

**INFINITE ELEMENTS FOR THE ANALYSIS
OF OPEN DIELECTRIC WAVEGUIDES**

Marc J. McDougall, B. Eng., B. Sc.

A thesis submitted to the Faculty of Graduate Studies and
Research in partial fulfillment of the requirements
for the degree of Master of Engineering

Department of Electrical Engineering
McGill University
Montréal, Canada
July, 1988

© Marc J. McDougall 1988

Abstract

The finite element method is extended to the modal analysis of open dielectric waveguides with the use of infinite elements which preserves the linearity of the method. Previously, infinite elements incorporating a single radially decaying trial function were used. Since the decay length has to be optimized for each mode by means of an outer iteration loop, these are called *optimized single decay* methods.

In the method of this thesis, called the *multiple fixed decay* method, each infinite element incorporates several radially decaying exponential trial functions. The outer optimization loop is eliminated and all modes corresponding to a given phase constant are calculated in one pass of the solver. In addition, these infinite elements improve the accuracy by providing more degrees of freedom with which to model each mode.

The three component curl-curl functional is used with a penalty term to eliminate spurious modes from the propagating mode frequency range.

Sommaire

La méthode des éléments finis est étendue à l'analyse modale de guides d'ondes ouverts diélectriques en utilisant des éléments infinis ce qui préserve la linéarité de la méthode.

Auparavant, les éléments infinis incorporant une seule fonction d'essai de décroissance radiale étaient utilisés. Etant donné que la longueur de la décroissance doit être optimisée pour chaque mode au moyen d'une boucle d'itération externe, celles-ci sont appelées méthodes optimisées de longueur unique.

Par la méthode présentée dans cette thèse, appelée la méthode de longueur multiple fixe, chaque élément infini incorpore plusieurs fonctions d'essai exponentielles à décroissance radiale. La boucle d'optimisation externe est donc éliminée et tous les modes correspondant à une constante de phase donnée sont calculés à la fois. De plus, ces éléments infinis améliorent la précision en offrant plus de degrés de liberté pour modéliser chaque mode.

Une formulation variationnelle à trois composantes est utilisée avec une fonction de pénalité afin d'éliminer les modes indésirables de l'intervalle de fréquence de modes propagés.

Acknowledgements

I would like to thank my supervisor, Dr. Jon Webb, for his invaluable assistance in the preparation of this thesis and his many contributions.

I would also like to thank Minou Mansouri for translating the abstract and Sussan Nassehi for providing software assistance. I also extend my thanks my colleagues in the Computational Analysis and Design Laboratory for their help and support.

Financial support from the Natural Science and Engineering Research Council and from the Centre de Recherches en Informatique de Montréal is gratefully acknowledged.

Table of Contents

Chapter One: Introduction and Literature Survey

1. Introduction	p.1
2. Literature Survey	p.2
2.1 Linear Methods	p.4
2.1.1 The Finite Element Method	p.4
The Virtual Boundary Technique	p.4
Scalar Ballooning	p.5
Conformal Mapping	p.6
Combined Finite Elements and Infinite Elements	p.6
2.1.2 Finite Differences	p.7
2.1.3 Telegraphist's Equations	p.7
2.2 Non-Linear Methods	p.8
2.2.1 Point Matching	p.8
2.2.2 Integral Equations	p.8
2.2.3 Finite Element Hybrid Methods	p.9
Combined Finite Elements and Function Expansion	p.9
Combined Finite Elements and Integral Equations	p.10
3. Thesis Summary	p.10

Chapter Two: Open Waveguide Theory

1. The Problem Definition	p.11
2. The Field Notation	p.12
3. Maxwell's Equations and the Constitutive Relations	p.14
4. The Perfect Electric and Magnetic Conductor Boundary Conditions	p.15
5. The Continuity Relations	p.15
6. The Curl-Curl Equation in \mathbf{H}	p.16
7. The Far-Field Boundary Conditions	p.16
8. The Variational Principle	p.18

Chapter Three: The Conventional Finite Element Method

1. The Isotropic Guide Case	p.21
2. The Unknown Field Components	p.21
3. The Raleigh-Ritz Method	p.23
4. The Finite Element Method	p.24
5. Triangular Finite Elements	p.24
6. The Field Vector Trial Functions	p.27
7. Local Matrix Assembly	p.29
8. The Application of the Boundary Conditions	p.31
9. Global Matrix Assembly	p.34

Chapter Four: Infinite Elements

1. Elements for the Infinite Regions	p.35
2. The Infinite Element Mapping and the Azimuthal Trial Function	p.35
3. The Radial Trial Functions	p.36
4. The Full Trial Function	p.39
5. The Integration of Trial Function Terms	p.40

Chapter Five: Results

1. The Fortran Program	p.44
2. The Slab Waveguide	p.46
3. The Circular Dielectric Waveguide	p.50
4. The Square Dielectric Waveguide	p.50
5. The Effect of Varying the Parameter s	p.60
6. A Comparison of the MFD and OSD Techniques	p.60

Chapter Six: Conclusions	p.62
---	------

Appendix I: Proof of Variational Principle 2.1	p.63
---	------

Appendix II: The Frequency Range of Spurious Modes	p.72
---	------

References	p.80
-----------------------------	------

Symbol List

β	the phase constant
f	the frequency
ω	the angular frequency $= 2\pi f$
ϵ, ϵ_r	the permittivity and relative permittivity
ϵ_0	the permittivity of free space
K	the relative permittivity tensor
μ, μ_r	the permeability and relative permeability
μ_0	the permeability of free space
c	the speed of light in vacuo $= \frac{1}{\sqrt{\epsilon_0 \mu_0}}$
k_0	the normalised frequency $= \frac{\omega}{c}$
Ω	the problem geometry in the x-y plane
$\partial\Omega_e$	the perfect electric conductor boundary contour
$\partial\Omega_o$	the perfect magnetic conductor boundary contour
$\partial\Omega_\infty$	the far-field boundary
x, y, z	cartesian coordinates
r, ϕ, z	cylindrical coordinates
$\zeta_1, \zeta_2, \zeta_3$	the finite element local coordinates
ξ, η	the infinite element local coordinates
e, h	the fields as a function of (x, y, z, t)
E, H	the fields as a function of (x, y, z)
E, H	the fields as a function of (x, y)
$\tilde{H} = [a_x \ a_y \ a_z]^T$	the trial function weight variables

Chapter One

Introduction and Literature Survey

1. Introduction

Waveguides which employ dielectric materials as a transmission media include devices of current interest such as optical fibres and planar guides. They are used as low-loss carriers of communication signals over long distances or as components of planar circuits at frequencies ranging from millimeter-wave to optical. Typical unbounded dielectric waveguide cross-sections are shown in Figure 1.1. The importance of these devices in the field of communications has provoked considerable research into the development of numerical methods for their analysis. Indeed, the task of designing a dielectric waveguide can be made easier and less costly if its high frequency behaviour can be predicted prior to its fabrication.

The electromagnetic waves which are carried by unbounded dielectric waveguides without radiation loss are called *surface* waves and they assume distinct field patterns called modes. Each mode has a cut-off frequency which is the lowest frequency at which it will propagate and remain guided. The mode with the lowest cut-off frequency is called the fundamental mode. The modal wave properties that are of the most interest to the designer of dielectric waveguides are the dispersion characteristics and the field distributions. The dispersion characteristics of a mode are usually presented in the form of a dispersion curve which is a plot of the normalised frequency (or free-space wavenumber) k_0 versus the phase constant β for each mode of interest (see Figures 5.2, 5.4, 5.6). Given the dispersion curves of the fundamental mode and the next lowest mode of a particular device for example, the frequency interval in which conversion of energy from the fundamental mode to higher modes cannot occur can be determined; it is the interval between the cut-off frequencies of these two modes. The phase and group velocities of the modes can also be determined from the dispersion curves.

A knowledge of the field distribution of a mode is also important; it can indicate for example the extent to which the energy of the surface wave supported by the guide

is concentrated in the core region. It can also be used to estimate ohmic losses suffered by the guide, if they are small, with a perturbation method (first the fields of an equivalent lossless guide are calculated, then these values are used to calculate the losses by considering any finite conductivity of the dielectric materials or existing conductors).

Various useful dielectric waveguide structures have been proposed or fabricated which because of the exploitation of desirable properties or the manufacturing process, have irregular shapes or permittivity distributions. Finding a closed-form analytical expression for the modes of most these unbounded guides is usually impossible unless the waveguide geometry and permittivity profile are identical with the coordinate curves of a coordinate system in which the Helmholtz equation is separable. The only recourse for solving the general dielectric waveguide problem is the application of a numerical method.

This thesis presents a numerical method for the modal analysis of a general class of translationally-symmetric dielectric waveguides. This class includes transversally unbounded dielectric guides with arbitrarily-shaped cross-sections and arbitrarily-inhomogeneous permittivity profiles. It is restricted however to sourceless, lossless, isotropic guides with uniform permeability $\mu = \mu_0$. All conductors are assumed to be perfect. For brevity, members of this class shall be referred to in this thesis simply as *open guides*.

The numerical method employed is the finite element method using a three component magnetic vector functional and using infinite elements to handle the unbounded nature of open guides. It yields the frequencies of the modes which correspond to a given phase constant value (these can be used to plot dispersion curves) and the field distribution.

2. Literature Survey

The body of literature related to the modal analysis of translationally symmetric dielectric waveguides is enormous. This survey will therefore be limited to numerical methods which are capable of modelling any open guide as defined above. The vari-

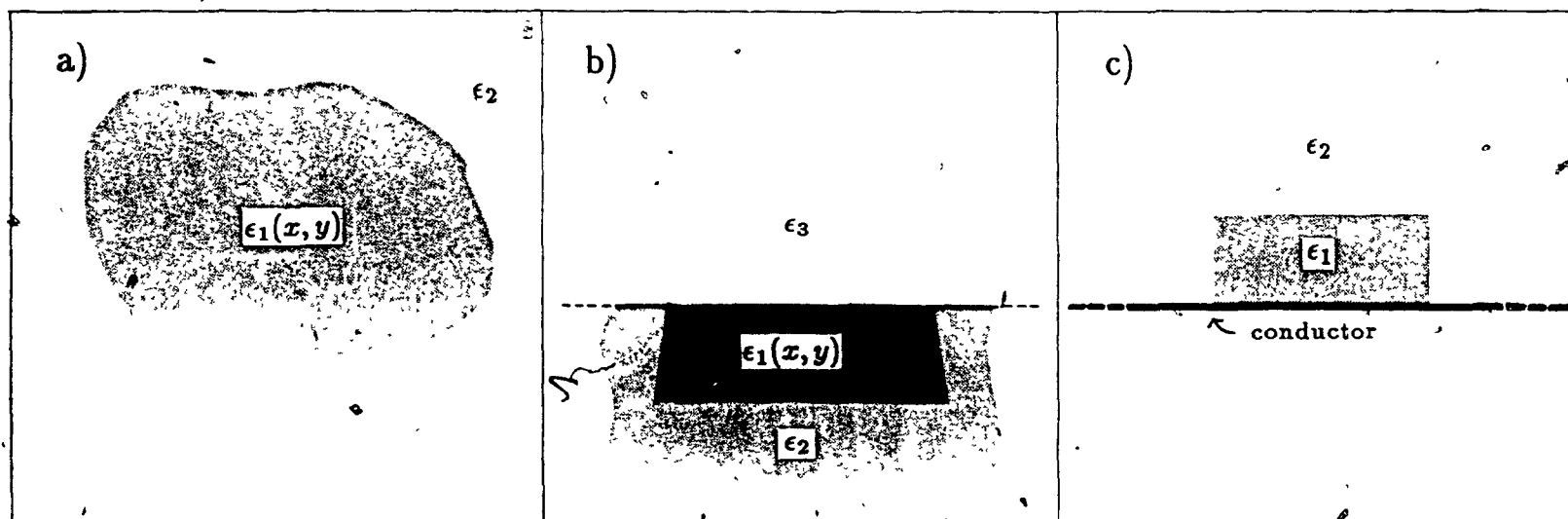


Figure 1.1 Three examples of unbounded dielectric waveguides:- a) arbitrarily-shaped arbitrarily-inhomogeneous fibre guide, b) diffused channel guide and c) image guide.

ous methods will be categorised according to whether the final approximate numerical problem is linear or non-linear (defined below), and according to the particular scheme adopted to model the infinite extent surrounding the core region.

In the following, linear methods are defined as those in which the final algebraic problem can be reduced to the generalised eigenvalue problem, $Ax = \lambda Bx$. These methods can benefit from specialised, efficient algorithms [1,2] which return an exhaustive list of the eigenvalues.

With non-linear methods, the final algebraic problem has the form $\det(C) = 0$ where the elements of the matrix C are non-linear functions of the free space wavenumber k_0 . Methods of this type must use computationally costly root-finding algorithms and the resulting solutions must then be checked carefully to ensure that none of the roots have been missed.

2.1 Linear Methods

2.1.1 The Finite Element Method

The finite element method (FEM) [3,4] consists in dividing the problem region into small sub-domains over which trial function expansions are defined. The unknown coefficients in the expansions are determined by a variational method in which the trial functions are inserted into a functional whose stationary points are solutions to the defining differential equations.

The FEM was applied to the modal analysis of closed homogeneous waveguide problems by Silvester [5] who used polynomial trial functions and a functional for the scalar Helmholtz equation. Because of its great flexibility and relative simplicity, the FEM has since been applied extensively to a wide variety of closed waveguide problems [6,7,8,9,10]. The following methods extend the FEM to open guides while preserving linearity.

The Virtual Boundary Technique

Yeh et al applied the unmodified FEM to the modal analysis of open waveguides in 1975 [11]. The modelling of the infinite extent was performed by simply surrounding

the guide with a perfectly conducting cylinder, called a virtual boundary, whose radius was large relative to the dimensions of the core. The entire problem, that is the interior of the core and the exterior region between the core and the conducting cylinder was then meshed with regular finite elements and the modes were solved using a $E_z - H_z$ functional. The authors noted that a gradual increase in the radius of the cylinder resulted in a convergence of the calculated modes to the actual solutions. This effect was greatest near cutoff where the exterior fields decay very slowly and thus the radius of the cylinder had to be made very large to obtain accurate results.

A similar method was employed by Ikeuchi et al. [12] who added an outer iterative loop which gradually increased the radius of the virtual boundary (on which E_z was constrained to be zero) until the maximum value of H_z on the boundary became small relative to its maximum value in the entire region. In this way the inaccuracy due to the presence of the artificial virtual boundary could be estimated and the solutions could be obtained to any desired level of accuracy except near cut-off where the region to be meshed became prohibitively large.

Welt and Webb [13] chose the location of the virtual boundary for arbitrarily shaped guides by first calculating the field decay rate for an equivalent homogeneous dielectric rod and placing the virtual boundary where its field was negligible. This method was found to be suitable for all the cases studied.

In all applications of the virtual boundary method, an improvement of accuracy is obtained at the expense of an increase in the mesh size, especially near cut-off. This and the concomitant increase in computing time and memory requirement are the major drawbacks of this method

Scalar Ballooning

The ballooning method was first applied by Silvester et al. [14] to open boundary Laplace field problems. In this technique, the finite interior region is meshed with regular finite elements while the exterior region is meshed iteratively in successive identical annulus-shaped regions. Each iteration doubles the exterior mesh area but does not increase the global matrix size. A quasi-infinite area can therefore be modelled in only

a few (about eight) iterations.

Chiang [15,16] applied the same method to weakly guiding fibres using an approximate scalar functional. The pre-supplied constant which is usually β or β/k_0 with the other methods is neither for this method. It is rather a variable chosen specially to eliminate the eigenvalue variable k_0 from the functional contributions of the exterior mesh. This allows for the recursive meshing and global matrix assembly of the outer annuli. At the present, this method cannot be applied to the general open guide problem which requires a vector functional and cannot be accurately approximated by a scalar functional.

Conformal Mapping

By conformally mapping the quarter plane into a half-circle, Wu and Chen [17] were able to transform the infinite extent surrounding a waveguide into a finite region. The FEM was then applied to discretise the finite region, and accurate results were obtained with an $E_z - H_z$ formulation. This particular application of the conformal mapping method is limited however to guides having a plane of symmetry.

Combined Finite Elements and Infinite Elements

Another method of solving unbounded waveguide problems is to define *infinite elements* which extend radially to infinity [18,19,20,21]. To correctly model the decaying behaviour of the actual solutions to the Helmholtz equation, each infinite element must incorporate one or more decaying trial functions in the infinite direction. This approach using one decaying exponential trial function in each infinite element was used by Yeh et al. [22] with an $E_z - H_z$ functional and later by Rahman and Davies [23] with a three component vector functional.

The main disadvantage of this method is that the correct value for the parameter which specifies the infinite element decay rate cannot be determined before the problem is solved, nor can it figure as one of the unknown variables to be solved for, because of the non-linear dependence of the functional on the decay parameter value. If a generalised eigenvalue solver is to be used, the decay parameter must be supplied beforehand as a

fixed constant. Since the accuracy of the solutions near the cut-off is critically dependent on the value of the decay parameter, Yeh and also Rahman and Davies found it necessary to optimize the decay parameter for each mode through the addition of an outer iteration loop to the FEM program. In this thesis, not one, but a range of exponential decaying trial functions is used in each infinite element, allowing the variational principle to weight them to match the actual decay as closely as possible without optimization.

2.1.2 Finite Differences

In the finite difference method the problem region is first discretised into a rectangular grid of nodes. Then the differential operators of the functional or the defining equations themselves are approximated using difference calculus which produces an algebraic system of equations. Schweig and Bridges [24] applied this method to open guides by approximating the integrand of the $E_z - H_z$ functional with finite differences. Since the problem was enclosed in a perfectly conducting box as in the virtual boundary method, the accuracy near cut-off was compromised.

A major disadvantage with the finite difference method is that if the mesh must be refined in a small region, extra nodes must be added along the length of the entire problem region which increases the matrix problem size unnecessarily. With the FEM, the mesh can be refined locally in a small region and the modelling of oblique boundaries is facilitated since the nodes are not constrained to lie in a rectangular grid.

2.1.3 Telegraphist's Equations

This method consists in enclosing the open guide in a large rectangular waveguide and approximating the modes by a linear combination of the modes of an equivalent homogeneous rectangular waveguide. The Telegraphist's equations are derived directly from Maxwell's equations and the orthogonality properties of the closed waveguide modes [25,26,27]. This method is similar to a virtual boundary technique in that the problem is enclosed in a perfectly conducting boundary; it is therefore inaccurate near the cut-off frequencies of the modes.

2.2 Non-Linear Methods

2.2.1 Point Matching

First applied to homogeneous unbounded rectangular waveguides by Goell in 1969 [28], the point matching method involves expressing the longitudinal electric and magnetic field components (E_z and H_z) in the interior and exterior regions of the guide in four truncated series of Bessel functions which are solutions to the Helmholtz equation. The unknown coefficients of the series are found by matching the tangential components of these field expressions at N points on the interface of the two regions and solving the resulting characteristic equation for the eigenvalues.

The method has been applied to unbounded guides having more exotic shapes [29,30,31,32,33] and composite dielectric profiles [34,35]. Its application to the general open guide is not straightforward however; increasing the number of matching points, which adds more terms to the truncated series representation of the fields, will not necessarily cause the eigenvalues to converge to the exact solutions in the general case [29,36]. Problems such as this in establishing the validity of the series representation make the point matching method difficult to apply to open guides.

2.2.2 Integral Equations

By constructing equivalent electric and magnetic current densities from the inhomogeneities of the dielectric fibre, de Ruiter [37] arrived at a set of surface integral equations similar to those representing free-space electromagnetic fields with volume source terms. To solve for the modal free-space wavenumbers, a system of equations is derived by applying the moment method and a point-matching technique.

Green's functions were used in the derivation, and therefore the integral equations contained Hankel functions whose argument was the free-space wavenumber. Thus the problem was non-linear and was solved by an iterative complex zero search algorithm.

For the analysis of an open guide consisting of one homogeneous dielectric cylinder of arbitrary shape embedded in a single medium, Su [38] devised a method in which the integral equations are derived from Green's second identity. As with previous scalar

integral equation methods, the resulting equations contain a contour integral which is difficult to evaluate since the closed contour includes Green's function singularities.

2.2.3 Finite Element Hybrid Methods

Various schemes have been proposed to couple the FEM with other methods more suited to unbounded regions. They combine the flexibility of the FEM in handling arbitrarily-shaped and arbitrarily-inhomogeneous finite regions with the superior ability of other methods in modelling homogeneous infinite regions.

Combined Finite Elements and Function Expansion

Oyamada and Okoshi [39] divided the fibre into two regions separated by a circular boundary. The finite region inside the circle contains all the inhomogeneities of the fibre while the exterior region is homogeneous and unbounded. An $E_z - H_z$ functional similar to that of Yeh [11] is used to solve for the modes of the combined region. The contribution of the interior region to the functional is supplied by regular finite elements. For the exterior region, the E_z and H_z variables are expanded into two truncated series of modified Bessel functions which are then substituted into the functional.

By relating the Bessel function weights to the nodal field values on the circle through a discrete Fourier transform, the functional for the combined region can be expressed in terms of the nodal field values in the interior region and the free space wavenumber only. Due to the presence of the Bessel functions, the functional is a non-linear function of the free-space wavenumber k_0 . For this reason, a non-linear root finding technique must be used to find the modal solutions which are the minima of the functional.

Instead of the $E_z - H_z$ functional, Wu and Chen [40] employed a variational principle based on the reaction concept [41][42], which solves for the scattered field from an obliquely illuminated dielectric waveguide. This method is otherwise very similar to that of Oyamada and Okoshi (above) except that the modal frequencies are the poles of the scattering coefficients.

Combined Finite Elements and Integral Equations

Integral equation methods such as the one presented by Su [38] (above) are able to efficiently model the infinite cladding surrounding a homogeneous guide but they cannot be applied to inhomogeneous guides for lack of a convenient Green's function. By using finite elements in the interior region combined with integral equations, Su [43] effectively eliminated two drawbacks of the previous method. First, inhomogeneous core regions could be accommodated with the application of the FEM, and second, the hybrid method did not require the integration of a contour integral containing Green's function singularities. The method, however, remains non-linear.

3. Thesis Summary

The remainder of this thesis is organised as follows. Chapter 2 presents, for the general lossless reciprocal anisotropic open guide case, the theory required later for the Finite/Infinite Element Method of Chapters 3 and 4 in which only isotropic guides are considered. Chapter 3 is an introduction to the finite element method for closed guides and Chapter 4 extends this method to open guides through the use of infinite elements. Chapter 5 contains results for several open guide examples; Chapter 6 is the conclusion. Appendices I and II contain proofs required by Chapter 2.

Chapter Two

Open Waveguide Theory

1. The Problem Definition

Consider a dielectric waveguide with a given cross-sectional geometry and permittivity profile. The guide is translationally symmetric along the z axis; the transverse plane is the $x-y$ plane. The geometry may include unbounded homogeneous regions (if the guide is open) and conductors. Figure 2.1 is an example of an open dielectric waveguide. In the transverse plane, the problem geometry is given the symbol Ω . In some cases, such as the one in Figure 1.1a, the Ω surface may consist of the entire $x-y$ plane.

The problem can include three types of boundaries. All conductors are assumed to be perfect; the skin depth is therefore zero and only the contours of the conductors, represented by $\partial\Omega_c$, need be considered. The problem can also include perfect magnetic conductors contours $\partial\Omega_o$ which are used to exploit symmetry and reduce the problem. The last contour, $\partial\Omega_\infty$, is used to constrain the fields at $r = \infty$ where $r = \sqrt{x^2 + y^2}$. It is defined as $\partial\Omega_\infty = \Omega \cap (\lim_{R \rightarrow \infty} r = R)$ which is the part of a circle of infinite radius, centered at the origin, that is inside the problem region.

The materials which compose the guide can be anisotropic but are assumed to be lossless and reciprocal everywhere. The permittivities of all materials are assumed to be independent of frequency over the range of frequencies considered. The permeability is uniform everywhere and equal to that of free-space.

In the transverse plane, the homogeneous regions can be either finite or infinite. The *core* region is defined simply as the smallest region which includes all finite homogeneous regions. The *guided modes* of the guide are defined to be the electromagnetic waves with time and z dependence $e^{j(\omega t - \beta z)}$ (where ω and β are real numbers) and whose power flow is in the positive z direction only. Also, the energy of guided modes is confined to the core region of the guide which implies that the energy densities of the magnetic and electric fields, and therefore the field components, decay with distance from the core. The guided modes can be thought of as being excited by a source at

$z = -\infty$ and propagating without loss since time $t = -\infty$.

Let $\mathbf{e}_i = \mathbf{e}_i(x, y, z, t)$ and $\mathbf{h}_i = \mathbf{h}_i(x, y, z, t)$ represent the electric and magnetic fields corresponding to the i^{th} mode supported by a given open guide. The mode has phase constant β , angular frequency ω_i , and normalised frequency k_{0i} , defined as follows:

$$\beta = \frac{2\pi}{\lambda} \quad (2.1)$$

$$\omega_i = 2\pi f_i \quad (2.2)$$

$$k_{0i} = \frac{\omega_i}{c} \quad (2.3)$$

where $c = 1/\sqrt{\mu_0\epsilon_0}$ is the speed of light in vacuo, f_i is the frequency of mode i and λ is the wavelength.

The problem considered in this thesis is as follows. Given a pre-supplied value of the phase constant β and a description of the geometry and material properties of an open guide, determine the normalised frequencies $(k_{01}, k_{02}, \dots, k_{0p})$ and magnetic field distributions $(\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_p)$ of the p lowest guided modes (i.e. $k_{01} \leq k_{02} \leq \dots \leq k_{0p}$). In developing the necessary theory and later determining the modes of a particular waveguide, it is assumed that the waveguide in question is indeed capable of supporting guided modes. A proof of the existence of guided modes for the general open guide case will not be attempted.

As the theory will later demonstrate, if the electric field distribution is required, it is more convenient to determine the modal magnetic field distributions $(\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_p)$ first and then to calculate from these the $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p)$ modal electric field distributions.

