A SECOND ORDER ISOPARAMETRIC FINITE ELEMENT ANALYSIS OF DIELECTRIC WAVEGUIDES WITH CURVED BOUNDARIES

Daniel WELT
Department of Electrical Engineering
McGill University, Montreal
APRIL , 1984

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of M.ENG.

Je tiens à remercier les personnes suivantes pour leur assistance :

John Webb, pour le support technique au niveau de la formulation mathématique et de la programmation.

Suzanne Lacroix et Jacques Bures pour leurs contributions à la compréhension du coupleur optique monomode.

A. Raab, pour l'équipement mis à ma disposition.

Viviane Leaune pour sa collaboration.

ABSTRACT

An analysis of the arbitrarily shaped and infinitely long dielectric waveguide is carried out using the Finite Element Method (FEM) in 2-dimensions. Emphasis is put on structures having curvilinear boundaries such as those associated with optical fibers.
The efficiency of the numerical process is greatly improved by using quadratic and isoparametric elements for the discretisation.

The analysis of known models shows the accuracy and the limits of the method. Propagation curves and the equipotential contour plots of all field components are available through a series of programs.

The special case of the monomode fiber-to-fiber tapered optical coupler is also considered, requiring an extremely high accuracy. This example introduces some new ideas concerning the concept of mode degeneracy and raises new questions about the power distribution between non degenerate modes of different polarizations in non-symmetric structures.

Abstract

Une analyse détaillée du guide d'onde diélectrique de forme arbitraire et infiniment long est présentée en utilisant la méthode dite des Eléments Finis dans un espace bi-dimensionnel.
On s'est particulièrement intéressé aux structures de forme curviligne comme celles que l'on retrouve dans les composantes construites avec des fibres optiques. Pour améliorer de façon considérable le processus numérique, des éléments quadratiques et isoparamétriques sont utilisés pour la discrétisation.

Des exemples connus faisant appel à des modèles de fibres existants mettent en évidence la précision des résultats et les limites de la méthode. Les courbes représentant la constante de propagation des premiers modes en fonction de la fréquence, ainsi que les tracés des contours équipotentiels de toutes les composantes du champ électromagnétique sont disponibles par l'intermédiaire d'une série de programmes.

On s'intéresse en particulier au cas du coupleur optique monomode fibre-fibre, dont l'analyse met à rude épreuve le processus numérique et exige une grande précision. Cet exemple apporte de nouvelles idées en ce qui concerne la notion de dégénérescence des modes de propagation et la question de la répartition énergétique entre les différentes polarisations dans des structures non symétriques.

TABLE OF CONTENTS

1. INTRODUCTION

2. THE OPTICAL GUIDING STRUCTURE
   - 2.1 Introduction
   - 2.2 Theory
   - 2.3 Summary of some Analytical Methods of Solutions

3. THE FINITE ELEMENT METHOD
   - 3.1 Detailed Formulation
     - 3.1.1 Literature Review
     - 3.1.2 Mathematical Formulation
     - 3.1.3 Continuity Condition and Boundary Conditions
   - 3.2 Interpolation Functions and Discretisation
   - 3.3 Mapping and Change of Coordinates
   - 3.4 Evaluation of the Discretised Integral Equation
     - 3.4.1 Scale change
     - 3.4.2 Mapping
     - 3.4.3 Numerical integration
   - 3.5 Computation of the Tranverse Field Components
   - 3.6 The Global Matrix Formulation
     - 3.6.1 The local matrix formulation
     - 3.6.2 The local to global matrix transformation
     - 3.6.3 Effect of the boundary conditions
     - 3.6.4 Final answer
   - 3.7 The Virtual Boundary
   - 3.8 Symmetry and Boundary Conditions

4. ANALYSIS OF KNOWN EXAMPLES WITH THE FEM PROGRAMS
   - 4.1 The Dielectric Rod Analysis
     - 4.1.1 Analytical solution
     - 4.1.2 FEM solution
     - 4.1.3 Higher order modes
   - 4.2 The Inhomogeneous waveguide
     - 4.2.1 Runge-Kutta differential analysis
     - 4.2.2 FEM solution
   - 4.3 The Elliptical Fiber
     - 4.3.1 Description
     - 4.3.2 The $\Delta\gamma$ characteristics
     - 4.3.3 Error Analysis

4.4 The Finite Cladding Fiber and the Cutoff Problem
4.4.1 Description
4.4.2 Example
4.5 Spurious Mode Detection and Analysis

5. ANALYSIS OF A BICONICAL-TAPERED-MONOMODE OPTICAL COUPLER
5.1 Description
5.2 Mode Superposition and coupling effect in 2-D
5.2.1 Introduction
5.2.2 Theoretical model

5.3 Mode Superposition and coupling effect in 3-D
5.3.1 A Three Dimensional Approximate Analysis
5.3.2 Simplifying assumptions
5.3.3 Coupling Mechanism
5.3.4 Coupler Geometry
5.3.5 Computation of the Coupling Coefficient for 2 modes interaction
5.3.6 Computation of the Coupling Coefficient for 4 modes interaction

5.4 Examples
5.4.1 Example 1
5.4.2 Example 2
5.4.3 Comparison with experimental data
5.4.4 Accuracy of numerical results
5.4.5 Equivalence with Snyder coupled-mode theory

6. PROGRAMS
6.1 Structure
6.2 The Pre-processors
6.2.1 Input program "DWIN"
6.2.2 The automatic mesh generator "ELF"
6.3 The Solvers
6.3.1 The Global matrices generator "DW1P"
6.3.2 The eigenvalue solvers "DW2PX", "DW2PXX"
6.4 The Post Processors
6.4.1 Computation of the transverse field components with "DW5P" program
6.4.2 The graphic program "DW6P"
6.4.2.1 Theory
6.4.2.2 Implementation

6.4.3 The coupler analyser "DW8P(2)"
6.4.3.1 The $\Delta\gamma$ characteristics and Error Analysis
6.4.3.2 Computation of the coupling coefficients

7. CONCLUSION

Appendix 1 : The continuity conditions on the radial field components.

Appendix 2 : Details of the discretisation process.

Appendix 3 : Analytical solution of the dielectric rod problem.

Appendix 4 : A simple method to find the extremum inside the reference element.

Appendix 5 : A variational formulation which removes the discontinuity at τ_i.

LIST OF FIGURES

1. The geometric transformation T_i.
2. Determination of the virtual boundary with the equivalent model of the dielectric rod.
3. Variation of the fundamental mode frequency at $\gamma = 1.45$ with respect to the location of the virtual boundary R_{max}.
4. Propagation characteristics for the first modes of a dielectric rod.
5. Dielectric rod example, discretisation with 62 elements and 159 nodes.
6. Fundamental mode of a dielectric rod : field plots.(HE_{11}) Propagation constant = 1.450, boundary conditions code = 2.
7. Dielectric rod, Hx field, linear scale, $\gamma = 1.45$, fundamental mode (HE_{11}). (Ey and Sz are similar).
8. Dielectric rod, Ex field, linear scale, $\gamma = 1.45$, fundamental mode (HE_{11}). (Hy is similar).
9. Dielectric rod, Hz field, higher order mode, $\gamma = 1.45$, linear scale, boundary conditions code = 2. (HE_{12})
10. Dielectric rod , Sz component, linear scale (x2), higher order mode, $\gamma = 1.45$, boundary conditions code = 2. (HE_{12})
11. Discretisation of a circular waveguide for the computation of higher order modes propagation characteristics (typical examples).
12. The inhomogeneous optical fiber. Step index model of the parabolic profile.
13. Elliptical waveguide discretisation. 46 elements, 119 nodes.
14. Elliptical waveguide example. Propagation characteristics of the two fundamental modes.
15. Elliptical waveguide example. Difference between the propagation characteristics of the first two modes (KBC 2 & 3).
16. Finite cladding fiber, discretisation.
17. Finite cladding fiber, Hz field, linear scale, $\gamma = 1.450$. (HE_{11})
18. Finite cladding optical fiber. Propagation characteristics of the fundamental mode near the cladding cutoff.
19. Dielectric rod example, first mode (spurious), linear scale, $\gamma = 1.450$.
20. Finite cladding fiber, high order mode, linear scale, $\gamma = 1.51$, boundary conditions code 2. (HE_{31})
21. Symmetry of modes and superposition.
22. Geometry of the coupler.

23. Coupler discretisation.
24. Coupler Example 1 .
    Propagation characteristics of the first 4 modes, computer interpolation.
25. A & B : same as 24, but with a different scale.
26. Coupler example 1.
    Propagation characteristics of the first 4 modes.
27. Coupled power versus coupler length L (example 1).
28. Coupler, Ez field, linear scale, boundary conditions code = 2, fundamental mode.
29. same as 28, but with different values of γ .
30. Coupler, Hz field, linear scale, boundary conditions code = 2. fundamental mode.
31. Coupler, Ez field, log. scale, γ = 1.490, boundary conditions code = 2, fundamental mode.
32. Coupler, Ez field, linear scale, γ = 1.463, boundary conditions code = 4, fundamental mode.
33. Coupler, Ez field, linear scale, boundary conditions code = 4. fundamental mode.
34. Coupler example 2.
    Propagation characteristics of the first 4 modes.
35. Flow chart (1)
36. " " (2)
37. Other programs.

LIST OF TABLES

1. The interpolation functions for isoparametric and quadratic elements.
2. Local $[A_i]$ matrix.
3. Local $[B_i]$ matrix.
4. The local to global matrix transformation.
5. Effect of the virtual boundary.
6. Definition of the "KBC boundary conditions code".
7. The dielectric rod example with mode number (n) = 0.
8. The dielectric rod example with mode number (n) = 1.
9. The dielectric rod example with mode number (n) = 2.
10. Numerical results for the inhomogeneous optical fiber.
11. The elliptical fiber propagation characteristics for the two quasi-degenerate dominant modes.
12. $\Delta\gamma$ characteristics of the elliptical fiber.
13. Near cutoff propagation characteristics of a finite-cladding fiber.
14. Example 1. FEM solution for the first modes.
15. Example 2. FEM solution for the first modes.
16. Examples 1 and 2.
    Difference between the normalised frequency versus γ .
17. List of programs.
18. Dimension of the discretised problem vs the radial and angular division.

1. INTRODUCTION

The analysis of electromagnetic wave propagation in optical guiding structures may be a very complicated task when the geometry involved is not simple (e.g. a circle, an ellipse or a rectangle).
When such a situation may not be avoided, an answer is the use of the Finite Element Method, widely used in many different fields : structural analysis, fluid mechanics, heat transfer, magnetostatics, etc.
The present work was undertaken primarly for the analysis of an optical fiber-to-fiber monomode coupler. However, the programs written for this purpose may be easily used for the analysis of other waveguiding structures at optical or microwave frequencies.
They have been designed for problems having curvilinear boundaries such as those encountered in devices built with optical fibers. A very direct approach has been chosen to obtain the accuracy and generality at the price of complexity, large computer programs and execution time.

Chapter 2 gives a complete theoretical analysis of the problem in terms of variational calculus, which is not the standard form of Maxwell's equations, but the necessary step for the application of the Finite Element Method. Careful attention has been paid to the exactness of the mathematics. Then, chapter 3 describes the relatively standard discretisation procedure of the 2-D area of interest and writes the equations derived in chapter 2 as a large eigenvalue problem. These first two chapters contain most of the theoretical background. Chapter 4 contains the results from the numerical analysis of four different problems with known solutions, which show the accuracy and the limits of the methods.
A separate section, Chapter 5, has been reserved for the analysis of the fiber-to-fiber single mode tapered optical coupler, which is one of the goals of this work.
Finally, the last chapter summarizes the software developed or

adapted for this thesis. Its detailed description would have taken too much space to be included here; but the flow charts included there indicate the process followed from the initial data to the final output.

2. THE OPTICAL GUIDING STRUCTURE

2.1 Introduction

An optical waveguide is a structure in which light travels with very low loss. It is usually made with two or more materials having a slightly different index of refraction. An analysis done with a simple geometry such as the dielectric rod shows a theoretically lossless propagation of electromagnetic fields in different "modes". These are the various solutions of the Maxwell's equations [see appendix 3]. The applications of optical waveguides may be arbitrarily divided in three parts :

- Image transmission : in this case, the fiber transmits incoming light from one location to another without any sort of signal processing. Hundreds of fibers may be put in bundles to carry a complete picture.
- Telecommunications : because of the very high bit rate capacity of optical links, their major application in the near future will be as point-to-point information carriers. On a long term basis, they will probably replace to a certain extent the usual copper telephone wires in each house. They will carry most of the traffic on high density long distance routes (like intercontinental links), the satellite becoming a more specialised tool for wireless needs such as mobile telecommunications and teledetection.
- Sensors : their very small size and high noise immunity, make them ideal for acoustic sensing [1], temperature sensing [2], etc.

2.2 Theory

As we have mentioned previously, an optical waveguide is made of several materials each having a different index of refraction.
The laws needed for the analysis are the classical Maxwell's equations. Because of their guiding purpose, the structures involved are usually very long compared to their other dimensions. The wavelength of the carrier is usually between 0.5 micron and 1.5 micron. This allows a first simplification, by assuming a quasi two-dimensional problem in which the waveguide is infinitely long.
The second important simplification comes from the type of solution we expect to get, which is a wave-like solution, with a given angular frequency (the operating frequency of the laser or the LED exciting the guide) and with a purely sinusoidal behavior along the direction of propagation.
Finally, we assume that the propagating media are isotropic, lossless and may be divided into homogeneous areas in which the Maxwell's equations in homogeneous media are verified. This last simplification is necessary for the type of formulation we use. Each element of the discretised area will be homogeneous.
With these assumptions, any vector field $\vec{F}$ of the electromagnetic field may be written :

$$\vec{F}(x,y,z,t) = \mathrm{Re}\{\vec{f}(x,y)\, e^{j(\omega t - \beta z)}\} \qquad (2.1)$$

Where : x,y and z are Cartesian space coordinates,
z is the direction of propagation,
β is the constant of propagation (unknown),
ω is the angular frequency of the wave,
t is time,
j is $\sqrt{-1}$,
$\vec{f}$ is a complex vector.

These are the usual assumptions in all waveguide problems. They lead to the simplified Maxwell's equations :

$$\nabla \times \vec{E} = - \frac{\partial \vec{B}}{\partial t} = - j \omega \mu \vec{H} \qquad (2.2)$$

$$\nabla \times \vec{H} = + \frac{\partial \vec{D}}{\partial t} = + j \omega \epsilon \vec{E} \qquad (2.3)$$

$$\nabla . \vec{B} = 0 \qquad (2.4)$$

$$\vec{B} = \mu \vec{H} \qquad (2.5)$$

$$\nabla . \vec{D} = 0 \qquad (2.6)$$

$$\vec{D} = \epsilon \vec{E} \qquad (2.7)$$

The usual mathematical manipulations [3] give the well known Helmholtz equations :

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k^2 \right) . \begin{Bmatrix} Ez \\ Hz \end{Bmatrix} = 0 \qquad (2.8)$$

Where : E_z is the z-component of the $\vec{E}$ vector
H_z is the z-component of the $\vec{H}$ vector

$$k^2 = \left(\frac{\omega}{c} \right)^2 . \left(\frac{\epsilon}{\epsilon_0} - \gamma^2 \right) \qquad (2.9)$$

$$\gamma = \frac{\beta c}{\omega} \qquad (2.10)$$

$$c = \text{speed of light in vacuum} = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \qquad (2.11)$$

β and ω have been previously defined.
ϵ is the permittivity of the medium.
ϵ_0 is the permittivity in vacuum.
μ_0 is the permeability in vacuum.

Note : The Helmholtz equations cannot be written in inhomogeneous media, where a term involving $\nabla\epsilon$ (grad ϵ) cannot be eliminated.

The knowledge of the z-components of $\vec{E}$ and $\vec{H}$ vectors is enough to solve completely any problem of that kind because the other components can be derived from them.

We have :

$$Ex = j \frac{\omega\mu_0}{k^2} \left(\frac{\partial Hz}{\partial y} + \gamma \frac{1}{\eta_0} \frac{\partial Ez}{\partial x} \right) \qquad (2.12)$$

$$Ey = j \frac{\omega\mu_0}{k^2} \left(- \frac{\partial Hz}{\partial x} + \gamma \frac{1}{\eta_0} \frac{\partial Ez}{\partial y} \right) \qquad (2.13)$$

$$Hx = j \frac{\omega\varepsilon_0}{k^2} \left(- \frac{\varepsilon}{\varepsilon_0} \frac{\partial Ez}{\partial y} + \gamma \eta_0 \frac{\partial Hz}{\partial x} \right) \qquad (2.14)$$

$$Hy = j \frac{\omega\varepsilon_0}{k^2} \left(\frac{\varepsilon}{\varepsilon_0} \frac{\partial Ez}{\partial x} + \gamma \eta_0 \frac{\partial Hz}{\partial y} \right) \qquad (2.15)$$

$$\eta_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} \qquad (2.16)$$

The transverse components above are in quadrature with the z-components. Therefore, the Poynting vector has only one real component (S_z), which indicates a real power flow in the z-direction only and no propagation in the transverse directions.

Finally, note that the two Helmholtz equations (2.8) are coupled by the continuity of tangential field at the interface between any region, and cannot be written and solved separately when an exact solution is required. Details are given in section 3.1 and in the references [10,11,12].

2.3 Summary of some analytical methods of Solutions

When the geometry is simple, the Helmholtz equations (2.8) are solvable directly using the separation of variables technique. For example, the exact analysis of propagation inside a dielectric rod is given in appendix 3. But when the shape of the guide becomes more complicated and/or when the medium is not homogeneous, it is necessary to find other methods.

If the problem has one axis of symmetry, it reduces to a single dimension and may be solved with differentiál techniques such as the Runge Kutta method [4], or the Predictor-Corrector method [5].In these approaches, the radial variations of the field components are computed step by step from the center and the tangential solutions are matched at every boundary. Complete descriptions of these methods are given in the references.

3. THE FINITE ELEMENT METHOD

The Finite Element Method (FEM) is very well documented. The steps missing in the formulation presented here may be easily found in one of these references : [6,7,8].

3.1 Detailed formulation

3.1.1 Literature Review

The Finite Element method in electromagnetics has mainly two application areas : (i) magnetostatic problems for machines and high power devices ; (ii) wave propagation problems.
The first category does not fall directly in our field of interest, but we can mention the use of curvilinear elements in reference [9], because these are not common in electrical engineering field analysis.
In the propagation category, the first finite element analyses were of waveguides, using 2-dimensional triangular elements well adapted for the usual rectangular components and geometries. The basic formulation for inhomogeneous waveguides may be found in reference [10], by S.Ahmed and P.Daly, who used first order interpolation functions in 1969. One year later, high order polynomials were implemented by Silvester and Csendes [11] to solve essentially the same problem.

The application of FEM to optical fiber or integrated optics problems may be found in reference [12]. In that analysis, C.Yeh, S.B. Dong, and W. Oliver used triangular elements and linear interpolation to solve various dielectric waveguide problems. The discretisation of circular structures (e.g. a quarter section of a fiber) requires a large number of elements. In order to keep the non-zero field area finite for computational feasibility, they use a virtual boundary on which true zero field values were enforced far enough from the high permittivity guiding region.
Later, on 1979 the same authors implement infinite elements [13]

with special decaying interpolation functions to model the truly infinite domain of integration required by the method. They replace the problem of locating the virtual boundary by the search for an appropriate decay factor.

All the papers previously mentioned were based on the longitudinal field components formulation for isotropic materials.
This constraint was removed by Mabaya, Lagasse and Vendenbulcke who extend the formulation to allow anisotropic materials [14]. They use triangular elements with interpolation polynomials of order 1 to 4.

In order to obtain directly the transverse field components which are useful for computing the transmitted power flow, J. Katz [15] proposed a different variational approach in terms of E_x and E_y. In the same paper, he proposed a model for the transition between layers of different permittivity by assuming a linear change over a short distance instead of a step transition.
Mention the recent paper by Rahman and Davies [16], who used a H-vector field formulation to solve anisotropic problems with triangular elements, and finally, Hano's solution [26] which suppresses spurious modes with rectangular elements (not very useful to model optical fibres).

3.1.2 Mathematical Formulation

In our analysis, the variational expression given in references [10, 11, 12 and 13] is chosen for the following reasons :
- The waveguides to be analysed are assumed isotropic.
- The formulation is well established.
The use of curvilinear elements is the original feature of that work.

If we split the x-y plane containing a cross section of a given dielectric waveguide in bounded areas called elements, and if we define the following quantities :

S_i = area of element "i" (3.1)

H_z^i = z-component of $\vec{H}$ in element "i" (3.2)

E_z^i = z-component of $\vec{E}$ in element "i" (3.3)

ε_i = Permittivity in element "i" (3.4)

$$\tau_i = \frac{\gamma^2 - 1}{\gamma^2 - (\varepsilon_i/\varepsilon_0)} \tag{3.5}$$

N = Number of elements ($1 \leq i \leq N$)

$$k_0^2 = (\frac{\omega}{c})^2 (\gamma^2 - 1) \tag{3.6}$$

The solution to the following variational equation :

$$\delta \sum_{i=1}^{N} \{ \int_{S_i} \tau_i |\nabla\phi_i|^2 + \gamma^2\tau_i (\frac{\varepsilon_i}{\varepsilon_0}) |\nabla\psi_i|^2 + 2\gamma^2\tau_i \vec{e}_z.(\nabla\psi_i \times \nabla\phi_i) + k_0^2|\phi_i|^2 + k_0^2\gamma^2(\frac{\varepsilon_i}{\varepsilon_0}) |\psi_i|^2 dS_i \} = 0 \tag{3.7}$$

Where :

$$\phi_i = H_z^i \tag{3.8}$$

$$\psi_i = \frac{1}{\gamma\eta_0} E_z^i \tag{3.9}$$

$\vec{e}_z$ = unit vector in the direction of propagation (z)

is the solution to the Maxwell's equations with the proper boundary conditions when the sum of all the elements fills the entire x-y plane.

