

AD-A273 074



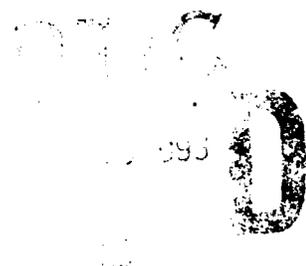
2

RL-TR-93-130
In-House Report
July 1993



HYBRID FINITE ELEMENT/WAVEGUIDE MODE ANALYSIS OF PASSIVE RF DEVICES

Daniel T. McGrath, Major, USAF



APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

93-28509



Rome Laboratory
Air Force Materiel Command
Griffiss Air Force Base, New York

93 11 22 1 27

This report has been reviewed by the Rome Laboratory Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RL-TR-93-130 has been reviewed and is approved for publication.

APPROVED:



ROBERT J. MAILLOUX, Acting Chief
Antenna & Components Division
Electromagnetics & Reliability Directorate

FOR THE COMMANDER:



JOHN K. SCHINDLER
Director of Electromagnetics & Reliability

If your address has changed or if you wish to be removed from the Rome Laboratory mailing list, or if the addressee is no longer employed by your organization, please notify RL(ER) Hanscom AFB MA 01731-5000. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE July 1993	3. REPORT TYPE AND DATES COVERED In-House, Jan 1992-Jan 1993
----------------------------------	------------------------------------	--

4. TITLE AND SUBTITLE Hybrid Finite Element/Waveguide Mode Analysis of Passive RF Devices	5. FUNDING NUMBERS PE: 61102F PR: 2304 TA: 13 WU: 02
---	--

6. AUTHOR(S) Daniel T. McGrath, Major, USAF	
---	--

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Dept. of Electrical and Computer Engineering Wright-Patterson AFB, OH 45433-6583	8. PERFORMING ORGANIZATION REPORT NUMBER
---	--

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Rome Laboratory (ERA) 31 Grenier St. Hanscom AFB, MA 01731-3010	10. SPONSORING / MONITORING AGENCY REPORT NUMBER RL-TR-93-130
---	---

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited	12b. DISTRIBUTION CODE
--	------------------------

13. ABSTRACT (Maximum 200 words)

A numerical solution for time-harmonic electromagnetic fields in two-port passive radio frequency (RF) devices has been developed, implemented in a computer code, and validated. Vector finite elements are used to represent the fields in the device interior and field continuity across waveguide apertures is enforced by matching the interior solution to a sum of waveguide modes. Consequently, the mesh may end at the aperture instead of extending into the waveguide. The report discusses the variational formulation and its reduction to a linear system using Galerkin's method. It describes the computer code, including its interface to commercial CAD software used for geometry generation. It presents validation results for waveguide discontinuities, coaxial transitions and microstrip circuits. They demonstrate that the method is an effective and versatile tool for predicting the performance of passive RF devices.

14. SUBJECT TERMS Electromagnetics Microwave circuits Waveguides Finite elements Numerical methods Computation RF components Waveguide discontinuities	15. NUMBER OF PAGES 88
16. PRICE CODE	

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited
--	---	--	--

Accession For	
NTIS GR&I	4
DTIC TAB	
Unannounced	
Justification	
By	
Date	
Dist	
A-1	

Contents

1.	INTRODUCTION AND BACKGROUND	1
2.	PROBLEM DESCRIPTION AND SOLUTION APPROACH	2
3.	VARIATIONAL STATEMENT AND BOUNDARY CONDITIONS	5
	3.1. Variational Statement	5
	3.2. Waveguide Field Representation	6
	3.3. Transverse Field Continuity at the Apertures	7
4.	DISCRETIZATION	9
	4.1. Scalar vs Vector Expansion Functions	9
	4.2. Linear Expansion Functions on Triangles and Tetrahedra	12
	4.3. Vector Edge-Based Expansion Functions	15
	4.4. Discretization Using Galerkin's Method	16
	4.5. Volume Integral Computations	18
	4.6. Boundary Integral Computations	19
	4.7. Matrix Structure and Solution Methods	21
5.	COMPUTER CODE DESIGN	24
	5.1. Computational Algorithms	25
	5.2. Sparse Storage Structure and Conjugate Gradient Algorithm	29
	5.3. I-DEAS TM Interface	30
6.	VALIDATION AND DEMONSTRATION	32
	6.1. Waveguides and Waveguide Discontinuities	33
	6.1.1. Waveguide Propagation	33
	6.1.2. Waveguide Irises	36

Contents

6.1.3. Circular Waveguide Mode Converter	37
6.2. Waveguide Transitions	40
6.2.1. Coax-Stripline Transition	40
6.2.2. Coax-Rectangular Waveguide Launcher	43
6.3. Printed Circuit Devices	45
6.3.1. Microstrip Transmission Line	45
6.3.2. Microstrip Meander Line	46
6.4. Modeling Requirements	48
6.4.1. Higher-Order Modes	48
6.4.2. Discretization Requirements	52
6.4.3. Matrix Solution	53
7. CONCLUSIONS AND RECOMMENDATIONS	54
REFERENCES	57
APPENDIX A: THE ELECTRIC FIELD FUNCTIONAL	59
A1. Variational Principle vs Weak Form	59
A2. The Adjoint Problem	61
A3. Continuity Conditions for Waveguide Apertures	63
A4. Galerkin's Method vs Rayleigh-Ritz	64
APPENDIX B: COAXIAL AND CIRCULAR WAVEGUIDE MODE FUNCTIONS	67
B1. COAXIAL	67
B1.1. Mode Functions	67
B1.2. Computations	70
B2. CIRCULAR WAVEGUIDE	72
APPENDIX C: RECTANGULAR WAVEGUIDE MODE FUNCTIONS AND INNER PRODUCT INTEGRATIONS	77

Illustrations

1. Generic Waveguide/Cavity Problem	3
2. Example Problems: (a) Waveguide Launcher; (b) Microstrip Circuit	4
3. Subdivision of a Volume Region (one quadrant of a disk) Into Tetrahedra: (a) interior edges visible; (b) interior edges hidden	10
4. A Two Dimensional Linear Expansion Function (defined over triangular cells adjacent to a node and zero on all other triangles)	13

Illustrations

5. A Three Dimensional Linear Expansion Function: (left) eight cells surrounding a node; (right) surfaces of constant function value in one cell	13
6. A Two-Dimensional Vector Expansion Function: (a) direction (arrow length indicates function magnitude); (b) constituent linear finite elements	16
7. Sample Matrix Structure (dark points are nonzero elements)	23
8. Convergence of Residual Norm and Reflection Coefficient using Conjugate Gradient Method	25
9. Flow Diagram for Program TWOPORT	26
10. Example Input Geometry File (as Created by Translator Program from CAD Universal File; italics are annotation)	27
11. Structure of Finite Element Interior Computations	29
12. Structure for Waveguide Computations	29
13. Coaxial Waveguide Section: (a) Solid Object; (b) Tetrahedron Mesh	31
14. Propagation Phase Through 4 mm APC-7 Coaxial Waveguide Section	33
15. Finite Element Mesh for a Rectangular Waveguide Section	34
16. Propagation Phase Delay, Rectangular Waveguide Section	35
17. Finite Element Mesh for a Circular Waveguide Section	35
18. Propagation Phase Through Circular Waveguide Section	36
19. Measured and Calculated Reflection Coefficient vs Frequency for Coaxial Irises (APC-7mm coax): (a) 180° Opening; (b) 270° Opening	38
20. Finite Element Mesh for Diagonal Iris in Rectangular Waveguide: (a) Complete Mesh; (b) Cells above $z=5$ mm removed	39
21. Measured and Calculated S_{11} for Diagonal Iris in Rectangular Waveguide	39
22. Finite Element Mesh for Circular Waveguide Step Discontinuities (Mode Converter), with $a_2=1.4$ in.: (a) $a_1=1.05$ in.; (b) $a_1=1.15$ in.	41
23. Mode Conversion Ratio for Circular Waveguide Mode Converters, Hybrid Finite Element (HFEM) and Multimode Calculations	41

Illustrations

24. Stripline Test Case: (a) Complete Problem Geometry; (b) Lower Half of Finite Element Mesh; and (c) Dimensions at Coax/Stripline Junction	42
25. Reflection and Transmission vs Frequency for 4 mm Long Stripline Section	44
26. Coaxial End-Wall Launcher for Rectangular Waveguide: (a) Geometry and Dimensions; (b) Cutaway with Right-half Mesh Cells;	44
27. Reflection Coefficient vs Frequency for Coax-Rectangular Waveguide Launcher	45
28. Finite Element Mesh for Microstrip Line with Coaxial Connectors	46
29. Microstrip Meander Line: (a) Wafer Metallization Dimensions (in μm); and (b) Finite Element Mesh for Substrate	47
30. Comparison of Measured and Calculated S_{21} Phase for Microstrip Meander Line	48
31. Mode Excitation Coefficients at Coaxial Waveguide Irises: (a) 180° Iris Opening; (b) 270° Iris Opening	50
32. Reflection Coefficient Magnitude vs Frequency Calculated with Varying Numbers of Higher Order Modes, 270° Coaxial Iris	51
33. Higher-Order Coaxial Mode Excitation Coefficients at Coax-Stripline Transition	51
34. Execution Time vs Number of Unknowns for LU Decomposition and Conjugate Gradient Method	53
B1. Transverse Electric Characteristic Functions	71
B2. Transverse Magnetic Characteristic Functions	71
B3. APC-7 Mode Cutoff Frequency vs Index n	73
B4. APC-7 Mode Cutoff Frequency vs Index m	73
B5. Electric Field Contours for TE ₂₁ Mode: top, ρ -component; bottom, ϕ -component	74
B6. Electric Field Contours for TM ₂₁ Mode: top, ρ -component; bottom, ϕ -component	75

Table

1. Summary of Finite Element Discretization for Test Cases	54
--	----

Hybrid Finite Element/Waveguide Mode Analysis of Passive RF Devices

1. INTRODUCTION AND BACKGROUND

A numerical solution for the time-harmonic electric fields inside a cavity fed by two waveguides at opposite ends has been implemented and validated. The generic problem is that of a two-port RF (radio frequency) circuit. Specific test results were obtained for a variety of problems, including waveguide discontinuities and microwave circuits. The emphasis of this report is on the analysis technique, which uses the finite element method in conjunction with waveguide modes in a hybrid scheme. The results of this work are both a demonstration of the capabilities that may be obtained using the finite element method and a computer code that is accurate and versatile.

This is an interim report in a project dealing with the analysis of antennas using the finite element method (FEM). The results described here demonstrate some of the key building blocks, especially the coupling of the FEM solution to higher-order waveguide modes. This will lead to an accurate feed description for antennas that are driven by waveguides. There are other methods available for waveguide discontinuity and RF circuit analysis, but the present approach is partly justified by its intended applica-

tion. Nonetheless, the novel approach of using higher-order waveguide modes in conjunction with the finite element method has, for many problems, a decided advantage over previous solutions that used only the lowest order waveguide mode and extended the finite element region into the waveguide far enough to make evanescent modes negligible.

This report will discuss the general cavity/waveguide problem and its formulation as a variational problem using weighted residuals and Galerkin's method (Sections 2 and 3). These two chapters are important to understanding how the physical problem is cast as a mathematical boundary value problem. Section 4 will show the details of the derivation of a matrix equation using vector edge-based finite elements, as well as the methods used to solve the equation and obtain S parameter information from the solution. That chapter is important to the reader who wishes to understand the details of the finite element and waveguide mode solutions. Section 5 will discuss the structure of the analysis code, including its interface to commercial software used for geometry definition and mesh generation. That information will be useful mainly to the reader interested in using or modifying the code, or in developing a similar code. Last, Section 6 will show the results of test cases that validate the code's accuracy and demonstrate its usefulness.

2. PROBLEM DESCRIPTION AND SOLUTION APPROACH

The general problem to be solved is illustrated in Figure 1. The cavity region, denoted Ω , has perfectly conducting walls except for apertures at $z=0$ and $z=d$. The apertures need not extend over the entire waveguide cross sections. The boundary of Ω is denoted Γ , and the two apertures' open areas are Γ_A and Γ_B . The two waveguides need not be the same size or even the same type. Derivations and results will be presented for coaxial, circular and rectangular waveguides. The apertures are in the waveguides' end walls.

The cavity interior may be regarded as an arbitrary two-port, passive, linear device. It may

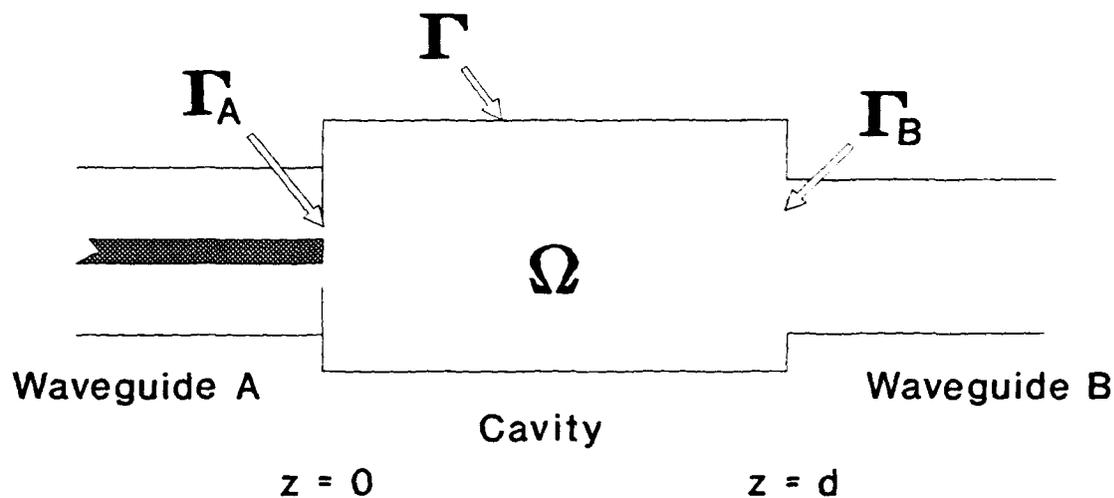
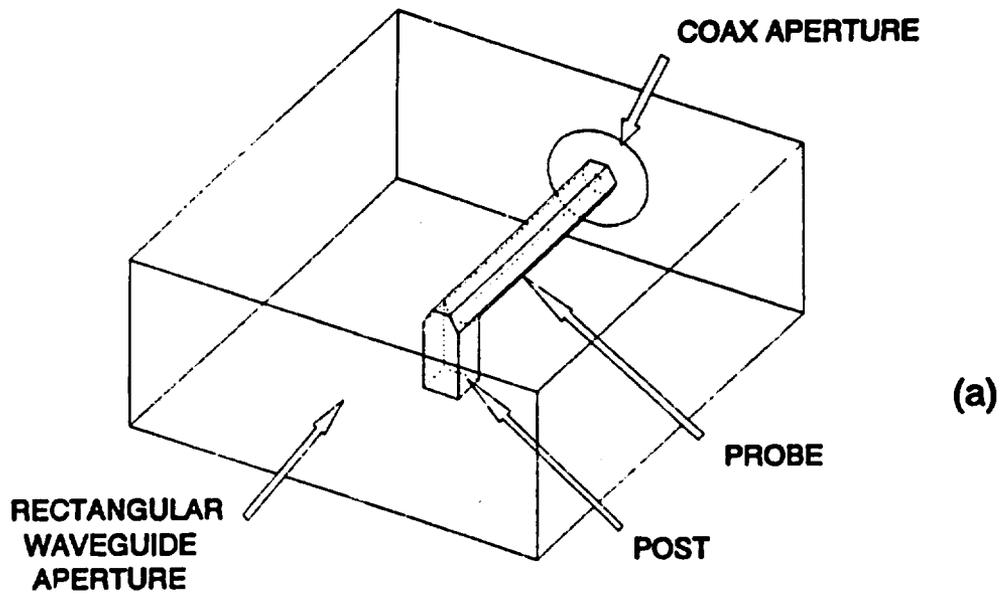


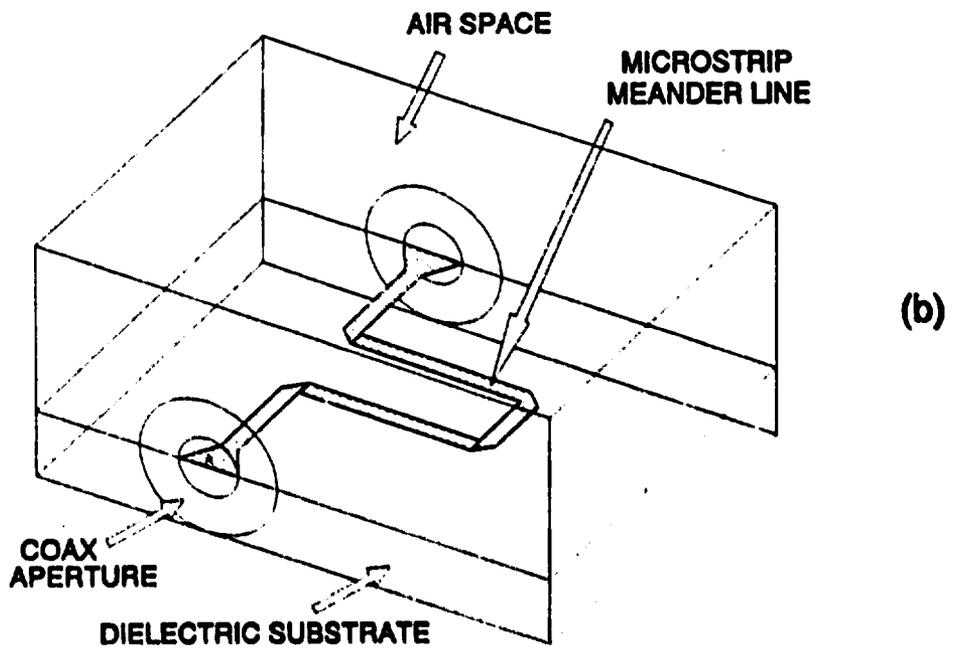
Figure 1. Generic Waveguide/Cavity Problem

contain any mixture of isotropic dielectric and ferrite materials, as well as perfectly conducting obstacles. Those features are adequate to describe a broad range of practical RF devices. Some examples are shown in Figure 2. In the first (Figure 2a) the inlet waveguide is a coaxial line whose center conductor extends into the cavity region, a short section of rectangular waveguide. The outlet waveguide is the same size rectangular waveguide. In the second example (Figure 2b) the cavity is a rectangular conducting box containing a dielectric slab that supports a microstrip conductor (a meander line). The inlet and outlet ports are both coaxial waveguides.

An overview of the solution procedure is as follows: (1) Represent the waveguide fields in terms of orthonormal mode functions; (2) Represent the field inside Ω by a variational functional; (3) Match the transverse magnetic field across the apertures by incorporating the waveguide mode fields in the functional's boundary term; (4) Generate a linear system by expanding the field in terms of known functions; and (5) compute the devices' S parameters from the solution of the linear system (matrix equation). The following two sections provide the details of steps 1 - 3 and 4 & 5, respectively.



(a)



(b)

Figure 2. Example Problems: (a) Waveguide Launcher; (b) Microstrip Circuit

3. VARIATIONAL STATEMENT AND BOUNDARY CONDITIONS

3.1. Variational Statement

The finite element method is usually applied to a *functional*, an integral that contains the unknown field as part of the integrand [1]. (A functional is a mapping of a function into a number. In other words, the result of the integration is a single number. This is in contrast to an integral equation, which also contains an unknown function in the integrand, but which results not in a number, but in another function.) The functional is also a *variational statement*, meaning that the correct solution for the unknown field is the function that minimizes the integral. In this case, the variational statement is a *weak form* of the time-harmonic wave equation for the electric field inside Ω . The source-free form is appropriate to this problem because all sources are assumed to be located at $z \ll 0$ in waveguide A. The weak form is the functional

$$F(\bar{E}) = \int_{\Omega} \left[\frac{1}{\mu_r} \nabla \times \bar{W}^* \cdot \nabla \times \bar{E} - k_0^2 \epsilon_r \bar{W}^* \cdot \bar{E} \right] dv - j k_0 \eta_0 \int_{\Gamma} \bar{W}^* \cdot \hat{n} \times \bar{H} ds = 0 \quad (1)$$

where \bar{E} is the unknown electric field, \bar{W} is a trial function, \bar{H} is the magnetic field, k_0 is the free space wavenumber, η_0 is the free space impedance and \hat{n} is the outward normal to Γ . This is applicable to linear, isotropic, inhomogeneous materials. Appendix A discusses how this is derived and why this form is more appropriate than others that appear in the electromagnetics literature. This functional is a statement of conservation of energy, with the volume integral representing stored energy and dissipation in materials, and the surface integral representing power flow into and out of the volume. For that reason, the perfectly conducting portions of Γ do not contribute to the surface integral, and it reduces to an integration over Γ_A and Γ_B , where it will be used to enforce continuity of tangential \bar{H} .

3.2. Waveguide Field Representation

The following derivation is appropriate for arbitrary waveguide modes. It is valid for any type of waveguide for which orthonormal mode functions are known: rectangular, circular, coaxial and elliptical. Appendix B gives the specific forms for coaxial and circular waveguides. Appendix C discusses rectangular waveguides.

In waveguide A, the electric field transverse to z may be written in terms of orthonormal mode functions \bar{g}_i [2],[3]:

$$\bar{E}_t^A = e^{-\gamma_0 z} \bar{g}_0 + \sum_{i=0}^{\infty} C_i e^{\gamma_i z} \bar{g}_i \quad (2)$$

The index i includes TE and TM modes (and TEM if applicable) with $i=0$ corresponding to the dominant mode. The first term represents the incident field, assumed to be of unit magnitude and in the dominant mode only. The unknown coefficients C_i are the excitation coefficients for the modes reflected by the device. C_0 is particularly significant: $C_0 - 1$ is S_{11} , the reflection coefficient seen by an observer in waveguide A. A similar expression holds for waveguide B, except that there is no incident field term, and the coefficient for the dominant mode will be proportional to S_{21} , the transmission coefficient. The mode functions are entirely transverse to the waveguide axis (the z axis). The propagation constants are related to the mode cutoff wavenumbers by

$$\gamma_i = \sqrt{k_c^2 - k_0^2} \quad (3)$$

which is positive imaginary for propagating modes and positive real for evanescent modes.

The transverse magnetic field in waveguide A is

$$\bar{H}_t^A = Y_0 e^{-\gamma_0 z} (\hat{z} \times \bar{g}_0) - \sum_{i=0}^{\infty} C_i Y_i e^{\gamma_i z} (\hat{z} \times \bar{g}_i) \quad (4)$$

where Y_i is the modal admittance:

$$Y_i = \begin{cases} \frac{-j\gamma_i}{k_0 \eta_0 \mu_r} & \text{(TE)} \\ \frac{jk_0 \epsilon_r}{\gamma_i \eta_0} & \text{(TM)} \\ \eta^{-1} & \text{(TEM)} \end{cases} \quad (5)$$

The fields in waveguide B are essentially identical, except that: (a) there is no incident field term; (b) the fields are propagating in the +z direction; and (c) its mode functions, modal admittances, and propagation constants need not be the same as waveguide A.

3.3. Transverse Field Continuity at the Apertures

At $z=0$ the outward normal to Ω is $-\hat{z}$, so from Eq.(4):

$$\hat{n} \times \bar{H}_t^A \Big|_{z=0} = Y_0 \bar{g}_0 - \sum_{i=0}^{\infty} C_i Y_i \bar{g}_i \quad (6)$$

The C_i 's are found by taking the inner product of Eq.(2) with every \bar{g}_i over the waveguide cross section and using the orthogonality property, with the result:

$$C_i = \int_{\Gamma_A} \bar{E}_t^A \Big|_{z=0} \cdot \bar{g}_i ds - \delta_{0i} \quad (7)$$

The limits of integration above are over the waveguide end wall aperture instead of its cross section because the transverse electric field is zero on the conducting surfaces at $z=0$.

Now let $(F_A - F_{inc})$ represent that part of the boundary integral in Eq.(1) over Γ_A and substitute Eqs.(6) and (7):

$$F_A - F_{inc} = -jk_0 \eta_0 \left\{ Y_0 \int_{\Gamma_A} \bar{W}^* \cdot \bar{g}_0 ds - \sum_{i=0}^{\infty} Y_i \int_{\Gamma_A} \bar{W}^* \cdot \bar{g}_i ds \left[\int_{\Gamma_A} E_i^A \cdot \bar{g}_i ds - \delta_{0i} \right] \right\} \quad (8)$$

$$= jk_0 \eta_0 \left\{ -2 Y_0 \int_{\Gamma_A} \bar{W}^* \cdot \bar{g}_0 ds + \sum_{i=0}^{\infty} Y_i \int_{\Gamma_A} \bar{W}^* \cdot \bar{g}_i ds \int_{\Gamma_A} E_i^A \cdot \bar{g}_i ds \right\} \quad (9)$$

The first term inside the braces in Eq.(9) is the negative of the incident field term, $-F_{inc}$.

In the same manner, the boundary functional for waveguide B can be written in terms of its mode functions \bar{h}_i and modal admittances Y_i' :

$$F_B = +jk_0 \eta_0 \sum_{i=0}^{\infty} Y_i' \int_{\Gamma_B} \bar{W}^* \cdot \bar{h}_i ds \int_{\Gamma_B} \bar{E}_i^B \cdot \bar{h}_i ds \quad (10)$$

Let F_I denote the volume integral in Eq.(1), which is the part of the functional dealing with the interior fields. Then the functional equation may be written symbolically as

$$F_{inc} = F_I + F_A + F_B \quad (11)$$

When the unknown and trial fields are expanded in series of known functions with unknown coefficients, Eq.(11) will become a matrix equation:

$$E^{inc} = [S^I + S^A + S^B] E \quad (12)$$

where E^{inc} is a column vector representing the incident field, E is a column vector representing the unknown field at locations inside the volume, and S^I , S^A and S^B are matrices. The following section gives the details of the discretization process by which the continuous functional is mapped into the linear system to give expressions for the entries in E^{inc} , S^I , S^A and S^B in terms of known quantities.