2. The Field Notation

Since the fields of all modes are harmonic in both time and the longitudinal coordinate z , \mathbf{e} and \mathbf{h} can be expressed as the real parts of complex fields as follows:

$$\begin{aligned} \mathbf{e}(x, y, z, t) &= \text{Re}\{\mathbf{E}(x, y, z)e^{j\omega t}\} = \text{Re}\{\mathbf{E}(x, y)e^{j(\omega t - \beta z)}\} \\ \mathbf{h}(x, y, z, t) &= \text{Re}\{\mathbf{H}(x, y, z)e^{j\omega t}\} = \text{Re}\{\mathbf{H}(x, y)e^{j(\omega t - \beta z)}\} \end{aligned} \quad (2.4)$$

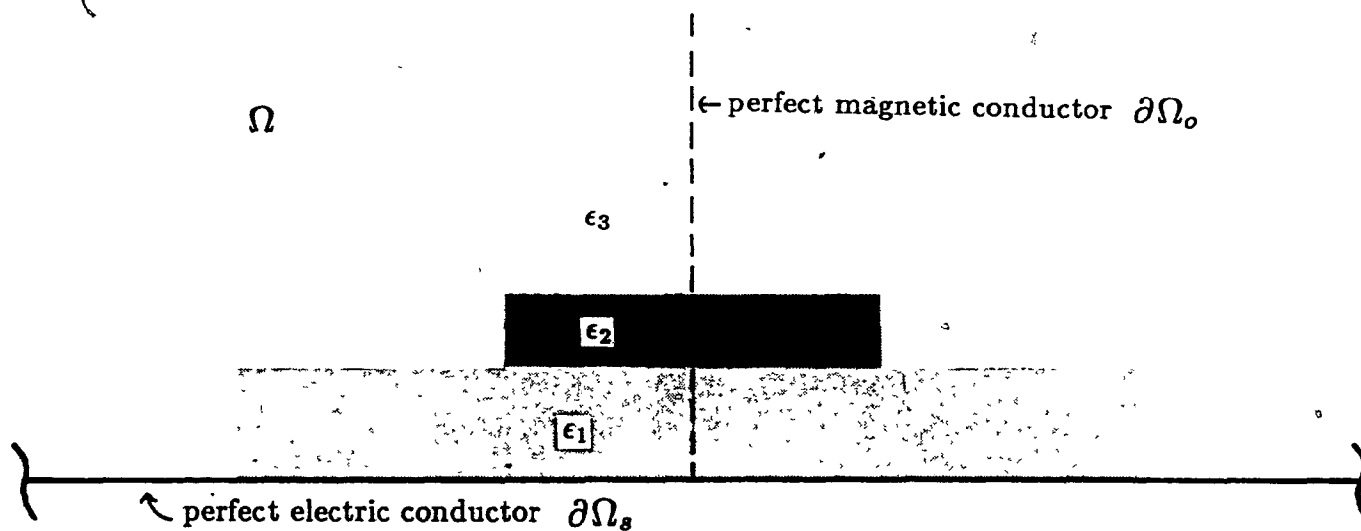


Figure 2.1 An open dielectric waveguide example.

In these relations, the variables (\mathbf{E}, \mathbf{H}) and (\mathbf{E}, \mathbf{H}) are implicitly defined as three component vectors having complex valued components and representing the (x, y, z) and (x, y) dependence of the fields respectively:

$$\mathbf{E}(x, y, z) = E_x(x, y, z)\mathbf{i} + E_y(x, y, z)\mathbf{j} + E_z(x, y, z)\mathbf{k}$$

$$\mathbf{H}(x, y, z) = H_x(x, y, z)\mathbf{i} + H_y(x, y, z)\mathbf{j} + H_z(x, y, z)\mathbf{k}$$

$$\mathbf{E}(x, y) = E_x(x, y)\mathbf{i} + E_y(x, y)\mathbf{j} + E_z(x, y)\mathbf{k}$$

$$\mathbf{H}(x, y) = H_x(x, y)\mathbf{i} + H_y(x, y)\mathbf{j} + H_z(x, y)\mathbf{k}$$

Although only the transverse dependence $(\mathbf{E}(x, y), \mathbf{H}(x, y))$ of the fields is unknown, the theory derived in this chapter will be in terms of (\mathbf{E}, \mathbf{H}) for notational convenience. The finite and infinite element theory of Chapter 3 will make use of the (\mathbf{E}, \mathbf{H}) functions. To facilitate integration later, the (\mathbf{E}, \mathbf{H}) fields are defined as being identically 0 outside Ω .

3. Maxwell's Equations and the Constitutive Relations

The theory of open waveguides begins with Maxwell's equations [44,p.17] for sourceless problems expressed in terms of the (\mathbf{E}, \mathbf{H}) variables:

$$\nabla \times \mathbf{E} = -j\omega\mathbf{B} \quad (2.5)$$

$$\nabla \times \mathbf{H} = j\omega\mathbf{D} \quad (2.6)$$

$$\nabla \cdot \mathbf{D} = 0 \quad (2.7)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2.8)$$

and the constitutive relations [44,p.21]:

$$\mathbf{B} = \mu \mathbf{H} = \mu_0 \mathbf{H} \quad (2.9)$$

$$\mathbf{D} = \begin{cases} \epsilon \mathbf{E} = \epsilon_r(x, y) \epsilon_0 \mathbf{E} & \text{for isotropic media} \\ K \mathbf{E} = K(x, y) \epsilon_0 \mathbf{E} & \text{for anisotropic media} \end{cases} \quad (2.10)$$

The relative permittivity tensor K is a real symmetric matrix [45,p.314] for lossless reciprocal dielectric media and therefore can be diagonalised by a rotation of the coordinate axes:

$$D_x = \kappa_{xx} E_x \quad D_y = \kappa_{yy} E_y \quad D_z = \kappa_{zz} E_z$$

The scalars $\kappa_{xx}, \kappa_{yy}, \kappa_{zz} > 0$ are called the *principal dielectric constants* of the material. They are also the eigenvalues of K , since the coordinate transformation is unitary [46,p.235]. The eigenvalues of K^{-1} are $\frac{1}{\kappa_{xx}}, \frac{1}{\kappa_{yy}}, \frac{1}{\kappa_{zz}} > 0$ (because $KE = \kappa E \Rightarrow K^{-1}E = \frac{1}{\kappa}E$). Since the eigenvalues of K^{-1} are non-zero and positive, it follows that K^{-1} is positive definite [46,p.250]. This result will be required by Variational Principle 2.1 (see below).

4. The Perfect Electric and Magnetic Conductor Boundary Conditions

The open guide cross-section Ω can include perfect electric conductor boundaries $\partial\Omega_s$ and perfect magnetic conductor boundaries $\partial\Omega_o$. The $\partial\Omega_s$ boundary condition follows from the physics of perfect conductors [44,p.36] whereas the $\partial\Omega_o$ boundary, also known as a *symmetry wall*, is an useful mathematical artifice which will be employed to reduce the problem size by exploiting symmetry.

In terms of the (\mathbf{E}, \mathbf{H}) variables, the boundary conditions are:

$$\mathbf{n} \times \mathbf{E} = 0 \quad \text{on } \partial\Omega_s \quad (2.11)$$

$$\mathbf{n} \cdot \mathbf{H} = 0 \quad \text{on } \partial\Omega_s \quad (2.12)$$

$$\mathbf{n} \cdot \mathbf{E} = 0 \quad \text{on } \partial\Omega_o \quad (2.13)$$

$$\mathbf{n} \times \mathbf{H} = 0 \quad \text{on } \partial\Omega_o \quad (2.14)$$

5. The Continuity Relations

From the integral forms of (2.5), (2.6), (2.7) and (2.8), and the underlying physics the continuity conditions

$$\mathbf{n} \times \mathbf{E}_1 = \mathbf{n} \times \mathbf{E}_2 \quad (2.15)$$

$$\mathbf{n} \times \mathbf{H}_1 = \mathbf{n} \times \mathbf{H}_2 \quad (2.16)$$

$$\mathbf{n} \cdot \mathbf{D}_1 = \mathbf{n} \cdot \mathbf{D}_2 \quad (2.17)$$

$$\mathbf{n} \cdot \mathbf{B}_1 = \mathbf{n} \cdot \mathbf{B}_2 \quad (2.18)$$

can be derived [47,p.149]. They apply along any sourceless boundary between two materials where \mathbf{n} is a unit vector normal to the boundary and $(\mathbf{E}_1, \mathbf{E}_2)$ and $(\mathbf{H}_1, \mathbf{H}_2)$ are

the fields infinitesimally close to the boundary on either side of it. The last two become, upon substitution of the constitutive relations (2.9) and (2.10):

$$\mathbf{n} \cdot (K_1 \mathbf{E}_1) = \mathbf{n} \cdot (K_2 \mathbf{E}_2) \quad (2.19)$$

$$\mathbf{n} \cdot (\mu_0 \mathbf{H}_1) = \mathbf{n} \cdot (\mu_0 \mathbf{H}_2) \quad (2.20)$$

where K_1, K_2 is the relative permittivity tensor on either side of the boundary. Since μ_0 is a constant, the second relation becomes:

$$\mathbf{n} \cdot \mathbf{H}_1 = \mathbf{n} \cdot \mathbf{H}_2 \quad (2.21)$$

Normal \mathbf{H} is therefore continuous along any material boundary whereas normal \mathbf{E} is discontinuous if the permittivity of the two materials is different.

6. The Curl-Curl Equation in \mathbf{H}

Since the \mathbf{H} field, from (2.16) and (2.21), is continuous across any material boundary, it is more convenient numerically to specify the problem in term of \mathbf{H} only. By substituting an expression for \mathbf{E} obtained from (2.6) into equation (2.5) applying the constitutive relations we can derive:

$$\nabla \times (K^{-1} \nabla \times \mathbf{H}) - k_0^2 \mathbf{H} = 0 \quad (2.22)$$

This is the anisotropic vector Helmholtz equation or curl-curl equation in \mathbf{H} .

7. The Far-Field Boundary Conditions

Consider the boundary value problem specified by the perfect electric and magnetic conductor boundary conditions and the vector Helmholtz equation in \mathbf{H} . This boundary value problem will admit solutions whose field components grow radially. These are not guided modes and they must be eliminated from the solution space by imposing additional constraints on the solutions in the far-field region.

To determine suitable far-field boundary conditions, consider any transverse slice $z_1 \leq z \leq z_2$ of a given isotropic open guide where $z_1 \neq z_2$. Let the energy stored in the

electric and magnetic fields inside the slice be denoted by \mathcal{E} . Since the mathematical model being developed is for physical guides excited by sources having finite power output, it is reasonable to require that \mathcal{E} be finite. The search for modes will therefore be restricted to those which satisfy

$$\begin{aligned}\mathcal{E} &= \int_{z=z_1}^{z=z_2} \int_{\Omega} \left(\epsilon \frac{\mathbf{E}^* \cdot \mathbf{E}}{2} + \mu \frac{\mathbf{H}^* \cdot \mathbf{H}}{2} \right) dx dy dz \\ &= \int_{z=z_1}^{z=z_2} \int_{\phi=0}^{\phi=2\pi} \int_{r=0}^{r=\infty} \left(\epsilon \frac{\mathbf{E}^* \cdot \mathbf{E}}{2} + \mu \frac{\mathbf{H}^* \cdot \mathbf{H}}{2} \right) r dr d\phi dz \\ &< \infty\end{aligned}$$

The integration is carried out over the entire transverse plane, not just Ω . This is justified if the fields outside Ω are defined to be zero. For this integral to converge, the following two conditions are sufficient, though not necessary:

$$\begin{aligned}\lim_{r \rightarrow \infty} \sqrt{r} \mathbf{E} = 0 \quad \Rightarrow \quad \lim_{r \rightarrow \infty} \sqrt{r} \nabla \times \mathbf{H} = 0 \quad \text{from (2.5) if } \omega \neq 0 \\ \text{and} \quad \lim_{r \rightarrow \infty} \sqrt{r} \mathbf{H} = 0\end{aligned} \tag{2.23}$$

In order to eliminate solutions which grow radially, the constraints of (2.23) will be adopted as the far-field boundary conditions since they only admit fields which decay radially. It will be assumed that these far-field constraints are valid for anisotropic open guides also and that their imposition does not result in the omission of valid guided modes.

It is also necessary to eliminate those solutions which propagate radially outside the core. For isotropic guides, this can be done by requiring that the transverse wavenumber $k_i^2 = \epsilon_{r_i} k_0^2 - \beta^2$ be purely imaginary in each infinite region i (see [48, p.292]). This relation is imposed on the solutions to the problem examples in Chapter 5 which are all isotropic. Setting $k_i^2 = jh$, we have

$$h^2 = \beta^2 - \epsilon_{r_i} k_0^2 > 0 \tag{2.24}$$

The case $h = 0$ corresponds to modes at cut-off and is not considered.

8. The Variational Principle

The solutions $(k_{01}, \mathbf{H}_1), (k_{02}, \mathbf{H}_2), \dots, (k_{0p}, \mathbf{H}_p)$ of the above curl-curl differential equation with boundary conditions (2.12), (2.14) and (2.23) can be obtained with the following variational principle [49,50]:

Variational Principle 2.1

Let \mathcal{D} be the space of complex three component vector functions $\mathbf{H}(x, y, z) = \mathbf{H}(x, y)e^{-j\beta z}$ ($\beta \neq 0$) which are infinitely differentiable in Ω , zero outside Ω , square integrable and satisfy the prescribed boundary conditions:

$$\mathbf{H} \times \mathbf{n} = \mathbf{0} \quad \text{on } \partial\Omega_o \quad (2.25a)$$

$$s \mathbf{H} \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega_s \quad (2.25b)$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \mathbf{H} = \mathbf{0} \quad (2.25c)$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \nabla \times \mathbf{H} = \mathbf{0} \quad (2.25d)$$

$$\lim_{r \rightarrow \infty} \sqrt{r} s \nabla \cdot \mathbf{H} = 0 \quad (2.25e)$$

Then the stationary points \mathbf{H}_i of the following functional for positive definite K , $s \geq 0$ and $\mathbf{H} \neq \mathbf{0}$

$$k_0^2(\mathbf{H}) = \frac{\int_{z=0}^{\lambda_0} \int_{\Omega} ((\nabla \times \mathbf{H})^* K^{-1} (\nabla \times \mathbf{H}) + s (\nabla \cdot \mathbf{H})^* (\nabla \cdot \mathbf{H})) dx dy dz}{\int_{z=0}^{\lambda_0} \int_{\Omega} (\mathbf{H}^* \mathbf{H}) dx dy dz} \quad (2.26)$$

are solutions to the following differential equation

$$\nabla \times K^{-1} \nabla \times \mathbf{H} - s \nabla (\nabla \cdot \mathbf{H}) - k_{0i}^2 \mathbf{H} = \mathbf{0} \quad \text{in } \Omega \quad (2.25f)$$

$$\text{where } k_{0i}^2 = k_{0i}^2(\mathbf{H}) \geq 0 \quad (2.25g)$$

and satisfy the following natural boundary conditions:

$$(\nabla \times \mathbf{H}_i) \times \mathbf{n} = \mathbf{0} \quad \text{on } \partial\Omega_s \quad (2.25h)$$

$$s \nabla \cdot \mathbf{H}_i = 0 \quad \text{on } \partial\Omega_o \quad (2.25i)$$

in addition to the above prescribed boundary conditions (2.25a,b,c,d,e)

The stationary points are the $\mathbf{H}(x, y, z)$ functions at which the first variation of $k_0^2(\mathbf{H})$ vanishes (see Gelfand and Fomin [51]). Note that if $k_0 \neq 0$, equation (2.25h) is equivalent to the electric wall boundary condition from (2.6). The proof of this variational principle is given in Appendix 1. If s is set to zero, then this reduces to Berk's variational principle [52,53] which, when used in the finite element method, is known to yield spurious, or unphysical, modes in addition to the physical modes [54]. To determine the source of these spurious modes, the comments made by Webb [55] concerning the stationary points of the above variational principle for the closed waveguide case will be adapted for the open waveguide case as follows. Consider the following two sets of equations:

$$\nabla \times K^{-1} \nabla \times \mathbf{H} - k_0^2 \mathbf{H} = \mathbf{0} \quad \text{in } \Omega \quad (2.27a)$$

$$\mathbf{H} \times \mathbf{n} = \mathbf{0} \quad \text{on } \partial\Omega_0 \quad (2.27b)$$

$$(\nabla \times \mathbf{H}) \times \mathbf{n} = \mathbf{0} \quad \text{on } \partial\Omega_s \quad (2.27c)$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \mathbf{H} = \mathbf{0} \quad (2.27d)$$

$$\lim_{r \rightarrow \infty} \sqrt{r} (\nabla \times \mathbf{H}) = \mathbf{0} \quad (2.27e)$$

$$k_0^2 > 0 \quad (2.27f)$$

and

$$s(\nabla(\nabla \cdot \mathbf{H}) + k_0^2 \mathbf{H}) = \mathbf{0} \quad \text{in } \Omega \quad (2.28a)$$

$$\mathbf{H} \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega_s \quad (2.28b)$$

$$\nabla \cdot \mathbf{H} = 0 \quad \text{on } \partial\Omega_0 \quad (2.28c)$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \mathbf{H} = \mathbf{0} \quad (2.28d)$$

$$\lim_{r \rightarrow \infty} \sqrt{r} (\nabla \cdot \mathbf{H}) = 0 \quad (2.28e)$$

$$k_0^2 > 0 \quad (2.28f)$$

As Webb demonstrates, the solutions of (2.27) and those of (2.28) are stationary points of Variational Principle 2.1. The static solutions for which $k_0^2 = 0$ are not considered. The divergence-free solutions of equation (2.27) are the physical modes of the guide whereas the solutions of (2.28) are irrotational and do not correspond to any physical

modes. Since the frequencies of the solutions to (2.28) scale with s , there will be infinitely many of these solutions as s approaches 0. This is the origin of the spurious modes which occur with Variational Principle 2.1 when s is set to zero.

From (2.24), the guided modes of isotropic guides satisfy $k_0^2 < \beta^2/\epsilon_{r_1}$. It is shown in Appendix 2 that the spurious mode satisfy the relation $k_0^2 > s\beta^2$. Therefore, if we set $s = 1/\epsilon_{r_1}$, the solutions obtained with Variational Principle 2.1 will be free of spurious modes in the interval $k_0^2 < \beta^2/\epsilon_{r_1}$. This technique will be used to eliminate spurious modes from the numerical solutions of isotropic guides.

It remains to justify the imposition of the boundary condition $\mathbf{H} \cdot \mathbf{n} = 0$ on $\partial\Omega_s$ on the physical modes in Variational Principle 2.1. This can be done as follows:

$$\begin{aligned}
 \text{From (2.11): } (\nabla \times \mathbf{H}) \times \mathbf{n} &= 0 \text{ on } \partial\Omega_s \\
 \Rightarrow K^{-1}(\nabla \times \mathbf{H}) \times \mathbf{n} &= 0 \text{ on } \partial\Omega_s \\
 \Rightarrow [\nabla \times K^{-1}(\nabla \times \mathbf{H})] \cdot \mathbf{n} &= 0 \text{ on } \partial\Omega_s \\
 \Rightarrow \mathbf{H} \cdot \mathbf{n} &= 0 \text{ on } \partial\Omega_s \text{ from (2.22) if } k_0^2 \neq 0
 \end{aligned}$$

This boundary condition is therefore consistent with the physical mode equation which are solutions to (2.27).

Chapter Three

The Conventional Finite Element Method

1. The Isotropic Guide Case

The finite element method (FEM) will be presented in this chapter. Chapter 3 applies the FEM only to the core region of isotropic dielectric guides and Chapter 4 extends the technique to handle the infinite regions by introducing a new type of *infinite element*. The theory of Chapter 2 is applicable to the more general open anisotropic guides, but because the anisotropic case was not implemented in code, the numerical method presented in the next two chapters is restricted to isotropic guides only. This restriction is not imposed by the FEM; this method has previously been applied to lossless anisotropic closed guides composed of materials having permittivity tensors that are diagonal [6,56] and non-diagonal [10,23,57].

The boundary value problem for isotropic open guides is the same for anisotropic open guides except that all occurrences of the relative permittivity tensor K are replaced by the scalar ϵ_r . The modes of isotropic guides can be also determined with Variational Principle 2.1, in which K is again replaced by ϵ_r .

The FEM theory pertaining to the vector Helmholtz problem is well established and therefore its presentation in this chapter is brief. Only the concepts that are relevant to this thesis are emphasised; many references to more detailed treatments in the literature are provided. For an introduction to the finite element method, the reader is referred to Zienkiewicz [3] or Silvester and Ferrari [4].

2. The Unknown Field Components

One advantage of restricting the consideration to isotropic guides is that the number of unknown field components is halved. To demonstrate this, we first decompose

the components \mathbf{H} into real and imaginary parts:

$$\mathbf{H} = (H_{xr} + jH_{xi})\mathbf{i} + (H_{yr} + jH_{yi})\mathbf{j} + (H_{zr} + jH_{zi})\mathbf{k} = \begin{bmatrix} H_{xr} + jH_{xi} \\ H_{yr} + jH_{yi} \\ H_{zr} + jH_{zi} \end{bmatrix}$$

where $H_{xr}, H_{xi}, H_{yr}, H_{yi}, H_{zr}$ and H_{zi} are real valued functions of (x, y, z) . Substituting these fields into the isotropic Helmholtz equation (2.22):

$$\begin{bmatrix} \left[\frac{\partial^2}{\partial y \partial x} H_{yr} + (-\frac{\partial^2}{\partial y^2} + \beta^2) H_{xr} - \beta \frac{\partial}{\partial x} (-H_{zi}) \right] + \\ j \left[\frac{\partial^2}{\partial y \partial x} H_{yi} + (-\frac{\partial^2}{\partial y^2} + \beta^2) H_{xi} - \beta \frac{\partial}{\partial x} H_{zr} \right] \\ \left[-\beta \frac{\partial}{\partial y} (-H_{zi}) + (-\beta^2 - \frac{\partial^2}{\partial x^2}) H_{yr} + \frac{\partial^2}{\partial x \partial y} H_{xr} \right] + \\ j \left[-\beta \frac{\partial}{\partial y} H_{zr} + (-\beta^2 - \frac{\partial^2}{\partial x^2}) H_{yi} + \frac{\partial^2}{\partial x \partial y} H_{xi} \right] \\ \left[\beta \frac{\partial}{\partial x} H_{xr} - (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}) (-H_{zi}) + \beta \frac{\partial}{\partial y} H_{yr} \right] + \\ j \left[\beta \frac{\partial}{\partial x} H_{xi} - (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}) (-H_{zr}) + \beta \frac{\partial}{\partial y} H_{yi} \right] \end{bmatrix} = \epsilon_r k_0^2 \begin{bmatrix} H_{xr} + jH_{xi} \\ H_{yr} + jH_{yi} \\ -H_{zi} + jH_{zr} \end{bmatrix}$$

The assumption of homogeneity (i.e. $\epsilon_r = \text{constant}$) has been made in this derivation since, as shown later, the FEM mesh defines a piecewise homogeneous problem region. By equating real and imaginary parts on both sides, we obtain identical equations in the variables $(H_{xr}, H_{yr}, -jH_{zi})$ and $(jH_{xi}, jH_{yi}, H_{zr})$. It is therefore necessary to solve for only half the components, and the transverse field \mathbf{H} (recall that $\mathbf{H} = \mathbf{H}e^{-j\beta z}$) is redefined as:

$$\mathbf{H}(x, y) = H_x(x, y)\mathbf{i} + H_y(x, y)\mathbf{j} - jH_z(x, y)\mathbf{k}$$

where H_x, H_y and H_z are real valued functions. This simplification is not possible in the general anisotropic guide case.

3. The Rayleigh-Ritz Method

The functional of Variational Principle 2.1 lends itself readily to the application of a numerical procedure called the *Rayleigh-Ritz method* or the *direct method* [58,p.162][51,p.193]. This technique is used to locate approximations to the stationary points of the curl-curl functional in the space spanned by the N linearly independent vector functions

$$\tilde{\mathbf{H}}_1, \tilde{\mathbf{H}}_2, \dots, \tilde{\mathbf{H}}_N$$

which are called *trial functions*. The trial functions must individually satisfy the essential boundary conditions stipulated by Variational Principle 2.1. In order to locate the stationary points, \mathbf{H} is written as a weighted sum of the trial functions:

$$\mathbf{H} = \sum_{i=1}^N a_i \tilde{\mathbf{H}}_i = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}^T \begin{bmatrix} \tilde{\mathbf{H}}_1 \\ \tilde{\mathbf{H}}_2 \\ \vdots \\ \tilde{\mathbf{H}}_N \end{bmatrix} = \mathbf{a}^T \begin{bmatrix} \tilde{\mathbf{H}}_1 \\ \tilde{\mathbf{H}}_2 \\ \vdots \\ \tilde{\mathbf{H}}_N \end{bmatrix}$$

and then substituted into the $k_0^2 (\mathbf{H} e^{-j\beta z})$ functional (2.26). The unknown weights a_i are determined by setting the derivatives $\frac{\partial k_0^2}{\partial a_1}, \frac{\partial k_0^2}{\partial a_2}, \dots, \frac{\partial k_0^2}{\partial a_N}$ to zero which is equivalent to taking the first variation of k_0^2 . From this a generalised eigenvalue problem $\mathbf{A}\mathbf{a} = \nu \mathbf{B}\mathbf{a}$ can be derived which yields N eigenvalues ν_n and eigenvectors \mathbf{a}_n :

$$\nu_i = k_0^2 (\mathbf{a}_i^T \begin{bmatrix} \tilde{\mathbf{H}}_1 \\ \tilde{\mathbf{H}}_2 \\ \vdots \\ \tilde{\mathbf{H}}_N \end{bmatrix} e^{-j\beta z}) ; \quad 1 \leq i \leq N$$

The objective of this thesis is to obtain the p lowest modal frequencies but the Rayleigh-Ritz method cannot guarantee this for any choice of trial functions. To demonstrate this, consider the trial function space spanned by the trial functions $(\tilde{\mathbf{H}}_i = \mathbf{H}_{i+1}; 1 \leq i \leq N)$, where $(\mathbf{H}_1, \mathbf{H}_2, \dots)$ are the actual solutions in increasing eigenvalue order. The eigenvalues obtained with the Rayleigh-Ritz method are then $(\nu_i = k_{0(i+1)}^2 = k_0^2(\mathbf{H}_{i+1} e^{-j\beta z}); 1 \leq i \leq N)$ which are not the lowest N eigenvalues (the fundamental mode is missing). As shown by this example, the ability of the Rayleigh-Ritz method to approximate the first p modes can depend strongly on the choice of trial functions.

When this method is applied to the analysis of open guides, the number and variety of trial functions should be such that each actual mode can be modelled to an arbitrarily close approximation.

4. The Finite Element Method

The FEM is a special case of the Rayleigh-Ritz method where a finite problem region Ω_f has been subdivided or *meshed* into L non-overlapping subregions such that

$$\Omega = \Omega_1 \cup \Omega_2 \cup \dots \cup \Omega_L$$

$$\text{and } \Omega_i \cap \Omega_j = 0 \text{ if } i \neq j$$

To each sub-region i there corresponds one or more trial functions which are zero outside Ω_i . The term *finite element* refers to a particular choice of the subregion shape and the accompanying trial functions defined in its interior. Many types of finite elements have been proposed in the literature (see Zienkiewicz). In this thesis, the triangular finite elements of Silvester [59] are implemented. As an example of problem region subdivision, the core region of the coated cylindrical open guide of Figure 3.1 a) is meshed with triangular elements in Figure 3.1 b). Note that the curved edges must be approximated with line segments. This introduces an error in the results which can be reduced by using smaller elements.