Proof : It is now important to prove that the integral formulation given above leads to the Maxwell's equations together with the usual boundary conditions. From equation (3.7), we can write the first variation of the functional with respect to ψ_i or with respect to ϕ_i .

In the first case we obtain :

$$\sum_{i=1}^{N} \{ \int_{Si} 2\gamma^2\tau_i \left(\frac{\varepsilon_i}{\varepsilon_0}\right) \nabla\psi_i . \nabla\delta\psi_i + 2\gamma^2\tau_i \vec{e}_z . (\nabla\delta\psi_i \times \nabla\phi_i) + 2 k_0^2 \gamma^2 \left(\frac{\varepsilon_i}{\varepsilon_0}\right) \psi_i \delta\psi_i \, dS_i \} = 0 \quad (3.10)$$

From vectorial calculus we have [17]

$$\nabla . (U\vec{A}) = (\nabla U) . \vec{A} + U (\nabla . \vec{A}) \quad (3.11)$$

Where U is a scalar function, $\vec{A}$ a vector function.

Then :

$$\nabla\psi . \nabla\phi = \nabla . (\psi\nabla\phi) - \psi (\nabla . \nabla\phi) \quad (3.12)$$

And we can write :

$$\nabla\psi_i . \nabla\delta\psi_i = \nabla . (\delta\psi_i \nabla\psi_i) - \delta\psi_i \nabla_t^2\psi_i \quad (3.13)$$

From vectorial calculus [17] we also have :

$$\nabla\phi \times \nabla\psi = \nabla \times (\phi\nabla\psi) \quad \text{because} \quad \nabla\times\nabla\psi = 0 \text{ by theorem.} \quad (3.14)$$

The equation (3.10) becomes :

$$\sum_{i=1}^{N} \int_{S_i} \gamma^2\tau_i \left(\frac{\varepsilon_i}{\varepsilon_0}\right) \left(-\nabla_t^2\psi_i + \frac{k_0^2}{\tau_i} \psi_i \right) \delta\psi_i \, dS_i + \int_{S_i} \gamma^2\tau_i \{ \left(\frac{\varepsilon_i}{\varepsilon_0}\right) \nabla . (\delta\psi_i \nabla\psi_i) + \vec{e}_z . (\nabla\times\delta\psi_i \nabla\phi_i) \} \, dS_i = 0 \quad (3.15)$$

The second part of the integral can be transformed in a line integral using the Divergence theorem and the Stokes theorem, both in 2 dimensions. We obtain :

$$\sum_{i=1}^{N} \int_{S_i} \gamma^2\tau_i \left(\frac{\varepsilon_i}{\varepsilon_0}\right) \left(-\nabla_t^2\psi_i + \frac{k_0^2}{\tau_i} \psi_i \right) \delta\psi_i \, dS_i + \int_{C_i} \gamma^2\tau_i \{ \left(\frac{\varepsilon_i}{\varepsilon_0}\right) \nabla\psi_i . \vec{n} + \nabla\phi_i . \vec{l}_i \} \delta\psi_i \, dl_i \quad (3.16)$$

Where $\vec{n}$ is the unit vector normal to the contour C_i, and $\vec{l}_i$ satisfies the relation : $\vec{n} \times \vec{l}_i = \vec{e}_z$ ($\vec{n}$ = outward normal).

Therefore, $\vec{l}_i = \vec{e}_z \times \vec{n}$ and the line integral becomes : (3.17)

$$\int_{C_i} \gamma^2 \tau_i \{ (\frac{\varepsilon_i}{\varepsilon_0}) (\nabla\psi_i . \vec{n}) - \vec{e}_z . (\nabla\phi_i \times \vec{n}) \} \delta\psi_i \, dl_i$$

We can also make the same development for the first variation of equation (3.7) with respect to ϕ_i:

$$\sum_{i=1}^{N} \int_{S_i} \tau_i (-\nabla_t^2 \phi_i + \frac{k_0^2}{\tau_i} \phi_i) \delta\phi_i \, dS_i \qquad (3.18)$$

$$+ \int_{C_i} \tau_i \{ \nabla\phi_i . \vec{n} + \gamma^2 \vec{e}_z . (\nabla\psi_i \times \vec{n}) \} \delta\phi_i \, dl_i = 0$$

The two line integrals can be related to the transverse field components using the equations (2.12) to (2.16) or reference 12.
We have :

$$\vec{\psi}_t^i = A \tau_i (\gamma^2 \nabla_t \psi_i - \vec{e}_z \times \nabla_t \phi_i) \qquad (3.19)$$

$$\vec{\phi}_t^i = \gamma A \tau_i \{ \nabla_t \phi_i + (\frac{\varepsilon_i}{\varepsilon_0}) \vec{e}_z \times \nabla_t \psi_i \} \qquad (3.20)$$

$$A = \frac{j}{\gamma (\frac{\omega a}{c}) (\gamma^2 - 1)} \qquad (3.21)$$

The gradients in equations (3.15) to (3.18) are in two dimensions (the subscript "t" was omitted for simplicity), and we can compare the line integrals with the tangential part of the transverse field components.
We have :

$$\vec{\psi}_t^i \times \vec{n} = A \tau_i \{ \gamma^2 \nabla_t \psi_i \times \vec{n} - (\vec{e}_z \times \nabla_t \phi_i) \times \vec{n} \} \qquad (3.22)$$

$$= A \tau_i \{ \gamma^2 \nabla_t \psi_i \times \vec{n} + \vec{e}_z (\nabla_t \phi_i . \vec{n}) \} \qquad (3.23)$$

$$= A \tau_i \vec{e}_z \{ \gamma^2 (\nabla_t \psi_i \times \vec{n}) . \vec{e}_z + \nabla_t \phi_i . \vec{n} \} \qquad (3.24)$$

The previous transformations are allowed because $\vec{\psi}_t^i$ and $\vec{n}$ are both in the x-y plane. Therefore, the line integral in equation (3.18) can be written :

$$(1/A) \quad \int_{C_i} \vec{e}_z \cdot (\vec{\psi}_t^i \times \vec{n}) \, \delta\phi_i \, dl_i \qquad (3.25)$$

Using the same arguments, we obtain :

$$\vec{\phi}_t^i \times \vec{n} = A\gamma\tau_i \{ (\nabla_t \phi_i \times \vec{n}) - (\frac{\varepsilon_i}{\varepsilon_0}) \vec{e}_z (\nabla_t \psi_i \cdot \vec{n}) \} \qquad (3.26)$$

$$= A\gamma\tau_i \vec{e}_z \{ \vec{e}_z \cdot (\nabla_t \phi_i \times \vec{n}) - (\frac{\varepsilon_i}{\varepsilon_0}) (\nabla_t \psi_i \cdot \vec{n}) \} \qquad (3.27)$$

and (3.17) becomes :

$$-(\gamma/A) \quad \int_{C_i} \vec{e}_z \cdot (\vec{\phi}_t^i \times \vec{n}) \, \delta\psi_i \, dl_i \qquad (3.28)$$

Equations (3.16) and (3.18) must be zero for any arbitrary function $\delta\psi_i$ and $\delta\phi_i$. Therefore, the surface integral and the line integral must vanish and we obtain :

- The Helmholtz equations (2.8),
- The continuity condition or the boundary conditions discussed in the following section.

3.1.3 Continuity condition and Boundary conditions

The vanishing of the line integrals in (3.16) and (3.18) for any variation $\delta\psi_i$ or $\delta\phi_i$ can only be obtained when the tangential part of the transverse field satisfies one of the following boundary conditions :

- If the integration path is an interelement boundary, the tangential field must be continuous, even if the two elements have a different index of refraction. With this condition, any line integral computed along one path is cancelled by the same integral done in the reverse direction. This is one of the required conditions which guarantees the equivalence between Maxwell's equations and the variational formulation (3.7).

- If the problem has natural geometrical axes of symmetry, it is very efficient for the computational point of view to split the domain in smaller parts. Therefore, this operation requires the definition of modes which use the symmetry and are still solutions of the variational equation. If we enforce $\psi = 0$ on the axis of symmetry, we obtain $\nabla\psi \times \vec{n} = 0$, where $\vec{n}$ is the unit vector normal to that axis on the integration contour. In order to keep the line integral zero, the other field ϕ must satisfy $\nabla\phi \cdot \vec{n} = 0$. This condition is called the natural boundary condition (The gradient of the non-zero field z-component on the axis of symmetry is parallel to that axis).

- Similarly, we can define on the same axis of symmetry another mode which satisfies the constraint $\phi = 0$, and the natural boundary condition becomes $\nabla\psi \cdot \vec{n} = 0$.

The last two conditions do not involve the Maxwell's equations directly, but allow considerable reduction in the size of the numerical problem which will be solved later on.

- As it is shown in the references (section 3.1.1), in a dielectric waveguide problem with no metallic boundary, the field extends in the whole plane, but with a fast decaying rate outside the guiding medium. Therefore, it is easy and efficient to create a virtual boundary on which all the fields are zero. This operation satisfies also the variational equation (3.7), but not Maxwell's equations and must be considered as a numerical approximation.

The other conditions on the normal field components are derived in Appendix 1.
Thus we have proved the validity of the variational expression and we can start the following step toward the numerical solution.

3.2 Interpolation functions and Discretisation

The mechanism involved in the Finite Element formulation is simple : the region studied is split into elements small enough to allow an approximation of the unknown field with interpolation polynomials and a set of nodal values (see references [6,7,8]).
Let ϕ be unknown inside the element "i". Let ϕ_{ij} be the unknown nodal values with : $1 < j < L_N(i)$; ($L_N(i)$ is the "local number of nodes in element "i"). Let N_{ij} be a set of interpolation functions having the required characteristics [6]. Then, we can approximate the unknown function ϕ by :

$$\phi = \sum_{j=1}^{LN(i)} N_{ij} \, \phi_{ij} \qquad (3.29)$$

ϕ and the N_{ij}'s are functions of the system of coordinates. If the variation of ϕ inside each element "i" is small, smooth and comparable with the variations of the N_{ij}'s, a very good accuracy may be obtained.
This step leads to the discretisation process in which the search for a continuous function is replaced by the search for discrete values. "j" is a local index, and the double index "ij" (local node j in element i) will be replaced by a global index "n" later on.

3.3 Mapping and Change of Coordinates

It is more practical and efficient to have only one set of interpolation functions N_j for all the elements of the problem. Then, each real element "i" is mapped into a reference element with a change of coordinates. This geometrical transformation, shown in Fig. 1 is denoted T_i.

There is a one-to-one equivalence between any point of the element "i" and a point of the reference element. The boundaries and the nodes are mapped exactly. The choice of the T_i's determines what type of boundaries in the real device will be perfectly matched.

Let X and Y be the coordinates in the real space, and ξ, η, the coordinates in the reference space.

T_i is defined such that :

$$X = \sum_{j=1}^{LN(i)} \bar{N}_j (\xi,\eta) X_{ij} \qquad (3.30)$$

$$Y = \sum_{j=1}^{LN(i)} \bar{N}_j (\xi,\eta) Y_{ij} \qquad (3.31)$$

The $\bar{N}_j$'s are the geometric transformation functions and are not necessarily equal to the interpolation functions N_j's. In fact the above equations (3.30 and 3.31) represent the inverse transformation which maps each point (ξ , η) of the reference element onto a point (X,Y) in the element "i". X_{ij} and Y_{ij} are the coordinates of the node "j" of the element "i".

It is then possible to map the nodal values of the unknown field function ϕ to the nodes of the reference element and a single set of interpolation functions (the N_j's) is needed (if all the elements "i" are of the same kind).

We can write :

$$\phi (\xi,\eta) = \sum_{j=1}^{LN} N_j \phi_{ij} \qquad (3.32)$$

The evaluation of the integral (3.7) will be done through the transformations T_i equivalent to a discretised change of coordinates.
In this work, the transformations T_i correspond to the choice of so-called quadratic-isoparametric elements. The boundaries of all the elements are quadratic, and the interpolation functions N_j are identical to the geometrical transformation functions $\bar{N}_j$. The main advantage of this type of element is the fact that it fits more accurately than straight edges in circular or elliptical geometries encountered in the optical waveguide analysis.
The N_j's are shown below, on table 1 [ref 6] :

j	$N_j = \bar{N}_j$
1	$-(1-\xi-\eta)(1-2(1-\xi-\eta))$
2	$4\xi\ (1-\xi-\eta)$
3	$-\xi\ (1-2\xi)$
4	$4\xi\eta$
5	$-\eta\ (1-2\eta)$
6	$4\eta\ (1-\xi-\eta)$

Each element has six (6) nodes.

TABLE 1
The interpolation functions for isoparametric and quadratic elements

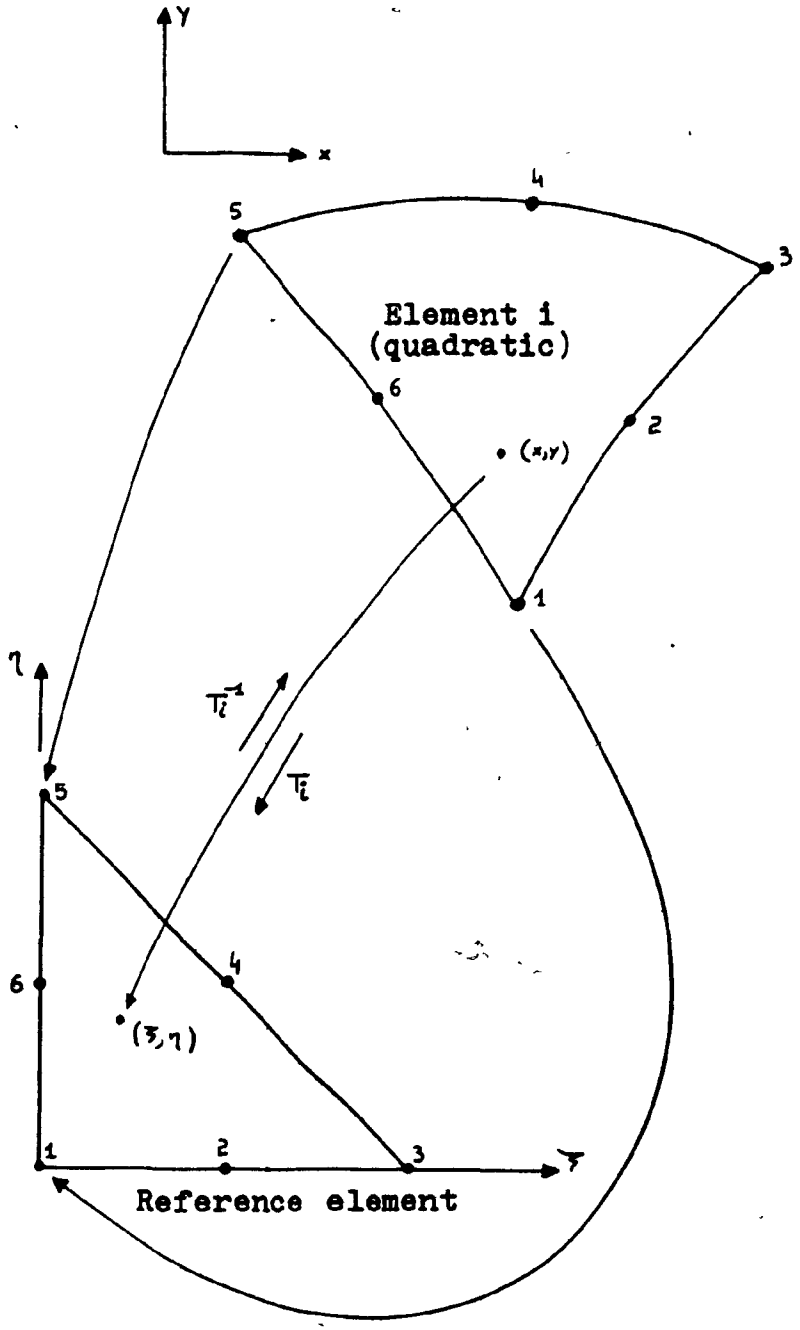


Fig. 1 : The geometric transformation "T_i".

3.4 Evaluation of the Discretised Integral Equation

The variational equation (3.7) may be written :

$$\delta \sum_{i=1}^{N} \int_{S_i} F (\phi_i, \psi_i) \, dS_i = 0 \qquad (3.33)$$

Two changes of coordinates are done :

- A scale change
- The mapping T_i

3.4.1 Scale change

If x and y are the coordinates in the real world, and if we use MKSA system, x and y are in meter(s) and this unit is not practical at optical wavelengths. In addition, it is always necessary to normalise the results as muchas possible. Then, the integral (3.7) is computed with $X = x/a$ and $Y = y/a$, where "a" is a normalisation constant (for example the radius of the fiber). It is easily found that equations (3.7) remains unchanged except for "ko" which becomes:

$$\left(\frac{\omega a}{c} \right) \left(\gamma^2 - 1 \right)^{\frac{1}{2}} \qquad (3.34)$$

and dS_i becomes dXdY instead of dxdy.

3.4.2 Mapping

From standard calculus theorems, we can write :

$$F \{ \phi_i (X,Y) , \psi_i (X,Y) \} = F \{ \phi_i (X(\xi,\eta),Y(\xi,\eta)) , \psi_i (X(\xi,\eta),Y(\xi,\eta)) \}$$

$$F (\phi_i, \psi_i) \, dX \, dY = F (\phi_i, \psi_i) \, |J_i| \, d\xi \, d\eta \qquad (3.35)$$

Where $[J_i]$ is the Jacobian matrix corresponding to the transformation T_i. From equations (3.30) and (3.31), we can write:

$$J^i_{11} = \sum_{j=1}^{LN} \left(\frac{\partial \bar{N}_j}{\partial \xi} \right) X_{ij} \qquad (3.36)$$

$$J^i_{12} = \sum_{j=1}^{LN} \left(\frac{\partial \bar{N}_j}{\partial \xi} \right) Y_{ij} \qquad (3.37)$$

$$J^i_{21} = \sum_{j=1}^{LN} \left(\frac{\partial \bar{N}_j}{\partial \eta} \right) X_{ij} \qquad (3.38)$$

$$J^i_{22} = \sum_{j=1}^{LN} \left(\frac{\partial \bar{N}_j}{\partial \eta} \right) Y_{ij} \qquad (3.39)$$

Where $\partial \bar{N}_j / \partial \xi$ and $\partial \bar{N}_j / \partial \eta$ are the partial derivatives of the functions $\bar{N}_j$ in the reference element and X_{ij}, Y_{ij} are the x-y coordinates of the nodes of the element "i". "LN" is the Local number of nodes, i.e. the number of nodes in the element "i".

Using the interpolation functions N_j on the reference element, we have :

$$\phi_i = \sum_{j=1}^{LN} N_j \phi_{ij} \qquad (3.40)$$

$$\psi_i = \sum_{j=1}^{LN} N_j \psi_{ij} \qquad (3.41)$$

ϕ_{ij} and ψ_{ij} are the nodal values of the field at the nodes of the element "i". For an isoparametric and quadratic element, $\bar{N}_j = N_j$. Using some straight forward algebraic manipulations, we can evaluate the integral (3.7) in terms of the nodal values of the field ϕ_{ij} and ψ_{ij} [see Appendix 2].
We obtain :

$$\delta \sum_{i=1}^{N} \{ \int_0^1 d\xi \int_0^{1-\xi} d\eta \; \tau_i <\phi_{ij}> [U_i] \{\phi_{ik}\} + \qquad (3.42)$$

$$\gamma^2 \tau_i \left(\frac{\epsilon_i}{\epsilon_0}\right) <\psi_{ij}> [U_i] \{\psi_{ik}\} + 2\gamma^2 \tau_i <\psi_{ij}> [V_i] \{\phi_{ik}\} +$$

$$k_0^2 <\phi_{ij}> [W_i] \{\phi_{ik}\} + k_0^2 \gamma^2 \left(\frac{\epsilon_i}{\epsilon_0}\right) <\psi_{ij}> [W_i] \{\psi_{ik}\} \quad = 0$$

with

$$< \phi_{ij} > = \text{row matrix} \quad < \phi_{i1}, \phi_{i2}, \ldots \phi_{i,LN} > \tag{3.43}$$

$$\{ \phi_{ik} \} = \text{vector} \quad \{ \phi_{i1}, \phi_{i2}, \ldots \phi_{i,LN} \} \tag{3.44}$$

$$< \psi_{ij} > = \text{row matrix} \quad < \psi_{i1}, \psi_{i2}, \ldots \psi_{i,LN} > \tag{3.45}$$

$$\{ \psi_{ik} \} = \text{vector} \quad \{ \psi_{i1}, \psi_{i2}, \ldots \psi_{i,LN} \} \tag{3.46}$$

$[U_i]$, $[V_i]$ and $[W_i]$ are three matrices defined by :

$$\text{Element } (j,k) \text{ of } [U_i] = \frac{1}{|J_i|} \{ (J^{i2}_{22} + J^{i2}_{21}) \frac{\partial N_j}{\partial \xi} \frac{\partial N_k}{\partial \xi} + \tag{3.47}$$

$$(J^{i2}_{12} + J^{i2}_{11}) \frac{\partial N_j}{\partial \eta} \frac{\partial N_k}{\partial \eta} - (J^{1}_{12} J^{1}_{22} + J^{1}_{21} J^{1}_{11}) (\frac{\partial N_j}{\partial \xi} \frac{\partial N_k}{\partial \eta} + \frac{\partial N_k}{\partial \xi} \frac{\partial N_j}{\partial \eta}) \}$$

$$\text{Element } (j,k) \text{ of } [V_i] = (\frac{\partial N_j}{\partial \xi} \frac{\partial N_k}{\partial \eta} - \frac{\partial N_k}{\partial \xi} \frac{\partial N_j}{\partial \eta}) \tag{3.48}$$

$$\text{Element } (j,k) \text{ of } [W_i] = N_j N_k \, | J_i | \tag{3.49}$$

Where the N_j and N_k are the interpolation functions in the reference element.