4. DISCRETIZATION

The finite element method generates a linear system of equations from a linear functional. The alternative used here is Galerkin's method, for reasons explained in Appendix A. It is quite similar to the method of moments, the distinction being that MOM applies weighted residuals to an integral equation rather than to a functional [Reference 4, p.18].

The first three parts of this section discuss the expansion functions that were chosen to represent the field and trial functions: the relative merits of scalar and vector expansion functions; the structure of the scalar functions and their representation in homogeneous coordinates; and the construction of the vector expansion functions from the scalar functions. The fourth part discusses the weighted residuals procedure leading to the matrix equation. The fifth and sixth parts derive explicit expressions for the matrix elements and the incident field terms. The last part discusses methods for solving the matrix.

4.1. Scalar vs Vector Expansion Functions

Figure 3 illustrates a gridding of a volume region into small elements. These elements or *cells* are tetrahedra, each of which has four faces and four vertices, or *nodes*. This figure illustrates an important feature of the finite element method: the grid density may be varied within the problem space so that it is coarse in regions where the fields are expected to be well-behaved; and fine in regions where the geometry is more complex or where fields may have singularities.

The electric field inside the region shown is to be expanded in terms of some form of known functions defined over individual cells with unknown coefficients representing the field strength and direction. The decision of whether to use scalar functions with vector coefficients or vector functions with scalar coefficients, as contrasted in the following two equations, is much more than a matter of preference.

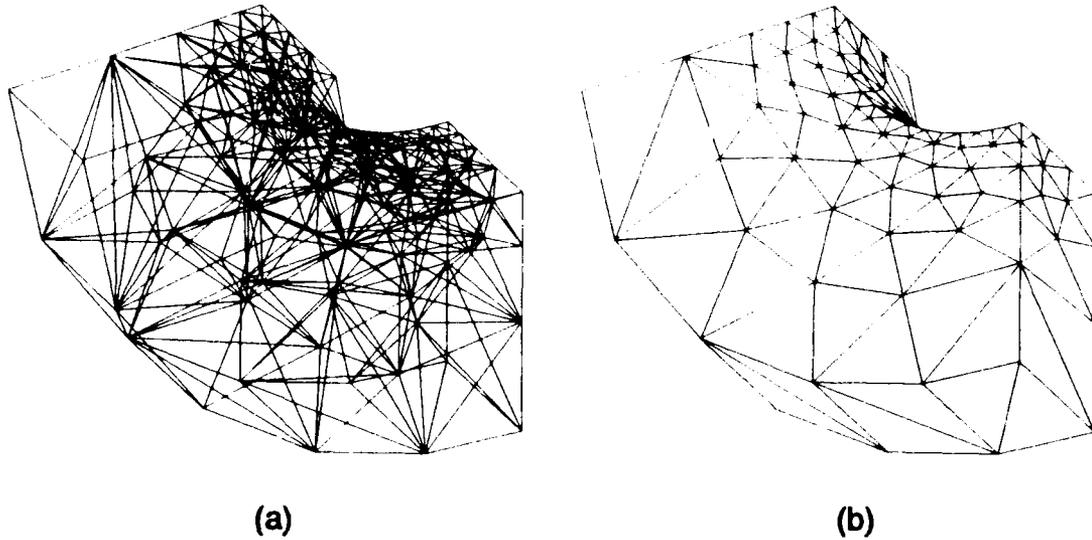


Figure 3. Subdivision of a Volume Region (one quadrant of a disk) Into Tetrahedra:
 (a) interior edges visible; (b) interior edges hidden

$$\bar{E} = \sum_{s=1}^M \bar{e}_s \phi_s(x, y, z) \quad (13)$$

$$\bar{E} = \sum_{s=1}^N e_s \bar{\psi}_s(x, y, z) \quad (14)$$

Traditional finite element calculations involve solving for the unknown function at the grid nodes. Hence, Eq.(13) seems the most natural expansion for the electric field. M is the total number of nodes in the volume. The function ϕ_s is a *subdomain* function, meaning that it is only defined over a small portion of the volume, specifically, those cells adjacent to node s . In contrast, the index s in Eq.(14) is over the grid *edges* (the lines joining nodes).

The node-based expansion in Eq.(13) has three important disadvantages:

- (a) It is difficult to enforce the boundary condition $\hat{n} \times \bar{E} = 0$ at nodes on perfect conductors (mainly in an algorithmic sense). Worse yet, if a node is on a conducting edge or tip, it is not

possible to enforce the condition because \hat{n} is not defined. The only remedy is to refine the mesh near edges so that their erroneous contributions are less significant.

(b) The expansion of Eq.(13) does not meet the divergence condition $\nabla \cdot (\epsilon \bar{E}) = 0$. It has been presumed that this could be circumvented by using a penalty function, but Boyse et al. point out that penalty methods are only justified for positive definite functionals [5], whereas Eq.(1) is indefinite.

(c) Suppose that the node s is on an interface between dielectrics. Then the component of \bar{E} normal to the interface must be discontinuous. Yet Eq.(13) implies that all components of \bar{E} are continuous at s . Hence the scalar functions over-prescribe the continuity at dielectric interfaces [6]. The remedy is to decrease the mesh cell size around dielectric boundaries so that the error contributed by the boundary nodes is less significant.

All three of these disadvantages are sources of inaccuracy and all three incur a computational cost: (a) and (c) in terms of matrix size; and (b) in terms of matrix element calculation time.

Several authors have reported success using vector expansion functions as in Eq.(14) in cases where the scalar functions failed [6],[7]. The specific form of the function $\bar{\psi}_s$ will be described in detail in Section 4.3. Its advantages are:

- (a) It meets the divergence condition.
- (b) It enforces only tangential field continuity across boundaries; and
- (c) It allows a simple yet effective method for implementing conductor boundary conditions, even at edges.

The first two are vital in terms of using the finite element method. These two qualities assure that the expansion and trial functions are *admissible*. The vector element's disadvantages are:

- (a) It is not compatible with most existing preprocessing and postprocessing software (usually CAD packages oriented to finite element computations for mechanical engineering), so additional

algorithms and computations are needed to transform the input geometry and output results; and

(b) There are generally 4-5 times as many edges as nodes in a tetrahedron mesh, and so the edge-based matrix will be $4/3 - 5/3$ larger. This is partly ameliorated by the fact that there is less need for fine sampling at material interfaces and conductor edges. Furthermore, the connectivity of edges is less than for nodes, so the edge-based matrix is considerably more sparse, and thus the total number of matrix entries may actually be less. The overall matrix size is a greater concern when using LU decomposition, but the total number of nonzero entries is the most important for an iterative solver such as the conjugate gradient method.

In summary, the edge-based vector functions were adopted because their disadvantages affect algorithm complexity, while their advantages in theoretical and mathematical correctness provided a path with considerably less risk.

4.2. Linear Expansion Functions on Triangles and Tetrahedra

Since the vector expansion functions $\vec{\psi}$ are defined in terms of the more conventional scalar functions ϕ , the latter are described here in some detail. Each ϕ_s is defined over the cells (tetrahedra) adjacent to node s . Its value is 1 at node s , and zero at the surrounding nodes and opposing faces. It may in general have a polynomial variation over the cells in which it is defined, but the present work uses linear functions only.

A two-dimensional linear function is shown in Figure 4. It is defined over the triangles adjacent to a single node, where its value is 1. This function may be regarded as being assembled from "linear finite elements" that are defined on individual triangles. Each triangle supports three finite elements, each of which is 1 at a single node and zero along the opposite edge. Figure 5 illustrates the three-dimensional equivalent: The node is at the center of a group of 8 tetrahedra that form a diamond shape. One tetrahedron is shown expanded with dashed lines indicating surfaces where the function has constant value. The

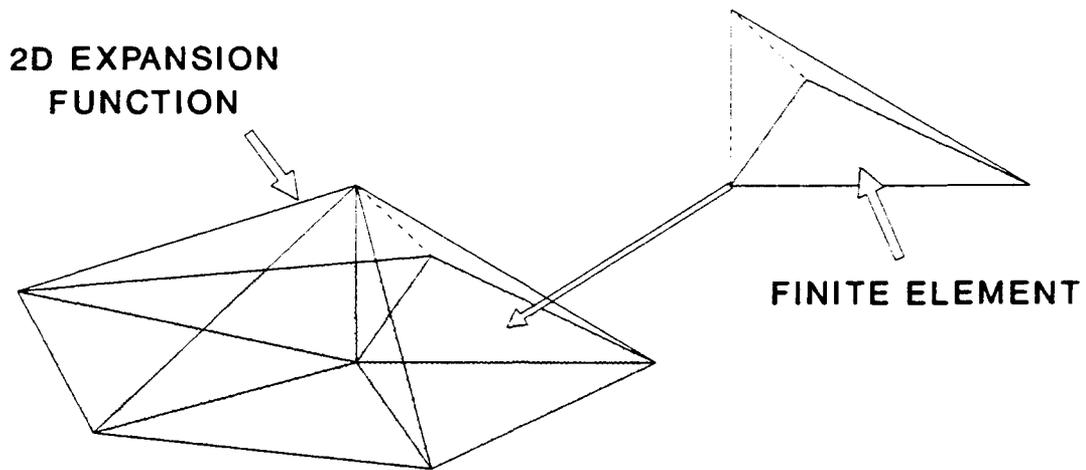


Figure 4. A Two Dimensional Linear Expansion Function (defined over triangular cells adjacent to a node and zero on all other triangles)

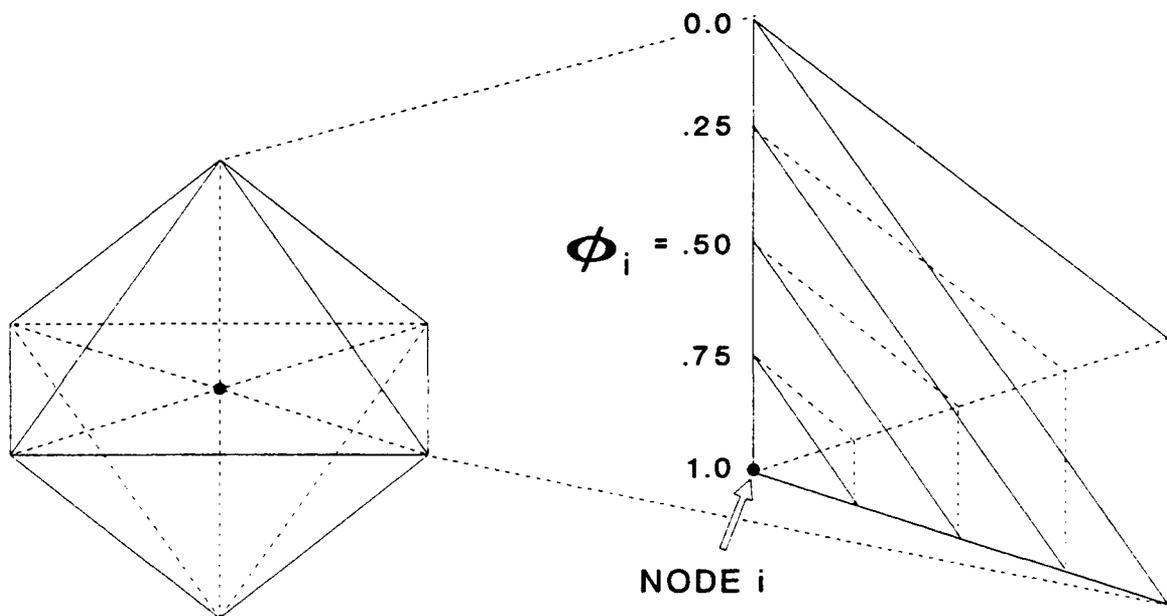


Figure 5. A Three Dimensional Linear Expansion Function: (left) eight cells surrounding a node; (right) surfaces of constant function value in one cell

three-dimensional ϕ_s is constructed from functions that are defined over individual tetrahedra. For example, within the tetrahedron with index k , there are four linear finite elements $f_i^{(k)}$, one for each node $i=1,2,3,4$. The index i is a local node index while s is a global node index. An expression for ϕ_s is

$$\phi_s = \sum_{k=1}^K \sum_{i=1}^4 f_i^{(k)} \delta(s;i,k) \quad (15)$$

The notation $\delta(s;i,k)$ is a Kronecker delta that equals 1 if the global node s is the same as the local node i within cell k , and equals zero otherwise. K is the total number of cells in the mesh. The notation in Eq.(15) is merely a formal representation for the *assembly* procedure that allows matrix element calculations to be performed one cell at a time, which is considerably more efficient than performing them one node at a time.

The integrations leading to matrix elements are products of these linear functions and their partial derivatives. Those integrations may be performed easily in closed form using *homogeneous* coordinates (also known as *simplex* coordinates) [Reference 8, pp.266-274]. The coordinate transformation is defined locally within each cell. In the tetrahedron there are four coordinates t_1, t_2, t_3 and t_4 . The coordinate t_i of a point anywhere within the cell is the normalized distance to node i from the opposing face (normalized to the height of the cell in that direction). The coordinate transform is given in terms of a matrix T :

$$\begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{bmatrix} = \frac{T}{6V} \begin{bmatrix} 1 \\ x \\ y \\ z \end{bmatrix} \quad (16)$$

where V is the cell volume. The entries in T are the 16 cofactors of the following matrix made up of the cell vertex coordinates:

$$U = \begin{bmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{bmatrix} \quad (17)$$

The two reasons these coordinates are so convenient are: (a) the linear expansion functions are simply

$$f_i(x, y, z) = t_i = \frac{1}{6V} [T_{i1} + xT_{i2} + yT_{i3} + zT_{i4}] \quad (18)$$

and (b), the limits of integration are simplified, leading to the following result for the inner product of any two linear expansion functions over a single cell:

$$\begin{aligned} \int_{\text{cell } k} \int \int f_i f_j dx dy dz &= 6V \int_0^1 dt_1 \int_0^{1-t_1} dt_2 \int_0^{1-t_1-t_2} t_i t_j dt_3 \\ &= \frac{V}{20} (1 + \delta_{ij}) \end{aligned} \quad (19)$$

This simplicity, one of the attractive features of FEM, applies equally to the vector expansion functions.

4.3. Vector Edge-Based Expansion Functions

The vector expansion function is defined relative to the edge joining nodes i and j [7]:

$$\vec{\psi}_{ij} = L_{ij} (f_i \nabla f_j - f_j \nabla f_i) \quad (20)$$

where f_i and f_j are the linear node-based functions and L_{ij} is the edge length. The scaling by L_{ij} ensures that the component of $\vec{\psi}_{ij}$ tangential to the edge is a unit vector. Figure 6 illustrates the two dimensional version of this function. Observe that ∇f_i is a vector in the direction of steepest ascent of f_i , so it is directed inward from the face opposite node i . This choice of function has the important properties that:

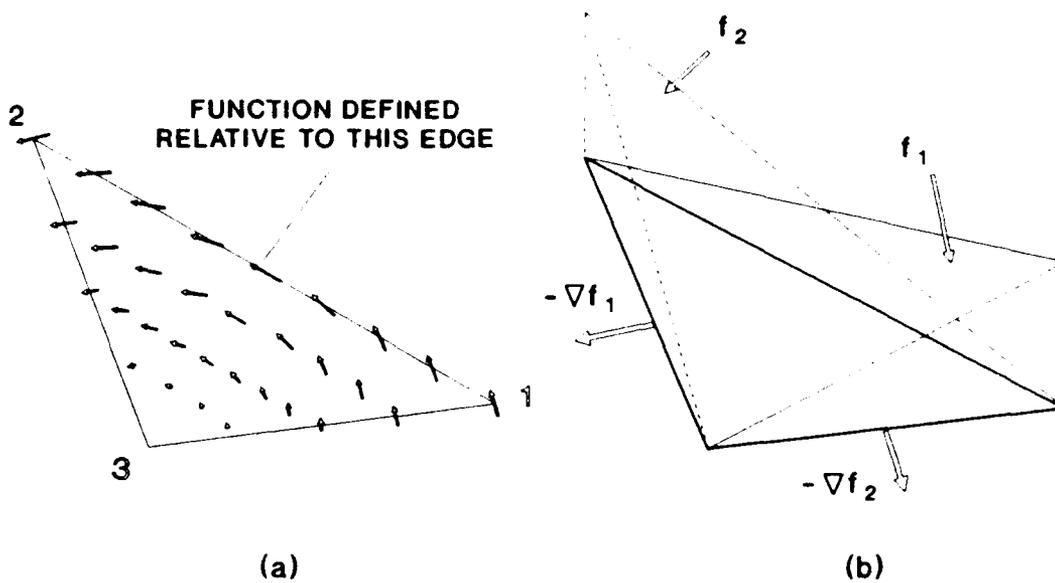


Figure 6. A Two-Dimensional Vector Expansion Function: (a) direction (arrow length indicates function magnitude); (b) constituent linear finite elements

(a) $\nabla \cdot \bar{\psi}_{ij} = 0$ so the divergence condition is automatically satisfied; and (b) the component along the edge $i-j$ has constant unit magnitude, ensuring tangential field continuity across element boundaries. One can represent the global expansion functions as being assembled from functions defined over individual cells:

$$\bar{\psi}_s = \sum_{k=1}^K \sum_{i=1}^4 \sum_{\substack{j=1 \\ j \neq i}}^4 \bar{\psi}_{ij}^{(k)} \delta(s; i, j, (k)) \quad (21)$$

where $\delta(s; i, j, (k))$ is a Kronecker delta equal to 1 when global edge s is the same as the edge connecting local nodes i and j within cell k . This completes the definition of the vector edge-based functions. Their use in the discretization of the functional is discussed next.

4.4. Discretization Using Galerkin's Method

Galerkin's method is the specialization of weighted residuals in which the trial functions are the

same as the expansion functions. Let $\bar{\psi}_r$ be the weighting function for edge r . Then substituting $\bar{W}^* = \bar{\psi}_r^* = \bar{\psi}_r$ for each $r=1, \dots, N$ into (1) will generate N simultaneous equations

$$F_r(\bar{E}) = \int_{\Omega} \left[\frac{1}{\mu_r} \nabla \times \bar{\psi}_r \cdot \nabla \times \bar{E} - k_0^2 \epsilon_r \bar{\psi}_r \cdot \bar{E} \right] dv - jk_0 \eta_0 \int_{\Gamma_r} \bar{\psi}_r \cdot \hat{n} \times \bar{H} ds = 0, \quad r=1, 2, \dots, N \quad (22)$$

(The subscript r on μ and ϵ still denotes *relative* permeability and permittivity.) Substituting the expansion for \bar{E} from Eq.(14) into the two terms of the volume integral:

$$F_{l,r} = \sum_{s=1}^N e_s \left[S_{rs}^{ll} + S_{rs}^{l2} \right] \quad (23)$$

$$S_{rs}^{ll} = \int_{\Omega} \frac{1}{\mu_r} \nabla \times \bar{\psi}_r \cdot \nabla \times \bar{\psi}_s dv \quad (24)$$

$$S_{rs}^{l2} = -k_0^2 \int_{\Omega} \epsilon_r \bar{\psi}_r \cdot \bar{\psi}_s dv \quad (25)$$

To simplify the surface integral terms, let Ψ_{ri} denote the inner product of the r 'th expansion function and the i 'th waveguide mode function:

$$\Psi_{ri} = \int_{\Gamma_A} \bar{\psi}_r \cdot \bar{g}_i ds \quad (26)$$

$$\Psi'_{ri} = \int_{\Gamma_B} \bar{\psi}_r \cdot \bar{h}_i ds \quad (27)$$

then from the first term of Eq.(9)

$$E_r^{\text{inc}} = 2jk_0 \eta_0 Y_0 \Psi_{r0} \quad (28)$$

and from the second term,

$$S_{rs}^A = jk_0\eta_0 \sum_{i=0}^{\infty} Y_i \Psi_{ri} \Psi_{si} \quad (29)$$

and similarly, from Eq.(10)

$$S_{rs}^B = jk_0\eta_0 \sum_{i=0}^{\infty} Y'_i \Psi'_{ri} \Psi'_{si} \quad (30)$$

Collecting the terms from Eqs.(24)-(30) gives the matrix equation (12). The next two sections give the details for performing the integrations in Eqs.(24)-(27).

4.5. Volume Integral Computations

The volume integral computations are carried out by visiting each cell once and adding its contribution to S_{rs} for each combination of edges shared by the cell, excluding those edges that are on perfect conductors. Let i,j and m,n be the local node indices defining edges r and s , respectively, $1 \leq i,j,m,n \leq 4$, $i \neq j$, $m \neq n$. After carrying out the coordinate transformation local to cell k , giving the matrix T :

$$f_i = \frac{1}{6V} [T_{i1} + xT_{i2} + yT_{i3} + zT_{i4}] \quad (31)$$

$$\nabla f_i = \frac{1}{6V} [\hat{x}T_{i2} + \hat{y}T_{i3} + \hat{z}T_{i4}] \quad (32)$$

Using the identity $\nabla \times (a\nabla b) = a\nabla \times \nabla b + \nabla a \times \nabla b$ and noting that the second derivatives of the linear functions f are all zero,

$$\begin{aligned} \nabla \times \bar{\Psi}_s &= 2L_s \nabla f_i \times \nabla f_j \\ \nabla \times \bar{\Psi}_r &= 2L_r \nabla f_m \times \nabla f_n \end{aligned} \quad (33)$$

The terms in Eq.(33) are all constants and may be taken outside the integral in Eq.(24).

$$\begin{aligned}
S_{rsk}^{I1} &= \frac{4V}{\mu_r} L_r L_s \nabla f_i \times \nabla f_j \cdot \nabla f_m \times \nabla f_n \\
&= \frac{4V L_r L_s}{\mu_r (6V)^4} \left[(T_{i3} T_{j4} - T_{i4} T_{j3})(T_{m3} T_{n4} - T_{m4} T_{n3}) \right. \\
&\quad + (T_{i4} T_{j2} - T_{i2} T_{j4})(T_{m4} T_{n2} - T_{m2} T_{n4}) \\
&\quad \left. + (T_{i2} T_{j3} - T_{i3} T_{j2})(T_{m2} T_{n3} - T_{m3} T_{n2}) \right]
\end{aligned} \tag{34}$$

The second volume integral's contribution from cell k is

$$\begin{aligned}
S_{rsk}^{I2} &= -k_0^2 \epsilon_r L_r L_s \int_{\Omega_k} \left[f_i f_m \nabla f_j \cdot \nabla f_n - f_j f_m \nabla f_i \cdot \nabla f_n \right. \\
&\quad \left. - f_i f_n \nabla f_j \cdot \nabla f_m + f_j f_n \nabla f_i \cdot \nabla f_m \right] dv
\end{aligned} \tag{35}$$

After breaking this integral into four parts, the gradient terms are constants that may be brought outside the integrals, which may then be evaluated using Eq.(19). The result is

$$\begin{aligned}
S_{rsk}^{I2} &= -\frac{k_0^2 \epsilon_r L_r L_s}{720V} \sum_{l=2}^4 \left[(1 + \delta_{im}) T_{jl} T_{nl} - (1 + \delta_{jm}) T_{il} T_{nl} \right. \\
&\quad \left. - (1 + \delta_{in}) T_{jl} T_{ml} + (1 + \delta_{jn}) T_{il} T_{ml} \right]
\end{aligned} \tag{36}$$

In Eqs.(34) and (36), the volume integral has been evaluated in closed form and written as an algebraic expression in terms of the geometry of a cell and the constitutive parameters contained within it. The following section derives similar expressions for the matrices representing the waveguide apertures.

4.6. Boundary Integral Computations

The only difficult computation involved in the boundary integrals Eqs.(26)-(28) are the inner products Ψ_{si} . When the functions g_i are circular or coaxial mode functions the integrals may not be evaluated in closed form, so they must be evaluated numerically.

As was the case with the volume integrals, the contributions to the surface integrals are carried

out one cell at a time. The two-dimensional simplex coordinates are used because the 2D analog of the vector expansion function evaluated over a triangular tetrahedron face is the same as the 3D function evaluated in that face. Hence the matrix T gives the coordinate transformation local to a triangle and the scalar expansion functions are

$$f_i = \frac{1}{2A} [T_{i1} + xT_{i2} + yT_{i3}] \quad (37)$$

where A is the triangle's area. The vector functions are still as defined by Eq.(20), except that the index i may only take on values 1,2 and 3.

The contribution of a triangle to Ψ_{si} will be approximated using Gaussian quadrature integration. Quadrature formulas only apply strictly to one dimension, but are easily extended to two dimensional integration over rectangular areas. To use these, the triangle's geometry is transformed to a unit square, using an approach suggested by Stroud & Secrest [9]. Note that the transformation to simplex coordinates mapped the arbitrary triangle into one whose vertex coordinates are (0,0), (0,1) and (1,0) in (t_1, t_2) coordinates. The second transformation is given by

$$u_1 = t_1$$

$$u_2 = \begin{cases} \frac{t_2}{(1-t_1)}, & t_1 \neq 1 \\ 1, & t_1 = 1 \end{cases} \quad (38)$$

The Jacobian of this transformation is $(1-t_1)^{-1}$, so that a typical integral term transforms as follows:

$$I_1 = \int_0^1 t_1 dt_1 \int_0^{1-t_1} \bar{g}(t_1, t_2) dt_2 = \int_0^1 u_1 du \int_0^1 (1-u_2)^2 \bar{g}(u_1, u_2) du_2 \quad (39)$$

Let u_k and u_m denote the one-dimensional quadrature sample points along u_1 and u_2 , respectively, with w_k and w_m the corresponding weights. Then the integral is approximated by the sum

$$I_1 \approx \sum_{k=1}^Q w_k u_k \sum_{m=1}^Q w_m (1-u_m)^2 \bar{g}(u_k, u_m) \quad (40)$$

where Q is the order of the quadrature formula. See, for example, Reference 9 or Reference 10, p.887 for tables of weights and quadrature points.

Finally, once $\Psi_{s,i}$ has been found for each edge s in Γ_A for mode i , then that mode's contribution to S^A may be computed using Eq.(27). The computation of S^B is identical. For rectangular waveguides the integrations have been carried out analytically and appropriate formulas are given in Appendix C.

