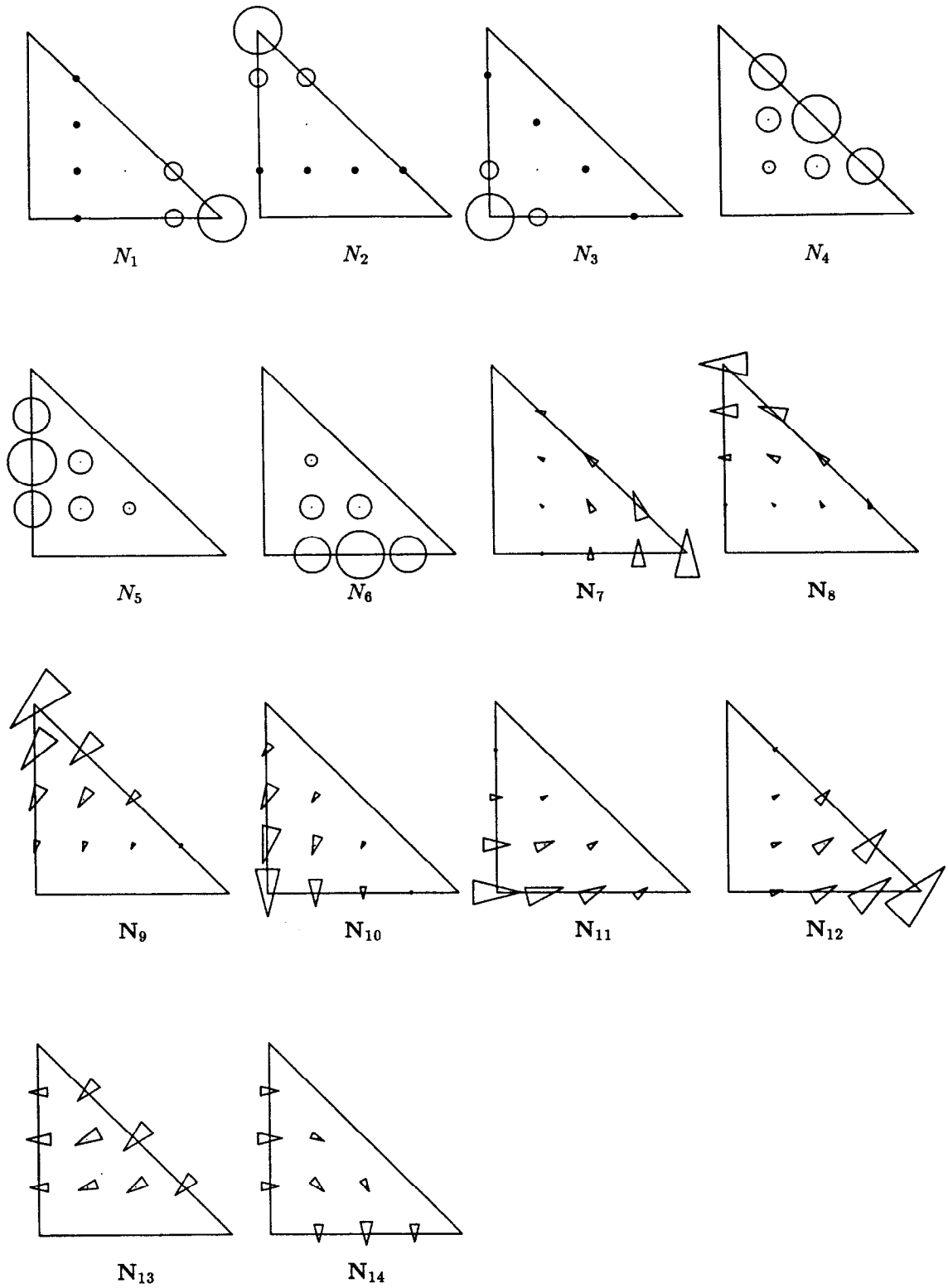


Figure 4.3: Quadratic local basis functions: N_1 through N_6 represent \bar{E}_ϕ , N_7 through N_{12} represent \mathbf{E}_t .

are:

$$\begin{aligned}
\mathbf{N}_7 &= r\mathbf{L}_1 & \mathbf{N}_{10} &= t\mathbf{L}_2 & \mathbf{N}_{13} &= -2s\mathbf{L}_3 \\
\mathbf{N}_8 &= s\mathbf{L}_1 & \mathbf{N}_{11} &= t\mathbf{L}_3 & \mathbf{N}_{14} &= -2t\mathbf{L}_1. \\
\mathbf{N}_9 &= s\mathbf{L}_2 & \mathbf{N}_{12} &= r\mathbf{L}_3
\end{aligned} \tag{4.13}$$

The first six vector basis functions, \mathbf{N}_7 to \mathbf{N}_{12} , have non-zero tangential field on one side of the element. This field must match the tangential field of the adjacent element, if there is one, in order to construct valid global basis functions. The last two vector basis functions, \mathbf{N}_{13} and \mathbf{N}_{14} , have no tangential field on the element sides. Each of the local vector basis functions \mathbf{N}_{13} and \mathbf{N}_{14} by themselves constitute valid global basis functions, thus they are internal degrees of freedom for the element.

4.3.3. Tests on an Annular Ring

As a test case consider the annular ring shown in figure 4.4. Let a and b be the inner and outer radii of the ring, respectively, and let l be the length of the ring. The annular ring problem can be solved analytically by treating it as a section of cylindrical (coaxial) waveguide which is shorted at $z = 0$ and $z = l$. The cross-section of the cylindrical waveguide is shown in figure 4.4b. The characteristic equation for the cutoff wavenumbers γ_m for the TM_m modes is

$$J_m(\gamma_m a)Y_m(\gamma_m b) - Y_m(\gamma_m a)J_m(\gamma_m b) = 0, \tag{4.14}$$

and the cutoff wavenumbers γ'_m for the TE_m modes satisfy

$$J'_m(\gamma'_m a)Y'_m(\gamma'_m b) - Y'_m(\gamma'_m a)J'_m(\gamma'_m b) = 0. \tag{4.15}$$

Let $\gamma_{m,n}$ be the n th solution of equation (4.14) and let $\gamma'_{m,n}$ be the n th solution of equation (4.15). Some approximate numerical values for the cutoff wavenumbers for the case $a = 1/2$, $b = 1$ and various m are listed in table 4.1. These values were obtained by finding the roots of the characteristic equations numerically and are correct to the given number of digits. Note that the $m = 0$ case is included despite the fact that the scalar formulations of the previous chapter are more efficient for $m = 0$ problems than the vector curl-curl formulation of this chapter. The vector curl-curl formulation is able to perform the $m = 0$ calculations so it will be tested.

Let p represent the number of half-wavelengths of field variation along the length l of the waveguide, the frequencies of the TM modes in the annular ring are

$$\text{TM}_{mnp} : k^2 = \gamma_{m,n}^2 + (p\pi/l)^2 \quad p = 0, 1, 2, \dots \tag{4.16}$$

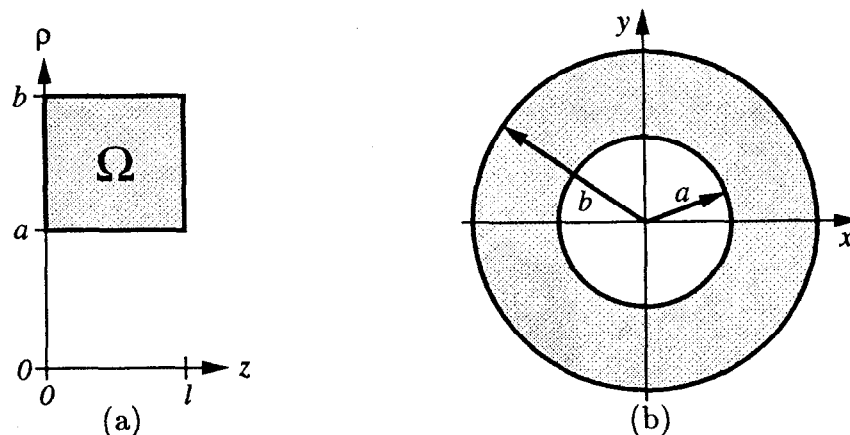


Figure 4.4: The region Ω for an annular ring with inner radius a , outer radius b and length l . The z - ρ (a) and x - y (b) cross-sections are shown.

	azimuthal symmetry number m			
	0	1	2	10
γ_i	6.2460618392 12.5468714280 18.8364150845 25.1228463711	6.3931567616 12.6246990207 18.8889298510 25.1624056202	6.8138428531 12.8555318452 19.0457045399 25.2807623390	14.5023698787 18.8240366105 23.5722353177 28.8469395663
γ'_i	0 6.3931567616 12.6246990207 18.8889298510 25.1624056202	1.3546720103 6.5649423823 12.7064223371 18.9426593061 25.2024870403	2.6812042867 7.0625816160 12.9494113826 19.1031584966 25.3224312292	11.7688573622 16.3031605700 19.6682488849 23.8402201120 28.9613713372

Table 4.1: Some numerical values for the cutoff wavenumbers $\gamma_{m,n}$ (for TM_m modes) and $\gamma'_{m,n}$ (for TE_m modes) of an annular ring with $a = 1/2$ and $b = 1$.

and the frequencies of the TE modes are

$$TE_{mnp} : k^2 = \gamma'_{m,n}{}^2 + (p\pi/l)^2 \quad p = 1, 2, 3, \dots \quad (4.17)$$

Note that the TEM modes of the waveguide are labelled here as the TE_{00} case, which has cutoff wavenumber $\gamma'_{0,0} = 0$.

Figure 4.5 shows five finite element meshes used for the test calculations on an annular ring. The radii are $a = 1/2$ and $b = 1$, and the length is $l = 1/2$. The coarsest mesh consists of only two elements and the finest mesh consists of 512 elements. The tests are performed with quadratic elements.

Fields plots of the lowest 10 modes with $m = 1$ are presented in figure 4.6. The plots were constructed from fields calculated using the finest mesh shown in figure 4.5 which contains 512 quadratic elements. The electric and magnetic fields

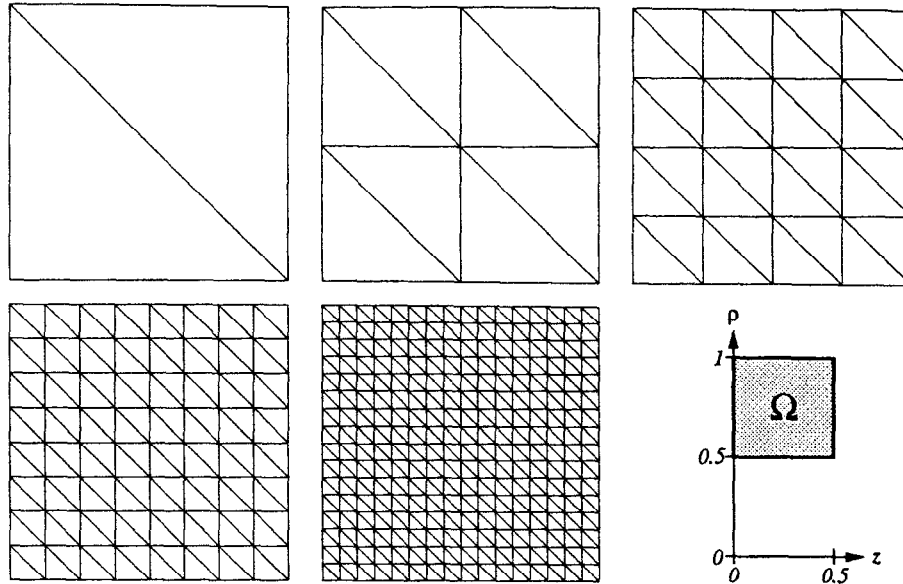


Figure 4.5: A course mesh (two elements) and four refined meshes for the annular ring problem. Each mesh covers the shaded region Ω .

were obtained from two separate calculations. The calculated eigenvalues and the corresponding relative errors for these modes are listed in table 4.2.

One may ask which field is better to use in order to obtain the most accurate answer. For most of the modes the accuracy is the same order of magnitude between the two field calculations, with some modes more accurately calculated using the electric field and a few modes more accurately calculated using the magnetic field. However, a few of the lowest modes (TM_{110} , TE_{111} and TM_{120}) are calculated with more than an order of magnitude better accuracy by using the electric field formulation instead of the magnetic field formulation. The reason for this can be found in the nature of the vector basis functions. The elements model field variation perpendicular to the field better than field variation along the field. For example, the TM_{120} mode has field E_z ($E_\rho = 0$) and the field is uniform along z . The field variation occurs with ρ which is perpendicular to the field. The magnetic field cannot be uniform along the direction of the magnetic field due to the boundary conditions at the walls. For example, look at H_z along the top of the magnetic field plot for the TM_{121} mode. It is sinusoidal (one half period), not uniform. Also notice that H_ρ also has one half period of variation with ρ . This situation is similar to the electric field of the TM_{111} mode, so the accuracies of the two calculations are the same.

Concerning individual modes, one field (the electric field, for example) of the mode is mostly azimuthal while the other field (in this example the magnetic field)

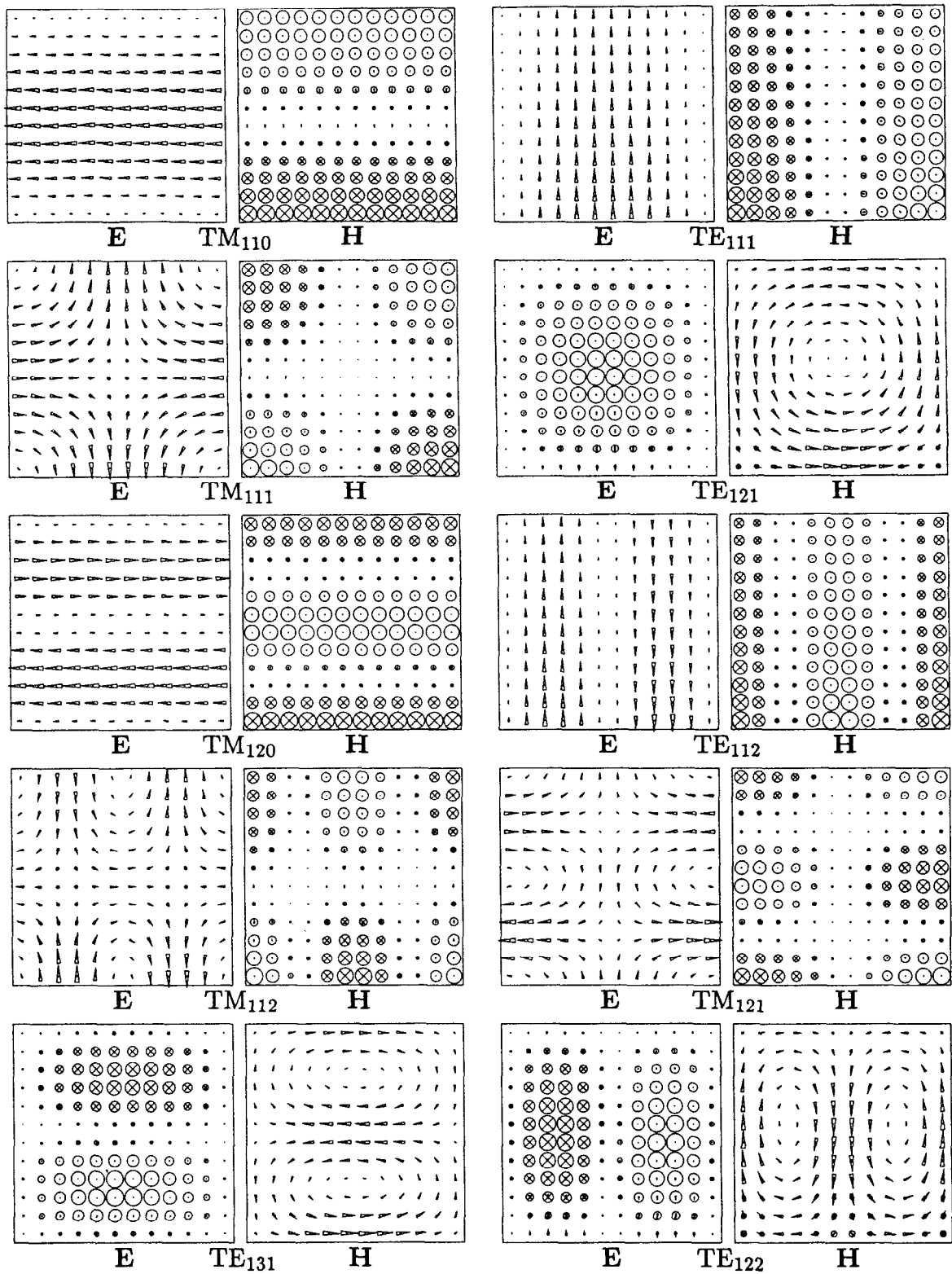


Figure 4.6: Field plots of the 10 lowest $m = 1$ modes of an annular ring. The ϕ -component of the fields is imaginary. The size of the circles are proportional to the ϕ -component of the field at the center of the circle. A cross in the circle indicates the ϕ -component is negative. The triangles indicate the size and direction of the z and ρ components of the field.

mode	k_{analytic}^2	k_{E}^2	relative error	k_{H}^2	relative error
TM ₁₁₀	40.872453	40.872449	-1×10^{-7}	40.872533	2×10^{-6}
TE ₁₁₁	41.313554	41.313557	8×10^{-8}	41.313633	2×10^{-6}
TM ₁₁₁	80.350871	80.351173	4×10^{-6}	80.351967	1×10^{-5}
TE ₁₂₁	82.576886	82.578057	1×10^{-5}	82.577194	4×10^{-6}
TM ₁₂₀	159.383025	159.383102	5×10^{-7}	159.388130	3×10^{-5}
TE ₁₁₂	159.748807	159.749047	2×10^{-6}	159.753861	3×10^{-5}
TM ₁₁₂	198.786124	198.789456	2×10^{-5}	198.799964	7×10^{-5}
TM ₁₂₁	198.861443	198.864440	2×10^{-5}	198.876238	7×10^{-5}
TE ₁₃₁	200.931586	200.946598	7×10^{-5}	200.934637	2×10^{-5}
TE ₁₂₂	201.012139	201.027047	7×10^{-5}	201.015726	2×10^{-5}

Table 4.2: Comparison between E and H calculations for the 10 lowest $m = 1$ modes of an annular ring with $a = 1/2$, $b = 1$ and $l = 1/2$. The mesh consists of 512 quadratic triangular elements.

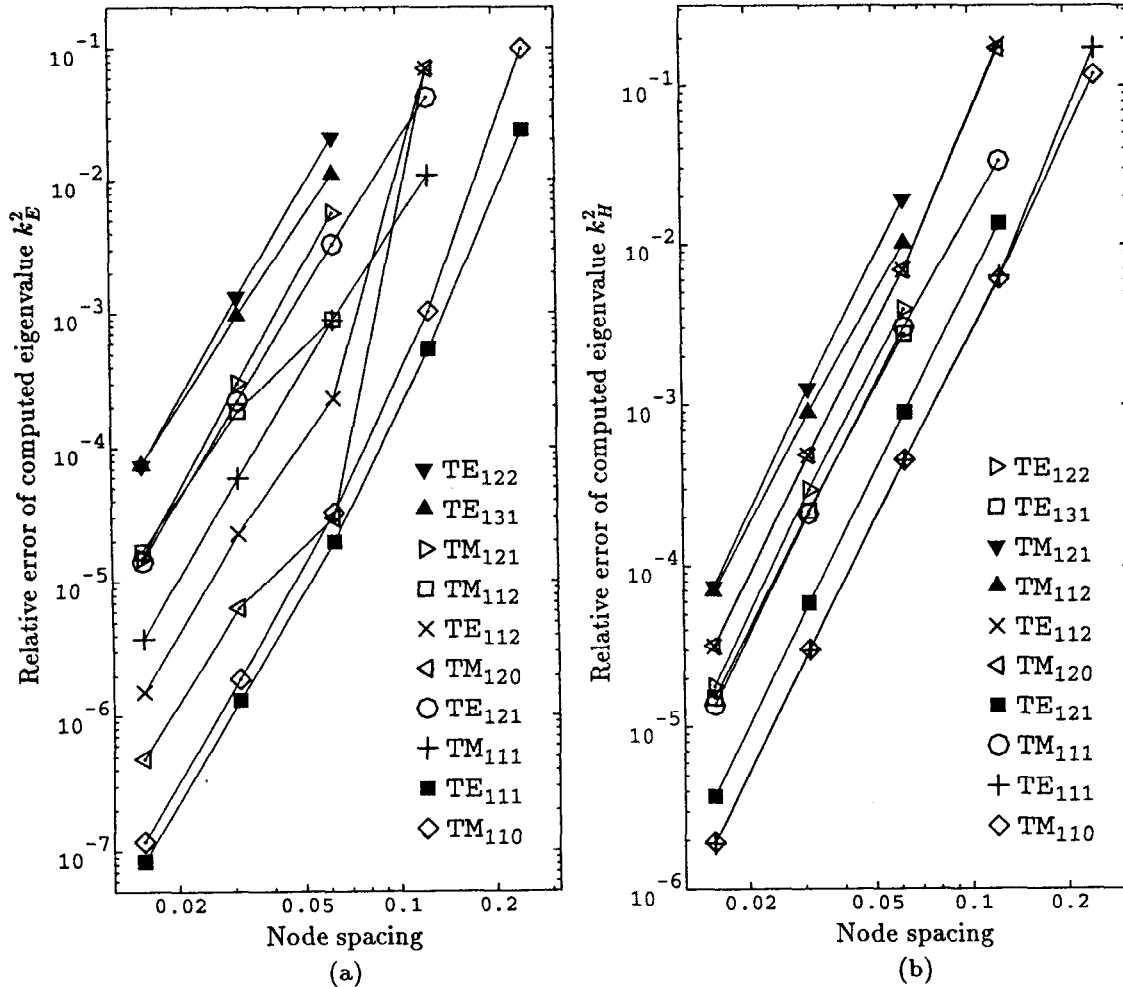


Figure 4.7: Relative error of the computed eigenvalues for the $m = 1$ modes of an annular ring as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} .

has very little azimuthal field. The field calculations which are mostly E_ϕ or H_ϕ are consistently less accurate than their counterparts which are mostly H_t or E_t , respectively. For the best accuracy choose the field which has the smallest azimuthal component.

The convergence of the eigenvalues as the mesh is refined is shown in figure 4.7. The eigenvalue error is $\mathcal{O}(h^4)$ where h is the size of the elements. This is the expected error for quadratic elements. Note that the convergence of the electric field calculations for the lowest mode, TM_{110} , is unusual in the sense that the exact eigenvalue is approached from below as the mesh is refined. This is indicated by the negative error listed in table 4.2. Usually the eigenvalue is overestimated, so the relative error $(k^2 - k_{\text{exact}}^2)/k_{\text{exact}}^2$ is positive and the computed eigenvalues decrease toward the exact eigenvalue as the mesh is refined. But the eigenvalue for the TM_{110} mode is always underestimated and increases toward the exact eigenvalue as the mesh is refined.

The spurious modes expected at $\omega^2/c^2 = 0$ had computed eigenvalues $\omega^2/c^2 \approx 10^{-13}$ for coarse meshes and $\omega^2/c^2 \approx 10^{-12}$ for the finest mesh. The separation between the physical and spurious modes is excellent since the spurious eigenvalues are many orders of magnitude smaller than the physical eigenvalues.

Figure 4.8 shows the mode frequencies as a function of the azimuthal symmetry number m . The TE_m and TM_m modes are plotted separately for clarity. Each curve is labelled by a radial mode index n and an axial index p . The points on the curves are the frequencies of the TE_{mnp} and TM_{mnp} modes, with n and p fixed.

Notice that points at $m = 0.5$, $m = 1.5$ and $m = 2.5$ are included in figure 4.8. The requirement that m be an integer is obtained from the demand that the field be single-valued: $\mathbf{E}(\rho, \phi, z) = \mathbf{E}(\rho, \phi + 2\pi n, z)$. Hence non-integral m are not modes of a truly axisymmetric structure.

However, some structures are partially axisymmetric in the following sense: the boundaries are a surface of revolution between two planes of constant ϕ . The revolution sweeps only a part of a circle, $\phi_1 \leq \phi \leq \phi_2$. Since the field expansion in azimuthal harmonics is intended to be valid over only a limited range of ϕ , there is no requirement that the field be single valued between ϕ and $\phi + 2\pi n$, and thus modes with non-integral m may occur. Chapter 7 describes an application of modes with non-integral m to the design of a 90° waveguide bend.

The dependence on m of the relative error of the computed eigenvalues is plotted in figure 4.9. The eigenvalues were computed using the finest mesh shown in figure 4.5. The relative error of a particular mode does not grow quickly for large m .

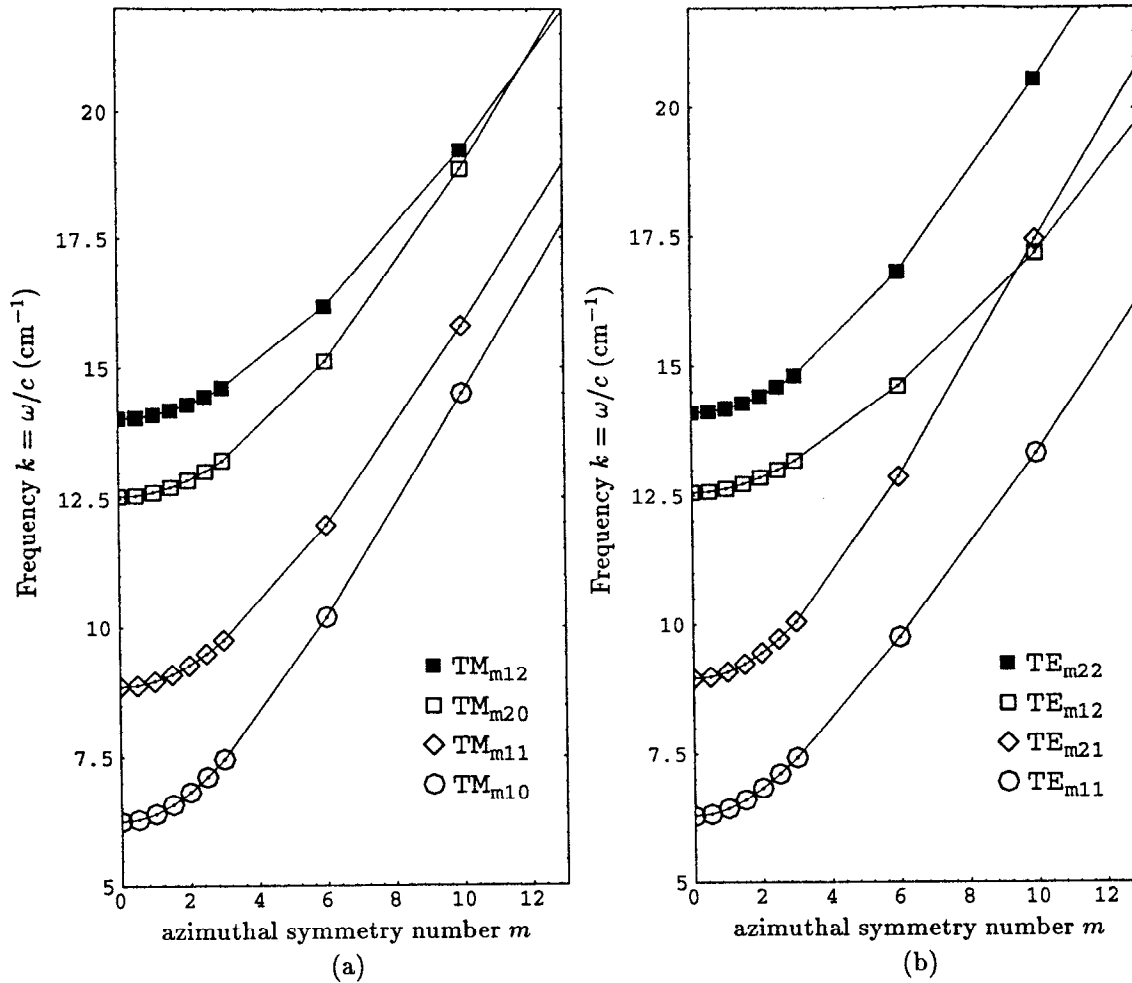


Figure 4.8: Frequency $k = \omega/c$ versus azimuthal symmetry number m for (a) four TM modes and (b) four TE modes.

Hence the field solver may be used effectively for any m , including non-integral m .

The dip in the TE_{m11} curve in figure 4.9c is related to an anomalous change in sign of the error. The error is negative for $m = 0$ and passes through zero near $m = 0.5$. For $m \geq 0.5$ the error is positive, as expected. The TM_{m10} mode error in figure 4.9a has similar anomalous behavior. For $m < 2.8$ the error is negative, while for $m > 2.8$ the error is positive.

The small bump at $m = 0.5$ for the TM_{m12} mode is due to the accidental degeneracy of this mode with the TM_{m21} mode at $m = 0.5$.

4.4. Axis Elements

Elements which touch the axis are restricted so that the global basis functions

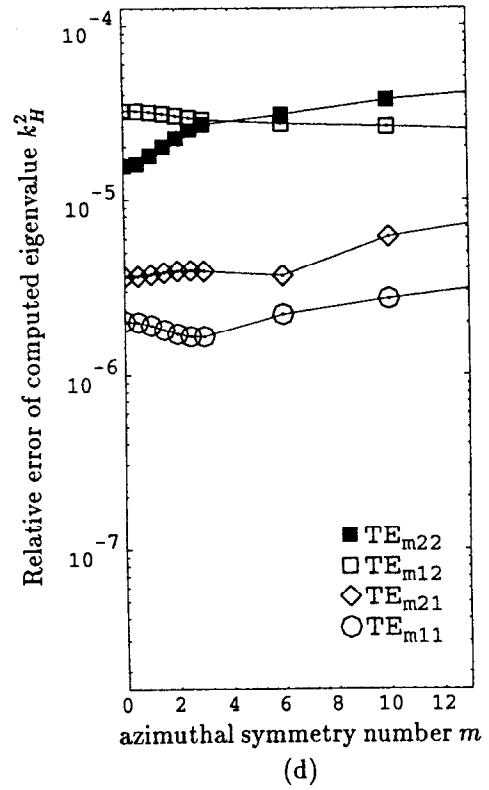
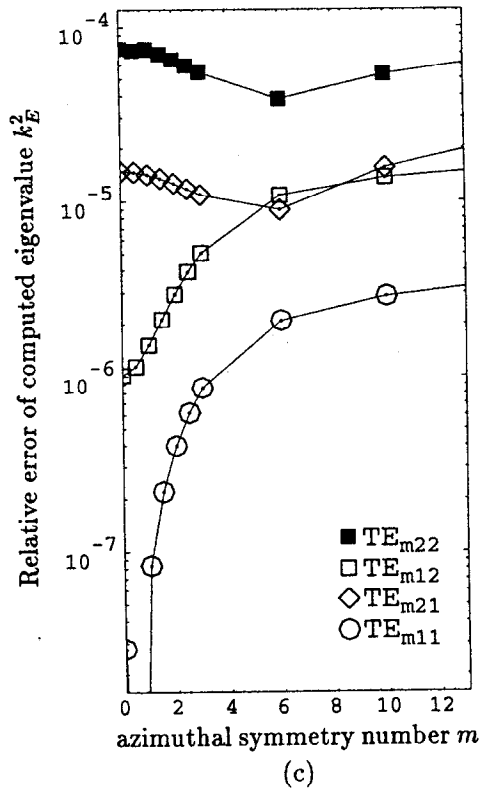
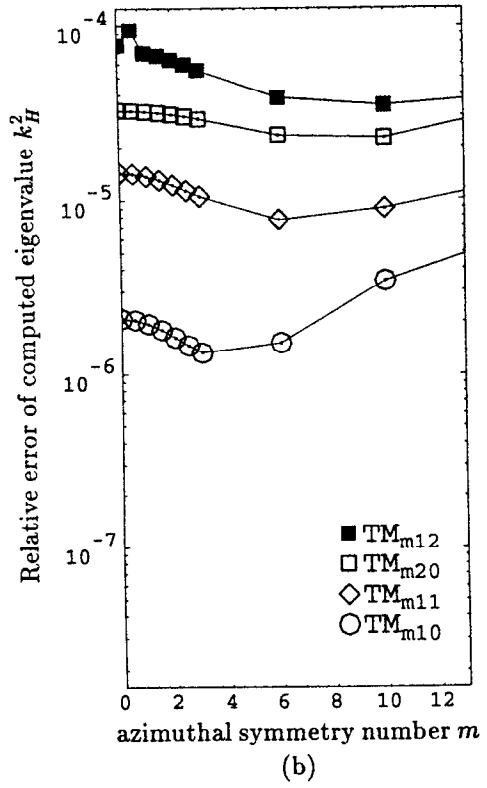
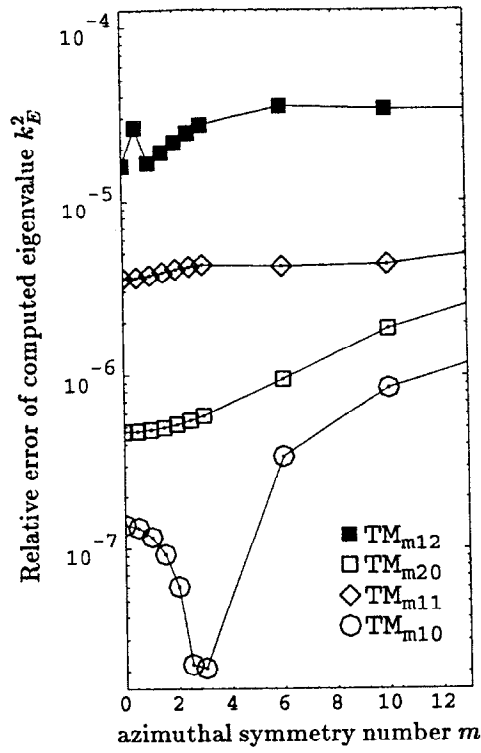


Figure 4.9: Relative error of the computed eigenvalue k^2 versus the azimuthal symmetry number m . The top two figures show the TM mode errors, the bottom two figures show the TE mode errors. The mesh consists of 512 quadratic elements.

satisfy condition (4.4c). In terms of the field components \tilde{E}_ϕ and \mathbf{E}_t , the integral

$$\int_{\tilde{\Omega}} \rho (\nabla \times \mathbf{E}_t)^* \cdot (\nabla \times \mathbf{E}_t) + \frac{1}{\rho} (m\mathbf{E}_t - \nabla_\perp \tilde{E}_\phi)^* \cdot (m\mathbf{E}_t - \nabla_\perp \tilde{E}_\phi) + \rho \mathbf{E}_t^* \cdot \mathbf{E}_t + \frac{1}{\rho} \tilde{E}_\phi^* \cdot \tilde{E}_\phi d\rho dz \quad (4.18)$$

must exist and be finite for any basis function $\mathbf{E} \in \mathcal{H}_{\text{curl}}(\tilde{\Omega})$. The terms which may cause difficulty are those weighted by $1/\rho$. The two terms will be integrable and finite if linear combinations of the basis functions described earlier are used with the following constraints on axis ($\rho = 0$):

$$\tilde{E}_\phi = 0 \quad (4.19a)$$

and

$$m\mathbf{E}_t - \nabla_\perp \tilde{E}_\phi = 0. \quad (4.19b)$$

If an element has an edge on axis then the z -component of the second constraint is $mE_z = 0$ since $\partial \tilde{E}_\phi / \partial z = 0$ is obtained from the first constraint. This leads to the condition $E_z = 0$ if $m \neq 0$. The same condition on E_z is obtained from demanding continuity (in the three dimensional sense) of the field near the axis. When $m = \pm 1$ the ρ -component of the second constraint is $\pm E_\rho = \partial \tilde{E}_\phi / \partial \rho = -iE_\phi$, which also agrees with the constraints obtained from demanding continuity of the field near the axis. For other m the constraints (4.19) are not the same as those obtained from demanding continuity near the axis. For example, demanding continuity near the axis requires $E_\rho = 0$ and $E_\phi = 0$ on axis whenever $m \neq \pm 1$.