The permittivity is constant inside each triangular element to facilitate the integration of the functional (see section 7). If the permittivity in any region of the original problem varies continuously, then it must be approximated in discrete steps by homogeneous elements. This approximation can again be improved by using smaller elements.

5. Triangular Finite Elements

Triangular elements possess a local coordinate system (see Figure 3.2) in which the position of a point P is given by the relative areas $(\zeta_1, \zeta_2, \zeta_3)$, called *zeta coordinates*, of the triangles formed by joining P to the three vertices with line segments. That is

$$\zeta_1 = \frac{\text{area of the triangle } 23P}{\text{area of the triangle } 123}$$

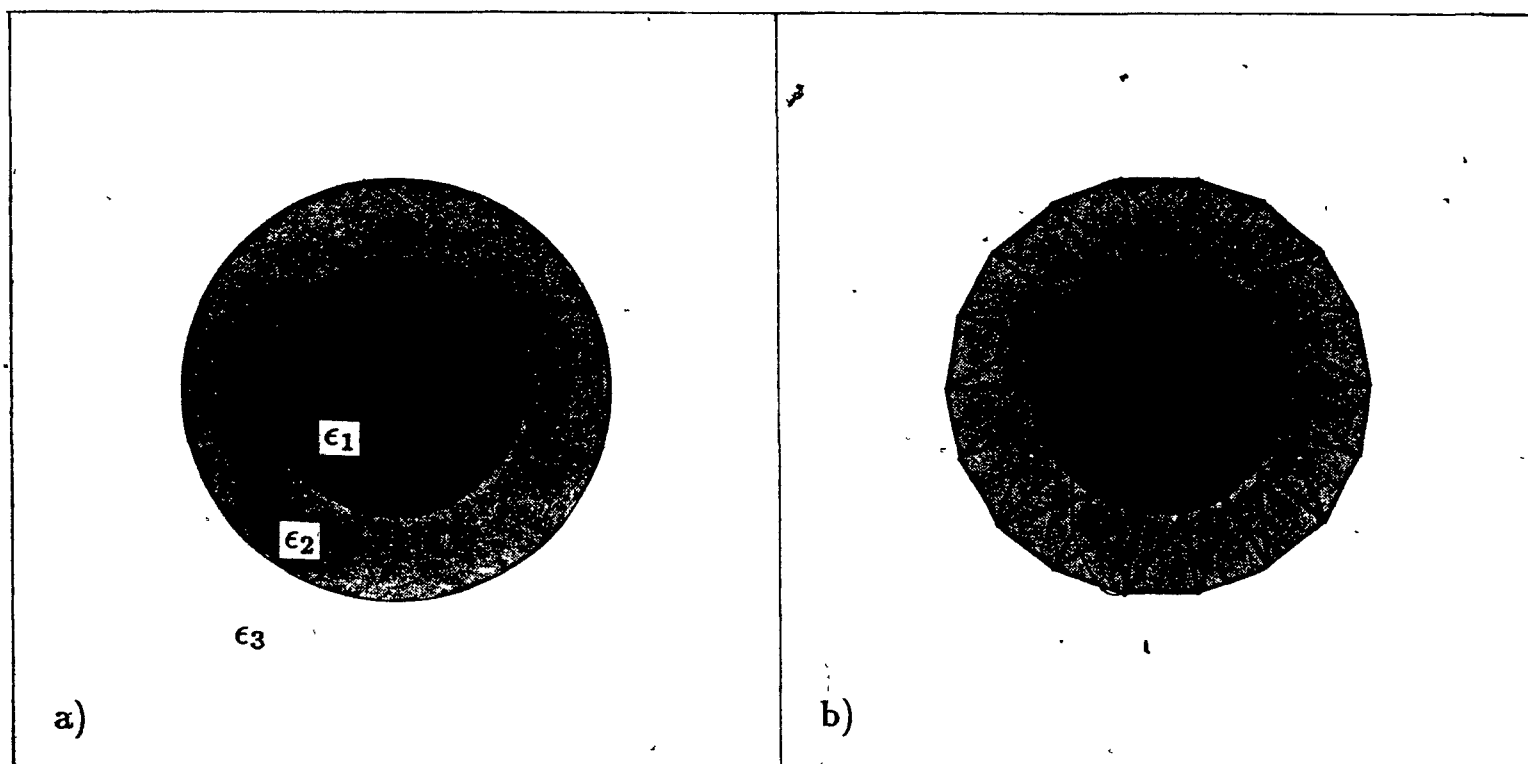


Figure 3.1 An example of problem region sub-division using triangular elements.

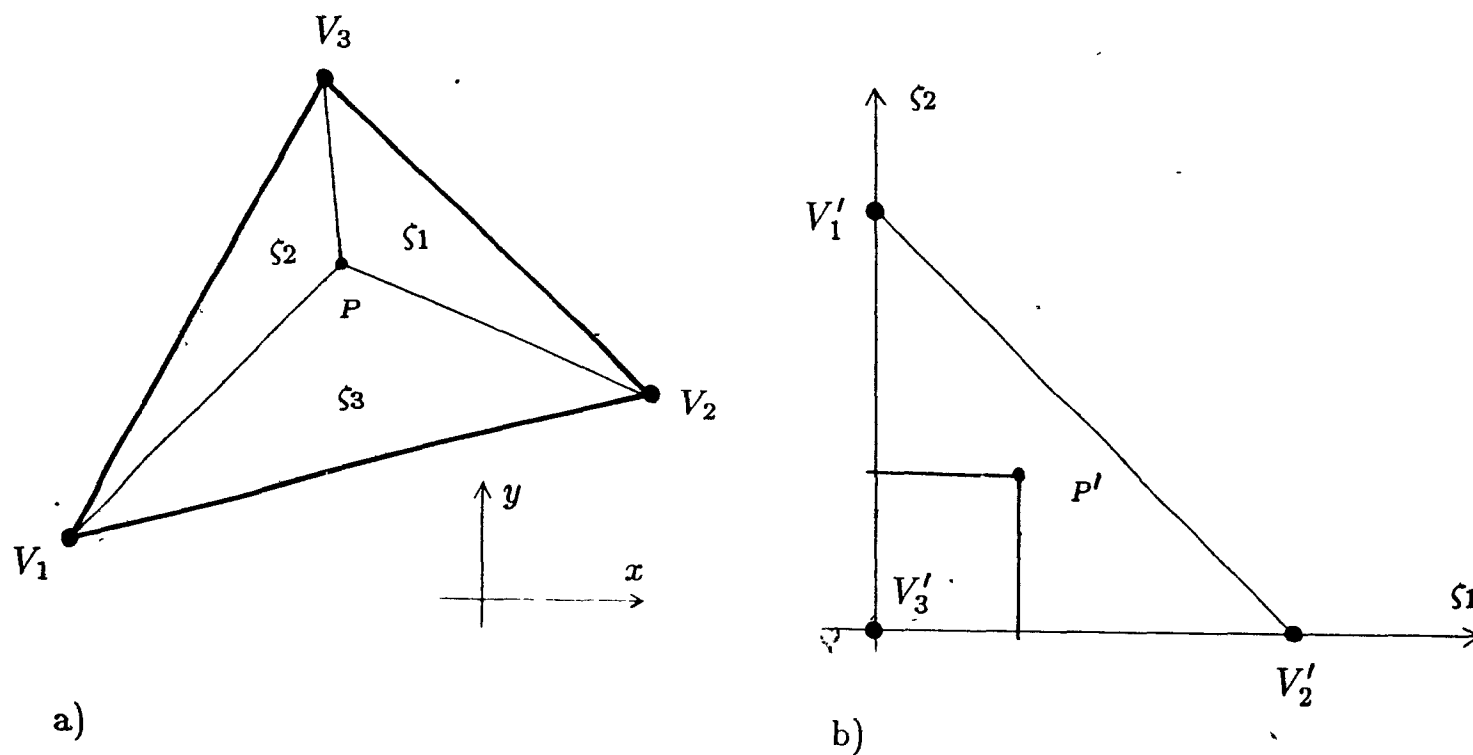


Figure 3.2 The zeta coordinates in a) the x - y plane and b) the ζ_1 - ζ_2 plane. The point P in the x - y plane maps to the point P' in the ζ_1 - ζ_2 plane.

$$\zeta_2 = \frac{\text{area of the triangle 31P}}{\text{area of the triangle 123}}$$

$$\zeta_3 = \frac{\text{area of the triangle 12P}}{\text{area of the triangle 123}}$$

The trial functions associated with triangular finite elements are called α -polynomials and each is a function of the zeta coordinates $(\zeta_1, \zeta_2, \zeta_3)$. The number of α -polynomials defined in each element subregion is $n_0 = (n+2)(n+1)/2$ where n is called the *order* of the element; n_0 is the number of polynomials required to form a complete set of n^{th} order trivariate polynomials in each element.

The α -polynomials functions are interpolatory at n_0 regularly spaced points called *nodes* in the triangle; that is each function has unity value at its particular node and is zero at all other nodes. Figure 3.3 illustrates how the nodes are positioned for element orders 1 to 4. Each node and corresponding α -polynomial is designated by both a single index and a triple index.

For each order n , the n_0 α -polynomials are defined [59] in terms of the triple indices as follows:

$$\alpha_{ijk} = P_i(\zeta_1) P_j(\zeta_2) P_k(\zeta_3) \quad (3.1)$$

$$\text{where } P_l(\zeta) = \begin{cases} \prod_{q=1}^l \left(\frac{n\zeta - q + 1}{q} \right) & \text{if } l \geq 1 \\ 1 & \text{if } l = 0 \end{cases}$$

6. The Field Vector Trial Functions

Inside each triangular element, the \mathbf{H} field is written as a weighted sum of n_0 vector trial functions $\tilde{\mathbf{H}}_i$ in which each of the three components of \mathbf{H} is assigned its own α -polynomial trial function at every node:

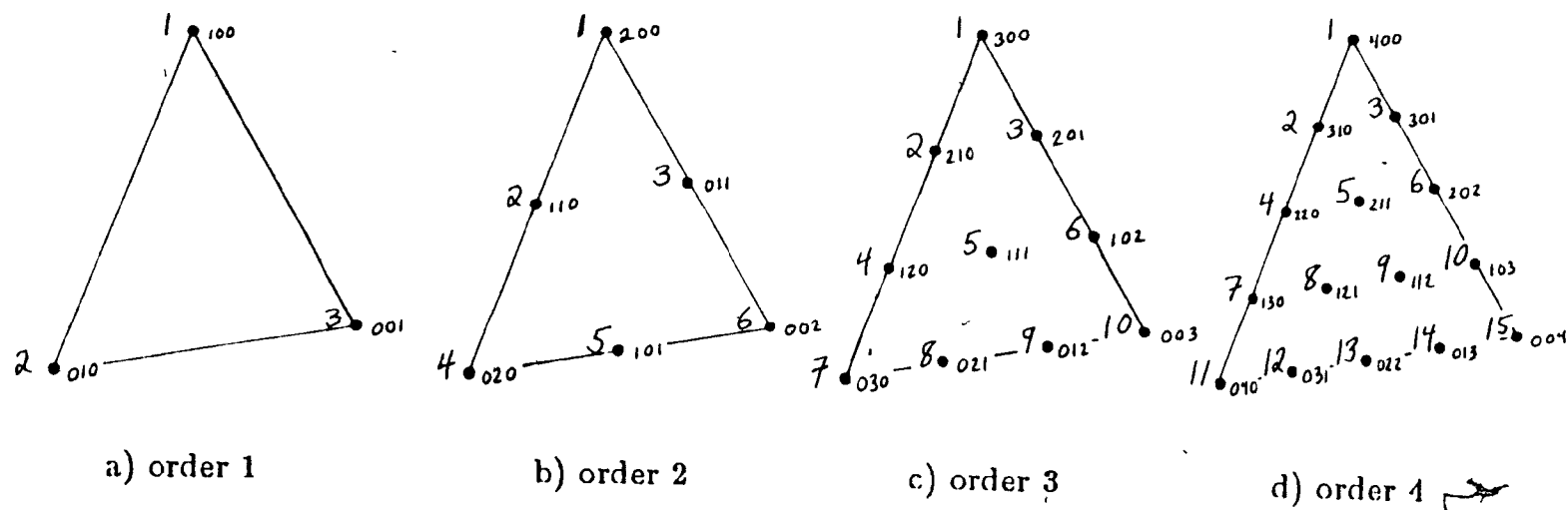


Figure 3.3 The single and triple indices for triangular element orders one to four.

$$\begin{aligned}
\mathbf{H} &= \sum_{i=1}^{n_0} \tilde{\mathbf{H}}_i \\
&= \sum_{i=1}^{n_0} [a_{ix} \alpha(\zeta_1, \zeta_2, \zeta_3) \mathbf{i} + a_{iy} \alpha(\zeta_1, \zeta_2, \zeta_3) \mathbf{j} - ja_{iz} \alpha(\zeta_1, \zeta_2, \zeta_3) \mathbf{k}] \\
&= \sum_{i=1}^{n_0} [a_{ix} \mathbf{i} + a_{iy} \mathbf{j} - ja_{iz} \mathbf{k}] \alpha(\zeta_1, \zeta_2, \zeta_3)
\end{aligned}$$

where the single index node numbering instead of the triple index is used (see Figure 3.3). The scalar components of the $[a_{ix} \mathbf{i} + a_{iy} \mathbf{j} - ja_{iz} \mathbf{k}]$ vector are the unknown weights. Since the α -polynomials are interpolatory, the vector $[a_{ix} \mathbf{i} + a_{iy} \mathbf{j} - ja_{iz} \mathbf{k}]$ is the vector value of \mathbf{H} at node i . Another property of the α -polynomial trial functions is that the value of \mathbf{H} along any edge depends only on the vector weights of the nodes located on that edge. Elements that abut each other therefore have the same trial function values along the entire length of the shared edge. Elements which satisfy this property, which is necessary for the continuity of \mathbf{H} , are called *conforming* elements.

7. Local Matrix Assembly

The next step in the Rayleigh-Ritz procedure is to insert the trial functions into the functional

$$\begin{aligned}
k_0^2(\mathbf{H}) &= k_0^2(\mathbf{H}e^{-j\beta z}) \\
&= \frac{\int_{z=0}^{\lambda_0} \int_{\Omega} \left[(\nabla \times (\mathbf{H}e^{-j\beta z}))^* \cdot \frac{1}{\epsilon_r} (\nabla \times (\mathbf{H}e^{-j\beta z})) + s(\nabla \cdot (\mathbf{H}e^{-j\beta z}))^* (\nabla \cdot (\mathbf{H}e^{-j\beta z})) \right] dx dy dz}{\int_{z=0}^{\lambda_0} \int_{\Omega} ((\mathbf{H}e^{-j\beta z})^* \cdot (\mathbf{H}e^{-j\beta z})) dx dy dz}
\end{aligned}$$

In the FEM code, the integration of the curl and divergence terms in the numerator of the functional and the magnitude term in the denominator are performed separately and one node pair at time. Consider for example the node pair (u, v) of element l ; insertion

of the trial functions associated with these nodes into the three integrals gives:

$$\int_{z=0}^{\lambda_0} \int_{\Omega_l} (\nabla \times (\tilde{\mathbf{H}}_u e^{-j\beta z}))^* \cdot \frac{1}{\epsilon_r} (\nabla \times (\tilde{\mathbf{H}}_v e^{-j\beta z})) dx dy dz = \begin{bmatrix} a_{ux} \\ a_{uy} \\ a_{uz} \end{bmatrix}^T M^{uv} \begin{bmatrix} a_{vx} \\ a_{vy} \\ a_{vz} \end{bmatrix}$$

$$\int_{z=0}^{\lambda_0} \int_{\Omega_l} (\nabla \cdot (\tilde{\mathbf{H}}_u e^{-j\beta z}))^* (\nabla \cdot (\tilde{\mathbf{H}}_v e^{-j\beta z})) dx dy dz = \begin{bmatrix} a_{ux} \\ a_{uy} \\ a_{uz} \end{bmatrix}^T N^{uv} \begin{bmatrix} a_{vx} \\ a_{vy} \\ a_{vz} \end{bmatrix}$$

$$\int_{z=0}^{\lambda_0} \int_{\Omega_l} (\tilde{\mathbf{H}}_u e^{-j\beta z})^* \cdot (\tilde{\mathbf{H}}_v e^{-j\beta z}) dx dy dz = \begin{bmatrix} a_{ux} \\ a_{uy} \\ a_{uz} \end{bmatrix}^T O^{uv} \begin{bmatrix} a_{vx} \\ a_{vy} \\ a_{vz} \end{bmatrix}$$

where Ω_l is the element subregion and the (M^{uv}, N^{uv}, O^{uv}) matrices are given by

$$M^{uv} = \begin{bmatrix} \frac{\lambda_0}{\epsilon_r} \int_{\Omega_l} \beta^2 \alpha_u \alpha_v dx dy & \frac{\lambda_0}{\epsilon_r} \int_{\Omega_l} (-\frac{\partial \alpha_u}{\partial y} \frac{\partial \alpha_v}{\partial x}) dx dy & \frac{\lambda_0}{\epsilon_r} \int_{\Omega_l} (\nabla \beta \alpha_u \frac{\partial \alpha_v}{\partial x}) dx dy \\ \frac{\lambda_0}{\epsilon_r} \int_{\Omega_l} (-\frac{\partial \alpha_u}{\partial x} \frac{\partial \alpha_v}{\partial y}) dx dy & \frac{\lambda_0}{\epsilon_r} \int_{\Omega_l} (\beta^2 \alpha_u \alpha_v + \frac{\partial \alpha_u}{\partial x} \frac{\partial \alpha_v}{\partial x}) dx dy & \frac{\lambda_0}{\epsilon_r} \int_{\Omega_l} (-\beta \alpha_u \frac{\partial \alpha_v}{\partial y}) dx dy \\ \frac{\lambda_0}{\epsilon_r} \int_{\Omega_l} (-\beta \frac{\partial \alpha_u}{\partial x} \alpha_v) dx dy & \frac{\lambda_0}{\epsilon_r} \int_{\Omega_l} (-\beta \frac{\partial \alpha_u}{\partial y} \alpha_v) dx dy & \frac{\lambda_0}{\epsilon_r} \int_{\Omega_l} (\frac{\partial \alpha_u}{\partial y} \frac{\partial \alpha_v}{\partial y} + \frac{\partial \alpha_u}{\partial x} \frac{\partial \alpha_v}{\partial x}) dx dy \end{bmatrix}$$

$$N^{uv} = \begin{bmatrix} \lambda_0 \int_{\Omega_l} \frac{\partial \alpha_u}{\partial x} \frac{\partial \alpha_v}{\partial x} dx dy & \lambda_0 \int_{\Omega_l} \frac{\partial \alpha_u}{\partial x} \frac{\partial \alpha_v}{\partial y} dx dy & \lambda_0 \int_{\Omega_l} (-\beta \frac{\partial \alpha_u}{\partial x} \alpha_v) dx dy \\ \lambda_0 \int_{\Omega_l} \frac{\partial \alpha_u}{\partial y} \frac{\partial \alpha_v}{\partial x} dx dy & \lambda_0 \int_{\Omega_l} \frac{\partial \alpha_u}{\partial y} \frac{\partial \alpha_v}{\partial y} dx dy & \lambda_0 \int_{\Omega_l} (-\beta \frac{\partial \alpha_u}{\partial y} \alpha_v) dx dy \\ \lambda_0 \int_{\Omega_l} (-\beta \alpha_u \frac{\partial \alpha_v}{\partial x}) dx dy & \lambda_0 \int_{\Omega_l} (-\beta \alpha_u \frac{\partial \alpha_v}{\partial y}) dx dy & \lambda_0 \int_{\Omega_l} \beta^2 \alpha_u \alpha_v dx dy \end{bmatrix}$$

$$O^{uv} = \begin{bmatrix} \lambda_0 \int_{\Omega_f} \alpha_u \alpha_v dx dy & 0 & 0 \\ 0 & \lambda_0 \int_{\Omega_f} \alpha_u \alpha_v dx dy & 0 \\ 0 & 0 & \lambda_0 \int_{\Omega_f} \alpha_u \alpha_v dx dy \end{bmatrix}$$

The elements of these matrices are integrated by converting the variables of integration from the (x, y) to the (ζ_1, ζ_2) variables and then making use of the following matrices:

$$T_{uv} = 2 \int_{\zeta_2=0}^1 \int_{\zeta_1=0}^{1-\zeta_2} \alpha_u \alpha_v d\zeta_1 d\zeta_2$$

$$K_{uv}^{(pq)} = 2 \int_{\zeta_2=0}^1 \int_{\zeta_1=0}^{1-\zeta_2} \frac{\partial \alpha_u}{\partial \zeta_p} \frac{\partial \alpha_v}{\partial \zeta_q} d\zeta_1 d\zeta_2$$

$$L_{uv}^{(p)} = 2 \int_{\zeta_2=0}^1 \int_{\zeta_1=0}^{1-\zeta_2} \frac{\partial \alpha_u}{\partial \zeta_p} \alpha_v d\zeta_1 d\zeta_2$$

The details of how these matrices, called *universal matrices*, can be calculated are given in references [59] and [60]. The universal matrices need only be computed once and then stored in a sequential disc file for retrieval during the execution of the FEM program.

8. The Application of the Boundary Conditions

The above \mathbf{H}_1 trial functions do not satisfy the essential boundary conditions of Variational Principle 2.1 which require that the trial fields be tangential to electric conductor surfaces and normal to magnetic conductor surfaces. In the finite element program, the imposition of these boundary conditions occurs only after the (M^{uv}, N^{uv}, O^{uv}) matrices are calculated because it is computationally more convenient. They are imposed by modifying these matrices to arrive at the same matrix equation that would have resulted if trial functions satisfying the boundary conditions had been used from the start.

As an illustrative example, consider the node pair (u, v) where the field at node u is not constrained by any boundary conditions but node v lies on a $\partial\Omega_o$ boundary surface. The integral of the curl term for instance yields:

$$\int_{z=0}^{\lambda_0} \int_{\Omega_l} (\nabla \times (\tilde{\mathbf{H}}_u e^{-j\beta z}))^* \cdot \frac{1}{\epsilon_r} (\nabla \times (\tilde{\mathbf{H}}_v e^{-j\beta z})) dx dy dz$$

$$= \begin{bmatrix} a_{ux} \\ a_{uy} \\ a_{uz} \end{bmatrix}^T \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} a_{vx} \\ a_{vy} \\ a_{vz} \end{bmatrix}$$

Since the \mathbf{H} field is normal to the $\partial\Omega_o$ boundary surface at node v , the unknown vector $[a_{vx}\mathbf{i} + a_{vy}\mathbf{j} - ja_{vz}\mathbf{k}]$ can be expressed as a normal vector:

$$\left. \mathbf{H} \right|_{\text{node } v} = a_{vx}\mathbf{i} + a_{vy}\mathbf{j} - ja_{vz}\mathbf{k} = a_{vn}\mathbf{n} = a_{vn}n_x\mathbf{i} + a_{vn}n_y\mathbf{j}$$

where $\mathbf{n} = n_x\mathbf{i} + n_y\mathbf{j}$. This implies that $a_{vx} = a_{vn}n_x$, $a_{vy} = a_{vn}n_y$, and $a_{vz} = 0$, so that by resolving the transverse components into a single normal component and eliminating a_{vz} , the integral becomes

$$\int_{z=0}^{\lambda_0} \int_{\Omega_l} (\nabla \times (\tilde{\mathbf{H}}_u e^{-j\beta z}))^* \cdot \frac{1}{\epsilon_r} (\nabla \times (\tilde{\mathbf{H}}_v e^{-j\beta z})) dx dy dz$$

$$= \begin{bmatrix} a_{ux} \\ a_{uy} \\ a_{uz} \end{bmatrix}^T \begin{bmatrix} M_{11}n_x + M_{12}n_y \\ M_{21}n_x + M_{22}n_y \\ M_{31}n_x + M_{32}n_y \end{bmatrix} a_{vn}$$

The modified M^{uv} matrix has only one column. For cases where the fields are constrained on the node u as well, a similar manipulation must be performed on the rows of the modified M^{uv} matrix. The type of manipulation depends on the applied boundary condition. All nodes are assumed to each belong to one of five categories labelled "F", "S", "Y", "O" and "Z", as shown in Figure 3.4. The "F" nodes are not constrained by any boundary surface and retain all three weight variables (a_{ux}, a_{uy}, a_{uz}). The "S" nodes are located on a straight electric conductor and therefore their transverse

- 33 -

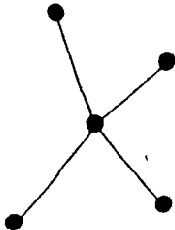
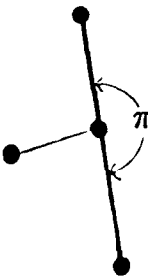
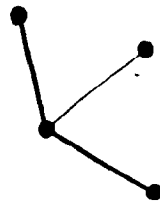
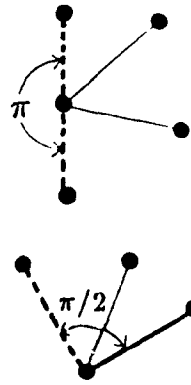
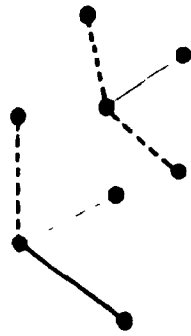
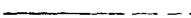


label	F	S	Y	O	Z
field components	(H_x, H_y, H_z)	(H_t, H_z)	(H_z)	(H_n)	none
examples					
legend :	<div style="display: flex; justify-content: space-around; align-items: center;"> <div>  non-boundary edge </div> <div>  $\partial\Omega_s$ perfect electric conductor </div> <div>  $\partial\Omega_o$ perfect magnetic conductor </div> </div>				

Figure 3.4 The five node categories.

components (a_{ux}, a_{uy}) are resolved into a single tangential component a_{ut} and the a_{uz} component is retained. The "Y" nodes also lie on a $\partial\Omega_s$ boundary but at a location where two boundary lines meet at angle other than 180° . Since two tangential components cannot co-exist at one point, the transverse components are simply eliminated for this case and only the a_{uz} component is kept (this is an approximation since the actual field may be singular at this point).

The "O" case occurs where two $\partial\Omega_o$ boundary lines meet at 180° or where a $\partial\Omega_o$ and a $\partial\Omega_s$ boundary line meet at 90° . In this case the transverse components are resolved into a single normal component a_{un} and the a_{uz} component is eliminated as was shown in the above example. Where two $\partial\Omega_o$ boundaries or a $\partial\Omega_o$ and a $\partial\Omega_s$ boundary meet at angles other than 180° or 90° respectively, the "Z" case occurs and all components are eliminated.