The incident field term computation is also the same, except that it involves only the dominant mode term, $\Psi_{s,0}$. It is also convenient to save $\Psi_{s,0}$ and $\Psi'_{s,0}$ for use later in finding the reflection and transmission coefficients. Solving the matrix gives the edge field coefficients e_s , from which the reflection coefficient in waveguide A and transmission coefficient in waveguide B are:

$$\rho_A = \sum_{s=1}^N e_s \Psi_{s,0} - 1 \quad (41)$$

$$\tau_B = \sqrt{Y'_0/Y_0} \sum_{s=1}^N e_s \Psi'_{s,0} \quad (42)$$

or S_{11} and S_{21} , respectively.

This completes the derivations for the expressions needed to perform the matrix fill computations. The next section will discuss how they are implemented algorithmically in the computer code.

4.7. Matrix Structure and Solution Methods

A few notes about the structure of the matrix S are important: (a) S_{rs}^{11} and S_{rs}^{12} are zero unless

edges r and s are shared by one or more cells, so they are sparse matrices; (b) Ψ_{ri} and Ψ'_{ri} are zero unless edge r is in Γ_A or Γ_B respectively, so that the entries in S^A and S^B are zero except when edges r and s are both in the same aperture; and (c) the incident field elements are zero except for those edges r in Γ_A . Suppose that the edges in the mesh are indexed sequentially starting with those in aperture A, continuing by increasing z coordinate, and ending with those in aperture B. Then: the matrix S^I will be banded; S^A will be zero except for a dense submatrix in the upper left corner of size $N_A \times N_A$ (N_A is the number of edges in aperture A); S^B will be zero except for a dense submatrix in the lower right corner of size $N_B \times N_B$ (N_B is the number of edges in aperture B); and only the first N_A elements of E^{inc} are nonzero.

Figure 7 shows the structure of a typical matrix, with black points indicating nonzero entries. Typical of most problems, the matrix is banded, but its bandwidth is quite large. Consequently, there is little to be gained by using a band storage or skyline storage format, since the number of zeros inside the band and/or profile is still quite large. With a "direct" solver such as LU decomposition, the amount of storage needed may be the limiting factor rather than the time needed to perform the decomposition. Even though most of the entries in S are zero, the same is not true of its L (lower) and U (upper) factors since there is a considerable amount of "fill in" during the factorization [11]. The problem that gave rise to Figure 7 had 964 unknowns. The matrix had 40,462 entries (4.4 percent full) while the LU factors had 348,180 entries (37.5 percent full). Heath et al. review the necessary steps required to efficiently use LU decomposition on this kind of matrix [12]: (a) execute a reordering algorithm that swaps rows and columns in an attempt to minimize the fill in during the subsequent factorization; (b) perform symbolic LU factorization to determine where the fill in will occur so that memory may be allocated for the new matrix elements; and (c) perform the factorization. But the reordering algorithms only lessen the amount of fill-in; they do not guarantee that it will be less than any percentage of the original matrix size. Hence, it is not generally possible to predict how much storage will be required for

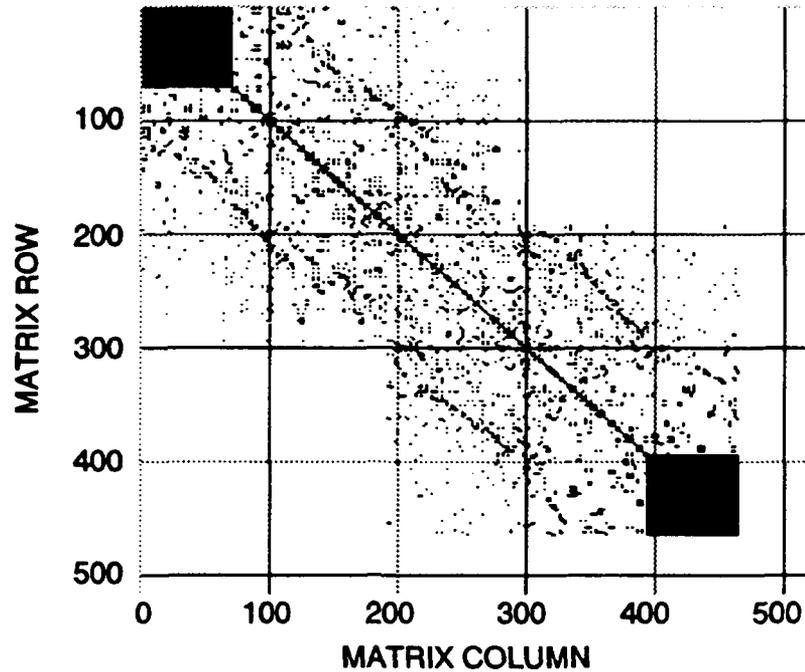


Figure 7. Sample Matrix Structure (dark points are nonzero elements)

a given problem aside from the upper bound of N^2 words.

To be able to handle larger problems, we chose to use the conjugate gradient algorithm. It uses a minimum of storage because it always works with the original matrix and does not generate any factors, and therefore there is no fill-in. The essential algorithm is given by Sarkar and Arvas [13]. The procedure to solve $AX = B$ for the unknown vector X begins with an initial guess X_0 , giving an initial residual error vector $R_0 = B - AX_0$. The initial search direction and gradient vectors P_0 and G_0 are

$$P_0 = G_0 = A^H R_0 \quad (43)$$

(The superscript H denotes Hermitian, that is, conjugate transpose.) The iteration consists of the following steps using $\|\cdot\|$ to denote the L_2 norm:

$$\alpha_i = \frac{\|G_i\|^2}{\|A P_i\|^2} \quad (44)$$

$$X_{i+1} = X_i + \alpha_i P_i \quad (45)$$

$$R_{i+1} = R_i - \alpha_i A P_i \quad (46)$$

$$G_{i+1} = A^H R_{i+1} \quad (47)$$

$$\beta_i = \frac{\|G_{i+1}\|^2}{\|G_i\|^2} \quad (48)$$

$$P_{i+1} = G_{i+1} + \beta_i P_i \quad (49)$$

A typical convergence test is to examine the ratio $\epsilon = (\|R_i\|^2 / \|R_0\|^2)$ at each iteration. A value of 10^{-4} or 10^{-5} is usually adequate to ensure convergence of the S parameters. This is demonstrated in Figure 8, which shows the value of both $|S_{11}|$ and ϵ for a test problem consisting of a microstrip circuit with coaxial inlet and outlet waveguides. The reflection coefficient magnitude has converged to one part in 1000 when the residual norm is 0.001 of its initial value (using $X_0=0$ as the initial guess). It is also important to note that the residual norm decreases monotonically, which is another important attribute for a good convergence criterion. The same is not true of the S parameters themselves which sometimes oscillate about their final values. For this particular problem the convergence was relatively slow, taking about $6N$ iterations, where N was the number of unknowns. N iterations is more typical, and this could probably be improved substantially using preconditioning.

5. COMPUTER CODE DESIGN

The computer program implementing the solution procedure is intended not only to demonstrate the validity and potential usefulness of the method, but also to serve as a design tool for a certain class of RF devices. As such, it must be *effective, reliable, efficient* and *versatile*. In other words, it must consistently provide accurate results for a reasonable cost for a wide range of problems. The first

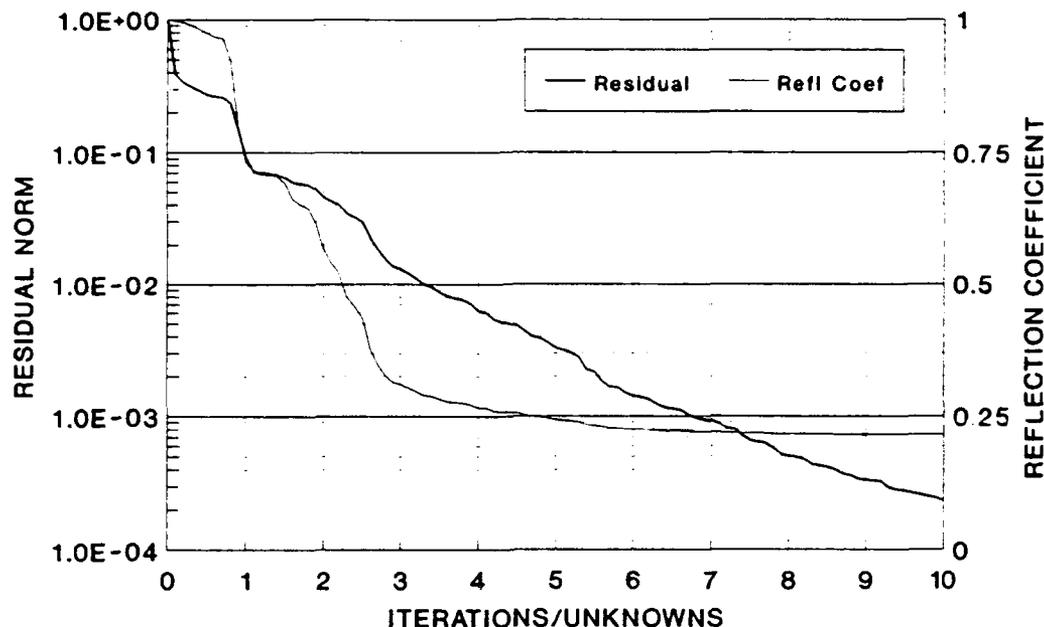


Figure 8. Convergence of Residual Norm and Reflection Coefficient using Conjugate Gradient Method

criterion, effectiveness, will be established by the validation cases to be discussed in Section 6. It is the result of a faithful implementation of the formulas from the preceding sections. The second criterion, efficiency, is partly a consequence of inherent properties of the solution method, but also results from design of algorithms to minimize storage and computation time. Those algorithms are discussed in Sections 5.1 and 5.2. The last criterion, versatility, is also partly a consequence of the method, but it cannot be realized without a means of defining complex geometries. That problem was solved by interfacing the code to commercial computer-aided design software, as discussed in Section 5.3.

5.1. Computational Algorithms

The overall design for the computer program TWOPORT is shown in Figure 9, most of which is self-explanatory. The main loop may be executed repeatedly for increments in frequency. The major

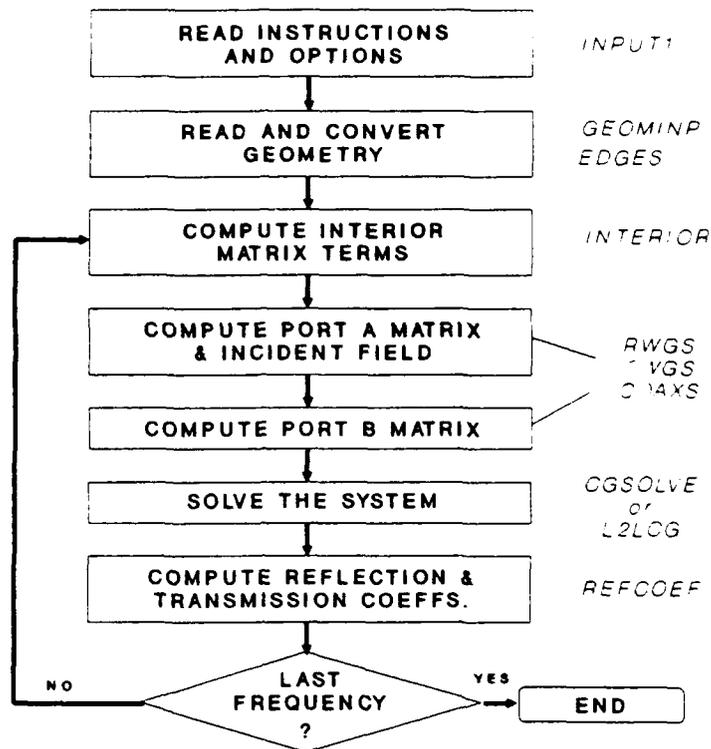


Figure 9. Flow Diagram for Program TWOPORT

subroutines involved in each step are annotated on the right. (The program is written in ANSI standard FORTRAN and has been executed on a VAX® 8650 minicomputer and a Cray Y-MP®.)

The serious work begins in the second block, which imports a geometry description that was produced by the CAD software. That description, illustrated in Figure 10, consists of a tetrahedron mesh defined by: (a) a list of vertices with their x,y,z coordinates, and flags that indicate whether they are on an aperture and/or on one or more conducting surfaces (allowing for the possibility that a node may be at the intersection of conducting surfaces); (b) a list of tetrahedra with the four nodes defining each and the index of the material filling it; and (c) a list of the materials with their permittivity and permeability. The geometry conversion is an additional step that is required to accommodate the edge-based vector expansion functions: (a) create a list of edges with their terminal node indices; and (b) create a new list

<i>Number of Nodes, Elements & Materials</i>									
181	506	2							
<i>Node Index</i>	<i>x</i>	<i>Node Coordinates</i>			<i>Port #</i>	<i>Conductor Surfaces</i>			
		<i>y</i>	<i>z</i>						
1	-0.17971E-02	0.10008E-02	-0.20000E-02	1	1	0	3	0	0
2	-0.20570E-02	0.00000E+00	-0.20000E-02	1	0	0	3	0	0
3	-0.17974E-02	-0.10003E-02	-0.20000E-02	1	1	0	3	0	0
4	0.17974E-02	-0.10003E-02	-0.20000E-02	1	1	0	3	0	0
		.							
		.							
178	-0.44520E-03	0.45279E-03	0.20000E-02	2	0	0	0	4	0
179	0.89846E-03	0.10000E-02	0.20000E-02	2	1	0	0	0	0
180	-0.11642E-09	0.10000E-02	0.20000E-02	2	1	0	0	0	0
181	-0.89846E-03	0.10000E-02	0.20000E-02	2	1	0	0	0	0
<i>Element Index</i>	<i>Node Indices</i>				<i>Material Index</i>				
1	37	27	53	31	2				
2	37	27	33	53	2				
3	37	28	55	32	1				
		.							
		.							
504	14	36	46	9	1				
505	14	48	60	46	1				
506	14	9	46	60	1				
<i>Material Index</i>	<i>Permittivity</i>			<i>Permeability</i>					
	<i>Real</i>	<i>Imag</i>		<i>Real</i>	<i>Imag</i>				
1	0.10000E+01	0.00000E+00		0.10000E+01	0.00000E+00				
2	0.12900E+02	0.00000E+00		0.10000E+01	0.00000E+00				

Figure 10. Example Input Geometry File (as Created by Translator Program from CAD Universal File; italics are annotation)

of tetrahedra with the six edges that define each one. Note that during this process, if two connected nodes are on the same conductor, then that edge is not included in the edge list, so it will be ignored in subsequent calculations. This mechanism forces the field along those edges to be zero, thereby enforcing conductor boundary conditions. However, when two nodes are on different conductors, there may be a field component along the edge linking them, and so that edge must be accounted for. That is the

reason why the geometry model must identify which specific conductors a node is on, and not simply whether or not it is on one.

The interior matrix computations are Eqs.(34) and (36). These are performed one cell at a time, as depicted in Figure 11.

When using row-column storage in the LUD version of the program, the waveguide contributions are not calculated as separate matrices as implied by Figure 9. They are merely added to the correct locations of the matrix that already contains the interior (FEM) results. This conserves storage by requiring only one matrix instead of three. As Figure 12 shows, these calculations are carried out one mode at a time for $mn=10,11,\dots$. For each mode the inner products Ψ_{smn} are generated for each edge and each polarization p,q . Then, in the last block, the matrix elements are updated for each pair r,s in the aperture. An important detail that is not reflected in Figure 10 is that each Ψ_{r0} (the dominant mode) is saved for later use in computing the reflection and transmission coefficients. A single array is used for both since edges in ports 1 and 2 are mutually exclusive, so their nonzero entries do not overlap.

5.2. Sparse Storage Structure and Conjugate Gradient Algorithm

The first version of program TWOPORT used an ordinary row-column storage format. Relatively small-scale problems can thus be solved using LU decomposition by standard library routines. It also allowed us to gain confidence in the code without the added uncertainty of matrix solver convergence. The LUD version is still recommended for problems having 1000 edges or less.

An existing CGM solver [14] was used in conjunction with the row-column storage to gather convergence data for some small test problems. However, taking full advantage of the CGM requires using a sparse storage format tailored to the form of matrix generated by this code. The structure that was adopted uses three separate arrays: one for the interior (sparse) matrix generated by the finite element computations; and one each for the waveguide apertures, which are dense. The dense matrices

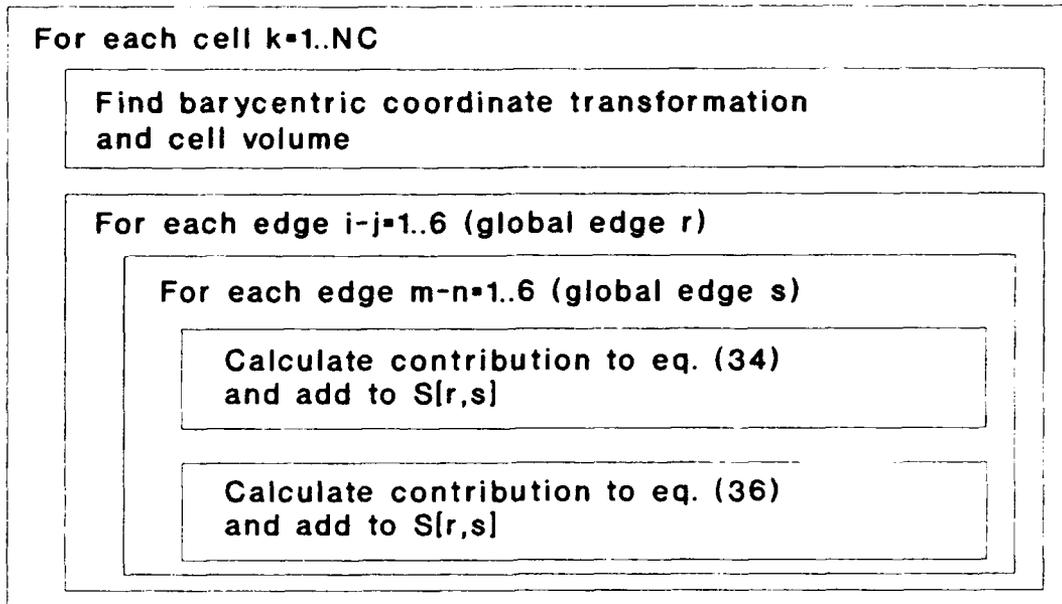


Figure 11. Structure of Finite Element Interior Computations

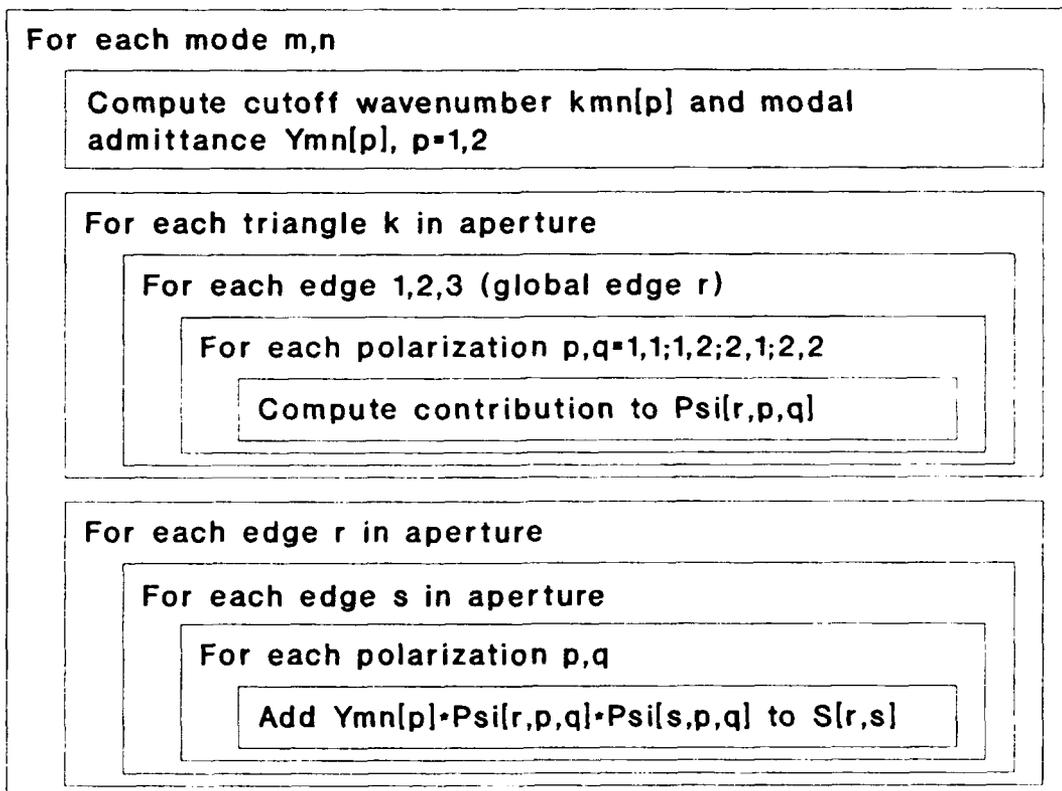


Figure 12. Structure for Waveguide Computations

are stored in row-column form.

The interior matrix entries are stored in a column array in the order in which they are created. That order depends only on the sequence in which edge pairs are encountered. For instance, if cell #1 includes edges 3,4, and 7 (although a tetrahedron has six edges, those that are on perfect conductors are not included in the edge list, and so a cell may have fewer than 6 numbered edges), then the first entries to be created are S_{33} , S_{34} , S_{37} , S_{43} , S_{44} , S_{47} , S_{73} , S_{74} , S_{77} , and these are stored in A(1) - A(9). If any of these pairs are shared by other cells, then those cells' contributions are added to the existing entries in A. Two integer arrays IA and JA keep track of the row and column locations, respectively, of the entries in A, that is: IA=[3,3,3,4,4,4,7,7,7,...] and JA=[3,4,7,3,4,7,3,4,7,...]. To aid in searching for existing entries, two arrays are maintained to point to the first and last entries for each row, for example: NF(3) = 1, NL(3)=3, NF(4)=4, NL(4)=6, etc. Since calculations for subsequent cells will involve searching A to determine whether a given edge pair already exists, these arrays NF and NL speed the process by bounding the search.

Implementing Eqs.(44)-(49) requires approximately $2MN$ multiplies, where N is the number of edges and M is the average edge connectivity (MN is the total number of nonzero entries in A). M is independent of the problem size, and is usually around 10. If convergence is obtained in N iterations, then the execution time is $O(2MN^2)$ operations. On the other hand, LU decomposition requires $O(N^3)$ operations.

The algorithm as programmed is essentially a straightforward implementation of Eqs.(43)-(49). The initial guess for the solution is the incident field for the first frequency. For subsequent frequencies, the converged result from the previous frequency is used as the initial X_0 .

5.3. I-DEASTM Interface

The commercial software I-DEASTM is a system of programs oriented toward structural analysis

[15]. Only a few of its components are used in this work, most notably, the solid object modeling and mesh generation. The "solid object" is the region denoted Ω . It is the entire cavity, excluding the interiors of conducting obstacles. For one test case the cavity was simply a short section of coaxial line, for which the solid object is the tube shown in Figure 13a.

Mesh generation is an automated process that creates nodes on the boundaries and in the interior of the solid object so that the distance between any node to its nearest neighbor is less than or equal to a user-specified global element edge length. Then the nodes are grouped into tetrahedral elements bounded by groups of four nodes. The result is shown in Figure 13b.

There are a few preliminary steps before the automatic mesh generation may be executed. First, points and curves are created from the solid object in the "geometry definition" task. Then the curves are linked together to form mesh areas in the "mesh generation" task. Last, the mesh areas are assembled into mesh volumes.

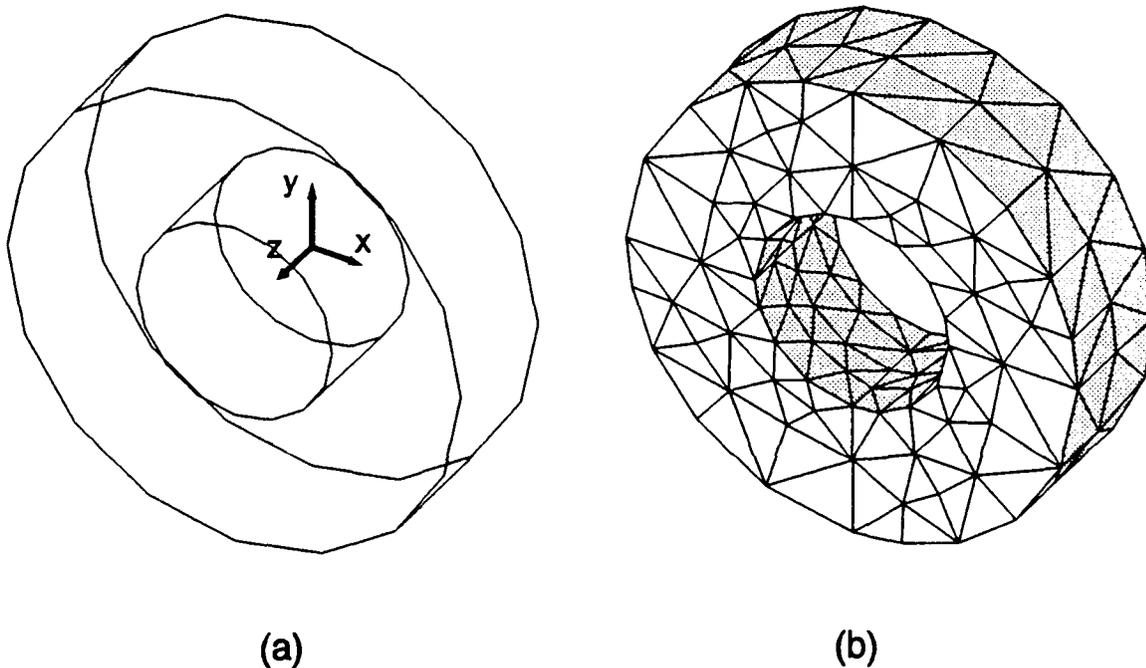


Figure 13. Coaxial Waveguide Section: (a) Solid Object; (b) Tetrahedron Mesh

The electromagnetic analysis code will need to have nodes identified by: (a) the port number, if they are in an aperture; and (b) conducting surface numbers(s) if applicable. The port identifier is conveyed through a boundary condition variable: the z component of "nodal force," set to 1 or 2 for waveguide A or B, respectively. The translator program (TRANSL8) will examine that variable in the I-DEAS universal file to determine which nodes are waveguide aperture nodes. The conducting surfaces are defined by grouping nodes into "permanent groups." The translator program looks for permanent groups in the universal file and uses those to identify each node's inclusion in up to six conductor surfaces. Last, the material properties in the cells must be identified. I-DEAS, being oriented to mechanics, does not provide menu options for defining permittivity and permeability. However, the "physical properties" table has nine general-purpose data fields called a "keyopt array," and the first four of these are used for $\text{Re}(\epsilon_r)$, $\text{Im}(\epsilon_r)$, $\text{Re}(\mu_r)$ and $\text{Im}(\mu_r)$. Any number of dielectric regions may be defined by grouping elements and assigning them to different physical properties tables.