4.4.1. Axis Elements for $m \neq 0$

Consider first a linear element with an edge, say edge 3, on axis. The task is to construct basis functions which are linear combinations of the six original basis functions (4.7) and (4.11) and which satisfy the constraints (4.19). The constraint $\tilde{E}_\phi = 0$ eliminates N_1 and N_3 and the constraint $E_z = 0$ eliminates N_6 . For the constraint $mE_\rho = \partial \tilde{E}_\phi / \partial \rho$ notice

$$\mathbf{N}_4 = r\mathbf{R}_3 = r \frac{\partial s}{\partial \rho}, \quad \mathbf{N}_5 = -t\mathbf{R}_3 = -(1-r) \frac{\partial s}{\partial \rho} \quad \text{and} \quad \frac{\partial N_2}{\partial \rho} = \frac{\partial s}{\partial \rho} \quad (4.20)$$

on the axis ($s = 0$). The only combination satisfying the constraint is $m\mathbf{E}_t = \mathbf{N}_4 - \mathbf{N}_5$ with $\tilde{E}_\phi = N_2$. Noting that $\mathbf{N}_4 - \mathbf{N}_5 = \mathbf{R}_3$, the single basis function for a linear element with an edge on axis is

$$\mathbf{N}_1^{\text{edge}} = \left(\frac{1}{m} \mathbf{R}_3, s \right). \quad (4.21)$$

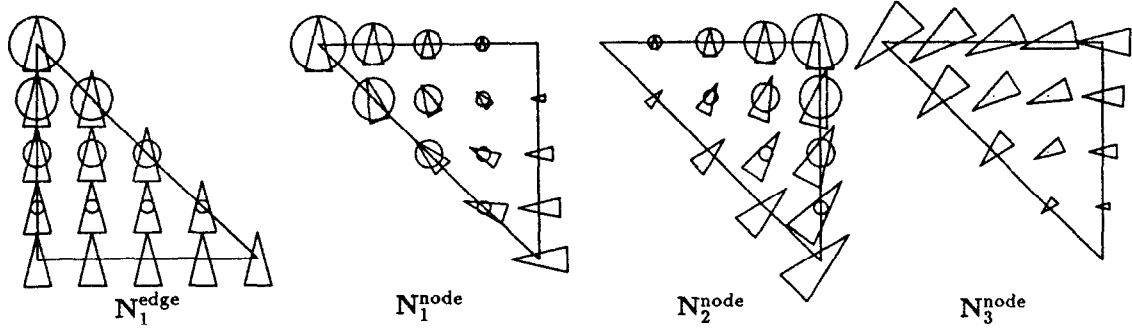


Figure 4.10: Local basis functions for the $m = 1$ problem for a linear element touching the axis. The bottom edge or node is on axis. The vectors (triangles) represent \mathbf{E}_t and the circles represent \tilde{E}_ϕ .

The notation for the basis function is $(\mathbf{E}_t, \tilde{E}_\phi)$. The superscript indicates the basis function is for an element with an edge (edge 3) on axis. An example is shown in figure 4.10.

Next consider a linear element with only a node, say node 2, on axis. The constraint $\tilde{E}_\phi = 0$ eliminates N_2 . Since only the node is on axis it is not necessary to have $E_z = 0$ at the node on axis. In other words, $\tilde{E}_\phi = 0$ at a single point (node 2) does not imply that the derivative $\partial \tilde{E}_\phi / \partial z = 0$ there. However, it is still necessary that $m\mathbf{E}_t - \nabla_\perp \tilde{E}_\phi = 0$ at the node on axis. The basis function $N_6 = L_3$ by itself satisfies the two constraints since it has no \mathbf{E}_t on axis and no \tilde{E}_ϕ , so this will be one basis function for a linear element with a node on axis. For the other basis functions notice that

$$\begin{aligned} N_4 = L_1 = -R_2 = -\nabla_\perp r, \quad N_5 = L_2 = R_1 = \nabla_\perp t, \\ \nabla_\perp N_1 = \nabla_\perp r \quad \text{and} \quad \nabla_\perp N_3 = \nabla_\perp t \end{aligned} \quad (4.22)$$

at the node on axis. Two node-on-axis basis functions can be constructed from these four basis functions. The three basis functions for the linear element with a node on axis are

$$N_1^{\text{node}} = \left(-\frac{1}{m}L_1, r\right), \quad N_2^{\text{node}} = \left(\frac{1}{m}L_2, t\right) \quad \text{and} \quad N_3^{\text{node}} = (L_3, 0). \quad (4.23)$$

The superscript indicates the basis functions are for an element with exactly one node (node 2) on axis. These local basis functions are plotted in figure 4.10.

In order to construct valid global basis functions from the on-axis local basis functions, the coefficients for the special local basis functions N_1^{edge} , N_1^{node} and N_2^{node} are associated with the nodes next to (not on) the axis. This obviously matches the scalar field \tilde{E}_ϕ . It also matches the component of \mathbf{E}_t which is tangential to the element boundaries.

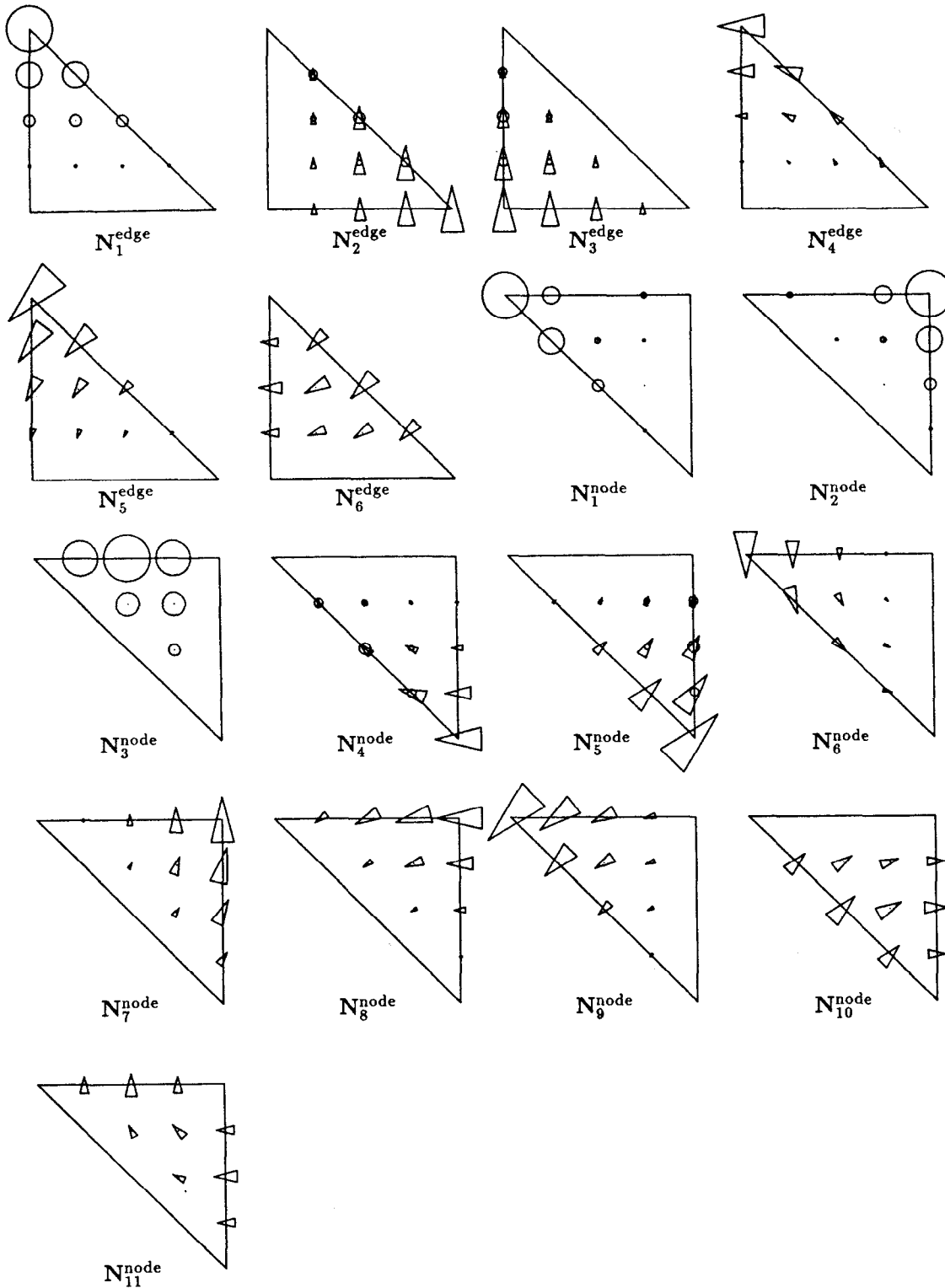


Figure 4.11: Local basis functions for the $m = 1$ problem for a quadratic element touching the axis. The bottom edge or node is on axis. The vectors (triangles) represent \mathbf{E}_t and the circles represent \tilde{E}_ϕ .

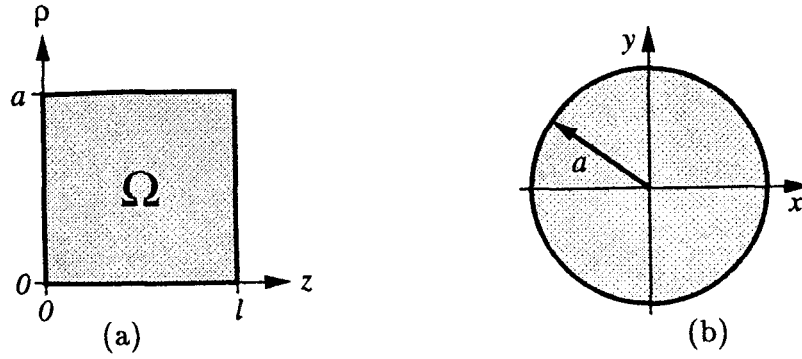


Figure 4.12: The region Ω for a pillbox with radius a and length l . The z - p (a) and x - y (b) cross-sections are shown.

	azimuthal symmetry number m			
	0	1	2	10
γ_i	2.4048255577 5.5200781103 8.6537279129 11.7915344390	3.8317059702 7.0155866698 10.1734681351 13.3236919363	5.1356223018 8.4172441404 11.6198411721 14.7959517824	14.4755006866 18.4334636670 22.0469853647 25.5094505542
γ'_i	3.8317059702 7.0155866698 10.1734681351	1.8411837813 5.3314427735 8.5363163663 11.7060049026	3.0542369282 6.7061331942 9.9694678231 13.1703708560	11.7708766750 16.4478527485 20.2230314127 23.7607158603

Table 4.3: Some numerical values for the cutoff wavenumbers $\gamma_{m,n}$ (for TM_m modes) and $\gamma'_{m,n}$ (for TE_m modes) of circular waveguide with radius $a = 1$.

Quadratic elements which touch the axis are constructed in the same manner. For an element with an edge (edge 3) on axis there are 6 local basis functions:

$$\begin{aligned}
 \mathbf{N}_1 &= (0, s^2) & \mathbf{N}_3 &= \left(\frac{t}{m} \mathbf{R}_3, st\right) & \mathbf{N}_5 &= (s \mathbf{L}_2, 0) \\
 \mathbf{N}_2 &= \left(\frac{r}{m} \mathbf{R}_3, rs\right) & \mathbf{N}_4 &= (s \mathbf{L}_1, 0) & \mathbf{N}_6 &= (-2s \mathbf{L}_3, 0).
 \end{aligned} \tag{4.24}$$

An element with exactly one node (node 2) on axis has 11 local basis functions:

$$\begin{aligned}
 \mathbf{N}_1 &= (0, r(2r-1) + rs) & \mathbf{N}_5 &= \left(\frac{s}{m} \mathbf{L}_2, st\right) & \mathbf{N}_9 &= (r \mathbf{L}_3, 0) \\
 \mathbf{N}_2 &= (0, t(2t-1) + st) & \mathbf{N}_6 &= (r \mathbf{L}_1, 0) & \mathbf{N}_{10} &= (-2s \mathbf{L}_3, 0) \\
 \mathbf{N}_3 &= (0, 4rt) & \mathbf{N}_7 &= (t \mathbf{L}_2, 0) & \mathbf{N}_{11} &= (-2t \mathbf{L}_1, 0). \\
 \mathbf{N}_4 &= \left(\frac{-s}{m} \mathbf{L}_1, rs\right) & \mathbf{N}_8 &= (t \mathbf{L}_3, 0)
 \end{aligned} \tag{4.25}$$

Examples of these basis functions are shown in figure 4.11.

4.4.2. Tests on a Pillbox

As a test case consider a pillbox with length l and radius a shown in figure 4.12a. This problem can be solved analytically by treating it as a section of cylindrical

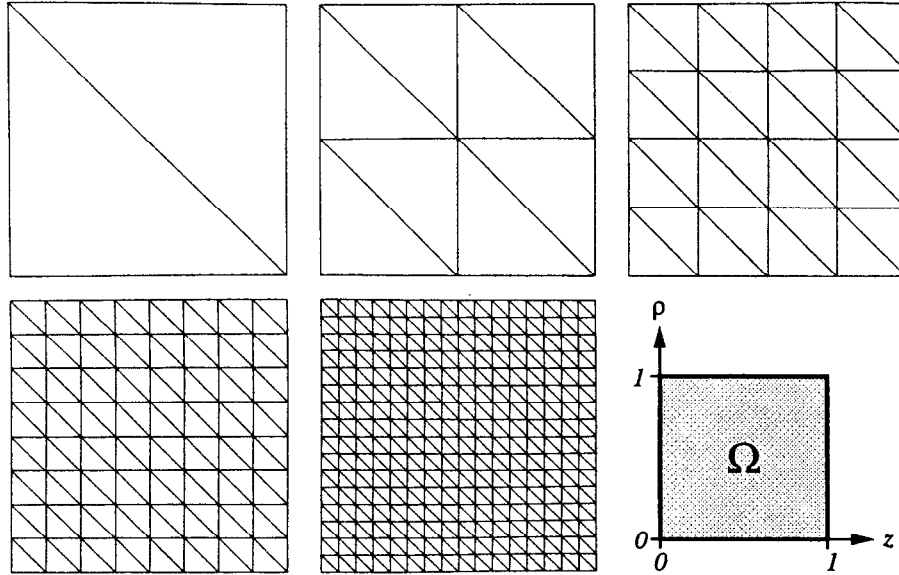


Figure 4.13: A course mesh (two elements) and four refined meshes for the pillbox problem. Each mesh covers the shaded region Ω .

waveguide shorted at the two planes $z = 0$ and $z = l$. The cross section of the cylindrical waveguide is circular as shown in figure 4.12b. The characteristic equation for the cutoff wavenumbers γ_m for the TM_m modes is

$$J_m(\gamma_m a) = 0, \quad (4.26)$$

and the characteristic equation for the cutoff wavenumbers γ'_m of the TE_m modes is

$$J'_m(\gamma'_m a) = 0. \quad (4.27)$$

Let $\gamma_{m,n}$ be the n th solution of (4.26) and let $\gamma'_{m,n}$ be the n th solution of (4.27). Some approximate numerical values for the cutoff wavenumbers for the case $a = 1$ and various m are listed in table 4.3. The wavenumbers were obtained by finding the roots of the characteristic equation numerically and are correct to the given number of digits. The $m = 0$ column is included for completeness only. The axis elements described in the previous section are not suitable for $m = 0$ calculations. Similar axis elements suitable for $m = 0$ problems are described later in section 4.4.4.

Let p represent the number of half-wavelengths of field variation along the length l of the waveguide. The frequencies of the TM modes in the pillbox are

$$\text{TM}_{mnp} : k^2 = \gamma_{m,n}^2 + (p\pi/l)^2 \quad p = 0, 1, 2, \dots \quad (4.28)$$

and the frequencies of the TE modes are

$$\text{TE}_{mnp} : k^2 = \gamma'_{m,n}{}^2 + (p\pi/l)^2 \quad p = 1, 2, 3, \dots \quad (4.29)$$

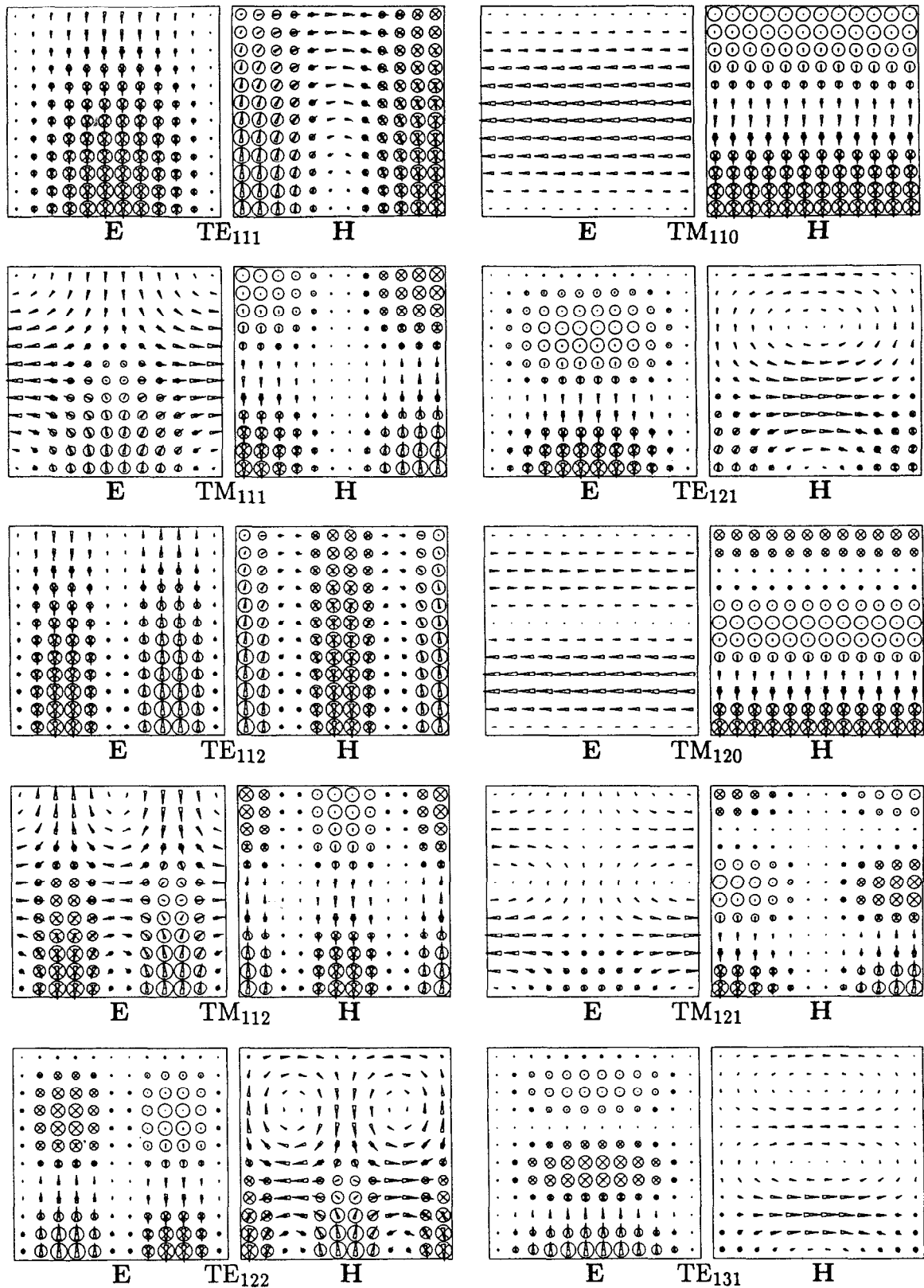


Figure 4.14: Field plots of the 10 lowest $m = 1$ modes of the pillbox. The ϕ -component of the fields is imaginary.

mode	k_{analytic}^2	k_E^2	relative error	k_H^2	relative error
TE ₁₁₁	13.259562	13.259697	1×10^{-5}	13.259680	9×10^{-6}
TM ₁₁₀	14.681971	14.681979	6×10^{-7}	14.682331	2×10^{-5}
TM ₁₁₁	24.551575	24.551897	1×10^{-5}	24.552431	3×10^{-5}
TE ₁₂₁	38.293886	38.297254	9×10^{-5}	38.294658	2×10^{-5}
TE ₁₁₂	42.868375	42.870229	4×10^{-5}	42.870595	5×10^{-5}
TM ₁₂₀	49.218456	49.218593	3×10^{-6}	49.225959	2×10^{-4}
TM ₁₁₂	54.160388	54.163750	6×10^{-5}	54.165748	1×10^{-4}
TM ₁₂₁	59.088061	59.089844	3×10^{-5}	59.099110	2×10^{-4}
TE ₁₂₂	67.902700	67.915516	2×10^{-4}	67.908651	9×10^{-5}
TE ₁₃₁	82.738302	82.766452	3×10^{-4}	82.741628	4×10^{-5}

Table 4.4: Relative error of the computed eigenvalues for the 10 lowest $m = 1$ modes of a pillbox with $a = 1$ and $l = 1$ and using a fine mesh (512 quadratic elements).

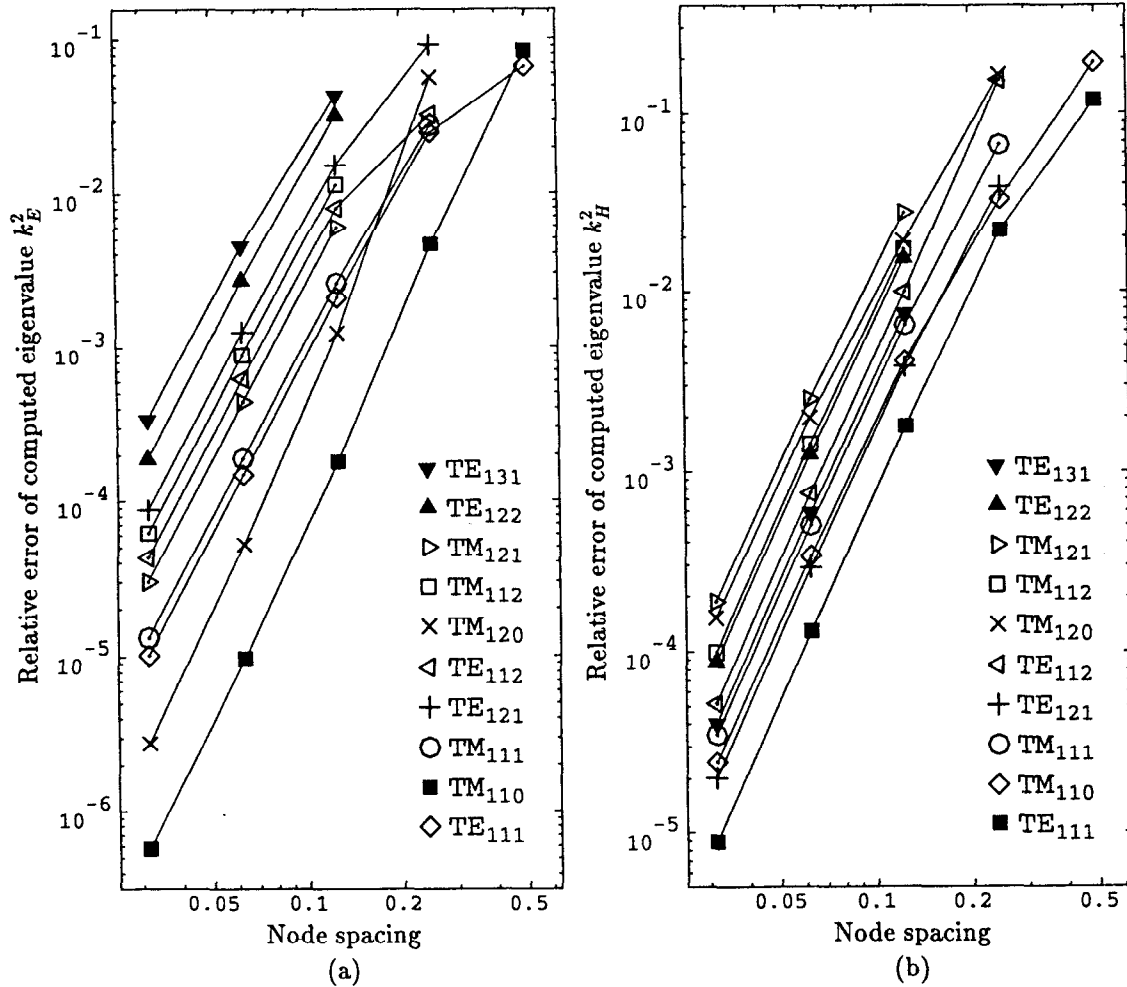


Figure 4.15: Relative error of the computed eigenvalues for the $m = 1$ modes of a pillbox as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} .

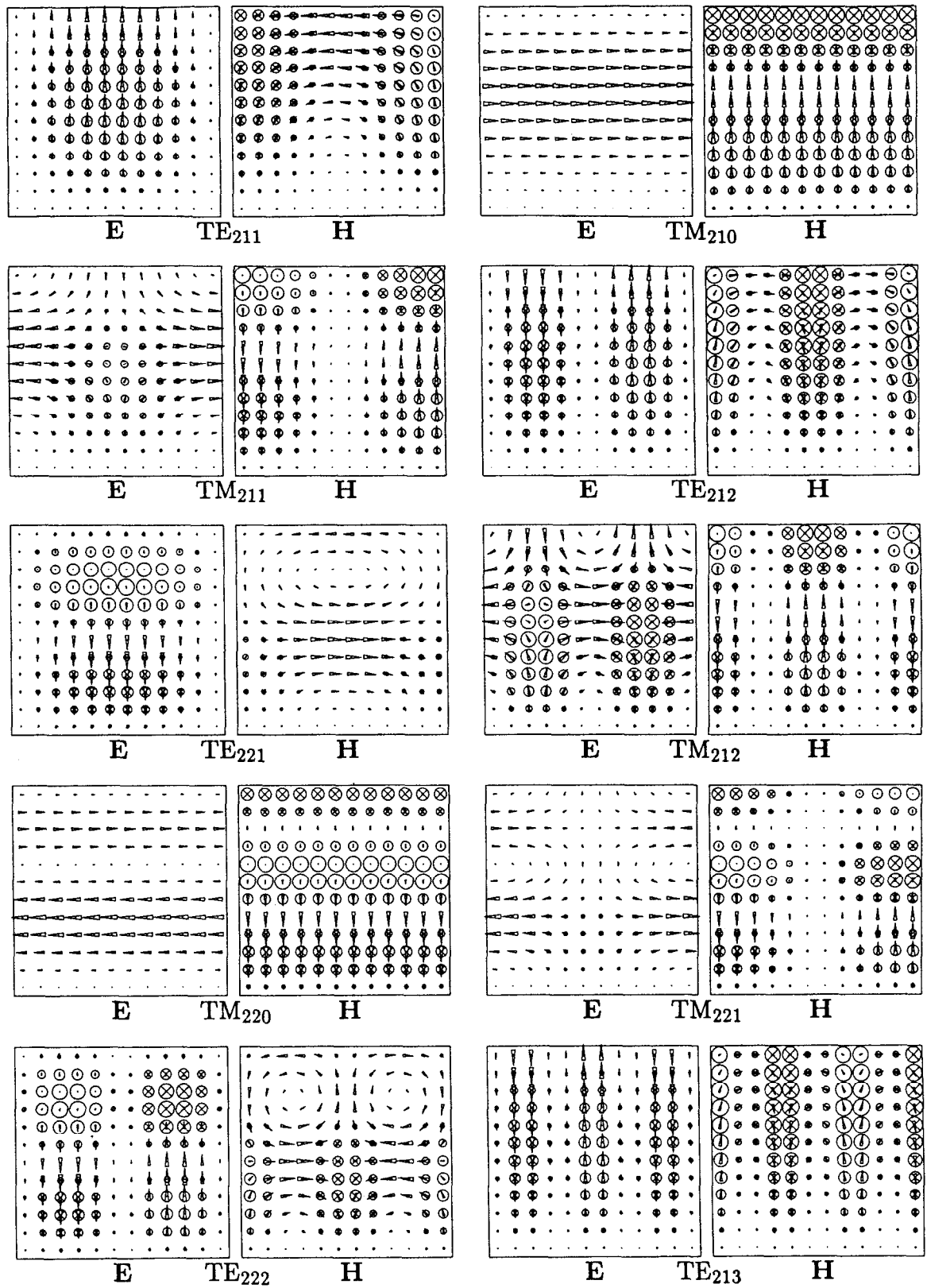


Figure 4.16: Field plots of the 10 lowest $m = 2$ modes of the pillbox. The ϕ -component of the fields is imaginary.

mode	k_{analytic}^2	k_E^2	relative error	k_H^2	relative error
TE ₂₁₁	19.197968	19.198095	7×10^{-6}	19.198081	6×10^{-6}
TM ₂₁₀	26.374616	26.374807	7×10^{-6}	26.374869	1×10^{-5}
TM ₂₁₁	36.244221	36.244688	1×10^{-5}	36.245058	2×10^{-5}
TE ₂₁₂	48.806781	48.807983	2×10^{-5}	48.808468	3×10^{-5}
TE ₂₂₁	54.841827	54.844857	6×10^{-5}	54.843164	2×10^{-5}
TM ₂₁₂	65.853034	65.856041	5×10^{-5}	65.857933	7×10^{-5}
TM ₂₂₀	70.849999	70.851913	3×10^{-5}	70.855221	7×10^{-5}
TM ₂₂₁	80.719603	80.723149	4×10^{-5}	80.729151	1×10^{-4}
TE ₂₂₂	84.450640	84.463415	2×10^{-4}	84.456630	7×10^{-5}
TE ₂₁₃	98.154803	98.162068	7×10^{-5}	98.168826	1×10^{-4}

Table 4.5: Relative error of the computed eigenvalues for the 10 lowest $m = 2$ modes of a pillbox with $a = 1$ and $l = 1$ and using a fine mesh (512 quadratic elements).

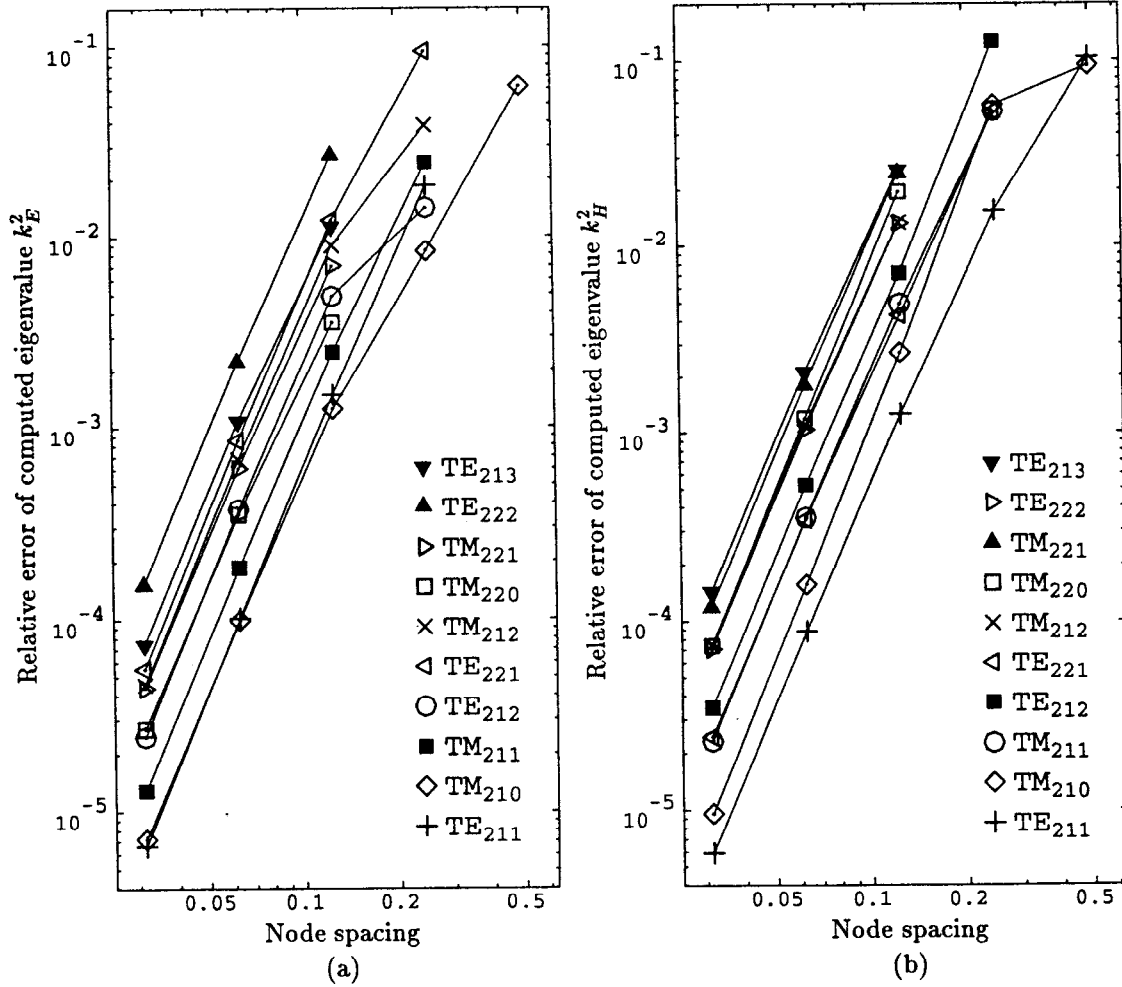


Figure 4.17: Relative error of the computed eigenvalues for the $m = 2$ modes of a pillbox as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} .