9. Global Matrix Assembly

By combining all the contributions from each node pair in every element, the k_0^2 functional for the entire finite problem region Ω_f can be obtained:

$$k_0^2 (\mathbf{H} e^{-j\beta z}) = \frac{\mathbf{a}^T \mathbf{A} \mathbf{a}}{\mathbf{a}^T \mathbf{B} \mathbf{a}}$$

where (\mathbf{A}, \mathbf{B}) are the *global matrices* of the discretised problem and

$$\mathbf{a} = [a_{1x} \ a_{1y} \ a_{1z} \ a_{2x} \ a_{2y} \ a_{2z} \ \dots]^T$$

is the vector consisting of all unknown x, y, z , normal or tangential component variables. Each unknown variable associated with a particular node is coupled algebraically only to the nodal variables of the other elements that contain the node. The \mathbf{A} and \mathbf{B} matrices are thus sparse. To locate the stationary points of this functional, the derivatives of k_0^2 with respect to each unknown variable are taken and set to zero; this yields the following generalised eigenvalue problem:

$$\mathbf{A} \mathbf{a} = k_0^2 \mathbf{B} \mathbf{a}$$

This is to be solved numerically by an eigenvalue solver program (see Chapter 5). The approximate modal normalised frequencies are the square roots of the eigenvalues and the \mathbf{H} field distribution is given by the eigenvectors.

Chapter Four

Infinite Elements

1. Elements for the Infinite Regions

Finite elements cannot be used to mesh the infinite transverse regions surrounding an open guide but infinite elements, which are elements with infinite area, can be used. Many infinite elements have been proposed [18,19,20,61,21,62] for solving deterministic (i.e. non-eigenvalue) unbounded problems in other disciplines of engineering such as fluid dynamics and mechanics. The shape most used in these references is the flared parametric infinite element shown in Figure 4.1 which results from the mapping of a semi-infinite strip in the ξ - η plane. The trial functions selected by each author are much more varied however, since they are specially selected to reflect the actual solution's asymptotic radial behaviour which depends strongly on the type of problem being solved. Infinite elements suitable for extending the FEM to open guides, which present additional difficulties, are discussed next. The resulting method is called the Finite/Infinite Element Method.

2. The Infinite Element Mapping and the Azimuthal Trial Function

Parametric infinite elements [19] are chosen for this thesis because their flexible shape can be used to mesh a wide range of problem geometries. To avoid having the semi-infinite edges 1 and 2 intersect, the angles θ_1 and θ_2 in Figure 4.1b are required to be greater than or equal to 90° . The mapping function which maps any point in the ξ - η plane into the x - y plane is:

$$\begin{aligned} x &= x_1 \left(\frac{1+\eta}{2} \right) (1-\xi) & y &= y_1 \left(\frac{1+\eta}{2} \right) (1-\xi) \\ &+ x_2 \left(\frac{1-\eta}{2} \right) (1-\xi) & &+ y_2 \left(\frac{1-\eta}{2} \right) (1-\xi) \\ &+ x_3 \left(\frac{1+\eta}{2} \right) \xi & &+ y_3 \left(\frac{1+\eta}{2} \right) \xi \\ &+ x_4 \left(\frac{1-\eta}{2} \right) \xi & &+ y_4 \left(\frac{1-\eta}{2} \right) \xi \end{aligned} \quad (4.1)$$

Because it makes the integration of the infinite elements easier (see section 5), the mapped infinite elements are constrained to be symmetrical, that is the angles θ_1 and

θ_2 must be equal. Since the infinite element adjoins a triangular element on an edge with $n + 1$ nodes, the infinite element also has $n + 1$ nodes along that edge numbered locally as shown in Figure 4.1b. Just as for triangular elements, one interpolatory vector trial function is associated with each node. For each component of the field at node i , the trial function is of the form $\psi_i(\eta) \gamma(\xi)$; the η and ξ dependence are thus separate. As shown later, $\gamma(0) = \text{constant}$ and therefore along the edge shared with the finite element, each component varies as $\psi_i(\eta)$. To achieve field continuity with the adjacent triangular element for any values of the nodal weights (a_{ix}, a_{iy}, a_{iz}) , each $\psi_i(\eta)$ function should have the same variation along the shared edge as the corresponding α -polynomial in the triangular element. By defining the local coordinates $\eta_1 = (1 + \eta)/2$ and $\eta_2 = (1 - \eta)/2$, which are equivalent to the corresponding triangular element zeta coordinates, the required ψ function can be written

$$\begin{aligned} \psi_i(\eta) &= P_{i-1}\left(\frac{1-\eta}{2}\right) P_{n+1-i}\left(\frac{1+\eta}{2}\right) \\ &= P_{i-1}^*(\eta_1) P_{n+1-i}(\eta_2) \quad i = 1, 2, \dots, (n+1) \end{aligned}$$

$$\text{where } P_l(v) = \begin{cases} \prod_{q=1}^l \left(\frac{nv - q + 1}{q} \right) & \text{if } l \geq 1 \\ 1 & \text{if } l = 0 \end{cases}$$

which is similar to equation (3.1). The remaining part of the variation of $\tilde{\mathbf{H}}$ is described by $\gamma(\xi)$ which is called the radial trial function.

3. The Radial Trial Functions

The choice of radial trial functions must take into consideration the solution's actual radial behaviour. The interval $0 \leq r \leq \infty$ can be thought of as consisting of two parts, near field and the far field or asymptotic region.

The main difficulty in solving the unbounded Helmholtz eigenproblem is that the actual asymptotic radial field behaviour is not known before the problem is solved. This difficulty is not encountered in open deterministic problems such as the Laplace problem for which the asymptotic field behaviour is known. If a finite/infinite element scheme is used, the correct radial trial functions can therefore be incorporated in the

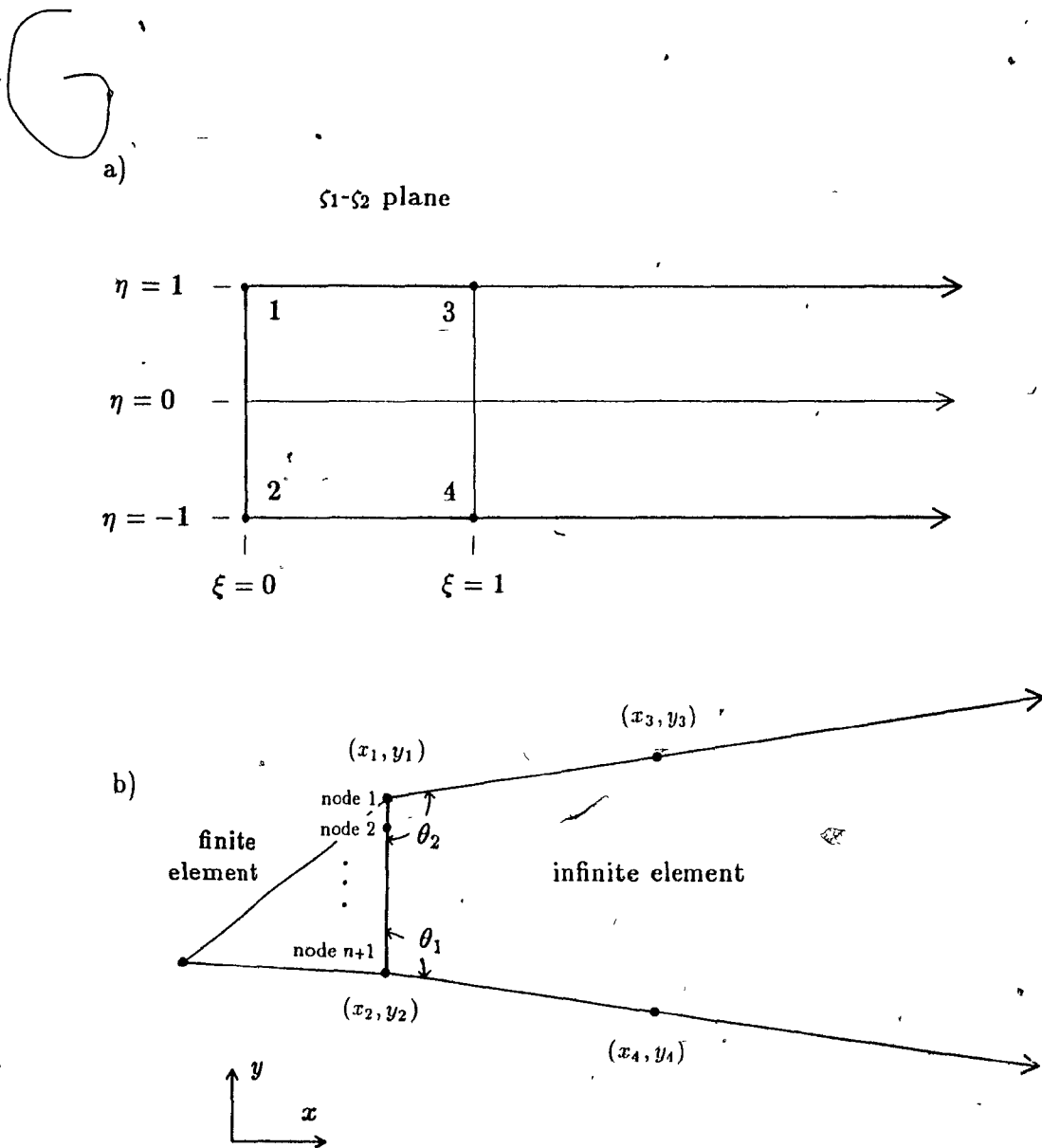


Figure 4.1 The flared parametric infinite element in a) the ξ - η plane and b) the x - y plane.

infinite elements. Alternatively, a number of other approaches can be employed, such as ballooning [14,16], or coupling the FEM to a surface integral method [63,64].

For the scalar Helmholtz eigenproblem described by the differential equation $\nabla^2 \phi + (\epsilon_r k_0^2 - \beta^2) \phi = 0$, this difficulty can be circumvented. It is possible to recast this problem [15] by making $h = \sqrt{\beta^2 - \epsilon_{re} k_0^2}$, the known quantity (instead of β) where ϵ_{re} is the relative permittivity in the exterior, and by modifying the scalar functional accordingly. This makes the asymptotic field behaviour known in advance, since it is specified by h . The same techniques available for handling the unbounded exterior region in the deterministic case can therefore be used.

This particular strategy is not applicable to the vector Helmholtz eigenproblem because the integrand of the functional contains terms in β as well as β^2 . To overcome this difficulty, Yeh et al. [22] used the FEM with parametric infinite elements which incorporated a radial trial function of the form $\gamma = e^{-\alpha r}$ where $\alpha > 0$. The asymptotic field behaviour is therefore specified by the decay length $1/\alpha$ which is globally defined for the whole problem. Because the correct decay to specify is unknown, an outer iteration loop was added to the FEM which optimized the α parameter for each mode of interest.

Rahman and Davies [23] employed similar elements whose decay was specified in the x direction by e^{-x/L_x} or the y direction by e^{-y/L_y} or both by $e^{-x/L_x - y/L_y}$. The total number of L decay lengths for the entire problem can be greater than one but only one was assigned to each element. The decay lengths were optimized in an outer iteration loop, as with Yeh et al.. Methods such as these which employ infinite elements incorporating a single optimized exponential decay in each element will be referred to as *optimized single decay* (OSD) methods.

Although it was formulated for the scalar Helmholtz problem, the two techniques proposed by Hayata et al. [65] are OSD methods which iterate to the optimal exponential decay parameter efficiently by using either the previous eigenvectors or the previous eigenvalue to calculate the next estimate of the decay length.

The radial trial functions used in this thesis make the above decay length optimization loop unnecessary and conserve the simplicity of the FEM. These trial functions

are of the form

$$\begin{aligned} \gamma(\xi, a_0, a_1, \dots, a_{q-1}) = & a_0 e^{-\xi/L_0} + a_1(e^{-\xi/L_1} - e^{-\xi/L_0}) \\ & + \dots + a_{(q-1)}(e^{-\xi/L_{(q-1)}} - e^{-\xi/L_0}) \end{aligned}$$

where the decay lengths $(L_0, L_1, \dots, L_{q-1})$ specify a sufficient range of decays to allow the γ trial function to adequately model the asymptotic behaviour of all modes of interest simultaneously. These decay lengths are not optimized so the method is called the *multiple fixed decay* (MFD) method. The addition of the unknown variables (a_1, \dots, a_{q-1}) does increase the problem size relative to the OSD method, and therefore also the execution time of the eigenvalue solver but this is offset by the removal of the outer optimization loop and the fact that all p modes are found for in one execution of the solver.

The MFD trial functions also provide superior modelling in general of the actual radial field behaviour because the fields are a linear combination of q exponential decays instead of just one in the case of the OSD method. For example, the modelling of field variation in the near field region is improved with the MFD method, since a certain number of radial trial functions with short decay lengths can be made available to supply flexibility in approximating the fields near the core. The superior modelling ability of the MFD method is particularly pronounced near the cut-off frequencies of the modes as the empirical results of section 5.6 demonstrate.

4. The Full Trial Function

Combining all the azimuthal $\psi_i(\eta)$ functions and the radial MFD $\gamma(\xi)$ function into a single expression, there obtains:

$$\mathbf{H}(\eta, \xi) = \sum_{i=1}^{n+1} \psi_i(\eta) (\gamma_{ix} \mathbf{i} + \gamma_{iy} \mathbf{j} - j\gamma_{iz} \mathbf{k}) \quad (4.2)$$

$$\text{where } \gamma_{ix} = a_{i0x} e^{-\xi/L_0} + \sum_{m=1}^{q-1} a_{imx} (e^{-\xi/L_m} - e^{-\xi/L_0})$$

$$\gamma_{iy} = a_{i0y} e^{-\xi/L_0} + \sum_{m=1}^{q-1} a_{imy} (e^{-\xi/L_m} - e^{-\xi/L_0})$$

$$\gamma_{iz} = a_{i0z} e^{-\xi/L_0} + \sum_{m=1}^{q-1} a_{imz} (e^{-\xi/L_m} - e^{-\xi/L_0})$$

For $\xi = 0$, the nodal trial functions reduce to $\gamma_{ix} = a_{i0x}$, $\gamma_{iy} = a_{i0y}$ and $\gamma_{iz} = a_{i0z}$. The unknown variables (a_{i0x} , a_{i0y} , a_{i0z}) are therefore the values of the field components at the $(n+1)$ nodes of each infinite element.

Just as for triangular elements, the trial functions are modified by the application of the electric or magnetic conducting wall boundary conditions on any of the three edges. As before, this modification occurs after the integration of the functional by resolving the unknown variable components into normal or tangential components, or the removal of components (see section 8 of Chapter 3). Note that these trial functions also satisfy the required far-field boundary conditions (2.25c), (2.25d) and (2.25e).

5. The Integration of Trial Function Terms

Let n be the order of the triangular elements. When the trial functions of (4.2) are inserted into the $k_0^2 (\mathbf{H}e^{-j\beta z})$ functional and expanded, both the numerator and denominator consist of a series of terms of the form

$$I = \int_{-1}^1 \int_0^\infty e^{-\xi/L} \xi^u \eta^v J^w d\xi d\eta$$

$$\begin{aligned} \text{where } u &= 0, 1, 2 \\ v &= 0, 1, \dots, 2n \\ w &= -1, 0, 1 \end{aligned}$$

and $J = \left| \frac{\partial(x,y)}{\partial(\xi,\eta)} \right|$ is the Jacobian of the mapping from the (x,y) to the (ξ,η) domain. For the mapping of equation (4.1), it can be shown that J is of the form

$$J = J_1 + J_2 \xi + J_3 \eta$$

where $J_1 > 0$ and $J_2 \geq 0$, and therefore the general functional term I is

$$I(L, u, v, w, J_1, J_2, J_3) = \int_{-1}^1 \int_0^\infty e^{-\xi/L} \xi^u \eta^v (J_1 + J_2 \xi + J_3 \eta)^w d\xi d\eta$$

A simple method of determining this integral does not exist as in the case of triangular elements for which universal matrices can be defined. However, if $w = 0$, $w = 1$ or $J_2, J_3 = 0$ the integration of I can be performed using the simple formula

$$\int_{-1}^1 \int_0^{\infty} e^{-\xi/L} \xi^u \eta^v d\xi d\eta = \begin{cases} 0 & \text{if } v \text{ is odd} \\ \frac{2 \Gamma(u+1)}{(v+1) L^{u-1}} & \text{if } v \text{ is even} \end{cases}$$

If $w = -1$, the I integral is

$$I(L, u, v, -1, J_1, J_2, J_3) = \int_{-1}^1 \int_0^{\infty} e^{-\xi/L} \xi^u \eta^v \frac{1}{J_1 + J_2 \xi + J_3 \eta} d\xi d\eta$$

For this integral, a Gauss cubature integration scheme [21] and a semi-closed exact formula were attempted. Both were found to be inadequate because their accuracy was not uniform over all possible values of L and $\left| \frac{J_1}{J_2} \right|$. The inaccuracy in the semi-closed exact formula was due to the round-off error that occurs in the summation of very large terms of near equal magnitude but opposite signs. The Gauss cubature method was also found to be too slow because of the high orders of Gauss-Legendre and Gauss-Laguerre quadrature that are required to achieve acceptable accuracy.

To resolve these difficulties, the integrand is simplified by imposing $J_3 = 0$ which is the same as requiring that the infinite element shape be symmetrical in the x - y plane. The following semi-closed formula can then be used [66,p.321]:

$$\begin{aligned} & \int_{-1}^1 \int_0^{\infty} e^{-\xi/L} \xi^u \eta^v \frac{1}{J_1 + J_2 \xi} d\xi d\eta \\ &= \begin{cases} 0 & \text{if } v \text{ is odd} \\ \frac{2}{(v+1)J_2} \left[(-1)^{u-1} \left(\frac{J_1}{J_2} \right)^u \exp\left(\frac{J_1}{LJ_2}\right) \text{Ei}\left(\frac{J_1}{LJ_2}\right) \right. \\ \quad \left. + \sum_{k=1}^u (k-1)! \left(-\frac{J_1}{J_2} \right)^{u-k} \left(\frac{1}{L} \right)^{-k} \right] & \text{if } v \text{ is even} \end{cases} \end{aligned}$$

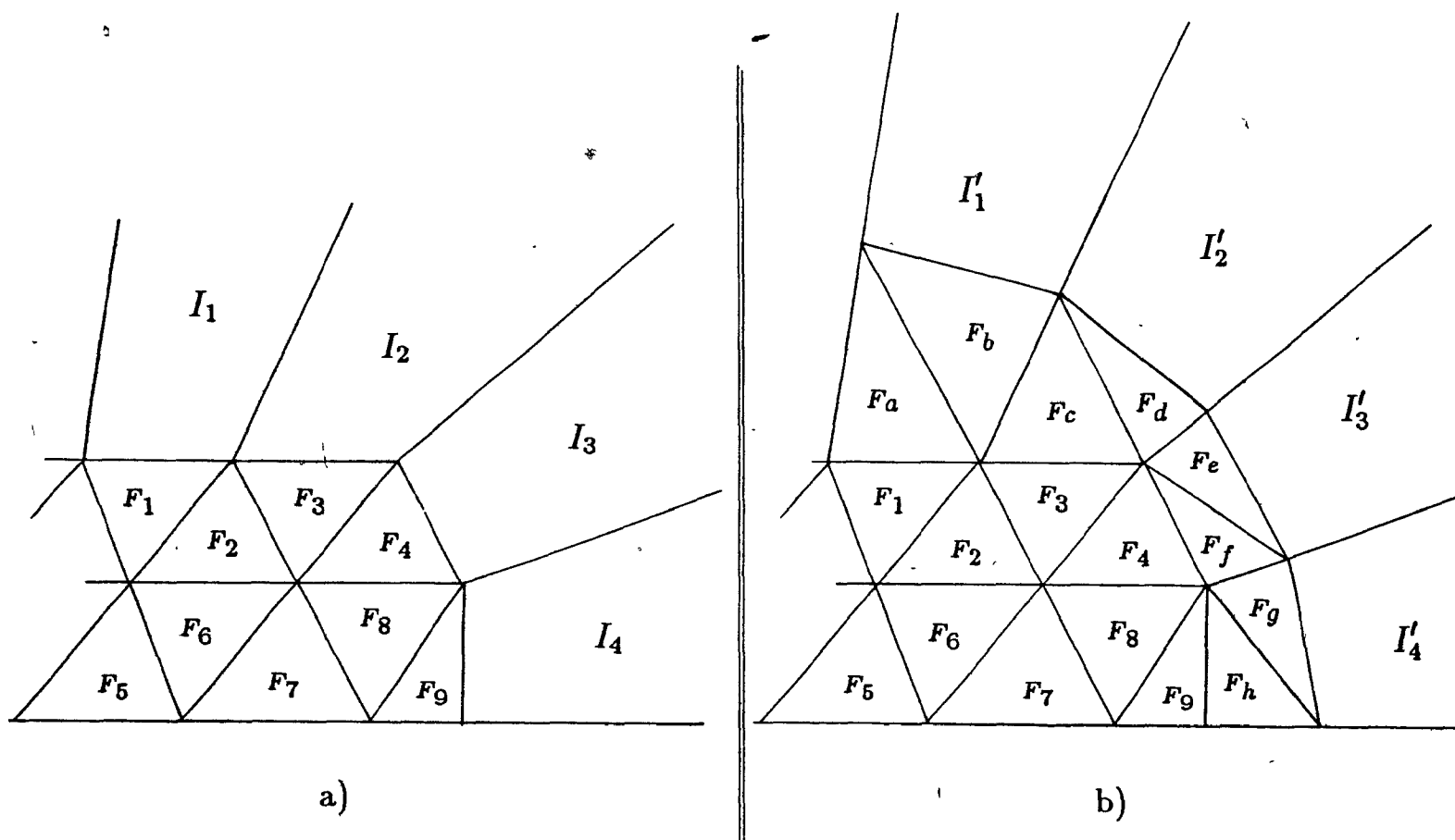


Figure 4.2. The asymmetrical infinite elements I_1 - I_4 in mesh a) are converted to the symmetrical infinite elements I'_1 - I'_4 in mesh b) by the addition of finite elements F_a - F_h .

where Ei is the *exponential integral* function [66,p.927]:

$$Ei(-x) = e^{-x} \sum_{k=1}^n (-1)^k \frac{(k-1)!}{x^k} + R_n \quad \text{for } x > 0$$

$$\text{where } |R_n| < \frac{n!}{x^{n+1}}$$

The use of symmetrical infinite elements does not restrict the problem geometry since any problem region meshed with asymmetrical infinite elements can be converted to a symmetrical infinite element mesh with the addition of triangular elements. Figure 4.2 shows an example of this conversion. Such a scheme increases the number of unknown variables but simplifies the integration code considerably and improves the accuracy.

Chapter Five

Results

1. The Fortran Program

The program "OMAX" was written in Fortran to implement the MFD method presented in Chapters 3 and 4. Two auxiliary programs, UNIMAT and MATPST, were also written to generate the required universal matrices.

The OMAX program accepts a data file as input which describes the β values, the problem mesh, the boundary conditions and the material permittivities. For each β value, it assembles the global matrices, solves the generalised eigenvalue problem and outputs a list of the first p eigenvalues. If requested, OMAX will also generate a file with plotting information that can be used to display field intensity graphs or transverse field vector plots.

The q fixed decay lengths (L_0, L_1, \dots, L_{q-1}) that are supplied to the trial functions of each infinite element are selected automatically by the OMAX program using the parameters q , C_d and N_{nt} (see below) which are supplied in the input file. In the examples, q is set to either 5 or 6 with good results. Increasing q causes OMAX to allocate more radial trial functions to each infinite element which tends to reduce the error in the results but also increases the computation time.

The decay length selection algorithm starts by calculating an estimate L_{min} of the shortest decay length required to model any mode for a particular choice of the phase constant β . Let ϵ_{rmax} and ϵ_{rmin} be the maximum relative permittivity of the problem (usually located in the core) and the minimum relative permittivity (usually in the infinite region) respectively. Now for an open guide consisting simply of a circular homogeneous core (of any radius) of relative permittivity ϵ_{rmax} surrounded by a cladding of relative permittivity ϵ_{rmin} , the following relation holds [67,p.367]:

$$[\max(h)]^2 = (\epsilon_{rmax} - \epsilon_{rmin}) k_0^2$$

where $h = \sqrt{\beta^2 - \epsilon_r k_0^2}$ is the transverse wavenumber anywhere in the cladding. Putting

this in terms of $\max(h)$ and β only,

$$[\max(h)]^2 = \left(1 - \frac{\epsilon_{r\min}}{\epsilon_{r\max}}\right) \beta^2$$

The asymptotic radial behaviour of the fields of this guide is $\frac{e^{-hr}}{\sqrt{hr}}$ [48,p.297], and therefore an estimate L_{\min} of the shortest decay length is

$$L_{\min} = \frac{1}{\max(h)} = \frac{1}{\left(1 - \frac{\epsilon_{r\min}}{\epsilon_{r\max}}\right)^{\frac{1}{2}} \beta}$$

Starting with this estimate, the decay lengths are generated by repeated multiplication by the coefficient C_d :

$$L_{i+1} = C_d L_i$$

The user selected parameter C_d therefore controls the distribution of the decay lengths. For all the test examples, it is set to the value 10. This value causes OMAX to specify decay lengths which are very large relative to the core dimensions of the guides, and thus permits the modelling of the modes very near to cut-off.

Since the fields may have features which are small in comparison to $\max(h)$, the user can specify a number N_{nf} of *near-field* decay lengths. These are also generated from L_{\min} by successive division by C_d . In summary, all decay lengths are generated by the formula:

$$L_i = L_{\min} (C_d)^{i-N_{nf}} \quad i = 0, 1, \dots, (q-1)$$

Most of the computation time is spent assembling the infinite element global matrix contribution and solving the generalised eigenvalue problem. Although the global matrices are sparse, a dense matrix solver is used. The solver is composed of EISPACK routines [2] which convert the generalised eigenvalue problem to a standard eigenvalue problem, tridiagonalise the resulting matrix and then use Sturm sequencing to locate the eigenvalues.