The last step inside I-DEAS is to write the universal file. That file contains a great deal of information not needed by the electromagnetics code. The translator program filters out the superfluous information and writes an output file in a form (see Figure 10) that is easier to inspect and modify.

6. VALIDATION AND DEMONSTRATION

This section presents test results for the computer program TWOPORT. The tests were intended to serve three purposes: first, to verify that the code produces correct results; second, to show that it is useful for problems that have practical applications; and third, to assess the modeling requirements such as mesh density and number of waveguide modes.

6.1. Waveguides and Waveguide Discontinuities

6.1.1. WAVEGUIDE PROPAGATION.

Referring again to Figure 1, suppose that the inlet and outlet waveguides are identical in size and shape, and that the "cavity" is another section of the same waveguide with length d . Then the magnitudes of S_{11} and S_{21} should be 0.0 and 1.0, respectively, since there are no discontinuities or lossy materials. Furthermore, the phase of S_{21} should be $360^\circ d/\lambda_g$, where λ_g is the guide wavelength of the dominant mode.

The coaxial waveguide section shown earlier in Figure 13 has the dimensions of an APC-7mm connector ($a=1.502$ mm, $b=3.5$ mm) and it is 4 mm long. The guide wavelength is the same as the free space wavelength since there is no dielectric and the dominant mode is TEM. Calculations by TWO-PORT over a range of frequencies are shown in Figure 14. The agreement with the theory is quite good,

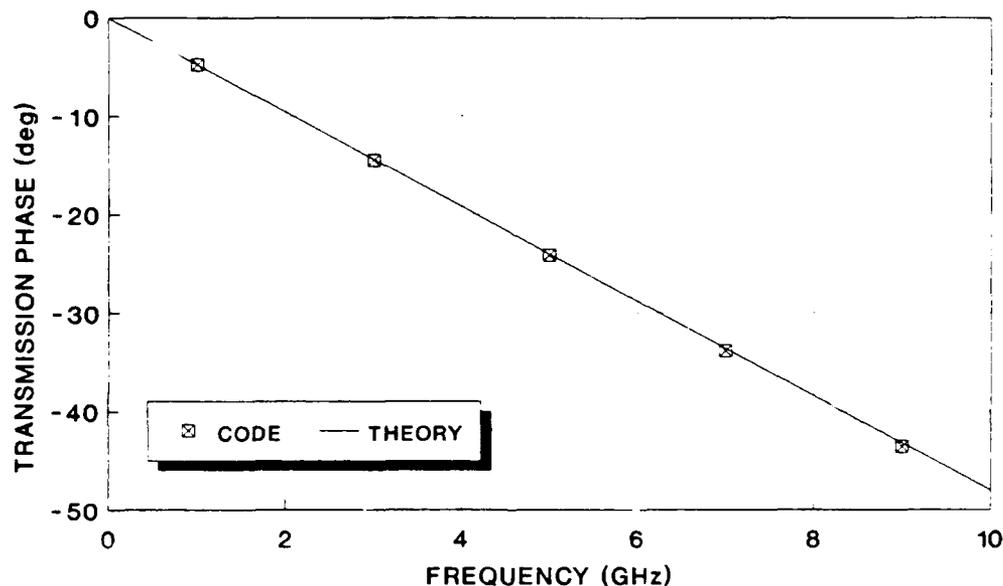


Figure 14. Propagation Phase Through 4 mm APC-7 Coaxial Waveguide Section

indicating that the finite element mesh does not introduce any dispersion. At all frequencies considered, $|S_{11}|$ and $|S_{21}|$ were within 1×10^{-4} of the correct values.

Figure 15 is the mesh for a section of standard X-band rectangular waveguide ($a=22.86$ mm, $b=10.16$ mm) with length $d=b/2=5.08$ mm. The S_{21} phase computed by TWOPORT, shown in Figure 16, agrees very well with the expected results for the TE_{10} mode given by

$$\angle S_{12} = -360^\circ \frac{d}{\lambda} \sqrt{1 - \left(\frac{f_c}{f}\right)^2} \quad (50)$$

For each frequency the magnitude of S_{11} was less than 0.01 and that of S_{21} was greater than 0.9999. The rectangular waveguide results are important because they demonstrate that the finite element implementation is correctly enforcing conductor boundary conditions at interior corners.

Figure 17 shows the mesh for a section of circular waveguide. The mesh generator placed the

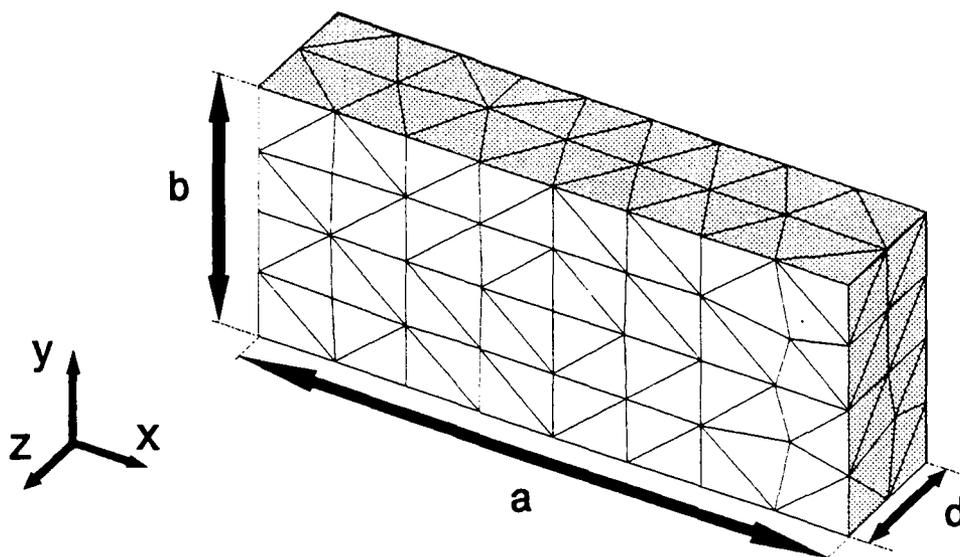


Figure 15. Finite Element Mesh for a Rectangular Waveguide Section
($a=22.86$ mm, $b=10.16$ mm, $d=5.08$ mm)

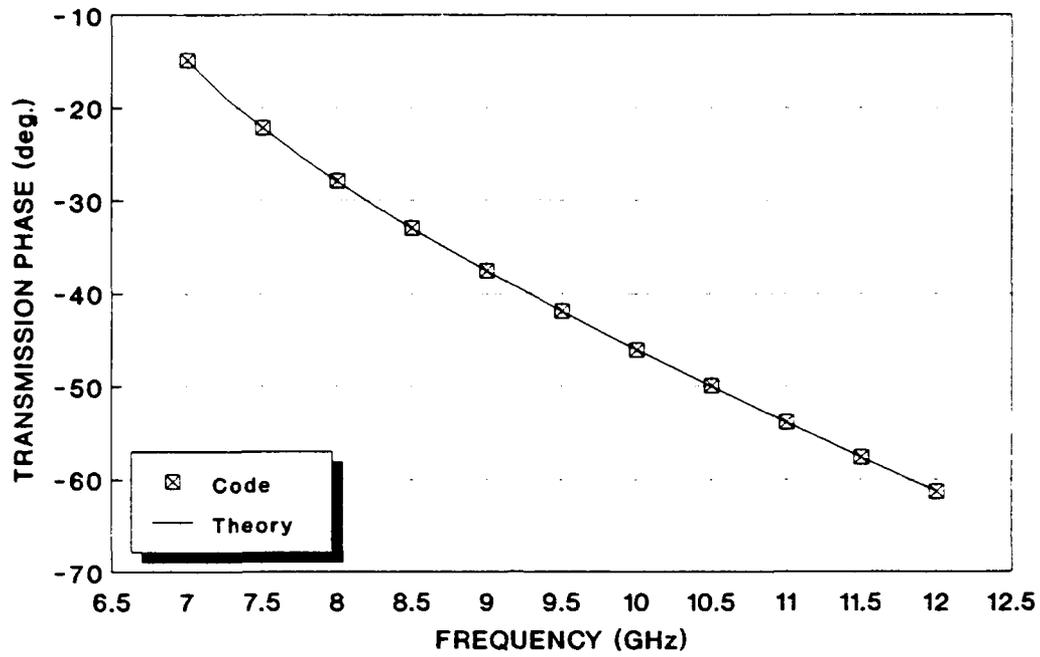


Figure 16. Propagation Phase Delay, Rectangular Waveguide Section

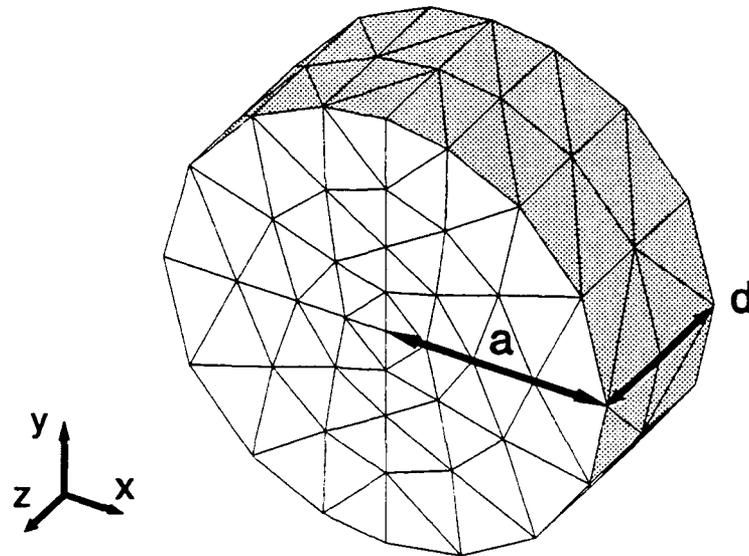


Figure 17. Finite Element Mesh for a Circular Waveguide Section
(radius=0.5 m, length=0.4 m)

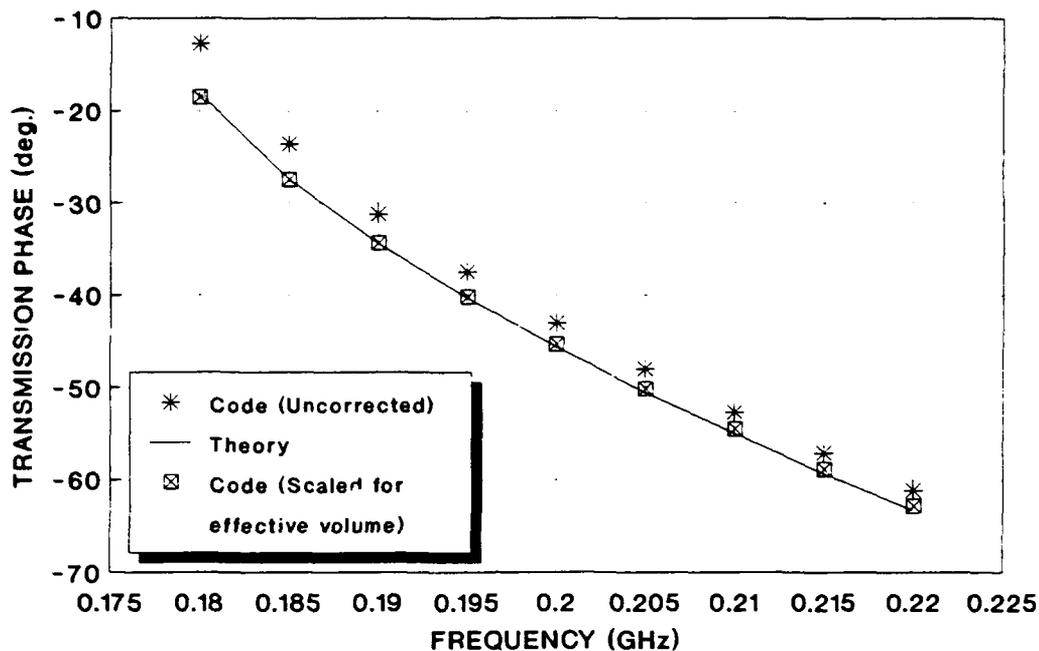


Figure 18. Propagation Phase Through Circular Waveguide Section (single mode propagation range is 176-231 MHz)

nodes on the surface of the cylinder, and consequently the collection of tetrahedra contains slightly less total volume than the original cylinder. It is necessary to scale the mesh in x and y to correct for that modeling error. (A scaling option is incorporated in the translator program.) As Figure 18 shows, the calculated and predicted phase of the dominant TE_{11} mode agree quite well when this volume correction is used.

6.1.2. WAVEGUIDE IRISES.

Thin metallic obstructions oriented transverse to the waveguide axis have applications as elements of microwave circuits, especially filters [16]. Our interest in them lies in their tendency to generate higher-order waveguide modes, making them useful as test cases for validating the aperture interactions through those higher order modes.

The first two iris test cases used a short section of APC-7 coaxial waveguide, just as in Figure 13, except that now part of the inlet aperture will be blocked. Network analyzer measurements were

made with foil obstructions placed between coaxial adapters, blocking either one quarter (a quadrant) or one half of their common aperture. Figures 19a and 19b show comparisons between measured and calculated S_{11} , both magnitude and phase, for these test cases.

The third iris test case was a diagonal obstruction in X-band rectangular waveguide. The finite element mesh model is shown in Figure 20a. Figure 20b shows the mesh with cells above $z=5$ mm removed, and shading added to identify the iris shape and location. Figure 21 is a comparison of the TWOPORT calculations and network analyzer measurements of S_{11} .

The close agreement between the measurements and calculations for these waveguide iris cases serves to validate the higher-order mode computations, which are fairly extensive, especially for the coaxial waveguide. They include calculations of: (a) cutoff wavelength; (b) mode functions; and (c) integration of inner products of mode functions and finite elements. The circular waveguide mode functions were validated by a step discontinuity problem, discussed in the following section.

6.1.3. CIRCULAR WAVEGUIDE MODE CONVERTER.

A step discontinuity in a circular waveguide is a fairly simple mode converter [17]. The outlet waveguide is slightly larger than the inlet waveguide so that it supports the TM_{11} mode as well as the TE_{11} mode. It is the abrupt change in radius that excites the TM_{11} mode.

Masterman and Clarricoats [17] give comparisons of multimode calculations against measured data for test cases in which the outlet waveguide has radius of 35.56 mm, and several different inlet waveguide radii. Their results are presented in terms of a "mode conversion ratio," M :

$$M = \frac{|D_2 h_{\rho 2}(\rho = a)|}{|D_1 h_{\rho 1}(\rho = a)|} \quad (51)$$

D_1 and D_2 are the excitation coefficients for the TE_{11} and TM_{11} modes, respectively, in the outlet waveguide, and $h_{\rho 1}$ and $h_{\rho 2}$ are the radial components of their orthonormal mode functions. Hence M is the relative strength of the two modes measured at the waveguide wall, $\rho = a$. The ratio of $h_{\rho 2}$ to $h_{\rho 1}$ is

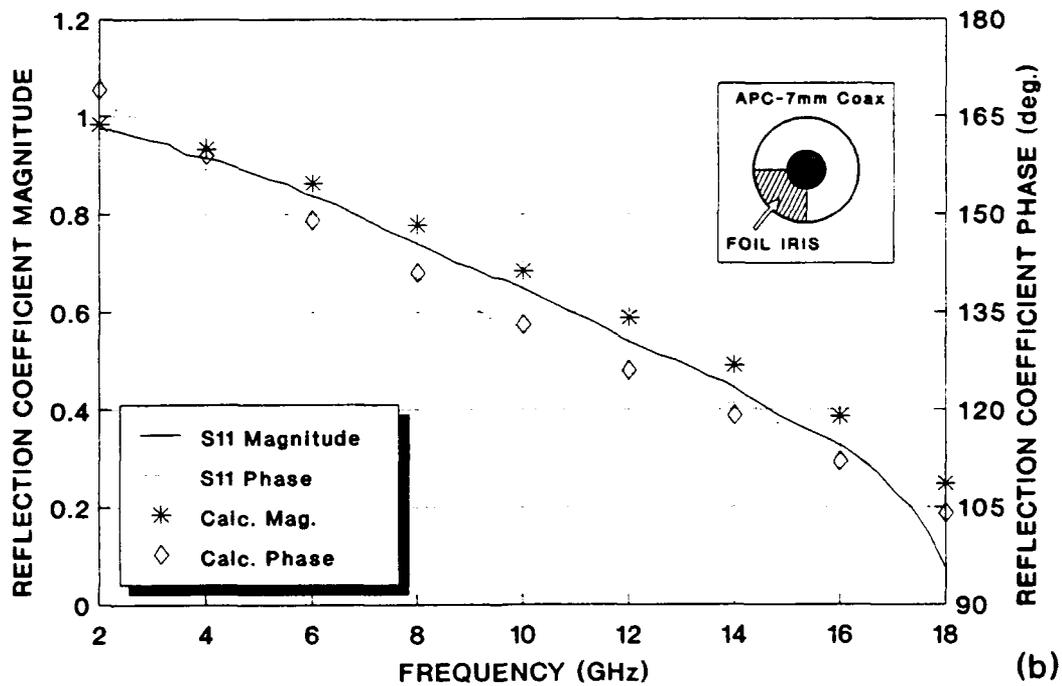
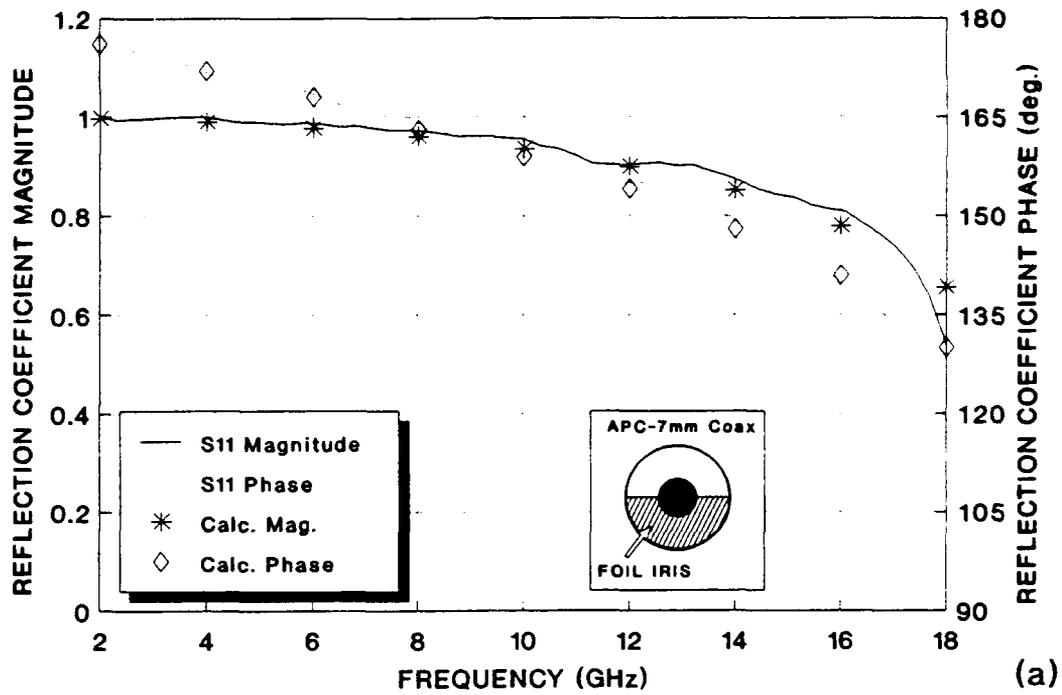


Figure 19. Measured and Calculated Reflection Coefficient vs. Frequency for Coaxial Irises (APC-7mm coax): (a) 180° Opening; (b) 270° Opening

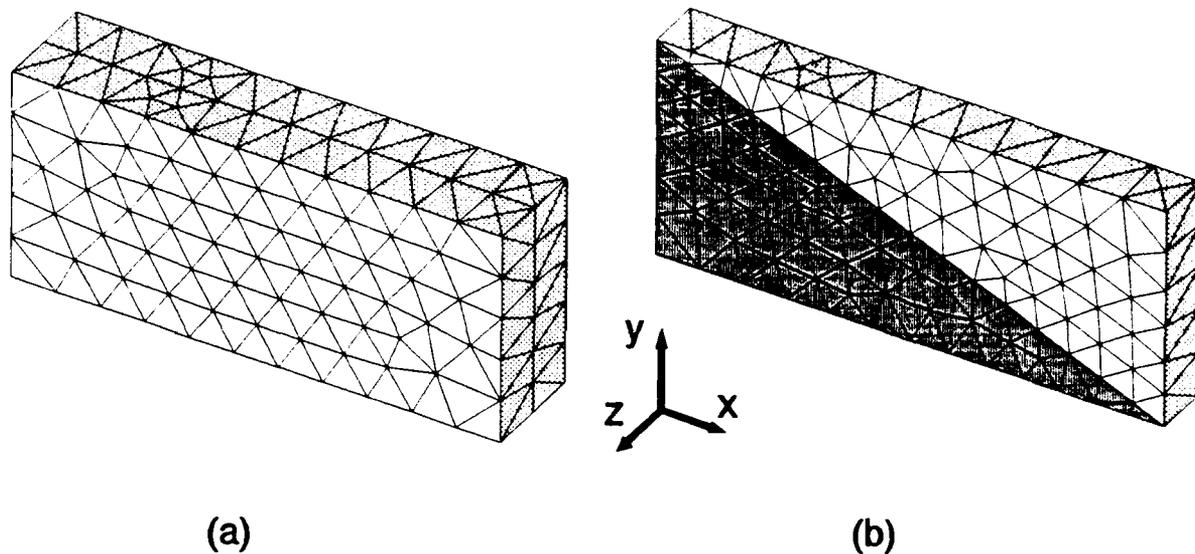


Figure 20. Finite Element Mesh for Diagonal Iris in Rectangular Waveguide: (a) Complete Mesh; (b) Cells above $z=5$ mm removed

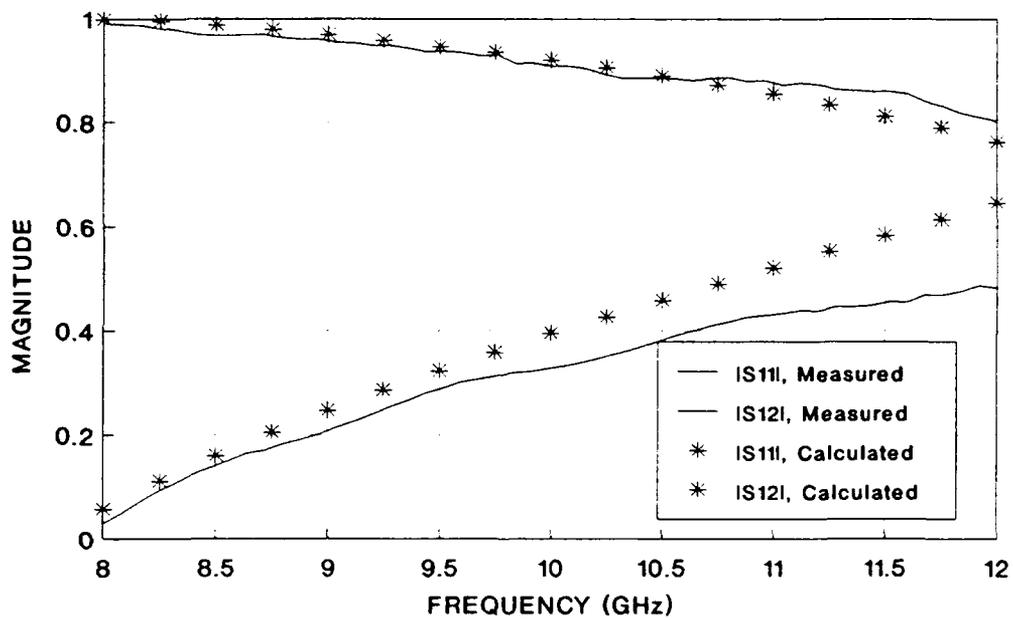


Figure 21. Measured and Calculated S_{11} for Diagonal Iris in Rectangular Waveguide (X-band waveguide)

independent of frequency, and is 3.79 dB for the 35.56 mm radius waveguide.

Figure 22 shows the finite element meshes used for two test cases, each with a different inlet waveguide radius: (a) 26.67 mm (1.05 in.); and (b) 29.21 mm (1.15 in.). The meshes produced by the CAD software were scaled appropriately to ensure that the included volume was correct. In both models, the total height in z is 10 mm, evenly divided between the two waveguide sections.

The excitation coefficients for the higher order modes are computed by the TWOPORT code, and from these it is a simple matter to find the mode conversion ratio. The results, shown in Figure 23, agree with the multimode calculations to within 1 dB over the entire frequency range. These results demonstrate a flexibility of our approach and the TWOPORT code for analyzing waveguide structures that may have many propagating modes.

6.2. Waveguide Transitions

Transitions between different types of waveguiding media are an important class of devices that can be modeled using the hybrid finite element method. The circular waveguide mode converter was a simple case. This section considers two structures that are somewhat more complex: a coax-stripline transition; and a coaxial end-wall launcher for rectangular waveguide.