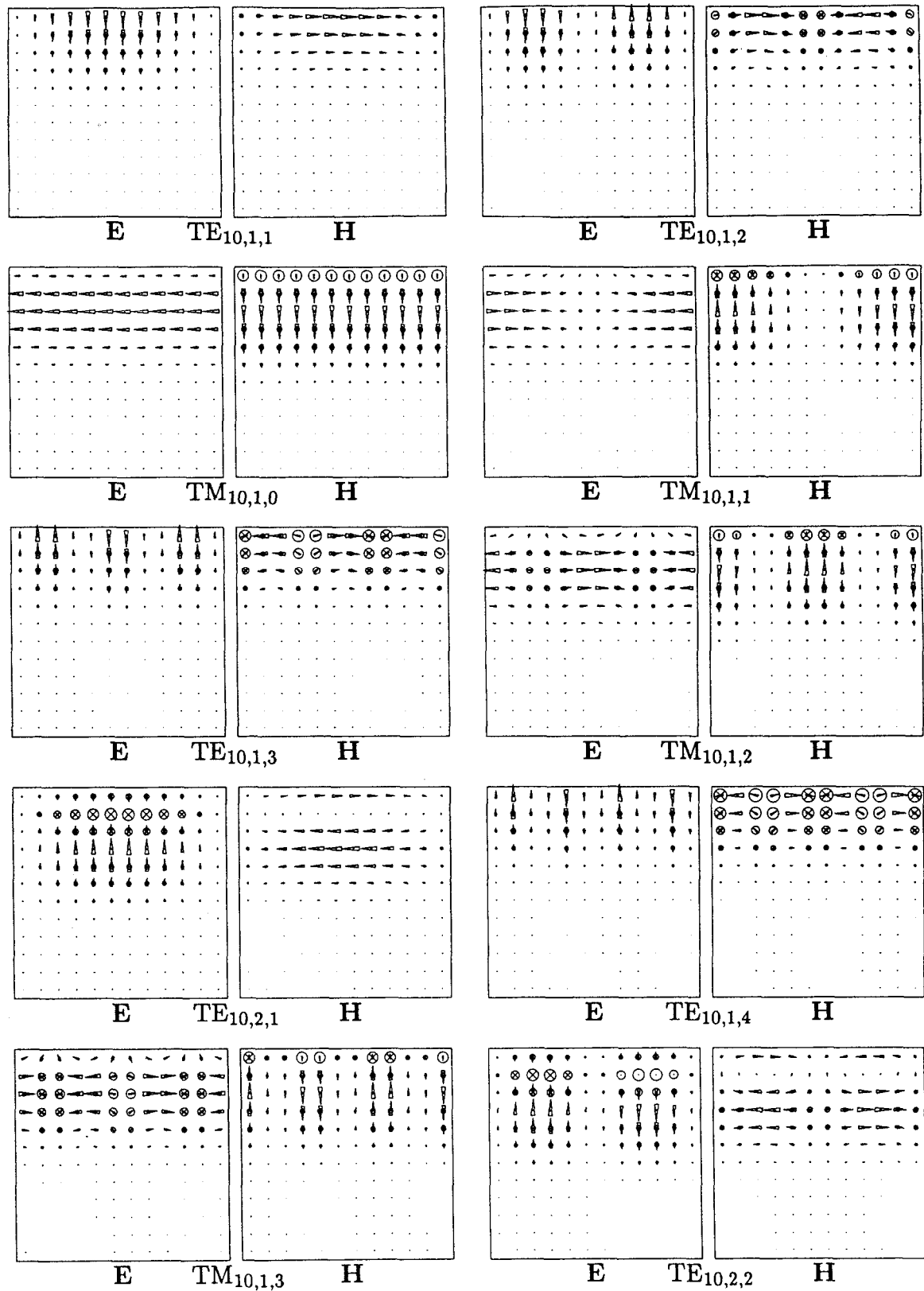


Figure 4.18: Field plots of the 10 lowest $m = 10$ modes of the pillbox. The ϕ -component of the fields is imaginary.

mode	k_{analytic}^2	k_E^2	relative error	k_H^2	relative error
TE _{10,1,1}	148.4231	148.4251	1×10^{-5}	148.4250	1×10^{-5}
TE _{10,1,2}	178.0320	178.0399	4×10^{-5}	178.0395	4×10^{-5}
TM _{10,1,0}	209.5401	209.5430	1×10^{-5}	209.5515	5×10^{-5}
TM _{10,1,1}	219.4097	219.4147	2×10^{-5}	219.4255	7×10^{-5}
TE _{10,1,3}	227.3800	227.4039	1×10^{-4}	227.4120	1×10^{-4}
TM _{10,1,2}	249.0185	249.0351	7×10^{-5}	249.0523	1×10^{-4}
TE _{10,2,1}	280.4015	280.4429	1×10^{-4}	280.4108	3×10^{-5}
TE _{10,1,4}	296.4672	296.5288	2×10^{-4}	296.5851	4×10^{-4}
TM _{10,1,3}	298.3666	298.4253	2×10^{-4}	298.4468	3×10^{-4}
TE _{10,2,2}	310.0103	310.0966	3×10^{-4}	310.0435	1×10^{-4}

Table 4.6: Relative error of the computed eigenvalues for the 10 lowest $m = 10$ modes of a pillbox with $a = 1$ and $l = 1$ and using a fine mesh (512 quadratic elements).

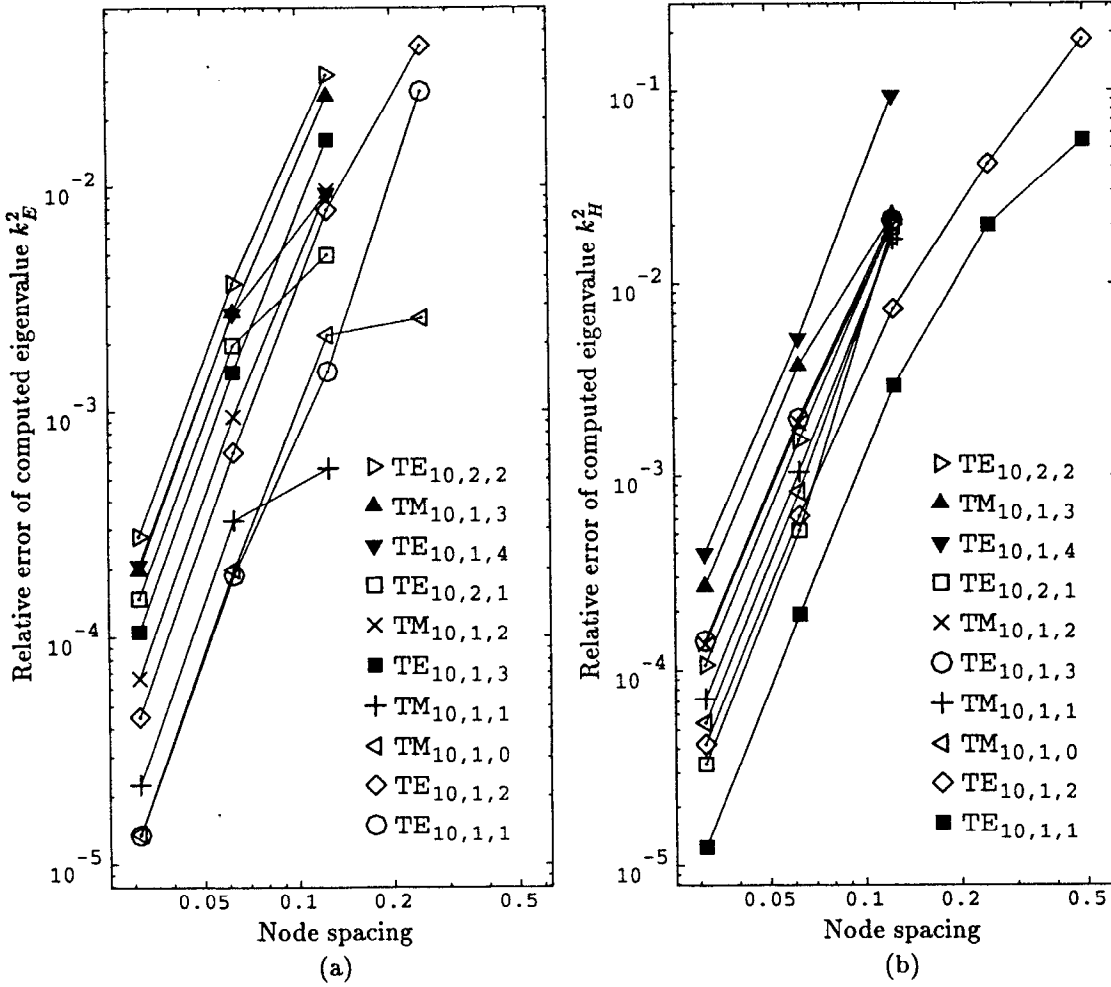


Figure 4.19: Relative error of the computed eigenvalues for the $m = 10$ modes of a pillbox as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} .

The lowest 10 modes with $m = 1$ are plotted in figure 4.14. The plots were constructed using the finest mesh of figure 4.13 which contains 512 quadratic triangular elements. The electric and magnetic fields were obtained from two separate calculations. The calculated eigenvalues and the corresponding relative errors for these modes are listed in table 4.4. In general the errors are typically higher than the annular ring case by about a factor of 4. The electric field calculations for the TM_{1n0} modes are surprisingly good compared to the other calculations. This is similar to results from the annular ring test since the TM_{1n0} modes have field E_z which is uniform in z . However, the electric field calculations for the TE_{111} mode might be expected to have similar accuracy since its field E_ρ is nearly uniform in ρ . This is not the case. The TE_{111} mode calculated using the electric field is an order of magnitude worse than the corresponding TM_{110} calculation. The best accuracy seems correlated with zero electric field on the axis, which indicates the axis elements may be inadequate. The convergence of the finite element formulation as the mesh is refined is shown in figure 4.15. As in the annular ring case, the convergence of the magnetic field formulation is smoother than the electric field formulation for coarse meshes. However, for both field calculations the error is $\mathcal{O}(h^{3.8})$. This is not as good as the annular ring case, where the error was $\mathcal{O}(h^4)$. This is another indication that the near-axis basis functions could be improved to deliver the same accuracy as the off-axis basis functions.

The separation between spurious and physical modes is still extremely good. The eigenvalues of the spurious modes are $\omega^2/c^2 < 10^{-10}$.

Figures 4.16 and 4.18 are plots of the $m = 2$ and $m = 10$ modes, respectively. The fields were computed using the finest mesh shown in figure 4.13. The field plots show no field on axis for these modes, as expected. The field on axis is not exactly zero, however. It is very small in fine mesh calculations, but for coarse mesh calculations the field on axis can become sufficiently large to be noticed in the field plots. The computed eigenvalues and their relative errors are listed in tables 4.5 and 4.6. Figures 4.17 and 4.19 show the convergence of the $m = 2$ and $m = 10$ mode calculations as the mesh is refined. The convergence is similar to the $m = 1$ mode calculations shown in figure 4.15. The error of the computed eigenvalue is typically $\mathcal{O}(h^{3.8})$ to $\mathcal{O}(h^{3.9})$, which is close to the theoretically expected $\mathcal{O}(h^4)$ error.

Figure 4.20 shows the mode frequencies as a function of the azimuthal symmetry number m . The TE_m and TM_m modes are plotted separately for clarity. Each curve is labelled by a radial mode index n and an axial index p . Except for the TE_0 modes, the points on the curves are the frequencies of the TE_{mnp} and TM_{mnp} modes, with

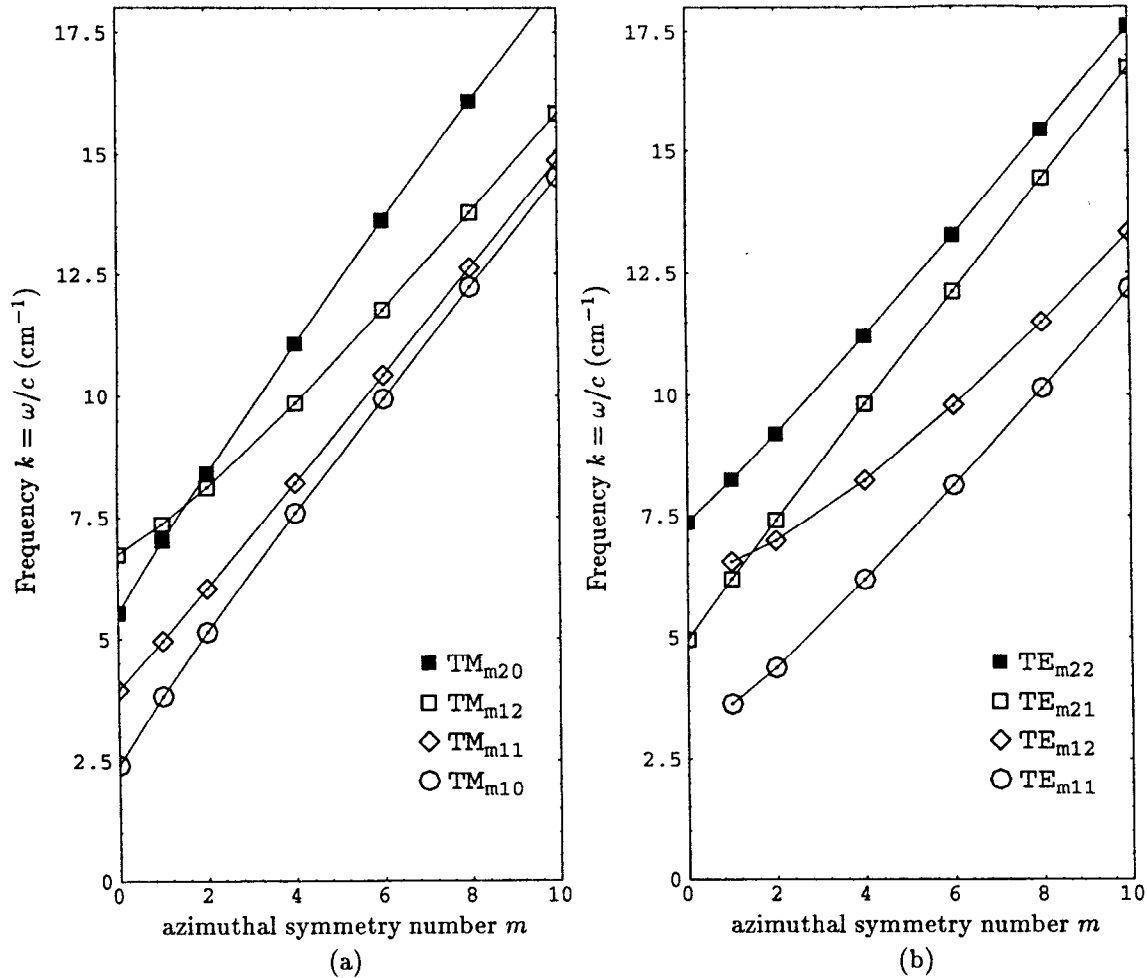


Figure 4.20: Frequency $k = \omega/c$ versus azimuthal symmetry number m for (a) four TM modes and (b) four TE modes.

n and p fixed.

Traditional labelling of the TE_0 modes does not lead to a smooth curve in figure 4.20. A smooth curve is obtained if the $m = 0$ point of the TE_{mnp} curve is the frequency of the $\text{TE}_{0,n-1,p}$ mode. Hence the TE_{m1p} curves do not include a point at $m = 0$. Comparison of the field plots of the TE_{mnp} modes in figures 4.28 ($m = 0$), 4.14 ($m = 1$) and 4.16 ($m = 2$) yields further evidence that this arrangement of the TE_0 modes is appropriate.

The dependence on m of the relative error of the computed eigenvalues is plotted in figure 4.21. The eigenvalues were computed using the finest mesh shown in figure 4.13. For $m > 2$ the relative error is increasing slowly as m increases. This is due to the field variations being confined to a small portion of the pillbox for large m , as shown in the $m = 10$ field plots in figure 4.18. The relative error for

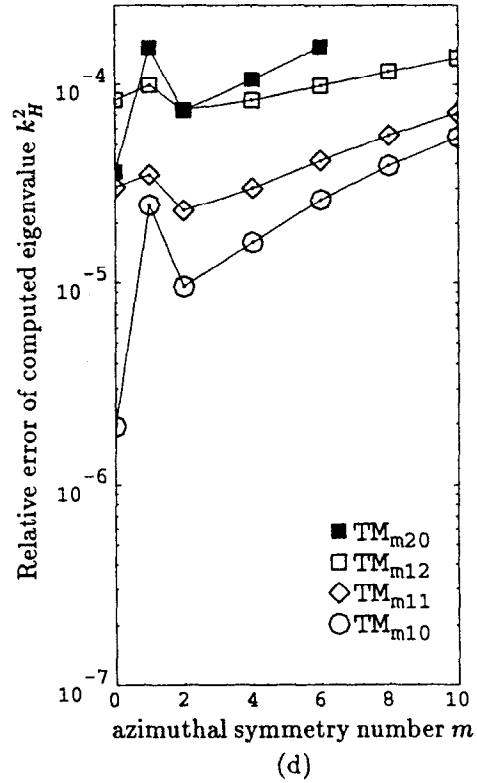
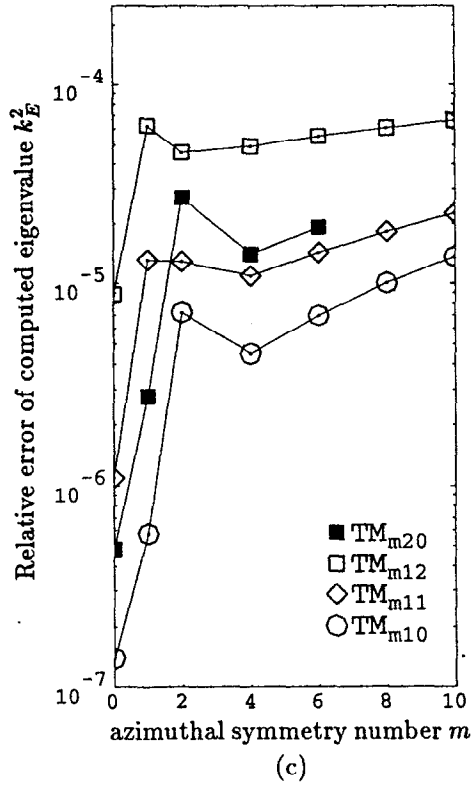
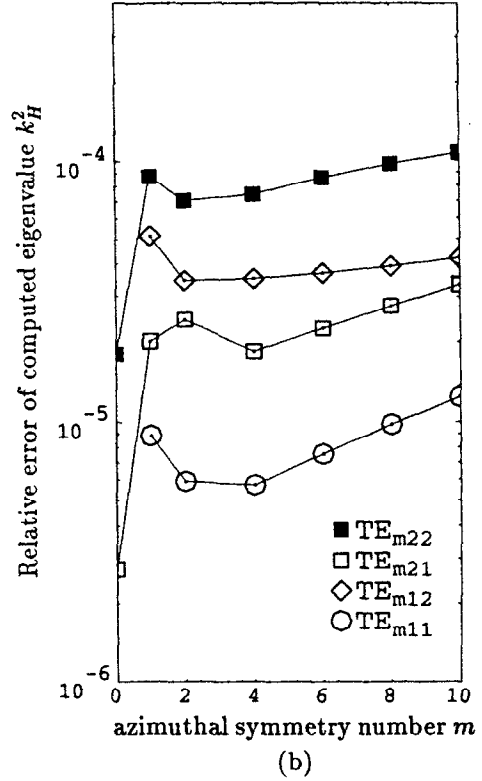
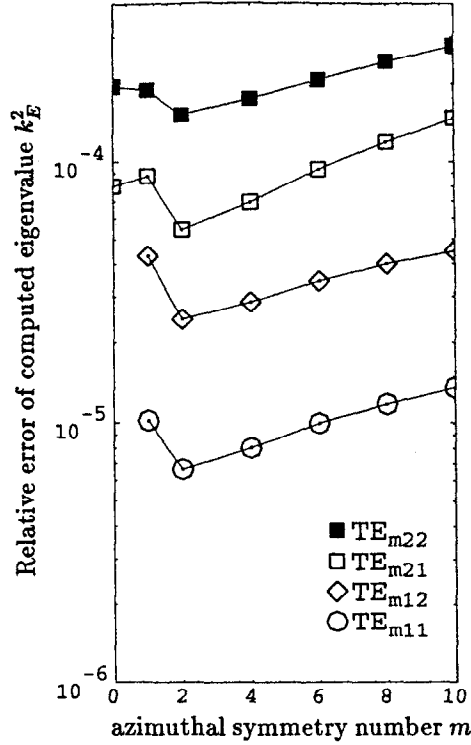


Figure 4.21: Relative error of the computed eigenvalue k^2 versus the azimuthal symmetry number m . The top two figures show the TE mode errors, the bottom two figures show the TM mode errors.

the $m = 1$ mode is typically significantly higher than the $m = 2$ mode, which is probably due to the presence of field on axis.

Two exceptions to this rule are the electric field calculations of the TM_{110} and TM_{120} modes. These calculations are significantly better than the TM_{210} and TM_{220} mode calculations. In these four calculations there is no field on axis and there is no variation of the field $\mathbf{E} = \hat{\mathbf{z}}E_z(\rho)$ along the field direction $\hat{\mathbf{z}}$. Contrast this with the electric field calculations of the TM_{m11} and TM_{m12} modes. The relative error of the $m = 1$ calculation is the same or higher as the $m = 2$ calculation. This correlates with field on axis and variation of the field along the direction of the field.

The computed $m = 0$ modes of a pillbox are discussed in the last section of this chapter. Many $m = 0$ modes are much more accurate than the corresponding $m = 1$ modes. The best accuracy is obtained in modes with no field variation along the direction of the field. Modes with field on axis, such as the electric field calculation of the TM_{010} mode, still have dramatically improved accuracy over their $m = 1$ counterparts. Comparing the magnetic field calculations of the TM_0 modes shows a correlation between poor improvement in accuracy from $m = 1$ to $m = 0$ with non-zero $\partial H_\phi / \partial z$.

4.4.3. Tests on a Sphere

As another test case consider a sphere with radius a as shown in figure 4.22. This will be an important test of how well the method handles curved boundaries. Using spherical coordinates r , θ and ϕ , the modes of a sphere can be found analytically in terms of the spherical harmonic functions $Y_{lm}(\theta, \phi)$ and the spherical bessel functions $j_l(kr)$. The modes separate into two distinct sets: the TE (magnetic multipole) modes which have no radial electric field and the TM (electric multipole) modes which have no radial magnetic field.

The characteristic equation for the wavenumber $k_l = \omega/c$ of the TE modes is

$$j_l(k_l a) = 0 \quad l = 1, 2, \dots \quad (4.30)$$

The characteristic equation for the TM modes is

$$\left. \frac{\partial}{\partial x} (x j_l(k_l' x)) \right|_{x=a} = 0 \quad l = 1, 2, \dots \quad (4.31)$$

Note that the characteristic equation is indexed by l and not m . For each solution k_l of one of the characteristic equations there are $2l + 1$ degenerate modes. The azimuthal symmetry numbers m of the modes are $-l, -l + 1, \dots, l - 1$ and l . Conversely, for a given m the modes will have index $l \geq |m|$. This corresponds to the

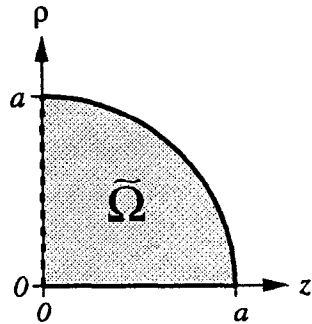


Figure 4.22: The region Ω for a sphere with radius a . The dashed boundary is a symmetry plane.

l	$k_{l,n}$	$k'_{l,n}$
1	4.4934094579	2.7437072700
	7.7252518369	6.1167642645
		9.3166156286
2	5.7634591969	3.8702385802
	9.0950113305	7.4430870540
3	6.9879320005	4.9734203508
		8.7217505135
4	8.1825614526	6.0619493630
		9.9675472302
5	9.3558121110	7.1402273640
6		8.2108419780
7		9.2754634855

Table 4.7: Some approximate numerical values for the wavenumbers $k_{l,n}$ (for TE modes) and $k'_{l,n}$ (for TM modes) of a sphere with radius $a = 1$.

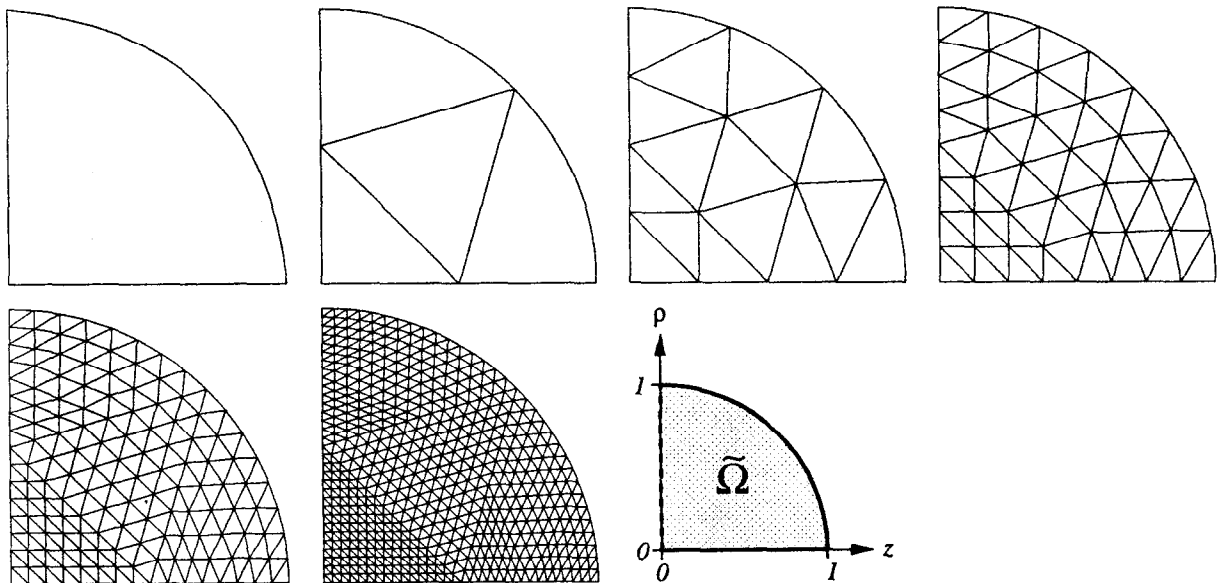


Figure 4.23: A course mesh (one element) and five refined meshes for the sphere problem. Each mesh covers the shaded region Ω .

field solver's situation. That is, the field solver finds modes which are constrained to have the azimuthal symmetry number m specified by the user. Also note that $l = 0$ is not allowed since the spherical bessel function j_0 is not a satisfactory solution at the origin. Hence the first $m = 0$ mode will have $l = 1$, not $l = 0$. Let $k_{l,n}$ be the n th solution of (4.30) and let $k'_{l,n}$ be the n th solution of (4.31). Some approximate numerical values for the case $a = 1$ and low l are listed in table 4.7. The wavenumbers were obtained by finding the roots of the characteristic equation numerically and are correct to the given number of digits. The table includes all of the solutions $k_l < 10$. The frequencies of the modes are

$$\text{TE}_{nlm} : k^2 = k_{l,n}^2 \quad \text{and} \quad \text{TM}_{nlm} : k^2 = k'_{l,n}{}^2. \quad (4.32)$$

The lowest 10 modes with $m = 1$ are plotted in figure 4.24. The plots were constructed using the fifth mesh of figure 4.23 which contains 256 quadratic triangular elements. The electric and magnetic fields were obtained from two separate calculations. The eigenvalues calculated using 1024 quadratic elements and the corresponding relative errors are listed in table 4.8. The errors are similar in magnitude to the pillbox test, and the dependence of the error on mesh size is $\mathcal{O}(h^{3.8})$, just like the pillbox. The error dependence on mesh size is shown in figure 4.25. Tests on a sphere for $m \neq 1$ show similar results. The electric field calculations for a sphere do not have a significant advantage over the magnetic field calculations. This behavior is unlike the pillbox test.

4.4.4. Axis Elements for $m = 0$

For completeness, a set of axis elements suitable for $m = 0$ calculations will be listed here. It should be noted, however, that for $m = 0$ problems the formulations of chapter 3 are more efficient than the formulations in this chapter.

The constraints on $m = 0$ axis elements are the same as on the $m \neq 0$ axis elements. Linear combinations of the basis functions of section 4.3 which satisfy the constraints (4.19a) and (4.19b) are desired. For $m = 0$ these constraints reduce to

$$\tilde{E}_\phi = 0 \quad \text{and} \quad \nabla_\perp \tilde{E}_\phi = 0 \quad \text{on axis.} \quad (4.33)$$

Notice that there are no constraints on E_t . Continuity of the field demands $E_\rho = 0$ on axis, but this is not a requirement in this formulation.

Consider first a linear element with edge 3 on axis. The constraint $\tilde{E}_\phi = 0$ eliminates N_1 and N_3 . Since $\nabla_\perp N_2 = \nabla_\perp s \neq 0$ on the axis, the second constraint eliminates N_2 . Thus, the field \tilde{E}_ϕ is identically zero over the entire element.

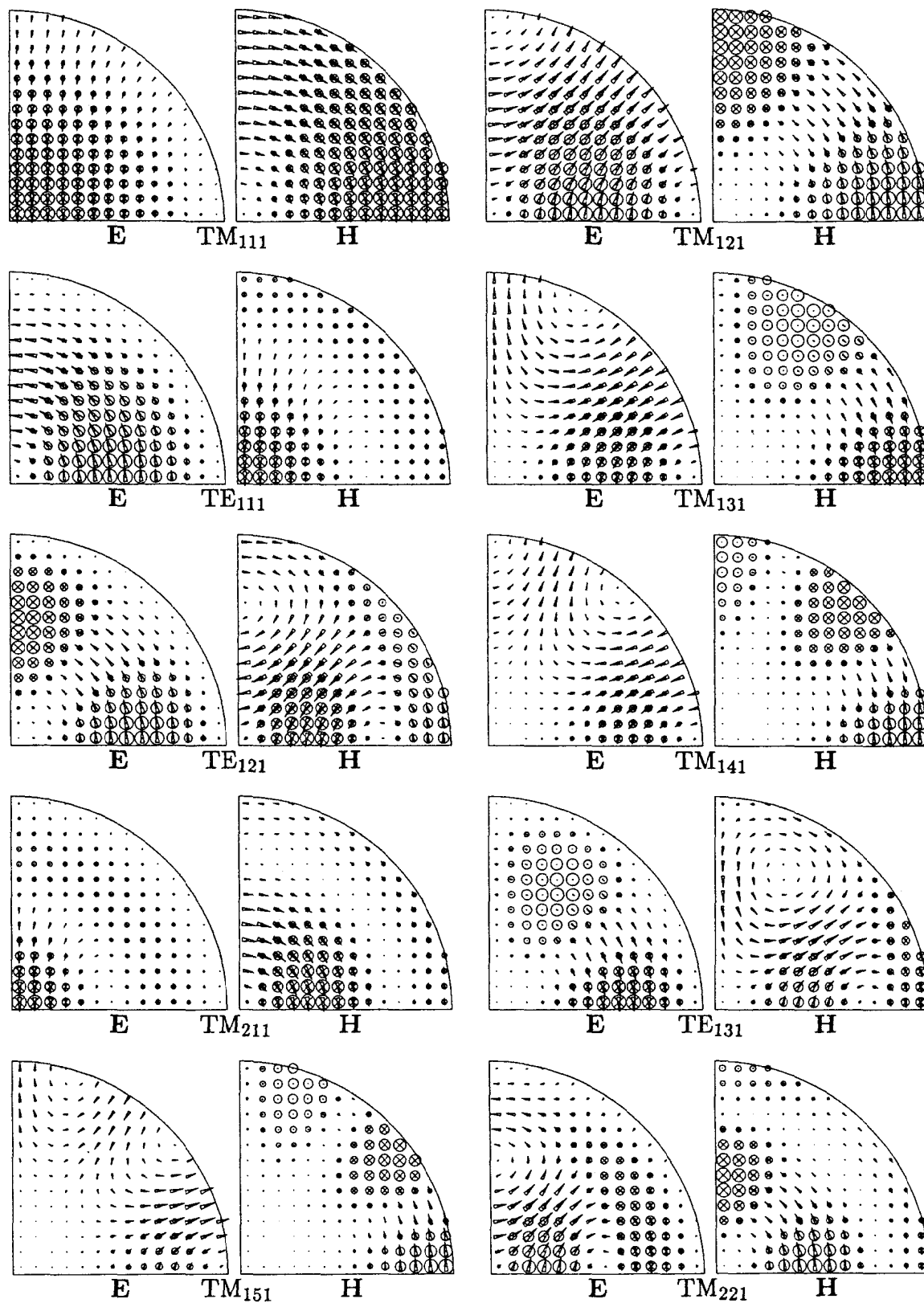


Figure 4.24: Field plots of the 10 lowest $m = 1$ modes of the sphere. The ϕ -component of the fields is imaginary.

mode	k^2_{analytic}	k^2_E	relative error	k^2_H	relative error
TM ₁₁₁	7.527930	7.527935	8×10^{-7}	7.527932	4×10^{-7}
TM ₁₂₁	14.978747	14.978759	8×10^{-7}	14.978781	2×10^{-6}
TE ₁₁₁	20.190729	20.190759	2×10^{-6}	20.190792	3×10^{-6}
TM ₁₃₁	24.734910	24.734943	1×10^{-6}	24.735072	7×10^{-6}
TE ₁₂₁	33.217462	33.217664	6×10^{-6}	33.217578	4×10^{-6}
TM ₁₄₁	36.747230	36.747308	2×10^{-6}	36.747745	1×10^{-5}
TM ₂₁₁	37.414805	37.415089	8×10^{-6}	37.414948	4×10^{-6}
TE ₁₃₁	48.831194	48.831872	1×10^{-5}	48.831466	6×10^{-6}
TM ₁₅₁	50.982847	50.983037	4×10^{-6}	50.984171	3×10^{-5}
TM ₂₂₁	55.399545	55.399939	7×10^{-6}	55.400238	1×10^{-5}

Table 4.8: Relative error of the computed eigenvalues for the 10 lowest $m = 1$ modes of a sphere with $a = 1$ and using a fine mesh (1024 quadratic elements).

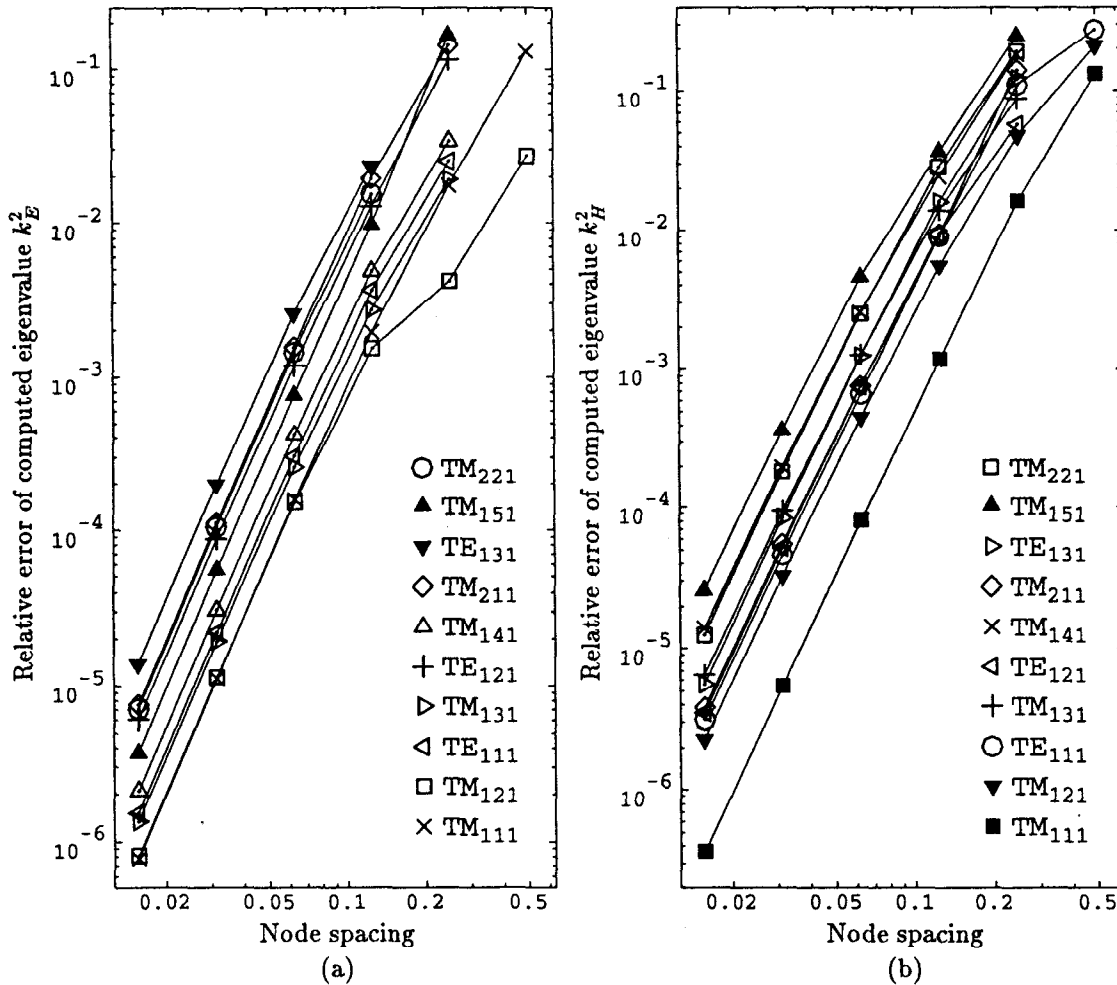


Figure 4.25: Relative error of the computed eigenvalues for the $m = 1$ modes of a sphere as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} .