2. The Slab Waveguide

As a simple example and test of the MFD technique, consider the problem of determining the TM modes with no y variation of the slab waveguide shown in Figure 5.1a. The guide consists of a homogeneous dielectric film with relative permittivity $\epsilon_{r1} = 3$ and thickness t , deposited on a conducting plane located at $x = 0$. The surrounding homogeneous cladding has relative permittivity $\epsilon_{r2} = 1$. An analytical expression for the E_z component of the TM solutions is [44,p.114]:

$$E_z = \begin{cases} \frac{\sin k_d x}{\sin k_d t} & \text{for } 0 \leq x < t \\ \frac{e^{-hx}}{e^{-ht}} & \text{for } x \geq t \end{cases} \quad (5.1a)$$

where k_d and h are the transverse wavenumbers satisfying the equations

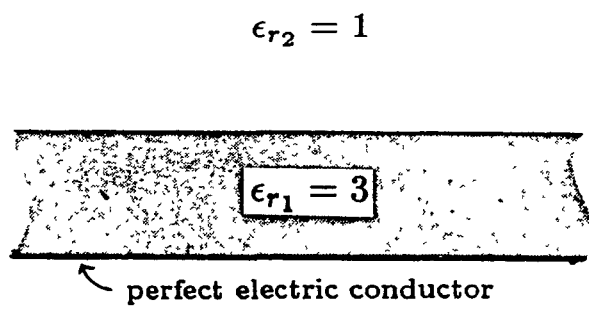
$$k_0^2 \epsilon_{r1} = \beta^2 + k_d^2 \quad (5.1b)$$

$$k_0^2 \epsilon_{r2} = \beta^2 - h^2 \quad (5.1c)$$

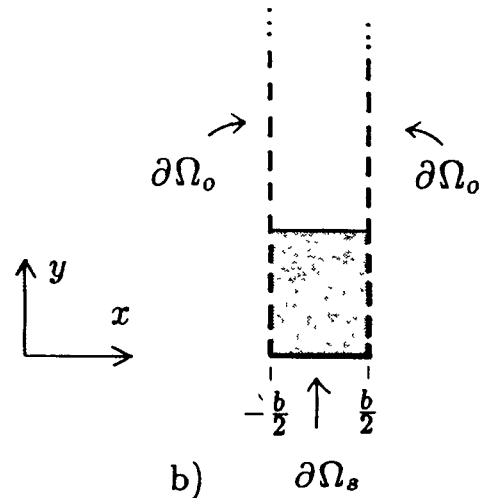
The modes are the simultaneous solutions to equations (5.1b), (5.1c) and the characteristic equation

$$k_d t \tan k_d t = \frac{\epsilon_{r1}}{\epsilon_{r2}} h t \quad (5.1d)$$

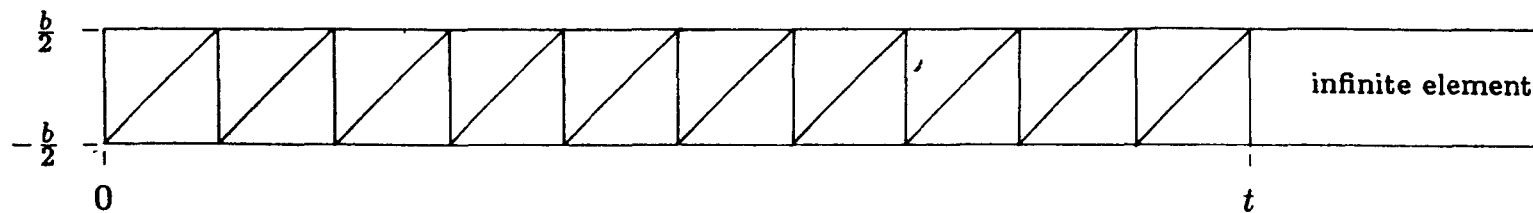
It is impossible to apply the infinite/finite element method to this problem in its unmodified form. The problem has no core region of finite size that can be meshed with finite elements and infinite elements that extend in the $+y$ or $-y$ direction cannot be used, since their trial fields decay in these directions whereas the actual solutions are y invariant from (5.1a). To overcome these difficulties, two perfect magnetic conductor boundaries ($\partial\Omega_o$) are placed as shown in Figure 5.1b at $y = \pm \frac{b}{2}$ which produces a core region and permits a mesh like that of Figure 5.1c with 20 finite elements and a single infinite element to be used. The actual solutions do decay in the $+x$ direction for $x \geq t$, so that infinite elements extending in that direction can be used. The modified guide supports the following TE and TM modes:



a)



b)



c)

Figure 5.1 The slab waveguide example: a) the original problem, b) the addition of magnetic walls and c) the mesh.

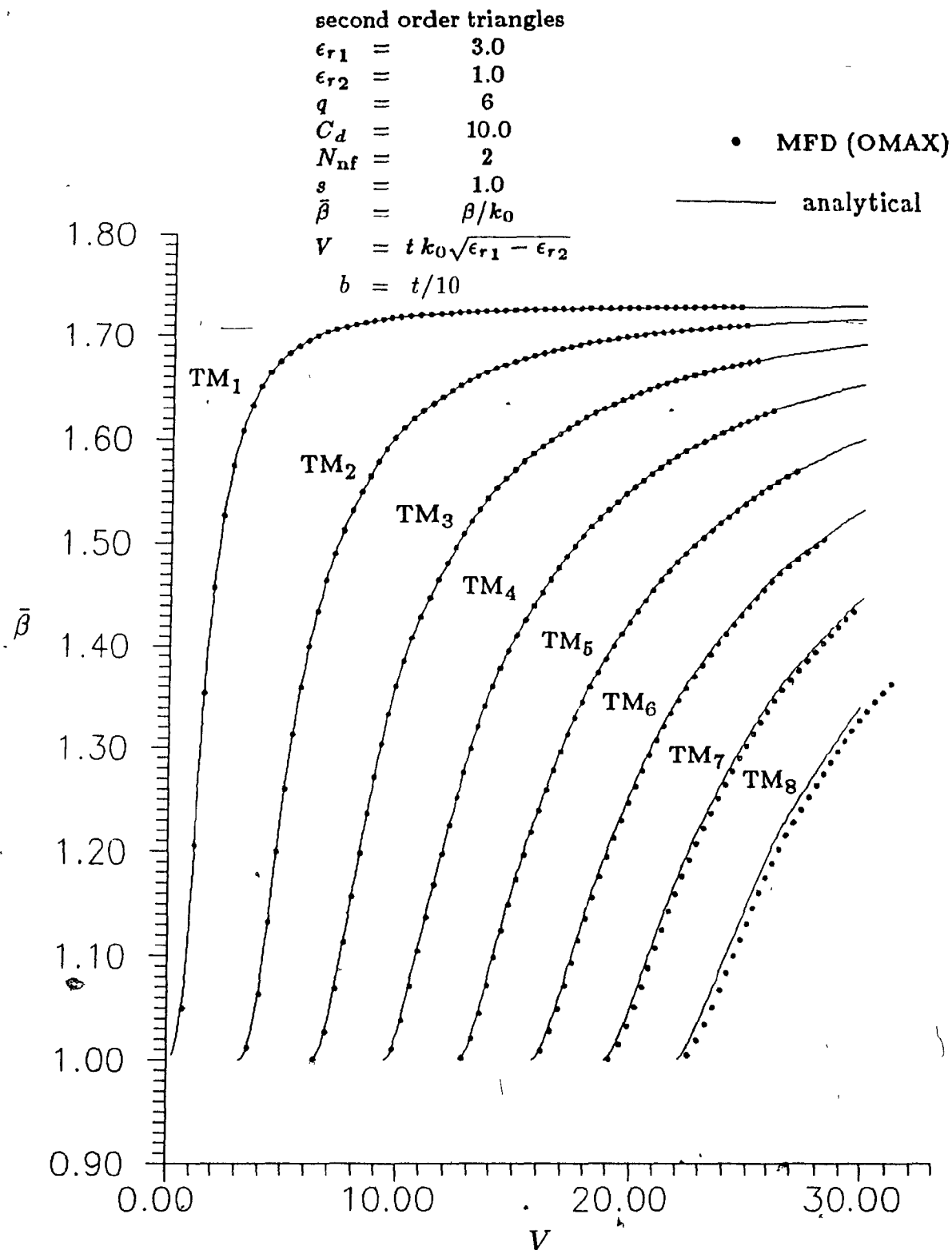


Figure 5.2 The slab waveguide results.

TM modes: $H_z = 0$

$$E_z = \begin{cases} \frac{\sin k_d x}{\sin k_d t} \cos k_m y & \text{for } 0 \leq x < t \\ \frac{e^{-hx}}{e^{-ht}} \cos k_m y & \text{for } x \geq t \end{cases} \quad m = 0, 1, 2, \dots \quad (5.2a)$$

TE modes: $E_z = 0$

$$H_z = \begin{cases} \frac{\cos k_d x}{\cos k_d t} \sin k_m y & \text{for } 0 \leq x < t \\ \frac{e^{-hx}}{e^{-ht}} \sin k_m y & \text{for } x \geq t \end{cases} \quad m = 1, 2, \dots \quad (5.2b)$$

$$\text{where } k_m = \frac{m\pi}{b} \quad (5.2c)$$

which are simultaneous solutions of the following equations

$$k_0^2 \epsilon_{r1} = \beta^2 + k_d^2 + k_m^2 \quad (5.2d)$$

$$k_0^2 \epsilon_{r2} = \beta^2 - h^2 + k_m^2 \quad (5.2e)$$

$$k_d t \tan k_d t = \frac{\epsilon_{r1}}{\epsilon_{r2}} h t \quad (\text{TM modes only}) \quad (5.2f)$$

$$k_d t \cot k_d t = -h t \quad (\text{TE modes only}) \quad (5.2g)$$

The frequencies of the TE and TM modes with $m > 0$ can be made arbitrarily large by reducing b (from (5.2c), this increases the value of the constant k_m in (5.2d) and (5.2e)). The lowest eigenvalues are thus the desired TM modes with no y dependence ($m = 0$).

The results from the OMAX program with second order triangles and 6 decays are shown in Figure 5.2; the normalised $\bar{\beta}$ versus V coordinates are used. The solutions to equations (5.2d,e,f) are also shown as the solid lines. Very close agreement is obtained for the lowest modes from cut-off to the highest frequencies shown on the graph.

3. The Circular Dielectric Waveguide

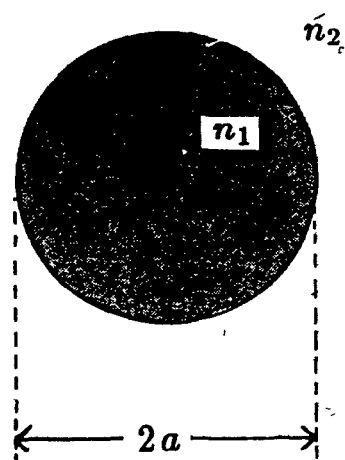
The next example is the weakly-guiding circular waveguide of Figure 5.3a with interior and exterior refractive indices of $n_1 = \sqrt{\epsilon_{r1}} = 1.53$ and $n_2 = \sqrt{\epsilon_{r2}} = 1.50$ respectively, and radius a . The closed-form analytical solutions to the modes of this guide, designated TM_{0u} , TE_{0u} , EH_{vu} and HE_{vu} (where $u, v = 1, 2, \dots$), are given in [68,p.225]. To reduce the problem size, only one quarter of the problem is meshed as illustrated in Figure 5.3b; the mesh is composed of 12 triangular and 4 infinite elements. All modal \mathbf{H} fields have ϕ dependence $e^{jv\phi}$ and therefore it can be shown that they all satisfy either the perfect electric conductor ($\partial\Omega_e$) or the perfect magnetic conductor ($\partial\Omega_o$) boundary conditions on the symmetry lines (the x and y axes). The correspondence between the applied boundary conditions and the resulting modes is shown in Table 5.1. The OMAX program was run with third order elements for each of the three pairs of boundary constraints in the table producing the results in Figures 5.4a,b,c. The largest global matrix size was 335 which occurred when both edges were set to the $\partial\Omega_e$ boundary type. The analytical solutions are shown as the solid lines. The results are most accurate at cut-off and towards the higher frequencies, and the maximum error in k_0 is 0.03%.

Table 5.1

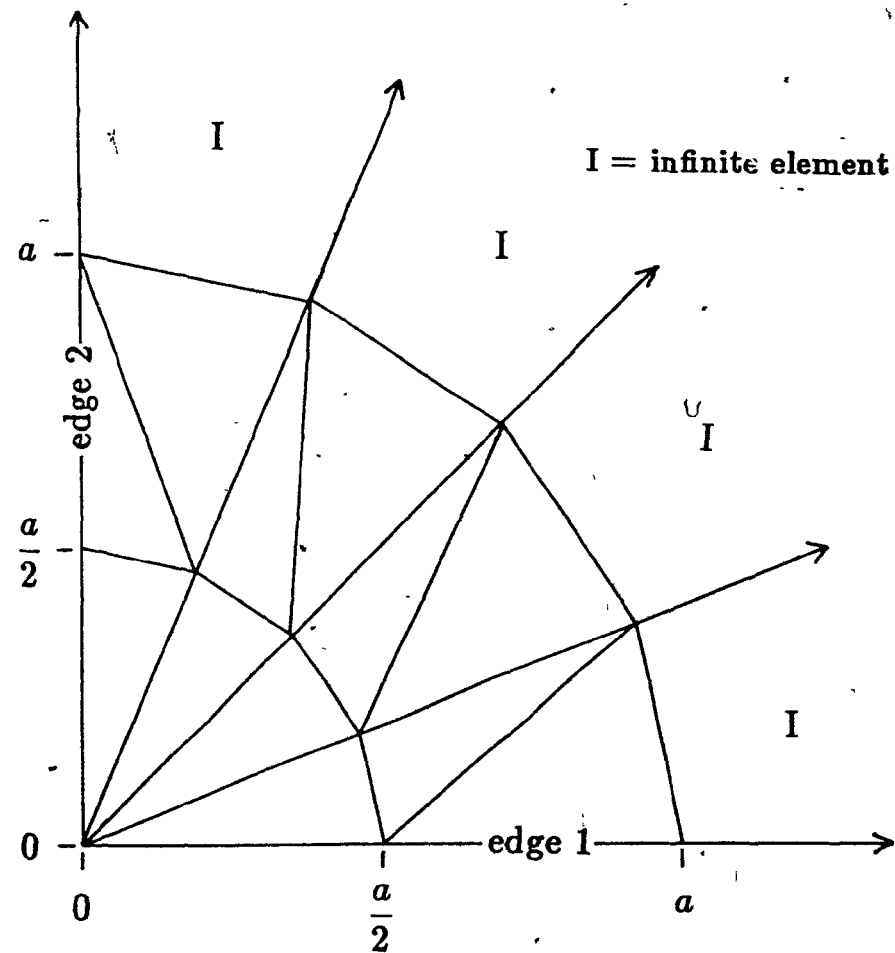
Boundary Conditions		Resulting Modes
edge 1	edge 2	
$\partial\Omega_o$	$\partial\Omega_o$	$TM_{0u}, EH_{vu}, HE_{vu} \quad (v \text{ even})$
$\partial\Omega_e$	$\partial\Omega_e$	$TE_{0u}, EH_{vu}, HE_{vu} \quad (v \text{ even})$
$\partial\Omega_o$	$\partial\Omega_e$	$EH_{vu}, HE_{vu} \quad (v \text{ odd})$

4. The Square Dielectric Waveguide

Using a point-matching method, Goell [28] calculated the dispersion curves of the modes of the square waveguide of Figure 5.5a for which no exact analytical solutions are available. As for the circular waveguide, a quarter mesh was used and the three pairs of edge 1 and 2 boundary conditions were applied separately. The point matching



a)



b)

Figure 5.3 The circular dielectric waveguide example: a) the problem and b) the quarter mesh.

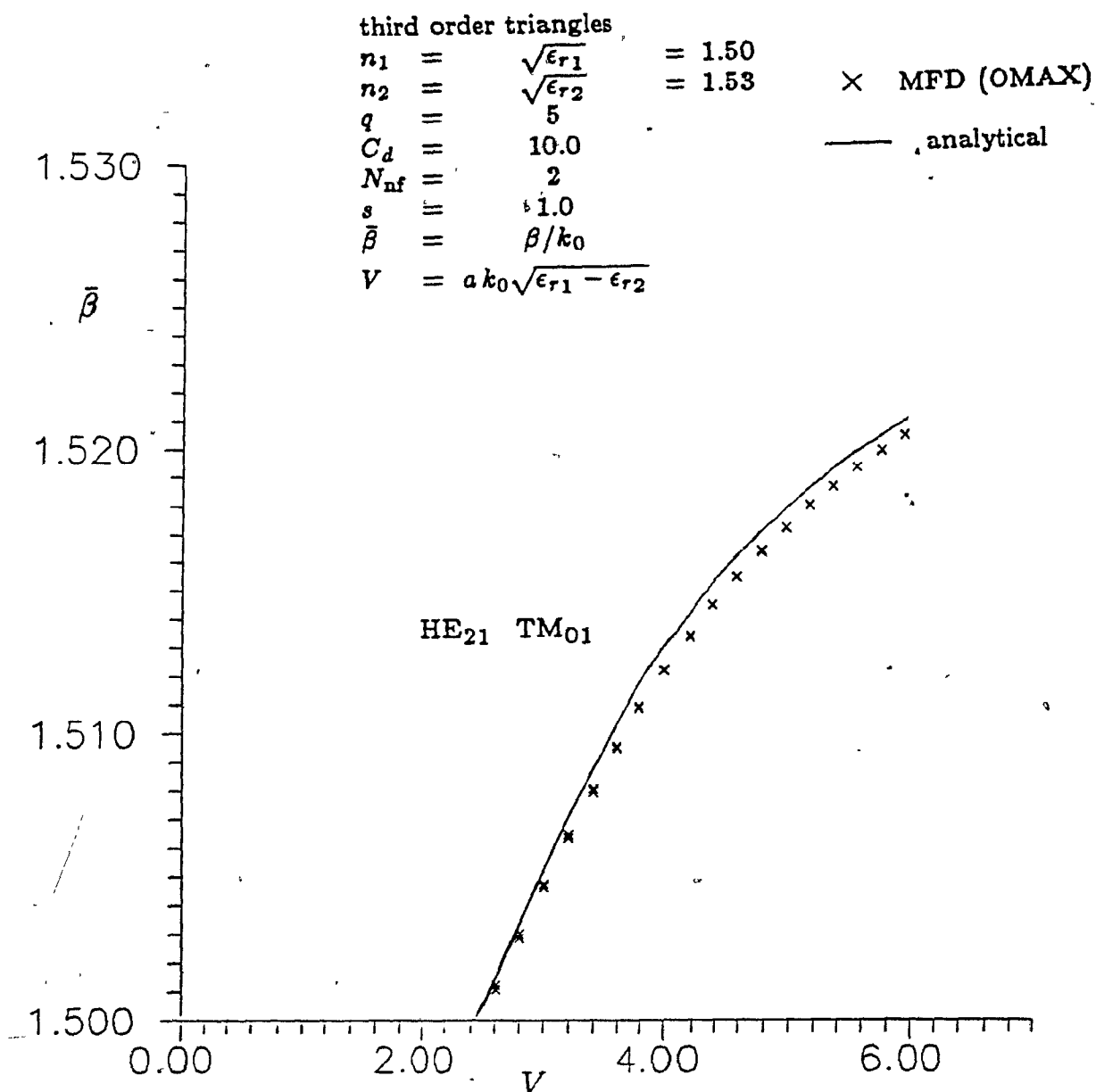


Figure 5.4a The circular dielectric waveguide results using the mesh of Figure 5.3b) with edge 1 = $\partial\Omega_o$ and edge 2 = $\partial\Omega_o$.

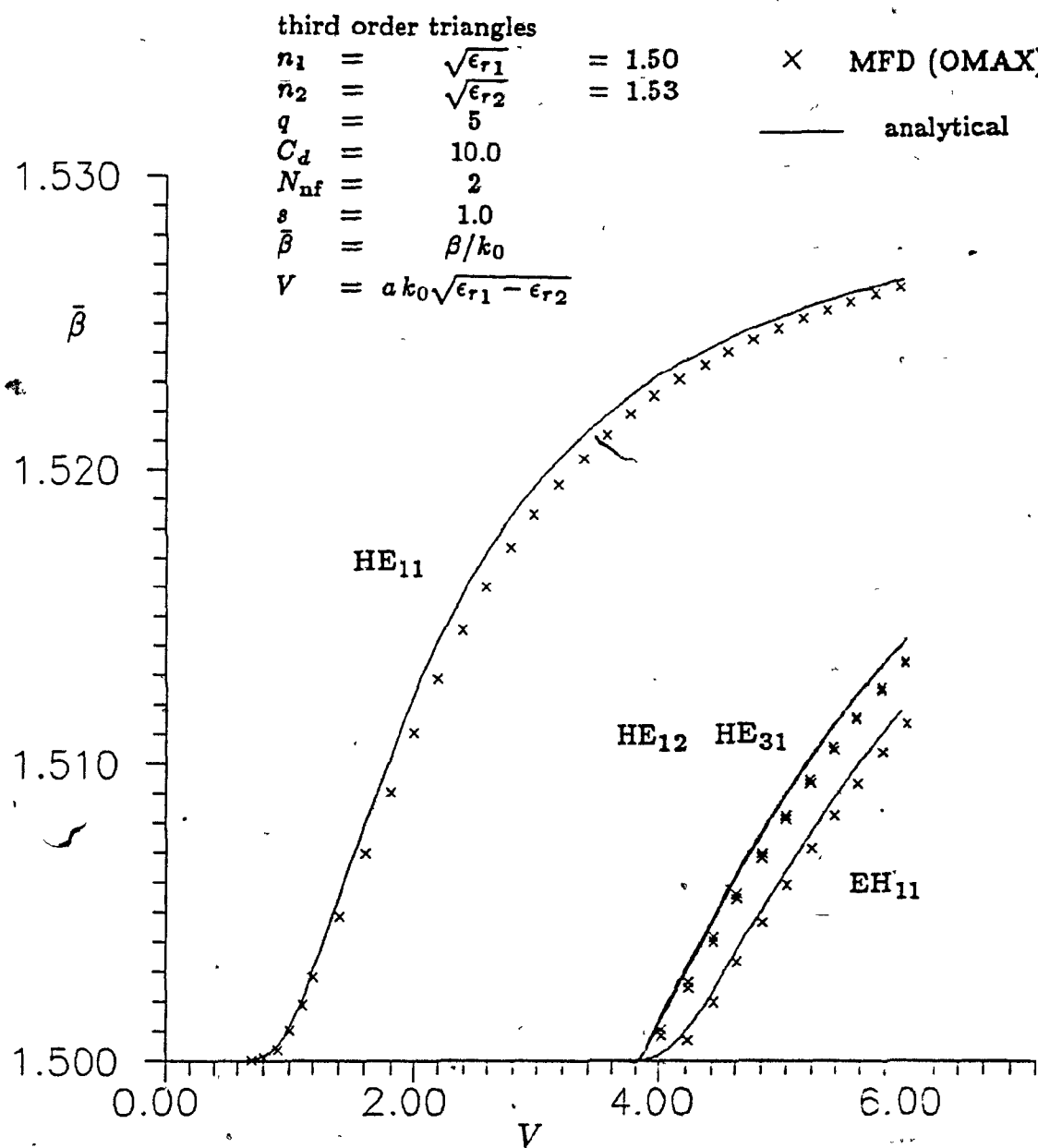


Figure 5.4b The circular dielectric waveguide results using the mesh of Figure 5.3b) with edge 1 = $\partial\Omega_o$ and edge 2 = $\partial\Omega_s$.

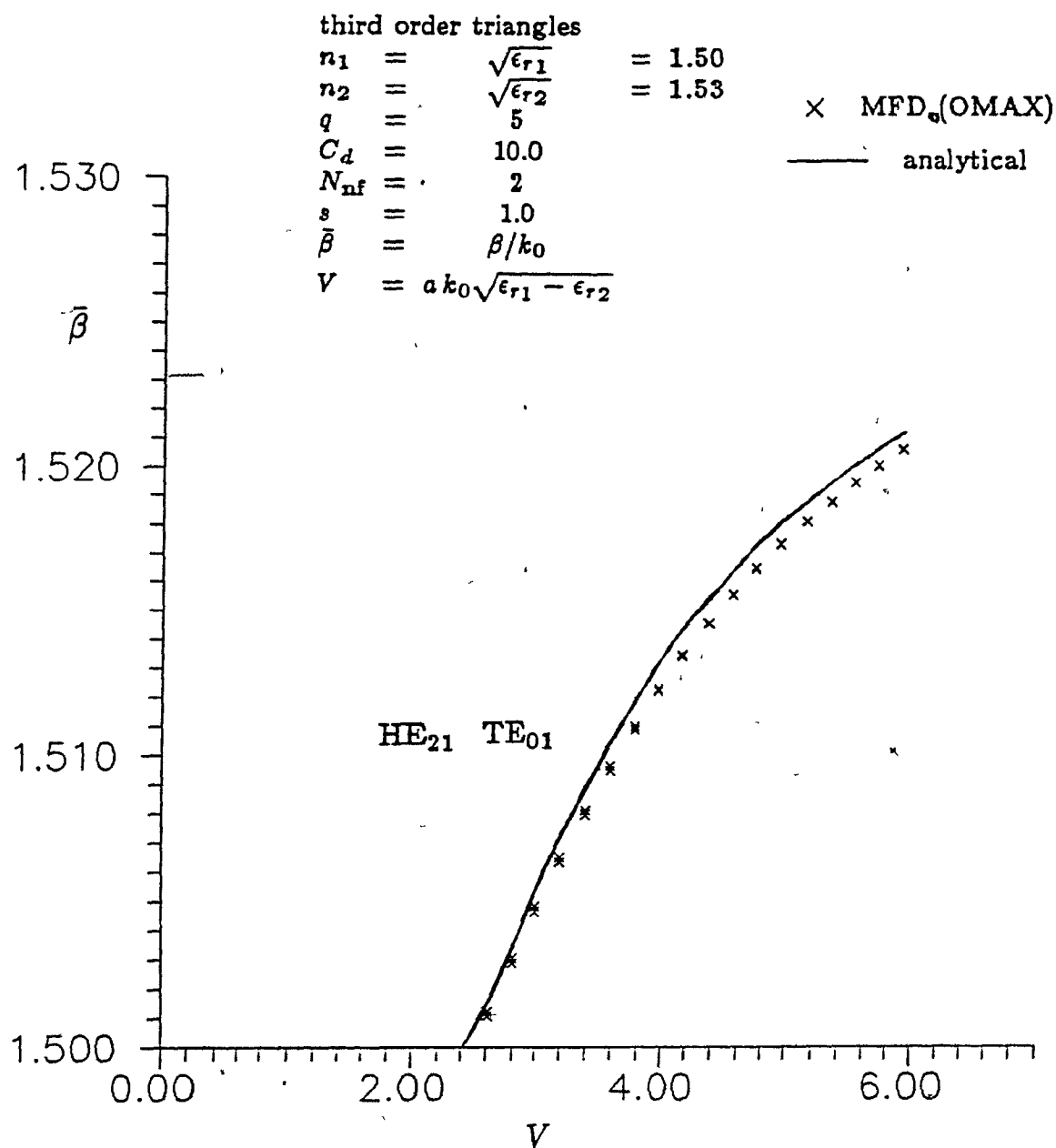
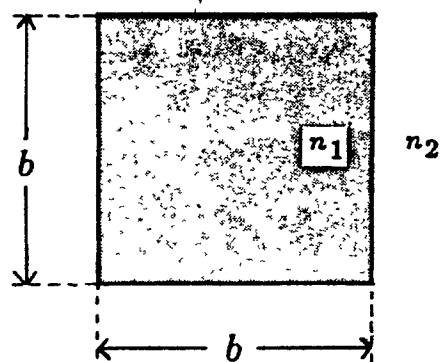
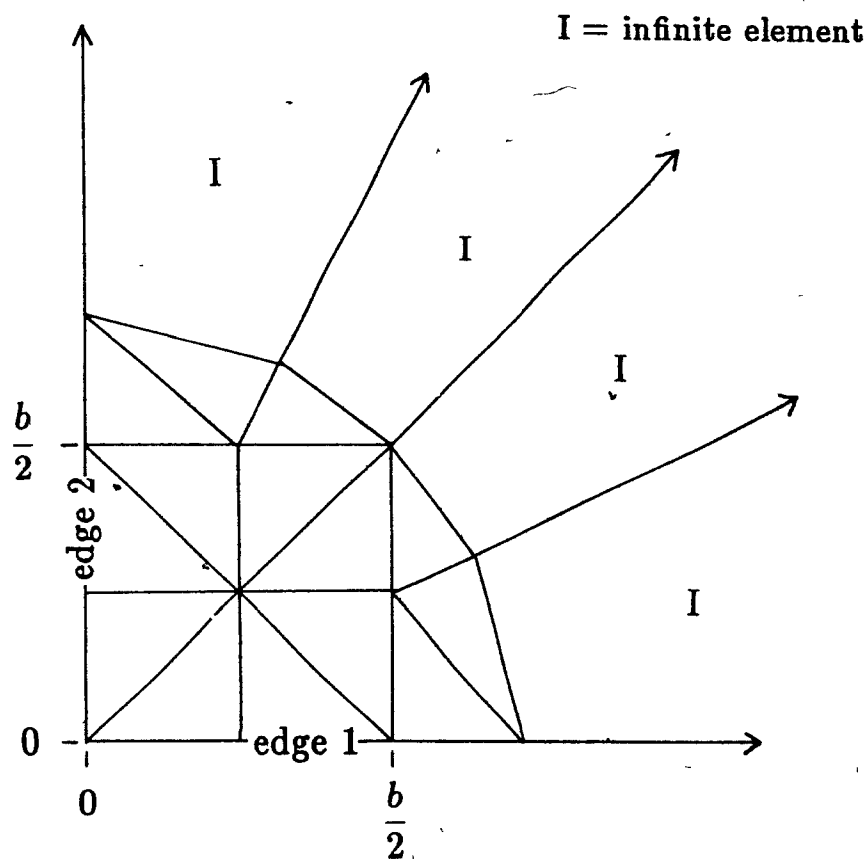


Figure 5.4c The circular dielectric waveguide results using the mesh of Figure 5.3b) with edge 1 = $\partial\Omega_s$ and edge 2 = $\partial\Omega_s$.



a)



b)

Figure 5.5 The square dielectric waveguide example: a) the problem and b) the quarter mesh.