$[J_i]$ is the Jacobian matrix associated with the element "i" (see equations 3.36 to 3.39).

$|J_i|$ is the determinant of $[J_i]$ and J^1_{11}, J^i_{12}, J^1_{21}, and J^i_{22} are the elements of $[J_i]$.

We can easily verify that all three matrices are symmetric and [V] is independent of the element.

3.4.3 Numerical integration

Integral (3.7) may be evaluated in terms of the unknown field values ϕ_{ij} and ψ_{ij} when the constants τ_i , γ and ε_i are known, and when the interpolation functions N_j have been defined by the type of element

selected. In our case, quadratic and isoparametric elements are used and the corresponding "N_J" functions have been given in table 1. The analytical determination of the integral is not practical because of the presence of polynomials in the denominator $1/|J_1|$.
Instead we use the common Gaussian method on the reference triangle. The integral :

$$\int_0^1 d\xi \int_0^{1-\xi} d\eta \; F(\phi_1, \psi_1) \; |J_i| \; d\xi \; d\eta \quad \text{becomes}$$

$$\sum_{R=1}^{NI} W_R \cdot F\{\phi_i(\xi_R, \eta_R), \psi_i(\xi_R, \eta_R)\} \; |J_1(\xi_R, \eta_R)| \qquad (3.50)$$

where :
NI is the number of intergration points. ξ_R and η_R are the coordinates of the integration points . W_R are the weights of the integration points. In our case, the value NI=6 have been chosen after several trials at the early stage of program development with 3 and 4 integration points. No numerical integration problems have been detected during the project and the accuracy was very satisfactory.

3.5 Computation of the transverse field components

We have seen in section 2.2 that every transverse field component can be found from the knowledge of the z-components and their partial derivatives. But the electromagnetic field is discretised, and the above quantities must be computed in the discretised space, using the Jacobian matrix of each element.

$$\text{If} \quad H_x = \phi_x \;, \; H_y = \phi_y \;, \; \frac{1}{\gamma\eta_0} E_x = \psi_x \quad \text{and} \quad \frac{1}{\gamma\eta_0} E_y = \psi_y \qquad (3.51)$$

Then, the equations (2.12 to 2.16) may be written in each element "i" :

$$\psi_x^1 = j \frac{1}{\gamma (\frac{\omega a}{c}) (\bar{\varepsilon}_i - \gamma^2)} \left(\frac{\partial \phi_i}{\partial y} + \gamma^2 \frac{\partial \psi_1}{\partial x} \right) \tag{3.52}$$

$$\psi_y^1 = j \frac{1}{\gamma (\frac{\omega a}{c}) (\bar{\varepsilon}_1 - \gamma^2)} \left(- \frac{\partial \phi_1}{\partial x} + \gamma^2 \frac{\partial \psi_i}{\partial y} \right) \tag{3.53}$$

$$\phi_x^1 = j \frac{\gamma}{(\frac{\omega a}{c}) (\bar{\varepsilon}_1 - \gamma^2)} \left(- \bar{\varepsilon}_i \frac{\partial \psi_1}{\partial y} + \frac{\partial \phi_1}{\partial x} \right) \tag{3.54}$$

$$\phi_y^1 = j \frac{\gamma}{(\frac{\omega a}{c}) (\bar{\varepsilon}_i - \gamma^2)} \left(\bar{\varepsilon}_1 \frac{\partial \psi_1}{\partial x} + \frac{\partial \phi_1}{\partial y} \right) \tag{3.55}$$

Where :

ϕ_1 and ψ_i are the usual z-components,

$\bar{\varepsilon}_1$ is the relative permittivity of the element ,

a is the usual normalising dimension.

The other variables are defined in section 2.2.

Then, we need to compute each partial derivative in each element with the Jacobian matrix.

We have :

$$\frac{\partial \phi_1}{\partial x} = \frac{1}{|J_i|} \left(J_{22}^1 \frac{\partial \phi_i}{\partial \xi} - J_{12}^i \frac{\partial \phi_i}{\partial \eta} \right) \tag{3.56}$$

$$\frac{\partial \phi_1}{\partial y} = \frac{1}{|J_i|} \left(-J_{21}^i \frac{\partial \phi_i}{\partial \xi} + J_{11}^i \frac{\partial \phi_i}{\partial \eta} \right) \tag{3.57}$$

The same relations hold for ψ.

$$| J_1 | = J_{11}^i J_{22}^i - J_{12}^i J_{21}^1 \tag{3.58}$$

Using the expression for the $[J_i]$ matrix components given in equations (3.36 to 3.39), we easily find :

$$\frac{\partial \phi_1}{\partial x} = \frac{1}{|J_i|} \left\{ \sum_j \sum_k \left(\frac{\partial N_j}{\partial \eta} \frac{\partial N_k}{\partial \xi} - \frac{\partial N_j}{\partial \xi} \frac{\partial N_k}{\partial \eta} \right) Y_{1j} \phi_{1k} \right\} \tag{3.59}$$

$$\frac{\partial \phi_i}{\partial y} = \frac{1}{|J_i|} \left\{ \sum_j \sum_k \left(\frac{\partial N_j}{\partial \xi} \frac{\partial N_k}{\partial \eta} - \frac{\partial N_j}{\partial \eta} \frac{\partial N_k}{\partial \xi} \right) X_{ij} \phi_{ik} \right\} \qquad (3.60)$$

"j" and "k" are the local index numbers inside each element.

These equations are similar for ψ_i.

$$\text{Let} \quad V_{jk} = \left(\frac{\partial N_j}{\partial \eta} \frac{\partial N_k}{\partial \xi} - \frac{\partial N_j}{\partial \xi} \frac{\partial N_k}{\partial \eta} \right) \qquad (3.61)$$

V_{jk} is independent of the element and may be computed at the beginning of the program and stored for the necessary points. Finally, we obtain :

$$\frac{\partial \phi_i}{\partial x} = \frac{1}{|J_i|} \sum_j \sum_k V_{jk} Y_{ij} \phi_{ik} \qquad (3.62)$$

$$\frac{\partial \psi_i}{\partial x} = \frac{1}{|J_i|} \sum_j \sum_k V_{jk} Y_{ij} \psi_{ik} \qquad (3.63)$$

$$\frac{\partial \phi_i}{\partial y} = \frac{1}{|J_i|} \sum_j \sum_k -V_{jk} X_{ij} \phi_{ik} \qquad (3.64)$$

$$\frac{\partial \psi_i}{\partial y} = \frac{1}{|J_i|} \sum_j \sum_k -V_{jk} X_{ij} \psi_{ik} \qquad (3.65)$$

$$|J_i| = \sum_j \sum_k -V_{jk} X_{ij} Y_{ik} \qquad (3.66)$$

X_{ij} and Y_{ik} (Y_{ij}) are the coordinates of the nodes of the element "i".

ϕ_{ik} and ψ_{ik} are the nodal values of the field. V_{jk} is a function of (ξ, η) in the reference element. The partial derivatives at every point (X,Y) in the real element may be found with the preceding relations when the corresponding coordinates (ξ,η) are known. In our case, we only look for the partial derivatives at the nodes. Therefore, the corresponding points in the reference element are the nodes of the triangle.

In our Finite Element approximation, the continuity of the derivatives is not mathematically exact at the boundaries between the homogeneous elements, but if these are small enough, the discon-

tinuity should be very small.

The problem of the non-uniqueness of the partial derivatives at the nodes may be solved by averaging them in the elements of same permittivity only. At the boundaries between regions with different index of refraction, two derivatives are computed and the transverse components will be discontinuous.

The discontinuities left in homogeneous regions are due to the discretisation process and occur in the low field intensity regions or for high order modes.

3.6 The Global Matrix Formulation

So far, we have written our integral equation in terms of the ϕ_{ij} and ψ_{ij} nodal values. This double index is not practical and is confusing. When a node is shared between several elements, different combinations of the i-j indices represent the same physical node called "n". By construction of the mesh, there exists a transformation which gives to any node labelled "ij" a single label "n". We will express the integral in terms of this unique node numbering system.

3.6.1 The local matrix formulation

$$\text{Let} \qquad <\theta_i> = <\phi_{i1},\psi_{i1},\phi_{i2},\psi_{i2}, \ldots> \qquad (3.67)$$

We want to write the integral (3.7) as :

$$\delta \sum_{i=1}^{N} <\theta_i> [A_i]\{\theta_i\} + k_0^2 <\theta_i> [B_i] <\theta_i> = 0 \qquad (3.68)$$

To determine each element A^i_{pq} (or B^i_{pq}) of the matrix $[A_i]$ (or $[B_i]$), we must find the coefficient of $\theta_p\theta_q$ from the equations (3.42 to 3.49). We obtain the tables below :

A^1_{pq}	p (odd) = 2j-1	p (even) = 2j
q (odd) = 2k-1	$\tau_1 \sum_R W_R U^i_{jk}(\xi_R,\eta_R)$	$\gamma^2\tau_i \sum_R W_R V_{jk}(\xi_R,\eta_R)$
q (even) = 2k	symmetric	$\gamma^2(\frac{\varepsilon_i}{\varepsilon_0})\ A^1_{p(odd)q(odd)}$

TABLE 2 : Local $[A_i]$ matrix.

B^1_{pq}	p (odd) = 2j-1	p (even) = 2j
q (odd) = 2k-1	$\sum_R W_R W^1_{jk}(\xi_R,\eta_R)$	0
q (even) = 2k	0	$\gamma^2(\frac{\varepsilon_1}{\varepsilon_0})\ B^1_{p(odd)q(odd)}$

TABLE 3 : Local $[B_i]$ matrix.

- $[A_i]$ and $[B_i]$ are symmetric.

- U^i_{jk}, V_{jk} and W^1_{jk} are given in equations (3.47 to 3.49).
- The summation in "R" represents the numerical integration using the Gaussian method.

3.6.2- The Local to Global Matrix Transformation

Each term U^1_{jk} or W^1_{jk} represents the contribution to the global matrix of the pair of nodes j-k in the element 1. If "n" is the global node number of the local node j of the element "i", and "m" the global node number of the local node "k" of the element "i", then it is possible to find a correspondance between the local position "pq" of each element A^i_{pq} (B^i_{pq}) of the LOCAL matrix and the global position "gh" of the same element in the global matrix $[A_G]$ ($[B_G]$).

Let n(i,j) and m(i,j) be the global node numbers as defined previously. Then, we obtain the following table :

	p(odd) = 2j-1	p(even) = 2j
q(odd) = 2k-1	g = 2n(i,j)-1 h = 2m(i,k)-1	g = 2n(i,j) h = 2m(i,k)-1
q(even) = 2k	g = 2n(i,j)-1 h = 2m(i,k)	g = 2n(i,j) h = 2m(i,k)

n(i,j) and m(i,k) are defined by the numbering system chosen by the user.

TABLE 4 : The local to global matrix transformation.

3.6.3 Effect of the boundary conditions

The enforcing of essential boundary conditions ($\phi = 0$ or $\psi = 0$ for some nodes) will slightly modify the global matrices and reduce their size.

If we have the initial matrix equation :

$$< \theta >[A_G]\{\theta\} + k_0^2 < \theta >[B_G]\{\theta\} = 0 \qquad (3.69)$$

and if $\theta_L = 0$, then the column "L" and the row "L" of both matrices $[A_G]$ and $[B_G]$ may be removed. The size of the system is reduced by one, and the process repeated for every node subject to an explicit boundary condition.

3.6.4 Final answer

The final result of this complex process is a very large matrix equation of the form :

$$\delta\{ \langle\Theta\rangle[A_G]\{\Theta\} + k_0^2 \langle\Theta\rangle[B_G]\{\Theta\} \} = 0 \qquad (3.70)$$

From reference [12], it has been proved that the stationary point of this expression is the solution of the following eigenvalue problem:

$$[A_G]\{\Theta\} + k_0^2 [B_G]\{\Theta\} = 0 \qquad (3.71)$$

(with $[A_G]$, $[B_G]$ large, sparse, symmetric matrices; $[A_G]$ is indefinite and $[B_G]$ is positive definite).

Where :

- $[A_G]$ and $[B_G]$ are the two global matrices assembled from the local matrices defined in table 2 and 3.
- $\{\Theta\}$ is the vector of all the nodal values of ϕ and ψ, the explicit boundary conditions being included.

$$- \quad k_0^2 = \left(\frac{\omega a}{c}\right)\left(\gamma^2 - 1\right)$$

ko is an independent parameter which does not appear in any coefficient of the matrices $[A_G]$ and $[B_G]$.

For a given geometry and a given value of γ (the normalised propagation constant), we can solve the problem for the eigenvalues ko corresponding to different modes of the structure. From these values, dispersion curves may be plotted. By convention, we call $S = \omega a/c$, the normalised frequency. "a" is a normalising dimension and "c" is the speed of light in vacuum (c = 299800 km/sec).

3.7 The Virtual Boundary

Outside the dielectric waveguide, and if there is no opaque coating, the electromagnetic field decays exponentially for a theoretically infinite distance. But it is obviously not possible to discretise this infinitely large area with a finite number of elements of finite size.

Some authors [13,16] have used special elements in the unbounded region with intrinsically decaying interpolation functions. In this approach, the user has to guess the decaying factor with some

accuracy to keep the results in acceptable range. In our case, we use the same type of element everywhere in a finite region, and set both the electric and magnetic fields to a true zero at a certain distance from the guide, on the so-called virtual boundary (see references [13,15]).

The determination of this boundary could be done iteratively, with step by step moves, expecting a convergence of the eigenvalues for sufficiently large values of the boundary distance. But this process would be very costly in term of computer time.

A better approach is to consider an equivalent simple model (the dielectric rod) for the arbitrarily shaped guide when it is seen from a relatively large distance (see fig. 2). It is then possible to compute an approximate field decay rate to obtain the location of the virtual boundary (see fig. 2). A separate program (DW3) analyses the dielectric rod with great accuracy. This method has been used successfully in the different cases studied in section 4.

The effect of the virtual boundary location on the accuracy of the answer has been numerically studied in one example. It shows an optimum range of values between which the results are very close to each other, within a so-called "numerical noise level" (see fig. 2). This expression refers to apparent random variation of the eigenvalues with small variations of relatively irrelevant parameters.

When the boundary is too close, the approximation of zero field leads to a fast-growing error. When the boundary is too far, the elements outside the waveguide become too large and the interpolation functions do not follow the fast decay rate.

Table 5 and Fig. 3 show an example of the virtual boundary location effect. The case studied is a dielectric rod with an internal index of 1.5, embedded in air. The parameter $R_{max} = 1$ corresponds to the waveguide interface. The propagation constant is 1.45. The number of elements is constant.

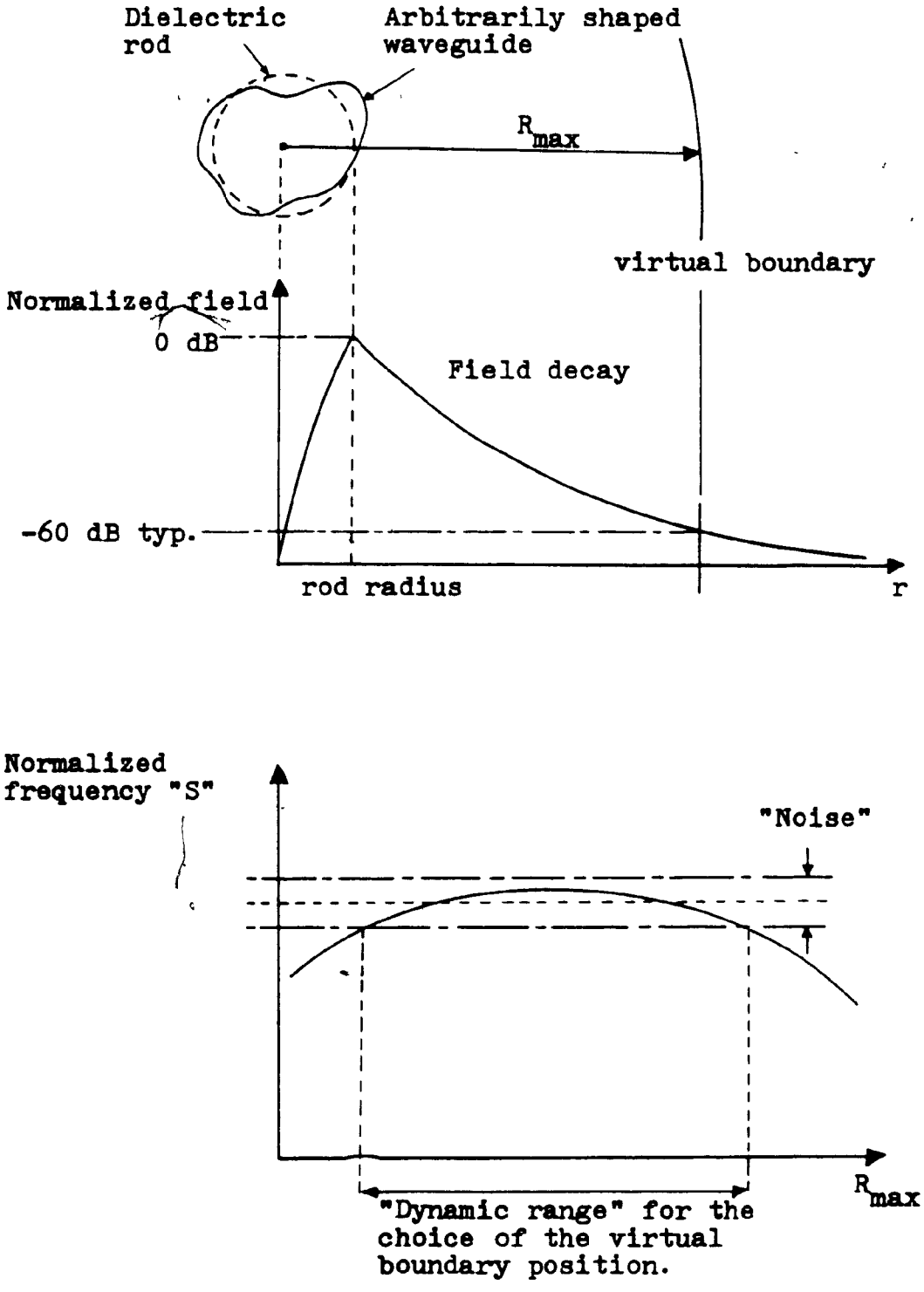


Fig. 2 : Determination of the virtual boundary with the equivalent model of the dielectric rod.

Virtual boundary location (1 = rod radius)	First real mode		First spurious mode	
	S	% Change	S	% Change
1.15	5.34032	-	none	-
1.21	5.51138	(+3.20)	1.44239	(-)
1.36	5.61845	(+1.94)	3.01922	(+109)
1.50	5.63368	(+0.27)	3.21124	(+6.4)
1.71	5.63676	(+0.05)	3.25987	(+1.5)
1.86	5.63710	(+0.01)	3.26134	(+0.05)
2.00	5.63705	(+0.00)	3.25728	(-0.12)
2.50	5.63682	(-0.01)	3.23414	(-0.71)
3.00	5.63640	(-0.01)	3.20989	(-0.75)
4.00	5.63588	(-0.01)	3.16335	(-1.45)
5.00	5.63508	(-0.01)	3.12113	(-1.33)
6.00	5.63446	(-0.01)	3.08247	(-1.24)
7.00	5.63377	(-0.01)	3.04762	(-1.13)
8.00	5.63308	(-0.01)	3.01551	(-1.05)
10.00	5.63171	(-0.02)	2.95873	(-1.88)

TABLE 5 : Effect of the virtual boundary

S is the normalised frequency $\omega a/c$.

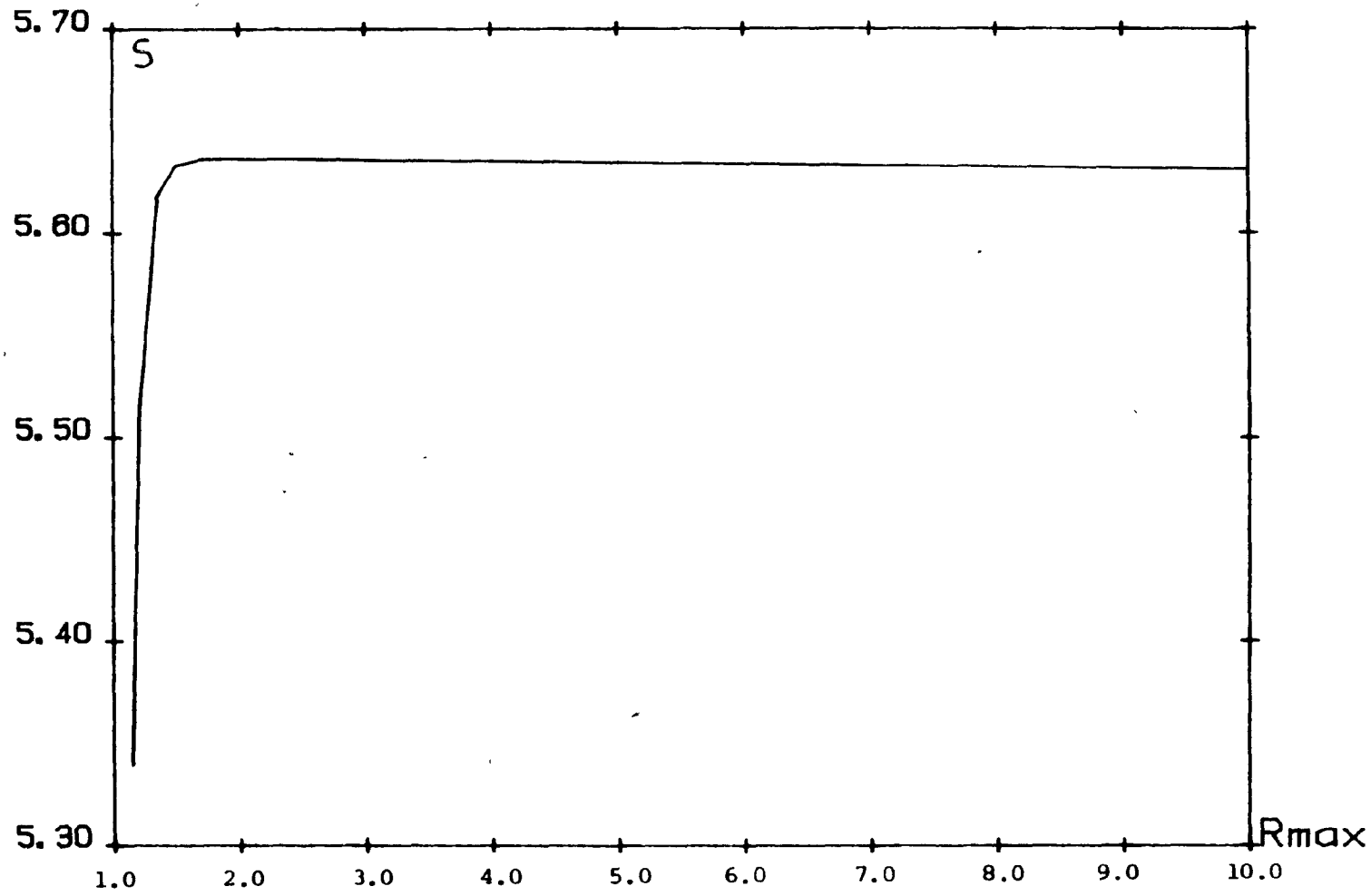


Fig. 3 : Variation of the fundamental mode frequency at $\gamma = 1.45$ with respect to the location of the virtual boundary Rmax.