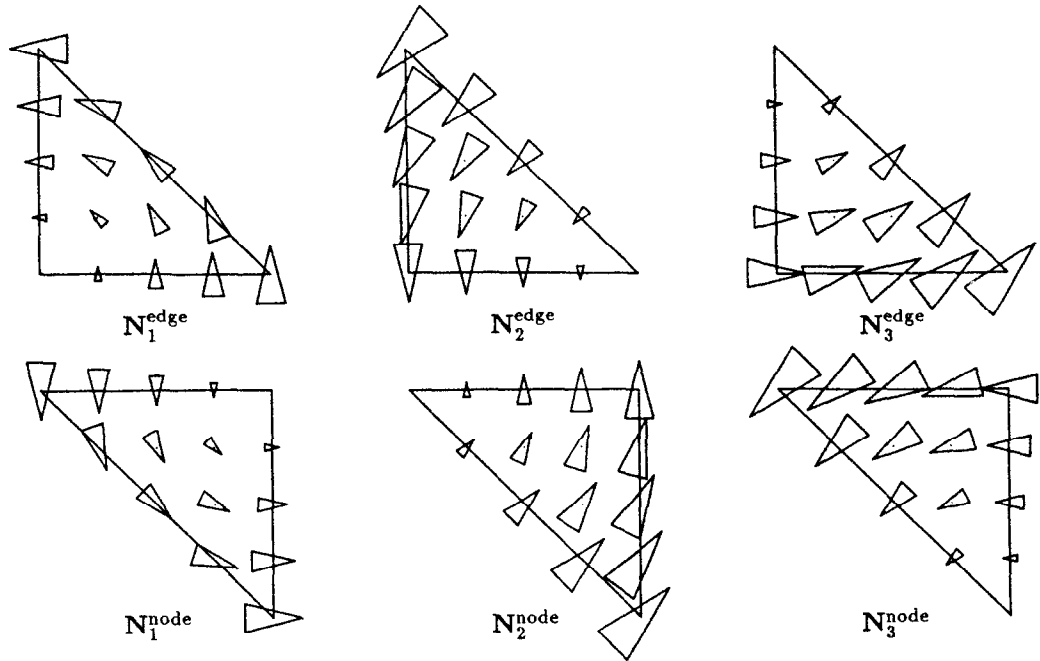


Figure 4.26: Local basis functions for the $m = 0$ problem for a linear element touching the axis. The bottom edge or node is on axis. The vectors (triangles) represent \mathbf{E}_t . There are no basis functions for \tilde{E}_ϕ .

Similarly, a linear element with only node 2 on axis has no basis functions for \tilde{E}_ϕ . The first constraint eliminates N_2 , while the second constraint eliminates N_1 and N_3 .

Both types of axis elements have only three basis functions. These are the usual linear basis functions for \mathbf{E}_t ,

$$N_1^{\text{edge/node}} = L_1, \quad N_2^{\text{edge/node}} = L_2 \quad \text{and} \quad N_3^{\text{edge/node}} = L_3. \quad (4.34)$$

These basis functions are plotted in figure 4.26.

A quadratic element with edge 3 on axis has one basis function,

$$N_1^{\text{edge}} = s^2, \quad (4.35)$$

for \tilde{E}_ϕ which satisfies the constraints (4.33). There is a full complement of 8 basis functions, N_2 through N_9 , for \mathbf{E}_t .

A quadratic element with only node 2 on axis has three basis functions,

$$N_1^{\text{node}} = r(2r - 1) + rs \quad N_2^{\text{node}} = t(2t - 1) + st \quad \text{and} \quad N_3^{\text{node}} = 4rt, \quad (4.36)$$

for \tilde{E}_ϕ . Again, there is a full complement of 8 basis functions, N_4 through N_{11} , for \mathbf{E}_t . The basis functions for the two quadratic $m = 0$ axis elements are shown in figure 4.27.

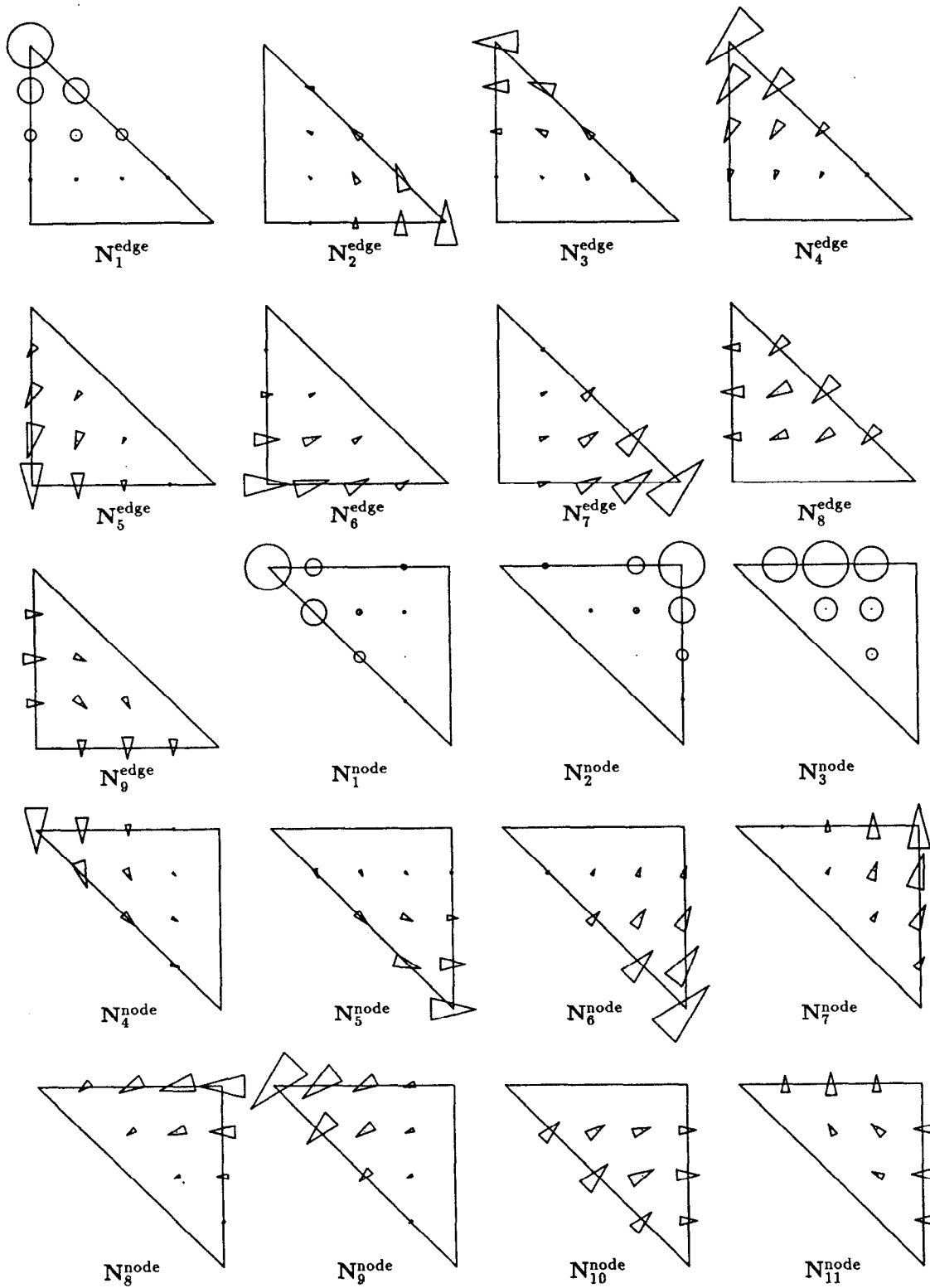


Figure 4.27: Local basis functions for the $m = 0$ problem for a quadratic element touching the axis. The bottom edge or node is on axis. The vectors (triangles) represent \vec{E}_t and the circles represent \vec{E}_ϕ .

The basis functions for \mathbf{E}_t and \tilde{E}_ϕ are independent of each other. If the magnetic permeability μ and the electric permittivity ϵ are diagonal then the equations for \mathbf{E}_t and \tilde{E}_ϕ are uncoupled when $m = 0$. The integral equation (4.4a), or similarly equation (4.6), splits into two separate integral equations,

$$\Re e \ 2\pi \int_{\tilde{\Omega}} \rho (\nabla_{\perp} \times \mathbf{F}_t)^* \cdot \mu_t^{-1} (\nabla_{\perp} \times \mathbf{E}_t) - \frac{\omega^2}{c^2} \rho \mathbf{F}_t^* \cdot \epsilon_t \mathbf{E}_t \, d\rho \, dz = 0 \quad (4.37a)$$

$$\text{and } \Re e \ 2\pi \int_{\tilde{\Omega}} \frac{1}{\rho} \nabla_{\perp} \tilde{F}_\phi^* \cdot \mu_\phi^{-1} \nabla_{\perp} \nabla_{\perp} \tilde{E}_\phi - \frac{\omega^2}{c^2} \frac{1}{\rho} \tilde{F}_\phi^* \cdot \epsilon_\phi \tilde{E}_\phi \, d\rho \, dz = 0, \quad (4.37b)$$

where μ_t and ϵ_t are the transverse portions of the permeability and permittivity matrices, and μ_ϕ and ϵ_ϕ are the ϕ -components of the constitutive matrices.

The TE_0 modes have $\mathbf{E}_t = 0$, while the TM_0 modes have $\tilde{E}_\phi = 0$. Hence the problem is reduced to only TE_0 modes if only \tilde{E}_ϕ is computed. Likewise, the problem is reduced to only TM_0 modes if only \mathbf{E}_t is computed. Magnetic field calculations can be separated similarly. The TE_0 modes are computed using only \mathbf{H}_t , while the TM_0 modes are computed using only \tilde{H}_ϕ .

The $m = 0$ modes of a pillbox were computed as a test case. Figure 4.28 shows the field patterns of the lowest 10 modes computed using the finest mesh of figure 4.13. None of the modes have noticeable E_ρ or H_ρ on axis. The fields E_ϕ and H_ϕ are automatically zero on axis. The electric and magnetic fields were obtained from separate calculations. The calculated eigenvalues and the corresponding relative errors are listed in table 4.9. The convergence of the eigenvalues as the mesh is refined is shown in figure 4.29. The error is $\mathcal{O}(h^4)$, where h is the element size.

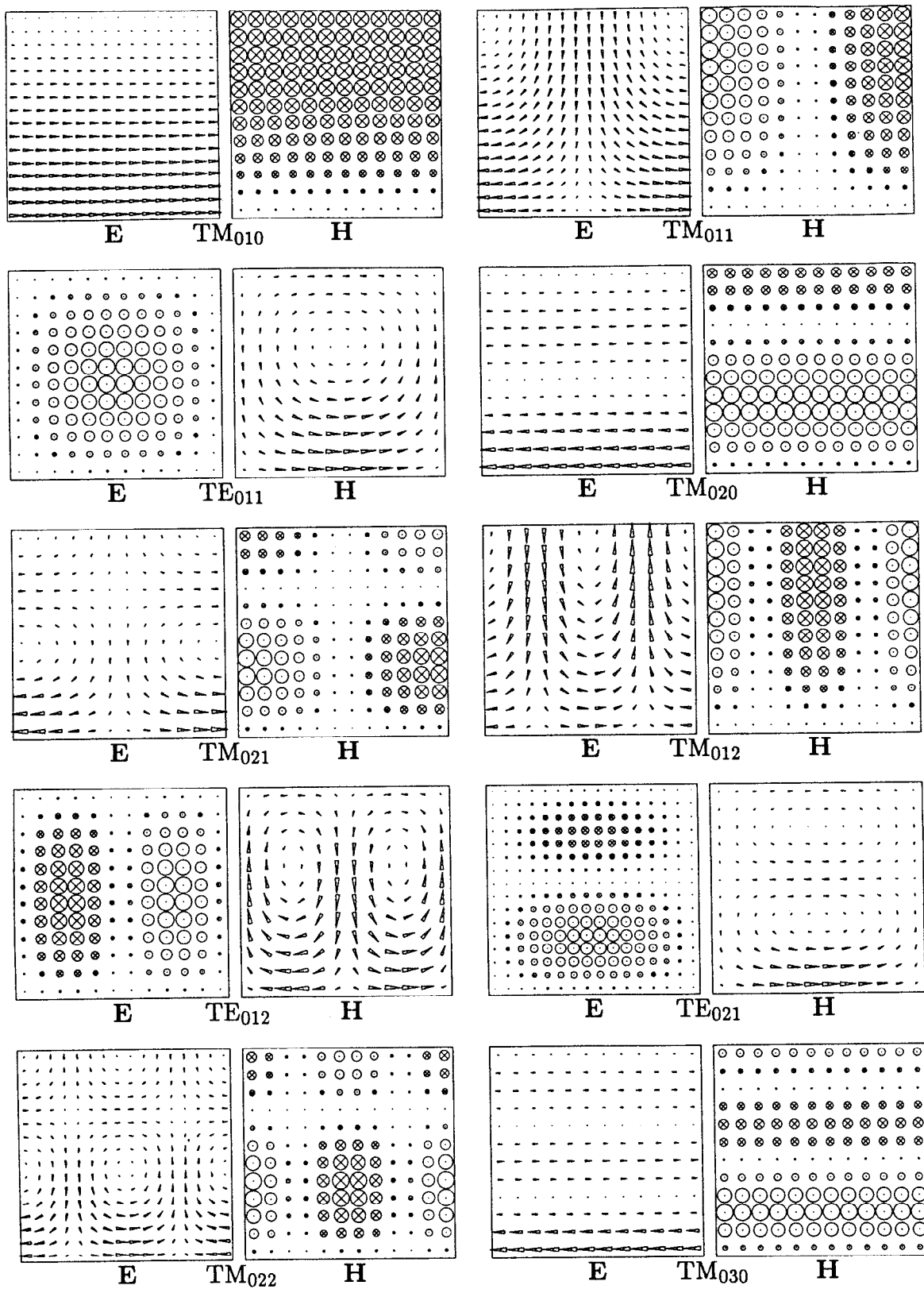


Figure 4.28: Field plots of the 10 lowest $m = 0$ modes of the pillbox. The ϕ -component of the fields is imaginary.

mode	k_{analytic}^2	k_E^2	relative error	k_H^2	relative error
TM ₀₁₀	5.783186	5.783187	1×10^{-7}	5.783197	2×10^{-6}
TM ₀₁₁	15.652790	15.652808	1×10^{-6}	15.653264	3×10^{-5}
TE ₀₁₁	24.551575	24.553544	8×10^{-5}	24.551642	3×10^{-6}
TM ₀₂₀	30.471262	30.471277	5×10^{-7}	30.472366	4×10^{-5}
TM ₀₂₁	40.340867	40.341104	6×10^{-6}	40.347125	2×10^{-4}
TM ₀₁₂	45.261604	45.262002	9×10^{-6}	45.265379	8×10^{-5}
TE ₀₁₂	54.160388	54.170944	2×10^{-4}	54.161366	2×10^{-5}
TE ₀₂₁	59.088061	59.103211	3×10^{-4}	59.088647	1×10^{-5}
TM ₀₂₂	69.949680	69.951961	3×10^{-5}	69.976447	4×10^{-4}
TM ₀₃₀	74.887007	74.887140	2×10^{-6}	74.901008	2×10^{-4}

Table 4.9: Relative error of the computed eigenvalues for the 10 lowest $m = 0$ modes of a pillbox with $a = 1$ and $l = 1$ and using a fine mesh (512 quadratic elements).

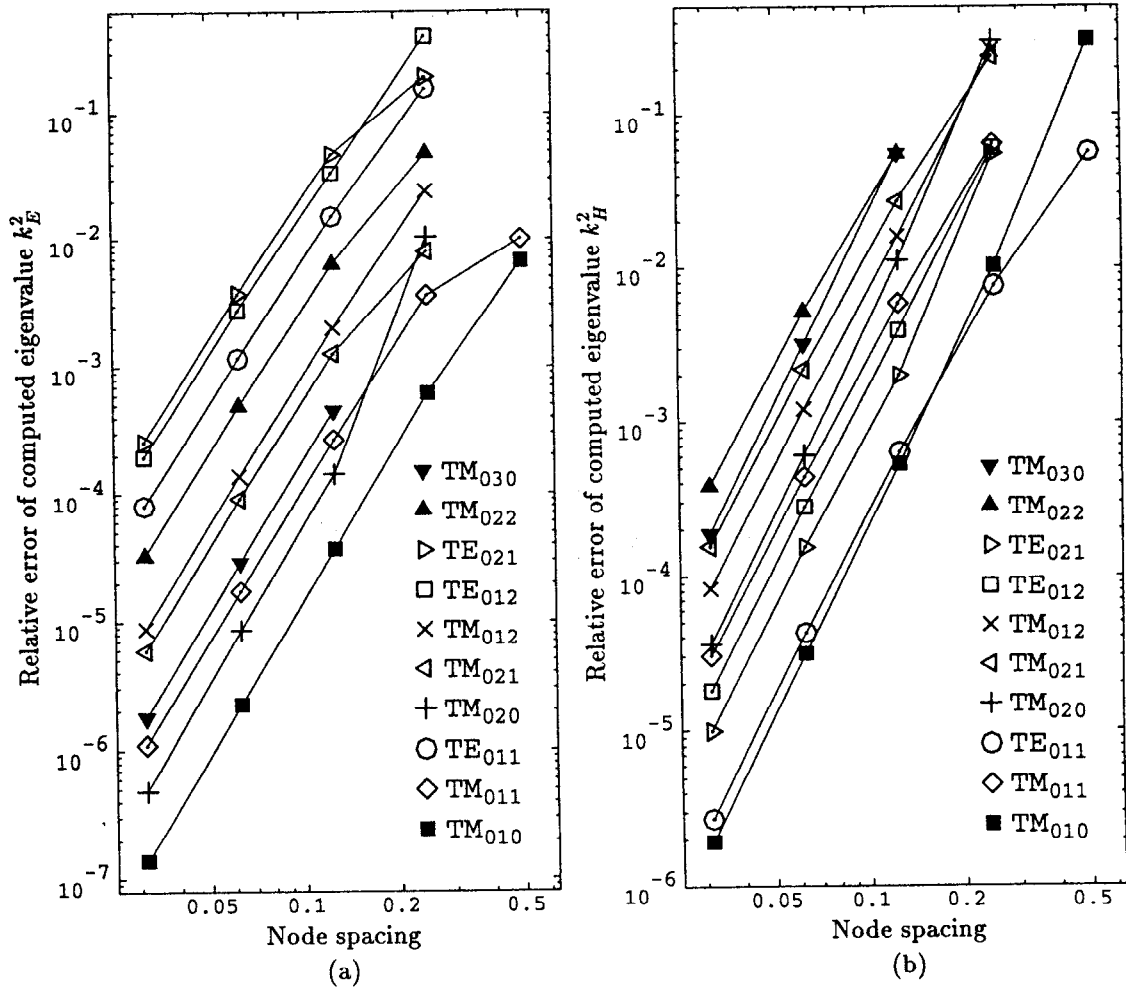


Figure 4.29: Relative error of the computed eigenvalues for the $m = 0$ modes of a pillbox as the mesh is refined. The computed fields are (a) \mathbf{E} and (b) \mathbf{H} .

Chapter 5

Cross-Field Amplifier

The field solver can be applied to the design of slow-wave circuits for cross-field amplifiers (CFAs) and magnetrons. A review of the mechanism by which the RF wave is amplified is given in [23]. A portion of this theory is reviewed in the first two sections below.

A diagram of a CFA is shown in figure 5.1. This CFA is a prototype which was built and tested at SLAC. An RF wave enters the input and is amplified as it passes clockwise through the interaction region. The amplified RF wave is extracted at the output. The sever section allows electrons to drift between the input and output cells while inhibiting the RF wave from propagating from the output cell back to the input cell. Such isolation is necessary to avoid making an oscillator instead of an amplifier.

The corrugation of the anode creates a slow-wave circuit through which the RF propagates. The circuit is called a slow-wave circuit because the phase velocity v_p of the wave is less than the speed of light.

Some slow-wave circuits can be modelled fairly well as two-dimensional structures with fields that are uniform with respect to the third dimension z . Indeed, CFAs and magnetrons are conceptually two-dimensional devices. The third dimension mainly enters as a complication which must be dealt with in real devices, usually in a manner which minimizes or avoids the effect of the third dimension.

Only moderate accuracy is required for the initial two-dimensional design of the slow-wave circuit. Accurate three-dimensional field solutions will be required at a later stage of the design process. In both cases, it is helpful if the field solver can exploit the periodicity of this structure. Examples of two-dimensional and three-dimensional field calculations for various slow-wave circuits are given in section 5.4. The examples illustrate the deficiency of existing finite difference codes when designing structures and the gains which may be made if better algorithms (e.g., finite element methods) are employed.

In order to facilitate the application of the equations in this chapter to real CFAs, a notation is employed to distinguish between MKSA and cgs-gaussian forms of the equations. The equations are written using the cgs-gaussian system of units. However, dropping the factors in brackets will yield equations suitable for the MKSA system of units. Note that $\epsilon_0 = 1$ in the cgs-gaussian system, while in the MKSA system $\epsilon_0 \cong 8.854 \times 10^{-12} \text{ F/m}$. This notation is the same as that employed in

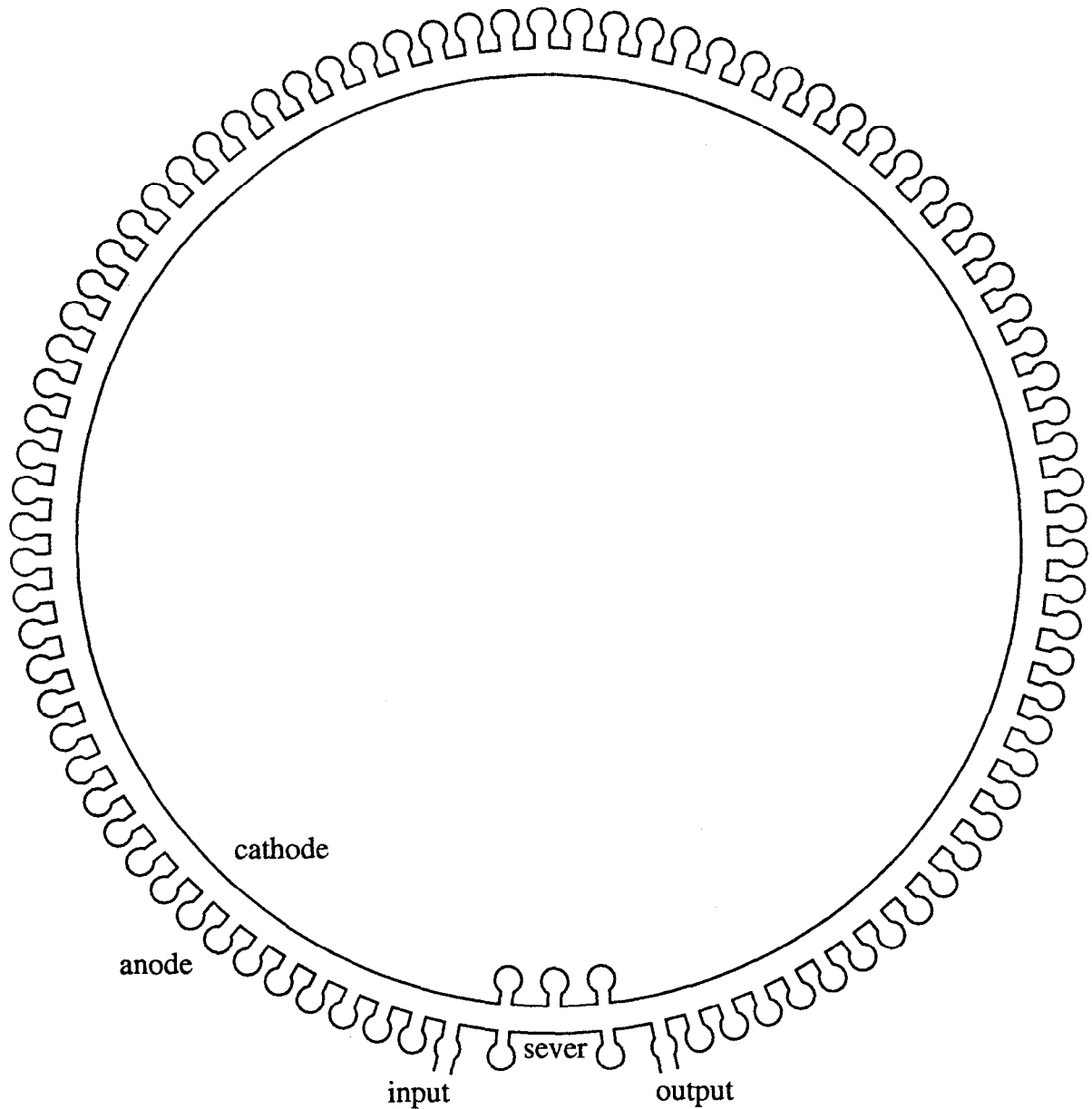


Figure 5.1: The prototype SLAC backward wave cross-field amplifier. The RF wave propagates clockwise, while the electrons travel counter-clockwise.

chapter 8 of Jackson.

5.1. DC Fields and the Brillouin Flow Model

A diagram of a portion of a typical cross-field amplifier is shown in figure 5.2. The interaction region is the space between the anode and the cathode. A slow-wave structure is incorporated into the anode. This particular slow-wave structure is called a hole and slot circuit, which is the structure used in the one CFA built at SLAC. This CFA will be used as an example in this chapter. Selected design

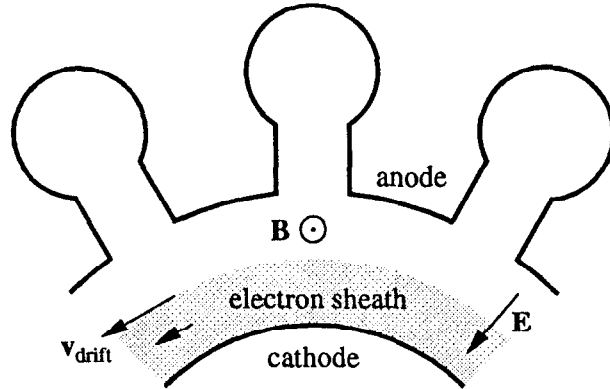


Figure 5.2: Schematic diagram of a cross-field amplifier in the x - y plane. The direction of the static electric and magnetic fields are shown. The electron drift velocity \mathbf{v} is also shown for two locations in the electron sheath.

parameters for this CFA are listed in table 5.1.

A voltage V is applied across the interaction region, and the corresponding static electric field \mathbf{E} pulls electrons toward the anode. A static magnetic field \mathbf{B} perpendicular to the plane of figure 5.2 is also applied. For convenience $B_z > 0$ will be assumed. The magnetic field \mathbf{B} is perpendicular to \mathbf{E} , hence this is called a crossed-field device, and the name cross-field amplifier follows.

Electrons leaving the cathode do not travel straight to the anode. Instead, in the absence of space charge and other RF fields, the magnetic field causes the electrons to follow cycloidal trajectories. If the voltage V does not exceed V_c (defined below) then the electrons do not reach the anode. The electrons are confined to an electron sheath next to the cathode, and the anode and cathode are magnetically insulated from each other.

A simple model of the beam in the CFA is the Brillouin flow model. This model assumes the RF fields are small, which is the situation the CFA or magnetron must start from. A non-relativistic version of the model will be reviewed here. Curvature will be neglected as well, so cartesian coordinates x and y will be used instead of polar coordinates ρ and ϕ . The cathode and anode are at $y = 0$ and $y = d$, respectively. This is shown in figure 5.3.

The electrons drift along the interaction region due to the $\mathbf{E} \times \mathbf{B}$ drift. Laminar flow is assumed in the Brillouin flow model. The electrons travel parallel to the anode and cathode with velocity

$$\mathbf{v}(y) = -\hat{\mathbf{x}}v(y). \quad (5.1)$$

The flow is assumed to be independent of x . That is, the effect of the anode slots

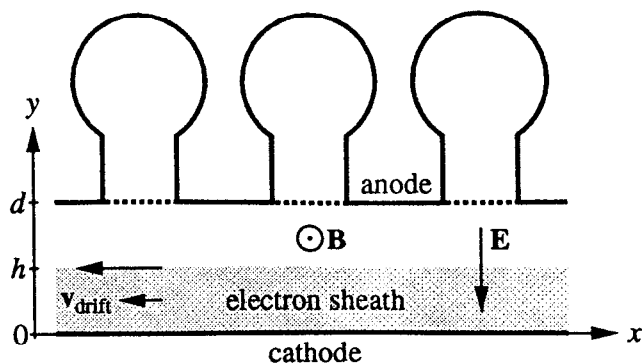


Figure 5.3: Schematic diagram of a straight cross-field amplifier. The dotted lines at the anode indicate the simplified model assumes the anode is smooth.

Frequency f	11.424 GHz
DC magnetic field B_z	5 kG
Anode voltage V	120 kV
Anode-cathode gap d	0.3175 cm
Phase velocity v_p/c	0.248
Phase advance per cell ψ	225°
Cell length L at anode	0.4064 cm
Anode radius r_a	5.8217 cm
Cell length L/r_a	4°
Planar non-relativistic laminar model	
Hull cutoff voltage V_c	220 kV
Electron sheath size h/d	0.32
Buneman-Hartree threshold voltage V_{BH}	102 kV

Table 5.1: Selected SLAC Backward Wave CFA Parameters.

and the RF fields on the electron flow are neglected. The electrons in figure 5.2 move in the $-\hat{x}$ direction (the $\hat{\phi}$ direction in polar coordinates) since $E_y < 0$ and $B_z > 0$.

The equilibrium drift velocity is obtained by balancing the electric and magnetic forces,

$$v(y) = -[c] \frac{E_y(y)}{B_z}. \quad (5.2)$$

Note that the radial electric field E_y is modified by the space charge in the electron sheath. The electric field is reduced near the cathode, so the beam near the cathode drifts slower than the beam in the middle of the interaction region. This is indicated by the relative size of the two arrows representing \mathbf{v} in figures 5.2 and 5.3. The current in the electron sheath also contributes to the magnetic field, but this contribution is neglected in this model. The magnetic field B_z is assumed to be

uniform.

Assuming the electrons leave the cathode with negligible kinetic energy, then conservation of energy gives

$$-|e|V + \frac{1}{2}mv^2 = 0 \quad (5.3)$$

where the electric potential $V(y)$ is

$$V(y) = - \int_0^y E_y(y') dy', \quad (5.4)$$

and e and m are the charge and rest mass of the electron, respectively. Combining equations (5.2) and (5.3) yields the electron velocity

$$v(y) = \left[\frac{1}{c} \right] \frac{|e|B_z}{m} y = \omega_c y \quad (5.5)$$

where ω_c is the cyclotron frequency,

$$\omega_c = \left[\frac{1}{c} \right] \frac{|e|B_z}{m}. \quad (5.6)$$

Then, from (5.3), the potential inside the electron sheath is

$$V(y) = \frac{1}{2} \frac{m\omega_c^2 y^2}{|e|}. \quad (5.7)$$

The corresponding electric field is

$$E_y(y) = \frac{m\omega_c^2 y}{|e|} \quad (5.8)$$

and the corresponding charge density is

$$\rho = \left[\frac{1}{4\pi} \right] \epsilon_0 \frac{m\omega_c^2}{|e|}. \quad (5.9)$$

Notice that the charge density is uniform in the electron sheath, and that the electric field rises linearly from zero at the cathode to its maximum value at the edge of the electron sheath. This is shown in figure 5.4 for the parameters listed in table 5.1.

For a given magnetic field B_z , the Hull cutoff voltage V_c is defined as the anode voltage below which electrons do not reach the anode. If the electron sheath fills the entire interaction region, $0 \leq y \leq d$, then the anode voltage $V = V(d)$ is the Hull cutoff voltage V_c . From (5.7) the Hull cutoff voltage is

$$V_c = \frac{1}{2} \frac{m\omega_c^2 d^2}{|e|} = \left[\frac{1}{c^2} \right] \frac{1}{2} \frac{|e|B_z^2 d^2}{m}. \quad (5.10)$$

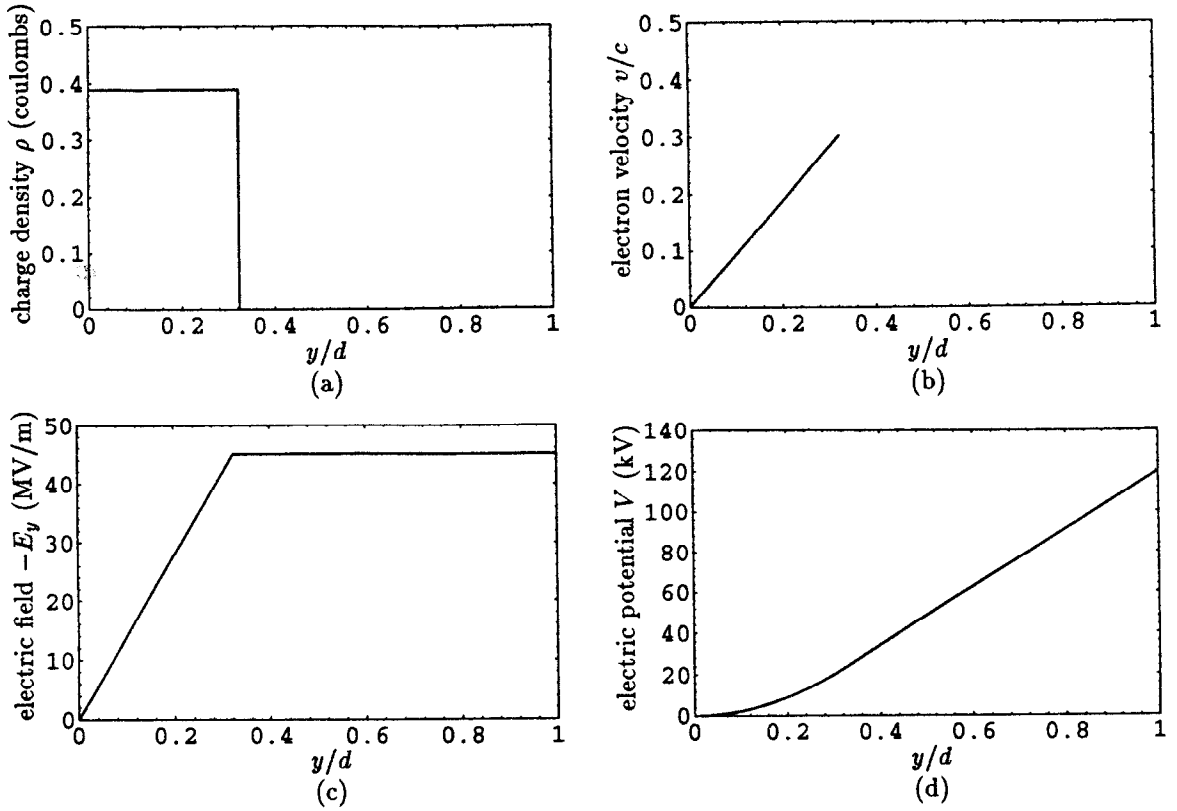


Figure 5.4: Planar non-relativistic Brillouin flow: (a) space charge density ρ , (b) electron velocity v/c , (c) electric field E_y and (d) electric potential $V(y)$.