6.2.1. COAX-STRIPLINE TRANSITION.

Figure 24a shows a short section of stripline (with air dielectric) enclosed in a rectangular conducting box. The openings at each end are the appropriate size for SMA coaxial adapters. The mesh geometry in Figure 24b shows only the cells below the stripline. Shading has been added to identify those cells that border on the strip conductor, the coax dielectric, and end of the coax center conductor, which is trimmed flat where it meets the stripline. The cross-sectional dimensions for both the coax and the stripline are given in Figure 24c. The stripline width was chosen for a characteristic impedance of 50Ω , using formulas from Gupta et. al. [Reference 18, pp. 57-60]. The computed reflection and trans-

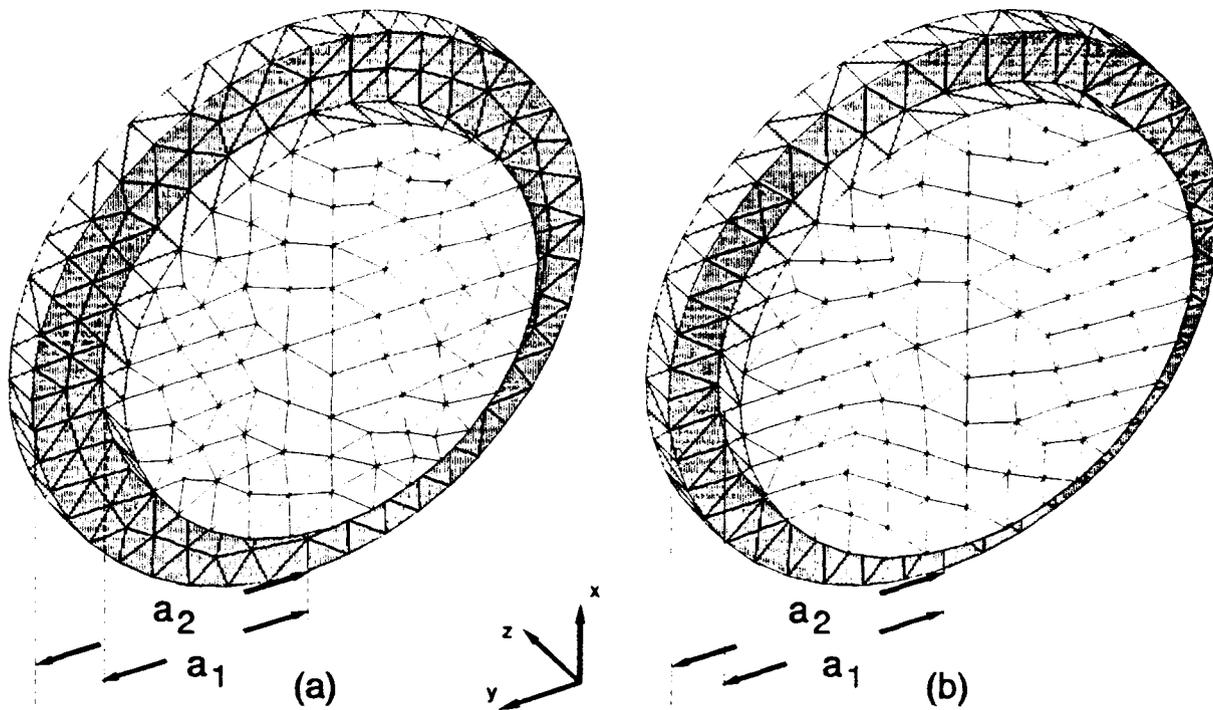


Figure 22. Finite Element Meshes for Circular Waveguide Step Discontinuities (Mode Converter), with $a_2 = 1.4$ in.: (a) $a_1 = 1.05$ in.; (b) $a_1 = 1.15$ in.

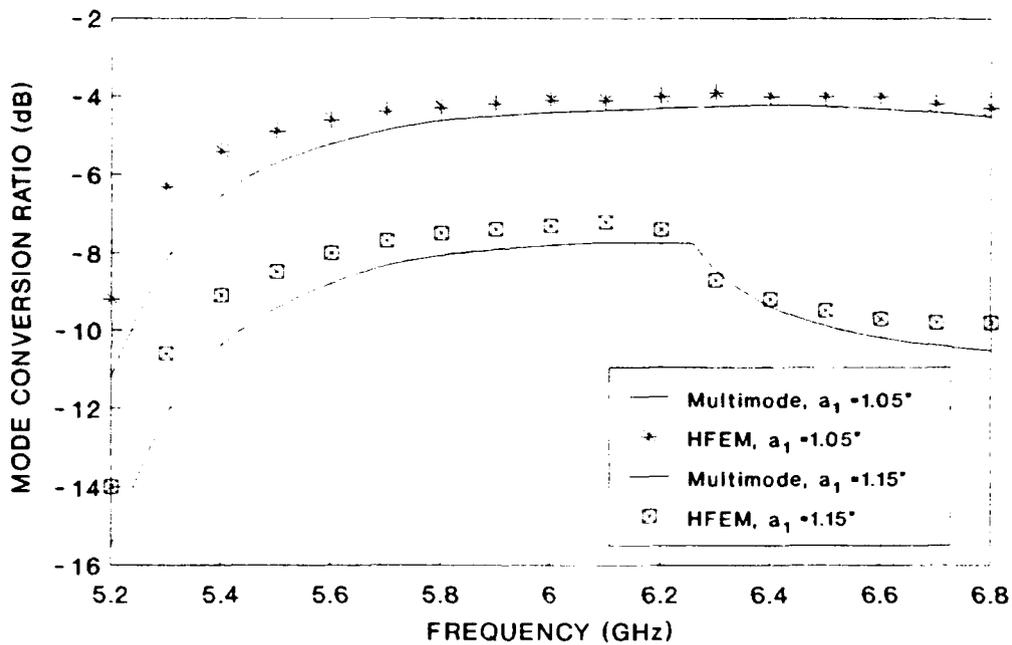
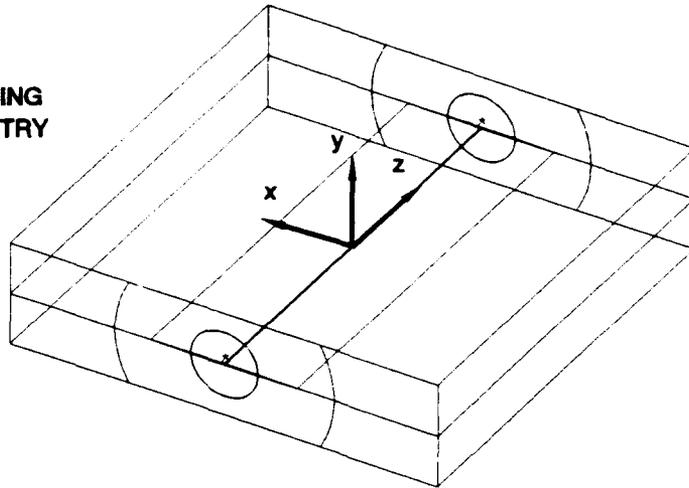
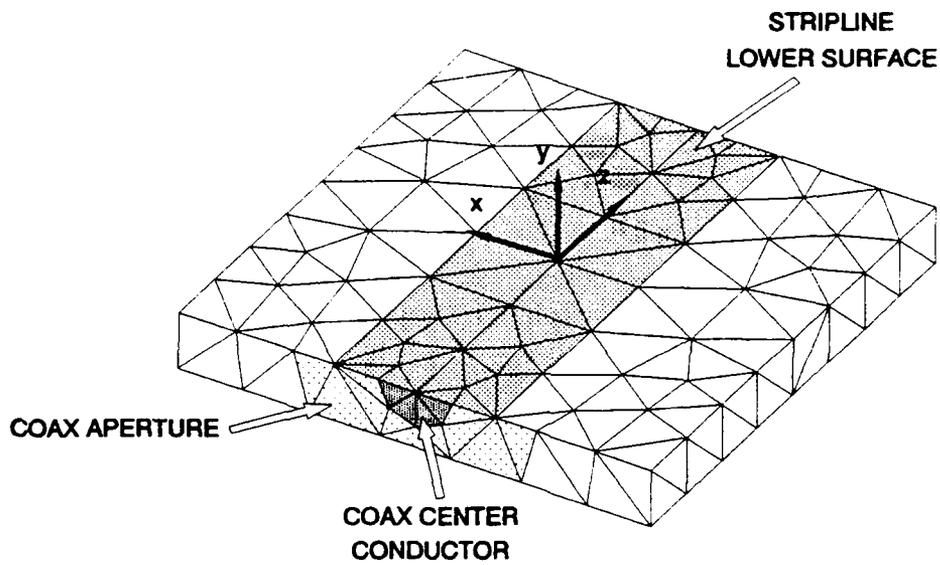


Figure 23. Mode Conversion Ratio for Circular Waveguide Mode Converters, Hybrid Finite Element (HFEM) and Multimode [17] Calculations

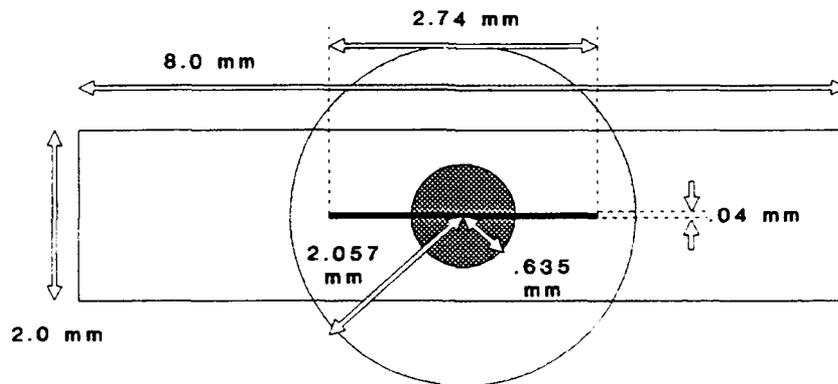
DEFINING
GEOMETRY



(a)



(b)



(c)

Figure 24. Stripline Test Case: (a) Complete Problem Geometry; (b) Lower Half of Finite Element Mesh; and (c) Dimensions at Coax/Stripline Junction

mission coefficients for a 4 mm long stripline are shown in Figure 25. As expected, the reflection coefficient is small at the lower frequencies because the coax and the stripline have the same characteristic impedance. But the junction's reactances become more significant with increasing frequency, as indicated by a gradual rise in $|S_{11}|$, until it decreases again approaching 18.75 GHz, where there is a resonant cancellation because the stripline is one quarter wavelength long at that frequency. These results are important because they demonstrate that conductor boundary conditions are correctly enforced at sharp conductor edges and at multi-conductor junctions.

6.2.2. COAX-RECTANGULAR WAVEGUIDE LAUNCHER.

A design described by Tang & Wong [19] for an end-wall transition between coax and rectangular waveguide is shown in Figure 26a. Their design used X-band rectangular waveguide (10 mm x 23 mm). The coax inner and outer radii were 1 mm and 2.3 mm, respectively. The coax port was centered in the y-direction (the waveguide height) but it was slightly off center in the x direction, by 1/10th the waveguide width. The 2 mm-thick post was centered 8.4 mm from the end wall. The probe (an extension of the coax center conductor) was 9.4 mm long overall. For convenience of modeling, the round conductors are modeled as eight-sided polygons and the corner between the probe and post was sliced at a 45°.

Figure 26b is a cutaway view of the finite element mesh for this object. It was fairly challenging in terms of the geometry generation requirements, and the success is attributable to the flexibility of the CAD software. The Figure illustrates that the interiors of solid conductors are represented as voids in the finite element mesh. As usual, the nodes that lie on those surfaces were appropriately tagged to identify them for the electromagnetic analysis code.

Figure 27 is the calculated reflection coefficient vs. frequency for this launcher design. As expected, it performs well (less than 2:1 VSWR) over most of the X-band (8 - 12 GHz).

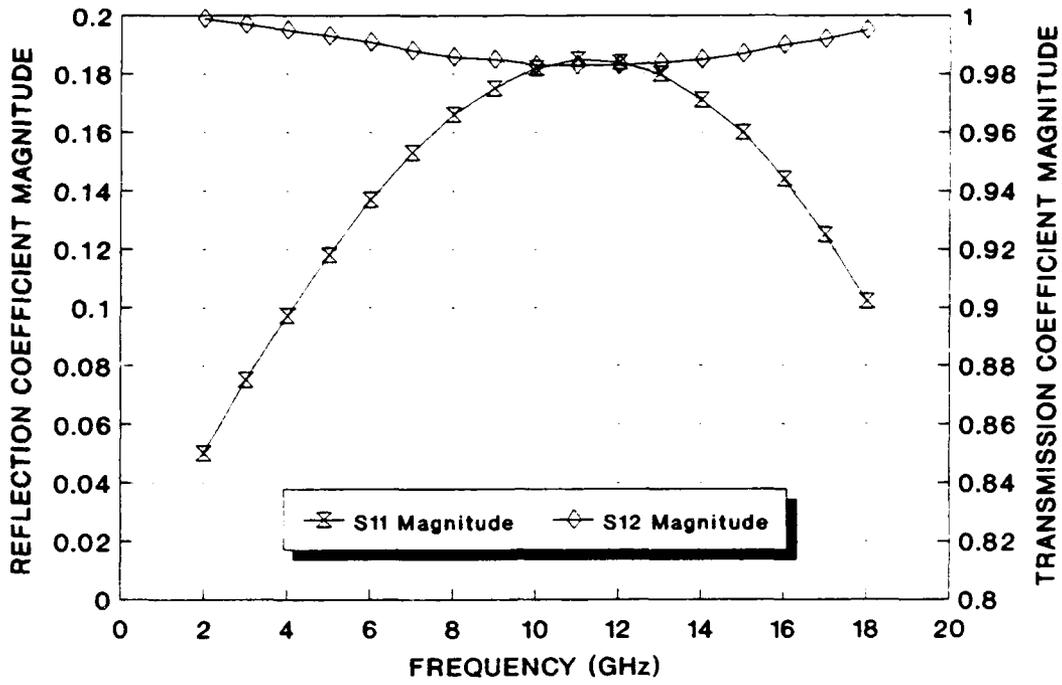


Figure 25. Reflection and Transmission vs. Frequency for 4 mm Long Stripline

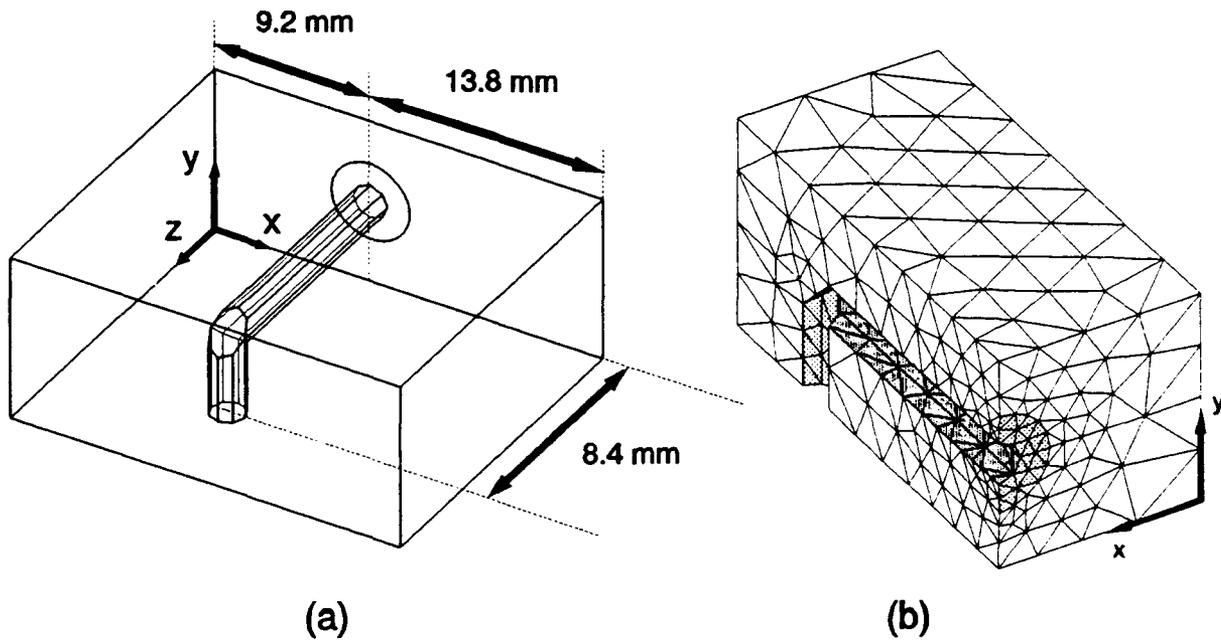


Figure 26. Coaxial End-Wall Launcher for Rectangular Waveguide:
 (a) Geometry and Dimensions; and (b) Cutaway with Right-half Mesh Cells;

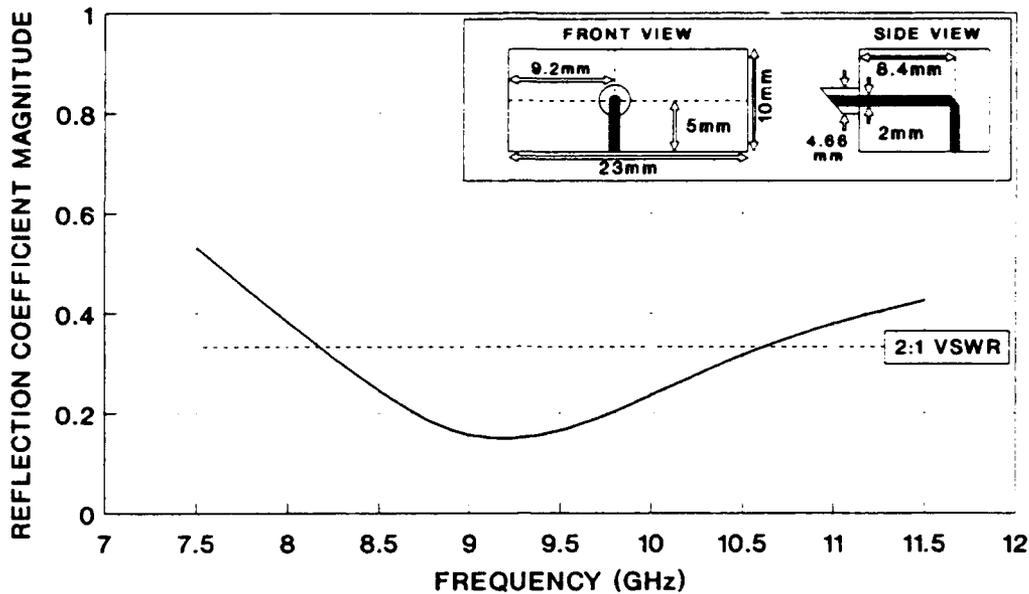


Figure 27. Reflection Coefficient vs Frequency for Coax-to-Rectangular Waveguide Launcher

6.3. Printed Circuit Devices

Up to now this report has only discussed structures with air dielectric. Yet one of the most important capabilities of the finite element method is its ability to deal with inhomogeneous dielectrics. This section will discuss results for two microstrip structures: first, a calculation of guide wavelength in a microstrip line; and second, calculations of S parameters for a microstrip meander line.

6.3.1. MICROSTRIP TRANSMISSION LINE.

Figure 28 shows the mesh for a length of $75 \mu\text{m}$ wide microstrip line on a $100 \mu\text{m}$ -thick substrate (Gallium Arsenide, $\epsilon_r=12.9$). The coaxial connector dimensions are $a = 43 \mu\text{m}$ and $b = 100 \mu\text{m}$, chosen for convenience and an impedance match to the 50Ω microstrip. The substrate is $1000 \mu\text{m}$ wide and the cavity height is $500 \mu\text{m}$ total, with $400 \mu\text{m}$ of air between the microstrip and the cavity's top wall. These dimensions ensure that there is a space of 4-5 line widths between the microstrip and the enclosure, which is adequate to ensure that the package does not influence the guide wavelength or

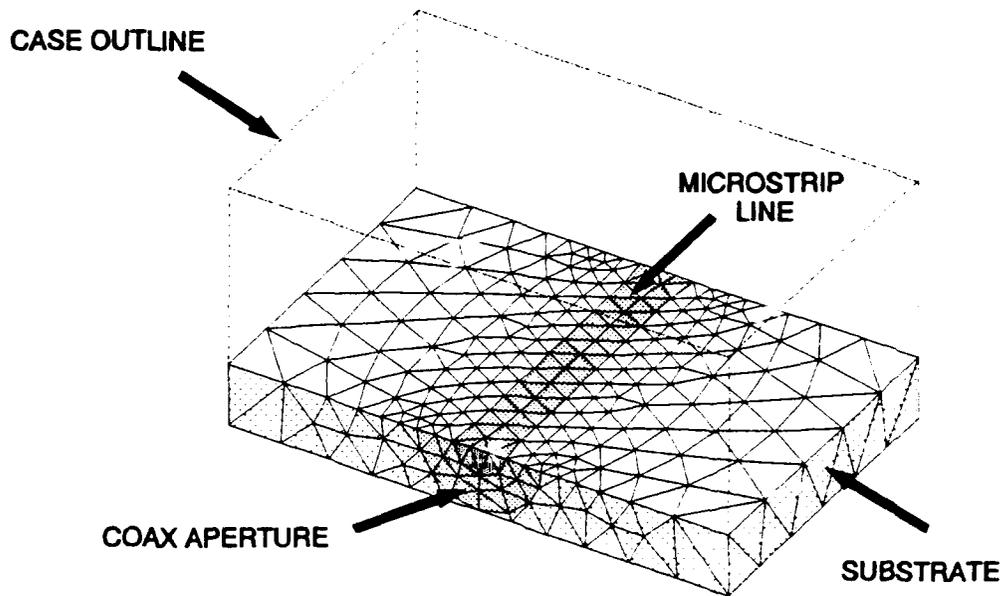


Figure 28. Finite Element Mesh for Microstrip Line with Coaxial Connectors

characteristic impedance [20].

Determining the properties of the microstrip by itself requires calibrating out, or "de-embedding" the effects of the coaxial connectors. Using a scaling option in the TWOPORT code, the S parameters were computed first for a line length of $500 \mu\text{m}$; then second with the same finite element mesh stretched in z to a length of $750 \mu\text{m}$. The differential phase in S_{21} was 35° for the test frequency of 40 GHz. This yields 2.57 mm for the guide wavelength, while quasi-static formulas developed by Wheeler [21] give 2.54 mm. This close agreement indicates that the model is correctly predicting the fields at interfaces between two dielectrics, even in the presence of sharp conducting edges.

6.3.2. MICROSTRIP MEANDER LINE.

Since the modeling method accurately accounts for all relevant boundary conditions for a simple microstrip line, it should be obvious that it can correctly predict the performance of most passive printed-circuit RF components. A microstrip meander line was chosen as a demonstration case since measurements were available. The circuit dimensions (in μm) and the finite element mesh (substrate cells only)

are shown in Figure 29.

The measurement reference planes are at the centers of the pads at each end of the circuit, while those used in the calculations are at the points where the lines begin to taper from 75 μm down to 25 μm , so the difference must be accounted for when comparing the data. The circuit layout shows the location of "via holes" that provide a ground for the measurement probes. At the contact points, the 75 μm center conductor and 50 μm -wide slots form a 50 Ω coplanar waveguide. Using formulas given by Rowe and Lao [22], the effective relative permittivity for that transmission line structure was found as $\epsilon_{\text{eff}}=6.29$. The measured S_{21} phase was corrected by adding $360^\circ \cdot (75 \mu\text{m})(\epsilon_{\text{eff}})^{1/2}/\lambda_0$. The calculated S_{21} phase was also corrected by subtracting the excess phase introduced by the coaxial connectors, computed from a straight length of microstrip line as discussed in the preceding section. Figure 30 shows that the measurements and calculations agree very well over the frequency range from 1 to 26 GHz. Note that the measurement apparatus uses coplanar waveguide probes, while the calculation simulates coaxial

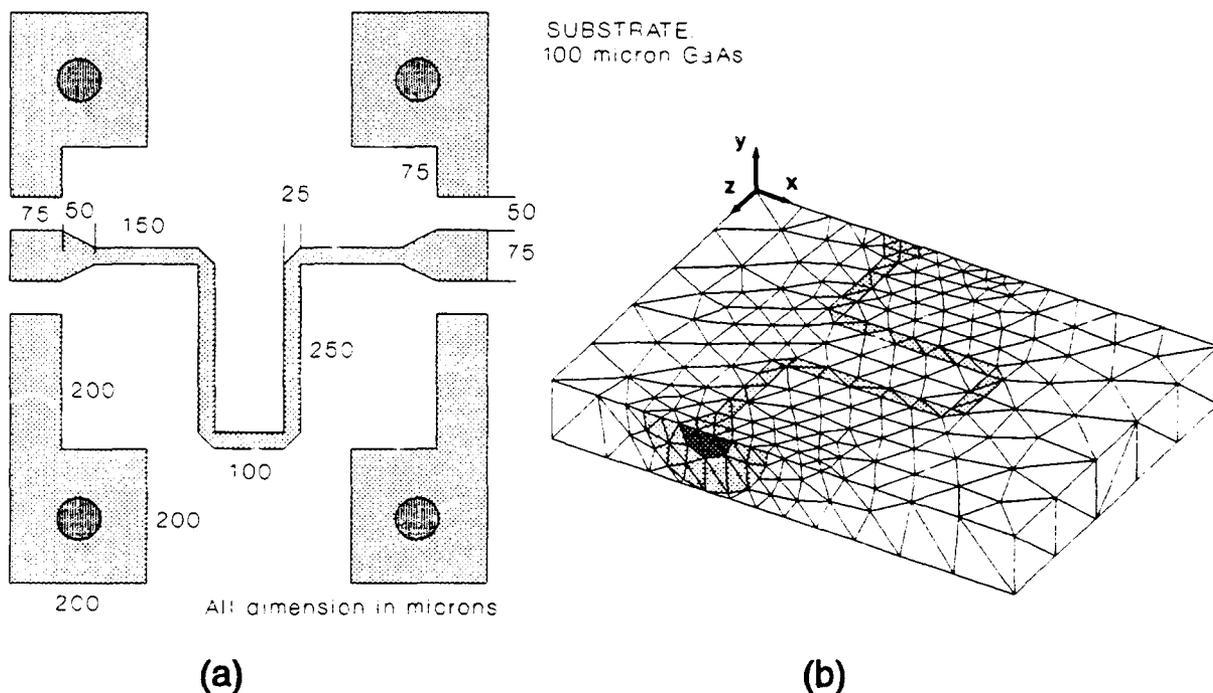


Figure 29. Microstrip Meander Line: (a) Wafer Metallization Dimensions (in μm); and (b) Finite Element Mesh for Substrate

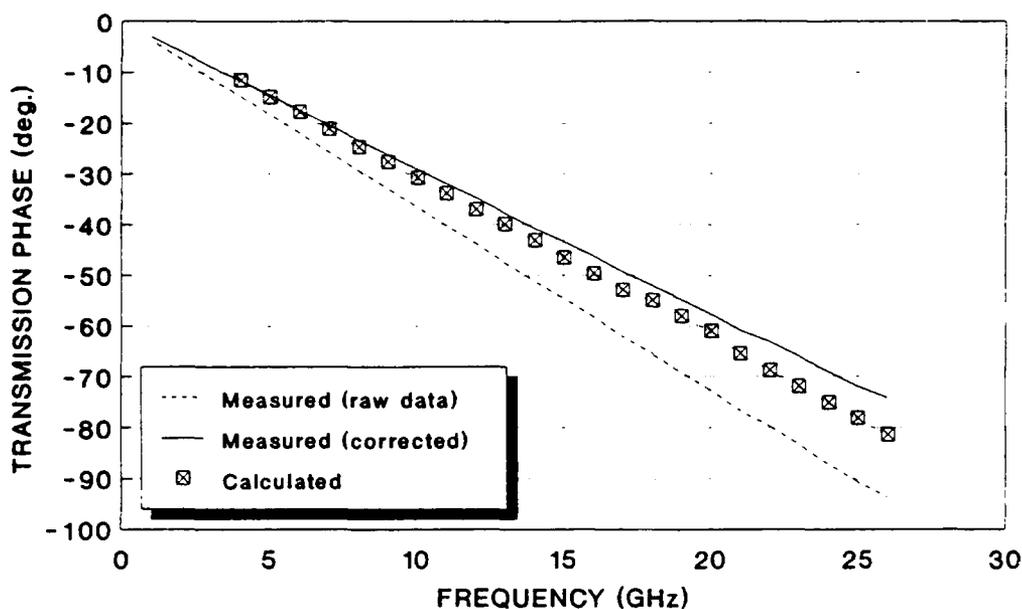


Figure 30. Comparison of Measured and Calculated S_{21} Phase for Microstrip Meander Line

connectors. The CPW-microstrip and coax-microstrip transitions have different reactances, which accounts for the difference in the phase slope between measurement and calculation in Figure 30.