3.8 Symmetry and Boundary Conditions

The use of symmetry in the analysis may greatly reduce the size of the discretised problems and must be included. When an axis of symmetry exists, only one half of the problem needs to be modelled, providing the correct boundary condition is imposed on the axis. The conditions f = 0 (f being one of the z-components) on one axis indicates that f is odd with respect to this axis. When no explicit boundary condition is set, the natural boundary conditions induced by the variational formulation holds and f is even with respect to the axis (see section 3.1). The electromagnetic field is different in each case, but not necessarily the characteristic curves of the propagation constant. When two axes of symmetry exist (Ox and Oy), a set of four boundary conditions may be defined which are :

"KBC" code	(Ox axis) $Ez = \psi = 0$	(Ox axis) $Hz = \phi = 0$
$Ez = \psi = 0$ (Oy axis)	4	3
$Hz = \phi = 0$ (Oy axis)	2	1

TABLE 6 : Definition of the "KBC boundary conditions code"

When two sets of boundary conditions lead to the same propagation constant (at any frequency), the two corresponding modes are said "degenerate". This degeneracy is the key phenomenon for the analysis of the coupler and will be discussed later.

4. ANALYSIS OF KNOWN EXAMPLES WITH THE FEM PROGRAMS

A set of computer programs described in section 6 has been written for the FEM analysis of dielectric waveguides. Four examples have been chosen to demonstrate the most important characteristics of the method. The results are given in this chapter.

4.1 The Dielectric rod analysis

The dielecric rod is the simplest dielectric waveguide. A cylinder with an index of refraction n1 is embedded in a region of index n2, and n1 > n2. The following proof indicates that guided waves will propagate inside this guide.

4.1.1 Analytical solution

The analytical solution of this problem is well known and is obtained in appendix 3. It is easily found in many Electromagnetic books [e.g. ref 18].
The normalised frequency $S = \omega a/c$ is the solution of the following transcendental equation :

$$\left\{n\gamma\left(\frac{1}{S_1^2}+\frac{1}{S_2^2}\right)\right\}^2 - \left\{n\left(\frac{1}{S_1^2}+\frac{1}{S_2^2}\right) - \frac{K_{n+1}(S_2)}{S_2K_n(S_2)} - \frac{J_{n+1}(S_1)}{S_1J_n(S_1)}\right\} \cdot$$

$$\left\{n\left(\frac{n_1^2}{S_1^2}+\frac{n_2^2}{S_2^2}\right) - \frac{n_2^2K_{n+1}(S_2)}{S_2K_n(S_2)} - \frac{n_1^2J_{n+1}(S_1)}{S_1J_n(S_1)}\right\} = 0 \qquad (4.1)$$

with $$S_1 = \left(\frac{\omega a}{c}\right)\sqrt{n_1^2 - \gamma^2} \qquad (4.2)$$

$$S_2 = \left(\frac{\omega a}{c}\right)\sqrt{\gamma^2 - n_2^2} \qquad (4.3)$$

and $n_1^2 = \varepsilon_1$ = dielectric constant inside the rod.

$n_2^2 = \varepsilon_2$ = dielectric constant outside the rod.

The electromagnetic field z-components are

$$Ez\ (inside) = A_1 J_n(k_1 r)\ \cos(n\phi)\ e^{j(\beta z - \omega t)} \qquad (4.4)$$

$$Hz\ (inside) = B_1 J_n(k_1 r)\ \sin(n\phi)\ e^{j(\beta z - \omega t)} \qquad (4.5)$$

$$Ez(outside) = A_2 K_n(k_2 r)\ \cos(n\phi)\ e^{j(\beta z - \omega t)} \qquad (4.6)$$

$$Hz(outside) = B_2 K_n(k_2 r)\ \sin(n\phi)\ e^{j(\beta z - \omega t)} \qquad (4.7)$$

$$k_1 = \left(\frac{\omega}{c}\right) \sqrt{n_1^2 - \gamma^2} \qquad (4.8)$$

$$k_2 = \left(\frac{\omega}{c}\right) \sqrt{\gamma^2 - n_2^2} \qquad (4.9)$$

The results are shown for some modes on Fig. 4.
"n" is the number which appears in the previous equations and determines the order of the Bessel functions.
For each "n" (0,1,2,...) there are an infinite number of frequencies which satisfy (4.1) and (4.2) at a given γ ($n1>\gamma>n2$). The values of ω and corresponding fields (4.4) to (4.7) are labelled with "n" and a second index "m" (0,1,2,...).
Usually, the first mode is defined as the one with the lowest frequency for a given value of the propagation constant. In a dielectric rod, it is an hybrid mode and all the field components are non-zero, even if some are clearly much larger than other. Such a mode is usually denoted HE_{nm}, where n and m have been defined previously.

4.1.2 FEM solution

The dielectric rod and the surrounding medium have been discretised with 50 quadratic elements and 129 nodes. Some results have been checked with 62 elements and 159 nodes and no appreciable difference had been noticed (see Fig. 5 for the discretisation). Tables 7, 8 and 9 give the numerical results and the corresponding errors for the normalised frequency as a function of the normalised propagation constant. Because of the very small error level, theoretical values

have been computed in double precision, with accurate determination of the Bessel functions using series expansions. The error is below 0.035% over the γ-range (1.05-1.49) for dominant HE_{11} mode. Some components of the electromagnetic field have been displayed with a contour plotter program (DW6P, see section 6.4.2) for γ = 1.45 [see Fig. 6 to 10]. The smoothness of the equipotential lines and the continuity of their partial derivatives at the intersection of elements (except at the boundary and for E_x and H_y) is another proof of the quality of the FEM discretisation and interpolation process. The poor quality of the contours of E_x and H_y can be explained by the very low relative amplitude level of E_x and H_y with respect to E_y and H_x. This is easily demonstrated by comparing the maximum field strength values of the x and y components. In fact, E_y and H_x are dominant for that set of boundary conditions (KBC=2), and become very small for the other set (KBC=3).

The number denoted "No" on the tables 7, 8 and 9 indicates the order in which the eigenvalues are classified when starting from the smallest positive one.

The poor aspect ratio of some elements exhibited in Fig. 5 does not appear to cause inaccuracy.

4.1.3 Higher order modes

The higher eigenvalues give also the normalised frequency of some of the high order modes, but with an accuracy which falls down very quickly because the rate of change of the field becomes very high and cannot be followed by the interpolation functions. For the next mode (HE_{12}), the accuracy remains very good, below 0.08% over the γ-range (1.05-1.49). Other modes exist between HE_{11} and HE_{12} which cannot be found with the set of boundary conditions (KBC=2) which has been used in this example.

An efficient way of computing modes such that the field periodicity of the z-components are of the form $\sin(n\phi)$ or $\cos(n\phi)$ with "n" greater then one is to simply put the elements in a section of disc $\pi/2n$ wide. This increases the density of elements for a better resolution at no extra cost [see Fig. 11]. This feature is succes-

sfully used to compute the modes HE_{11}, HE_{12} and HE_{13} in the next example.

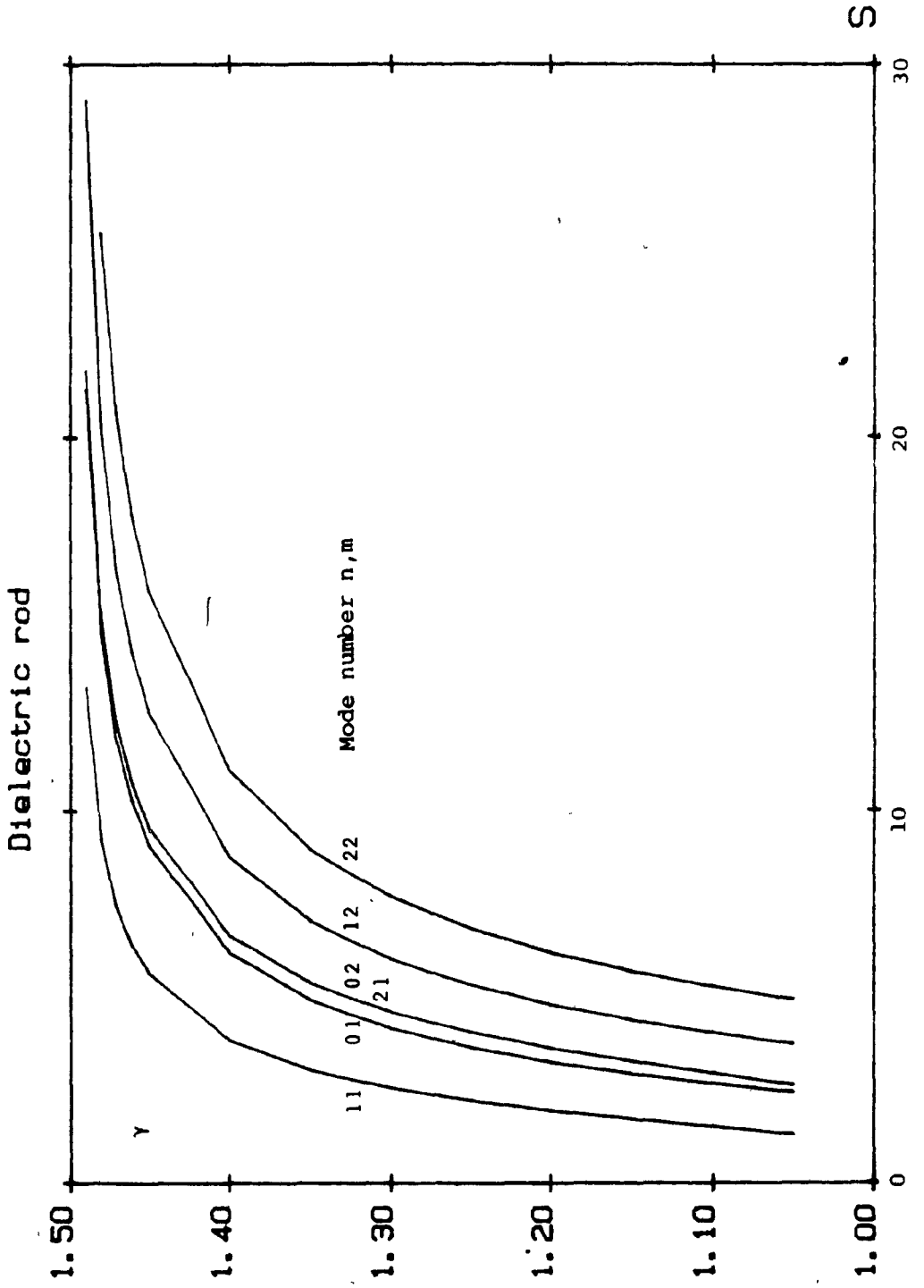
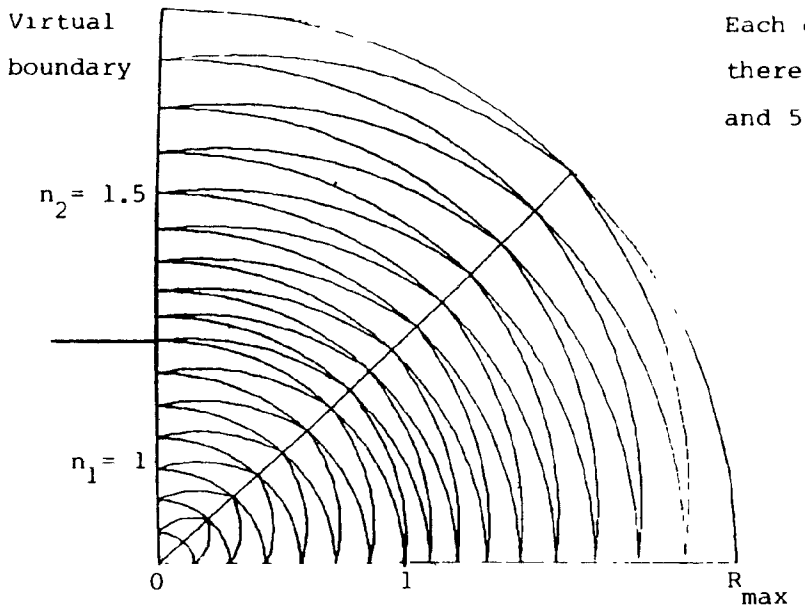


Fig. 4 : Propagation characteristics for the first modes of a dielectric rod.



Each element has 6 nodes, there are 33 radial divisions and 5 angular divisions.

Fig. 5 : Dielectric rod example, discretisation with 62 elements and 159 nodes.

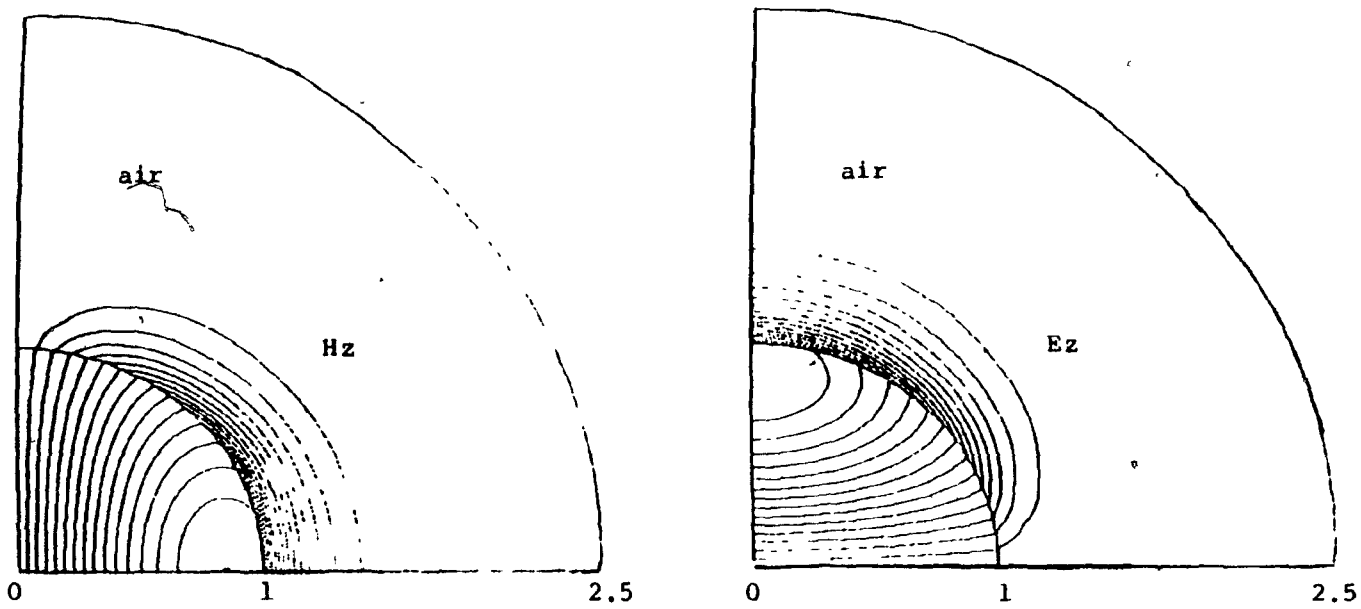


Fig. 6 : Fundamental mode of a dielectric rod : Field plots.(HE_{11})
Propagation constant = 1.450, boundary conditions code = 2.

γ	m=1 S (Th.)	No	S (FEM) *1	% Error	γ	m=2 S (Th.)	No	S (FEM) *2	% Error
1,05	2.41752	1	2.41941	+.08	1.05	2.61070	1	2.61281	+.08
1.10	2.65381	2	2.65588	+.08	1.10	2.93867	2	2.94106	+.08
1.15	2.91655	1	2.91877	+.08	1.15	3.26071	1	3.26343	+.08
1.20	3.22905	1	3.23171	+.08	1.20	3.61461	2	3.61756	+.08
1.25	3.62159	1	3.62439	+.08	1.25	4.03736	2	4.04073	+.08
1.30	4.14601	1	4.14935	+.08	1.30	4.58452	2	4.58837	+.08
1.35	4.90871	1	4.91241	+.08	1.35	5.36480	2	5.36932	+.08
1.40	6.18377	2	6.18889	+.08	1.40	6.65442	3	6.66011	+.09
1.45	9.06490	1	9.07118	+.07	1.45	9.54682	2	9.55472	+.08
1.46	10.2273	1	10.2338	+.06	1.46	10.7115	2	10.7206	+.08
1.47	11.9319	1	11.9403	+.07	1.47	12.4184	2	12.4288	+.08
1.48	14.7940	2	14.8033	+.06	1.48	15.2829	3	15.2961	+.09
1.49	21.2615	2	21.2791	+.08	1.49	21.7531	3	21.7718	+.09

FEM Solution : 50 elements, 129 nodes. Note *1 : KBC = 4 (m = 1)

No = Eigenvalue number. Note *2 : KBC = 1 (m = 2)

TABLE 7 : The dielectric rod example with mode number (n) = 0

	HE_{11}					HE_{12}			
	m=1					m=2			
γ	S (Th.)	No*2	S (FEM)*1	% Error	γ	S (Th.)	No*2	S (FEM)*1	% Error
1.05	1.290371	1	1.290175	-.015	1.05	3.71919	2	3.72199	+.075
1.10	1.510383	1	1.510468	+.006	1.10	4.02084	2	4.02389	+.076
1.15	1.71591	2	1.71619	+.016	1.15	4.36556	3	4.36900	+.079
1.20	1.938383	1	1.938836	+.023	1.20	4.78074	2	4.78424	+.073
1.25	2.20246	1	2.20302	+.026	1.25	5.30594	2	5.30988	+.074
1.30	2.54322	1	2.54397	+.030	1.30	6.01055	2	6.01514	+.076
1.35	3.02859	1	3.02956	+.032	1.35	7.03786	3	7.04330	+.077
1.40	3.83111	1	3.83237	+.033	1.40	8.75663	2	8.76278	+.070
1.45	5.63521	2	5.63680	+.028	1.45	12.6354	3	12.6440	+.068
1.46	6.36240	2	6.36446	+.032	1.46	14.1989	3	14.2074	+.060
1.47	7.42959	1	7.43224	+.036	1.47	16.4904	3	16.5036	+.080
1.48	9.22224	1	9.22550	+.035	1.48	20.3352	3	20.3491	+.069
1.49	13.2758	1	13.2804	+.035	1.49	29.0159	3	29.0342	+.063

*1 FEM solution with 50 elements and 129 nodes.

*2 Eigenvalue number.

TABLE 8 : The dielectric rod example with mode number (n) = 1

	HE_{21}					HE_{22}			
		m=1					m=2		
γ	S (Th.)	No*2	S (FEM)*1	% Error	γ	S (Th.)	No*2	S (FEM)*1	% Error
1.05	2.78442	2	2.7533	-1.1	1.05	4.91063	4	4.9377	+0.6
1.10	3.02123	3	3.0108	-0.3	1.10	5.26356	5	5.2954	+0.6
1.15	3.28023	2	3.2749	-0.2	1.15	5.67710	3	5.6962	+0.3
1.20	3.58649	2	3.58901	+0.07	1.20	6.18157	3	6.2177	+0.6
1.25	3.97093	2	3.97471	+0.1	1.25	6.82447	3	6.8582	+0.5
1.30	4.48556	2	4.49332	+0.2	1.30	7.69076	4	7.7297	+0.5
1.35	5.23657	2	5.24896	+0.2	1.35	8.95834	3	9.0030	+0.5
1.40	6.49729	3	6.52200	+0.4	1.40	11.0842	4	11.139	+0.5
1.45	9.35970	2	9.35312	-0.07	1.45	15.8898	4	15.966	+0.5
1.46	10.5172	2	10.5202	+0.03	1.46	17.8287	4	17.911	+0.5
1.47	12.2163	2	12.2295	+0.1	1.47	20.6715	4	20.778	+0.5
1.48	15.0719	3	15.0915	+0.1	1.48	25.4430	5	25.567	+0.5
1.49	21.5308	3	21.5817	+0.2	1.49	36.2206	5	36.479	+0.7

*1 FEM solution with 50 elements and 129 nodes.

*2 Eigenvalue number.