When the anode voltage V is less than the Hull cutoff voltage there is a gap between the electron sheath and the anode. In the gap there is no charge, so the electric field E_y is constant,

$$E_y = -\frac{m\omega_c^2 h}{|e|} \quad \text{for } h \leq y \leq d, \quad (5.11)$$

where h is the width of the electron sheath. The anode voltage V is

$$V = \frac{m\omega_c^2 h}{|e|} \left(d - \frac{h}{2}\right) = \left[\frac{1}{c^2}\right] \frac{|e|B_z^2 h}{m} \left(d - \frac{h}{2}\right) = \frac{2V_c h}{d} \left(1 - \frac{h}{2d}\right). \quad (5.12)$$

This can be solved for the width h of the electron sheath as a function of anode voltage V and the applied magnetic field B_z ,

$$\frac{h}{d} = 1 - \sqrt{1 - \frac{V}{V_c}} = 1 - \sqrt{1 - \frac{B_H^2}{B_z^2}}, \quad (5.13)$$

where

$$B_H = [c] \sqrt{\frac{2mV}{|e|d^2}} \quad (5.14)$$

is the Hull cutoff magnetic field. If the magnetic field B_z is larger than the Hull cutoff magnetic field B_H (for a given anode voltage V) then the anode and cathode are magnetically insulated from each other.

When the anode voltage V is fixed, then the Hull cutoff magnetic field B_H is the appropriate cutoff to consider. The magnetic field should be $B_z > B_H$. If the magnetic field B_z is fixed instead, then the Hull cutoff voltage V_c is the appropriate cutoff to consider. An anode voltage $V < V_c$ is desired.

5.2. Interaction with the RF Field

The input RF induces a travelling wave in the slow-wave circuit. The electrons interact with this travelling wave, amplifying the wave as it propagates toward the output. Treating the slow-wave circuit as a periodic structure, the travelling wave has a phase advance ψ per cell, and thus the wave in the interaction region can be written as a sum of space harmonics. For the planar structure,

$$B_z = \sum_{n=-\infty}^{\infty} B_{zn} \cosh(\kappa_{yn}y) e^{ik_{zn}x}, \quad (5.15a)$$

$$E_x = \left[\frac{1}{c^2} \right] \frac{ic^2}{\omega} \sum_{n=-\infty}^{\infty} \kappa_{yn} B_{zn} \sinh(\kappa_{yn}y) e^{ik_{zn}x} \quad \text{and} \quad (5.15b)$$

$$E_y = \left[\frac{1}{c^2} \right] \frac{c^2}{\omega} \sum_{n=-\infty}^{\infty} \frac{\psi + 2\pi n}{L} B_{zn} \cosh(\kappa_{yn}y) e^{ik_{zn}x}, \quad (5.15c)$$

where $k_{zn} = (\psi + 2\pi n)/L$ is the propagation constant for the n th space harmonic along the slow-wave structure, and $\kappa_{yn} = \sqrt{k_{zn}^2 - \omega^2/c^2}$ is the corresponding decay constant. The space harmonics are large near the anode and decay to small fields near the cathode.

The space harmonic expansion above is written in terms of the magnetic field amplitudes B_{zn} because the field solver computes the magnetic field B_z . The fields E_x and E_y are derived from the computed B_z . However, the electrons are affected by the RF electric field more than they are by the RF magnetic field in this CFA, and the PIC simulations described later also describe the travelling wave by its electric field, so an expansion in terms of electric field amplitudes E_{zn} is useful,

$$B_z = -[c] \frac{i\omega}{c^2} \sum_{n=-\infty}^{\infty} \frac{E_{zn}}{\kappa_{yn}} \cosh(\kappa_{yn}y) e^{ik_{zn}x}, \quad (5.16a)$$

$$E_x = \sum_{n=-\infty}^{\infty} E_{zn} \sinh(\kappa_{yn}y) e^{ik_{zn}x} \quad \text{and} \quad (5.16b)$$

$$E_y = -i \sum_{n=-\infty}^{\infty} \frac{k_{zn} E_{zn}}{\kappa_{yn}} \cosh(\kappa_{yn}y) e^{ik_{zn}x}. \quad (5.16c)$$

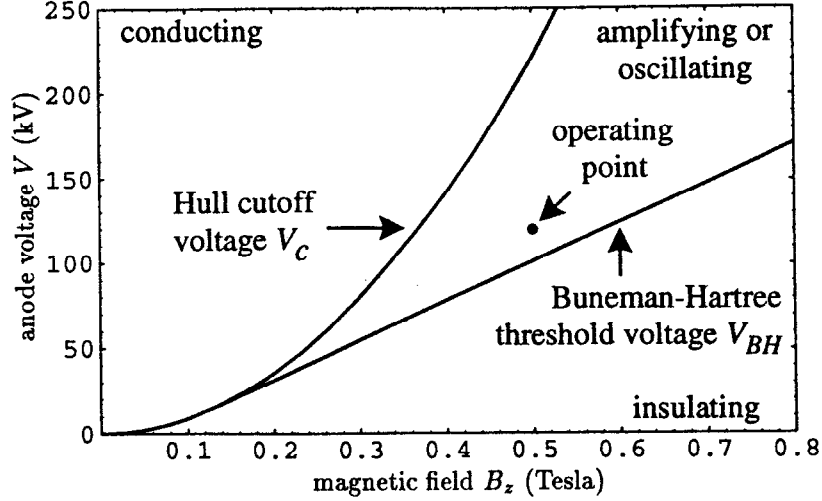


Figure 5.5: Operating point for the SLAC CFA. The Hull cutoff voltage V_c and Buneman-Hartree threshold voltage V_{BH} delimit the three regions where the CFA is expected to (1) conduct, (2) amplify or (3) insulate. The curves are obtained from the planar non-relativistic Brillouin model.

When the RF field is small only one space harmonic affects the electron sheath strongly. The effect of the other space harmonics tends to cancel when averaged over many cycles. However, when the RF field becomes large additional space harmonics may become important. An example of this will be mentioned later.

In order for the n th space harmonic of the travelling wave to interact strongly with an electron, the space harmonic must be synchronous with the electron drifting in the CFA. That is, the electron should be travelling with speed v near v_p , where $v_p = \omega/k_{zn}$ is the phase velocity of the n th space harmonic. The fastest electrons are at the edge of the electron sheath nearest the anode. The Buneman-Hartree threshold voltage V_{BH} is the anode voltage at which the speed $v(h)$ of the fastest electrons equals the phase velocity v_p of the desired space harmonic. At this threshold the electrons at the edge of the sheath barely keep up with the travelling wave. The remaining electrons are slower, and thus lag behind the wave and do not interact strongly with the wave.

From (5.5), $h = v_p/\omega_c$ at the Buneman-Hartree threshold. Plugging this height into (5.12) yields the Buneman-Hartree threshold voltage

$$V_{BH} = \left[\frac{1}{c} \right] B_z v_p d - \frac{m}{2|e|} v_p^2. \quad (5.17)$$

The CFA is operated with $V_{BH} < V < V_c$.

Figure 5.5 shows the operating point in the V - B_z plane for the SLAC CFA. The figure labels the region $V > V_c$ where the CFA is conducting, and the region

$V < V_{BH}$ where the CFA is insulating. The two curves are obtained from the planar non-relativistic Brillouin flow model. Between the two curves is the region where the CFA is expected to amplify the incoming RF. This region is discussed further below. The operating point is about 20% above V_{BH} , which is typical for CFAs and magnetrons.

Figure 5.6a shows the electric field for one period of the synchronous space harmonic of the travelling wave in the SLAC CFA. Note that other space harmonics of the travelling wave are omitted from the field plot. For comparison, figure 5.6b shows the static electric field and figure 5.6c shows the combined field when the magnitude of the electric field of the space harmonic is 2/3 of the DC electric field at the anode. When the RF electric field is this large, the peak RF magnetic field is only 0.025 Tesla. This is small compared to the DC magnetic field of 0.5 Tesla, so the RF magnetic field will be neglected when considering the effect of the RF field on the particles.

The amplification mechanism in the CFA can be understood by following the trajectories of four electrons labelled A, B, C and D in figure 5.6. The four electrons are near the edge, or beyond the edge, of the electron sheath. They are drifting to the left with a speed close to the phase velocity v_p of the space harmonic of the travelling wave.

Consider electrons B and D. Electron B experiences a decelerating electric field, and thus does work on the RF field. That is, the amplitude of the RF field is being increased at the expense of the electron's energy. Since the particle is synchronous with the wave, the particle experiences this field for a long time. Also note that the $E \times B$ drift has a small vertical component which pushes the particle toward the anode. As the particle drifts toward the anode the magnitude of the decelerating RF field increases, which further increases the interaction between the electron and the RF field.

Electron D is at the opposite RF phase relative to electron B, thus electron D is accelerated by the RF field. That is, energy is transferred from the RF field to the electron. This transfer of energy is limited by the small vertical component of the $E \times B$ drift. This component pushes the electron toward the cathode, where the magnitude of the RF field is smaller, and thus the interaction between the electron and the field is reduced.

Electron D may be pushed into the middle of the sheath, where the reduced static electric field implies a reduced drift velocity, and thus the electron will no longer be synchronous with the travelling wave. As the electron loses synchronism

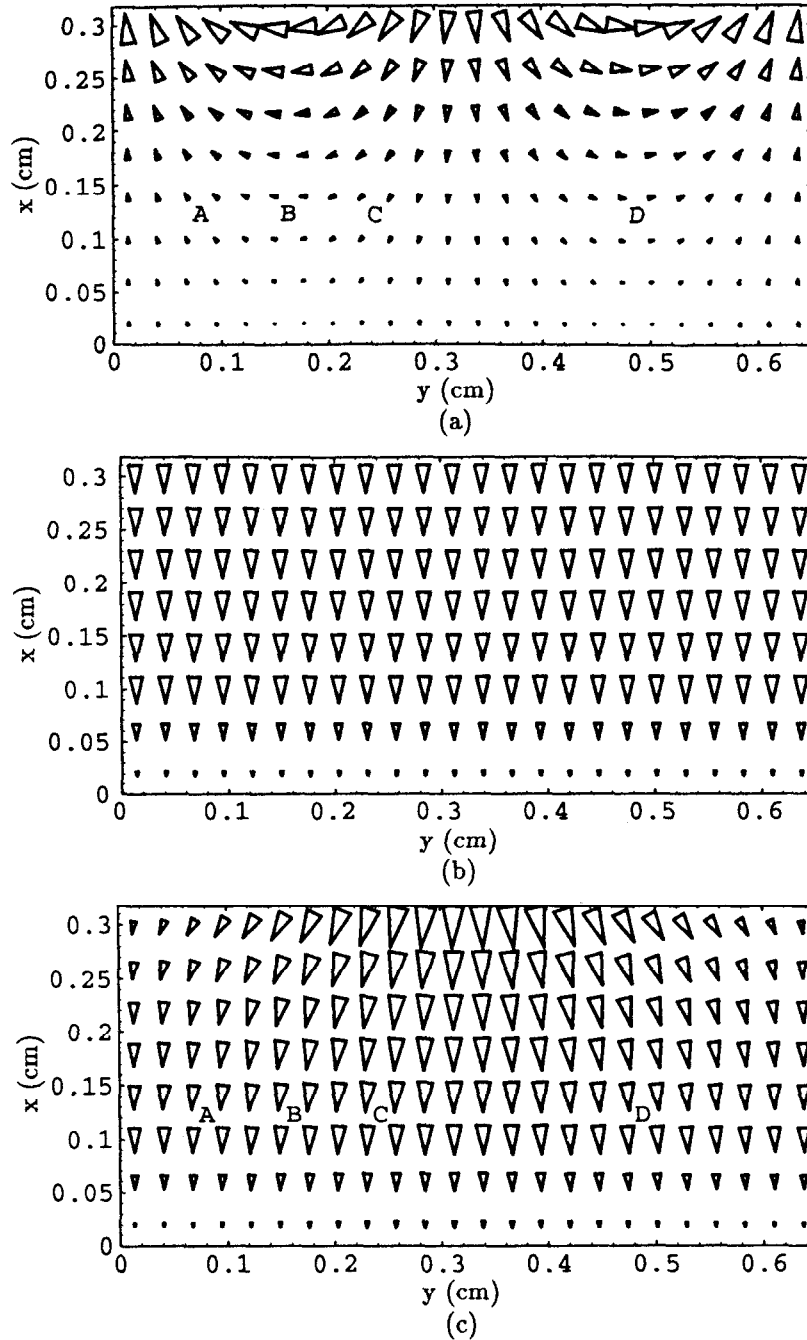


Figure 5.6: One wavelength of the electric field of a single space harmonic of the RF travelling wave: (a) RF field, (b) static field and (c) combined field.

with the travelling wave, the interaction between the electron and the wave is greatly reduced. This electron may appear later at a phase corresponding to electron B due to phase focussing, which will be discussed shortly.

Electron D may also be pushed into the cathode proper, which removes the electron from the interaction region, and no further interaction between this electron

and the RF wave occurs. If the electron strikes the cathode within a certain energy range, then the cathode will emit secondary electrons. Low energy electrons scatter few of the cathode's electrons, while high energy electrons scatter many of the cathode's electrons, but these are buried so deep in the cathode that the scattered electrons are stopped in the cathode and do not escape.

This leads to a short digression on the creation of electron sheaths in CFAs and magnetrons. In the SLAC CFA secondary emission is the mechanism which creates the electron sheath. Some electrons, such as those near electron D, are accelerated into the cathode. Most of the electrons striking the cathode have an energy in the range which generates multiple secondary electrons. Thus, there is a gain mechanism which rapidly builds the electron sheath from a few electrons. The original few electrons are obtained, for example, from field emission.

Many CFAs and magnetrons use a thermionic cathode instead of a secondary emission cathode. A hot cathode emits electrons due to thermionic emission, and it is these electrons which constitute the electronic sheath.

Now consider electrons A and C. At electron A the RF electric field opposes the DC electric field, so the drift velocity of electron A is reduced slightly below the phase velocity of the space harmonic and the velocity of electron B. This means the electron A will lag behind the travelling wave, changing its phase with respect to the RF until the electron is near electron B in figure 5.6.

Likewise, at electron C the RF electric field enhances the DC electric field, so the drift velocity of electron C is increased slightly above the velocity of electron B and the phase velocity of the space harmonic. Electron C will change phase with respect to the travelling wave, catching up to electron B.

Both electrons A and C are focussed toward the phase of RF where electron B resides. This is called phase focussing. Also notice that electrons A and C have a small vertical component of $E \times B$ drift velocity which pushes them toward the anode, just like electron B is pushed toward the anode.

The combination of electron selection and phase focussing leads to amplification of the travelling wave. Electrons which take energy from the RF field, such as electron D, are removed from the interaction region, while other electrons are focussed into a phase relative to the RF where they give their energy to the RF field. These electrons maintain synchronism with the travelling wave as they drift toward the anode, so a spoke composed of electrons forms at a phase near electron B. The spoke travels with the RF field and extends from the cathode to the anode. Electrons in the spoke are drifting from the electron sheath to the anode,

converting their potential energy into energy in the travelling wave.

5.3. Simulations of the Electron-Circuit Interaction

The simple model of the CFA above is qualitatively correct, but quantitative results of the electron-circuit interaction have only been obtained through the use of particle-in-cell (PIC) simulations of the interaction region. The PIC simulations for the SLAC CFA are described in [24] and [25]. The simulations model only the interaction region, not the whole slow-wave circuit. Instead, eigenmodes of the slow-wave circuit were computed separately using, for example, the finite element field solver of this work. Then selected RF parameters of the travelling wave modes were incorporated into the PIC simulation, as will be described below.

An integral number of wavelengths of the space harmonics are included in the simulation. If $0 \leq x \leq l$ is the simulation region and k_{xn} is the propagation constant for the synchronous space harmonic, then $k_{xn}l = 2\pi p$ for some integer p is required. If a second space harmonic (the m th harmonic) is included in the simulation then $k_{xm}l = 2\pi p'$ for some integer p' is also required. Periodic boundary conditions are employed at the ends of the simulation. Some simulations modelled the RF field as a single space harmonic, while others modelled the RF field with two space harmonics.

As in any PIC simulation, for each time step the momentum of each particle is modified according to the electric and magnetic fields experienced by the particle. The particles are pushed according to their momentum, and the field is advanced according to Maxwell's equations.

The advancement of the field from one time step to the next requires knowledge of the interaction between the electrons and the fields of the slow-wave circuit. The simulations employ boundary conditions at the anode which mimic the slow-wave structure. At $y = d$ the tangential electric field is forced to be a travelling wave. This travelling wave corresponds to one or two space harmonics of the RF wave. To update the amplitude and phase of the RF wave, the work by the the electrons on the field is computed by

$$P = \int_{\Omega} \mathbf{E}_{rf} \cdot \mathbf{J} d\Omega, \quad (5.18)$$

where \mathbf{E}_{rf} is the electric field of the travelling wave, \mathbf{J} is the current density, and Ω is the simulation region. The power P is the power delivered to the RF field in the simulation region. The field \mathbf{E}_{rf} is considered a complex field with the imaginary component differing from the real component by one quarter of a cycle. Hence the power P has both real and imaginary components. The real part represents a

change in the amplitude of the wave, while the imaginary part represents a change in the phase of the wave.

The relation between the power P and the change of the field amplitude and phase is described by the impedance Z of the slow-wave circuit. The power flow through the circuit is

$$P_{rf} = \frac{V_n^2}{2Z} = \frac{E_{xn}^2}{2k_{xn}^2 Z}, \quad (5.19)$$

where $V_n = E_{xn}/k_{xn}$ is a voltage obtained from the space harmonic electric field amplitude E_{xn} . Other definitions for the impedance Z are possible, but this definition is convenient because it relates power flow through the circuit to the synchronous space harmonic field amplitude. The final product of the CFA is supposed to be power flowing out the output waveguide. The power flow through the circuit is also obtained from

$$P_{rf} = \frac{v_g}{l} \mathcal{U}, \quad (5.20)$$

where \mathcal{U} is the energy stored in the RF wave in the simulation region and v_g is the group velocity of the RF wave. From the last two equations,

$$P = \frac{\partial \mathcal{U}}{\partial t} = \frac{l}{v_g} \frac{V_n}{Z} \frac{\partial V_n}{\partial t}, \quad (5.21)$$

or

$$\frac{\partial E_{xn}}{\partial t} = v_g Z \frac{k_{xn}^2}{E_{xn}} \frac{P}{l}. \quad (5.22)$$

5.4. Tests on a Hole and Slot Circuit

In addition to designing slow-wave circuits with the appropriate phase velocity, the eigenmode field solver can be used to compute the RF parameters of the slow-wave circuit for use in the PIC simulation. These parameters include the group velocity v_g and the impedance Z .

Figure 5.7 shows the measured dispersion curve for the prototype SLAC CFA. Dispersion curves computed by two finite difference programs are also shown. The finite difference program QUAP models the planar structure in two dimensions, while the finite difference field solver ARGUS models the planar structure in three dimensions. The version of ARGUS used here was modified to include quasi-periodic boundaries, and thus handle periodic structures.

The agreement between the two finite difference calculations is accidental. The two methods are not expected to agree so well. The difference between the ARGUS

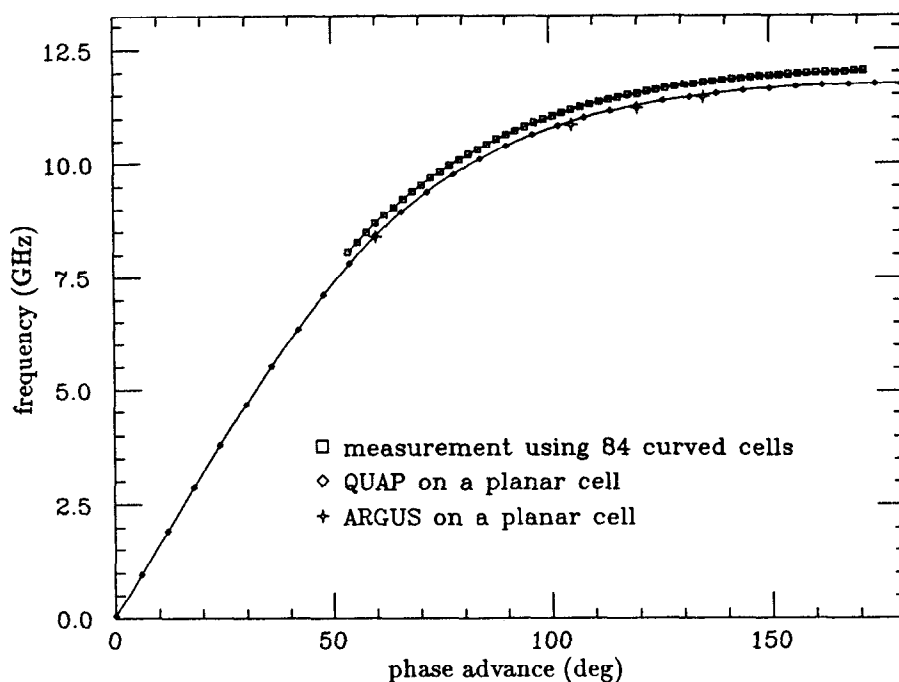


Figure 5.7: Comparison of computed and measured dispersion diagrams for a slow-wave circuit.

calculations and measurement, however, is significant. Two-dimensional simulations of straight and curved structures show that the effect of curvature in this problem is very small, so the discrepancy between the dispersion curves is due to discretization errors in the computer model. In particular, the finite difference mesh does not model the boundary of the structure well. Reduction of the discretization error to levels comparable to dimensional tolerances requires a very fine finite difference mesh. Such problems require very fast computers with hundreds of megabytes of core memory. Alternatively, a field solver with a mesh which models the three-dimensional structure well (e.g., a finite element mesh) could compute the modes of the slow-wave circuit with much better accuracy and perhaps on only a modest computer.

Both finite difference and finite element field solvers have been employed to solve the planar hole and slot structure in two dimensions. This slow-wave structure is similar to the structure shown in figure 5.3. Figure 5.8 shows calculations of the lowest travelling wave mode at a phase advance $\psi = 135^\circ$.

Dispersion diagrams are conventionally drawn for phase advances $0 \leq \psi \leq \pi$, or $0^\circ \leq \psi \leq 180^\circ$. Any other phase advance can be obtained from the symmetry of the dispersion diagram. For example, the mode with $\psi = 135^\circ$ includes a $\psi = -225^\circ$ space harmonic. Simply reversing the direction of propagation (i.e., conjugating

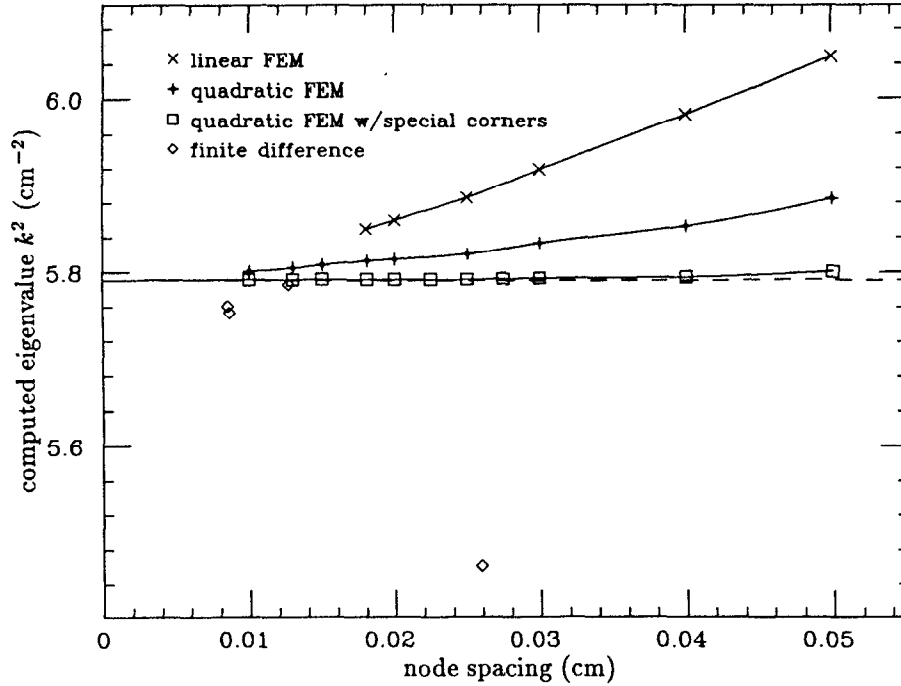


Figure 5.8: Computed eigenvalue k^2 versus mesh refinement for the 135° mode of a hole and slot circuit for the SLAC CFA. Finite difference and various finite element methods are shown. The dashed line is the extrapolation of the quadratic FEM calculations with special corner elements to zero node spacing.

the complex field) gives the mode with the desired synchronous space harmonic at $\psi = 225^\circ$.

The finite difference calculations suffer from errors much more than the finite element calculations. Compared to the linear finite element example, much of the finite difference error is due to the inability of the finite difference mesh to conform to the geometry. Such errors are detrimental because the group velocity of the structure is small near the operating point on the dispersion diagram. An error in the frequency at a desired phase advance corresponds to a large error in the phase advance and phase velocity at the drive frequency.

Even the finite element methods have some difficulty because the design model for the hole and slot structure has sharp corners in it. Sharp corners deteriorate the accuracy of the field solvers. However, the utility of the quadratic finite element field solver with special corner elements is apparent in figure 5.8.

The effect of curvature should also be included in the model. This is easily included in the finite element simulation, while in the finite difference simulations much additional work is required. However, for this slow-wave circuit the curvature is small. Finite element method calculations show that the $\psi = 135^\circ$ frequency in

the curved structure is only 0.07% higher than the planar structure.

5.5. Other Slow-Wave Circuits

The PIC simulations described earlier found problems with the prototype SLAC CFA design. When the amplitude of the RF travelling wave becomes sufficiently large, the electrons accelerated into the cathode are so energetic that the scattered electrons do not escape the cathode. That is, the secondary emission ratio at high energies is insufficient to maintain the electron sheath. This problem has been attributed to the presence of the fundamental -135° space harmonic. The fields of the fundamental space harmonic are much larger at the cathode than the fields of the synchronous space harmonic. Thus, the fundamental harmonic may have a larger effect on the electrons near the cathode than the synchronous harmonic.

Consider two solutions to this problem. First, the synchronous harmonic and the fundamental harmonic can be made the same harmonic by using a slow-wave circuit with a different period and phase advance. This yields a forward wave structure, where the electron sheath and the travelling wave propagate in the same direction. The prototype CFA is a backward wave device. Simulations show that the RF field threshold at which the electron sheath is no longer maintained by secondary emission is lower in the backward wave device than in the corresponding forward wave device.

Backward wave devices have the convenient property that the electron sheath velocity is never synchronous with the π -mode as the anode voltage is applied to the CFA. As the anode voltage is increased from zero to the operating voltage, the drift speed of the electrons starts from zero and increases to the operating phase velocity. In the forward wave device, the drift speed must sweep through the phase velocity of the π -mode, which is a mode particularly susceptible to oscillation. These oscillations can be avoided in the forward wave device by having only a short rise time on the anode voltage. Then the π -mode oscillation will not have time to build.

Even when a forward wave device is employed, there is still a limit to the RF fields. Instead of electrons being slammed back into the cathode with the help of other space harmonics, the forward wave device strips the electron sheath from the cathode, leaving no electrons in the interaction region to regenerate the sheath via secondary emission.

Since the amplitude of the RF field is limited, a tapered impedance structure is desired. Near the input, where the RF power is low, the impedance is high. Near the output, where high RF power is expected, the impedance is low. The RF fields

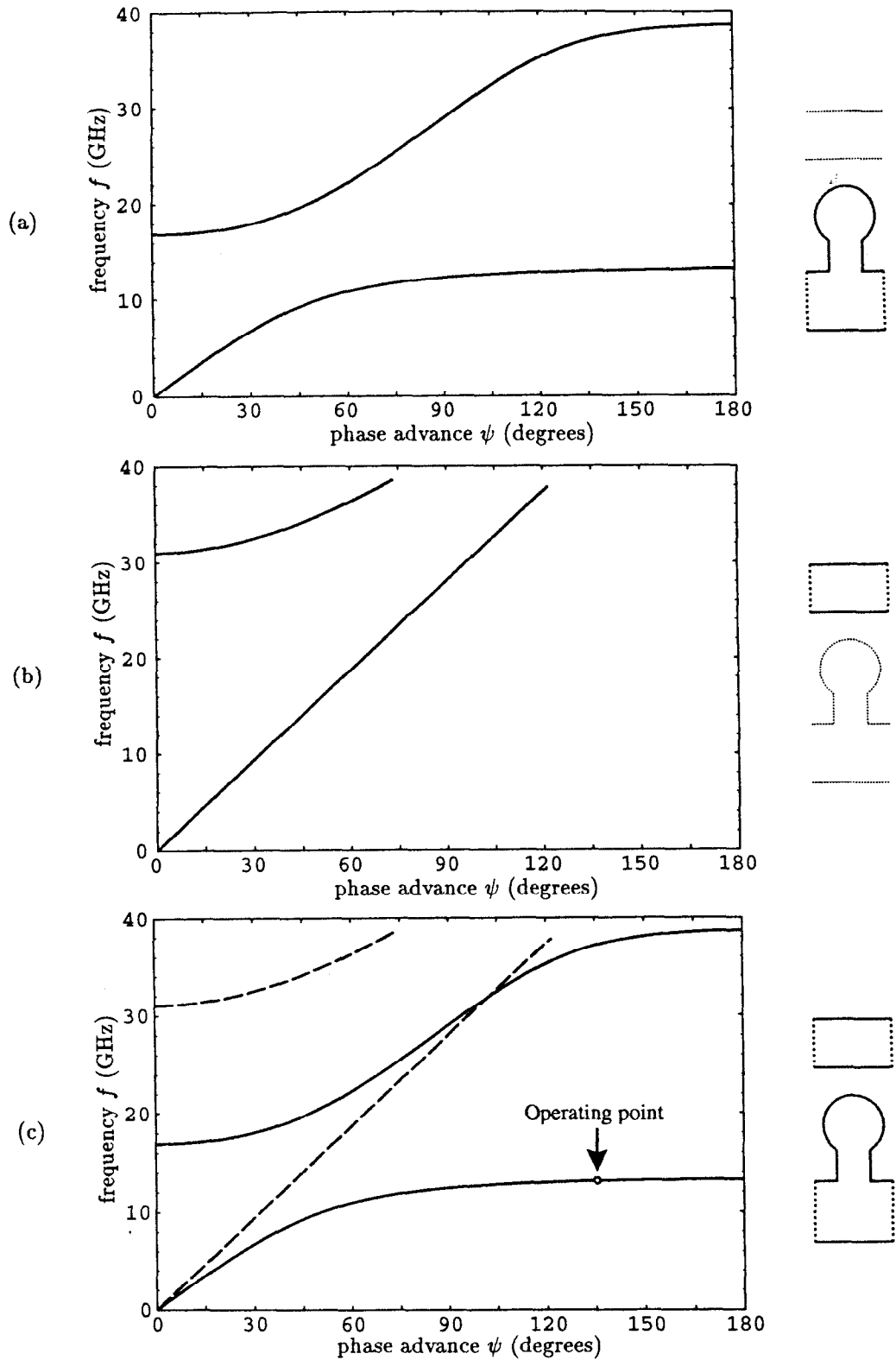


Figure 5.9: Dispersion curves for (a) a slow-wave circuit and (b) a smooth waveguide. The operating point and both dispersion curves are shown in (c). The dashed line in (c) is the smooth waveguide.

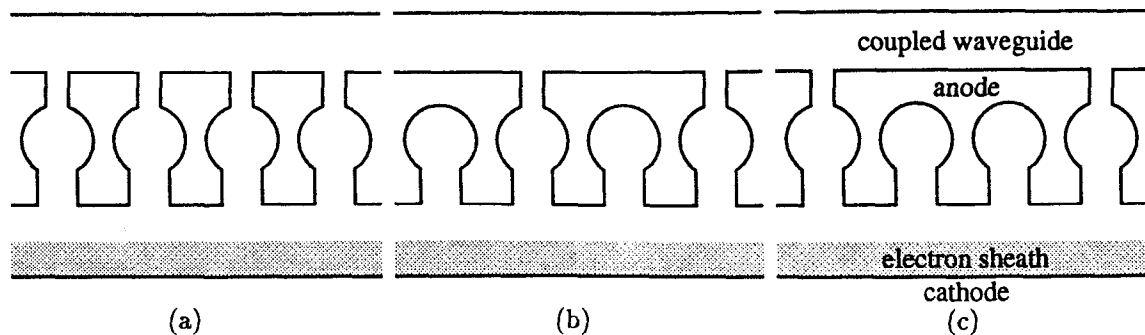


Figure 5.10: Waveguide coupled slow-wave circuits. The guides are coupled (a) every cell, (b) every other cell and (c) every third cell.

can be kept roughly constant throughout the device.

In order to vary the impedance over a wide range, the slow-wave circuit is coupled to a waveguide as shown in figure 5.10a. Near the input the waveguide is narrow, so only a little RF power is propagated in the waveguide. This implies an impedance comparable to the uncoupled slow-wave circuit. Near the output the waveguide is wide. Most of the RF power is propagated in the waveguide. Only a little RF power is propagated in the interaction region portion of the slow-wave circuit. This implies a very low impedance.

Coupling the slow-wave circuit to the waveguide is not trivial because the phase velocity of a smooth waveguide is necessarily above the speed of light, c , while the phase velocity of the desired slow-wave circuit is much less than c . In order for the slow-wave circuit to couple well to the waveguide, the dispersion curves of the slow-wave circuit and the waveguide should cross near the desired operating point. As shown in figure 5.9c, this is not likely.

One possible solution to this coupling problem is to couple to another slow-wave guide instead of to a smooth waveguide. The phase velocities of the two structures can be made identical at the operating point.

Another solution, which will be discussed in some detail here, is to couple the slow-wave circuit to the waveguide at only every other cell, or every third cell. Examples of such structures are shown in figures 5.10b and 5.10c. The period of these structures is double or triple the period of the uncoupled structure. One period of these structures will be called a supercell. Only one hole of the two or three holes in a supercell are coupled to the waveguide.

Dispersion curves for the slow-wave circuit interpreted as a two slot or three slot circuit are shown in figures 5.11a and 5.11b. The supercells are composed of repetitions of the single hole-and-slot cell and the corresponding length of guide.

The solid lines are the slow-wave circuit dispersion curves, while the dashed lines are the waveguide dispersion curves. The curves are identical to figure 5.9c, except that the horizontal axis has been scaled and then folded back upon itself two or three times. The number of folds is the number of cells in the supercell.

The operating mode, which formerly had phase advance ψ , now has phase advance 2ψ or 3ψ . This means the phase velocity of the operating mode can now be greater than c , while the synchronous space harmonic still has the same phase velocity $v_p < c$.

Now consider the operating mode with phase advance $\psi = 135^\circ$ across one cell of the slow wave structure. Across two cells the phase advance is $\psi = 270^\circ = -90^\circ$. At the same frequency the waveguide could have a mode at $\psi = -45^\circ$. Over two cells the phase advance is $\psi = -90^\circ$, the same phase advance seen in two cells of the slow wave structure. If the slow-wave circuit is coupled to the waveguide at every other cell, then the dispersion diagrams cross at this frequency and phase advance, as shown in figure 5.11a. The operating mode in the slow-wave circuit can be coupled well to the waveguide.