This concludes discussion of the validation results. Together they demonstrate the effectiveness, reliability and versatility of the hybrid finite element method. The final part of this section presents conclusions that may be drawn from these tests regarding rule-of-thumb modeling requirements such as number of waveguide modes, number of finite elements and typical execution time.

6.4. Modeling Requirements

6.4.1. HIGHER-ORDER MODES.

All of the validation cases used many more waveguide modes than were actually required. The additional matrix fill calculations are not very significant for rectangular waveguides, but they can be

somewhat expensive for coax. If computation time is critical, it may be important to know how many modes are actually needed to obtain an accurate answer.

Approximate matrix fill execution times are as follows for the VAX 8650 computer, given in terms of CPU seconds per mode per aperture edge: rectangular, 2.5 ms; circular, 25 ms; and coaxial, 250 ms. The circular and coaxial mode function integrations both use a 16-point quadrature rule, meaning that there are 16 integration sample points in each aperture triangle.

The TWOPORT code includes a calculation of the excitation coefficients for all of the modes included in the matrix calculation. These coefficients are the C_i 's from Eq.(7). Figure 31 shows the results for the two coaxial irises discussed in Section 6.1.2. Apparently, there is a "noise floor" at -30 to -40 dB due to gridding (discretization) effects. These iris discontinuities evidently only excite TE modes, and only higher orders of m . This is consistent with expectations, since higher orders of n correspond to functions with more cycles in the radial direction and should not be excited by radially-invariant obstructions. Since coaxial and circular modes degenerate into $\sin(m\phi)$ and $\cos(m\phi)$ components (see Appendix B), both must be included to represent the fields at *arbitrary* obstructions, which may not be aligned with the coordinate axes.

Evidently, only modes $mn=01,11$, and 21 are needed for these coaxial irises. This conclusion is reinforced by Figure 32, which shows the computed values of $|S_{11}|$ for different numbers of modes. Adding modes higher than 21 did not produce any change in the reflection coefficient, so they are not shown in Figure 32.

The coax-stripline transition discussed in Section 6.2.1 is, by contrast, a longitudinal (parallel to z) rather than a lateral (transverse to z) discontinuity. Consequently, it excites higher order TM modes. Its excitation coefficients are shown in Figure 33. The strength of these modes is abnormally high because the coax diameter is so much larger than the cavity height, so these results are not representative of coaxial transitions to stripline or microstrip. In fact, the microstrip circuits in Sections 6.3.1 and 6.3.2

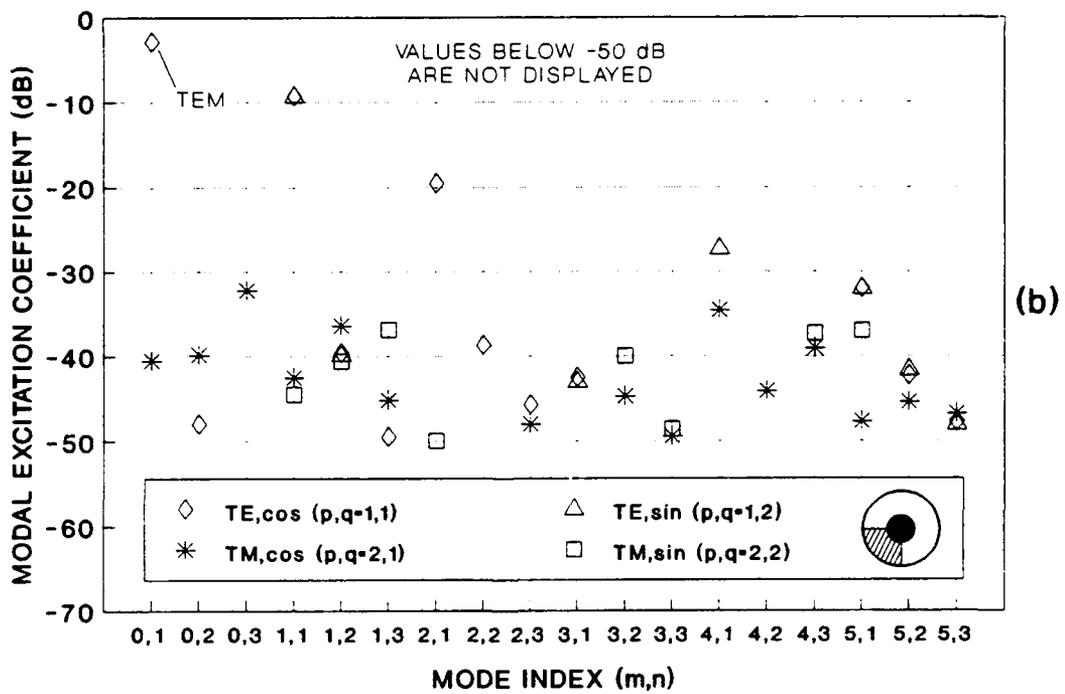
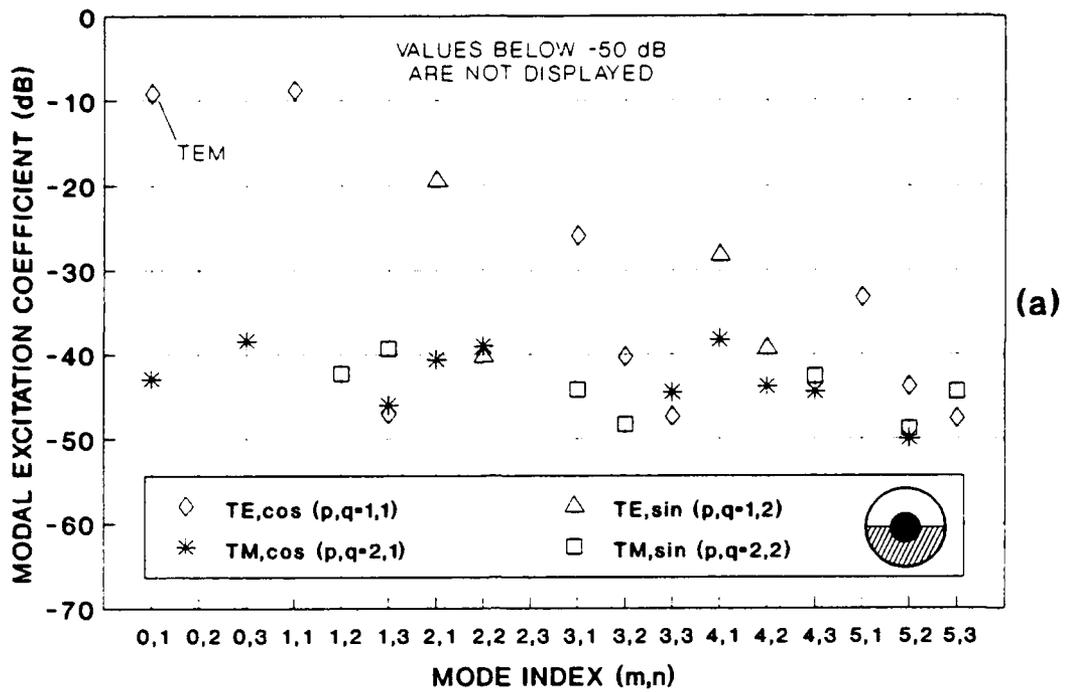


Figure 31. Mode Excitation Coefficients at Coaxial Waveguide Irises:
 (a) 180° Iris Opening; (b) 270° Iris Opening

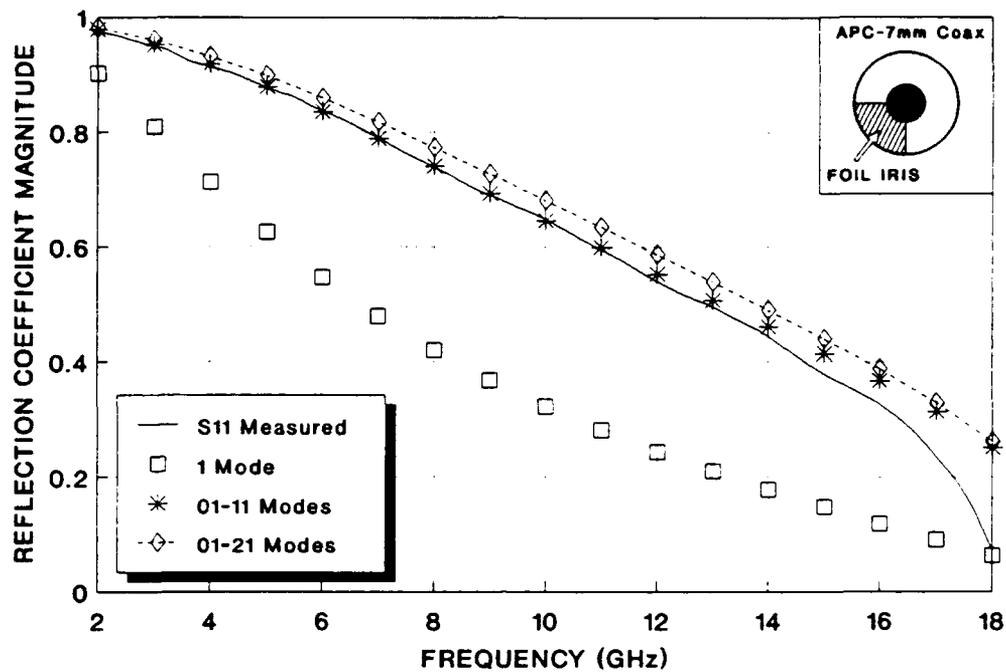


Figure 32. Reflection Coefficient Magnitude vs. Frequency Calculated with Varying Numbers of Higher Order Modes, 270° Coaxial Iris

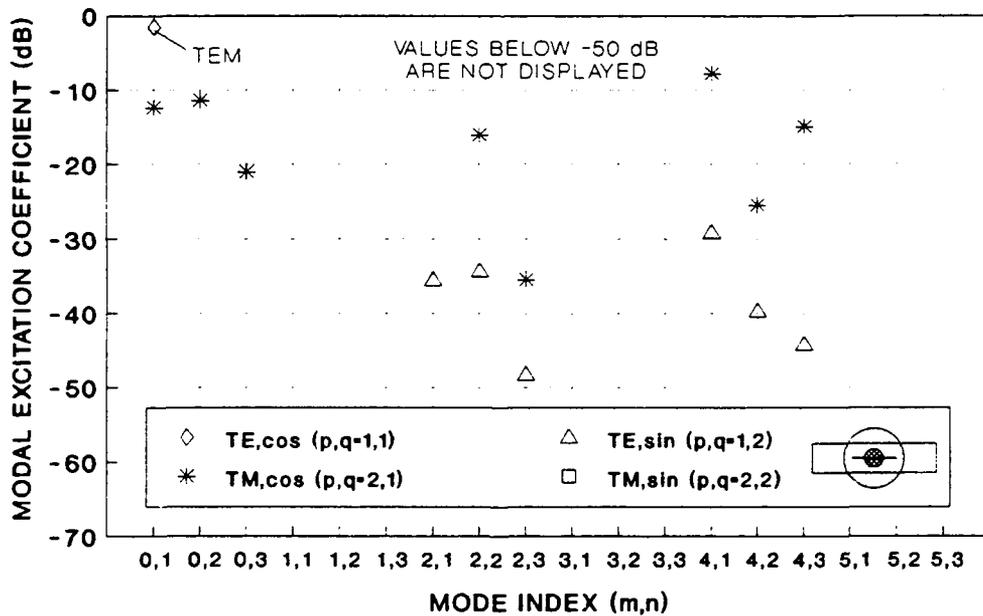


Figure 33. Higher-Order Coaxial Mode Excitation Coefficients at Coax-Stripline Transition

did not generate any appreciable higher order modes because the coax outer radius was the same as the substrate height.

The last issue to be addressed is whether there is in fact any benefit to using higher-order modes. The alternative is, of course, to extend the finite element mesh into the connecting waveguides far enough that any higher modes excited by the cavity's contents and/or apertures have decayed to insignificance [23]. The answer depends on two factors: first, the mode excitation strengths; and second, their attenuation constants in the waveguides, which depend mainly on the ratio of the frequency to the mode cutoff frequency. Consider the following two extreme examples: First, the microstrip circuits do not generate higher-order modes in the coaxial waveguides. Therefore, the mesh may be terminated at the coax apertures even when only the dominant mode is used. At the other extreme, the one-quadrant coaxial iris (Figure 30b) excites the TE_{11} mode at about -10 dB. Its cutoff frequency in APC-7 coax is about 20 GHz. At 18 GHz it decays by 29 dB per wavelength, and for it to be negligible (below -30 dB), the mesh would need to extend roughly $\frac{2}{3}\lambda$ into the waveguide in each direction. The finite element mesh and the interaction matrix would then contain an additional 500-1500 edges and unknowns, respectively. Thus it is clear that in some cases at least, the use of higher-order modes is very beneficial.

6.4.2. DISCRETIZATION REQUIREMENTS.

The discretization for most of the test cases is summarized in Table 1. It lists the total volume in cubic wavelengths at the highest frequency, the total number of cells and the number of cells per cubic wavelength. Accurate results for S parameters may be obtained with as few as 10 nodes or edges per wavelength in each linear direction. Finer sampling may be required in some instances. For example, the circular waveguide mesh in Figure 13 was not adequate for the mode converter problem. It produced fairly good answers for the case with the smaller inlet waveguide radius, but its results were off by as much as 4 dB for the other case. The reason is that the number of integration points over the collection of triangles in the aperture was not enough to give an accurate result for inner products with the higher-

order mode functions. No more than 20 nodes per wavelength in any linear direction should be required, but in dense dielectrics, such as GaAs, the wavelength is much smaller than in free space. In many instances the mesh density was dominated by fine details in the geometry. I-DEAS attempts to make each tetrahedron as nearly equilateral as possible, which does not allow the node spacing to "relax" very abruptly. The rationale is that highly elongated cells may cause ill conditioning in certain finite element applications. The consequence here is that there are many more cells than are required to make accurate field computations.

6.4.3. MATRIX SOLUTION.

Figure 34 summarizes the execution times observed for the various test cases. All of these were on the VAX 8650 using single precision. The LU decomposition was performed by the IMSL routine L2LCG, which is optimized to use a vector floating point coprocessor. LU decomposition was also executed on a Cray Y-MP, and typically ran about 60 times as fast as it did on the VAX.

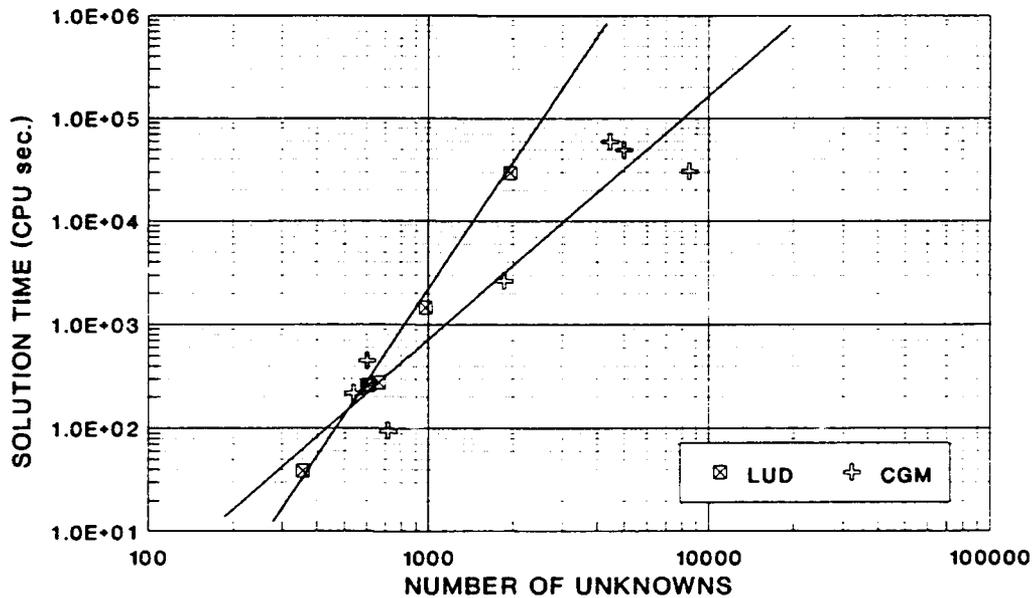


Figure 34. Execution Time vs Number of Unknowns for LU Decomposition and Conjugate Gradient Method (VAX 8650)

Table 1. Summary of Finite Element Discretization for Test Cases

Test Case	Reference Figure No.	$\frac{\text{Volume}}{\lambda^3}$	No. Cells in Mesh	$\frac{\text{Cells}}{\lambda^3}$
APC-7 Coax	13	0.0135	847	63×10^3
Rectangular Waveguide	15	0.149	308	2067
Circular Waveguide	17	0.120	342	2850
Circular W/G Mode Converter	22a	0.376	1504	4000
Stripline	24b	0.0138	506	37×10^3
Launcher	26	0.2944	2442	8294
Microstrip Line	27	5.7×10^{-4}	4662	8.2×10^6
Meander Line	28	9.4×10^{-5}	4190	4.5×10^7

The conjugate gradient solver is not vectorized, yet for small numbers of unknowns it runs almost as fast as LU decomposition. It is clearly a better choice for large problems, but the scattering of points in Figure 34 reflects the disadvantage that the solution time is difficult to predict *a priori*.

7. CONCLUSIONS AND RECOMMENDATIONS

The hybrid finite element/waveguide mode approach is clearly an effective tool for modeling the electromagnetic fields in a wide variety of radio frequency devices. The success of the implementation presented in this report indicates that a number of features of this approach are appropriate and correct, specifically: (a) the representation of the boundary value problem as a weak form of the Helmholtz

equation for the electric field; (b) the enforcement of transverse magnetic field continuity across waveguide apertures using the boundary term of the functional; (c) the representation of waveguide fields by mode sums and the implementation of the higher-order mode functions; (d) the enforcement of the divergence condition and conductor boundary conditions using vector finite elements; (e) the mapping of the problem to a matrix equation using Galerkin's method; (f) the solution of the system of equations using LU decomposition or the conjugate gradient method; and finally, (g) the implementation of all of the above in a general-purpose computer code.

There are numerous possible extensions and improvements of the work presented here. The most important of these are: (a) improving the efficiency of the matrix solution by using preconditioning in conjunction with the conjugate gradient method; and (b) implementation of a "combined source" analogue for the aperture field continuity to prevent the kind of resonance effects observed with the coax-rectangular launcher problem. A few others that would be quite straightforward code modifications are: (a) increasing the number of ports; and (b) allowing ports to be located in the side and/or top walls of the cavity. Others that are desirable, but will require some theoretical development as well, are anisotropic materials and active devices. However, as stated in the introduction, this project's goal is to apply the hybrid finite element method to phased array antenna analysis, and our work for the immediate future will continue in that direction.

References

1. Strang, G. and Fix, G.R. (1973) An Analysis of the Finite Element Method, Prentice-Hall, Englewood Cliffs, NJ.
2. Marcuvitz, N. (1951) Waveguide Handbook, New York: McGraw-Hill.
3. Harrington, R.F. and Mautz, J.R. (1975) A Generalized Network Formulation for Aperture Problems, Air Force Cambridge Research Laboratory, Hanscom AFB, MA, AFCRL-TR-75-0589, ADA019690.
4. Harrington, R.F. (1982) Field Computation by Moment Methods, Krieger, Malabar, FL.
5. Boyse, W.E., Lynch, D.R., Paulsen, K.D. and Minerbo, G.N. (1992) Nodal-Based Finite Element Modeling of Maxwell's Equations, IEEE Trans. Antennas Propagat., **AP-40**, pp. 642-651.
6. Barton, M.L. and Cendes, Z.J. (1987) New Vector Finite Elements for Three-Dimensional Magnetic Field Computation, J. Appl. Phys., **61**, pp. 3919-3921.
7. D'Angelo, J.D. and Mayergoyz, I.D. (1990) Finite Element Methods for the Solution of RF Radiation and Scattering Problems, Electromagnetics, **10**, pp. 177-199.
8. Silvester, P.P. and Ferrari, R.L. (1990) Finite Elements for Electrical Engineers, 2nd ed., Cambridge Univ. Press, 1990.
9. Stroud, A.H. and Secrest, D. (1966) Gaussian Quadrature Formulas, New York: Prentice-Hall.
10. Abramowitz, M. and Stegun, I.E. (1972) Handbook of Mathematical Functions, National Bureau of Standards.
11. Pissanetsky, S. (1984) Sparse Matrix Technology, New York: Academic Press.
12. Heath, M.T., Ng, E. and Peyton, B.W. (1991) Parallel Algorithms for Sparse Linear Systems, Siam Review, **33**, pp. 420-460.

13. Sarkar, T.K. and Arvas, E. (1985) On a Class of Finite Step Iterative Methods (Conjugate Directions) for the Solution of an Operator Equation Arising in Electromagnetics, IEEE Trans. Antennas Propagat., AP-33, pp. 1058-1066.
14. Woodworth, M.B. (1988) Large Matrix Solution Techniques Applied to an Electromagnetic Scattering Problem, RADC-TR-88-268, ADA206917.
15. Structural Dynamics Research Corporation (1990) Integrated Design Engineering Analysis Software Users' Manual, (Version V), SDRC, Milford, OH.
16. Montgomery, C.G., Dicke, R.H., and Purcell, E.M. (1948) Principles of Microwave Circuits, New York: McGraw-Hill.
17. Masterman, P.H. and Clarricoats, P.J.B. (1971) Computer Field-Matching Solution of Waveguide Transverse Discontinuities, Proc. IEE, 188, pp. 51-63.
18. Gupta, K.C., Gharg, R. and Chadha, R. (1981) Computer-Aided Design of Microwave Circuits, Dedham MA: Artech House.
19. Tang, R. and Wong, N.S. (1968) Multimode Phased Array Element for Wide Scan Angle Impedance Matching, Proc. IEEE, 56, pp. 1951-1959.
20. Kowalski, G. and Pregla, R. (1971) Dispersion Characteristics of Shielded Microstrips with Finite Thickness, Arch. Elek. Ubertragung, 25, pp. 193-196; see also (1989) Numerical Methods for Passive Microwave and Millimeter Wave Structures, ed. R. Sorrentino, IEEE Press.
21. Wheeler, H.A., (1983) Transmission Line Properties of a Strip on a Dielectric Sheet on a Plane, IEEE Trans. Microwave Theory Tech., MTT-25, pp. 911-915.
22. Rowe, D.A. and Lao, B.Y. (1983) Numerical Analysis of Shielded Coplanar Waveguide, IEEE Trans. Microwave Theory Tech., MTT-31, pp. 911-915.
23. Webb, J.P., Maile, G.L. and Ferrari, R.L. (1983) Finite Element Solution of Three Dimensional Electromagnetic Problems, IEE Proceedings, 130, pp. 153-159.
24. Jin, J-M and Volakis. J.L. (1991) A Hybrid Finite Element Method for Scattering and Radiation by Microstrip Patch Antennas and Arrays Residing in a Cavity, IEEE Trans. Antennas Propagat., AP-39, pp. 1598-1604.
25. Mikhlin, S.G. (1974) Variational Methods in Mathematical Physics, New York: MacMillan.
26. Harrington, R.F. (1961) Time-Harmonic Electromagnetic Fields, New York: McGraw-Hill.
27. Naylor, A.W. and Sell, G.R. (1982) Linear Operator Theory in Engineering and Science, New York: Springer-Verlag.
28. Whinnery, J.R., Jamieson, H.W. and Robbins, T.E. (1944) Coaxial-Line Discontinuities, Proc. IRE, pp. 695-709.