TABLE 9 : The dielectric rod example with mode number (n) = 2

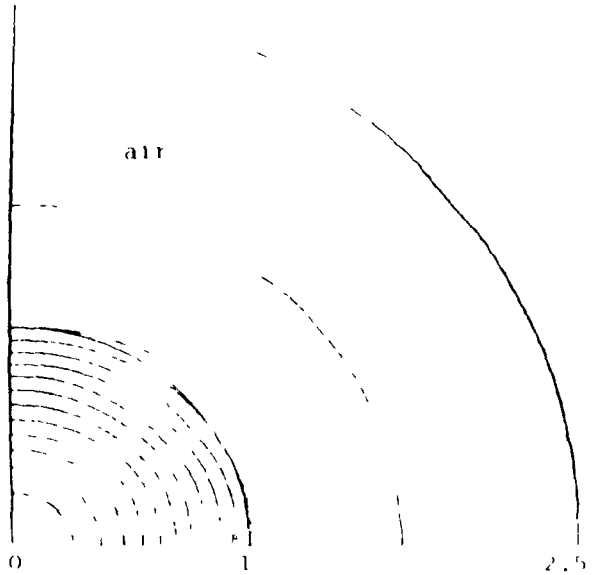


Fig. 7 · Dielectric rod, Hx field, linear scale, γ = 1.450 , fundamental mode. (HE_{11})
(Fy and Sz are similar)

This component is very small with respect to the dominant one. Therefore, the FEM solution was not accurate enough to give proper contours.

Fig. 8 Dielectric rod, Ex field, linear scale, γ 1.450 , fundamental mode. (HE_{11})
(Hy is similar, field values are small)

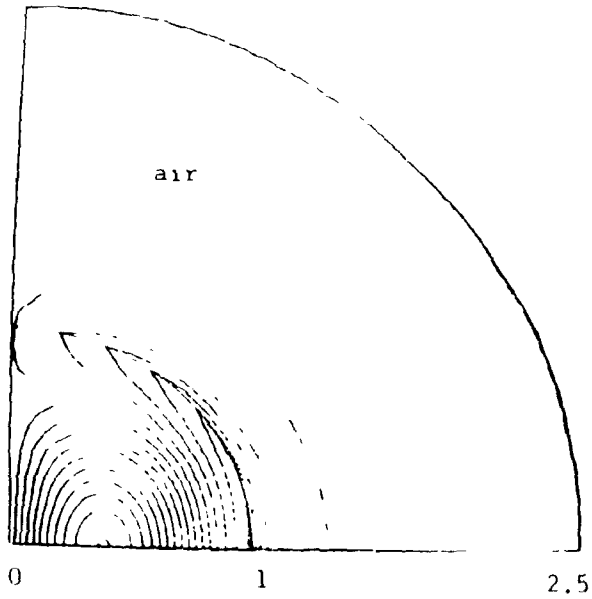


Fig. 9 : Dielectric rod, Hz field, higher order mode, γ = 1.450, linear scale, boundary conditions code = 2. (HF_{12})

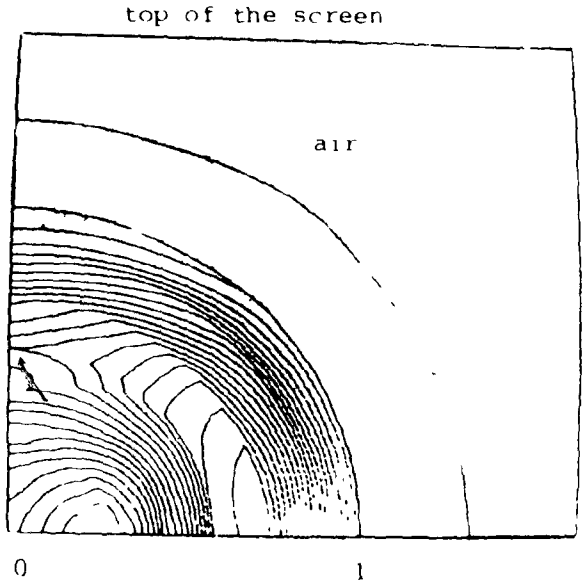
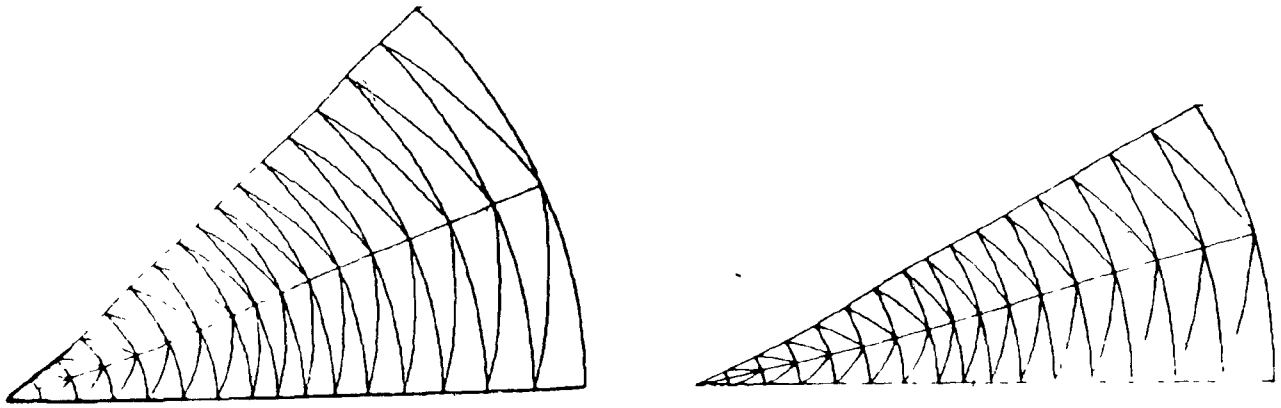


Fig. 10 : Dielectric rod, Sz component, linear scale (x2), higher order mode, γ = 1.450, boundary conditions code = 2. (HF_{12})



1/8 circle

1/12 circle

Fig. 11 : Discretisation of a circular waveguide for the computation of higher order modes propagation characteristics (typical examples).

4.2 The Inhomogeneous Waveguide

4.2.1 Runge-Kutta differential analysis

Great attention has been paid to the inhomogeneous waveguide in the past due to its dispersion-reduction properties in multimode applications. Various analytical and variational methods [4,5,19,20] have been developed for the analysis. The inhomogeneous fiber has a slowly varying index of refraction in the radial direction (Fig.12). One of the methods used a differential Runge-Kutta step by step approach which was developed and implemented some years ago at McGill University. Because it is a one-dimensional problem, the 2-D FEM method is obviously not needed for routine analysis, but the comparison between the results from these two methods greatly improves the confidence in the FEM programs. The detailed Runge-Kutta analysis is available in reference [19].

4.2.2 FEM solution

A step index model shown in Fig. 12 is used to approximate the parabolic profile of an inhomogeneous fiber with a finite cladding. 62 elements and 159 nodes are generated to map the structure. Table 10 gives the results from both methods and shows a excellent accuracy, except for the first value of ν. The Runge-Kutta results are computed in double precision. At very low frequency, the field covers a large area and the number of elements required outside for a good accuracy should be greater. But this lower edge of the spectrum is not usually reached at optical wavelength.

For $\lambda = 1\mu$, $\omega a/c = 2\pi a/\lambda$, a normalised frequency of one gives a radius "a" of .16 μ which is extremely small.

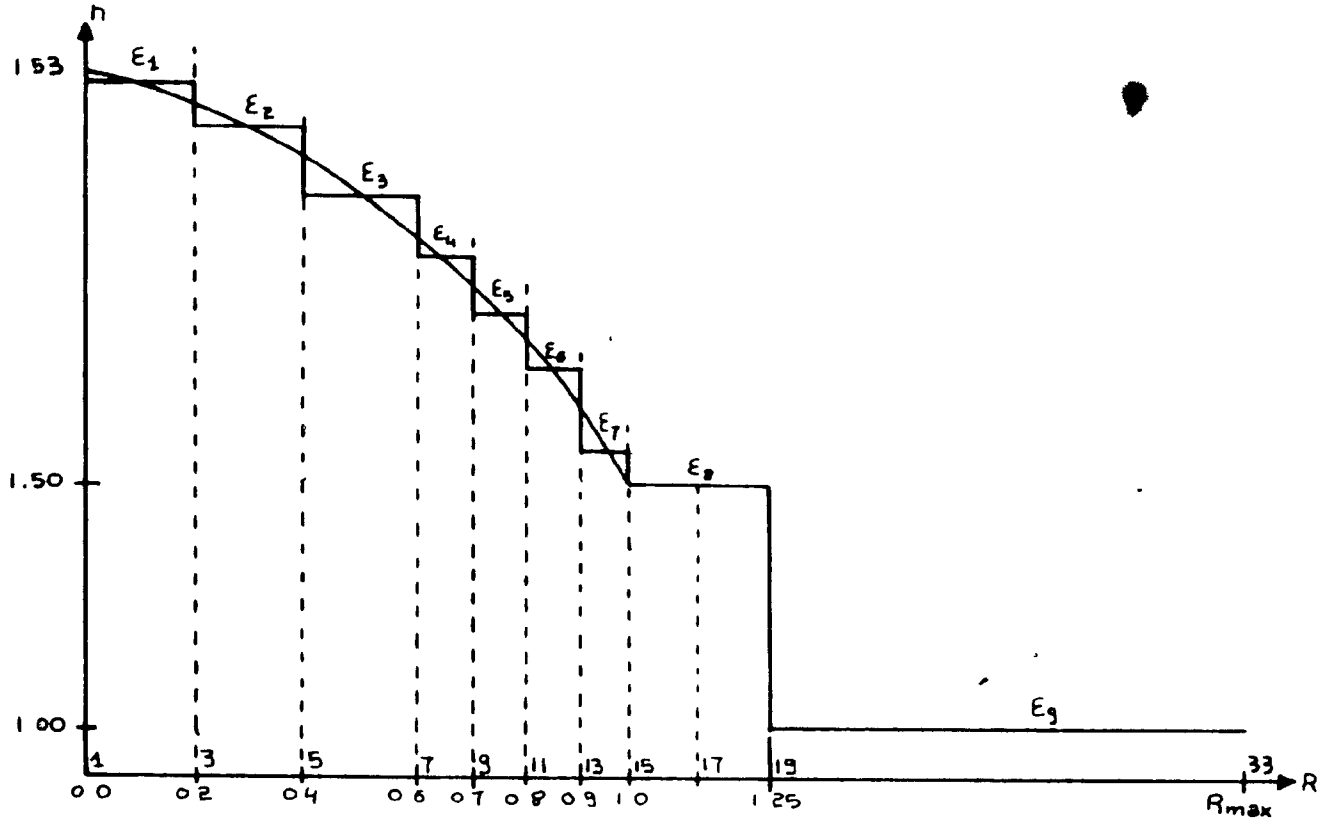


Fig. 12 : **The inhomogeneous optical fiber.**
Step index model of the parabolic profile.

	HE_{11}			HE_{21}			HE_{31}		
γ	S (RK) *4	S (FEM) *1	% Error	S (RK) *4	S (FEM) *2	% Error	S (RK) *4	S (FEM) *3	% Error
1.04	0.977610	0.976640	-.100	2.170030	2.164290	-.260	3.20710	3.19167	-.48
1.10	1.187251	1.187124	-.011	2.389231	2.386839	-.100	3.458954	3.452893	-.180
1.14	1.311962	1.311868	-.007	2.545459	2.543810	-.065	3.652222	3.649044	-.087
1.20	1.510205	1.510257	+.003	2.819214	2.817969	-.044	4.003891	4.001578	-.058
1.24	1.662580	1.662602	+.001	3.043603	3.042554	-.034	4.299130	4.297761	-.032
1.30	1.949522	1.949422	-.005	3.485835	3.485218	-.018	4.891114	4.890539	-.012
1.34	2.208118	2.208048	-.003	3.898263	3.895336	-.075	-	5.447305	-
1.40	2.807512	2.807164	-.012	4.885428	-	-	6.811314	6.807586	-.055

*1 : Mode numbers n=1, m=1 ; computed with 62 elements and 159 nodes in 1/4 fiber.

*2 : Mode numbers n=2, m=1 ; " " " " " " " " 1/8 fiber.

*3 : Mode numbers n=3, m=1 ; " " " " " " " " 1/12 fiber.

*4 : "RK" solution computed in double precision

TABLE 10 : Numerical results for the inhomogeneous optical fiber.

4.3 The Elliptical Fiber

4.3.1 Description

The elliptical fiber example chosen for this section is fully described and analysed with another method in reference [21]. It is made of 3 dielectric layers, the last one being infinitely wide. The two boundaries have a different ellipticity [The ellipticity ϵ is defined such that any point x-y of the ellipse verifies the following relation : $x = R\cos(\phi)$, $y = \epsilon R \sin(\phi)$ and ϕ is a parameter between 0 and 360 degrees.].
The virtual boundary is given the same ellipticity as the outer edge of the cladding. Fig. 13 shows one quarter of this structure.
The important result in this example is the $\Delta\gamma$ characteristic curve between the two quasi-degenerate modes corresponding to the two sets of boundary conditions KBC=2 and KBC=3. Tables 11, 12 and Fig. 14, 15 give the results from both methods and exhibit a good agreement between them with only 46 elements and 119 nodes.

4.3.2 The $\Delta\gamma$ characteristics

The difference between the propagation constants of two modes is not directly available from the output of the eigenvalue solver. This one gives the normalised frequency as a function of the normalised propagation constant for only a small number of points.
The method which computes the $\Delta\gamma$ characteristics at any frequency S is the following :

Each curve (γ versus S) is approximated with a piecewise quadratic function. For three data points, we can find three parameters A, B and C such that :

$$\gamma = AS^2 + BS + C \quad , \quad S = \frac{\omega a}{c} \tag{4.10}$$

If the three data points are (γ_1, S_1), (γ_2, S_2) and (γ_3, S_3), we easily find the following relations :

$$A = \frac{(S_2-S_1)(\gamma_3-\gamma_1) + (S_1-S_3)(\gamma_2-\gamma_1)}{(S_3-S_1)(S_2-S_1)(S_3-S_2)} \qquad (4.11)$$

$$B = \frac{\gamma_2-\gamma_1}{S_2-S_1} - A(S_1+S_2) \qquad (4.12)$$

$$C = \gamma_1 - AS_1^2 - BS_1 \qquad (4.13)$$

The complete curve γ versus S is completely defined by a set of A, B and C coefficients. Therefore, for any S within the range covered by the data points, we first find in which segment $[S_i, S_{i+2}]$ it belongs and using the corresponding coefficients, we can compute the quadratic approximation of γ.

4.3.3 Error analysis

Consider two modes at a given frequency S with propagation constants γ_1 and γ_2. Let $\Delta\gamma = \gamma_1 - \gamma_2$.
The $\Delta\gamma$ values may be relatively very small and the analysis of the worst case error is very important to determine the accuracy of the results.
The error is computed with respect to the expected error on the S-value (ΔS) from the solver. Examples previously studied in section 4 give us some upper limit for ΔS.
If :

$$\Delta(\Delta\gamma) = \Delta\gamma_1 + \Delta\gamma_2 \qquad (4.14)$$

$\Delta(\Delta\gamma)$ is the absolute error on $\Delta\gamma$, $\Delta\gamma_1$ and $\Delta\gamma_2$ are the absolute errors on γ_1 and γ_2 respectively.

then $$\Delta(\Delta\gamma) = \left(\frac{d\gamma_1}{dS}\right)\Delta S_1 + \left(\frac{d\gamma_2}{dS}\right)\Delta S_2 \qquad (4.15)$$

Define $\frac{\Delta(\Delta\gamma)}{\Delta\gamma}$ as the relative error on $\Delta\gamma$, $\frac{\Delta S}{S}$ as the relative error on S, and assume $\Delta S_1 = \Delta S_2$.

Then $$\frac{\Delta(\Delta\gamma)}{\Delta\gamma} = \left\{ \frac{S}{\Delta\gamma} \left(\frac{d\gamma_1}{dS} + \frac{d\gamma_2}{dS} \right) \right\} \frac{\Delta S}{S} \qquad (4.16)$$

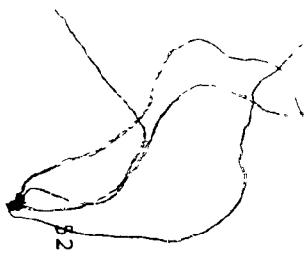
The coefficient $\frac{S}{\Delta\gamma} \left(\frac{d\gamma_1}{dS} + \frac{d\gamma_2}{dS} \right)$ is the error coefficient "EPS".

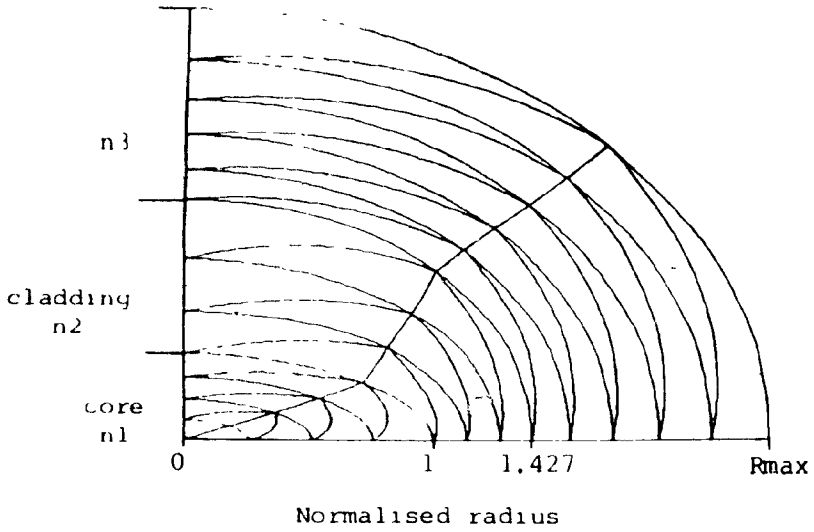
The quantities $d\gamma_1/dS$ and $d\gamma_2/dS$ are computed using the same quadratic interpolation defined previously. We obtain :

$$d\gamma/dS = 2AS + B \qquad (4.17)$$

This is true for any mode.

The $\Delta\gamma$ analysis is done in the program called "DBETAS".





Core index = 1.460

Cladding index = 1.401

Outside index = 1.343

Cladding ellipticity = 0.7616

Core ellipticity = 0.3800

Fig. 13 : Elliptical waveguide discretisation.

46 elements, 119 nodes.

γ	S (KBC = 2)	S (KBC = 3)
1.353	2.264655	2.226960
1.358	2.530127	2.485246
1.363	2.783134	2.731492
1.368	3.064107	3.011099
1.373	3.297490	3.230801
1.378	3.579265	3.504433
1.383	3.886182	3.802208
1.388	4.222381	4.128036
1.393	4.603401	4.497282
1.398	5.019148	4.900267
1.403	5.453403	5.318555
1.413	6.691547	6.519348
1.418	7.47180	7.27709
1.423	8.40982	8.18948
1.428	9.5630	9.31616
1.433	11.0155	10.7412
1.438	12.9433	12.6428
1.443	15.5582	15.2284
1.448	19.6661	19.3089

TABLE 11

The Elliptical fiber propagation characteristics for the two quasi-degenerate dominant modes.
See TABLE 6 for the definition of "KBC".

V(±5%) *2	$\frac{\Delta\gamma \text{ (ref)}}{(n1-n2)}$ *1	$\frac{\Delta\gamma \text{ (FEM)}}{(n1-n2)}$	% error	% *4 error coefficient
0.279	0.0819	-	-	-
0.382	0.255	0.241	-5.5	-115
0.498	0.365	0.351*3	-3.8	-100
0.616	0.398	0.386	-3.0	- 90
0.747	0.398	0.396	-0.5	- 85
0.900	0.380	0.362	-4.7	- 80
0.993	0.365	0.358	-1.9	- 80
1.105	0.346	0.338	-2.3	- 75
1.195	0.331	0.325	-1.8	- 75
1.274	0.318	0.316	-0.6	- 75

*1 From graphic estimation, ref [21]

*2 With : $V = 0.1561 \times S$, $n1 = 1.46$, $n2 = 1.401$, $n3 = 1.343$

*3 From the average of points close to $V = 0.499$

*4 See section 4.3.3 : The error coefficient is defined as the maximum relative error on $\Delta\gamma$ in % for 1% error on the corresponding eigenvalue.

TABLE 12 : $\Delta\gamma$ characteristics of the elliptical fiber

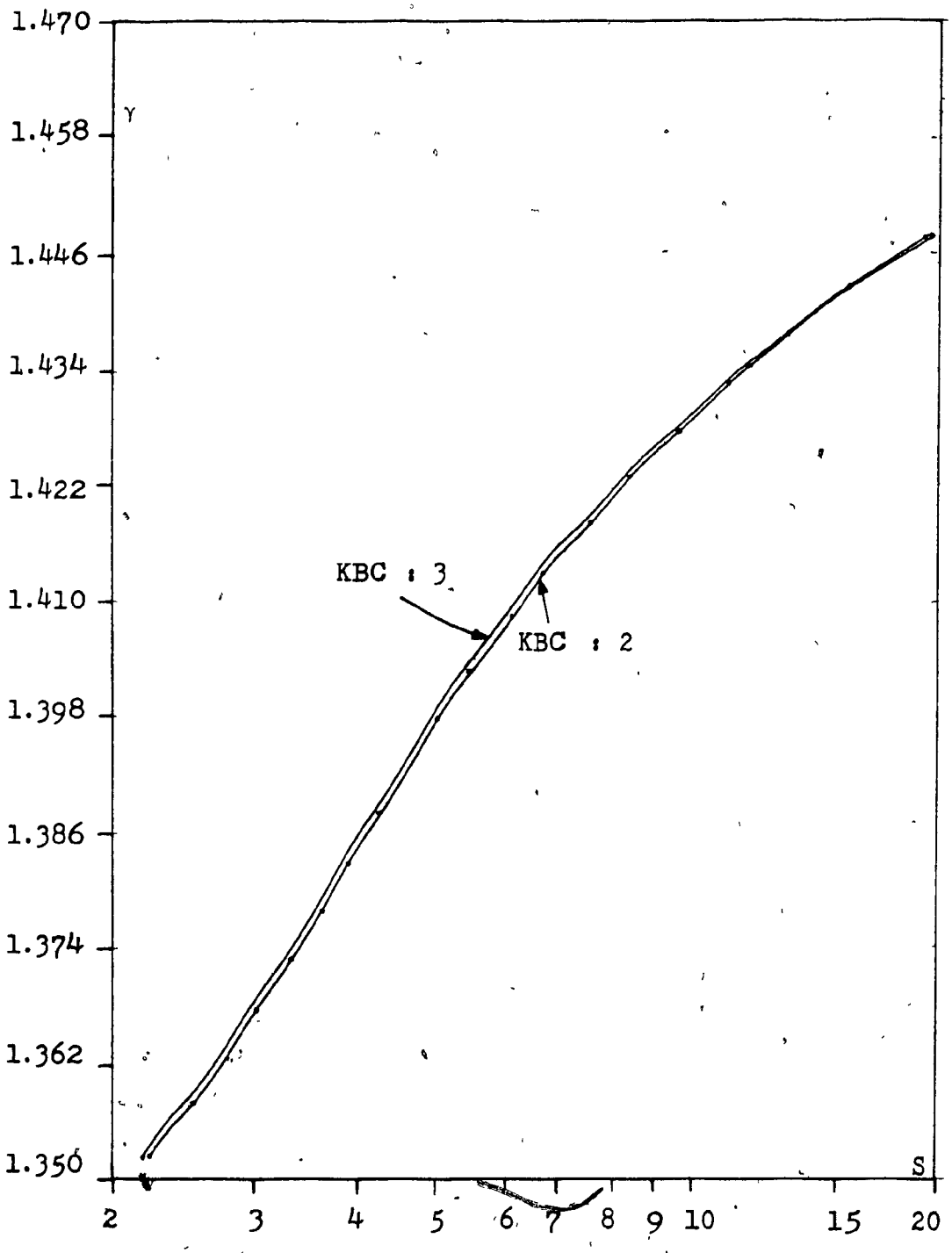


Fig. 14 : Elliptical waveguide example.