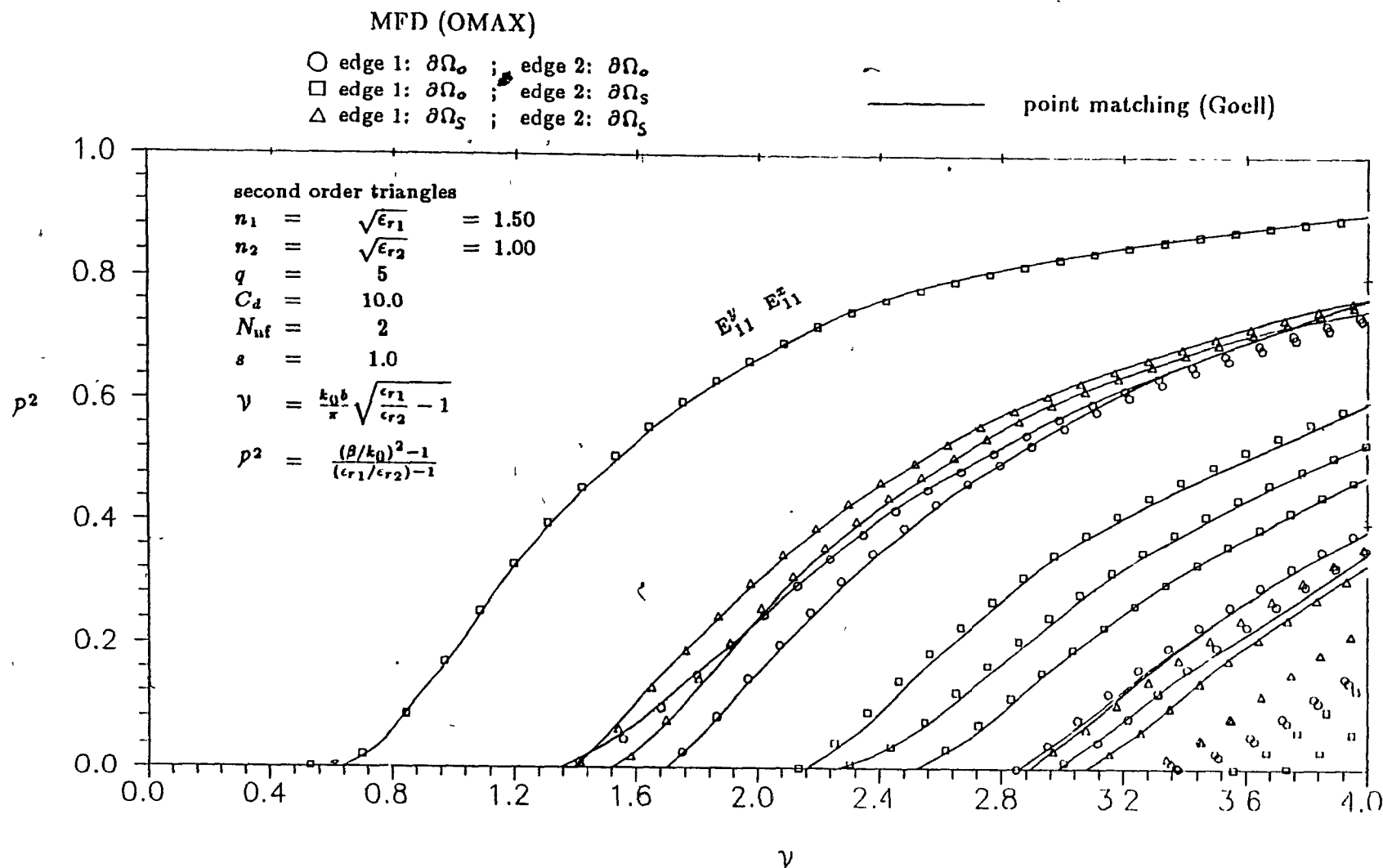


Figure 5.6 The square dielectric waveguide results using the mesh of Figure 5.5b).

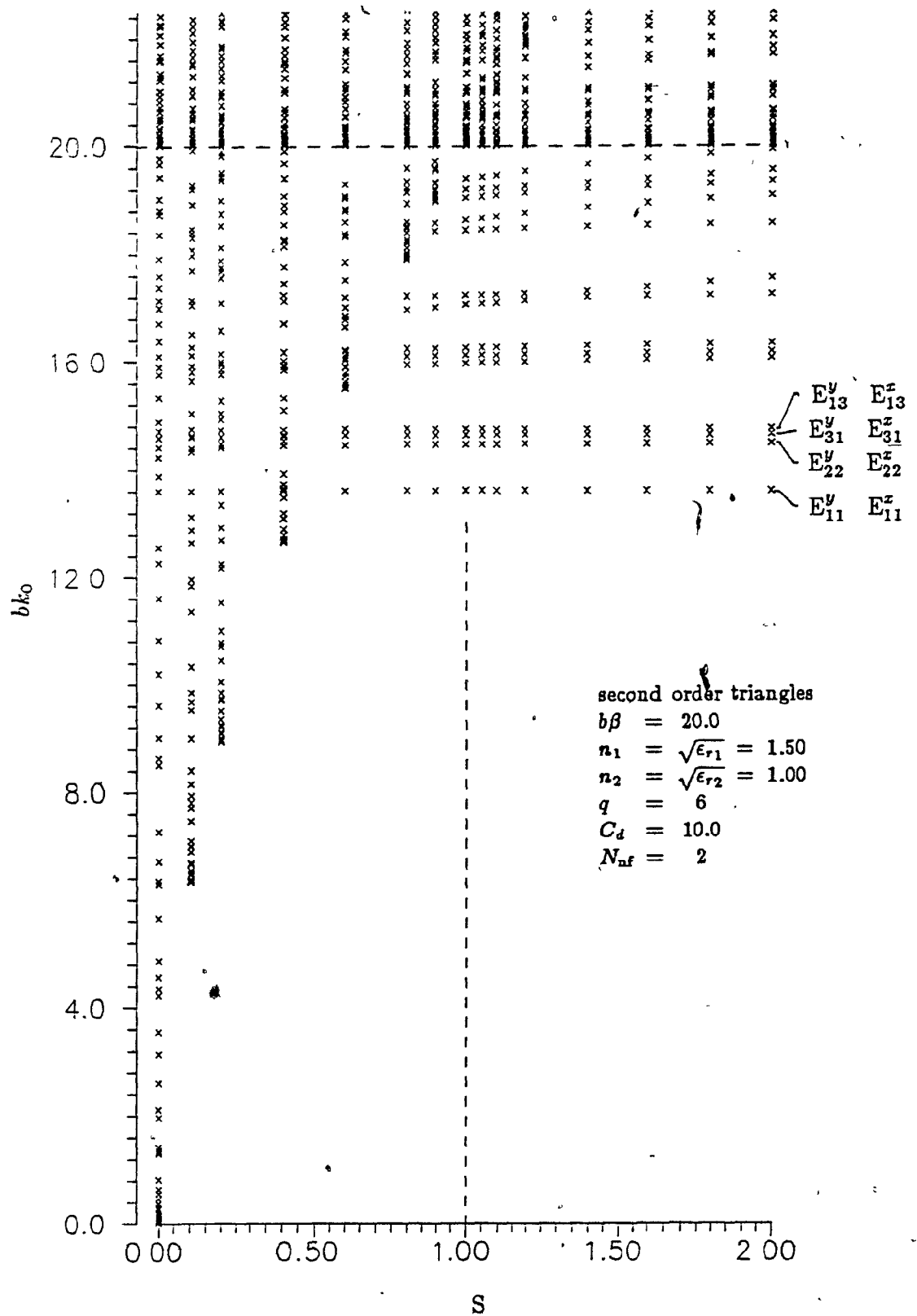


Figure 5.7 The effect of varying s on the spurious mode using the mesh of Figure 5.5b) with edge 1 = $\partial\Omega_o$ and edge 2 = $\partial\Omega_s$.

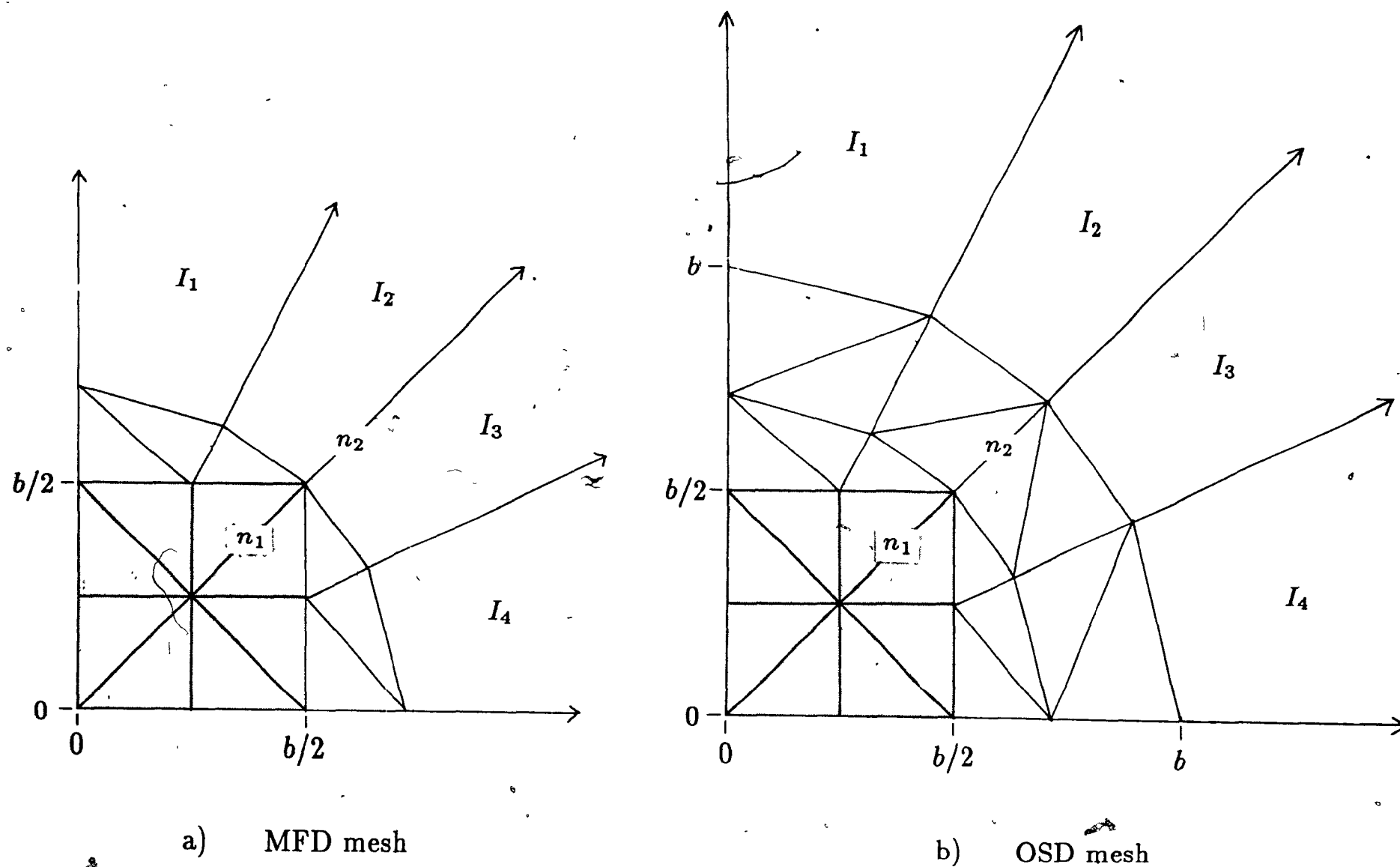
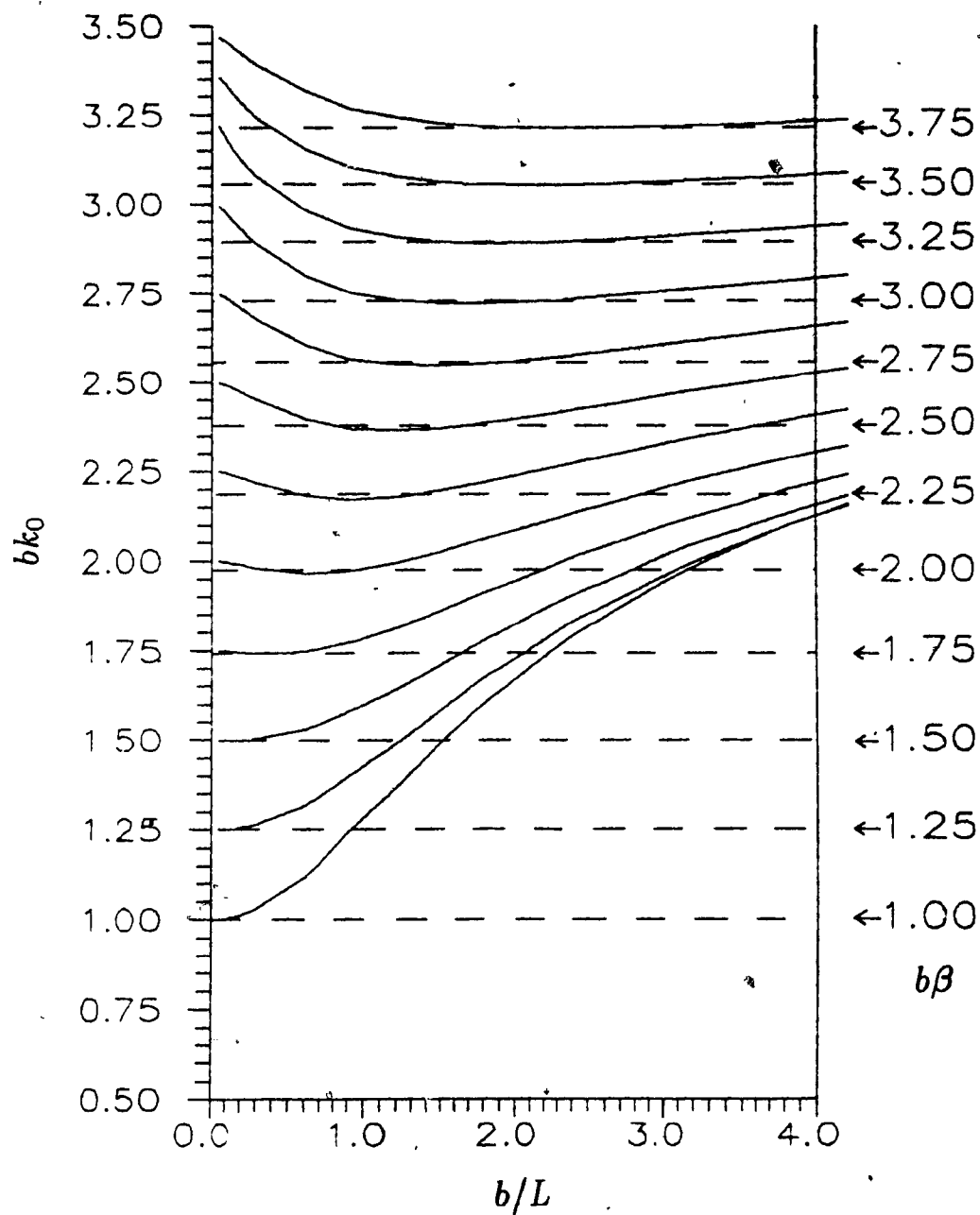


Figure 5.8 The square waveguide meshes used to calculate the results shown in Figure 5.9.



— — — — MFD

————— OSD

mesh of Figure 5.8a)
second order triangles

$$n_1 = \sqrt{\epsilon_{r1}} = 1.50$$

$$n_2 = \sqrt{\epsilon_{r2}} = 1.00$$

$$q = 6$$

$$C_d = 10.0$$

$$N_{nf} = 2$$

$$s = 1.0$$

mesh of Figure 5.8b)
second order triangles

$$n_1 = \sqrt{\epsilon_{r1}} = 1.50$$

$$n_2 = \sqrt{\epsilon_{r2}} = 1.00$$

$$q = 1$$

$$s = 1.0$$

Figure 5.9 A comparison of the OSD and MFD techniques for the fundamental mode of the square waveguide. The solid curves illustrate the variation of $b k_0$ with b/L using the OSD technique for several values of normalised phase constant $b\beta$. The dashed lines represent the results of the MFD technique with 6 fixed decays for the same values of $b\beta$.

results are shown in Figure 5.6 as the solid lines while the OMAX results appear as the circles, squares and triangles.

With the boundary conditions on edges 1 and 2 both set to $\partial\Omega_o$, the execution time for each β value cycle was 12 minutes on a VAX 8650 and the size of the global matrices was 338.

5. The Effect of Varying the Parameter s

It is mentioned in Chapter 2 and proved in Appendix 2 that spurious modes satisfy $k_0 > \sqrt{s}\beta$. To confirm this relation, the OMAX program was executed for the square waveguide problem of the previous example by fixing the value of β at 20 while varying the s parameter between 0 and 2. In Figure 5.7, each column of crosses represents the modes, both spurious and physical, that result from the execution of OMAX for a particular value of s . The total number of modes in each column is 217 (this is the order of the global matrices) but only those which satisfy $bk_0 < 22.4$ are shown.

Figure 5.7 shows how the spurious modes grow in frequency as s increases, whereas the frequencies of the physical modes remain nearly constant. Since $\sqrt{\epsilon_{r1}} = 1.5$, $\sqrt{\epsilon_{r2}} = 1$ and $b = 1$, the spurious modes are eliminated completely from the propagating range $b\beta/\sqrt{\epsilon_{r1}} < bk_0 < b\beta/\sqrt{\epsilon_{r2}}$, or $13.33 < bk_0 < 20$, when s reaches unity. Reference lines at $s = 1$ and $bk_0 = 20$ are included in the figure.

Note that for this elimination method to be effective, all the essential boundary conditions of Variational Principle 2.1 must be applied. Using a similar method, Rahman and Davies [49] describe an open channel waveguide example (on page 927) which requires that s be set to 10^4 for the elimination of spurious modes. This may be due to an incomplete application of boundary conditions.

6. A Comparison of the MFD and OSD Techniques

The drawback of the OSD method is that the single decay length L must be optimized, particularly near cut-off. This is illustrated in Figure 5.9 in which the nor-

malised frequency of the fundamental mode is plotted versus the decay length L (solid lines) for the OSD method applied to the square waveguide mesh of Figure 5.8b. Each curve represents a different value of β which varies between 1 and 3.75. The need for optimization is made apparent by the severe variation of the frequencies with decay length.

The results of the MFD method applied to the mesh of Figure 5.8a are included for comparison (horizontal dashed lines). Even though the mesh is smaller, the results of the MFD technique compare very well with those of OSD.

Chapter Six

Conclusions

Computational analysis and design tools for arbitrarily-shaped and arbitrarily-inhomogeneous unbounded dielectric waveguides (i.e. open guides) are of current interest since analytical closed form solutions are usually not available and determining the required modal characteristics through experiment is expensive and laborious. This thesis presents a numerical technique for the modal analysis of open guides in which infinite elements are used to extend the finite element method to unbounded guides. This technique preserves the linearity of the finite element method and allows the problem to be transformed algebraically into a generalised eigenvalue problem which yields the first p modal eigenfrequencies in one pass of the solver.

Previously, the *optimized single decay* (OSD) method was employed which has the disadvantage of requiring that the decay length be optimized separately for each mode of interest. In this thesis, the *multiple fixed decay* (MFD) method is introduced which supplies each infinite element with a weighted sum of decaying trial functions with fixed decay length. The outer optimization loop is eliminated and each mode is approximated more closely since the infinite elements dispose of more degrees of freedom.

Further improvements can be made to the MFD technique by incorporating a sparse solver into the code, since the matrices generated by the integration of the functional are sparse. The integration itself could be improved by developing an error bound for the exact integration of asymmetrical elements, thus removing the present restriction of symmetrical shape on the infinite elements.

Appendix I

Proof of Variational Principle 2.1

Let \mathcal{D} be the space of complex three component vector functions $\mathbf{H}(x, y, z)$ of the form

$$\mathbf{H}(x, y, z) = \mathbf{H}(x, y)e^{-j\beta z} = [H_x(x, y)\mathbf{i} + H_y(x, y)\mathbf{j} + H_z(x, y)\mathbf{k}]e^{-j\beta z} \quad \beta \neq 0$$

which are infinitely differentiable in Ω , zero outside Ω , square integrable and satisfy boundary conditions (A1.3) (see below). The norm is defined as

$$\|\mathbf{H}\| = \left[\int_{z=0}^{\lambda_0} \int_{\Omega} (\mathbf{H}^* \cdot \mathbf{H}) dx dy dz \right]^{\frac{1}{2}}$$

Two preliminary results that are required for the main proof are presented first. Let $F(\mathbf{H})$, $Q(\mathbf{H})$ and $N(\mathbf{H})$ be functionals which map from \mathcal{D} into the set of real numbers. In particular, let Q and N be

$$\begin{aligned} Q(\mathbf{H}) &= q(\mathbf{H}; \mathbf{H}) \\ N(\mathbf{H}) &= n(\mathbf{H}; \mathbf{H}) \end{aligned} \tag{A1.1}$$

where q and n are *sesquilinear* forms and therefore satisfy the properties

$$\begin{aligned} q(\mathbf{A}; \mathbf{B}) &= q(\mathbf{B}; \mathbf{A})^* \\ n(\mathbf{A}; \mathbf{B}) &= n(\mathbf{B}; \mathbf{A})^* \end{aligned} \tag{A1.2}$$

where \mathbf{A} and \mathbf{B} are arbitrary vector functions in \mathcal{D} .