Appendix A

The Electric Field Functional

A1. Variational Principle vs Weak Form

Publications dealing with electromagnetic finite element applications usually begin with one of two forms of *functionals*¹ with little or no discussion as to why one is used and not the other. This appendix discusses their origins and the circumstances in which each is appropriate. The conclusion is that the differences in the results of the two methods are inconsequential for typical electromagnetic radiation and scattering problems even though only one of them is considered to be rigorously justified.

The first alternative is

$$F_v(\bar{E}) = \frac{1}{2} \iiint_{\Omega} \left[\frac{1}{\mu_r} \nabla \times \bar{E} \cdot \nabla \times \bar{E} - k_0^2 \epsilon_r \bar{E} \cdot \bar{E} \right] dv + jk_0 \eta_0 \iint_{\Gamma} (\bar{E} \times \bar{H}) \cdot \hat{n} ds \quad (A1)$$

(see, for example, Jin & Volakis [24]) where Ω is the volume region over which the unknown field is

¹ A functional is an integral whose integrand includes an unknown function. It maps the function into a number, that is, the result of the integration is a single number, as opposed to an integral equation, which maps the function into another function.

to be found and Γ is its enclosing boundary. The subscript v denotes that this functional is the *variational principle*.² The second alternative is

$$F_w(\bar{E}) = \iiint_{\Omega} \left[\frac{1}{\mu_r} \nabla \times \bar{W} \cdot \nabla \times \bar{E} - k_0^2 \epsilon_r \bar{W} \cdot \bar{E} \right] dv + jk_0 \eta_0 \iint_{\Gamma} (\bar{W} \times \bar{H}) \cdot \hat{n} ds \quad (A2)$$

(see, for example, D'Angelo & Mayergoyz [7]). Here, \bar{W} is a trial function whose form is yet to be determined. The subscript w on F denotes the fact that this is the *weak form*.

The variational principle F_v has been deduced from a general form of energy functional [25] and is used in conjunction with the Rayleigh-Ritz principle to form a system of equations. The variational principle is usually avoided in situations where the problem is not *self-adjoint* because the functional F_v cannot then be proven to be stationary about the solution to the operator equation. The weak form is derived more directly by simply taking the inner product of the operator equation with the trial function. It will be used in conjunction with the method of weighted residuals to form the system of equations. When Galerkin's method (a specialization of weighted residuals) is used, the two forms may give exactly the same system of equations. But, to use Galerkin's method, the expansion function that is admissible in the original problem must also be admissible in the *adjoint problem*.

The following section will discuss the meaning of the adjoint problem and what the conditions are for self-adjointness as applied to the Helmholtz equation. The third section will then show that typical waveguide continuity conditions represent non-self-adjoint boundary conditions. The fourth will show that under the assumption that Galerkin's method is applicable, the two formulations generate identical systems of equations.

² A *variational statement* may be either a *variational principle* or a *weak form*.

A2. The Adjoint Problem

To decide whether to use Eq.(A1) or Eq.(A2) one must know whether or not the operator equation is self-adjoint. If it is not, then the properties of the adjoint problem must be found in order to ensure that the trial functions are capable of representing its solutions.

The operator equation is the time-harmonic, source-free Helmholtz equation for the electric field in a linear, isotropic, inhomogeneous region:

$$L(\bar{E}) = \nabla \times \frac{1}{\mu_r} \nabla \times \bar{E} - k_0^2 \epsilon_r \bar{E} = 0 \quad (\text{A3})$$

Its inner product with an arbitrary complex function \bar{W} is

$$\langle L(\bar{E}), \bar{W} \rangle = \iiint_{\Omega} \left[\nabla \times \frac{1}{\mu_r} \nabla \times \bar{E} - k_0^2 \epsilon_r \bar{E} \right] \cdot \bar{W}^* dv = 0 \quad (\text{A4})$$

Using a Green's identity (integration by parts) twice, we shift all of the derivatives from \bar{E} to \bar{W} :

$$\begin{aligned} \langle L(\bar{E}), \bar{W} \rangle &= \iiint_{\Omega} \left[\frac{1}{\mu_r} \nabla \times \bar{E} \cdot \nabla \times \bar{W}^* - k_0^2 \epsilon_r \bar{E} \cdot \bar{W}^* \right] dv \\ &\quad - \iint_{\Gamma} \frac{1}{\mu_r} \bar{W}^* \times \nabla \times \bar{E} \cdot \hat{n} ds \end{aligned} \quad (\text{A5})$$

$$\begin{aligned} &= \iiint_{\Omega} \bar{E} \cdot \left[\nabla \times \frac{1}{\mu_r} \nabla \times \bar{W} - k_0^2 \epsilon_r \bar{W} \right]^* dv \\ &\quad - \iint_{\Gamma} \frac{1}{\mu_r} [\bar{W}^* \times \nabla \times \bar{E} - \bar{E} \times \nabla \times \bar{W}^*] \cdot \hat{n} ds \end{aligned} \quad (\text{A6})$$

From the definition

$$\langle L(\bar{E}), \bar{W} \rangle = \langle \bar{E}, L^a(\bar{W}) \rangle \quad (\text{A7})$$

it is evident that the term in brackets in the first integral of (A6) is L^a , the *adjoint operator*. It is simply

the Helmholtz equation with the constitutive parameters replaced by their complex conjugates. It is now clear that \bar{W} must be in the domain of the adjoint operator.

The definition of self-adjointness is $L=L^a$, which obviously cannot be true if the problem includes lossy materials. But it also depends on whether the boundary conditions are such as to make the surface integral in Eq.(A6) vanish, that is,

$$\iint_{\Gamma} \frac{1}{\mu_r} \bar{E}^{a*} \times \nabla \times \bar{E} \cdot \hat{n} ds = \iint_{\Gamma} \frac{1}{\mu_r} \bar{E} \times \nabla \times \bar{E}^{a*} \cdot \hat{n} ds \quad (\text{A8})$$

$$\iint_{\Gamma} \frac{\mu_r}{\mu_r} \bar{E}^{a*} \times \bar{H} \cdot \hat{n} ds = \iint_{\Gamma} \frac{\mu_r}{\mu_r} \bar{E} \times \bar{H}^{a*} \cdot \hat{n} ds \quad (\text{A9})$$

Suppose that all lossy magnetic materials are confined inside Ω so that along the boundary Γ , the permeability is entirely real. Then Eq.(A9) is satisfied when $\bar{E}^a = \bar{E}^*$ and $\bar{H}^a = \bar{H}^*$. Recall that the time-harmonic and time-dependent fields are related by [Reference 26, p.15]

$$\bar{\mathbf{E}}(x, y, z, t) = \sqrt{2} \operatorname{Re}(\bar{\mathbf{E}} e^{j\omega t}) \quad (\text{A10})$$

(boldface represents the time-dependent quantity; the expression for magnetic field is identical). Evaluating Eq.(A10) at any point \bar{r} along Γ , with \bar{E}_0 denoting $\bar{\mathbf{E}}(\bar{r})$:

$$\bar{\mathbf{E}}(\bar{r}, t) = \sqrt{2} \left[\operatorname{Re}(\bar{E}_0) \cos \omega t - \operatorname{Im}(\bar{E}_0) \sin \omega t \right] \quad (\text{A11})$$

and assuming that $\bar{E}^a = \bar{E}^*$

$$\begin{aligned}
\bar{E}^a(\bar{r}, t) &= \sqrt{2} \operatorname{Re} \left[\bar{E}_0^* e^{j\omega t} \right] = \sqrt{2} \left[\operatorname{Re}(\bar{E}_0) \cos \omega t + \operatorname{Im}(\bar{E}_0) \sin \omega t \right] \\
&= \sqrt{2} \operatorname{Re} \left[\bar{E}_0 e^{-j\omega t} \right] \\
&= \bar{E}(\bar{r}, -t)
\end{aligned} \tag{A12}$$

In other words, the adjoint fields are time-reversed versions of the original fields. They carry power across the boundary in the opposite direction and they encounter materials that have gain instead of loss. This is the physical interpretation of the adjoint problem, and is consistent with the property that if the operator L is *causal*, then the operator L^a is *anti-causal* [Reference 27, p.356].

Notice that if the boundary Γ is comprised entirely of perfect electric and perfect magnetic conductors, then Eq.(A9) is always satisfied because there cannot be any transfer of power across such boundaries. Thus we are led to suspect that the boundary conditions that will cause non-self-adjointness are those open boundaries where conditions of field continuity are to be enforced. The next section will demonstrate that this is true for the waveguide/cavity apertures that are considered in the main body of this report.

A3. Continuity Conditions for Waveguide Apertures

For the waveguide continuity problem, the boundary Γ reduces to the waveguide aperture Γ_A . Taking $\bar{E}^a = \bar{E}^*$ and $\bar{H}^a = \bar{H}^*$, the following form is equivalent to Eq.(A9):

$$\iint_{\Gamma_A} \bar{E}^* \cdot (\hat{n} \times \bar{H}) ds = \iint_{\Gamma_A} \bar{E} \cdot (\hat{n} \times \bar{H}^*) ds \tag{A13}$$

(again assuming that μ_r is real along Γ). Consider a waveguide whose axis is the z axis and which joins the volume Ω through an aperture in its end wall at $z=0$. It is assumed to be match-terminated at $z \ll 0$. The field in the aperture due to the dominant mode and its complex conjugate are

$$\begin{aligned}\bar{E} &= \bar{g}_0(1+C_0) \\ \bar{E}^* &= \bar{g}_0(1+C_0^*)\end{aligned}\quad (\text{A14})$$

$$\begin{aligned}\bar{H} &= \hat{z} \times \bar{g}_0 Y_0(1-C_0) \\ \bar{H}^* &= \hat{z} \times \bar{g}_0 Y_0(1-C_0^*)\end{aligned}\quad (\text{A15})$$

where \bar{g}_0 is a transverse mode function, Y_0 is the modal admittance and C_0 is an unknown coefficient.

The outward normal to Ω is $\hat{n} = -\hat{z}$, so the left and right sides of Eq.(A13) are

$$\iint_{\Gamma_A} \bar{E}^* \cdot (\hat{n} \times \bar{H}) ds = Y_0(1+C_0^*)(1-C_0) \iint_{\Gamma_A} |\bar{g}_0|^2 ds \quad (\text{A16})$$

$$\iint_{\Gamma_A} \bar{E} \cdot (\hat{n} \times \bar{H}^*) ds = Y_0(1+C_0)(1-C_0^*) \iint_{\Gamma_A} |\bar{g}_0|^2 ds \quad (\text{A17})$$

respectively. They are only equal if C_0 is real, which is generally false. Therefore, the continuity condition across a waveguide aperture will render the operator equation non self-adjoint; the functional Eq.(A1) is not appropriate even when there are no lossy materials; and the weak form should be used.

A4. Galerkin's Method vs Rayleigh-Ritz

In the main body of this report, the finite element method is used to produce a system of equations from the weak form functional. The field is represented by a summation of unknown complex coefficients with known, linear vector functions:

$$\bar{E} \approx \hat{E} = \sum_{s=1}^N e_s \bar{\psi}_s(x, y, z) \quad (\text{A18})$$

For the expansion functions to be *admissible*, they must be in the domain of the functional. Linear

functions meet that requirement since their first derivatives are continuous (integrable).

The residual error is defined as $\bar{R}=L(\bar{E})$. N weighting functions will be chosen, each of which is orthogonal to the residual so that $\langle \bar{R}, \bar{w}_r \rangle = 0$, giving the system of N equations

$$0 = \sum_{s=1}^N e_s \iiint_{\Omega} \left[\frac{1}{\mu_r} \nabla \times \bar{w}_r \cdot \nabla \times \bar{\psi}_s - k_0^2 \epsilon_r \bar{w}_r \cdot \bar{\psi}_s \right] dv + jk_0 \eta_0 \iint_{\Gamma} (\bar{w}_r \times \bar{H}) \cdot \hat{n} ds, \quad r=1, 2, \dots, N \quad (\text{A19})$$

The choice $\bar{w}_r = \bar{\psi}_r$ satisfies the orthogonality requirement, but it is also necessary that $\bar{\psi}_r$ be an admissible expansion function for the adjoint electric field. Starting with the adjoint operator equation, forming $\langle \bar{E}, L^a(\bar{W}) \rangle$ and using the Green's identity once gives

$$\langle \bar{E}, L^a(\bar{W}) \rangle = \iiint_{\Omega} \left[\frac{1}{\mu_r} \nabla \times \bar{W}^* \cdot \nabla \times \bar{E} - k_0^2 \epsilon_r \bar{W}^* \cdot \bar{E} \right] dv - \iint_{\Gamma} \frac{1}{\mu_r} \bar{E} \times \nabla \times \bar{W}^* \cdot \hat{n} ds \quad (\text{A20})$$

which makes it evident that expansion functions for \bar{W} must have continuous first derivatives. Therefore, the same expansion functions are admissible in both the original and the adjoint problems. It is this fact that allows Galerkin's method to be employed.

Consider the variational functional, Eq.(A1), with the expansion Eq.(A18) for the electric field:

$$F_v(\bar{E}) = \frac{1}{2} \sum_{r=1}^N e_r \sum_{s=1}^N e_s \iiint_{\Omega} \left[\frac{1}{\mu_r} \nabla \times \bar{\psi}_r \cdot \nabla \times \bar{\psi}_s - k_0^2 \epsilon_r \bar{\psi}_r \cdot \bar{\psi}_s \right] dv + jk_0 \eta_0 \sum_{r=1}^N e_r \iint_{\Gamma} (\bar{\psi}_r \times \bar{H}) \cdot \hat{n} ds, \quad r=1, 2, \dots, N \quad (\text{A21})$$

The Rayleigh-Ritz method equates the stationary point of F_v to the minimization of the above with respect to each of the coefficients e_r , that is,

$$\delta F_v(\bar{E}) = 0 \Leftrightarrow \frac{\partial F_v(\bar{E})}{\partial e_r} = 0, \quad \forall r \quad (\text{A22})$$

Carrying out the partial derivatives in Eq.(A21) gives

$$\begin{aligned} 0 = \sum_{s=1}^N e_s \iiint_{\Omega} \left[\frac{1}{\mu_r} \nabla \times \bar{\psi}_r \cdot \nabla \times \bar{\psi}_s - k_0^2 \epsilon_r \bar{\psi}_r \cdot \bar{\psi}_s \right] dv \\ + j k_0 \eta_0 \iint_{\Gamma} (\bar{\psi}_r \times \bar{H}) \cdot \bar{n} ds, \quad r=1,2,\dots,N \end{aligned} \quad (\text{A23})$$

which is identical to Eq.(A19) with \bar{w}_r replaced by $\bar{\psi}_r$. This shows that the systems of equations resulting from the variational principle and from the weak form are identical. Thus, under at least some circumstances the distinction between the two forms is inconsequential, the adjoint form may be ignored, and the variational principle may be used in situations where it is not rigorously justified.

Appendix B

Coaxial and Circular Waveguide Mode Functions

B.1. COAXIAL

In most instances the properties of circular coaxial waveguide are governed by the TEM mode because higher-order modes are strongly attenuated. Near discontinuities, however, those higher order TE and TM modes contribute significantly to the fields, and so their effects must be included to accurately predict the properties of discontinuities. This appendix discusses the nature of coaxial modes and gives formulas defining the orthonormal mode functions (extracted or derived from Marcuvitz [2]). It also describes a set of computer subroutines that implement the formulas, as well as some calculations that validate their accuracy.

B1.1. Mode Functions

The circular coaxial waveguide has inner and outer radii a and b , respectively. The formulas that follow will be given in terms of polar coordinates, with the waveguide axis coinciding with the z axis. The higher order modes have indices m and n , which are, respectively, the number of cycles in the

circumferential direction, and the number of half-cycles in the radial direction. For each combination m,n there are four distinct mode types: TE or TM; and $\cos(m\phi)$ or $\sin(m\phi)$. All four must be included to represent an arbitrary field. The mode types will be represented in the following by the subscripts p and q where $p=1$ for TE or 2 for TM, and $q=1$ for $\cos(m\phi)$ or 2 for $\sin(m\phi)$.

Waveguide fields are represented in terms of transverse (no \hat{z} component) mode functions $\bar{g}(\rho, \phi)$. The transverse component of an arbitrary field with propagation in the $+z$ direction is represented as a sum of modes with unknown excitation coefficients C_{pqmn} :

$$\bar{E}_t(\rho, \phi, z) = C_0 e^{\gamma_0 z} \bar{g}_0 + \sum_{p=1}^2 \sum_{q=1}^2 \sum_{m=0}^{\infty} \sum_{n=n_1}^{\infty} C_{pqmn} e^{-\gamma_{pmn} z} \bar{g}_{pqmn} \quad (\text{B1})$$

The starting index for n is $n_1=2$ for $p=1$ and $m=0$ or $n_1=1$ otherwise, because there is no TE_{01} mode - the first TE mode is $mn=02$. The propagation constants are

$$\gamma_{pmn} = \begin{cases} j \left[k^2 - k_{mn}^2 \right]^{1/2}, & k > k_{mn} \text{ (propagating)} \\ \left[k_{mn}^2 - k^2 \right]^{1/2}, & k < k_{mn} \text{ (evanescent)} \end{cases} \quad (\text{B2})$$

where the cutoff wavenumber k_{mn} is the n 'th root of one of the characteristic equations

$$J'_n(k_{mn}a) Y'_n(k_{mn}b) = Y'_n(k_{mn}a) J'_n(k_{mn}b) \quad (\text{TE}) \quad (\text{B3})$$

$$J_n(k_{mn}a) Y_n(k_{mn}b) = Y_n(k_{mn}a) J_n(k_{mn}b) \quad (\text{TM}) \quad (\text{B4})$$

The higher-order mode functions are given by the following ten equations:

$$\bar{g}_{11mn} = \hat{\rho} \frac{m}{\rho} Z_{1mn}(k_{mn}\rho) \sin(m\phi) + \hat{\phi} k_{mn} Z'_{1mn}(k_{mn}\rho) \cos(m\phi) \quad (\text{B5})$$

$$\bar{g}_{12mn} = -\hat{\rho} \frac{m}{\rho} Z_{1mn}(k_{mn}\rho) \cos(m\phi) + \hat{\phi} k_{mn} Z'_{1mn}(k_{mn}\rho) \sin(m\phi) \quad (\text{B6})$$

$$\bar{g}_{21mn} = -\hat{\rho} k_{mn} Z'_{2mn}(k_{mn}\rho) \cos(m\phi) + \hat{\phi} \frac{m}{\rho} Z_{2mn}(k_{mn}\rho) \sin(m\phi) \quad (\text{B7})$$

$$\bar{g}_{22mn} = -\hat{\rho} k_{mn} Z'_{2mn}(k_{mn}\rho) \sin(m\phi) - \hat{\phi} \frac{m}{\rho} Z_{2mn}(k_{mn}\rho) \cos(m\phi) \quad (\text{B8})$$

$$Z_{1mn}(k_{mn}\rho) = C_{1mn} \left[J_m(k_{mn}\rho) Y'_m(k_{mn}a) - Y_m(k_{mn}\rho) J'_m(k_{mn}a) \right] \quad (\text{B9})$$

$$Z_{2mn}(k_{mn}\rho) = C_{2mn} \left[J_m(k_{mn}\rho) Y_m(k_{mn}a) - Y_m(k_{mn}\rho) J_m(k_{mn}a) \right] \quad (\text{B10})$$

$$Z'_{1mn}(k_{mn}\rho) = C_{1mn} \left[J'_m(k_{mn}\rho) Y'_m(k_{mn}a) - Y'_m(k_{mn}\rho) J'_m(k_{mn}a) \right] \quad (\text{B11})$$

$$Z'_{2mn}(k_{mn}\rho) = C_{2mn} \left[J'_m(k_{mn}\rho) Y_m(k_{mn}a) - Y'_m(k_{mn}\rho) J_m(k_{mn}a) \right] \quad (\text{B12})$$

$$C_{1mn} = \frac{\sqrt{\pi\epsilon_m}}{2} \left\{ \left[\frac{J'_m(k_{mn}a)}{J'_m(k_{mn}b)} \right]^2 \left[1 - \left[\frac{m}{k_{mn}b} \right]^2 \right] - \left[1 - \left[\frac{m}{k_{mn}a} \right]^2 \right] \right\}^{-1/2} \quad (\text{B13})$$

$$C_{2mn} = \frac{\sqrt{\pi\epsilon_m}}{2} \left\{ \frac{J_m^2(k_{mn}a)}{J_m^2(k_{mn}b)} - 1 \right\}^{-1/2} \quad (\text{B14})$$

The TEM mode function is

$$\bar{g}_0 = \frac{\hat{\rho}}{\rho} \left[2\pi \ln\left(\frac{b}{a}\right) \right]^{-1/2} \quad (\text{B15})$$

Finally, it is convenient to represent the transverse magnetic fields as

$$\bar{H}_i = Y_0 C_0 e^{\gamma_0 z} \times \bar{g}_0 - \sum_{p=1}^2 \sum_{q=1}^2 \sum_{m=0}^{\infty} \sum_{n=n_1}^{\infty} Y_{pmn} C_{pqmn} e^{-\gamma_{pmn} z} \times \bar{g}_{pqmn} \quad (\text{B16})$$

where Y_{pmn} is a modal admittance

$$Y_{pmn} = \frac{1}{\eta_0} \begin{cases} -\frac{j\gamma_{pmn}}{k_0 \mu_r}, & p=1 \text{ (TE)} \\ \frac{jk_0 \epsilon_r}{\gamma_{pmn}}, & p=2 \text{ (TM)} \end{cases} \quad (\text{B17})$$

The dominant mode's admittance is simply $Y_0=1/\eta$. A subtle but important point is that this modal admittance is not the same as the characteristic admittance [28]. (The characteristic admittance of the dominant mode in a coaxial cable is $2\pi/[\eta \ln(b/a)]$.)

B1.2. Computations

The IMSL[®] library routines are used to compute Bessel functions. Additional routines were created to implement formulas for the Bessel function derivatives.

Figures B1 and B2 show the TE and TM characteristic equations for various orders m (the ratio $b/a=7/3$ is typical of 50 Ω characteristic impedance). The problem of finding the roots, indicated by the circles, turned out to be non-trivial. Abramowitz & Stegun [Reference 10, p.374] give three terms of a series for the roots. Evidently, more terms are needed for orders higher than $m=3$ and/or $n=4$, because in those instances the calculated root was not close enough to the actual root for a Newton's method iteration to converge to the actual root. Unfortunately it is not possible to deduce what the remaining terms in the series are from the ones that are given. It is probably no coincidence that the tables of roots given in Marcuvitz [Reference 2, pp.74,78] contain blanks for those values where the 3-term series fails.

Marcuvitz gives approximate formulas for $m=0$ that are fairly accurate for all n . The procedure that was adopted uses this as a seed for Newton's method in order to find the $0,n$ root. That is used in

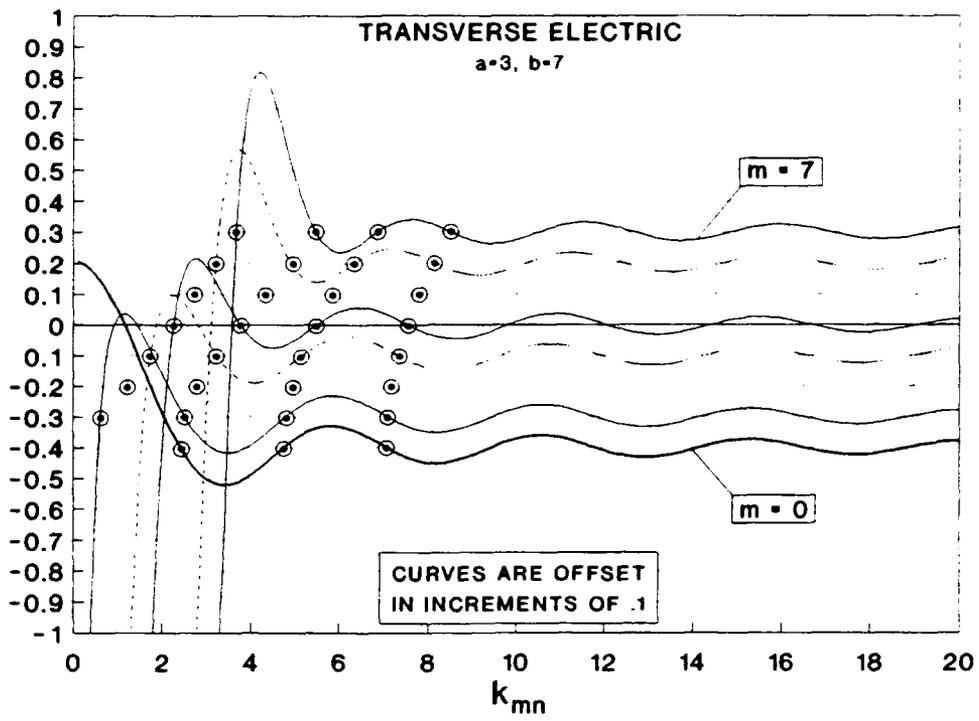


Figure B1. Transverse Electric Characteristic Functions

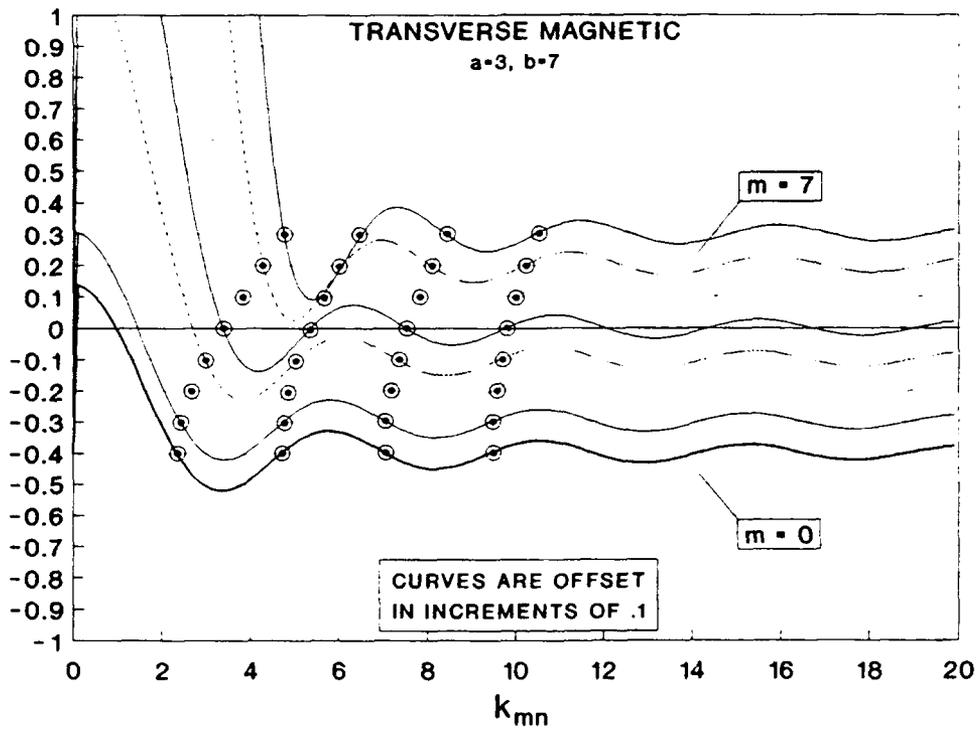


Figure B2. Transverse Magnetic Characteristic Functions

turn as a seed to find the $1,n$ root; and so on up to the desired m,n root. This method has reliably found all of the roots up to $m,n=6,10$. Figures B3 and B4 are results obtained using that routine to find the cutoff frequencies of higher-order modes in APC-7 coax ($a=1.502$ mm, $b=3.5$ mm). As expected, the first mode to become propagating is the TE_{11} mode, but its cutoff frequency of 19.5 GHz is well above the specified operating limit of 18 GHz for the APC-7 coaxial connectors.