Propagation characteristics of the two fundamental modes.

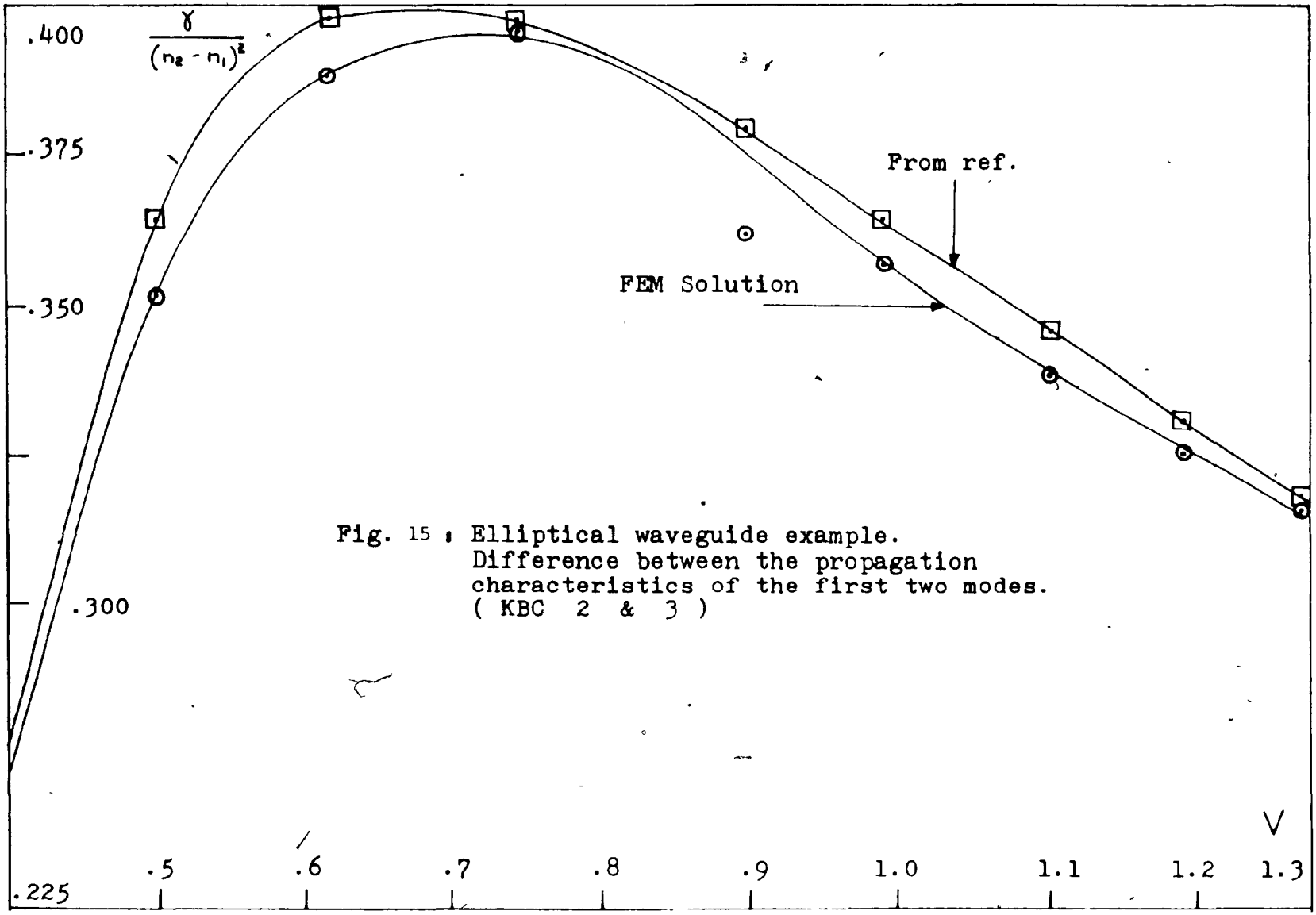


Fig. 15 : Elliptical waveguide example.
Difference between the propagation
characteristics of the first two modes.
(KBC 2 & 3)

4.4 The Finite Cladding Fiber and the cutoff problem

4.4.1 Description

This example have been chosen to show one limit of the formulation given in detail in section 3. Equation (3.7) becomes undefined when goes to infinity, when ε1/ε0 = 1 (This point is defined as the cladding cutoff). As it is shown in this example, values of close to the relative permittivity of one of the dielectric materials lead to inaccurate or erroneous results. This behavior may be explained with a slightly different formulation of the variational expression [ref. 10], which can be written :

$$\sum_{i=1}^{N} \int_{S_i} \left[\nabla_t \psi_i \cdot \nabla_t \psi_i + \left(\frac{\varepsilon_i}{\varepsilon_0}\right) \psi_i \nabla_t^2 \psi_i - k_0^2 \psi_i^2 + \left(\frac{\varepsilon_i}{\varepsilon_0}\right) \psi_i^2 \right] dS = 0 \qquad (4.18)$$

In the media of permittivity such that ε1/ε0 = 1, the Helmolhtz's equations (2.8) become the static equations :

$$\nabla_t^2 \psi_i = 0 \qquad (4.19)$$

When ε1 goes to infinity, $\nabla_t^2 \psi_i$ and $\nabla_t \psi_i$ go to zero and the products occuring in (4.18) are undefined.

One solution to eliminate this discontinuity is to use another integral formulation with the transverse components (see Appendix 5), but these are discontinuous. In the special case of the finite cladding fiber, the solution may be found using differential methods, replacing the usual Bessel functions by the solutions of the Laplace equation in the regions where ε1/ε0 = 1. Such a solution for the z-components are :

$$E_z = (Ar^n + Br^{-n}) \cos (n\phi) \qquad (4.20)$$

$$H_z = (Cr^n + Dr^{-n}) \sin (n\phi) \qquad (4.21)$$

Where A, B,C and D are constants, r is the radius, φ the angle and n the mode number.

4.4.2 Example

The fiber is made with three materials having respectively a relative index of refraction of 1.53, 1.50 and 1.0 (air). The cladding to core dimension ratio is 5. It has been discretised with 25 radial divisions and 5 angular divisions as it is shown on Fig. 16. Only a few points of the fundamental mode close to the virtual value of $\gamma = 1.50$ have been studied. Numerical values have been checked with the results given in reference [22]. A noticeable error is detected when the absolute value of $(\gamma - \sqrt{\varepsilon_1 / \varepsilon_0})$ is smaller than 0.002, and the results when γ is inside this interval are not valid.
Details are given on Table 13 and Fig. 18.
Fig. 17 represents the Hz field for $\gamma = 1.45$, far from the cutoff.

γ	$S(HE_{11})$
1.5100	29.1642
1.5050	22.5605
1.5020	18.9013
1.5010	17.6086
1.5005	16.6580
1.5002	14.9684
1.4998	18.8415
1.4995	17.3660
1.4990	16.5441
1.4980	15.5109
1.4950	13.3127
1.4900	10.7572

TABLE 13

Near cutoff propagation characteristics of a finite-cladding fiber

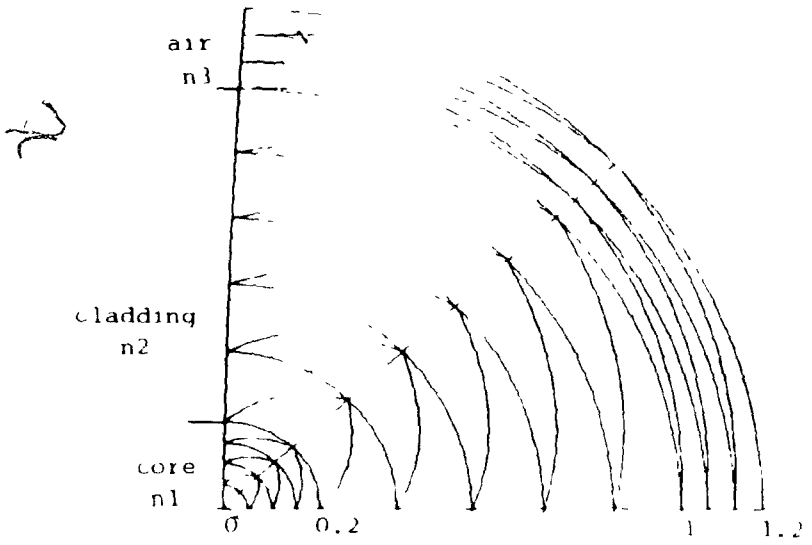


Fig. 16 - Finite cladding fiber, discretisation.

Core index n1 = 1.530

Cladding index n2 = 1.500

Outside index n3 = 1.000

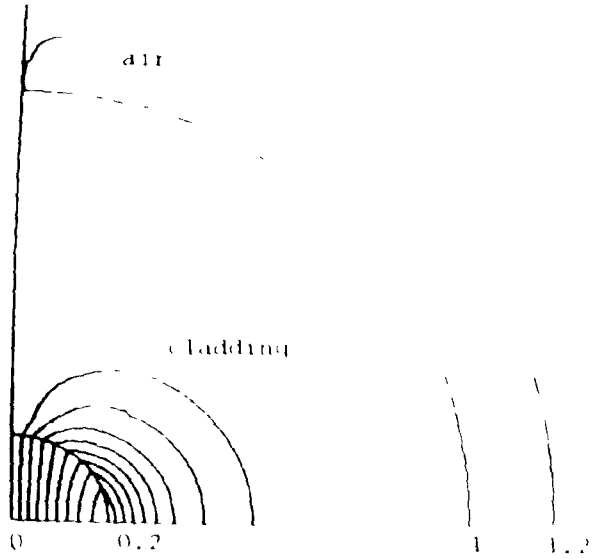


Fig. 17 Finite cladding fiber, Hz field, linear scale, $\gamma = 1.450$ (HE_{11})

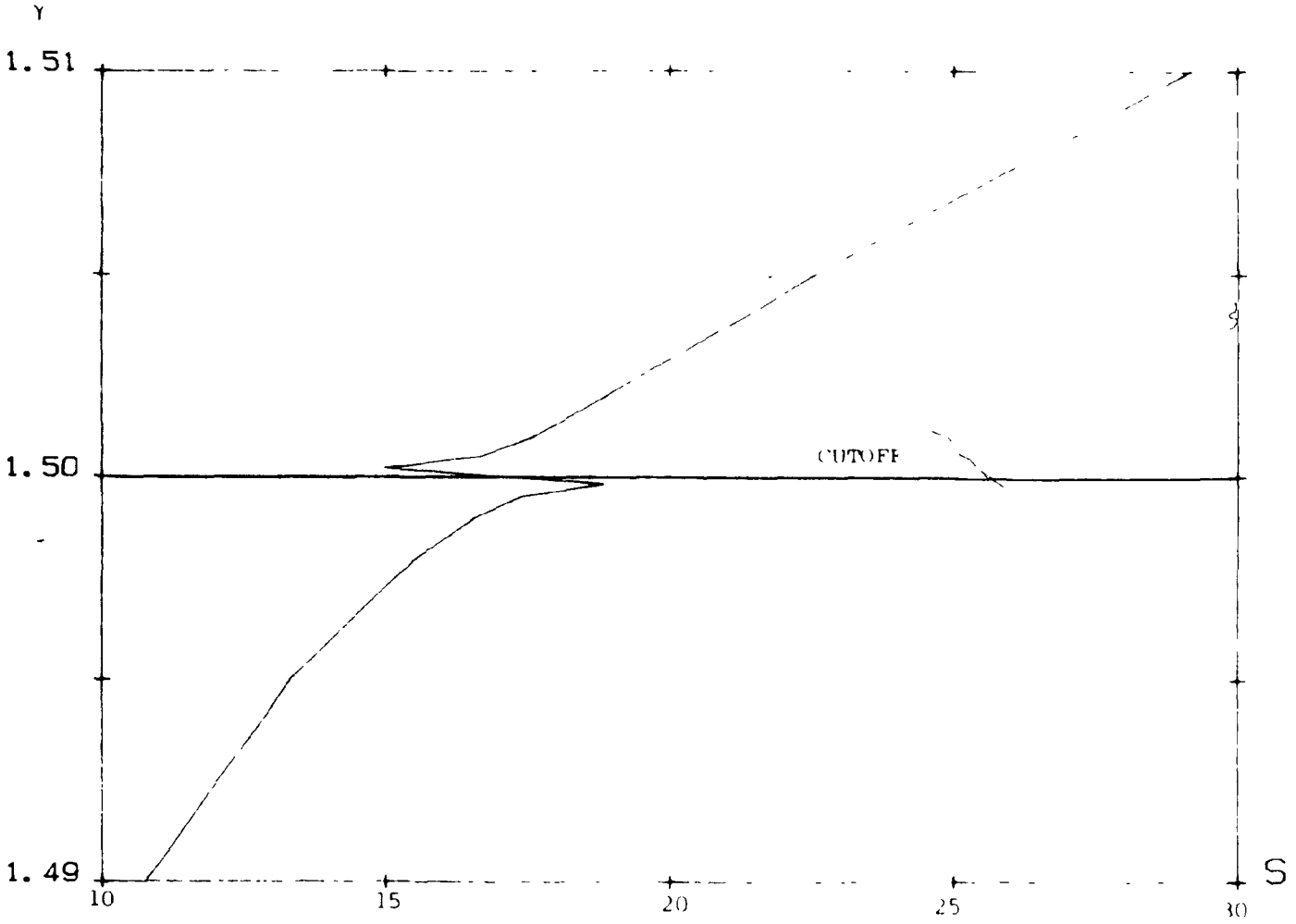


Fig. 18 : Finite cladding optical fiber. Propagation characteristics of the fundamental mode near the cladding cutoff.

4.5 Spurious Mode Detection and Analysis

A spurious mode is an eigenvalue-eigenvector pair of the numerical scheme which does not represent a real propagating mode in the waveguide . It has been observed by many authors working with Finite Elements in that field [ref. 14, 16 and 23]. Its detection is obviously easy when the exact answer is known from another source, but may be difficult in a real unknown problem, when many modes have propagation characteristics close to each other.

As we show in Fig. 19 to 20, an efficient way to solve this detection problem, which gives also a better understanding of the phenomenon, is to plot the equipotential lines of suspected modes.

A spurious mode has three main characteristics :

- The equipotential lines are not as smooth as they are for real modes. Sharp contours are usually absent from the real low order mode plots.
- The effect of the discretisation is important. Many sharp edges in the contours correspond to a transition between elements with the same permittivity. These are usually absent in the regions of constant permittivity of real modes.
- Some additional periodicity appears, indicating an approximation to a highly distorted high order mode.

From the analysis of many contours, it seems that the so called spurious modes are usually very bad approximation of high order modes which have extremely inaccurate propagation constants.

Finally, note that the true modes are relatively insensitive to the virtual boundary location ; this is not so for the spurious modes, as it is shown in table 5. This characteristic may be also used for detection, but the graphic solution appears to be a faster and more reliable way.

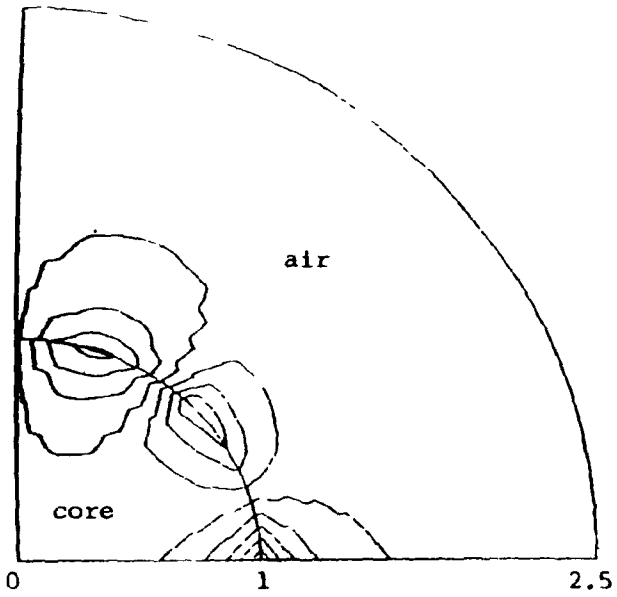


Fig. 19 : Dielectric rod example, first mode (spurious), linear scale, $\gamma = 1.450$

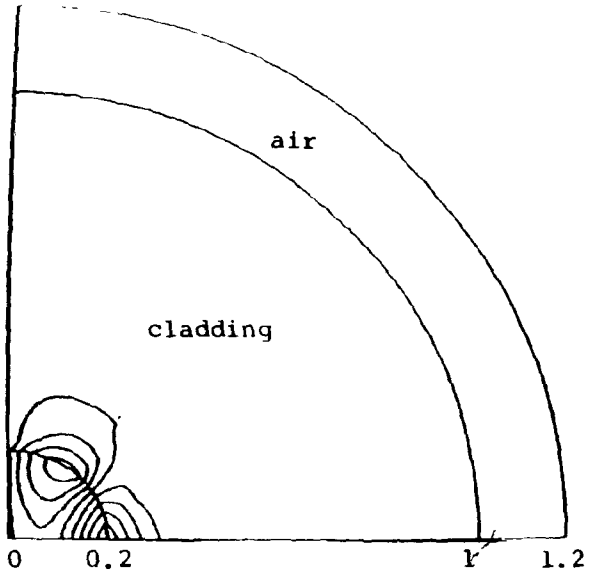


Fig. 20 : Finite cladding example, higher order mode, linear scale, $\gamma = 1.510$, boundary conditions code = 2.(HE_{31})

5. ANALYSIS OF A BICONICAL-TAPERED-MONOMODE OPTICAL COUPLER

5.1 Description

It has been experimentally proved that two monomode fibers fused and stretched may act as a directional coupler if the dimensions of the cross section of the tapered region are small enough (see ref. [24]). An asymptotic analysis using Snyder's coupled-mode theory [25] has been carried out. The physical interpretation of the coupling effect is relatively simple : in the reduced fibers, the electromagnetic field can escape from the core and is spread in the common cladding. The effect of the core becomes secondary and the cladding becomes the major guiding structure in which coupling is allowed.

A more accurate method of analysis is to consider the complete coupler region as a guiding medium of complex shape and to compute its propagation characteristics with the FEM.

A longitudinal view and a sample of the discretised cross section of the coupler are shown in Fig. 22 and Fig. 23.

5.2 Mode Superposition and Coupling Effect in 2-D

5.2.1 Introduction

As we have seen previously, in any structure with two axes of symmetry, we need only model one quarter of the structure if we use four sets of boundary conditons.

In that case, we analyse only one quarter of the x-y plane. The four lowest modes of the structure turn out to be the lowest modes of each type of symmetry. When the fibers which constitute the coupler are separated by a large distance, these four lowest modes are degenerated and have the same propagation characteristics, even if their field distributions are not similar.

If we use the definition given in table 6 in accordance with the "KBC" code, we obtain the following relations :

For two separate fibers :

- Modes 2 and 3 have the same propagation characteristics and field distributions identical with a 90 degrees rotation. Therefore, they correspond to two different polarizations.
- Modes 1 and 3, or 2 and 4 are such that their sum gives the fundamental (and degenerate) mode of one isolated fiber, and their difference, the fundamental mode of the other . Fig 21 illustrates this behavior with simplified equipotential contours . It is the consequence of the fiber separation and the fast decaying rate of the field in the cladding between them.

When the two fibers are brought close to each other, the degeneracy slowly disappears and each mode becomes different from the other with distinct propagation characteristics (see Fig. 26).
This behavior is the key to understanding the coupling effect observed in the laboratory. Even in a single mode coupler (where each fiber carries a single mode), the fundamental mode which carries all the power is split by the loss of symmetry and a modal superposition occurs.

5.2.2 Theoretical Model

In this section, we consider a structure uniform in the z-direction (i.e. ignoring the coupler taper). If we consider $\vec{E}$ (or $\vec{H}$), the total field, we can write :

$$\vec{E} = \vec{E}_1 e^{-j\beta_1 z} + \vec{E}_2 e^{-j\beta_2 z} + \vec{E}_3 e^{-j\beta_2 z} + \vec{E}_4 e^{-j\beta_4 z} \qquad (5.1)$$

Where :

$\vec{E}_1$ (x,y) is the field of one of the quasi-degenerate modes (i=1,2,3 or 4) over the cross section of the coupler for a wave propagating in the positive direction.