The dispersion diagram of the coupled structure is not identical to the dispersion diagrams of the uncoupled slow-wave circuit and waveguide. In the coupled circuit the dispersion curves avoid crossing one another. This can be seen comparing figure 5.11b with figure 5.11c. The slow-wave circuit couples to the waveguide at points where the dispersion curves in figure 5.11b cross. The amount of coupling is related to the distance separating the dispersion curves at the corresponding point in figure 5.11c. Large separation corresponds to large coupling, while small separation corresponds to small coupling.

A potential difficulty with coupled structures is the presence of additional competing modes. In the uncoupled backward wave CFA, only one mode is ever synchronous with the electron sheath, so only one mode needs to be considered. As mentioned earlier, the π mode is a concern in forward wave CFAs. In coupled structures there are many modes synchronous with the electron sheath and with a frequency near the operating frequency. These modes potentially interfere with the operating mode, so the interaction of all of these modes with the electron sheath must be considered in the design of a coupled CFA circuit.

Competing modes also limit the height h of the slow-wave circuit. A large height is desired to obtain as much power as possible, but as the height increases modes with one or more wavelengths of vertical field variation have a frequency which is close to the operating frequency. Thus these modes will compete with the

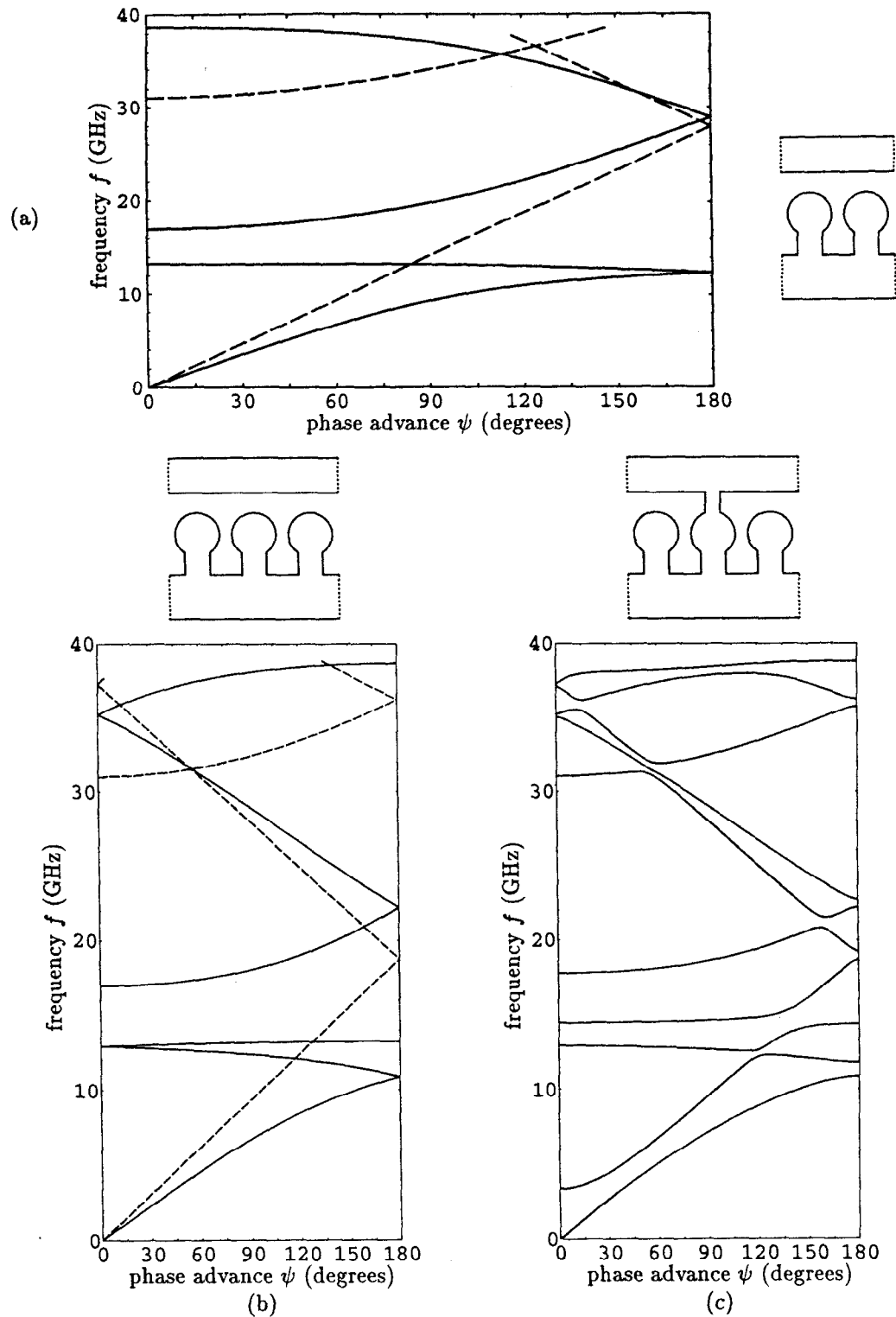


Figure 5.11: Dispersion diagrams of multiple cell slow-wave circuits. Diagrams (a) and (b) treat the uncoupled structure as a periodic structure with a period twice and three times the length of the original cell. The dashed lines are the dispersion curves of the waveguide. Diagram (c) is the dispersion diagram of the structure coupled every third cell.

operating mode if the height of the slow-wave circuit is too large. Three-dimensional field solvers can be used to compute these competing modes, as well as to verify the vertical uniformity of the field in the operating mode.

Chapter 6

X-band Accelerator Structure

In collaboration[26] with Juwen Wang, the field solver YAP was used as an integral part of the design procedure for the detuned accelerator structure for SLAC's Next Linear Collider (NLC) program[27]. The design procedure and some results are described in this chapter.

6.1. Accelerating Mode

The operating frequency of the accelerator is chosen to be $f = 11.424$ GHz, which is 4 times the frequency of the SLAC linac. The phase advance per cell is chosen to be $\psi = 2\pi/3$.

The phase advance is chosen to be a simple fraction of π so that standing wave test structures can be built which are only a few cells long. In this case one wavelength of the accelerating mode is exactly three cells of the structure. Assuming the cells have a symmetry plane or are nearly symmetric, then a short standing wave structure consisting of some multiple of $1\frac{1}{2}$ cells and shorted at symmetry planes at both ends will have a standing wave mode which corresponds to two travelling wave modes with phase advance $\psi = 2\pi/3$ propagating in opposite directions.

A $\pi/2$ phase advance is often chosen for travelling wave structures, particularly if a large bandwidth is desired. For this accelerator structure $\psi = 2\pi/3$ is chosen because the accelerating gradient for a given input power is better at $\psi = 2\pi/3$ than at the common phase advance $\psi = \pi/2$.

The travelling wave which accelerates the beam must be synchronous with the beam. That is, the phase velocity v_ϕ of the travelling wave must be the same as the velocity v of the beam. For the NLC the electron and positron beams are ultrarelativistic, so the beam velocity is the speed of light, c , and the synchronism condition is $v_\phi = c$.

Given the desired operating frequency f , phase advance ψ and phase velocity c of the travelling wave, then the length, or period L , of the cells of the accelerator structure is fixed at $L = c\psi/2\pi f = 0.8748$ cm

Synchronism with the beam is a constraint on the remaining dimensions of a cell. The remaining dimensions are the disk aperture $2a$, the cell diameter $2b$ and the disk thickness t . As shown in figure 6.1, the inner edge of the disks are round with full radius, not flat. The RF parameters of a cell can be computed by treating the cell as part of a periodic structure. Let $f_o(2a, 2b, t)$ be the accelerating mode

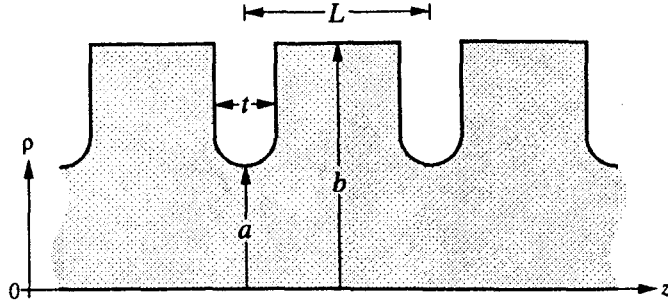


Figure 6.1: Disk loaded waveguide in the z - ρ plane.

frequency at $\phi = 2\pi/3$. Then the synchronism constraint is

$$f_o(2a, 2b, t) = f = 11.424 \text{ GHz.} \quad (6.1)$$

This can be considered an implicit formula for $2b(2a, t)$ for the cells of the accelerator structure. That is, given the parameters $2a$ and t of a cell, solve (6.1) for $2b$.

The finite element field solver YAP was used to compute the accelerating mode frequency f_o for various cell dimensions covering the range

$$\begin{aligned} 0.7493 \text{ cm} &\leq 2a \leq 1.1684 \text{ cm} \\ \text{and } 0.1016 \text{ cm} &\leq t \leq 0.2540 \text{ cm.} \end{aligned}$$

The range of cell diameters was

$$2.0828 \text{ cm} \leq 2b \leq 2.3622 \text{ cm,}$$

although the range of $2b$ was not covered uniformly. As indicated below, cell diameters $2b$ were chosen such that f_o was within about 200 MHz of the drive frequency 11.424 GHz.

For a given set of dimensions ($2a$, $2b$ and t), meshes for one cell of a periodic structure were constructed using triangular elements. Two cases are shown in figure 6.2. Notice that the size of the elements is far from uniform. There are more elements near the disk edges in order to model the boundary and the fields better. The bottom of the mesh is the axis, and the right and left sides of the mesh are quasi-periodic boundaries. The number of elements and the topology of the meshes are independent of the cell parameters, so the mesh depends smoothly on the cell dimensions. Then the discretization error can be considered a systematic error when comparing two cells with slightly different dimensions.

The accelerating mode frequency f_o was calculated by YAP using four successively refined meshes composed of quadratic elements. The calculations converged

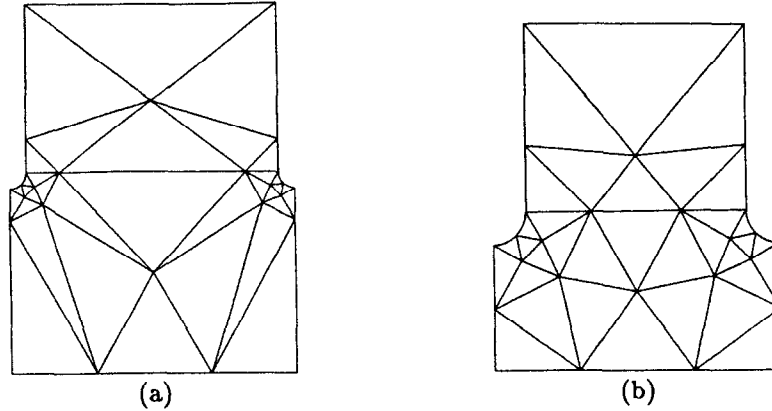


Figure 6.2: Two coarse meshes for cells of the disk loaded waveguide. The meshes are composed of quadratic triangular elements. Mesh (a) is for the first cell of the structure and mesh (b) is for the last cell of the structure.

smoothly as the mesh was refined, although the error appeared to be $\mathcal{O}(h^{3.3})$ instead of $\mathcal{O}(h^4)$. This is probably due to the non-uniform nature of the mesh. The estimated relative accuracy of the frequency calculation for the finest mesh is $\sim 10^{-6}$. Further accuracy was obtained by extrapolating the four calculations to zero element size (an infinitely refined mesh). Conservative error estimates for the extrapolated f_o range from 5 kHz for large t cases to 30 kHz for small t cases.

The calculations were fit to the polynomial

$$f_o(2a, 2b, t) = \sum_{i=-1}^1 \sum_{j=0}^2 \sum_{k=0}^4 \tilde{f}_{ijk}(2b)^i (2a)^j t^k. \quad (6.2)$$

Notice that the leading dependence of the frequency on the cell diameter $2b$ is

$$f_o(2a, 2b, t) = \frac{K}{2b} + \dots \quad (6.3)$$

where K is a constant depending on $2a$ and t . This is similar to the frequency of a pillbox model of the cell. Only 165 calculations with $|f_o - 11.424 \text{ GHz}| < 150 \text{ MHz}$ were included in the least squares fit for the 45 parameters. The polynomial approximates the calculations with residual errors $< 80 \text{ kHz}$. The polynomial can be used to solve (6.1) for the cell diameter $2b$ with an error $\sim 1 \mu\text{m}$. This error is comparable to the skin depth in copper at the operating frequency and about an order of magnitude smaller than available machining tolerances. A contour plot of $2b(2a, t)$ is shown in figure 6.3.

The logarithmic derivative $-b/f \partial f_o / \partial b$ is plotted in figure 6.4. If the cell were a simple pillbox with no aperture connecting the cells, then the expected derivative

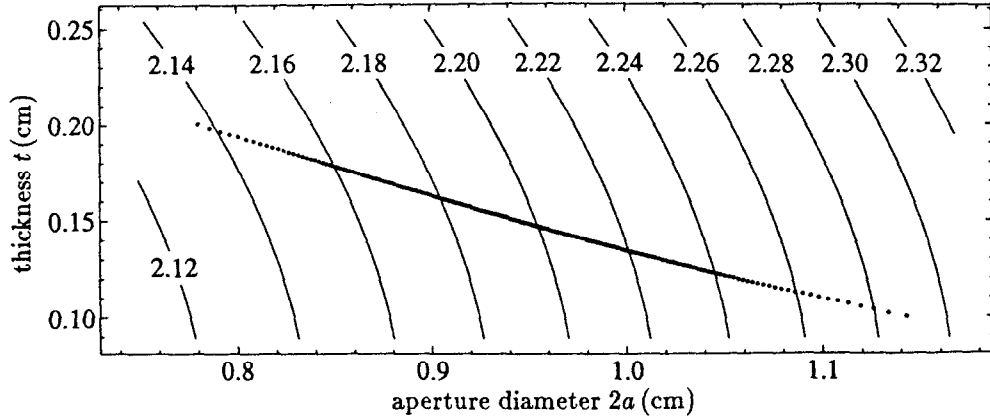


Figure 6.3: The cell diameter $2b$ (in cm) which yields synchronism with the beam ($v_\phi = c$). The dots are the 206 cells of the detuned accelerator structure. The rightmost dot is the first (input) cell and the leftmost dot is the last (output) cell.

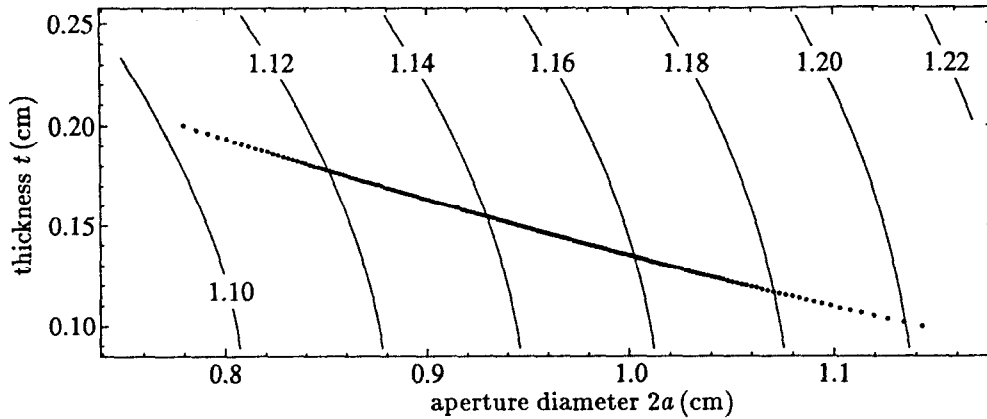


Figure 6.4: The logarithmic derivative $-b/f \partial f_o / \partial b$ with $2b$ determined from the synchronism condition. The dots are the 206 cells of the detuned accelerator structure.

would be unity. This simple model gives the logarithmic derivative within 20% of the correct answer.

The error in the estimate of $2b$ can be obtained from the estimated error of the frequency f_o using the derivative $\partial f_o / \partial b$. Since the residual frequency error of the polynomial fit is 80 kHz and only a moderate number of points were used in the fit, the maximum error of the fit should be less than 200 kHz. The discretization error may add a systematic frequency error less than 30 kHz. This indicates the error estimate for $2b$ as computed by the polynomial fit is $0.4 \mu\text{m}$. This is an order of magnitude smaller than machining tolerances. In fact, the skin depth for pure copper at 11.424 GHz is

$$\delta = \frac{c}{\sqrt{2\pi\mu\omega\sigma}} = 0.7 \mu\text{m} \quad (6.4)$$

which is larger than the estimated error for b . If the cell diameter $2b$ must be known to better than a micron then the finite conductivity of the walls must be taken into account.

6.2. Dipole Modes

In order to reduce the effect of wakefields it is useful to detune or spread out the frequencies of the undesired modes[28]. A gaussian distribution of the modes leads to very good cancellation of the wakefield effects. The wakefield decoheres in a time comparable to the reciprocal of the width of the frequency distribution of the modes.

The largest contribution to the wakefield comes from the lowest band of dipole modes. Higher dipole modes have less effect and can be effectively detuned by varying the disk thickness t along the structure such that the distribution of thicknesses is nearly gaussian[29]. The disk thickness t has little effect on the lowest dipole mode, so the aperture diameter $2a$ is varied along the structure such that the distribution of lowest synchronous dipole mode frequencies f_1 is nearly gaussian. While the synchronous dipole mode frequencies of the cells are not the same as the dipole mode frequencies of the structure, the distributions of the frequencies are similar according to equivalent coupled-circuit models of the structure[30].

A gaussian distribution with rms width $2\sigma = 0.7$ GHz provides decoherence of the wakefield effects by the time the following bunch arrives 1.4 ns later. To obtain good cancellation of wakefield effects over the whole bunch train, perhaps 90 bunches, it is necessary to have a good distribution of modes. Beam dynamics simulations[31] for the NLC indicate the tolerance for systematic relative frequency errors is approximately 10^{-4} , hence accurate calculations of f_1 are important for the design of detuned accelerator structures.

The lowest synchronous dipole mode frequency f_1 was calculated using YAP. The cell diameter $2b$ was fixed using the beam synchronism constraint (6.1), so two parameters remain for the dipole mode calculations: $2a$ and t . Frequencies at two phase advances ψ close to the synchronous phase and 0.02 radians apart were computed for 15 cells. Calculations on three successively refined meshes were extrapolated to zero mesh size, with conservative error estimates ranging from 120 kHz for large t cases to 400 kHz for small t cases. Then the lowest synchronous dipole mode frequency and phase advance were obtained using linear interpolation. An estimate of the group velocity of the lowest synchronous dipole mode was also obtained by computing the differences of the frequencies at the two phase advances.

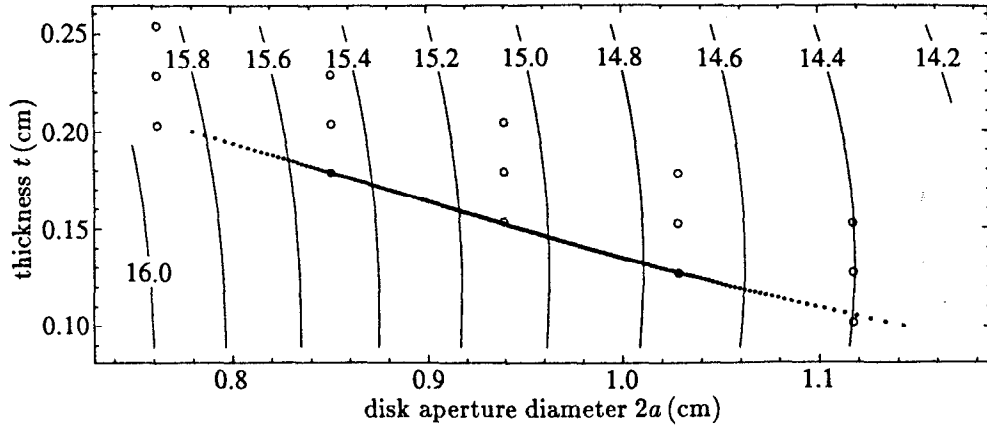


Figure 6.5: The lowest synchronous dipole mode frequency f_1 (in GHz). The cell diameter $2b$ is determined from the synchronism condition. The open circles are the dimensions at which f_1 was computed. The dots are the 206 cells of the detuned accelerator structure.

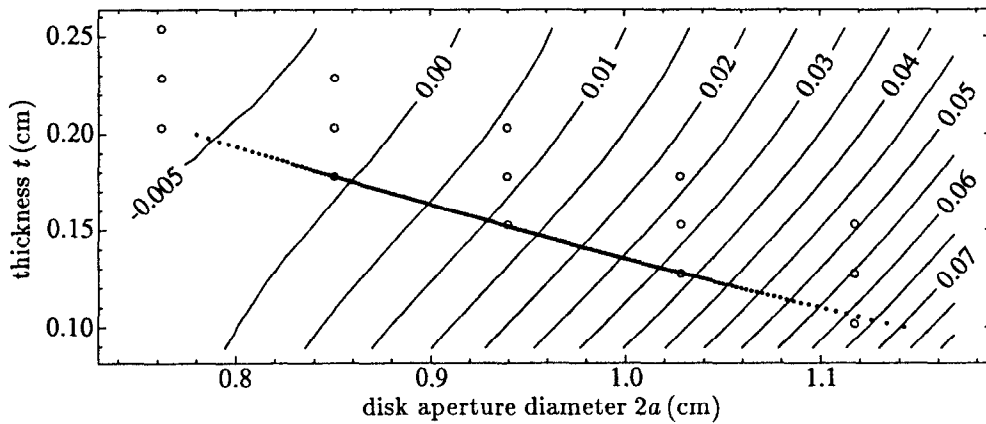


Figure 6.6: Group velocity v_g/c for the lowest synchronous dipole mode. The cell diameter $2b$ is determined from the synchronism condition. The open circles are the dimensions at which f_1 was computed. The dots are the 206 cells of the detuned accelerator structure.

The frequencies were fit to a polynomial quadratic in $2a$ and t . The estimated error of the fit is 600 kHz. The fit is shown in figure 6.5 and is expected to be good only near the region encompassed by the 15 calculated points. A similar fit for the group velocity is shown in figure 6.6. The calculated points are shown as open circles in the two figures.

The gaussian distributions for t and f_1 for 206 cells were truncated at $\pm 2\sigma$. The range of the disk thickness t was chosen to be 0.1 cm to 0.2 cm. This is sufficient to specify t for all cells. The relative range of f_1 was chosen to be 10.1%. With this information only one free parameter for the whole structure remains. In practice the parameter was the aperture diameter $2a$ of the first cell, but the average lowest synchronous dipole mode frequency is also a viable free parameter. In either case

the free parameter determines f_1 for each cell. Given f_1 and t for a cell the aperture diameter $2a$ was found using the quadratic polynomial approximation to the lowest synchronous dipole mode frequency. Then the cell diameter $2b$ was obtained from $2a$, t and the synchronism constraint (6.1). Note that cell-to-cell variation of the transverse kick factor, which represents the strength of the transverse wakefield, has not been included in this design procedure.

6.3. Structure RF Parameters

RF parameters like the shunt impedance, quality factor Q , group velocity v_g and peak surface field to accelerating gradient ratio $E_{s,\max}/E_{\text{acc}}$ were computed for the accelerating mode by SUPERFISH for a few of the cells. While the code calculates RF parameters for periodic structures, the calculations are still valid locally for the detuned structure since the cell dimensions vary slowly. A polynomial fit was used to obtain the RF parameters for the remaining cells. The filling time of the structure was computed from

$$T_f = \int_0^{1.8\text{m}} \frac{dz}{v_g(z)}. \quad (6.5)$$

The free parameter (e.g., the aperture diameter $2a$ of the first cell) of the structure was varied to achieve $T_f = 100$ nsec. The corresponding cell dimensions are plotted in the contour plots like figures 6.3 and 6.5.

Dispersion curves for the two lowest dipole modes of three cells (first, middle and last) of the structure were computed using YAP and are shown in figure 6.7. All modes up to 50 GHz are shown in figure 6.8.

The accelerating gradient along the length of the structure was computed[32] by first finding the attenuation $\tau(z)$ along the structure,

$$\tau(z) = \int_0^z \frac{\omega dz}{2Q(z)v_g(z)} \quad (6.6)$$

where $\omega = 2\pi f_o$ is the angular frequency. Then the power flow $P(z)$ in the structure is

$$P(z) = P_{\text{in}} e^{-2\tau(z)} \quad (6.7)$$

and the accelerating gradient is

$$E_{\text{acc}}(z) = \sqrt{r(z) \frac{dP}{dz}(z)} \quad (6.8)$$

where $r(z)$ is the shunt impedance per unit length. The gradient is shown in figure 6.9 along with gradients for a conventional constant gradient structure and a

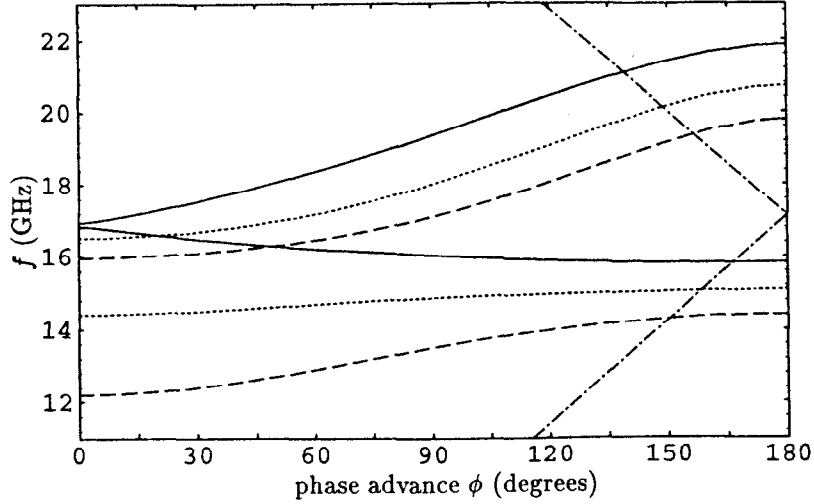


Figure 6.7: Dispersion diagram for the two lowest dipole modes of three different cells of the structure. The dashed line is the first (input) cell, the dotted line is the middle cell and the solid line is the last (output) cell. The dot-dash line is the velocity of light line.

constant impedance structure. All three structures in figure 6.9 have the same attenuation constant τ .

It is noteworthy that the peak surface field $E_{s,max}$ on the disk edges varies little along the structure as shown in figure 6.10. The first cell has a thin disk and large aperture, so $E_{s,max}/E_{acc}$ is high ($\cong 3$) but E_{acc} is small due to the low shunt impedance and high group velocity. For the last cell the situation is just the opposite. Other RF parameters are listed in table 6.1.

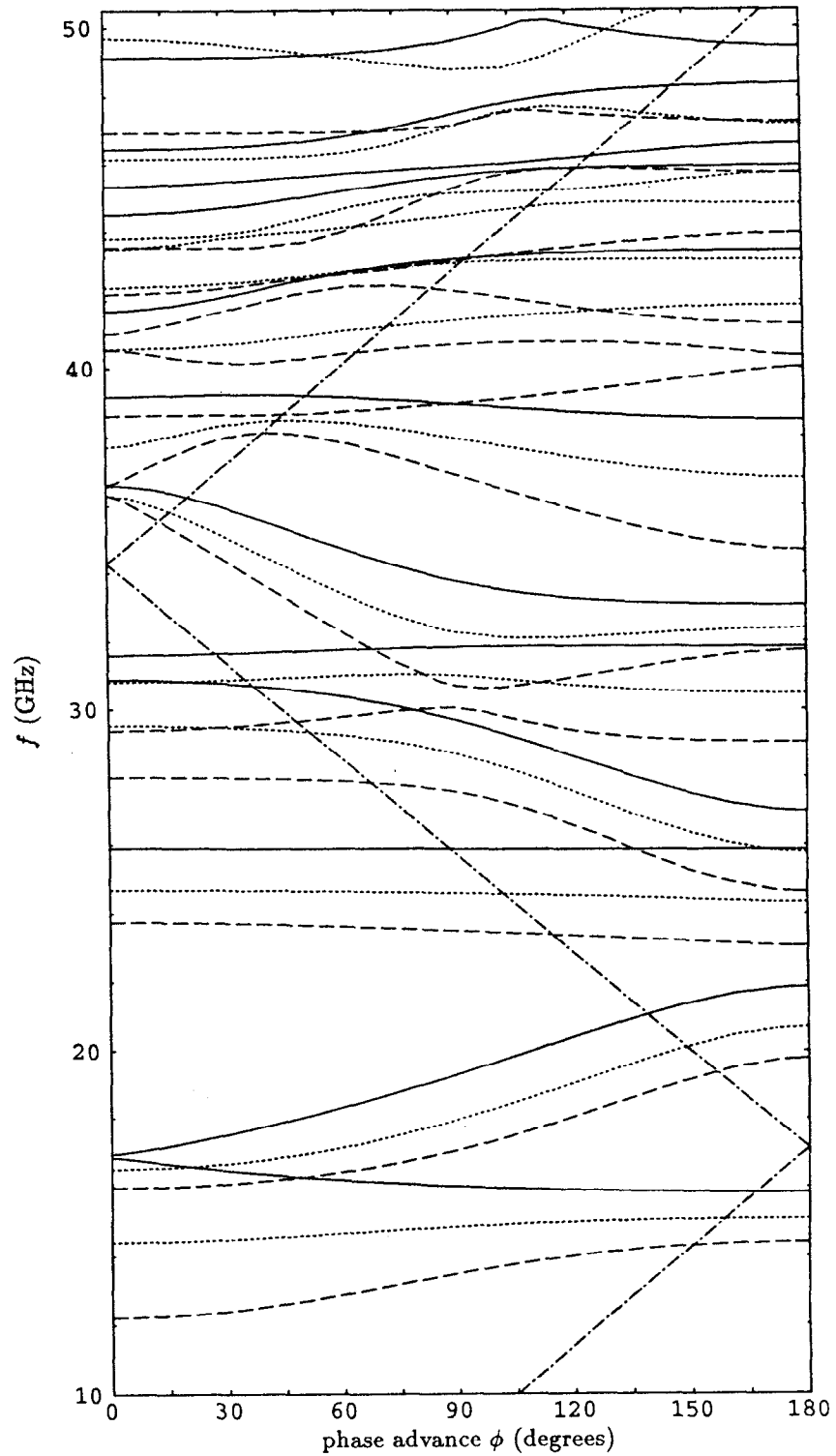


Figure 6.8: Dispersion diagram for the dipole modes of three different cells of the structure. The dashed line is the first (input) cell, the dotted line is the middle cell and the solid line is the last (output) cell. The dot-dash line is the velocity of light line.

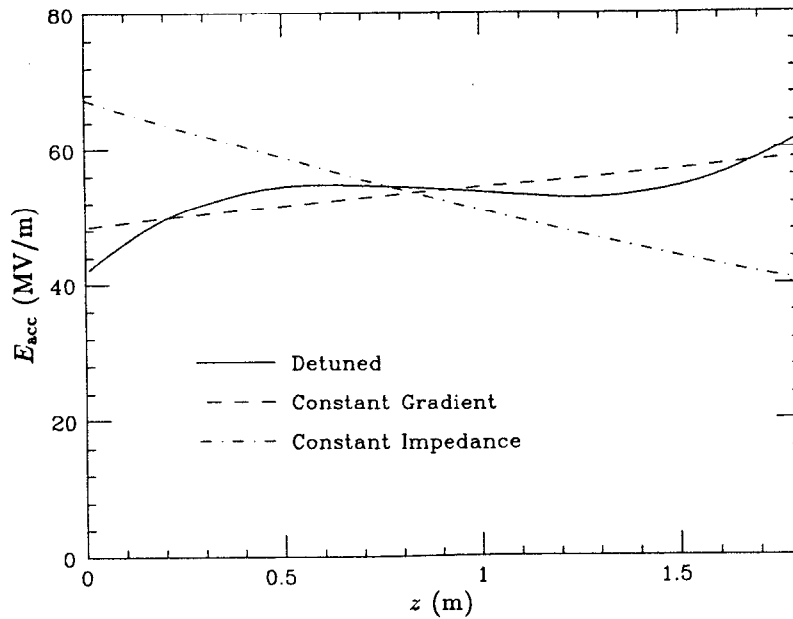


Figure 6.9: Accelerating gradient along the length of the structure for 100 MW input power and various structure types.

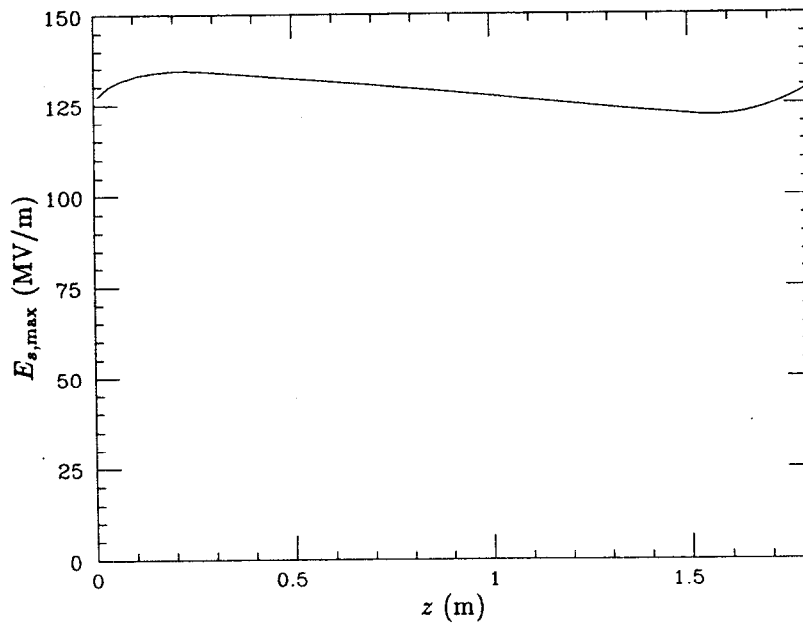


Figure 6.10: Peak surface gradient along the length of the detuned accelerator structure for 100 MW input power.

Section length	1.8 m
Phase advance per cell	$2\pi/3$ radians
Iris aperture:	
radius	5.72–3.91 mm
normalized radius	0.218λ – 0.149λ
Group velocity	$0.12c$ – $0.03c$
Filling time	100 ns
Unloaded time constant	207–186 ns
Attenuation constant	0.517 nepers
Shunt impedance	67.5–88.0 M Ω /m
Elastance	853–946 V/pC/m
For 50 MV/m average gradient:	
Peak input power/(1.8 m)	48.1 MW/m
Peak power per feed	86.5 MW
Average power dissipation for 250 ns pulses, 180 pps	1.4 kW/m

Table 6.1: Structure RF parameters.

Chapter 7

90° Overmoded Circular Waveguide Bend

Some designs for the Next Linear Collider (NLC) transmit power from the source (a klystron or the output of a pulse compressor) to the accelerator structure in the TE₀₁ mode of overmoded circular waveguide in order to have small transmission loss. The waveguide run from the source to the accelerator includes some 90° bends. Ideally these bends would be loss-less.