Once the cutoff wavenumbers are known it is a tedious but straightforward matter to find the mode functions using Eqs.(B5)-(B14). Figures B5 and B6 are the results of example calculations. They are contour plots of the electric field lines for two higher order modes. These results agree with figures in Marcuvitz [Reference 2, pp.76,79].

The important feature of orthonormality was also verified by numerically integrating function products for various values of the indices p,q,m,n . The integral of any pair of mode functions over the entire waveguide cross section should equal 1 if the two functions are the same, and should equal 0 otherwise. The success of this test for all of the 30 mode pairs attempted verifies the accuracy of the normalization factors (B13) and (B14).

B2. CIRCULAR WAVEGUIDE

The procedures used for circular waveguide functions are substantially the same as those used for coaxial waveguides. There are again four types of mode functions, TE and TM with $\sin(m\phi)$ or $\cos(m\phi)$ dependence. The cutoff wavenumbers for the m,n modes, k_{mnp} , are the n 'th nonzero roots of $J'_m(ka)$ for TE ($p=1$), or of $J_m(ka)$ for TM ($p=2$), where a is the waveguide's inside wall radius. The same root-finding procedure discussed in the preceding section was used to determine the cutoff wavelengths. The mode functions are as follows:

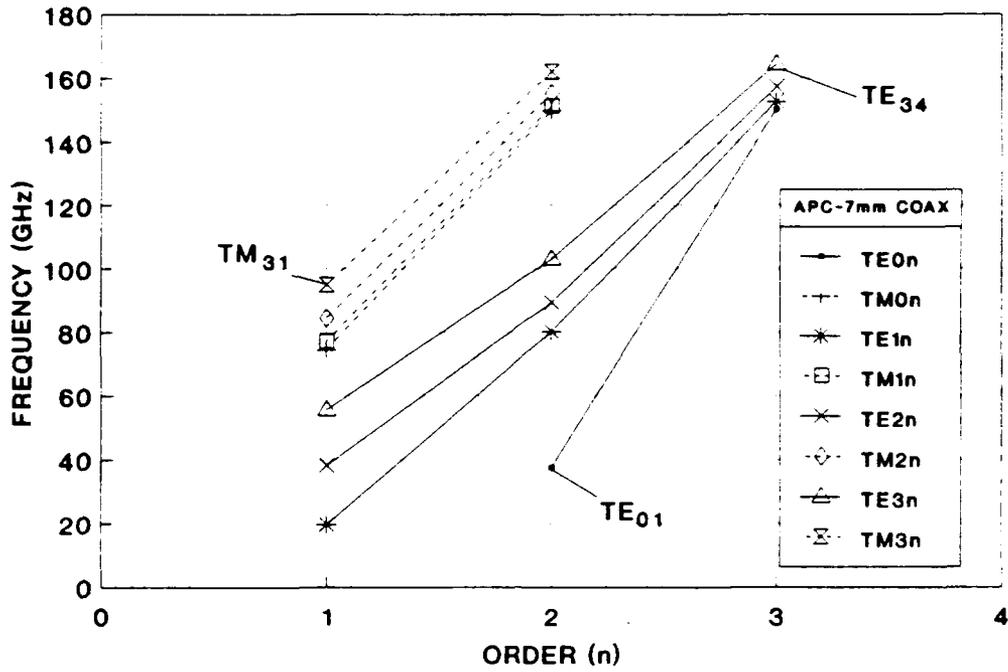


Figure B3. APC-7 Mode Cutoff Frequency vs. Index n

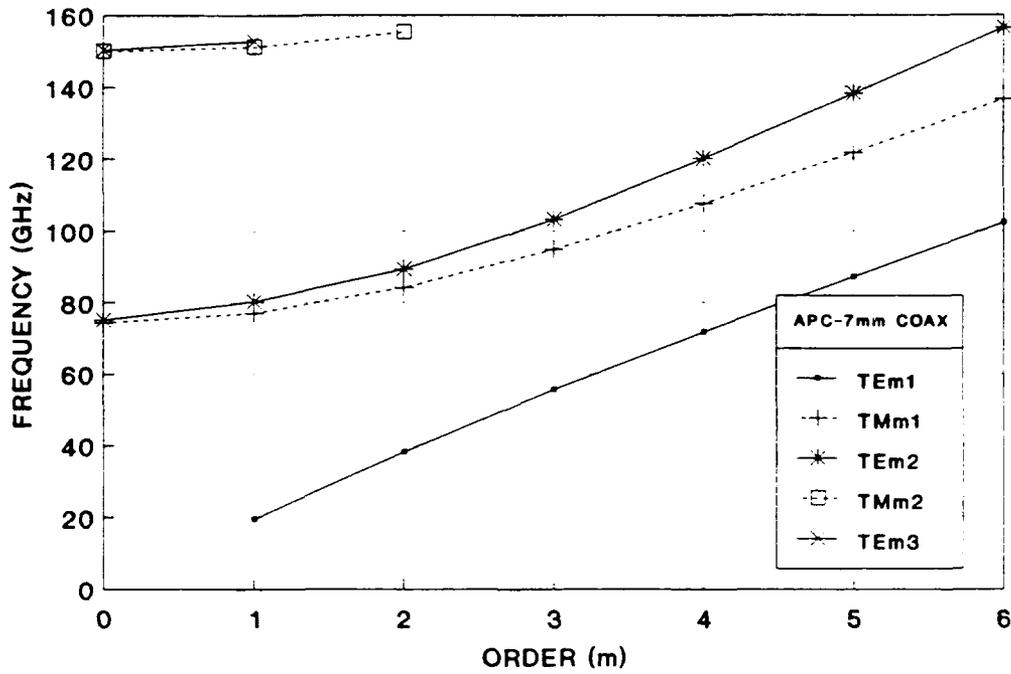


Figure B4. APC-7 Mode Cutoff Frequency vs. Index m

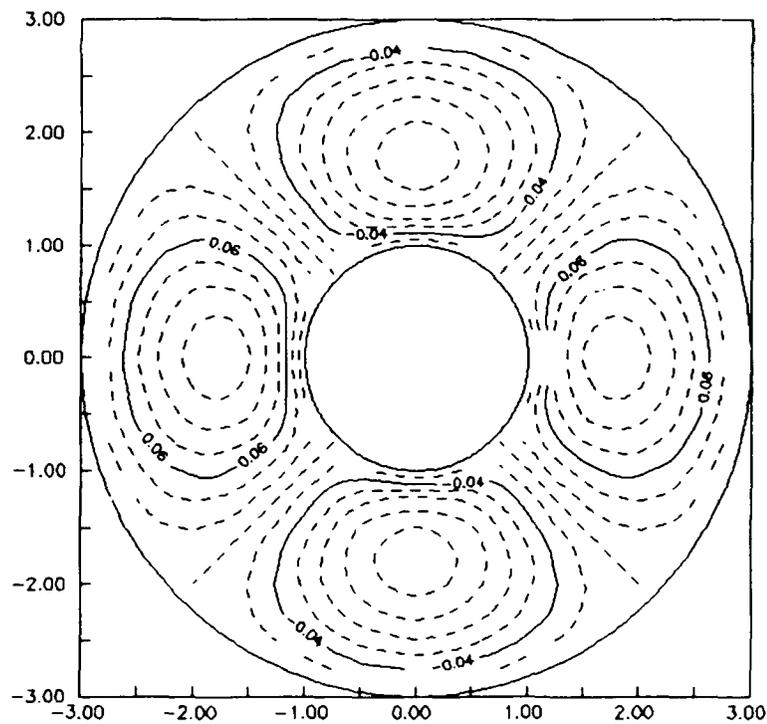
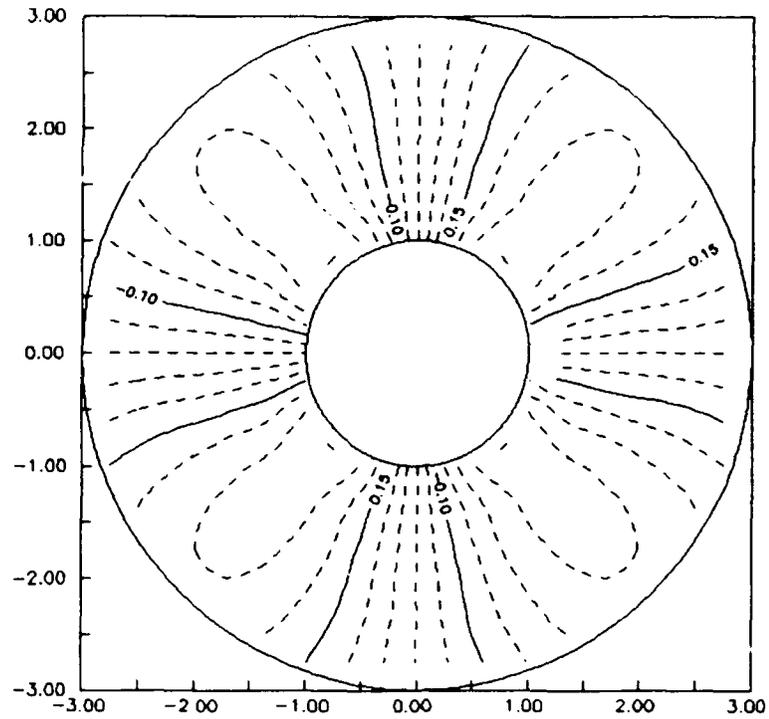


Figure B5. Electric Field Contours for TE21 Mode:
top, ρ -component; bottom, ϕ -component

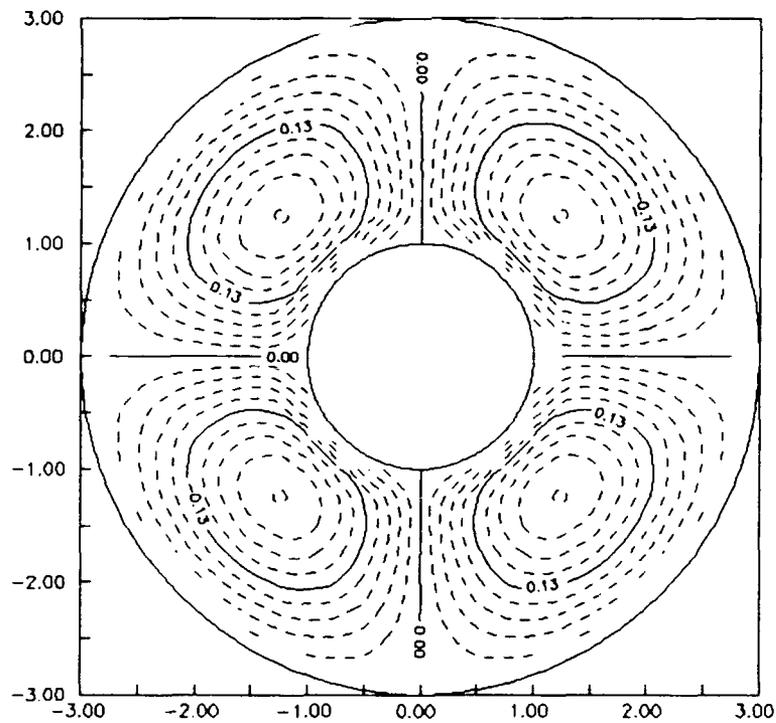
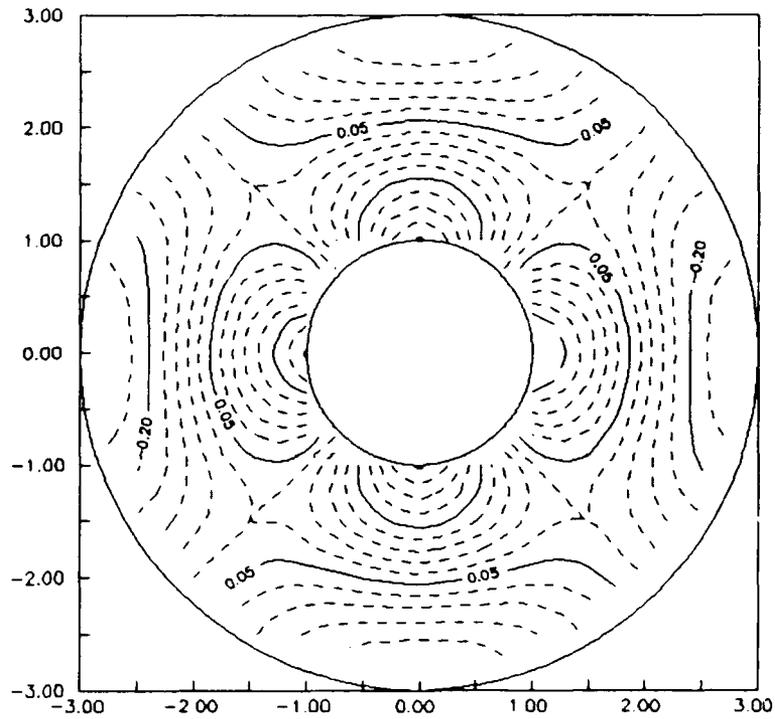


Figure B6. Electric Field Contours for TM₂₁ Mode:
top, ρ -component; bottom, ϕ -component

$$\bar{g}_{11mn} = C_{1mn} \left[\hat{\rho} \frac{m}{\rho} J_m(k_{mn}\rho) \sin(m\phi) + \hat{\phi} k_{mn} J'_m(k_{mn}\rho) \cos(m\phi) \right] \quad (\text{B18})$$

$$\bar{g}_{12mn} = C_{1mn} \left[-\hat{\rho} \frac{m}{\rho} J_m(k_{mn}\rho) \cos(m\phi) + \hat{\phi} k_{mn} J'_m(k_{mn}\rho) \sin(m\phi) \right] \quad (\text{B19})$$

$$\bar{g}_{21mn} = C_{2mn} \left[-\hat{\rho} k_{mn} J'_m(k_{mn}\rho) \cos(m\phi) + \hat{\phi} \frac{m}{\rho} J_m(k_{mn}\rho) \sin(m\phi) \right] \quad (\text{B20})$$

$$\bar{g}_{22mn} = C_{2mn} \left[-\hat{\rho} k_{mn} J'_m(k_{mn}\rho) \sin(m\phi) - \hat{\phi} \frac{m}{\rho} J_m(k_{mn}\rho) \cos(m\phi) \right] \quad (\text{B21})$$

$$C_{1mn} = \sqrt{\epsilon_m / \pi} \left[k_{mn}^2 a^2 - m^2 \right]^{-1/2} \frac{1}{J_m(k_{mn}a)} \quad (\text{B22})$$

$$C_{2mn} = \sqrt{\epsilon_m / \pi} \frac{1}{k_{mn} a J_{m+1}(k_{mn}a)} \quad (\text{B23})$$

Modes exist for all combinations of $m \geq 0$ and $n \geq 1$. The dominant mode is TE_{11} . Since it may have both $\sin(m\phi)$ and $\cos(m\phi)$ components, it is necessary to specify the polarization of the incident field, and the TWOPORT code includes an input parameter for that purpose.

Appendix C

Rectangular Waveguide Mode Functions and Inner Product Integrations

The rectangular waveguide mode functions are defined for a waveguide whose side walls are parallel to the x and y axes, with lengths a and b respectively. The inside cross sectional area is confined to the first quadrant so that two of the inner walls are in the x - z and y - z planes. The matrix terms relating interior electric field to waveguide modes are given in terms of products of integrals of the form

$$\Psi_{rmnp} = \int_{\square} \bar{\psi}_r \cdot \bar{g}_{mnp} ds \quad (\text{C.1})$$

where \square denotes the open portion of the waveguide end wall aperture, $\bar{\psi}_r$ is the vector expansion function associated with edge r and \bar{g}_{mnp} is the transverse mode function for the mn mode with $p=1$ for TE and $p=2$ for TM. Since $\bar{\psi}_r$ is only defined over the triangles adjacent to edge r (2 at most), the region of integration reduces to those triangles. Over a single triangle with index k , suppose that the edge r 's endpoints are the local nodes 1 and 2. Then substituting the definition of the expansion function Eq.(20) gives

$$\Psi_{rmnp}^{(k)} = L_r \int_{\Delta_k} (f_1 \nabla f_2 - f_2 \nabla f_1) \cdot \bar{g}_{mnp} ds \quad (C2)$$

where f_1 and f_2 are the linear scalar expansion functions associated with nodes 1 and 2, respectively.

Transforming to homogeneous coordinates gives $f_1 = t_1$ and $f_2 = t_2$, $dx dy = 2A dt_1 dt_2$ and

$$\nabla f_1 = \frac{1}{2A} (xT_{12} + yT_{13}) \quad (C3)$$

$$\nabla f_2 = \frac{1}{2A} (xT_{22} + yT_{23}) \quad (C4)$$

where A is the triangle area. The matrix T gives the coordinate transformation as in Eq. (16). The inverse of this transform gives x and y in terms of t_1 and t_2 as

$$x = \alpha_0 + \alpha_1 t_1 + \alpha_2 t_2 \quad (C5)$$

$$y = \beta_0 + \beta_1 t_1 + \beta_2 t_2 \quad (C6)$$

The transverse mode functions and normalization coefficients are given by Marcuvitz [Reference 15, pp.57-60]:

$$g_{mnp_x} = C_{mnp_x} \cos\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) \quad (C7)$$

$$g_{mnp_y} = C_{mnp_y} \sin\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right) \quad (C8)$$

$$C_{mn1x} = \left(\frac{n}{b}\right) \left[\frac{n^2 a}{2 b \epsilon_m} + \frac{m^2 b}{2 a \epsilon_n} \right]^{-1/2} \quad (C9)$$

$$C_{mn1y} = -\left(\frac{m}{a}\right) \left[\frac{n^2 a}{2 b \epsilon_m} + \frac{m^2 b}{2 a \epsilon_n} \right]^{-1/2} \quad (C10)$$

$$C_{mn2x} = -\left(\frac{m}{a}\right) \left[\frac{n^2 a}{4b} + \frac{m^2 b}{4a} \right]^{-1/2} \quad (C11)$$

$$C_{mn2y} = -\left(\frac{n}{b}\right) \left[\frac{n^2 a}{4b} + \frac{m^2 b}{4a} \right]^{-1/2} \quad (C12)$$

where ϵ_ν is Neumann's number.

Let the coefficients ξ_i and η_i represent the following combinations of α_i and β_i from Eqs.(C5) and (C6):

$$\xi_i = \frac{m\pi\alpha_i}{a} + \frac{n\pi\beta_i}{b} \quad (C13)$$

$$\eta_i = \frac{m\pi\alpha_i}{a} - \frac{n\pi\beta_i}{b} \quad (C14)$$

which allows the trigonometric functions in Eqs.(C7) and (C8) to be expressed in the convenient form

$$\cos\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) = \frac{1}{2} \sin(\xi_0 + \xi_1 t_1 + \xi_2 t_2) - \frac{1}{2} \sin(\eta_0 + \eta_1 t_1 + \eta_2 t_2) \quad (C15)$$

$$\sin\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right) = \frac{1}{2} \sin(\xi_0 + \xi_1 t_1 + \xi_2 t_2) + \frac{1}{2} \sin(\eta_0 + \eta_1 t_1 + \eta_2 t_2) \quad (C16)$$

Let H_ξ denote the integral

$$H_\xi = \frac{1}{2} \int_0^1 t_1 dt_1 \int_0^{1-t_1} \sin(\xi_0 + \xi_1 t_1 + \xi_2 t_2) dt_2 \quad (C17)$$

H_η is identical with H_ξ replacing ξ_i everywhere. H'_ξ and H'_η are the same, but with ξ_1 and ξ_2 or η_1 and η_2 reversed. Then finally, the inner product integral may be written as

$$\Psi_{r m n p}^{(k)} = L_r \left\{ T_{22} C_{m n p x} (H_\xi - H_\eta) + T_{23} C_{m n p y} (H_\xi + H_\eta) \right. \\ \left. - T_{12} C_{m n p x} (H'_\xi - H'_\eta) - T_{13} C_{m n p y} (H'_\xi + H'_\eta) \right\} \quad (C18)$$

(The transformation of the integral to homogeneous coordinates introduces a scale factor of $2A$ that cancels the $1/2A$ in Eqs.(C3) and (C4).) The integral in Eq.(C17) may be carried out in closed form, but it must be evaluated separately for each of the following cases (a) through (e):

(a) general case:

$$H_\xi = \frac{1}{2\xi_2} \left\{ \frac{1}{\xi_1^2} [\cos(\xi_0 + \xi_1) - \cos \xi_0 + \xi_1 \sin(\xi_0 + \xi_1)] \right. \\ \left. - \frac{1}{(\xi_1 - \xi_2)^2} [\cos(\xi_0 + \xi_1) - \cos(\xi_0 + \xi_2) + (\xi_1 - \xi_2) \sin(\xi_0 + \xi_1)] \right\} \quad (C19a)$$

(b) $\xi_2 \neq 0$, $\xi_1 \neq 0$, $|\xi_2 - \xi_1| \ll 1$ (for small values of $\xi_2 - \xi_1$, Eq.(C.19a) is numerically unstable, susceptible to large roundoff error and overflow):

$$H_\xi \approx \frac{1}{2\xi_2} \left\{ \frac{1}{\xi_1^2} [\cos(\xi_0 + \xi_1) - \cos \xi_0 + \xi_1 \sin(\xi_0 + \xi_1)] \right. \\ \left. - \frac{1}{2} \cos(\xi_0 + \xi_2) + \frac{1}{3}(\xi_1 - \xi_2) \sin(\xi_0 + \xi_1) \right\} \quad (C19b)$$

(c) $|\xi_2| \ll 1$:

$$H_\xi = \frac{1}{4\xi_1^4} \left\{ (4\xi_1 + 6\xi_2 - \xi_1^2 \xi_2) \cos(\xi_0) - (4\xi_1 + 6\xi_2) \cos(\xi_0 + \xi_1) \right. \\ \left. - (2\xi_1^2 + 4\xi_1 \xi_2) \sin(\xi_0) - 2(\xi_1^2 + \xi_1 \xi_2) \sin(\xi_0 + \xi_1) \right\} \quad (C19c)$$

(d) $|\xi_1| \ll 1$:

$$H_{\xi} \approx \frac{1}{2 \xi_2^4} \left\{ \frac{1}{2} \cos \xi_0 - \frac{1}{3} \xi_1 \sin \xi_0 \right. \\ \left. - \frac{1}{(\xi_1 - \xi_2)^2} [\cos(\xi_0 + \xi_1) - \cos(\xi_0 + \xi_2) + (\xi_1 - \xi_2) \sin(\xi_0 + \xi_1)] \right\} \quad (\text{C19d})$$

(e) $|\xi_2| \ll 1, |\xi_1| \ll 1$:

$$H_{\xi} = \frac{1}{12} \left\{ \sin \xi_0 + \left(\frac{\xi_1}{2} + \frac{\xi_2}{4} \right) \cos(\xi_0) \right\} \quad (\text{C19e})$$

Note that the last four are correct in the limits as those quantities specified as much less than unity go to zero. The computer code implementation gives precedence to the five formulas in reverse order, checking for condition (e) first, and executing (a) only when none of the others (d), (c) or (b) are true.

**MISSION
OF
ROME LABORATORY**

Rome Laboratory plans and executes an interdisciplinary program in research, development, test, and technology transition in support of Air Force Command, Control, Communications and Intelligence (C³I) activities for all Air Force platforms. It also executes selected acquisition programs in several areas of expertise. Technical and engineering support within areas of competence is provided to ESD Program Offices (POs) and other ESD elements to perform effective acquisition of C³I systems. In addition, Rome Laboratory's technology supports other AFSC Product Divisions, the Air Force user community, and other DOD and non-DOD agencies. Rome Laboratory maintains technical competence and research programs in areas including, but not limited to, communications, command and control, battle management, intelligence information processing, computational sciences and software producibility, wide area surveillance/sensors, signal processing, solid state sciences, photonics, electromagnetic technology, superconductivity, and electronic reliability/maintainability and testability.