β_1 is the corresponding propagation constant.

A similar equation may be written for $\vec{H}$.

In the experimental set up, only the power P_S delivered to a detector

having a finite receiving surface S can be measured. Therefore, it is necessary to compute this quantity in terms of the longitudinal components.

The power flow is obtained from the Poynting vector :

$$\vec{S} = \vec{E} \times \vec{H}^{*} \tag{5.2}$$

$\vec{E}$ and $\vec{H}$ represent the total complex electromagnetic field vectors. From equations (2.12) to (2.15), we know that the transverse field components are in phase quadrature with the longitudinal z-components. Therefore, the real power flow propagating along the z-axis is :

$$\mathrm{Re}\ (S_z) = \mathrm{Re}\ (\vec{E}_t \times \vec{H}_t^{\,*}).\vec{e}_z \tag{5.3}$$

where $\vec{E}_t$ and $\vec{H}_t$ are transverse (in the x-y plane), and $\vec{e}_z$ is the unit vector in the direction of propagation.

If we assume that the total field is the sum of the 4 lower modal fields obtained from the four sets of boundary conditions, we have :

$$\vec{E}_t = \sum_{i=1}^{4} \vec{E}_t^{\,i} \; e^{-j\beta_i z} \tag{5.4}$$

$$\vec{H}_t = \sum_{i=1}^{4} \vec{H}_t^{\,i} \; e^{-j\beta_i z} \tag{5.5}$$

where the β_i's are the propagation constants of each mode.

The time factor $e^{+j\omega t}$ is omitted for simplicity.

$\vec{E}_t^{\,i}$ and $\vec{H}_t^{\,i}$ are in phase and real.

Therefore :

$$\mathrm{Re}\ (S_z) = \sum_{i=1}^{4} \sum_{j=1}^{4} \vec{e}_z.(\vec{E}_t^{\,i} \times \vec{H}_t^{\,j}) \cos(\beta_i - \beta_j) z \tag{5.6}$$

The total power flow over an area S in the x-y plane can be written :

$$P_S = \sum_{i=1}^{4} \sum_{j=1}^{4} P_{ij} \cos(\beta_i - \beta_j) z \tag{5.7}$$

where : $$P_{ij} = \tfrac{1}{2} \int_S \vec{e}_z . (\vec{E}_t^i \times \vec{H}_t^j) \, dS \qquad (5.8)$$

Next, we write P_{ij} in terms of the z-components ψ_i and ϕ_j previously defined and which are computed by the FEM program. From the definitions given in section 3.5, equation (3.51), we can write :

$$\vec{e}_z . (\vec{E}_t^i \times \vec{H}_t^j) = \gamma_i \eta_0 \, \vec{e}_z . (\vec{\psi}_t^i \times \vec{\phi}_t^j) \qquad (5.9)$$

Using equations (3.19) and (3.20), (5.9) becomes :

$$A_i A_j \gamma_i \gamma_j \tau_r^i \tau_r^j \eta_0 \{ (\nabla_t \phi_i . \nabla_t \phi_j) + \gamma_i^2 (\frac{\varepsilon_r}{\varepsilon_0}) (\nabla_t \psi_i . \nabla_t \psi_j) + \gamma_i^2 \vec{e}_z . (\nabla_t \psi_i \times \nabla_t \phi_j) + (\frac{\varepsilon_r}{\varepsilon_0}) (\nabla_t \psi_j \times \nabla_t \phi_i) . \vec{e}_z \} \qquad (5.10)$$

Subscripts i and j refer to different modes. Subscript r refers to elements. Therefore :

γ_i and γ_j are the propagation constants of modes i and j.

τ_r^i and τ_r^j are the corresponding constants defined in equation (3.5), but with the new subscripts.

ε_r is the permittivity in element "r".

A_i and A_j are the corresponding constants given in equation (3.21).

This transformation is possible only because the gradients are in the x-y plane. The vectorial manipulations are done with the standard results given in reference (17), page 118.

If we use the following identities :

$$\nabla . (f \nabla g) = (\nabla f . \nabla g) + f \nabla^2 g \qquad (5.11)$$

$$\nabla \times (f \nabla g) = \nabla f \times \nabla g \qquad (5.12)$$

we obtain :

$$\int_S (\nabla f . \nabla g) \, dS = \tfrac{1}{2} \int_S -(f \nabla^2 g + g \nabla^2 f) + \nabla . (f \nabla g + g \nabla f) \, dS \qquad (5.13)$$

$$\int_S (\nabla f \times \nabla g) . d\vec{S} = \tfrac{1}{2} \int_S \nabla \times (f \nabla g - g \nabla f) . d\vec{S} \qquad (5.14)$$

The Divergence and the Stokes's theorems can be used only if the gradients are continuous in each area of integration. Therefore, we write P_{ij} in terms of the sum of the power flows going through each element. We obtain :

$$P_{ij} \sim \sum_r \int_{S_r} \{ (\nabla_t \phi_i^r \cdot \nabla_t \phi_j^r) + \gamma_i^2 (\frac{\varepsilon_r}{\varepsilon_0}) (\nabla_t \psi_i^r \cdot \nabla_t \psi_j^r) + \gamma_i^2 \vec{e}_z . (\nabla_t \psi_i^r \times \nabla_t \phi_j^r) + (\frac{\varepsilon_r}{\varepsilon_0}) (\nabla_t \psi_j^r \times \nabla_t \phi_i^r) . \vec{e}_z \} dS_r \quad (5.15)$$

where : $\sum_r S_r = S$. ("$\sim$" = proportional to)

Using equations (5.13), (5.14) and the Helmholtz equation (2.8), we obtain :

$$P_{ij} \sim \sum_r \int_{S_r} - k_o^2 (\frac{1}{\tau_r^j} + \frac{1}{\tau_r^i}) \{ \phi_i^r . \phi_j^r + \gamma_i^2 (\frac{\varepsilon_r}{\varepsilon_0}) \psi_j^r \psi_i^r \} \, dS_r \quad (5.16)$$

$$+ \int_{C_r} \{ \phi_i^r \nabla_t \phi_j^r + \phi_j^r \nabla_t \phi_i^r + (\gamma_i^2 \frac{\varepsilon_r}{\varepsilon_0}) (\psi_i^r \nabla_t \psi_j^r + \psi_j^r \nabla_t \psi_i^r) \} . \vec{n} \, dl_r$$

$$+ \int_{C_r} \{ \gamma_i^2 (\psi_i^r \nabla_t \phi_j^r - \phi_j^r \nabla_t \psi_i^r) + (\frac{\varepsilon_r}{\varepsilon_0}) (\psi_j^r \nabla_t \phi_i^r - \phi_i^r \nabla_t \psi_j^r) \} . \vec{l} \, dl_r$$

The line integral can be written : (5.17)

$$\sum_r \int_{C_r} \phi_i^r \{ \nabla_t \phi_j^r . \vec{n} - (\frac{\varepsilon_r}{\varepsilon_0}) \nabla_t \psi_j^r . \vec{l} \} + \phi_j^r \{ \nabla_t \phi_i^r . \vec{n} - \gamma_i^2 \nabla_t \psi_i^r . \vec{l} \}$$

$$+ \gamma_i^2 \psi_i^r \{ (\frac{\varepsilon_r}{\varepsilon_0}) \nabla_t \psi_j^r . \vec{n} + \nabla_t \phi_j^r . \vec{l} \} + (\frac{\varepsilon_r}{\varepsilon_0}) \psi_j^r \{ \gamma_i^2 \nabla_t \psi_i^r . \vec{n} + \nabla_t \phi_i^r . \vec{l} \} dl_r$$

Using the following relations :

$(\vec{e}_z \times \vec{f}) . \vec{n} = \vec{e}_z . (\vec{f} \times \vec{n})$ for any function $\vec{f}$ in x-y plane;

$\vec{l} = \vec{e}_z \times \vec{n}$ and :

$\vec{f} . \vec{l} = \vec{f} . (\vec{e}_z \times \vec{n}) = - (\vec{e}_z \times \vec{f}) . \vec{n} = - \vec{e}_z . (\vec{f} \times \vec{n})$ when $\vec{f}$ is in x-y plane, (5.17) can be written in terms of the transverse field components computed in section 3.1.2 (equations 3.24 and 3.27) and in appendix 1.

We obtain : (5.18)

$$\sum_r \int_{C_r} A_1 \phi_1^r (\vec{\phi}_t^{jr} . \vec{n}) + \{A_2 \phi_j^r (\vec{\psi}_t^{1r} \times \vec{n}) + A_3 \gamma_i^2 \psi_1^r (\vec{\phi}_t^{jr} \times \vec{n})\} . \vec{e}_z + A_4 \left(\frac{\varepsilon_r}{\varepsilon_0}\right) \psi_j^r (\vec{\psi}_t^{1r} . \vec{n}) \; dl_r$$

where A_1, A_2, A_3 and A_4 are independent of "r" and do not need to be evaluated.

The continuity of the fields (see section 3.1 and appendix 1) leads to a concellation of the parts of the integral (5.18) along interelement boundaries.

The contribution of the virtual boundary is also zero because all the field components vanish there.

Finally, the only path on which the field must be analysed is the axis of symmetry.

On that axis, we always have one of the following boundary conditions :

$\phi_1 = 0$	or $\psi_1 = 0$ (for a mode "1")
Then : $\nabla_t \phi_1 \times \vec{n} = 0$	or $\nabla_t \psi_1 \times \vec{n} = 0$
and : $\nabla_t \psi_i . \vec{n} = 0$	or $\nabla_t \phi_1 . \vec{n} = 0$
$\vec{\psi}_t^{i} . \vec{n} = 0$	or $\vec{\phi}_t^{1} . \vec{n} = 0$
$\vec{\phi}_t^{1} \times \vec{n} = 0$	or $\vec{\psi}_t^{i} \times \vec{n} = 0$

"j" can replace "1" in the above relations.

P_{ij} becomes : (5.19)

$$P_{ij} \simeq - k_o^2 \left(\frac{1}{\tau_i} + \frac{1}{\tau_j} \right) \sum_r \int_{S_r} \phi_i^r \phi_j^r + \gamma_1^2 \left(\frac{\varepsilon_r}{\varepsilon_0}\right) \psi_j^r \psi_i^r \; dS_r$$

$$+ \int_{\text{Axes of symmetry}} A_1 \phi_1 (\vec{\phi}_t^{j} . \vec{n}) + A_2 \phi_j \vec{e}_z . (\vec{\psi}_t^{1} \times \vec{n}) + A_3 \gamma_i^2 \psi_1 \vec{e}_z . (\vec{\phi}_t^{j} \times \vec{n}) + A_4 \left(\frac{\varepsilon_r}{\varepsilon_0}\right) \psi_j (\vec{\psi}_t^{i} . \vec{n}) \; dl$$

According to the definition of the modes given in table 6, we also have the following boundary conditions with respect to the x-axis :

ϕ_i , $\vec{\psi}_t^i \cdot \vec{n}$, $\vec{e}_z.(\vec{\phi}_t^i \times \vec{n})$ are antisymmetric when i = 1 or 3.

ψ_i , $\vec{\phi}_t^i \cdot \vec{n}$, $\vec{e}_z.(\vec{\psi}_t^i \times \vec{n})$ are antisymmetric when i = 1 or 3.

When i = 2 or 4, these quantities are symmetric.

When a detector is put at the output of one arm of the coupler, the measured power flow is the integrated power over one half plane (defined by the Oy-axis with our convention).

Taking S to be this half plane (x > 0), the surface integral in (5.19) has the following properties :

- It vanishes when i and j are such that :

  (i,j) = (1,2), (2,1), (1,4), (4,1), (2,3), (3,2), (3,4) or (4,3)

This is due to the difference of symmetry between the modes i and j with respect to the Ox-axis.

- It does not vanish when i = j or when :

  (i,j) = (1,3), (3,1), (2,4) or (4,2)

The line integral in (5.19) has the following properties :

- It vanishes when $\phi_i = \phi_j = 0$ (or $\psi_i = \psi_j = 0$) on the Oy-axis,

  e.g. when i = j or when (i,j) = (1,4), (4,1), (2,3) or (3,2).

- It vanishes when (i,j) = (1,4), (4,1), (2,3) or (3,2) because of the difference of symmetry between the modes i and j with respect to the Ox-axis.

- It does not vanish when (i,j) = (1,3), (3,1), (2,4) or (4,2).

Finally, the only non zero terms are :

P_{11} , P_{22} , P_{33} , P_{44} , P_{13} , P_{31} , P_{24} and P_{42}

and equation (5.7) becomes :

$$P_S = P_{11} + P_{22} + P_{33} + P_{44} + (P_{13} + P_{31})\cos(\beta_1-\beta_3)z + (P_{24} + P_{42})\cos(\beta_2-\beta_4)z \qquad (5.20)$$

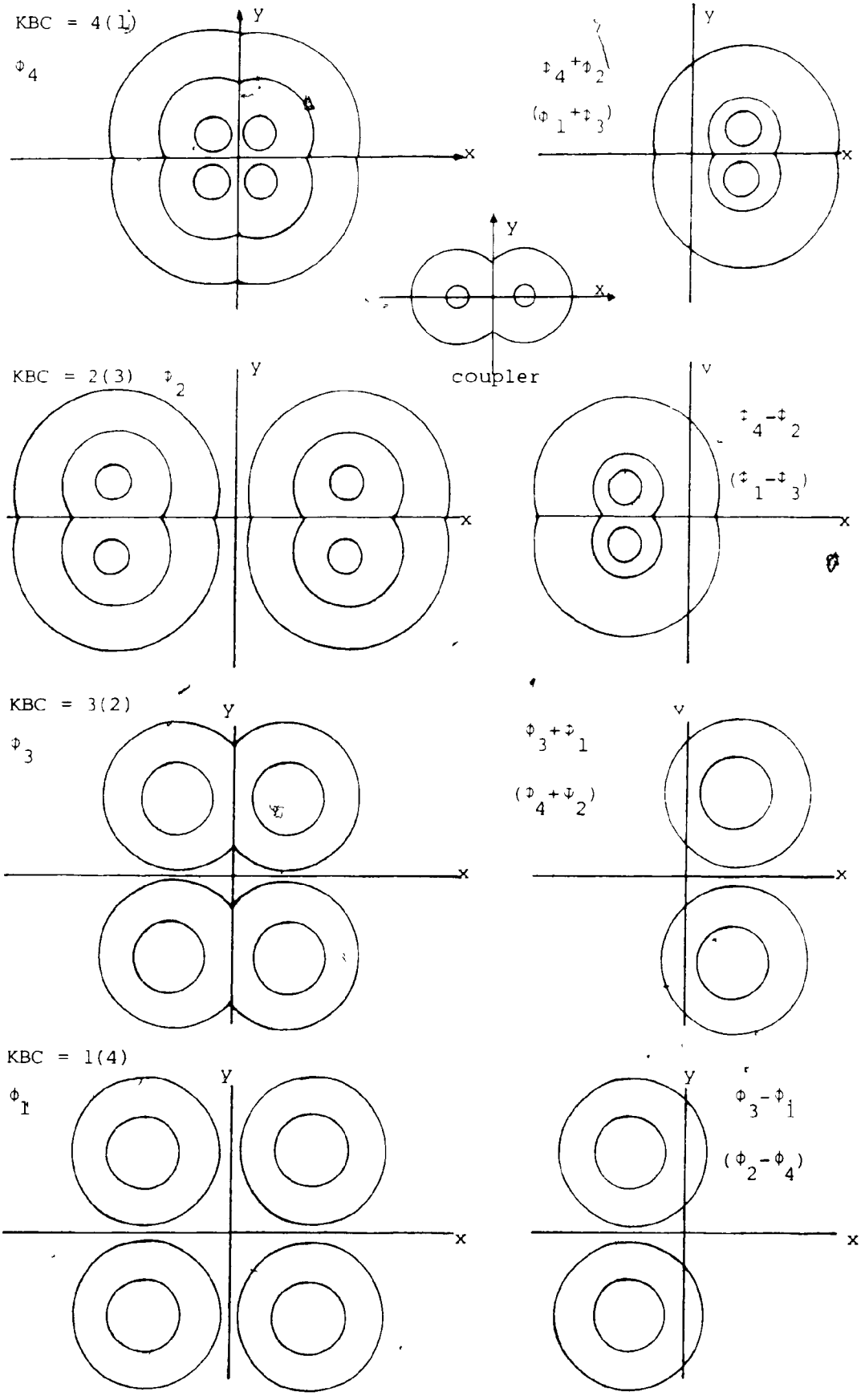


Fig. 21 : Symmetry of modes and superposition.

5.3 Modes Superposition and Coupling Effect in 3-D

5.3.1 A Three Dimensional Approximate Analysis

As shown on Fig. 22, the coupler is truly a 3-D structure, but with a very slow variation of its shape along the longitudinal z-direction. According to the experimental results (reference [24]), we can assume that the taper does not create any appreciable reflection to the source, and the isolation is very good from the adjacent input.

Now :

$$\vec{E} = \tilde{E}(x,y)\, e^{-j\beta z} \tag{5.21A}$$

$$\vec{H} = \tilde{H}(x,y)\, e^{-j\beta z} \tag{5.21B}$$

is a solution to :

$$\nabla \times \vec{H} = j\omega\varepsilon(x,y)\, \vec{E} \tag{5.22A}$$

$$\nabla \times \vec{E} = -j\omega\mu(x,y)\, \vec{H} \tag{5.22B}$$

then, equivalently :

$\tilde{E}(x,y)$, $\tilde{H}(x,y)$, β is a solution to :

$$\nabla' \times \tilde{H} = j\,\omega\varepsilon(x,y)\, \tilde{E} \tag{5.23A}$$

$$\nabla' \times \tilde{E} = -j\,\omega\mu(x,y)\, \tilde{H} \tag{5.23B}$$

where : $\nabla' = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, -j\beta \right)$ (5.24)

Now, suppose ε and μ vary with z.

Let :

$\tilde{E}(x,y,z)$, $\tilde{H}(x,y,z)$, $\beta(z)$ be a solution to (5.23) at a given z, i.e. a solution to :

$$\nabla' \times \tilde{H} = j\,\omega\varepsilon(x,y,z)\, \tilde{E} \tag{5.25A}$$

$$\nabla' \times \tilde{E} = -j\,\omega\mu(x,y,z)\, \tilde{H} \tag{5.25B}$$

$\tilde{\vec{E}}(x,y,z)$ and $\tilde{\vec{H}}(x,y,z)$ are exactly the modal fields computed for a uniform waveguide with cross section defined by $\varepsilon(x,y,z)$, $\mu(x,y,z)$. Then, in the tapered guide, an approximate solution to Maxwell's equations is :

$$\vec{E}(x,y,z) = \tilde{\vec{E}}(x,y,z)\, e^{-j\phi(z)} \qquad (5.26A)$$

$$\vec{H}(x,y,z) = \tilde{\vec{H}}(x,y,z)\, e^{-j\phi(z)} \qquad (5.26B)$$

where :

$$\phi(z) = \int_{z_o}^{z} \beta(\xi)\, d\xi \qquad (5.27)$$

(z_o corresponds to a phase reference at the end of the coupler, where the z-variation of the cross section disappears.)

... providing the following is true :

$$\left| e^{-j\phi(z)} \frac{\partial \tilde{E}_x}{\partial z} \right| \ll \left| \tilde{E}_x \frac{\partial}{\partial z} e^{-j\phi(z)} \right| \qquad (5.28)$$

at each point (x,y,z).

This relation comes from the differentiation with respect to z of $\tilde{E}_x$. Similar relations must hold for $\tilde{E}_y$, $\tilde{H}_x$ and $\tilde{H}_y$.

But :

$$\frac{\partial}{\partial z} e^{-j\phi(z)} = -j\, e^{-j\phi(z)} \frac{\partial}{\partial z} \phi(z) \qquad (5.29)$$

$$= -j\, e^{-j\phi(z)} \frac{\partial}{\partial z} \int_{z_o}^{z} \beta(\xi)\, d\xi \qquad (5.30)$$

$$= -j\, e^{-j\phi(z)}\, \beta(z) \qquad (5.31)$$

and (5.28) becomes :

$$\left| \frac{\partial \tilde{E}_x}{\partial z} \right| \ll \left| \beta(z)\, \tilde{E}_x \right| \qquad (5.32)$$

If we call λ the free space wavelength of the beam propagating inside the coupler and $\Delta\tilde{E}_x$ the change in $\tilde{E}_x$ at (x,y) in a distance Δz, equation (5.32) can be written

$$\frac{\Delta\tilde{E}_x}{\tilde{E}_x} \ll \frac{2\pi\gamma}{\lambda}\,\Delta z \qquad (5.33)$$

and similarly for $\tilde{E}_y$, $\tilde{H}_x$ and $\tilde{H}_y$.

This last inequality gives the approximate conditions which validate equations (5.26A) and (5.26B).

In the example 1 which follows in section 5.4.1, we have :

$$\lambda \simeq 0.63\ \mu$$

and the coupling occurs when

$$\gamma \simeq 1.46$$

With the geometric model developed in section 5.3.4, we can also evaluate the maximum change in radius R of the coupler with respect to z. Using equations (5.40) to (5.42) and the parameters given in section 5.4 and Fig. 27, we obtain a maximum rate $|\Delta R / \Delta z|$ of about 30 μ per mm with a typical coupler length of 5mm.

Therefore, if Δz is a guide wavelength $\lambda_g = \frac{\lambda}{\gamma} \simeq \frac{1}{2}$ micron, the change in radius of the coupler is only 0.015μ in the worst case.

With the same model and the same parameters, we can also compute the maximum value of $\Delta R / R$ for the same Δz. We obtain :

$$\left|\frac{\Delta R}{R}\right|_{max} \simeq 1.1\,\Delta z \simeq 0.055\ \% \quad \text{when} \quad \Delta z = 0.5\mu\ .$$

Even if these numbers are only an approximation, they are small enough to validate equation (5.33).