In collaboration[33] with Norman Kroll and Chris Nantista, a design algorithm and some results for one type of overmoded waveguide bend were found. A curved section of waveguide connects two straight sections as shown in figure 7.1a. The curvature in the bend is constant so the waveguide follows a 90° arc with radius of curvature ρ_c between the two straight sections. The cross-section of the waveguide is uniform throughout the curved section, but the cross-section is not simply a circle. The cross-section and radius of curvature ρ_c will be chosen so that the incoming wave propagates through the curved section with negligible mode conversion. This is the principal form of loss considered here. Reflection and wall losses are only considered heuristically. The straight sections are adiabatic tapers from and to circular waveguide.

7.1. Modes in Curved Guide

A curved guide can be treated as a portion of a cylindrically symmetric structure. For the 90° bend the structure starts at $\phi = 0$ and ends at $\phi = \pi/2$. The fields in the waveguide can be decomposed into modes with azimuthal dependence $e^{im\phi}$. In the axisymmetric waveguide paradigm the waves propagate along ϕ with propagation constant m . Compare this with the phase $e^{i\beta z}$ for waves propagating along z with propagation constant β in straight waveguides. The curved guide does not close on itself so there is no requirement that m be an integer.

The finite element field solver YAP is capable of computing the frequencies of the modes of axisymmetric structures for any real m . Non-integral m is allowed. YAP was used to compute dispersion diagrams for curved guide with various cross-sections. One such dispersion diagram is shown in figure 7.2. A dispersion diagram for curved guide looks similar to dispersion diagrams for straight guide. However, the simple dispersion formula $\omega^2/c^2 = k_c^2 + \beta^2$ for a straight waveguide containing no media does not apply to curved guide. This can be seen best in figure 7.2, where

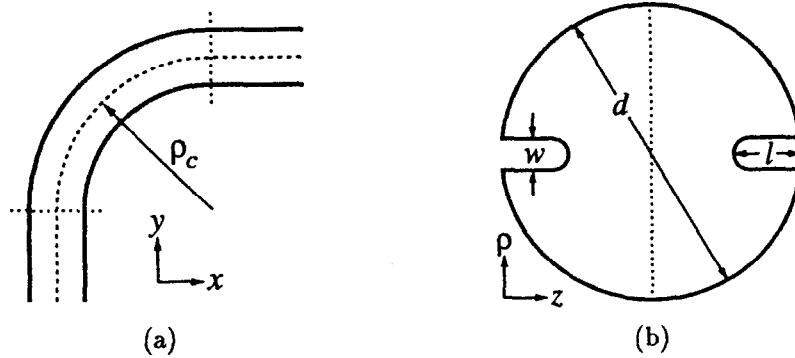


Figure 7.1: Outer geometry (a) and cross-section (b) of the bend. The cross-section's dashed line is a symmetry plane.

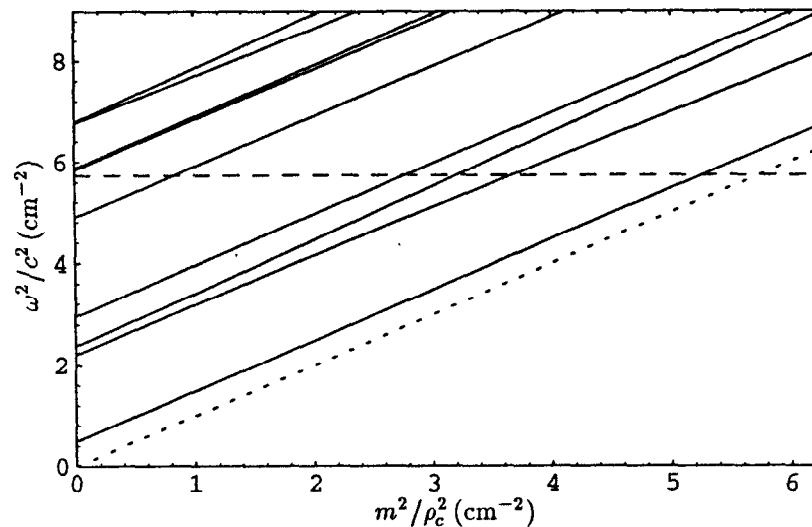


Figure 7.2: Dispersion diagram of the curved guide for the first design listed in table 7.1. The dashed line is the drive frequency 11.424 GHz. The dotted line corresponds to the speed of light along the center of the guide.

the dispersion curves are not parallel lines. A power series of the form

$$\frac{\omega^2}{c^2} = k_c^2 + \alpha_1 \left(\frac{m}{\rho_c}\right)^2 + \alpha_2 \left(\frac{m}{\rho_c}\right)^4 + \dots \quad (7.1)$$

approximates the dispersion curves well. The cutoff k_c^2 and the coefficients α_i depend on ρ_c and on the cross-section Ω of the guide. When ρ_c is large then $\alpha_1 \cong 1$ and the cutoffs k_c^2 are approximately the same between straight and curved guide with the same cross-section.

7.2. Scattering at the Interface

There is potentially some reflection at the interface between the straight waveguide and the curved waveguide. A generalized scattering matrix S_i for the propa-

gating modes in the straight and curved guides can be constructed.

As an example, the scattering matrix for the straight-to-curved interface in an overmoded rectangular H-plane waveguide bend was computed using a mode-matching method. Only TE_{n0} modes were considered so the fields are uniform vertically. In the straight guide propagating along y the modes are $E_z \propto \sin(2\pi nx/w)$ where $0 \leq x \leq w$ is the horizontal domain of the waveguide. In the curved guide the modes involve Bessel functions. They are $E_z \propto AJ_m(k\rho) + BY_m(k\rho)$ where $\rho_c - w/2 \leq \rho \leq \rho_c + w/2$ and $k = \omega/c$ is the drive frequency. Note that m is real for propagating modes and imaginary for evanescent modes.

The boundary conditions $E_z = 0$ at $\rho = \rho_c \pm w/2$ yield a characteristic equation for the propagation constants m . Solutions were obtained by numerically integrating Bessel's equations and using a shooting method to match the boundary conditions. This yielded numerical values for m^2 for both propagating ($m^2 > 0$) and evanescent ($m^2 < 0$) modes. The field E_z for each mode was obtained similarly.

The normalized generalized scattering matrix S_i was computed for an example with $w/\lambda = 1.36$ and $\rho_c/\lambda = 3.87$, where λ is the free space wavelength. There are two propagating modes in the guides. Using 14 modes for the field expansion on each side of the interface, the computed scattering matrix for the interface is

$$S_i = \begin{bmatrix} 4 \cdot 10^{-5} \angle -8^\circ & 8 \cdot 10^{-4} \angle -2^\circ & 0.982 & 0.190 \\ 8 \cdot 10^{-4} \angle -2^\circ & 8 \cdot 10^{-4} \angle -122^\circ & -0.190 & 0.982 \\ 0.982 & -0.190 & 3 \cdot 10^{-4} \angle -4^\circ & 8 \cdot 10^{-4} \angle -4^\circ \\ 0.190 & 0.982 & 8 \cdot 10^{-4} \angle -4^\circ & 6 \cdot 10^{-4} \angle -82^\circ \end{bmatrix} \quad (7.2)$$

where $[a_{s1}, a_{s2}, a_{c1}, a_{c2}]^T$ is the incoming wave vector. The wave amplitudes a_{sn} and a_{cn} are for the modes in the straight and curved guides, respectively.

Notice that the reflection amplitude is less than 10^{-3} . If one assumes the reflections are similar for bends with different cross-sections but similar curvature, then reflection at the straight-to-curved interface can be neglected. The reflected power will be negligible as long as resonances are avoided. The principal concern, then, is mode conversion.

7.3. Around the Bend

The scattering matrix S_b for a bend over angle ϕ_b can be easily computed given S_i for the straight-to-curved interface and the propagation constants m_1 and m_2 for the two propagating modes in the curved guide. The example above has $m_1 = 22.85$ and $m_2 = 16.18$. The next mode is evanescent with $m_3 = i11.38$. The transmission coefficient $s_{2f,1f}$ for the (straight guide) fundamental mode for

various bend angles ϕ_b was computed. This is shown in figure 7.3. At $\phi_b = 2\pi/(m_1 - m_2) = 0.941$ the transmission is nearly perfect. At this bend angle the two propagating waves in the curved guide arrive at the output end of the bend with the same relative phases they had at the input end of the bend. The propagating field at the output is the same as at the input except for an overall phase, so waves are faithfully transmitted through the bend with no mode conversion.

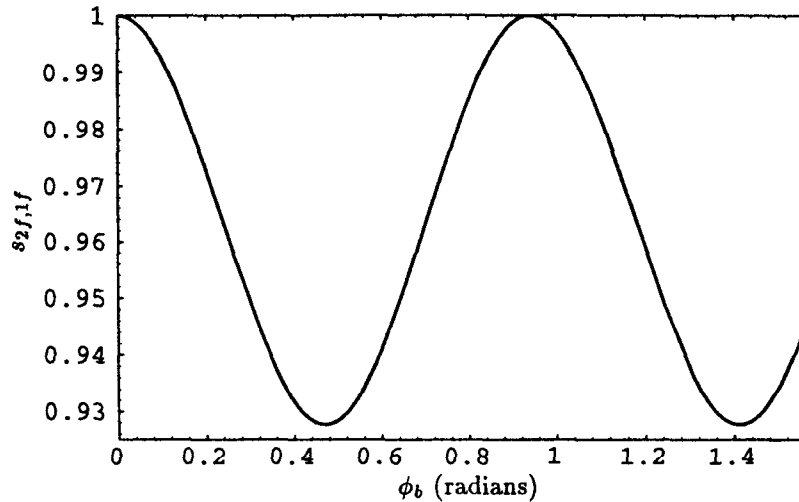


Figure 7.3: Transmission of the fundamental mode through a rectangular waveguide bend over angle ϕ_b . Loss is due to conversion into higher modes.

The evanescent waves at the interfaces have decayed sufficiently in the curved guide so that they can be neglected in the transmission calculations for $\phi_b = 0.941$.

This example leads to the principal design criterion for this type of overmoded waveguide bend: the phases $e^{m_i\phi_b}$ must be identical for all modes propagating in the curved guide. In addition, evanescent modes should be sufficiently above cutoff so that they decay well over the length of the bend, and thus can be neglected.

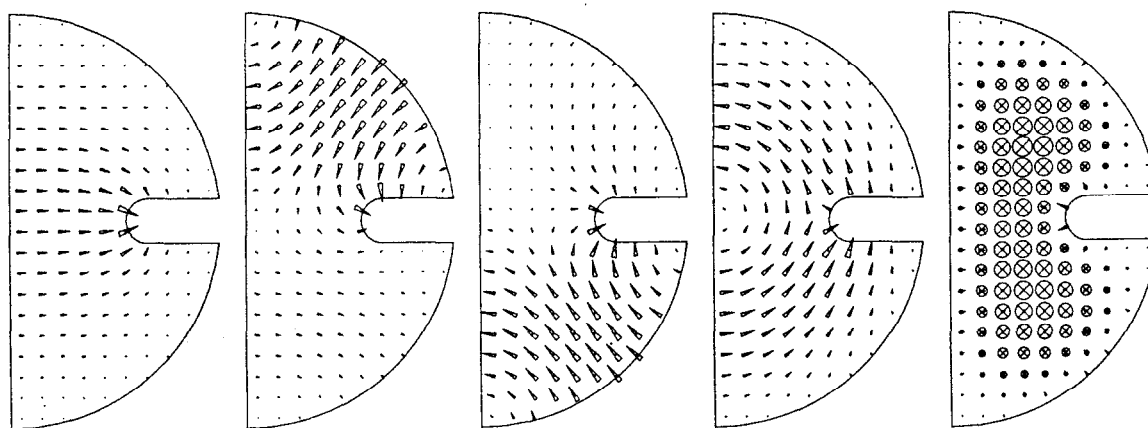
7.4. 90° Bend Design

Designs for a 90° bend with a cross-section as shown in figure 7.1b were computed. The phases $e^{m_i\pi/2}$ for the five lowest propagating modes excited by the incoming wave can be fixed relative to each other by adjusting the four parameters: d , ρ_c , l and w . Propagating modes not excited by the incoming wave (due to symmetry) are neglected. Dispersion diagrams were computed using YAP and the bend parameters were adjusted so that the phases were the same. This corresponds to the propagation constants m_i differing from one another by multiples of 4. The cutoff ($m = 0$) frequencies of higher order modes were computed in order to discard

90° Overmoded Waveguide Bends

d (cm)	l (cm)	w (cm)	ρ_c (cm)	m_1	m_2	m_3	m_4	m_5	f_{c6} (GHz)
4.372	0.986	0.465	31.786	72.873	60.873	56.873	52.873	28.874	11.536
4.275	0.971	0.611	36.655	83.867	67.867	63.867	59.867	23.868	11.819
4.358	1.054	0.593	38.754	89.034	73.034	69.034	65.034	25.033	11.579
3.940	0.765	0.476	23.891	53.870	41.870	37.870	33.870	9.871	12.726
4.157	0.904	0.622	33.894	77.212	61.212	57.212	53.212	17.213	12.163

Table 7.1: Dimensions and propagation constants for various bend designs. The cutoff frequency for the sixth mode is also listed.



Mode 1 (“TE₁₁”) Mode 2 (“TE₂₁”) Mode 3 (“TE₃₁”) Mode 4 (“TE₀₁”) Mode 5 (“TM₁₁”)

Figure 7.4: Electric field patterns for the five propagating modes of the first design in table 7.1.

designs with more than five propagating modes at 11.424 GHz. Table 7.1 lists the parameters for five solutions. It also lists the propagation constants for the five lowest modes and the cutoff frequency f_{c6} for the sixth lowest mode.

The cross-section in figure 7.1 and the dispersion diagrams in figure 7.2 correspond to the first design in table 7.1. The field patterns for the propagating modes are shown in figure 7.4. At cutoff the field patterns for the modes in curved guide are similar to the corresponding modes in straight guide, but for large m the second and third modes are mixed. This is evident in the field plots and in the dispersion diagram, where it appears that the second and third curves are repelling each other. These modes arise, with the introduction of the septa, from the TE₂₁ and TE₃₁ modes of circular guide. The incoming wave is similar to the fourth mode, which is a TE₀₁-like mode.

The cutoff frequency for the sixth mode of the first design appears close to cutoff. The estimated propagation constant using the straight guide formula is $m_6 \cong i10.7$ and the decay amplitude over the length of the waveguide is $e^{im_6\pi/2} = 5 \times 10^{-8}$.

This decay is satisfactory.

7.5. Further Work

Further designs can be found, perhaps with smaller radii of curvature and shorter septa so that the bend will have smaller wall losses and be easier to manufacture.

A variation of the YAP field solver will compute the evanescent modes in curved guide. With these modes a mode-matching algorithm can be employed to calculate the scattering matrix S_i for the straight-to-curved guide interface, and then verify that reflections are negligible and that the design criterion is appropriate.

Calculation of the wall losses through the bend and mode-conversion losses (due to manufacturing errors) also requires knowledge of S_i in order to obtain the mode amplitudes in the bend as well as the evanescent fields near the interface.

Chapter 8

Conclusion

The finite element method has been successfully applied to the electromagnetic eigenmode problem in axisymmetric structures. Very accurate frequencies are now available for the multipole modes of axisymmetric structures.

Similar finite element methods can and should be applied to three-dimensional problems. Also, these finite elements should be applied to other electromagnetic problems, such as the driven (inhomogenous) cavity problems, time domain simulations, and scattering matrix problems.

Appendix A

Implementation of YAP

The finite element formulations described in this thesis have been implemented in a set of codes collectively called YAP. There is a set of codes on the SLAC IBM/VM mainframe which implement the planar and $m = 0$ formulations. There is a separate set of codes on a NeXTstation which implement the $m \neq 0$ formulation. The two sets of codes have different input formats which are not compatible with each other.

A note of caution to the potential YAP user: while this implementation is excellent for delivering accurate mode frequencies, it is still far from an ideal field solver. There are few post-processing capabilities, generating input files is tedious, and the algebraic eigenvalue solver is not optimal. The user of this implementation has my sympathy, because it is not user friendly. It is not easy. You have been warned! It is my sincere hope that this implementation is not final, and that more convenient implementations will arise.

A.1. Planar and Monopole Formulations

The two programs which constitute YAP on the IBM/VM mainframe are written in C. The first program is a mesh generator. The second program is the field solver. The relationship between these two programs and selected files and input parameters is shown in figure A.1.

A.1.1. Mesh Generator and Boundary Model Input Format

The mesh generator is called **mg1**. It constructs a mesh composed of quadratic triangular elements given a boundary model and a desired element size. The actual size of the elements in the output mesh is, of course, only approximately the desired element size.

The boundary model specifies the geometry and physical dimensions of the problem. The boundary model specifies if the structure is planar or axisymmetric. It does not specify anything concerning electromagnetic fields.

In the boundary model, the interior of the structure is covered by one or more *regions*. In principle, different physical properties (e.g., dielectric constant) could be associated with different regions. However, at this time YAP only employs regions for computational reasons: different element sizes can be specified for different regions.

Each region is bounded by one or more closed curves called *loops*. A region has one outer loop and zero or more inner loops. Inner loops are used to describe

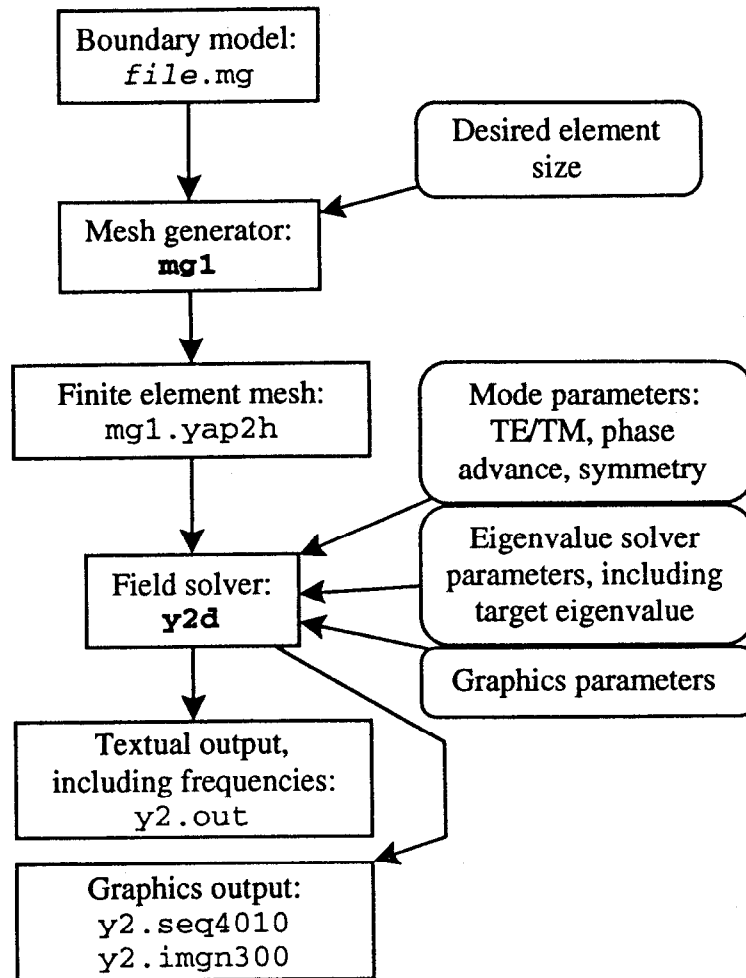


Figure A.1: Relationship between programs, files and input parameters for the YAP implementation of the planar and $m = 0$ formulations. Programs and files are on the left, input parameters are on the right.

holes in a region. A loop is an ordered list of curve segments called *halfedges*. Each halfedge represents the portion of the loop between two points called *nodes*. The halfedges of an outer loop are listed in counterclockwise order, while the halfedges of an inner loop are listed in clockwise order.

Curve segments are also called *edges*. In this boundary model an edge is an object with two sides. If the edge is a boundary between two regions, then the edge is composed of two halfedges, one from each of the two adjacent regions. Other edges are composed of a single halfedge (from the adjacent region) and a boundary specification. The boundary specification may be a metal boundary, an axis boundary, a symmetry boundary or a periodic boundary. A curve segment between two nodes may be either a straight line or an arc of a circle.

```

boundModel :
    | boundModel nodeList
    | boundModel heList
    | boundModel edgeList
    | boundModel loopList
    | boundModel regionList
    | boundModel defDimStat
    | boundModel spaceDesc
nodeList  : NODES
    | nodeList INTEGER length length
    | nodeList defDimStat
heList    : HALFEDGES
    | heList INTEGER INTEGER INTEGER
edgeList  : EDGES
    | edgeList INTEGER boundary boundary
    | edgeList INTEGER boundary boundary curvParam
    | edgeList defDimStat
bndtype   : METAL
    | SYMMETRY
    | PERIODIC
    | AXIS
boundary  : INTEGER
    | bndtype
    | bndtype ( INTEGER )
    | bndtype ( INTEGER INTEGER )
curvParam : CIRCLE ( direct length length )
direct    : DIRECTCW
    | DIRECTCCW
loopList  : LOOPS
    | loopList INTEGER INTEGER INTEGER
regionList : REGIONS
    | regionList INTEGER INTEGER
spaceDesc : XY | PLANAR
    | RZ | AXISYMMETRIC | CYLINDRICAL
defDimStat : DIMENSION dimension
dimension  : MILS
    | INCHES
    | MILLIMETERS | MM
    | CENTIMETERS | CM
    | METERS
length    : real
    | real dimension
real      : REAL
    | INTEGER

```

Figure A.2: Boundary model grammar in a format similar to the format for the YACC parser generator. Uppercase words are terminal symbols. Except for INTEGER and REAL, the terminal symbols are literal.

The first line of the boundary model file is always a title line. The remaining lines follow the grammar listed in figure A.2. Selected parts of the grammar are described in detail below. An example of a boundary model is shown in figure A.3.

Comments are allowed in the boundary model. Comments start with “/*” and end with “*/”, and they may span multiple lines.

The first token in figure A.2 represents the entire boundary model:

```
boundModel :
    | boundModel nodeList
    | boundModel heList
    | boundModel edgeList
    | boundModel loopList
    | boundModel regionList
    | boundModel defDimStat
    | boundModel spaceDesc
```

The boundary model is composed of node lists, halfedge lists, edge lists, loop lists and region lists. It may also include statements specifying the default units for node coordinates and the geometry: planar or axisymmetric. There is no fixed order for these lists and statements. There must be at least one of each type of list. There may be more than one of a type of list.

A region list is:

```
regionList : REGIONS
    | regionList INTEGER INTEGER
```

The keyword REGIONS is followed by pairs of integers: one pair for each region in the list. The first integer is a unique region index assigned by the user. The second integer is the index of the outer loop of the region. The inner loops of the region are found by following a linked list of loops. The example in figure A.3 has two regions, labelled 1 and 2.

A loop list is:

```
loopList : LOOPS
    | loopList INTEGER INTEGER INTEGER
```

The keyword LOOPS is followed by triplets of integers: one triplet per loop in the list. The first integer is a unique loop index assigned by the user. The second integer is any index of a halfedge in the loop. Other halfedges in the loop can be found by following the linked list of halfedges.

The third integer of a loop specification is the index of the next loop of the region. Recall that a region with one or more holes requires at least two loops: an outer loop and one or more inner loops. The purpose of this index is to form linked lists of loops: one for each region. The head of a linked list is the outer loop of a region. The remaining loops of a linked list are inner loops. The order of the inner loops is not important. The last loop in the linked list points to the outer loop.

If a region has only one loop then the next loop of the region is the same as the first loop of the region. That is, the first and third integers of the triplet will be identical. This is what occurs in the two regions of the example in figure A.3. Each of the two regions in the example has only one loop.

A halfedge list is:

```
heList      : HALFEDGES
             | heList INTEGER INTEGER INTEGER
```

The keyword `HALFEDGES` is followed by triplets of integers: one triplet per halfedge. The first integer is a unique halfedge index assigned by the user. The second integer is the index of the node from which the halfedge starts. Recall that halfedges represent curve segments between two nodes, and that loops, and hence halfedges, are traversed counterclockwise for outer loops and clockwise for inner loops. Only the starting node of the halfedge is specified. The ending node of a halfedge is the starting node of the next halfedge of the loop. The third integer is the index of the next halfedge in the loop. The purpose of the third index is to form linked lists of halfedges: one list per loop. The order of halfedges in the linked list is important. In the example in figure A.3, there are 10 halfedges: 6 in the first loop and 4 in the second loop.

A node list is:

```
nodeList    : NODES
             | nodeList INTEGER length length
             | nodeList defDimStat
```

The keyword `NODES` is followed by triplets: one triplet per node. The first component of the triplet is an integer. It is a unique node index assigned by the user. The last two components are the coordinates of the node: either x followed by y or z followed by ρ . Each component may include units (e.g., `INCHES`). If no unit is specified then the default unit is assumed. This default unit may be changed in the middle of the list, between triplets. The example in figure A.3 has 8 nodes.

An edge list is:

```
edgeList    : EDGES
             | edgeList INTEGER boundary boundary
             | edgeList INTEGER boundary boundary curvParam
             | edgeList defDimStat
```

The keyword `EDGES` is followed by edges. Each edge starts with an unique integer index assigned by the user, followed by two side specifications: one for each side of the edge. The first side specification is the left side of the edge, the second side specification is the right side of the edge. Left and right are only important when the edge has an optional curve or if the edge is a periodic boundary. The side specification may be the index of a halfedge. It may also be one of the boundary

conditions described below. There are 9 edges in the example in figure A.3.

Each edge optionally has a curve specification. If there is no curve specification for the edge, then the edge is a straight line between the start and end nodes. The only alternative curve specification available at this time is an arc of a circle between the start and end nodes. A curve specification has the form

```
    curvParam  : CIRCLE ( direct length length )
    direct     : CW
               | CCW
```

That is, it is the keyword **CIRCLE** followed by a triplet in parenthesis. The last two components of the triplet are the center of the circle. The first component is one of the keywords **CW** or **CCW**, indicating the circle is to be traversed clockwise or counterclockwise between the start and end nodes of the edge. Four of the edges in the example in figure A.3 have curve specifications.

Note that the halfedges associated with an edge determine the start and end nodes of the edge. The start node of the edge is the start node of the left halfedge or the end node of the right halfedge.

A boundary specification is:

```
    bndtype    : METAL
               | SYMMETRY
               | PERIODIC
               | AXIS
    boundary   : INTEGER
               | bndtype
               | bndtype ( INTEGER )
               | bndtype ( INTEGER INTEGER )
```

There are four types of boundaries. A boundary specification starts with one of the keywords **METAL**, **SYMMETRY**, **PERIODIC** or **AXIS**.

Metal and symmetry boundaries optionally have a positive integer index. This index is enclosed in parenthesis. The index for metal boundaries has no practical effect in this implementation, but an electrostatic field solver would find distinctions between metal boundaries useful. The metal boundaries in the example in figure A.3 are divided into two groups: one for the cathode and one for the anode.

The index for symmetry boundaries is used to assign the boundary to distinct symmetry planes. Up to four different symmetry planes, labelled 1 through 4, are allowed. Symmetry plane 1 is the default. Note that the symmetry boundary does not specify the symmetry of the fields in the field solver. It only specifies that the structure itself is symmetric and that this boundary is on a symmetry plane.

Periodic boundaries are followed by two integers in parenthesis. The first integer is the index of the edge one period away from this edge. Note the edge one period

away must also refer to this edge in a similar manner. The second integer is a flag holding two bits of information. The default flag is 0. Up to two periodicities are allowed, hence the boundary model can represent a two-dimensional lattice. Add 1 to the flag if this periodic boundary belongs to the second periodicity instead of the first periodicity. The second bit of information indicates the direction of the periodicity's rigid motion. Of two edges one period away from each other, the first edge is mapped to the second by the rigid motion. The second edge is mapped to the first by the inverse of the rigid motion. Add 2 to the flag if this edge is the second edge.

The programs say the second edge is on the "positive" side of the periodic boundary. The corresponding first edge is on the "negative" side of the periodic boundary. In terms of the notation used in this thesis, the second edge is on the "right" periodic boundary, while the first edge is on the "left" periodic boundary.

Note that the halfedge associated with the edge determines the start and end nodes of the edge. The start node of the edge is the start node of the left halfedge or the end node of the right halfedge. The rigid motion maps the start node of the first edge to the start node of the second edge, and the end node of the first edge to the end node of the second edge. Careless ordering of the halfedge and the periodic boundary in the edge will lead to an undesired reflection in the rigid motion. The example in figure A.3 shows the common ordering for periodic boundaries. In edge 2 the periodic boundary is the right side of the edge, while in edge 9 the periodic boundary is the left side of the edge.

A default unit specification is:

```
defDimStat : DIMENSION dimension
dimension  : MILS
            | INCHES
            | MILLIMETERS | MM
            | CENTIMETERS | CM
            | METERS
```

Specifying the default unit affects the interpretation of the coordinates which appear later in the file. It does not affect coordinates appearing before the default unit specification. Note that the default unit is overridden for a coordinate if a unit is specified for that coordinate. Initially the default unit is centimeters. The mesh generator converts coordinates to centimeters and internally stores the coordinates in centimeters. The second line of example in figure A.3 sets the default unit to mils.


```

CFA hole+slot circuit w/exaggerated curvature
DIMENSION MILS
NODES /* (node number) (x coordinate) (y coordinate) */
1 -79.87 -4.00
2 79.87 -4.00
3 92.35 120.38
4 35.00 124.34
5 -35.00 124.34
6 -92.35 120.38
7 35.00 188.22
8 -35.00 188.22
HALFEDGES /* (halfedge number) (start node) (next halfedge of loop) */
1 1 2 /* Interaction region. */
2 2 3
3 3 4
4 4 5
5 5 6
6 6 1
7 5 8 /* Hole and slot. */
8 4 9
9 7 10
10 8 7
EDGES /* (edge number) (left halfedge/bound) (right halfedge/bound) */
1 1 METAL(1) CIRCLE(CW 0.0 -800.0)
2 2 PERIODIC(9 2)
3 3 METAL(2) CIRCLE(CCW 0.0 -800.0)
4 4 7
5 5 METAL(2) CIRCLE(CCW 0.0 -800.0)
6 10 METAL(2)
7 8 METAL(2)
8 9 METAL(2) CIRCLE(CCW 0.0 240.0)
9 PERIODIC(2 0) 6
LOOPS /* (loop number) (a halfedge) (next loop of region) */
1 1 1
2 9 2
REGIONS /* (region number) (outer loop) */
1 1 /* Interaction region. */
2 2 /* Hole and slot. */

```

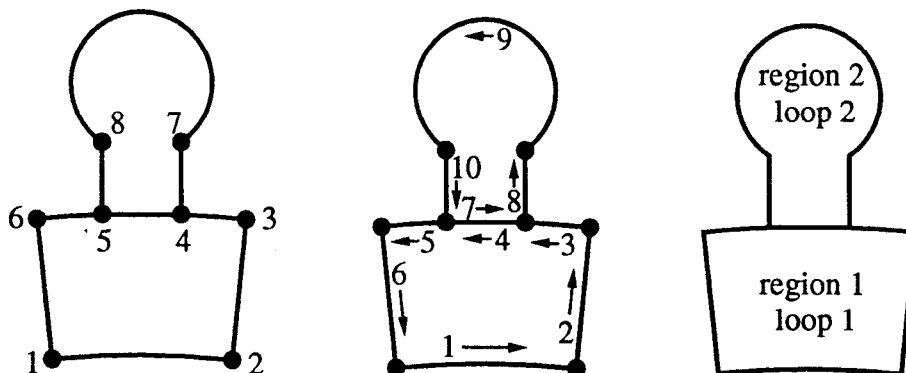


Figure A.3: An example of a boundary model. The three figures show the nodes (left), halfedges (middle), loops and regions (right). The halfedge arrows point to the next halfedge of the loop.

A geometry specification is:

```
spaceDesc : XY | PLANAR  
           | RZ | AXISYMMETRIC | CYLINDRICAL
```

The default geometry is planar. The only alternative is axisymmetric, which may be specified by including any of the three keywords RZ, AXISYMMETRIC or CYLINDRICAL in the boundary model file.

To run the mesh generator, type the command

```
mg1 boundary-model-file element-size
```

where *boundary-model-file* is the filename of the boundary model. The filetype is mg. The parameter *element-size* gives the desired element size in centimeters. If *element-size* is zero, then the mesh generator prompts for the element size along each edge and in each region of the boundary model. Hence the mesh may be made finer in some regions and coarser in others. The output mesh is mg1.yap2h. Additional output files provide diagnostic listings and some graphics which may be used to check the boundary model and the mesh.

In brief, the mesh generation algorithm first generates a mesh of linear triangular elements. Midside nodes are added to the elements to generate a quadratic triangular mesh. Care is taken with element sides on curved boundaries to ensure the midside node is on the curved boundary.

The linear triangular mesh generation algorithm[38] first approximates the edges of the boundary model by a series of nodes connected by line segments. Each line segment corresponds to a side of one element of the finite element mesh. The mesh is built region by region. For each region a list of nodes is constructed in the interior of the region. The spacing between the nodes is approximately the desired element size. The algorithm for constructing the interior node list is described below. The elements are constructed by starting with a line segment on the region boundary. The interior and boundary nodes are searched to find a good third corner for a linear triangular element. The other two corners of the triangle are the two boundary nodes at the ends of the line segment on the boundary. This element is removed from the region, and the process is repeated with another line segment on the boundary of the reduced region. This repeats until the whole region has been triangulated. There are, of course, many cases to handle. For example, removing a triangular element from the region occasionally splits the region into two subregions. Then the triangulation of one subregion continues. When this subregion is done then triangulation starts on the second subregion.

The interior nodes of a region are generated on horizontal (constant y or ρ) lines

```

title
nodes elements qp-nodes-1 qp-nodes-2 corner-elements flags
i1  x1  y1  f1  }
i2  x2  y2  f2  } nodes
⋮    ⋮    ⋮    ⋮    }

i1  n1,1  n2,1  n3,1  n4,1  n5,1  n6,1  s1,1  s2,1  s3,1  }
i2  n1,2  n2,2  n3,2  n4,2  n5,2  n6,2  s1,2  s2,2  s3,2  } elements
⋮    ⋮    ⋮    ⋮    ⋮    ⋮    ⋮    ⋮    ⋮    ⋮    }

nR1  nL1  }
nR2  nL2  } qp-1 nodes
⋮    ⋮    }

nR1  nL1  }
nR2  nL2  } qp-2 nodes
⋮    ⋮    }

e1  β1  }
e2  β2  } sharp corner elements
⋮    ⋮    }

```

Figure A.4: Format of the finite element mesh file `mg1.yap2h`.

between the top and bottom limits of the region. The spacing between these lines is approximately the desired element size. The intersections between each horizontal line and the boundary (as approximated by line segments) is found. A list of horizontal line segments is built. The list represents the portion of the horizontal lines in the interior of the region. For each horizontal line segment, interior nodes are placed equidistantly between the left and right ends of the segment. The horizontal spacing of the nodes is approximately the desired element size.

A.1.2. Finite Element Mesh File Format

The format of the file `mg1.yap2h` is shown in figure A.4. The first line of the file is the title line. The following line lists the number of nodes, the number of elements, the number of nodes on the right side of the first periodic boundary, the number of nodes on the right side of the second periodic boundary, the number of elements touching a sharp corner and a flag. The flag is 1 if the structure is axisymmetric.