Let the symbol δF represent the first variation of $F(\mathbf{H})$ with respect to an arbitrary change $\mathbf{H}_a \in \mathcal{D}$, that is

$$F(\mathbf{H} + \mathbf{H}_a) - F(\mathbf{H}) = \delta F + O(\|\mathbf{H}_a\|^2)$$

Lemma A1.1

The first variations of Q and N are given by:

$$\delta Q = \operatorname{Re}\{q(\mathbf{H}_a; \mathbf{H})\}$$

$$\delta N = \operatorname{Re}\{n(\mathbf{H}_a; \mathbf{H})\}$$

Proof:

The change in Q due to a change \mathbf{H}_a is:

$$\begin{aligned} Q(\mathbf{H} + \mathbf{H}_a) - Q(\mathbf{H}) &= q(\mathbf{H}_a + \mathbf{H}; \mathbf{H}_a + \mathbf{H}) - q(\mathbf{H}; \mathbf{H}) \\ &= q(\mathbf{H}; \mathbf{H}) + q(\mathbf{H}; \mathbf{H}_a) + q(\mathbf{H}_a; \mathbf{H}) \\ &\quad + q(\mathbf{H}_a; \mathbf{H}_a) - q(\mathbf{H}; \mathbf{H}) \\ &= q(\mathbf{H}_a; \mathbf{H})^* + q(\mathbf{H}_a; \mathbf{H}) + q(\mathbf{H}_a; \mathbf{H}_a) \quad (\text{from A1.2}) \\ &= 2\operatorname{Re}\{q(\mathbf{H}_a; \mathbf{H})\} + q(\mathbf{H}_a; \mathbf{H}_a) \\ &\equiv \delta Q + O(\|\mathbf{H}_a\|^2) \\ \Rightarrow \delta Q &= 2\operatorname{Re}\{q(\mathbf{H}_a; \mathbf{H})\} \quad \text{Q.E.D.} \end{aligned}$$

Lemma A1.3

If $F = Q/N$ ($N \neq 0$) then the first variation of F is:

$$\delta F = \frac{\delta Q N - Q \delta N}{N^2}$$

Proof:

Writing $Q = FN$, the change in Q due to a change \mathbf{H}_a is:

$$\begin{aligned} Q(\mathbf{H} + \mathbf{H}_a) - Q(\mathbf{H}) &= F(\mathbf{H} + \mathbf{H}_a)N(\mathbf{H} + \mathbf{H}_a) - F(\mathbf{H})N(\mathbf{H}) \\ &= [F(\mathbf{H}) + \delta F + O(\|\mathbf{H}_a\|^2)][N(\mathbf{H}) + \delta N + O(\|\mathbf{H}_a\|^2)] - F(\mathbf{H})N(\mathbf{H}) \\ &= F\delta N + \delta F N + \delta F \delta N + O(\|\mathbf{H}_a\|^2) \\ &\equiv \delta Q + O(\|\mathbf{H}_a\|^2) \end{aligned}$$

and since the term $\delta F \delta N$ is of orders of $\|\mathbf{H}_a\|$ greater than one, we obtain

$$\begin{aligned} \delta Q &= F\delta N + \delta F N \\ \Rightarrow \delta F &= \frac{(\delta Q - F\delta N)}{N} \quad (N \neq 0) \\ &= \frac{\delta Q N - Q\delta N}{N^2} \quad (N \neq 0) \quad \text{Q.E.D.} \end{aligned}$$

We can now proceed with the main result of this Appendix, the proof of Variational Principle 2.1. The boundary conditions that are satisfied by any $\mathbf{H} \in \mathcal{D}$ are

$$\mathbf{H} \times \mathbf{n} = 0 \quad \text{on } \partial\Omega_o \quad (\text{A1.3a})$$

$$s \mathbf{H} \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega_s \quad (\text{A1.3b})$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \mathbf{H} = 0 \quad (\text{A1.3c})$$

$$\lim_{r \rightarrow \infty} \sqrt{r} (\nabla \times \mathbf{H}) = 0 \quad (\text{A1.3d})$$

$$\lim_{r \rightarrow \infty} s \sqrt{r} (\nabla \cdot \mathbf{H}) = 0 \quad (\text{A1.3e})$$

The functional of Variational Principle 2.1 is

$$k_0^2(\mathbf{H}) = \frac{Q(\mathbf{H})}{N(\mathbf{H})} = \frac{q(\mathbf{H}; \mathbf{H})}{n(\mathbf{H}; \mathbf{H})}$$

$$= \frac{\int_{z=0}^{\lambda_0} \int_{\Omega} ((\nabla \times \mathbf{H})^* \cdot K^{-1}(\nabla \times \mathbf{H}) + s(\nabla \cdot \mathbf{H})^*(\nabla \cdot \mathbf{H})) dx dy dz}{\int_{z=0}^{\lambda_0} \int_{\Omega} (\mathbf{H}^* \cdot \mathbf{H}) dx dy dz}$$

where

$$q(\mathbf{A}; \mathbf{B}) = \int_{z=0}^{\lambda_0} \int_{\Omega} ((\nabla \times \mathbf{A})^* \cdot K^{-1}(\nabla \times \mathbf{B}) + s(\nabla \cdot \mathbf{A})^*(\nabla \cdot \mathbf{B})) dx dy dz$$

and

$$n(\mathbf{A}; \mathbf{B}) = \int_{z=0}^{\lambda_0} \int_{\Omega} (\mathbf{A}^* \cdot \mathbf{B}) dx dy dz$$

The matrix K^{-1} is positive definite and $s \geq 0$; the functional $Q(\mathbf{H}) = q(\mathbf{H}; \mathbf{H})$ is therefore positive semidefinite. The functional $N(\mathbf{H}) = n(\mathbf{H}; \mathbf{H})$ is identical to the norm which is positive definite. We therefore have $k_0^2(\mathbf{H}) \geq 0$.

Now since K is real symmetric and therefore Hermitian, we have $K^H = K$ and $KK^{-1} = I \Rightarrow K^H K^{-1} = I \Rightarrow K(K^{-1})^H = I \Rightarrow K^{-1} = (K^{-1})^H$, and

therefore K^{-1} is Hermitian as well. Using this result, it can be shown that q and n satisfy (A1.2):

$$\begin{aligned} q(A; B) &= \int_{z=0}^{\lambda_0} \int_{\Omega} ((\nabla \times A)^* \cdot K^{-1}(\nabla \times B) + s(\nabla \cdot A)^*(\nabla \cdot B)) dx dy dz \\ &= \left[\int_{z=0}^{\lambda_0} \int_{\Omega} ((\nabla \times B)^* \cdot K^{-1}(\nabla \times A) + s(\nabla \cdot A)^*(\nabla \cdot B)) dx dy dz \right]^* \\ &= q(B; A)^* \end{aligned}$$

and

$$\begin{aligned} n(A; B) &= \int_{z=0}^{\lambda_0} \int_{\Omega} (A^* \cdot B) dx dy dz \\ &= \left[\int_{z=0}^{\lambda_0} \int_{\Omega} (B^* \cdot A) dx dy dz \right]^* \\ &= n(B; A)^* \end{aligned}$$

Hence from Theorem A1.1, the first variations of Q and N are

$$\begin{aligned} \delta Q &= 2 \operatorname{Re} \{ q(H_a; H) \} \\ \delta N &= 2 \operatorname{Re} \{ n(H_a; H) \} \end{aligned}$$

and the first variation of $k_0^2(H)$ is

$$\delta k_0^2 = \frac{\delta Q N - Q \delta N}{N^2}$$

from theorem A1.2. Now if H_s is a stationary point of the functional k_0^2 then

$$\left. \delta k_0^2 \right|_{H=H_s} = 0 \Rightarrow$$

$$\frac{1}{N(H_s)} (2 \operatorname{Re} \{ q(H_a; H_s) \} - \frac{Q(H_s)}{N(H_s)} 2 \operatorname{Re} \{ n(H_a; H_s) \}) = 0$$

$$\Rightarrow \operatorname{Re} \{ q(H_a; H_s) \} - k_{0s}^2 \operatorname{Re} \{ n(H_a; H_s) \} = 0$$

$$\Rightarrow \operatorname{Re} \{ q(H_a; H_s) - k_{0s}^2 n(H_a; H_s) \} = 0$$

where $k_{0s}^2 = k_0^2(\mathbf{H}_s)$. When $q(\mathbf{H}_a; \mathbf{H}_s)$ and $n(\mathbf{H}_a; \mathbf{H}_s)$ are expanded, this becomes

$$\begin{aligned} \text{Re} \left\{ \int_{z=0}^{\lambda_0} \int_{\Omega} ((\nabla \times \mathbf{H}_a)^* \cdot (K^{-1} \nabla \times \mathbf{H}_s) + s(\nabla \cdot \mathbf{H}_a)^* (\nabla \cdot \mathbf{H}_s)) dx dy dz \right. \\ \left. - k_{0s}^2 \int_{z=0}^{\lambda_0} \int_{\Omega} (\mathbf{H}_a^* \cdot \mathbf{H}_s) dx dy dz \right\} = 0 \end{aligned}$$

Using the vector identities

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B})$$

$$(\nabla U) \cdot \mathbf{A} = \nabla \cdot (U \mathbf{A}) - U (\nabla \cdot \mathbf{A})$$

we obtain

$$\begin{aligned} \text{Re} \left\{ \int_{z=0}^{\lambda_0} \int_{\Omega} (\nabla \cdot (\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s) + \mathbf{H}_a^* \cdot (\nabla \times (K^{-1} \nabla \times \mathbf{H}_s)) \right. \\ \left. + s [\nabla \cdot ((\nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) - (\nabla (\nabla \cdot \mathbf{H}_s)) \cdot \mathbf{H}_a^*]) dx dy dz \right. \\ \left. - k_{0s}^2 \int_{z=0}^{\lambda_0} \int_{\Omega} (\mathbf{H}_a^* \cdot \mathbf{H}_s) dx dy dz \right\} = 0 \end{aligned}$$

an application of the divergence theorem yields

$$\begin{aligned} \text{Re} \left\{ \int_{z=0}^{\lambda_0} \int_{\Omega} (\mathbf{H}_a^* \cdot (\nabla \times (K^{-1} \nabla \times \mathbf{H}_s)) - s(\nabla (\nabla \cdot \mathbf{H}_s)) \cdot \mathbf{H}_a^* - k_{0s}^2 \mathbf{H}_s \cdot \mathbf{H}_a^*) dx dy dz \right. \\ \left. + \oint_S (\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) \cdot \mathbf{n} dS \right\} = 0 \quad (\text{A2.4}) \end{aligned}$$

where S is the surface enclosing the volume of integration.

We will assume for the moment that \mathbf{H}_a^* is zero on S but otherwise an arbitrary vector in \mathcal{D} . The surface integral is zero in this case and

$$\begin{aligned} \text{Re} \left\{ \int_{z=0}^{\lambda_0} \int_{\Omega} (\mathbf{H}_a^* \cdot (\nabla \times K^{-1} (\nabla \times \mathbf{H}_s)) \right. \\ \left. - s(\nabla (\nabla \cdot \mathbf{H}_s)) \cdot \mathbf{H}_a^* - k_{0s}^2 (\mathbf{H}_s \cdot \mathbf{H}_a^*) dx dy dz \right\} = 0 \end{aligned}$$

Now since \mathbf{H}_a^* can assume any complex value in the volume ($\Omega, 0 \leq z \leq \lambda_0$) this implies:

$$\nabla \times (K^{-1}(\nabla \times \mathbf{H}_s)) - s(\nabla(\nabla \cdot \mathbf{H}_s)) - k_0^2 \mathbf{H}_s = 0 \quad \text{in } \Omega$$

The volume integral of equation (A1.4) therefore vanishes and if we now let \mathbf{H}_a assume any value on the surface S , then equation (A1.4) becomes

$$\operatorname{Re} \left\{ \oint_S (\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) \cdot \mathbf{n} dS \right\} = 0$$

This integral can be considered to consist of three parts:

$$\begin{aligned} & \operatorname{Re} \left\{ \oint_S (\mathbf{H}_a^* \times K^{-1}(\nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) \cdot \mathbf{n} dS \right\} \\ &= \operatorname{Re} \left\{ \int_{\substack{\Omega \\ z=0}} (\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) \cdot \mathbf{n} dS \right\} \\ &+ \operatorname{Re} \left\{ \int_{\substack{\Omega \\ \lambda_0}} (\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) \cdot \mathbf{n} dS \right\} \\ &+ \operatorname{Re} \left\{ \int_{\substack{\partial \Omega \\ 0 \leq z \leq \lambda_0}} (\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) \cdot \mathbf{n} dS \right\} \\ &= 0 \end{aligned} \tag{A2.5}$$

Since

$$\mathbf{H}(x, y, \lambda_0) = \mathbf{H}(x, y) e^{-j\beta\lambda_0} = \mathbf{H}(x, y) e^{-j2\pi} = \mathbf{H}(x, y) e^{-j0} = \mathbf{H}(x, y, 0)$$

the first and second integrals on the R.H.S. of (A1.5) are equal in magnitude but of opposite sign (because the unit normal vector reverses direction), and therefore cancel each other.

We have therefore

$$\begin{aligned} & \operatorname{Re} \left\{ \oint_S (\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s)) \cdot \mathbf{n} dS \right\} \\ &= \operatorname{Re} \left\{ \int_{\substack{\partial \Omega \\ 0 \leq z \leq \lambda_0}} (\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s)) \cdot \mathbf{n} dS \right\} \\ &= 0 \end{aligned}$$

Dividing this integral is into three integrals for each of the surface boundaries that enclose the region Ω :

$$\begin{aligned}
& \operatorname{Re} \left\{ \int_{\substack{\partial\Omega \\ 0 \leq z \leq \lambda_0}} ((\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) \cdot \mathbf{n} dS \right\} \\
&= \operatorname{Re} \left\{ \int_{\substack{\partial\Omega_s \\ 0 \leq z \leq \lambda_0}} (\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) \cdot \mathbf{n} dS \right\} \\
&+ \operatorname{Re} \left\{ \int_{\substack{\partial\Omega_o \\ 0 \leq z \leq \lambda_0}} ((\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) \cdot \mathbf{n} dS \right\} \\
&+ \operatorname{Re} \left\{ \lim_{R \rightarrow \infty} \int_{\substack{r=R \\ 0 \leq \phi \leq 2\pi \\ 0 \leq z \leq \lambda_0}} (\mathbf{H}_a^* \times (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) \mathbf{H}_a^*) \cdot \mathbf{r} r d\phi dz \right\} \\
&= \operatorname{Re} \left\{ \int_{\substack{\partial\Omega_s \\ 0 \leq z \leq \lambda_0}} ((\mathbf{n} \times \mathbf{H}_a^*) \cdot (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) (\mathbf{H}_a^* \cdot \mathbf{n})) dS \right\} \\
&+ \operatorname{Re} \left\{ \int_{\substack{\partial\Omega_o \\ 0 \leq z \leq \lambda_0}} ((\mathbf{n} \times \mathbf{H}_a^*) \cdot (K^{-1} \nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) (\mathbf{H}_a^* \cdot \mathbf{n})) dS \right\} \\
&+ \operatorname{Re} \left\{ \int_{\substack{r=R \\ 0 \leq \phi \leq 2\pi \\ 0 \leq z \leq \lambda_0}} ((\lim_{R \rightarrow \infty} (\sqrt{r} \mathbf{H}_a^*)) \times (\lim_{R \rightarrow \infty} (\sqrt{r} K^{-1} \nabla \times \mathbf{H}_s))) \right. \\
&\quad \left. + ((\lim_{R \rightarrow \infty} (\sqrt{r} s \nabla \cdot \mathbf{H}_s)) (\lim_{R \rightarrow \infty} (\sqrt{r} \mathbf{H}_a^* \cdot \mathbf{n}))) \cdot \mathbf{r} d\phi dz \right\} \\
&= 0
\end{aligned}$$

Note that for the third integral, the integration is performed over the entire $\lim_{R \rightarrow \infty} r = R$ surface, even though certain sections of it may not belong to Ω . This is possible since the \mathbf{H} field is defined as zero outside Ω as illustrated by the example in Figure A1.1. The integrand of the third surface integral is zero in Ω as well because $\mathbf{H}_a^*, \mathbf{H}_s \in \mathcal{D}$ which implies $\lim_{r \rightarrow \infty} (\sqrt{r} \mathbf{H}_a^*) = 0$, $\lim_{r \rightarrow \infty} (\sqrt{r} \nabla \times \mathbf{H}_s) = 0$ and $\lim_{r \rightarrow \infty} (\sqrt{r} s \nabla \cdot \mathbf{H}_s) = 0$. If we make \mathbf{H}_a^* zero on $\partial\Omega_o$, but otherwise an arbitrary vector in \mathcal{D} , then the second integral vanishes and

$$\operatorname{Re} \left\{ \int_{\substack{\partial\Omega_s \\ 0 \leq z \leq \lambda_0}} ((\mathbf{n} \times \mathbf{H}_a^*) \cdot K^{-1} (\nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s) (\mathbf{H}_a^* \cdot \mathbf{n})) dS \right\} = 0$$

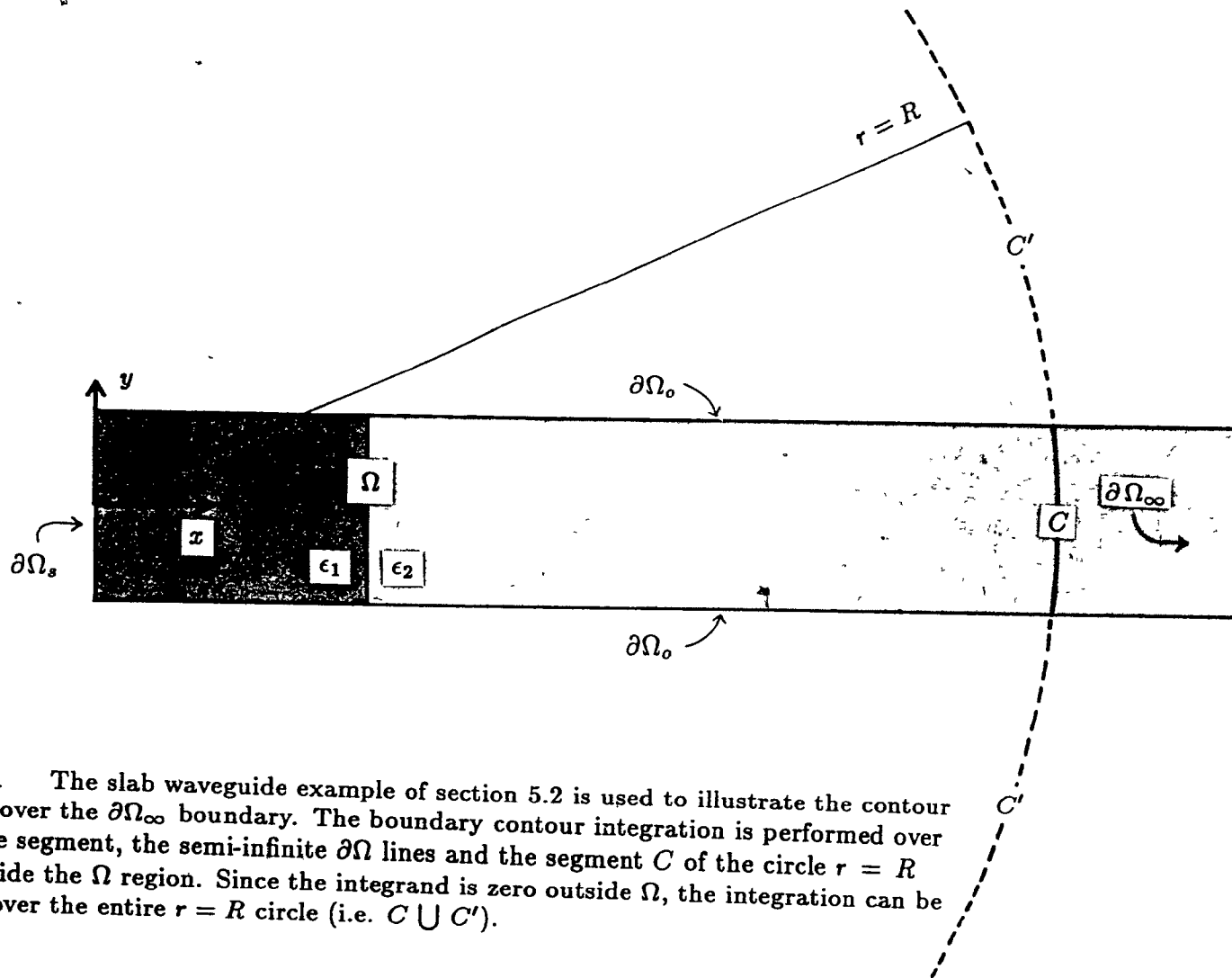


Figure A1.1 The slab waveguide example of section 5.2 is used to illustrate the contour integration over the $\partial\Omega_\infty$ boundary. The boundary contour integration is performed over the $\partial\Omega_s$ line segment, the semi-infinite $\partial\Omega$ lines and the segment C of the circle $r = R$ which is inside the Ω region. Since the integrand is zero outside Ω , the integration can be performed over the entire $r = R$ circle (i.e. $C \cup C'$).

Since $s \mathbf{H}_a^* \cdot \mathbf{n} = 0$ on $\partial\Omega_s$, the second term in the integrand vanishes:

$$\begin{aligned} \operatorname{Re} \left\{ \int_{\substack{\partial\Omega_s \\ 0 \leq z \leq \lambda_0}} ((\mathbf{n} \times \mathbf{H}_a^*) \cdot (K^{-1} \nabla \times \mathbf{H}_s)) dS \right\} &= 0 \\ \Rightarrow \operatorname{Re} \left\{ \int_{\substack{\partial\Omega_s \\ 0 \leq z \leq \lambda_0}} (\mathbf{H}_a^* \cdot ((K^{-1} \nabla \times \mathbf{H}_s) \times \mathbf{n})) dS \right\} &= 0 \end{aligned}$$

Since \mathbf{H}_a^* is arbitrary, this implies that the tangential component of $\nabla \times \mathbf{H}_s$ is zero:

$$(\nabla \times \mathbf{H}_s) \times \mathbf{n} = 0 \quad \text{on } \partial\Omega_s$$

This is the natural boundary condition on $\partial\Omega_s$.

If we now make \mathbf{H}_a^* zero on $\partial\Omega_s$, but otherwise an arbitrary complex vector in \mathcal{D} , then the first integral vanishes and

$$\operatorname{Re} \left\{ \int_{\substack{\partial\Omega_o \\ 0 \leq z \leq \lambda_0}} ((\mathbf{n} \times \mathbf{H}_a^*) \cdot K^{-1}(\nabla \times \mathbf{H}_s) + (s \nabla \cdot \mathbf{H}_s)(\mathbf{H}_a^* \cdot \mathbf{n})) dS \right\} = 0$$

Since $\mathbf{H}_a^* \times \mathbf{n} = 0$ on $\partial\Omega_o$, the first term in the integrand vanishes:

$$\operatorname{Re} \left\{ \int_{\substack{\partial\Omega_o \\ 0 \leq z \leq \lambda_0}} ((s \nabla \cdot \mathbf{H}_s)(\mathbf{H}_a^* \cdot \mathbf{n})) dS \right\} = 0$$

Also, $\mathbf{H}_a^* \times \mathbf{n} = 0$ on $\partial\Omega_o$ implies that \mathbf{H}_a^* is normal to the $\partial\Omega_o$ boundary. This makes $\mathbf{H}_a^* \cdot \mathbf{n}$ an arbitrary non-zero complex number, and hence

$$s \nabla \cdot \mathbf{H}_s = 0 \quad \text{on } \partial\Omega_o$$

This is the natural boundary condition on $\partial\Omega_o$.

Q.E.D.

Appendix II

The Frequency Range of Spurious Modes

In this appendix, the result $k_0 > \sqrt{s}\beta$ for the spurious solutions will be derived. The spurious modes are the solutions to

$$s \nabla(\nabla \cdot \mathbf{H}) + k_0^2 \mathbf{H} = 0 \quad \text{in } \Omega \quad (\text{A2.1a})$$

$$\mathbf{H} \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega_s \quad (\text{A2.1b})$$

$$\nabla \cdot \mathbf{H} = 0 \quad \text{on } \partial\Omega_\sigma \quad (\text{A2.1c})$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \mathbf{H} = 0 \quad (\text{A2.1d})$$

$$\lim_{r \rightarrow \infty} \sqrt{r} (\nabla \cdot \mathbf{H}) = 0 \quad (\text{A2.1e})$$

$$k_0^2 > 0 \quad (\text{A2.1f})$$

where $s > 0$ and $\mathbf{H} \in \mathcal{D}$ (see equation (2.28)). Taking the curl of (A2.1a) and noting that the curl of a gradient is zero, we obtain

$$\begin{aligned} s \nabla \times \nabla(\nabla \cdot \mathbf{H}) + k_0^2 \nabla \times \mathbf{H} &= 0 \quad \text{in } \Omega \\ \Rightarrow \nabla \times \mathbf{H} &= 0 \quad \text{in } \Omega \end{aligned}$$

The solutions to (A2.1) are therefore irrotational and can be written as the gradient of a complex scalar function Ψ as follows:

$$\begin{aligned} \mathbf{H}(x, y, z) &= C_1 + \Psi(x, y, z) \\ &= C_1 + \psi(x, y)e^{-j\beta z} \quad \beta \neq 0 \end{aligned}$$

Any value of C_1 will allow \mathbf{H} to satisfy (A2.1); the value $C_1 = 0$ is selected. Equations (A2.1a) through (A2.1e) can therefore be restated as follows:

$$\begin{aligned} \begin{aligned} &s \nabla(\nabla^2 \Psi) + k_0^2 \nabla \Psi = 0 \quad \text{in } \Omega \\ \Rightarrow &\nabla(\nabla^2 \Psi + \frac{k_0^2}{s} \Psi) = 0 \quad \text{in } \Omega \\ \Rightarrow &\nabla^2 \Psi + \frac{k_0^2}{s} \Psi = \text{constant} = C_2 \quad \text{in } \Omega \\ \Rightarrow &[\nabla_i^2 \psi + (\frac{k_0^2}{s} - \beta^2) \psi] e^{-j\beta z} = C_2 \quad \text{in } \Omega \end{aligned} & \quad (\text{A2.2a}) \\ (\nabla_i \equiv \partial/\partial x \mathbf{i} + \partial/\partial y \mathbf{j}) & \end{aligned}$$

$$\begin{aligned}
\nabla \Psi \cdot \mathbf{n} &= \frac{\partial \Psi}{\partial n} = \frac{\partial}{\partial n}(\psi e^{-j\beta z}) \\
&= \left(\frac{\partial \psi}{\partial n} \right) e^{-j\beta z} = 0 \quad \text{on } \partial\Omega_s \\
&\Rightarrow \frac{\partial \psi}{\partial n} = 0 \quad \text{on } \partial\Omega_s \quad (\text{A2.2b})
\end{aligned}$$

$$\begin{aligned}
\nabla^2 \Psi &= (\nabla_t^2 \psi - \beta^2 \psi) e^{-j\beta z} = 0 \quad \text{on } \partial\Omega_o \\
&\Rightarrow \nabla_t^2 \psi - \beta^2 \psi = 0 \quad \text{on } \partial\Omega_o \quad (\text{A2.2c})
\end{aligned}$$

$$\begin{aligned}
\lim_{r \rightarrow \infty} \sqrt{r} \nabla \Psi &= \lim_{r \rightarrow \infty} \sqrt{r} (\nabla_t \psi - j\beta \psi \mathbf{k}) e^{-j\beta z} = 0 \\
&\Rightarrow \lim_{r \rightarrow \infty} \sqrt{r} (\nabla_t \psi - j\beta \psi \mathbf{k}) = 0 \quad (\text{A2.2d})
\end{aligned}$$

$$\begin{aligned}
\lim_{r \rightarrow \infty} \sqrt{r} \nabla^2 \Psi &= \lim_{r \rightarrow \infty} \sqrt{r} (\nabla_t^2 \psi - \beta^2 \psi) e^{-j\beta z} = 0 \\
&\Rightarrow \lim_{r \rightarrow \infty} \sqrt{r} (\nabla_t^2 \psi - \beta^2 \psi) = 0 \quad (\text{A2.2e})
\end{aligned}$$

$$k_0^2 > 0 \quad (\text{A2.2f})$$

Equation (A2.2d) can be split into two equations:

$$\lim_{r \rightarrow \infty} \sqrt{r} \nabla_t \psi = 0 \quad (\text{A2.5d})$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \psi = 0 \quad (\text{A2.3})$$

Substituting (A2.3) into (A2.2e), we obtain also

$$\begin{aligned}
&\lim_{r \rightarrow \infty} \sqrt{r} (\nabla_t^2 \psi - \beta^2 \psi) = 0 \\
&\Rightarrow \lim_{r \rightarrow \infty} (\sqrt{r} \nabla_t^2 \psi) - \beta^2 \lim_{r \rightarrow \infty} (\sqrt{r} \psi) = 0 \\
&\Rightarrow \lim_{r \rightarrow \infty} \sqrt{r} \nabla_t^2 \psi = 0 \quad (\text{A2.4})
\end{aligned}$$

Multiplying both sides of (A2.2a) by \sqrt{r} , taking the limit as $r \rightarrow \infty$ and then substituting (A2.5d) and (A2.3):

$$\begin{aligned}
&\lim_{r \rightarrow \infty} \sqrt{r} ([\nabla_t^2 \psi + (\frac{k_0^2}{s} - \beta^2) \psi] e^{-j\beta z}) = \lim_{r \rightarrow \infty} \sqrt{r} C_2 \\
&\Rightarrow [(\lim_{r \rightarrow \infty} \sqrt{r} \nabla_t^2 \psi) + (\frac{k_0^2}{s} - \beta^2)(\lim_{r \rightarrow \infty} \sqrt{r} \psi)] e^{-j\beta z} = \lim_{r \rightarrow \infty} \sqrt{r} C_2 \\
&\Rightarrow 0 = \lim_{r \rightarrow \infty} \sqrt{r} C_2 \\
&\Rightarrow C_2 = 0
\end{aligned}$$

The constant C_2 must vanish or else the R.H.S. is infinite in magnitude. Equation (A2.2a) becomes:

$$\begin{aligned} [\nabla_t^2 \psi + (\frac{k_0^2}{s} - \beta^2) \psi] e^{-j\beta z} &= 0 \quad \text{in } \Omega \\ \Rightarrow \nabla_t^2 \psi + (\frac{k_0^2}{s} - \beta^2) \psi &= 0 \quad \text{in } \Omega \end{aligned} \quad (\text{A2.5a})$$

Further simplifications to (A2.2) can be made using

$$\begin{aligned} \nabla_t^2 \psi - \beta^2 \psi &= -(\frac{k_0^2}{s} - \beta^2) \psi + \beta^2 \psi \quad (\text{from A2.5a}) \\ &= -\frac{k_0^2}{s} \psi \end{aligned}$$

Making use of this relation and $k_0 \neq 0$, equations (A2.2c) and (A2.2e) become

$$\begin{aligned} \nabla_t^2 \psi - \beta^2 \psi &= 0 \quad \text{on } \partial\Omega_s \\ \Rightarrow \frac{k_0^2}{s} \psi &= 0 \quad \text{on } \partial\Omega_s \\ \Rightarrow \psi &= 0 \quad \text{on } \partial\Omega_s \end{aligned} \quad (\text{A2.3c})$$

and

$$\begin{aligned} \lim_{r \rightarrow \infty} \sqrt{r} (\nabla_t^2 \psi - \beta^2 \psi) &= 0 \\ \Rightarrow \lim_{r \rightarrow \infty} \sqrt{r} \frac{k_0^2}{s} \psi &= 0 \\ \Rightarrow \lim_{r \rightarrow \infty} \sqrt{r} \psi &= 0 \end{aligned} \quad (\text{A2.3e})$$

In summary, equation (A2.2) simplifies to

$$\nabla_t^2 \psi + (\frac{k_0^2}{s} - \beta^2) \psi = 0 \quad \text{in } \Omega \quad (\text{A2.5a})$$

$$\frac{\partial \psi}{\partial n} = 0 \quad \text{on } \partial\Omega_s \quad (\text{A2.5b})$$

$$\psi = 0 \quad \text{on } \partial\Omega_o \quad (\text{A2.5c})$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \nabla_t \psi = 0 \quad (\text{A2.5d})$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \psi = 0 \quad (\text{A2.5e})$$

$$k_0^2 > 0 \quad (\text{A2.5f})$$

For the special case of closed guides which are completely enclosed by $\partial\Omega$, or $\partial\Omega_0$ boundaries, it has been noted [50] that the solutions to equations (A2.5a), (A2.5b) and (A2.5c) only are equivalent to the TE modes of an equivalent homogeneous guide. These TE modes are known to satisfy $k_0 > \sqrt{s}\beta$. To prove this relation for open guides, consider now the following variational principle:

Variational Principle A2.1

Let S be the space of scalar functions $u(x, y)$ which are infinitely integrable in Ω , satisfy $\nabla_t u = 0$ outside Ω , are square integrable and satisfy the prescribed boundary conditions:

$$u = 0 \quad \text{on } \partial\Omega_0 \quad (\text{A2.6c})$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \nabla_t u = 0 \quad (\text{A2.6d})$$

$$\lim_{r \rightarrow \infty} \sqrt{r} u = 0 \quad (\text{A2.6e})$$

then the stationary points u_i of the functional

$$F(u) = \frac{\int_{\Omega} (\nabla_t u)^* \cdot (\nabla_t u) dx dy dz}{\int_{\Omega} u^* u dx dy dz} ; \quad u \neq 0 \quad (\text{A2.7})$$

are solutions to the differential equation:

$$\nabla_t u + F_i(u) = 0 \quad \text{in } \Omega \quad (\text{A2.6a})$$

$$\text{where } F_i = F(u_i) > 0 \quad (\text{A2.6f})$$

and satisfy the natural boundary condition

$$\frac{\partial u_i}{\partial n} = 0 \quad \text{on } \partial\Omega_0 \quad (\text{A2.6b})$$

in addition to the above prescribed boundary conditions (A2.6c), (A2.6d) and (A2.6e).

Proof:

First, we define

$$F(u) = \frac{Q(u)}{N(u)} = \frac{q(u; u)}{n(u; u)}$$

where

$$q(a; b) = \int_{\Omega} (\nabla_t a)^* \cdot (\nabla_t b) dx dy$$

$$n(a; b) = \int_{\Omega} a^* b dx dy$$

Since Q is positive semidefinite and N is positive definite, we have $F(u) \geq 0$. Since $q(a; b) = [q(a; b)]^*$ and $n(a; b) = [n(a; b)]^*$, the first variations of Q and N with respect to an arbitrary change u_a are:

$$\delta Q = 2 \operatorname{Re}\{q(u_a; u)\}$$

$$\delta N = 2 \operatorname{Re}\{n(u_a; u)\}$$

also, the first variation of F is

$$\delta F = \frac{1}{N} (\delta Q - \frac{Q}{N} \delta N) = 0 \quad (N \neq 0)$$

$$\Rightarrow \delta Q - \frac{Q}{N} \delta N = 0$$

$$\Rightarrow \operatorname{Re}\left\{\int_{\Omega} (\nabla_t u_a)^* \cdot (\nabla_t u_s) - F(u_s) u_a^* u_s dx dy\right\} = 0$$

the proofs of these three results are similar to those of Lemmas A1.1 and A1.2, with H and H_a replaced with u and u_a respectively. Setting $\delta F = 0$ to find the stationary points u_s , there obtains

$$\delta F = \frac{1}{N} (\delta Q - \frac{Q}{N} \delta N) = 0 \quad (N \neq 0)$$

$$\Rightarrow \delta Q - \frac{Q}{N} \delta N = 0$$

$$\Rightarrow \operatorname{Re}\left\{\int_{\Omega} (\nabla_t u_a)^* \cdot (\nabla_t u_s) - F(u_s) u_a^* u_s dx dy\right\} = 0$$

using the vector identity $\nabla_t \cdot (UA) = (\nabla_t U) \cdot A + U \nabla_t \cdot A$ with $U = u_a^*$ and $A = \nabla_t u_s$, so that

$$\begin{aligned} \nabla_t \cdot (u_a^* \nabla_t u_s) &= (\nabla_t u_a^*) \cdot \nabla_t u_s + u_a^* \nabla_t^2 u_s \\ \Rightarrow (\nabla_t u_a^*) \cdot \nabla_t u_s &= \nabla_t \cdot (u_a^* \nabla_t u_s) - u_a^* \nabla_t^2 u_s \end{aligned} \quad (A2.7)$$

there obtains

$$\begin{aligned} \operatorname{Re} \left\{ \int_{\Omega} (\nabla_t \cdot (u_a^* \nabla_t u_s) - u_a^* \nabla_t^2 u_s - F(u_s) u_a^* u_s) dx dy \right\} &= 0 \\ \Rightarrow \operatorname{Re} \left\{ \int_{\Omega} u_a^* (\nabla_t^2 u_s + F(u_s) u_s) dx dy - \oint_C (u_a^* \nabla_t u_s) \cdot \mathbf{n} dl \right\} &= 0 \end{aligned} \quad (\text{A2.8})$$

where the divergence theorem has been used and C is the boundary contour enclosing Ω . Making u_a^* zero on the boundary but otherwise arbitrary in Ω , the surface integral vanishes and we have that all u_s satisfy the differential equation

$$\nabla_t^2 u + F(u_s) u = 0 \quad (\text{A2.6a})$$

And therefore from (A2.8) and (A2.6a)

$$\oint_C (u_a^* \nabla_t u_s) \cdot \mathbf{n} dl = 0$$

This integral can be divided into three integrals for the different boundary types:

$$\begin{aligned} \oint_C (u_a^* \nabla_t u_s) \cdot \mathbf{n} dl &= \int_{\partial\Omega_0} (u_a^* \nabla_t u_s) \cdot \mathbf{n} dl \\ &+ \int_{\partial\Omega_s} (u_a^* \nabla_t u_s) \cdot \mathbf{n} dl \\ &+ \lim_{R \rightarrow \infty} \int_{r=R} (u_a^* \nabla_t u_s) \cdot \mathbf{n} dl \\ &= 0 \end{aligned}$$

Since $\nabla_t u$ is 0 outside Ω , the full circle $r = R$ has been included in the third integral, even though certain sections of this circle may not be in Ω (see Figure A2.1). The integrand of the first integral is zero since $u_a^* = 0$ on $\partial\Omega_s$. The integrand of the third integral is also zero since $(\lim_{r \rightarrow \infty} u_a^* \nabla_t u) r = (\lim_{r \rightarrow \infty} \sqrt{r} u_a^*) (\lim_{r \rightarrow \infty} \sqrt{r} \nabla_t u) = 0$. We are therefore left with

$$\int_{\partial\Omega_s} (u_a^* \nabla_t u) \cdot \mathbf{n} dl = 0$$

Since u_a^* is arbitrary, this implies the natural boundary condition

$$(\nabla_t u) \cdot \mathbf{n} = \frac{\partial u}{\partial n} = 0 \quad \text{on } \partial\Omega_s$$

Q.E.D.

Before stating the main result of this Appendix, it still remains to prove that for all admissible ψ we have $\psi \in \mathcal{S}$. Since \mathbf{H} is infinitely integrable in Ω and zero outside Ω and there are no ψ functions for which $\mathbf{H} = \nabla_t \psi e^{-\beta z}$ does not exist, it follows that ψ is infinitely integrable in Ω and that $\nabla_t \psi = 0$ outside Ω . Now from the square integrable property of \mathbf{H} ,

$$\begin{aligned} & \int_{z=z_1}^{z_2} \int_{\Omega} \mathbf{H}^* \cdot \mathbf{H} \, dx \, dy \, dz < \infty \quad (z_1 \neq z_2) \\ \Rightarrow & \int_{z=z_1}^{z_2} \int_{\Omega} (\nabla \Psi)^* \cdot \nabla \Psi \, dx \, dy \, dz \\ & = \int_{z=z_1}^{z_2} \int_{\Omega} (\nabla(\psi e^{-j\beta z}))^* \cdot \nabla(\psi e^{-j\beta z}) \, dx \, dy \, dz \\ & = \int_{z=z_1}^{z_2} \int_{\Omega} ((\nabla_t \psi - j\beta \psi) e^{-j\beta z})^* \cdot (\nabla_t \psi - j\beta \psi) e^{-j\beta z} \, dx \, dy \, dz \\ & = (z_1 - z_2) \int_{\Omega} ((\nabla_t \psi)^* \cdot (\nabla_t \psi) + \beta^2 \psi^* \psi) \, dx \, dy < \infty \quad (z_1 \neq z_2) \\ \Rightarrow & \int_{\Omega} ((\nabla_t \psi)^* \cdot (\nabla_t \psi) + \beta^2 \psi^* \psi) \, dx \, dy < \infty \end{aligned}$$

substituting (A2.7) and (A2.6a), then applying the divergence theorem,

$$\begin{aligned} & \int_{\Omega} ((\nabla_t \psi)^* \cdot (\nabla_t \psi) + \beta^2 \psi^* \psi) \, dx \, dy \\ & = \int_{\Omega} (\nabla_t \cdot (\psi^* \nabla_t \psi) - \psi^* \nabla_t^2 \psi + \beta^2 \psi^* \psi) \, dx \, dy \\ & = \int_{\Omega} (\nabla_t \cdot (\psi^* \nabla_t \psi) - \psi^* (\nabla_t^2 \psi - \beta^2 \psi)) \, dx \, dy \end{aligned}$$

$$\begin{aligned}
&= \int_{\Omega} (\nabla_t \cdot (\psi^* \nabla_t \psi) - \frac{k_0^2}{s} \psi^* \psi) dx dy \\
&= \oint_C (\psi^* \nabla_t \psi) \cdot \mathbf{n} dl - \int_{\Omega} \frac{k_0^2}{s} \psi^* \psi dx dy < \infty
\end{aligned}$$

In a manner similar to that in the proof of Variational Principle A2.1, using (A2.5b), (A2.5c), (A2.5d) and (A2.5e), it can be shown that the contour integral vanishes. Since $k_0^2 > 0$ (from (A2.5f)),

$$\int_{\Omega} \psi^* \psi dx dy < \infty$$

so ψ is square integrable. And since equations (A2.5c,d,e) and (A2.6c,d,e) are identical, therefore $\psi \in S$ and all solutions to (A2.5) are also stationary points of Variational Principle A2.1. The trivial solution $\psi = 0$ is excluded since $\psi = 0 \Rightarrow \nabla_t \mathbf{H} = 0$. For any spurious solution we have therefore

$$\begin{aligned}
\frac{k_0^2}{s} - \beta^2 &\equiv F(\psi) > 0 \\
\Rightarrow k_0 &> \sqrt{s} \beta
\end{aligned}$$

Q.E.D.

References

- [1] Bathe, K.J., Wilson, E.L.,
"Solutions methods for eigenvalue problems in structural mechanics",
Int. J. Num. Met. Engng., vol. 6, pp. 213-226, 1973
- [2] Garbow, B.S., Boyle, J.M., Dongarra, J.J., Moler, C.B.,
Matrix Eigensystem Routines - EISPACK Guide Extension,
Berlin, Germany: Springer-Verlag, 1977
- [3] Zienkiewicz, O.C.,
The finite element method,
3rd ed., London, England: McGraw-Hill, 1977
- [4] Silvester, P.P., Ferrari, R.L.,
Finite elements for electrical engineers,
Cambridge, England: Cambridge University Press, 1983
- [5] Silvester, P.,
"A general high-order finite-element waveguide analysis program",
IEEE Trans. Microwave Theory Tech., MTT-17, no. 4, pp. 204-210, April 1969
- [6] Hano, M.,
"Finite-element analysis of dielectric-loaded waveguides",
IEEE Trans. Microwave Theory Tech., MTT-32, no. 10, pp. 1275-1279, October 1984
- [7] Ahmed, S., Daly, P.,
"Finite element methods for inhomogeneous waveguides",
Proc. IEE, vol. 116, pp. 1661-1664, October 1969
- [8] Daly, P.,
"Hybrid-mode analysis of microstrip by finite-element methods",
IEEE Trans. Microwave Theory Tech., MTT-19, pp. 19-25, January 1971
- [9] Csendes, Z.J., Silvester, P.,
"Numerical solution of dielectric loaded waveguides: I-Finite-element analysis",
IEEE Trans. Microwave Theory Tech., MTT-18, no. 12, pp. 1124-1131, December 1970
- [10] Koshiba, M., Hayata, K., Suzuki, M.,
"Finite-element formulation in terms of the electric-field vector for electromagnetic waveguide problems",
IEEE Trans. Microwave Theory Tech., MTT-33, no. 10, pp. 900-905, October 1985

- [11] Yeh, C., Dong S.B., Oliver, W.,
"Arbitrarily shaped inhomogeneous optical fiber or integrated opt. waveguides",
J. Appl. Physics, Vol. 46, No. 5, May 1975
- [12] Ikeuchi, M., Sawami, H., Niki, H.,
"Analysis of open-type dielectric waveguides by the finite-element iterative method",
IEEE Trans. Microwave Theory Tech., MTT-29, no. 3, pp. 234-239, March 1981
- [13] Welt, D., Webb, J.,
"Finite element analysis of dielectric waveguides with curved boundaries",
IEEE Trans. Microwave Theory Tech., MTT-33, no. 7, pp. 576-585, July 1985
- [14] Silvester, P., Lowther, D.A., Carpenter, C.J., Wyatt, E.A.,
"Exterior finite elements for 2D field problems with open boundaries",
Proc. IEE, vol. 124, no. 12, pp.1267-1270, December 1977
- [15] Chiang, K.S.,
"Finite-element analysis of optical fibres with iterative treatment of the infinite 2-D space",
Optical and Quantum Electronics, vol. 17, no. 6, pp. 381-391, 1985
- [16] Chiang, K.S.,
"Finite element method for cutoff frequencies of weakly guiding fibres of arbitrary cross-section",
Opt. Quantum Electron., vol. 16, no. 6, pp. 487-493, November 1984
- [17] Wu, R.B., Chen C.H.,
"A variational analysis of dielectric waveguides by the conformal mapping technique",
IEEE Trans. Microwave Theory Tech., MTT-33, no. 8, August 1985
- [18] Zienkiewicz, O.C., Emson, C., Bettess, P.,
"A novel boundary infinite element",
Int. J. Num. Met. Engng., vol. 19, pp. 393-404, 1983
- [19] Bettess, P.,
"Infinite elements",
Int. J. Num. Met. Engng., vol. 11, pp. 53-64, 1977
- [20] Bettess, P.,
"More on infinite elements",
Int. J. Num. Met. Engng., vol. 15, pp. 1613-1626, 1980
- [21] Medina, F., Taylor, R.,
"Finite element techniques for problems of unbounded domains",
Int. J. Num. Met. Engng., vol. 19, pp. 1209-1226, 1983

- [22] Yeh, C., Ha, K., Dong, S.B., Brown, W.P.,
"Single-mode optical waveguides",
Applied Optics, vol. 18, no. 10, pp. 1490-1504 May 1979
- [23] Rahman, B.M.A., Davies, J.B.,
"Finite element analysis of optical and microwave problems",
IEEE Trans. Microwave Theory Tech., MTT-32, no. 1, pp. 20-28, January 1984
- [24] Schweig, E., Bridges, W.B.,
"Computer analysis of dielectric waveguides: A finite-difference method",
IEEE Trans. Microwave Theory Tech., MTT-32, no. 5, pp. 531-541, May 1984
- [25] Ogusu, K.,
"Numerical analysis of the rectangular dielectric waveguide and its modifications",
IEEE Trans. Microwave Theory Tech., MTT-25, pp. 874-885, Nov. 1977
- [26] Shinonaga, H., Kurazono, S.,
"Y dielectric waveguide for millimeter- and submillimeter-wave",
IEEE Trans. Microwave Theory Tech., MTT-29, pp. 542-546, June 1981
- [27] Schelkunoff, S.A.,
"Generalized telegraphist's equations for waveguides",
Bell Syst. Tech. J., vol. 31, pp. 784-801, July 1952
- [28] Goell, J.E.,
"A circular-harmonic computer analysis of rectangular dielectric waveguides",
Bell Syst. Tech. J., vol. 48, pp. 2133-2160, Sept. 1969
- [29] Jamés, J.R., Gallett, I.N.L.,
"Point-matched solutions for propagating modes on arbitrarily-shaped dielectric rods",
Radio Electron. Eng., vol. 42, no. 3, pp. 103-113, March 1972
- [30] Yamashita, E., Atsuki, K., Hashimoto, O., Kamijo, K.,
"Modal analysis of homogeneous optical fibers with deformed boundaries",
IEEE Trans. Microwave Theory Tech., MTT-27, no. 4, pp. 352-356, April 1979
- [31] Su, C.C.,
"Cutoff frequency of a homogeneous optical fiber with arbitrary cross section",
IEEE Trans. Microwave Theory Tech., MTT-33, no. 11, pp. 1101-1105, November 1985
- [32] Saad, S.M.,
"On the higher order modes of elliptical optical fibers",
IEEE Trans. Microwave Theory Tech., MTT-33, no. 11, pp. 1110-1113, November 1985

- [33] James, J.R., Gallett, I.N.L.,
"Modal analysis of triangular-cored glass-fibre waveguide",
Proc. Inst. Elec. Eng., vol. 120, no.11, pp.1362-1370, Dec. 1973
- [34] Yamashita, E., Atsuki, K., Kuzuya, R.,
"Composite dielectric waveguides",
IEEE Trans. Microwave Theory Tech., MTT-28, no. 9, pp. 986-990, Sep. 1980
- [35] Yamashita, E., Atsuki, K., Kuzuya, R.,
"Composite dielectric waveguides with two elliptic-cylinder boundaries",
IEEE Trans. Microwave Theory Tech., MTT-29, no. 9, pp. 987-990, Sep. 1981
- [36] Bates, R.H.T, James, J.R., Gallett, I.N.L., Millar, R.F.,
"An overview of point matching",
Radio Electron. Eng., vol. 43, no. 3, pp. 193-200, March 1973
- [37] de Ruiter, H.M.,
"Integral-equation approach to the computation of modes in an optical waveguide",
J. Opt. Soc. Am., vol. 70, no. 12, pp. 1519-1524, December 1980
- [38] Su, C.C.,
"A surface integral equations method for homogeneous optical fibers and coupled image lines of arbitrary cross section",
IEEE Trans. Microwave Theory Tech., MTT-33, no. 11, pp. 1115-1119, November 1985
- [39] Oyamada, K., Okoshi, T.,
"Two-dimensional finite-element method of propagation characteristics of axially non-symmetrical optical fibers",
Radio Science, vol. 17, no. 1, pp. 109-116, January-February 1982
- [40] Wu, R., Chen, C.,
"On the variational reaction theory for dielectric waveguides",
IEEE Trans. Microwave Theory Tech., MTT-33, no. 6, pp. 477-483, June 1985
- [41] Rumsey, V.H.,
"Reaction concept in electromagnetic theory",
Physical Review, vol. 94, no. 6, pp. 1483-1491, June 1954
- [42] Jeng, S.K., Chen, C.H.,
"On variational electromagnetics: Theory and application",
IEEE Trans. Antennas Propagat., vol. AP-32, no. 9, pp. 902-907, Sep. 1984

- [43] Su, C.C.,
 "A combined method for dielectric waveguides using the finite-element technique and the surface integral equations method",
IEEE Trans. Microwave Theory Tech., MTT-34, no. 11, pp. 1440-1446, Nov. 1986
- [44] Collin, R.E.,
Foundations for microwave engineering,
 New York: McGraw-Hill Book Company, 1966
- [45] Landau, L.D., Lipshitz, E.M.,
 "Electrodynamics of Continuous Media",
Oxford: Pergamon Press Ltd., 1960
- [46] Strang, S.,
Linear Algebra and its Applications, 2nd ed.,
 Orlando, Florida: Academic Press, Inc., 1980
- [47]^a Ramo, S., Whinnery, J.R., Van Duzer, T.,
Fields and waves in communications electronics,
 New York: John Wiley & Sons, 1984
- [48] Marcuse, D.,
Light transmission optics,
 New York: Van Nostrand Reinhold, 1982
- [49] Rahman, B.M.A., Davies, J.B.,
 "Penalty function improvement of waveguide solution by finite elements",
IEEE Trans. Microwave Theory Tech., MTT-32, no. 8, pp. 922-928, August 1984
- [50] Koshiba, M., Kazuya, H., Suzuki, M.,
 "Improved finite-element formulation in term of the magnetic field vector for dielectric waveguides",
IEEE Trans. Microwave Theory Tech., MTT-33, no. 3, pp. 227-233, March 1985
- [51] Gelfand, I.M., Fomin, S.V.,
Calculus of Variations,
 Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963
- [52] Berk, A.D.,
 "Variational principles for electromagnetic resonators and waveguides",
IRE Trans. on Antennas and Propagation, vol. AP-4, pp. 104-111, April 1956
- [53] Konrad, A.,
 "Vector variational formulation of electromagnetic fields in anisotropic media",
IEEE Trans. Microwave Theory Tech., MTT-24, pp. 553-559, September 1976

- [54] Davies, J.B., Fernandez, F.A., Philippou, G.Y.,
 "Finite element analysis of all modes in cavities with circular symmetry",
IEEE Trans. Microwave Theory Tech., MTT-30, no. 11, pp. 1975-1980, November 1982
- [55] Webb, J.P.,
 "The finite element method for finding modes of dielectric-loaded cavities",
IEEE Trans. Microwave Theory Tech., MTT-33, no. 7, pp. 635-639, July 1985
- [56] Mabaya, N., Lagasse, P.E., Vandenbulcke, P.,
 "Finite element analysis of optical waveguides",
IEEE Trans. Microwave Theory Tech., MTT-29, no. 6, pp. 600-605, June 1981
- [57] Hayata, K., Koshiha, M., Suzuki, M.,
 "Finite-element solution of anisotropic waveguides with arbitrary tensor permittivity",
Journal of Lightwave Technology, vol. LT-4, no. 2, pp. 121-125, February 1986
- [58] Dettman, John W.,
Mathematical Methods in Physics and Engineering,
 New York: McGraw-Hill Book Company, 1969
- [59] Silvester, P.,
 "High order polynomial triangular finite elements for potential problems",
Int. J. Engng. Sci., vol. 7 pp. 849-861, 1969
- [60] Silvester, P.,
 "Construction of triangular finite element universal matrices",
Int. J. Num. Met. Engng., vol. 12, pp. 237-244, 1978
- [61] Beer, G., Meek, J.L.,
 "Infinite domain elements",
Int. J. Num. Met. Engng., vol. 17, pp. 43-52, 1981
- [62] Zienkiewicz, O.C., Bando, K., Bettess, P., Emson, C., Chiam, T.C.,
 "Mapped infinite elements for exterior wave problems",
Int. J. Num. Met. Engng., vol. 21, 1229-1251, 1985
- [63] McDonald, B.H., Wexler, A.,
 "Finite element solution of unbounded field problems",
IEEE Trans. Microwave Theory Tech., MTT-20, no. 12, pp. 841-847, December 1972
- [64] Washisu, S., Fukai, I. Michio, S.,
 "Extension of finite-element method to unbounded field problems",
Electronic Letters, vol. 15, pp 772-774, 22 Nov., 1979

- [65] Hayata, K, Eguchi, M., Koshiha, M.,
"Self-consistent finite/infinite element scheme for unbounded guided wave problems",
IEEE Trans. Microwave Theory Tech., MTT-36, no. 3, pp. 614-616, Sep. 1988
- [66] Gradshteyn, I.S., Ryzhik, I.M.,
Tables of Integrals, Series, and Products,
corrected and enlarged ed., New York: Academic Press, 1980
- [67] Jackson, J.D.,
Classical electrodynamics,
New York: John Wiley & Sons, 1975
- [68] Adams, M.J.,
An Introduction to Optical Waveguides,
New York: John Wiley & Sons, 1981