The list of nodes follows. For each node, a node index i , the two coordinates x and y , and a node flag f are listed. The node coordinates are in centimeters. The flag is hexadecimal and indicates which boundary, if any, the node lies on. The flags are listed in table A.1. Note that flags may lie on more than one boundary.

flag	boundary
01	metal boundary
02	right side of first periodic boundary
04	right side of second periodic boundary
08	axis boundary
10	first symmetry boundary
20	second symmetry boundary
40	third symmetry boundary
80	fourth symmetry boundary

Table A.1: Node boundary flags in mg1.yap2h.

flag	boundary
-1	metal boundary
-2	right side of first periodic boundary
-3	right side of second periodic boundary
-4	axis boundary
-5	first symmetry boundary
-6	second symmetry boundary
-7	third symmetry boundary
-8	fourth symmetry boundary
-9	left side of first periodic boundary
-10	left side of second periodic boundary

Table A.2: Side boundary flags in mg1.yap2h.

Following the list of nodes is the list of elements. For each element an element index i , six node indices n_1 through n_6 , and three boundary/side indices s_1 through s_3 are listed. If an element side is adjacent to another element, then the boundary/side index for the side is the (positive) index of the adjacent element. If a side is on the boundary then the boundary/side index is a negative number indicating the appropriate boundary condition. These negative numbers are listed in table A.2.

Following the list of elements is a list of pairs of node indices for the two periodic boundaries. The first node index n_R is a node on the right periodic boundary and the second node index n_L is the corresponding node on the left periodic boundary. The nodes on the first periodic boundary are listed first, followed by the nodes on the second periodic boundary. These lists will be empty if the structure is not periodic.

Following the periodic boundary node lists is a list of elements touching sharp corners. For each element touching a sharp corner the element index e and the angle β of the associated sharp corner is listed. This list will be missing if there are

```

e=<target>    target eigenvalue (cm-2) (default is 0).
i=<fn>        specifies input filename.
k=<plotflag>  plot flag (hexadecimal).
    0x8000    graphics output for IMAGEN (default is Tek 4010).
    0x0008    time snapshots of complex field.
    0x0004    real field.
    0x0002    elements of mesh.
    0x0001    boundary of mesh.
l=<listflag>  listing flag (hexadecimal).
    0x8000    mesh size parameters.
    0x0020    execution times of field solver.
    0x0010    nearest eigenvalue estimate.
    0x0008    high eigenvalue estimate.
    0x0004    sparse matrices (with encoded b.c.'s).
    0x0002    equation numbers <-> node index.
    0x0001    number of equations.
m=<modeflag>  mode/boundary flag (hexadecimal).
    0x8000    force use of complex field solver.
    0x4000    suppress use of corner elements.
    0x2000    suppress all field-solving.
    0x1000    use conjugate gradient matrix solver (default ma28).
    0x0010    symmetry plane 4 is dirichlet (0).
    0x0008    symmetry plane 3 is dirichlet (0).
    0x0004    symmetry plane 2 is dirichlet (0).
    0x0002    symmetry plane 1 is dirichlet (0).
    0x0001    solve for TE mode (TM is default).
p=<phase1>    phase advance of 1st quasi-periodic boundary (degrees).
q=<phase2>    phase advance of 2nd quasi-periodic boundary (degrees).
t=<nitol>     numerical integration tolerance for corners (default 1e-4).
w=<phases>    list of phases "omega tee" for time snapshots of complex
              fields. The list is comma delimited with no spaces between
              the phases, and may contain up to 8 phases (degrees).

```

Figure A.5: Command line options for field solver `y2d`, from the online help file.

no sharp corners in the structure.

A.1.3. Field Solver

The finite element field solver is called `y2d`. The field solver reads a mesh file like `mg1.yap2h` and obtains all other parameters from the command line. The parameters are listed in figure A.5. The command line

```
y2d i=mg1 l=8039 k=8 m=4000 p=120.0 w=0,45,90,135
```

is an example invoking the field solver. Note that the flags for the desired options are added together. The listing flag `l=8039` indicates the user wants the output listing to include mesh size parameters, execution times, nearest eigenvalue estimate, high eigenvalue estimate and the number of equations. The flag indicates the user does not want, for example, a listing of the sparse matrices.

Given the desired mode type (TE vs TM) and boundary conditions, the field

solver loops over the nodes of the mesh, assigning an index to each node whose global basis function is appropriate for the problem. For example, a node on a metal boundary of a TM mode problem would not be assigned a global basis function index. Special indices are assigned to nodes without their own global basis function. Note that some of these nodes involve quasi-periodic boundary conditions, and the special indices reflect this.

The number of global basis functions M is known after this process. The coefficients of the global basis functions in the solution will be complex if complex fields are needed.

The global matrices \mathbf{A} and \mathbf{B} of order M are initialized. These matrices are stored as symmetric sparse matrices. The components of the stored matrices are real. A flag stored with each component of the matrices indicates the quasi-periodic phase factor associated with the component, if any. This allows the phase advances to be changed easily when computing dispersion diagrams of periodic structures: it is not necessary to compute the matrices \mathbf{A} and \mathbf{B} from scratch for each phase advance.

The field solver loops over the elements of the problem, computing the two element matrices \mathbf{A}^e and \mathbf{B}^e , and accumulating the dense element matrices to the appropriate components of the sparse global matrices.

For a given phase advance, the field solver finds an eigenvector \mathbf{x} and eigenvalue λ of the generalized algebraic eigenvalue problem $\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}$ using a shifted inverse power method. This is an iterative method. Starting from a trial vector \mathbf{x}_i , the eigenvalue solver computes \mathbf{x}_{i+1} by solving

$$(\mathbf{A} - \bar{\lambda}\mathbf{B})\mathbf{x}_{i+1} = \mathbf{B}\mathbf{x}_i, \quad (\text{A.1})$$

where $\bar{\lambda}$ is the target eigenvalue. The vectors \mathbf{x}_i converge to the eigenvector \mathbf{x} with eigenvalue λ closest to the target eigenvalue $\bar{\lambda}$.

Two methods are available to solve the matrix equation at each iteration step. One method uses the unsymmetric sparse matrix solver `ma28`, which is a direct method producing a sparse LU factorization for the sparse matrix $\mathbf{A} - \bar{\lambda}\mathbf{B}$. The other method uses a conjugate gradient matrix solver, which is an iterative method. Both methods work well. The conjugate gradient matrix solver requires less memory space, however, so it is used when the matrices become very large.

A.2. Multipole Mode Formulation

There are a number of programs which constitute YAP on the NeXTstation. Some of the programs are written in Objective-C. Others are written in C++.

A.2.1. Mesh Format and Refinement

There is no mesh generation program in YAP on the NeXTstation. Instead, there are programs which refine a finite element mesh. One program, **refineAll**, splits each element of a mesh into four smaller elements. This is called *h*-refinement. Another program, **pRefine**, converts linear elements to quadratic elements. This is called *p*-refinement. In both programs, care is taken to follow curved boundaries well.

Starting with a coarse mesh, **file.10**, composed of linear elements, a series of successively refined meshes, **file.qn**, composed of quadratic elements can be constructed. This is illustrated in figure A.6. Examples of unix commands to accomplish this are also listed. Notice that the two refinement programs read a mesh from the standard input and write a refined mesh to the standard output.

The finite element mesh files for the programs on the NeXTstation are incompatible with the programs on VM. A third program, **genYap2h**, converts a quadratic mesh file to a format suitable for the programs on the IBM/VM mainframe. This is shown in figure A.7.

The programs **refineAll**, **pRefine** and **genYap2h** are written in Objective-C, which is an object-oriented extension of C.

The grammar for a finite element mesh file is listed in figure A.8 in a format similar to the format required by the parser generator YACC. Selected portions of this grammar will be described below. Two examples of finite element meshes are listed in figure A.9.

The first token in figure A.8 represents the entire mesh:

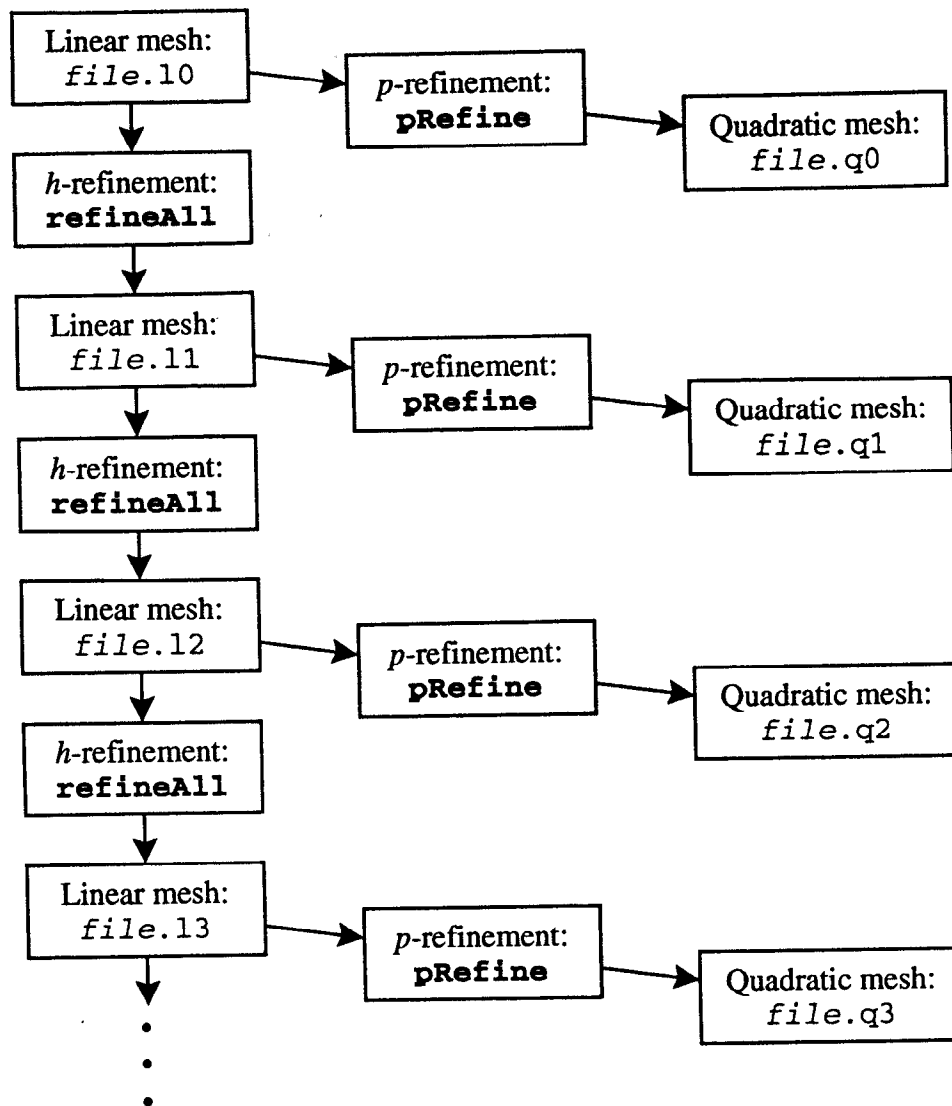
```
mesh :
    | mesh title
    | mesh defaultUnit
    | mesh geometryType
    | mesh nodeList
    | mesh elementList
    | mesh curveList
```

A mesh has three types of lists: node lists, element lists and curve lists. It may also have a title, default unit, or geometry statements. There is no fixed order for these lists and statements. A mesh must have at least a node list and an element list.

A node list is:

```
nodeList : NODES
          | nodeList aLength aLength
```

The keyword **nodes** is followed by pairs of node coordinates: first the *z* coordinate and then the ρ coordinate. For planar structures the *x* coordinate is first, followed



```

pRefine <file.10 >file.q0
refineAll <file.10 >file.11
pRefine <file.11 >file.q1
refineAll <file.11 >file.12
pRefine <file.12 >file.q2
:

```

Figure A.6: Construction of a series of successively refined quadratic meshes from a single coarse mesh of linear elements. Sample command lines are given below the flow diagram.

by the y coordinate. Each coordinate may include a unit. If no unit is specified, the default unit is assumed.



genYap2h <file.qn >file.yap2h

Figure A.7: Conversion of a quadratic finite element mesh for the NeXTstation (multipole) suite of programs to a quadratic finite element mesh for the IBM/VM (monopole and planar) programs. A sample command is given below the flow diagram.

An element list is:

```

elementList : ELEMENTS
              | elementList anElement
anElement   : LINTRI INTEGER INTEGER INTEGER
              | sideInfo sideInfo sideInfo
              | QUADTRI INTEGER INTEGER INTEGER INTEGER INTEGER INTEGER
              | sideInfo sideInfo sideInfo
  
```

The keyword `elements` is followed by a list of elements. A linear triangular element starts with the keyword `linTri` and a quadratic triangular element starts with the keyword `quadTri`. Following this keyword is a list of node indices: 3 for the linear element and 6 for the quadratic element. The nodes of the mesh are labelled sequentially starting with 1. The elements and curves of the mesh are labelled similarly. Following the node indices are 3 side specifications: one for each side of the triangular elements. The order of the node indices and side specifications is important. The order is shown in figures 2.3 and 2.6.

The pillbox example in figure A.9 has 4 nodes and 2 elements. The sphere example has 3 nodes and 1 element. The elements are linear triangular elements.

A side specification describes the boundary condition or element adjacent to the side. It optionally describes the curve which the side follows. This curve is used for mesh refinement only. It is not used in finite element calculations. The grammar for a side specification is:

```

sideInfo      : sideInfoSimple
               | sideInfoSimple sideCurveSpec
sideInfoSimple : INTEGER INTEGER
               | boundaryType;
  
```

If the side is adjacent to another element, then the side specification is the index of the adjacent element followed by the index of the adjacent side in the adjacent element. Two examples of this adjacent element side specification appear in the pillbox example in figure A.9. If the the side is on a boundary then a boundary specification is given:

```

mesh      :
           | mesh title
           | mesh defaultUnit
           | mesh geometryType
           | mesh nodeList
           | mesh elementList
           | mesh curveList
nodeList  : NODES
           | nodeList aLength aLength
elementList : ELEMENTS
           | elementList anElement
anElement : LINTRI INTEGER INTEGER INTEGER
           | sideInfo sideInfo sideInfo
           | QUADTRI INTEGER INTEGER INTEGER INTEGER INTEGER INTEGER
           | sideInfo sideInfo sideInfo
sideInfo  : sideInfoSimple
           | sideInfoSimple sideCurveSpec
sideInfoSimple : INTEGER INTEGER
boundaryType : metalBoundary
              | AXIS
              | periodicBoundary
              | symmetryBoundary;
metalBoundary : METAL
              | METAL ( INTEGER )
periodicBoundary : PERIODIC ( INTEGER INTEGER PM )
                 | PERIODIC ( INTEGER INTEGER INTEGER PM )
symmetryBoundary : SYMMETRY
                 | SYMMETRY ( INTEGER )
                 | SYMMETRY ( PM )
                 | SYMMETRY ( INTEGER PM )
sideCurveSpec : CURVE ( real real INTEGER )
              | CURVE ( real real aCurve )
curveList     : CURVES
              | curveList aCurve
aCurve        : NURBS INTEGER pointList weightList knotList
pointList     : POINTS
              | pointList aLength aLength
weightList    : WEIGHTS
              | weightList real
knotList      : KNOTS
              | knotList real
title         : TITLE QUOTEDSTRING
defaultUnit   : DIMENSION lengthUnit
lengthUnit    : CM | METER | INCH | MIL
aLength       : real
              | real lengthUnit
real          : REAL
              | INTEGER
geometryType  : AXISYMMETRIC | CARTESIAN

```

Figure A.8: Mesh grammar in a format similar to the format for the YACC parser generator. Uppercase words are terminal symbols. Except for INTEGER, REAL, PM and QUOTEDSTRING, the terminal symbols are literal.

```

title "1cm by 1cm pillbox"
dimension cm
axisymmetric
nodes /* z rho */
0.0 0.0
1.0 0.0
0.0 1.0
1.0 1.0
elements
linTri 2 3 1 2 1 metal axis
linTri 3 2 4 1 1 metal metal

```

```

title "sphere r=1cm"
dimension cm
axisymmetric
nodes /* z rho */
0.0 0.0
1.0 0.0
0.0 1.0
elements
linTri 2 3 1 metal curve(0.0 1.0 1) symmetry axis
curves
/* nurbs degree points weights knots */
nurbs 2 points 1.0 0.0 1.0 1.0 0.0 1.0 weights 1 0.707106781 1 knots 0 0 1 1

```

Figure A.9: Two examples of linear finite element meshes. The top example is a pillbox. The bottom example is a sphere.

```

boundaryType      : metalBoundary
                   | AXIS
                   | periodicBoundary
                   | symmetryBoundary;
metalBoundary     : METAL
                   | METAL ( INTEGER )
periodicBoundary  : PERIODIC ( INTEGER INTEGER PM )
                   | PERIODIC ( INTEGER INTEGER INTEGER PM )
symmetryBoundary  : SYMMETRY
                   | SYMMETRY ( INTEGER )
                   | SYMMETRY ( PM )
                   | SYMMETRY ( INTEGER PM )

```

The boundary specification is a keyword **METAL**, **AXIS**, **PERIODIC** or **SYMMETRY**. A metal boundary may be followed by an optional index in parenthesis, although this index is not used in the code at the moment.

A periodic boundary is followed by the element index and side index of the element side which is mapped onto this side by the periodicity's rigid motion or its inverse. The first index in parenthesis is the element index, the second index in parenthesis is the side index. If there are only three items in parenthesis, then the third item is a + or a -. One side of the cell should be +, while the other side of the

cell should be $-$. The rigid motion maps the $-$ boundary to the $+$ boundary. This case assumes the boundary belongs to the first periodicity. When there are four items in parenthesis, the third item is 0 or 1, indicating the periodicity to which this periodic boundary belongs. The fourth item is $+$ or $-$, as described above.

A symmetry boundary may be followed by an integer between 0 and 3 to indicate which symmetry plane the boundary lies on. The default symmetry plane is 0: the first symmetry plane.

A $+$ or $-$ in a symmetry boundary specification indicates the boundary is part of a symmetric periodic boundary. One symmetry plane is $+$ while the other symmetry plane is $-$.

A side specification allows an optional side curve specification:

```
sideCurveSpec : CURVE ( real real INTEGER )
               | CURVE ( real real aCurve )
```

The keyword `curve` is followed by the starting and ending parameter of a curve. These are the first two numbers in the parenthesis. The curve is specified by an index of a curve in the curve list. This is the third number in the parenthesis. A side curve specification can be seen in the sphere example in figure A.9. The curve may also be listed directly in side curve specification.

A curve list is:

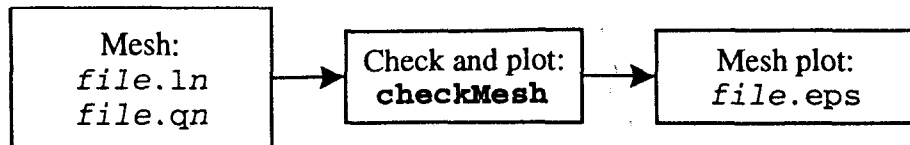
```
curveList : CURVES
           | curveList aCurve
aCurve    : NURBS INTEGER pointList weightList knotList
pointList : POINTS
           | pointList aLength aLength
weightList : WEIGHTS
           | weightList real
knotList   : KNOTS
           | knotList real
```

The keyword `curves` is followed by a list of curves. The only curve available at this time is a non-uniform rational B-spline, or `nurbs`. A `nurbs` curve has the flexibility to represent any conic section. The sphere example in figure A.9 shows such a curve. More information on `nurbs` curves can be found, for example, in [39].

The mesh file may be checked using a program `checkMesh` written in C++. This program can generate a plot of the mesh or its boundary. This is illustrated in figure A.10. The options for this program are also listed in the figure.

A.2.2. Field Solver

The relationship between the field solving codes and selected files and input parameters is shown in figure A.11. One program, `axiMnz`, constructs the matrices **A** and **B** for the multipole mode problem and generates field plots given a



```
checkMesh -gfile.eps file.qn
```

```
Usage: checkMesh [opts] [input-mesh-file]
  -B : do not plot interior edges
  -b : list boundaries present
  -e : calculate extent
  -g <file> : draw mesh to postscript file
  -o <file> : regurgitate input mesh to file
  -p <number> : draw field points - specify max along any axis
```

Figure A.10: Check and plot of a mesh. A sample command is given below the flow diagram, and the command options are listed below that.

solution eigenvector. The other program, **ma28Ges**, solves the generalized algebraic eigenvalue problem. Both programs are written in C++, although the **ma28** code on which the eigenvalue solver is based is in FORTRAN.

The program **axiMnz** constructs the matrices in a manner similar to **y2d** described earlier. However, there are two types of sparse symmetric matrices involved: quasi-periodic matrices and symmetric quasi-periodic matrices. These matrices store components as real numbers and have flags to represent quasi-periodic factors. Hence dispersion diagrams can be computed without recomputing the matrices from scratch.

The program writes the sparse symmetric matrices **A** and **B** to an output file. This file is read by **ma28Ges**. The options to **axiMnz** are listed in figure A.12. Notice the user is allowed to model either **E** or **H**, and that the user must specify the azimuthal symmetry number m . The quasi-periodic phase advance ψ is not specified in this program.

The program **ma28Ges** reads the matrix file generated by **axiMnz** and solves the generalized algebraic eigenvalue problem. The method is similar to **y2d** described earlier. However, **ma28Ges** uses a block shifted inverse power method, so it can find more than one eigenvalue and eigenvector simultaneously. The trial vectors converge to the eigenvectors with eigenvalue closest to the specified target eigenvalue. The iterations continue until a specified number of modes have a residual error smaller than the specified limit. The options for **ma28Ges** are listed in figure A.13. The examples in the same figure demand 3 good modes with residual error less than 10^{-25} . There are 6 trial vectors in the example, hence 3 of the trial

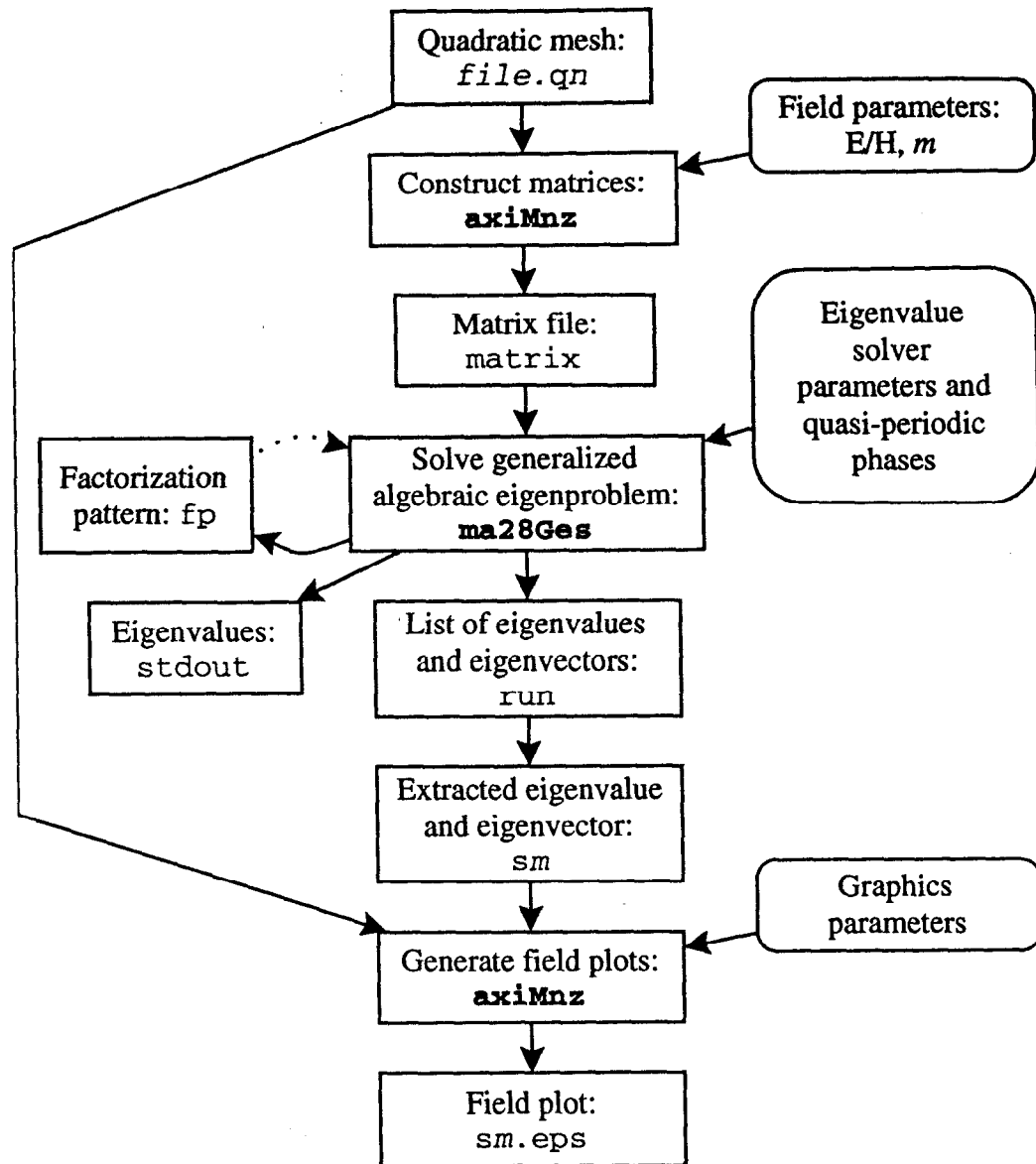


Figure A.11: Relationship between programs, files and input parameters for the YAP implementation of the multipole mode formulation. Programs and files are on the left, input parameters are on the right.

vectors are expected to be poor at the end of the calculation. These extra trial vectors are typically included in the calculation to improve the convergence rate of the iterative eigenvalue solver.

The program `ma28Ges` factorizes a matrix in order to solve the eigenvalue problem. Much of the time consumed by the `ma28` matrix solver is spent building a factorization pattern. The factorization pattern for the matrix may be written to a file and saved for later use. The same factorization pattern may be used for

```

axiMnz -m2 -Mmatrix -f1 file.qn
axiMnz -m2 -f1 -ss1 -gs1.eps file.qn

```

```

Usage: axiMnz [opts] [input-mesh-file]
-b <file> : basis index/flag output file
-f <flags> : flags for field and symmetries
  1 : model electric field (default is magnetic)
  2 : sym plane 0 is dirichlet - tangential field is zero
  4 : sym plane 1 is dirichlet - tangential field is zero
  8 : use sym-qp matrix if possible
-G : basis index/flag output is in genM format (obsolete)
-g <file> : graphics output file
-M <file> : matrix output file
-m <number> : azimuthal symmetry number (default m=1)
-P <phase> : plot phase in radians (default 0)
-p <integer> : max field points along any axis
-s <file> : solution input file
-v : only plot rho-z vector field
-z : only plot phi field

```

Figure A.12: A sample command and command options for the program `axiMnz`. The first command example generates a matrix, while the second example generates a field plot from a solution vector.

```

ma28Ges -iMatrix -M50 -r1e-25 -ffp -Erun -n6 -g3 -e1.5
ma28Ges -iMatrix -M50 -r1e-25 -Ffp -Erun -n6 -g3 -e3.5

```

```

Usage: ma28Ges [opts]
-c : force matrices to be complex
-d <file> : specify diagnostics output file
-E <file> : specify eigenvector output file
-e <eigenvalue> : specify target eigenvalue
-F <file> : specify factorization pattern input file
-f <file> : specify factorization pattern output file
-g <modes> : specify number of good modes
-i <file> : specify matrix input file
-M <iterations> : specify max iterations
-n <modes> : specify number of modes
-p <phase> : specify matrix phase (in radians)
-R <residual> : specify desired solve precision
-r <residual> : specify desired residual
-7 : matrix is in 'amniTest7' format

```

Figure A.13: A sample command and command options for the program `ma28Ges`. The first command writes a factorization pattern `fp`, while second command reads the factorization pattern `fp`.

various target eigenvalues, various phase advances, various m and various node coordinates in the mesh. Reusing the previously computed factorization pattern can yield tremendous savings in computational time when doing design studies or dispersion diagrams. The sample commands in figure A.13 demonstrate saving and

reusing a factorization pattern. Notice that a new target eigenvalue was chosen for the second sample command.

The eigenvector output file `run` includes all of the computed eigenvectors, not just the good ones. Using a text or stream editor, individual good eigenvectors can be extracted from `run` into separate files, say `s1`, `s2`, etc. The eigenvalue and phase advance is stored in the same file as the eigenvector.

Once a solution vector has been obtained, the program `axiMnz` can be run again to generate field plots. A sample command is shown in figure A.12.

REFERENCES

- [1] J. D. Jackson, *Classical Electrodynamics*, 2nd ed. New York: Wiley, 1975.
- [2] Robert E. Collin, *Field Theory of Guided Waves*, 2nd ed. New York: IEEE Press, 1991.
- [3] L. M. Falicov, *Group Theory and Its Physical Applications*. Chicago: University of Chicago Press, 1966.
- [4] A. G. Daikovsky, Yu. I. Portugalov and A. D. Ryabov, "PRUD-W: A New Code to Compute and Design Accelerating Structures," *Particle Accelerators*, vol. 20, pp. 23-43, 1986.
- [5] Claes Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge: Cambridge University Press, 1987.
- [6] Thomas J. R. Hughes, *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [7] Gilbert Strang and George J. Fix, *An Analysis of the Finite Element Method*, Englewood Cliffs, NJ: Prentice-Hall, 1973.
- [8] K. Halbach and R. F. Holsinger, "SUPERFISH—A Computer Program for Evaluation of RF Cavities with Cylindrical Symmetry," *Particle Accelerators*, vol. 7, pp. 213-222, 1976.
- [9] T. Weiland, "On the Computation of Resonant Modes in Cylindrically Symmetric Cavities," *Nuclear Instruments and Methods*, vol. 216, pp. 329-348, 1983.
- [10] U. van Rienen and T. Weiland, "Triangular Discretization Method for the Evaluation of RF Fields in Waveguides and Cylindrically Symmetric Cavities," *Particle Accelerators*, vol. 20, pp. 239-265, 1987.
- [11] T. J. R. Hughes and J. E. Akin, "Techniques for Developing 'Special' Finite Element Shape Functions with Particular Reference to Singularities," *Int. J. Num. Meth. Eng.*, vol. 15, pp. 733-751, 1980.
- [12] Daniel R. Lynch and Keith D. Paulsen, "Origin of Vector Parasites in Numerical Maxwell Solutions," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-39, pp. 383-394, Mar. 1991.
- [13] C. W. Crowley, P. P. Silvester and H. Hurwitz Jr., "Covariant Projections Elements for 3D Vector Field Problem," *IEEE Trans. Magnetics*, vol. MAG-24, pp. 397-400, Jan. 1988.

- [14] A. R. Pinchuk, C. W. Crowley and P. P. Silvester, "Spurious Solutions to Vector Diffusion and Wave Field Problems," *IEEE Trans. Magnetics*, vol. MAG-24, pp. 158-161, Jan. 1988.
- [15] J. P. Webb and Ruth Miniowitz, "Analysis of 3-D Microwave Resonators using Covariant-Projection Elements," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-39, pp. 1895-1899, Nov. 1991.
- [16] Ruth Miniowitz and J. P. Webb, "Covariant-Projection Quadrilateral Elements for the Analysis of Waveguides with Sharp Edges," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-39, pp. 501-505, Mar. 1991.
- [17] C. W. Crowley, "Mixed Order Covariant Projection Finite Elements for Vector Fields," Ph.D. dissertation, McGill University, Montreal, PQ, Canada, 1988.
- [18] J. C. Nedelec, "Mixed Finite Elements in \mathbf{R}^3 ," *Num. Math.*, vol. 35, pp. 315-341, 1980.
- [19] Jin-Fa Lee, Din-Kow Sun and Zoltan J. Cendes, "Full-Wave Analysis of Dielectric Waveguides Using Tangential Vector Finite Elements," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-39, pp. 1262-1271, Aug. 1991.
- [20] Jian-She Wang and Nathan Ida, "Eigenvalue Analysis in Anisotropically Loaded Electromagnetic Cavities Using Edge Finite Elements," *IEEE Trans. Magnetics*, vol. MAG-28, pp. 1438-1441, Mar. 1992.
- [21] Mitsuo Hano, "Finite-Element Analysis of Dielectric-Loaded Waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-32, pp. 1275-1279, Oct. 1984.
- [22] Kiyoshi Ise, Kazuhiro Inoue and Masanori Koshiba, "Three-Dimensional Finite-Element Method with Edge Elements for Electromagnetic Waveguide Discontinuities," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-39, pp. 1289-1295, Aug. 1991.
- [23] E. Okress, *Crossed-Field Microwave Devices*, New York: Academic Press, 1961.
- [24] K. Eppley, "Numerical Simulation of Cross Field Amplifiers," contributed to *Conference on Computer Codes and the Linear Accelerator Community*, Los Alamos, NM, January 22-25, 1990; also SLAC-PUB-5183.
- [25] K. Eppley and K. Ko, "Design of a High Power Cross Field Amplifier at X-Band with an Internally Coupled Waveguide," contributed to *SPIE's Symposium on High Power Lasers*, Los Angeles, CA, January 20-25, 1991; also SLAC-PUB-5416.

- [26] J. W. Wang and E. M. Nelson, "Design of the Detuned Accelerator Structure," in *Proc. IEEE Particle Accelerator Conf.*, 1993; also SLAC-PUB-6142.
- [27] R. D. Ruth, "The Development of the Next Linear Collider at SLAC," in *Proc., Physics and Experiments with Linear Colliders*, vol. 2, pp. 629-662, Saariselkæ, Finland, 1991; also SLAC-PUB-5729.
- [28] H. Deruyter, *et. al.*, "Damped and Detuned Accelerator Structures," *1990 Linear Accelerator Conf. Proc.*, Albuquerque, NM; also SLAC-PUB-5322.
- [29] K. Bane, private communication.
- [30] K. Bane and R. Gluckstern, "The Transverse Wakefield of a Detuned X Band Accelerator Structure," SLAC-PUB-5783, submitted to *Particle Accelerators*.
- [31] K. A. Thompson, C. Adolphsen and K.L.F. Bane, "Multibunch Beam Break-up in Detuned Structures," in *Proc. IEEE Particle Accelerator Conf.*, 1993; also SLAC-PUB-6153.
- [32] J.W. Wang and B.W. Littmann, "Design Study on Quasi-Constant Gradient Accelerator Structure," SLAC/AP-92, Sep. 1991.
- [33] C. Nantista, N.M. Kroll and E. M. Nelson, "Design of a 90° Overmoded Waveguide Bend," in *Proc. IEEE Particle Accelerator Conf.*, 1993; also SLAC-PUB-6141.
- [34] T. Weiland, "On the Unique Numerical Solution of Maxwellian Eigenvalue Problems in Three Dimensions," *Particle Accelerators*, vol. 17, pp. 227-242, 1985.
- [35] A. G. Daikovsky, Yu. I. Portugalov and A. D. Ryabov, "PRUD-Code for Calculation of the Nonsymmetric Modes in Axial Symmetric Cavities," *Particle Accelerators*, vol. 12, pp. 59-64, 1982.
- [36] S. P. Yu, G. P. Kooyers and O. Buneman, "Time-Dependent Computer Analysis of Electron-Wave Interaction in Crossed-Fields," *J. Appl. Phys.*, vol. 36, pp. 2550-2559, Aug. 1965.
- [37] J. Feinstein and R. J. Collier, "A Class of Waveguide-Coupled Slow-Wave Structures," *IRE Trans. Electron Devices*, p. 9, Jan. 1959.
- [38] S. H. Lo, "A New Mesh Generation Scheme for Arbitrary Planar Domains," *Int. J. Num. Meth. Eng.*, vol. 21, pp. 1403-1426, 1985.
- [39] Gerald Farin, *Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide*, 2nd ed., New York: Academic Press, 1990.