Therefore, the total field at each point (x,y,z) can be written :

(for $\vec{E}$ field; $\vec{H}$ is similar)

$$\vec{E}(x,y,z) = \sum_i \vec{E}_i(x,y,z)\, e^{-j\int_0^z \beta_i(\zeta)\,d\zeta} \qquad (5.34)$$

and the power becomes:

$$P(x,y,z) = \sum_i \sum_j P_{ij}(x,y,z)\, \cos\left[\int_0^z \beta_i(\zeta)\,d\zeta - \int_0^z \beta_j(\zeta)\,d\zeta\right] \qquad (5.35)$$

Even if the field distribution $\vec{E}$ varies much with z as shown on the contour plots Fig. 29 to 33, it is reasonable to assume that the power functions are not very affected by the z-dependence in the coupling section and $P_{ij}(x,y,z) \simeq P_{ij}(x,y)$.
Therefore, all the mathematical formulation of the previous section is still valid to a first approximation by replacing all the phase factors "βz" with the proper integral :

$$\int_{\zeta=-\infty}^{\zeta=z} \beta(\zeta)\,d\zeta \qquad (5.36)$$

In that model, β is a function of the longitudinal dimension. The lower bound $\zeta = -\infty$ corresponds to the input of the coupler. But in practice, it will not be necessary to know β over such a wide interval because only the difference between the β's is needed. This difference is zero when the modes are degenerate and this is true along the longest part of the coupler.

5.3.2 Simplifying assumptions

At this point, several assumptions may be made concerning the relative power carried by each mode.

- Assumption 1

The power carried by each mode of each polarization is equal : $P_{11} = P_{33}$ and $P_{22} = P_{44}$ over the whole x-y plane. Because of the symmetry relations, this is also verified for each half plane. This is necessarily true at the input of the coupler because at that level, all the power is in one branch only ($\phi_1 + \phi_3$ is zero in one half plane, $\phi_1 - \phi_3$ is zero in the other ; the same situation is repeated for ϕ_2 and ϕ_4 ; see Fig. 21). It is not proved at the present time

that this situation is kept for any z and especially in the coupling area of the coupler.

With this assumption, equation (5.20) becomes :

(5.37)

$$P_S = 2P_{11} + (P_{13}+P_{31})\cos\int(\beta_1-\beta_3)\,d\xi + 2P_{22} + (P_{24}+P_{42})\cos\int(\beta_2-\beta_4)\,d\xi$$

- Assumption 2

When the coupling begins, the cross coupling terms $P_{13} + P_{31}$ ($P_{24} + P_{42}$) is almost equal to the initial power $2P_{11}$ ($2P_{22}$) over one half plane. This allows 100% coupling for both polarizations.

Justifications: The amplitude distribution of the field for the boundary conditions 1 and 3 (2 and 4) does not differ by much except near the y-axis (see Fig. 28, 29 32 and 33).

Therefore, the product $\psi_1\psi_3$ ($\psi_2\psi_4$, $\psi_3\psi_1$, $\psi_4\psi_2$) is close to the product $\psi_1\psi_3$ ($\psi_2\psi_4$, $\psi_3\psi_1$, $\psi_4\psi_2$) and the integration over one half plane gives almost identical results.

On the y-axis, the contribution of the line integral (5.18) is zero for the self coupling term ($\psi_1\psi_1$) and can be neglected for the cross coupling term when the field is still weak at that place, even if the coupling becomes strong.

P_S becomes :

$$P_S = 2P_{11}\{1 + \cos\int(\beta_1-\beta_3)\,d\xi + 2P_{22}\{1 + \cos\int(\beta_2-\beta_4)\,d\xi \qquad (5.38)$$

- Assumption 3

Finally, we may assume an equal distribution of power in each polarization : $P_{11} = P_{22} = Po/4$ (Po = total power).

This is verified on an averaged basis in perfectly circular and well separated waveguides where the degeneracy is total and the polarization not stable, but it in not proved that this feature is kept for any z in an highly non symmetric device like the coupler.

This last hypothesis leads to the following expression :

(5.39)

$$P_S = Po\{1 + \cos\int\tfrac{1}{2}(\beta_1-\beta_3+\beta_2-\beta_4)\,d\xi \cdot \cos\int\tfrac{1}{2}(\beta_1-\beta_3-\beta_2+\beta_4)\,d\xi\}$$

5.3.3 Coupling Mechanism

The coupling occurs when $\beta_1 \neq \beta_3$ or $\beta_2 \neq \beta_4$. This is achieved when the normalised frequency $V = ka/c$ becomes sufficiently small, when the normalised propagation constant reaches a value close to the refractive index of the cladding. The corresponding value of V is defined as the cladding cutoff. The only variable which can be changed is the scale factor "a", the frequency k being set by the laser source. The scale reduction is done in the tapered coupling section, which must be smooth to avoid losses and reflections.

The relation between ξ and z comes from the relations between b (the normalized propagation constant) and V (the normalized frequency ka/c). The factor "a" is a function of z, because the radius of the coupler changes continuously along z. The mathematical model which describes this variation is given in section 5.3.4 below. The power intercepted by the detector is given by equation (5.20) with the phase factor (5.36) in the general case, and in equations (5.37), (5.38) and (5.39), depending on the validity of the simplifications. Because of the very short length of the coupler, we assume a lossless propagation. This may not be experimentally verified if the shape of the taper is not good and presents some discontinuity.

5.3.4 Coupler Geometry

The geometrical model of the coupler has been developed in reference [24] and is shown on Fig. 22. Fig. 23 shows a sample of a discretised quarter of the cross section. The small elements are in the core. The radial boundaries are perfectly circular (The picture taken on a terminal screen shows some distortion). It is assumed that the shape is simply scaled along the z-axis. Even if this is not true when the fibers are completely separated, this assumption is reasonable for the coupling section.
R_{in} is the initial radius of the coupler and R_0 the radius at the plane of symmetry.

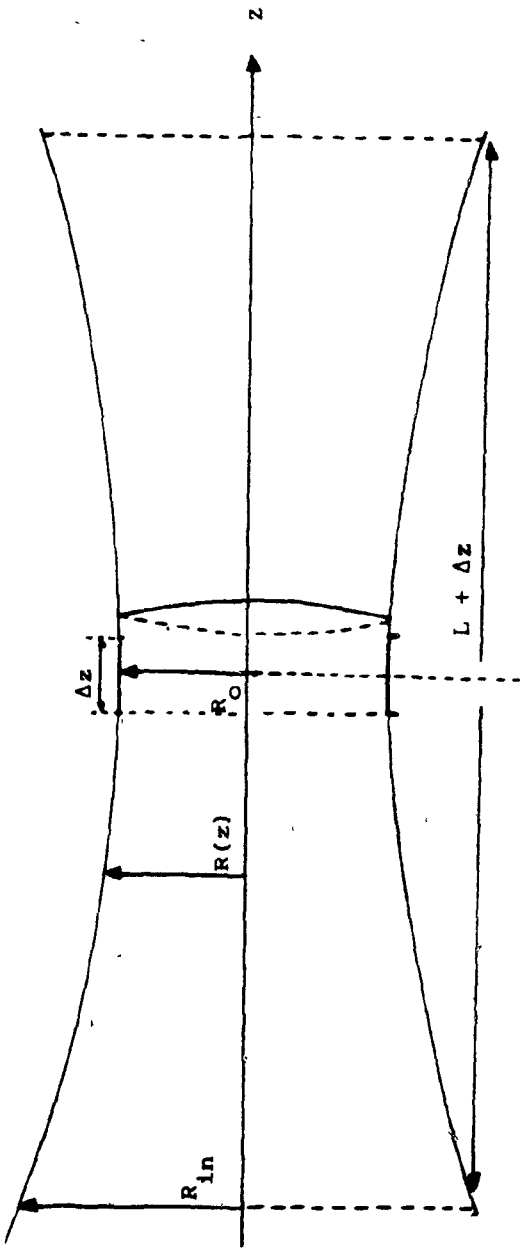


Fig. 22 : Geometry of the coupler.

From reference [24], we can write :

$$R(z) = R_{\cdot} \left(1 + \gamma z^2 \right) \tag{5.40}$$

$$R_{\cdot} = R_{in} e^{-\frac{L}{2\Delta z}} \tag{5.41}$$

$$\gamma = 4 \frac{\left(e^{\frac{L}{2\Delta z}} - 1 \right)}{(L + \Delta z)^2} \tag{5.42}$$

$L + \Delta z$ is the total length of the coupling region.

The choice of the normalising dimension being arbitrary, we have chosen the fiber radius $R(z)$ and we can write :

$$a(z) = R(z) \tag{5.43}$$

The 3 parameters R_{in}, L and Δz define the coupler geometry.

5.3.5 Computation of the coupling coefficient for 2 modes interaction.

Define a power ratio P_{ij}/Po as :

$$P_{ij}/Po = \cos^2 I_{ij}/2 = \tfrac{1}{2} \left(1 + \cos I_{ij} \right) \tag{5.44}$$

$$\text{with : } I_{ij} = \int_{-\frac{1}{2}(L+\Delta z)}^{\frac{1}{2}(L+\Delta z)} \left(\beta_i - \beta_j \right) d\xi \tag{5.45}$$

β_i, β_j are two propagation constants (not normalised) as a function of z.

L and Δz are defined in Fig. 22.

Using the symmetry of the geometry, we can write :

$$I_{ij} = 2 \int_{0}^{\frac{1}{2}(L+\Delta z)} \left(\beta_i - \beta_j \right) d\xi \tag{5.46}$$

$$= 2 \int_{R_o}^{R_{in}} \{ (\beta_i - \beta_j) (R) \} \left(\frac{d\xi}{dR} \right) dR \tag{5.47}$$

Using the model developed in reference [24] and equations (5.40) to (5.42), we can write :

$$I_{ij} = \int_{Ro}^{R_{in}} (\beta_i - \beta_j)(R) \{ \Gamma(R - Ro) Ro \}^{-\frac{1}{2}} dR \qquad (5.48)$$

Using the following transformations : $S = \frac{\omega R}{c}$ and :

$\gamma = \frac{\beta c}{\omega}$, $\beta_i - \beta_j = \frac{2\pi}{\lambda} (\gamma_i - \gamma_j)$, λ = free space wavelength of the beam,

We finally get : $\Delta\gamma = \gamma_i - \gamma_j$,

$$I_{ij} = \frac{2\pi}{\lambda\sqrt{\Gamma S_0}} \int_{S_0}^{S_{in}} \frac{\Delta\gamma(S)}{\sqrt{S - S_0}} dS \qquad (5.49)$$

with $$S_{in} = \frac{\omega R_{in}}{c} = \frac{2\pi}{\lambda} R_{in} \qquad (5.50)$$

$$S_0 = S_{in} e^{-L/2\Delta z} \qquad (5.51)$$

The integration can be done analytically for each γ using piecewise quadratic interpolation for γ_i and γ_j :

Let $$\gamma_i = A_i^k S^2 + B_i^k S + C_i^k \qquad (5.52)$$

$$\gamma_j = A_j^k S^2 + B_j^k S + C_j^k \qquad (5.53)$$

on a given segment "k" .

We can write $$I_{ij} = \frac{2\pi}{\lambda\sqrt{\Gamma S_0}} (I_i - I_j)$$

$$= \frac{2\pi}{\lambda\sqrt{\Gamma S_0}} \{ \int_{S_0}^{S_{in}} \frac{\gamma_i(S)}{\sqrt{S-S_0}} dS - \int_{S_0}^{S_{in}} \frac{\gamma_j(S)}{\sqrt{S-S_0}} dS \} \qquad (5.54)$$

$$I_{i*} = \sum_k \int_{S_k^-}^{S_k^+} \frac{A_i^k S^2 + B_i^k S + C_i^k}{\sqrt{S-S_0}} dS \qquad (5.55)$$

Where "k" is the label of one segment in the $[S_0, S_{in}]$ interval, S_k^- is the lower end of segment "k" and S_k^+ its upper end. Using standard integration formula , we obtain :

(5.56)

$$I_i = \sum_k \left| \left\{ \frac{2A_i^k}{15}(3S^2+4S_0S+8S_0^2) + \frac{2B_i^k}{3}(S+2S_0) + 2C_i^k \right\} \sqrt{S-S_0} \right|_{S=S_k^-}^{S=S_k^+}$$

I_j is similar for the other mode and I_{ij} can be evaluated.
This analytic result avoids other numerical errors in a very long and delicate process.

When two modes are involved, and if assumption 1 holds (see 5.3.2), we can use for example equation (5.37) with P_{22}, P_{24} and $P_{42} = 0$. Therefore :

$$P_S = 2P_{11} + (P_{13} + P_{31}) \cos I_{13} \quad (5.57)$$

If assumption 2 holds :

$$P_S = 2P_{11} (1 + \cos I_{13}) \quad (5.58)$$

$$= 4P_{11} \cos^2 I_{13}/2 \quad (5.59)$$

If we define $Po = 4P_{11}$, then

$$P_S/Po = \cos^2 I_{13}/2 \quad (5.60)$$

5.3.6 Computation of the coupling coefficient for 4 modes interaction.

The same principle using four piecewise quadratic approximations instead of two is used to obtain the results shown in section 5.4 and computed from the equation (5.39 in section 5.3.2). The geometric model and the previously defined transformations are unchanged.
Equation (5.39) can be written :

$$P_S = 2P_{11}(1 + \cos I_{13} + 1 + \cos I_{24}) \quad (5.61)$$

$$= 4P_{11}(\cos^2 I_{13}/2 + \cos^2 I_{24}/2) \quad (5.62)$$

If we define $Po = 4P_{11}$ as before, then :

$$P_S/Po = \cos^2 I_{13}/2 + \cos^2 I_{24}/2 \quad (5.63)$$

5.4 Examples

5.4.1 Example 1

The quarter of the cross section of the coupler is shown in Fig. 23. The core index is 1.506 and the cladding index 1.46. The propagation characteristics have been computed and are shown on table 14 and Fig. 23 to 26. From these results, we obtain the following conclusions :

- The near degeneracy of the modes 1, 2, 3 and 4 is very well observed for values of γ which are above 1.462 (The cladding cutoff is at γ = 1.46).

This behavior is related to a change in the field solution inside the cladding which becomes slowly non-decaying. It is equivalent to a change of the Bessel's functions type in the analysis of a finite cladding optical fiber (ref. [22]).

- The contours (Fig. 28, 31 and 32) associated with these γ-values confirm the degeneracy. In addition, the fundamental mode field concentrated inside the core is very similar to the field of an isolated fiber shown in section 4.1, Fig. 6.
- The end of the degeneracy is very fast when γ falls down below the cladding cutoff and in that case the field fills the cladding along complicated equipotential lines. This corresponds to the coupling effect, when the superposition of non-degenerate modes gives a beat effect.

The ratio of the power coupled into the second arm (P_S) with respect to the incoming power is computed according to the geometric model of the taper described in section 5.3.4, and the assumptions given in section 5.3.2 and is shown on Fig. 27 as a function of the length of the coupling section. The slow variation at the beginning of the curve corresponds to the end of the degeneracy between the modes 1-3 and 2-4. Then, this variation becomes very fast, but is modulated by the relatively small degeneracy between the modes 2 and 3. The modes 1 and 4 are almost completely degenerate for γ as low as 1.43. Some experimental points from reference [24] are shown, even if the data were taken with a fiber having a core index of 1.4687.

Corresponding results from the same reference show the approximate solution from the Coupled Mode theory (see section 5.4.5).

5.4.2 Example 2

This example is similar to the previous one, except for the index of the core which is much smaller (1.4687). This makes the analysis with our method slightly less accurate for γ's above the cladding cutoff.

The propagation characteristics are given on Fig. 34 and table 15. No contour is shown because of the similarity with the previous example.

As it is expected, the degeneracy occurs at a higher frequency and the transition is smoother than in the previous case. The smaller gap between the two indices allows the field to spread sooner into the cladding.

At low γ, the two examples become similar, and the importance of the cores becomes negligible in the field distribution. Table 16 shows how the propagation characteristics come closer to each other.

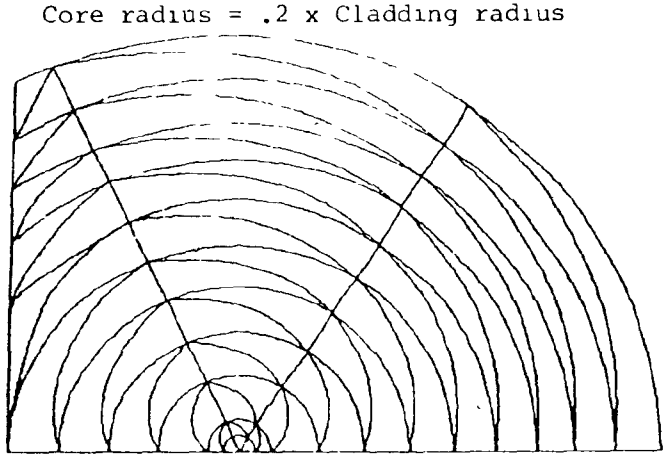
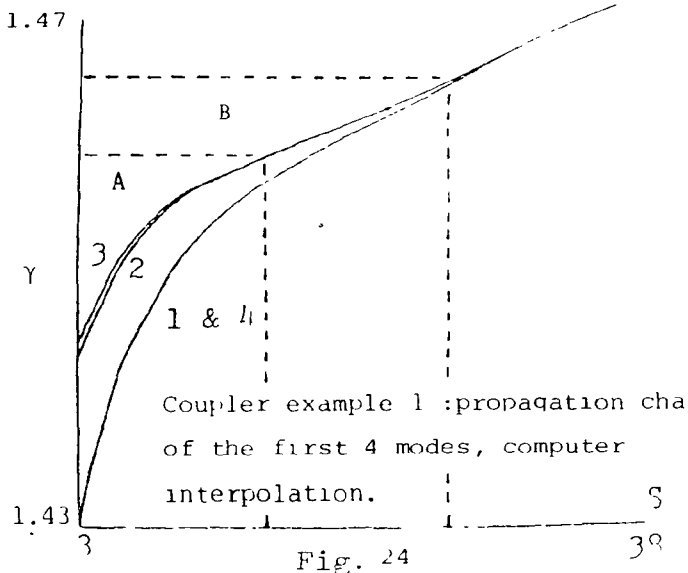
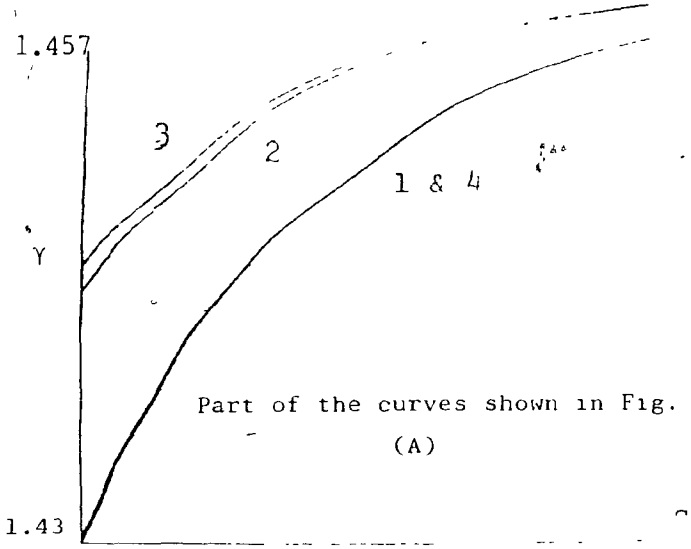


Fig. 23, : Coupler discretisation



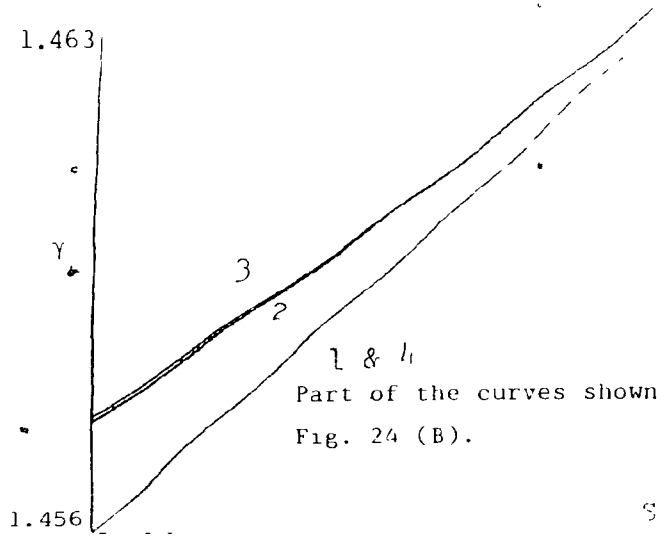
Coupler example 1 :propagation char. of the first 4 modes, computer interpolation.

Fig. 24



Part of the curves shown in Fig. 24. (A)

Fig. 25A



Part of the curves shown in Fig. 24 (B).

Fig. 25B

γ	S KBC=1	S KBC=2	S KBC=3	S KBC=4
1.490	77.098	77.084	77.097	77.084
1.489	62.605	62.768	62.605	62.768
1.480	53.585	53.585	53.585	53.584
1.475	45.421	45.426	45.422	45.425
1.470	37.929	37.931	37.927	37.934
1.465	30.705	30.612	30.609	30.707
1.463	26.896	26.593	26.580	26.901
1.462	23.959	23.402	23.379	23.962
1.458	22.654	20.337	20.326	22.677
1.457	20.418	17.574	17.446	20.360
1.455	17.217	14.414	14.207	17.226
1.450	13.385	10.617	10,354	13.422
1.445	11.148	8.738	8.454	11.187
1.440	9.726	7.558	7.262	9.765
1.435	8.720	-	-	8.759
1.430	7.957	6.109	5.798	7.994

Note : The example 4.4 has demonstrated the poor performance of the programs near the cladding cutoff. Values close to 1.46 +/- 0.002 are ignored.

TABLE 14 : Example 1
FEM solution for the first modes.

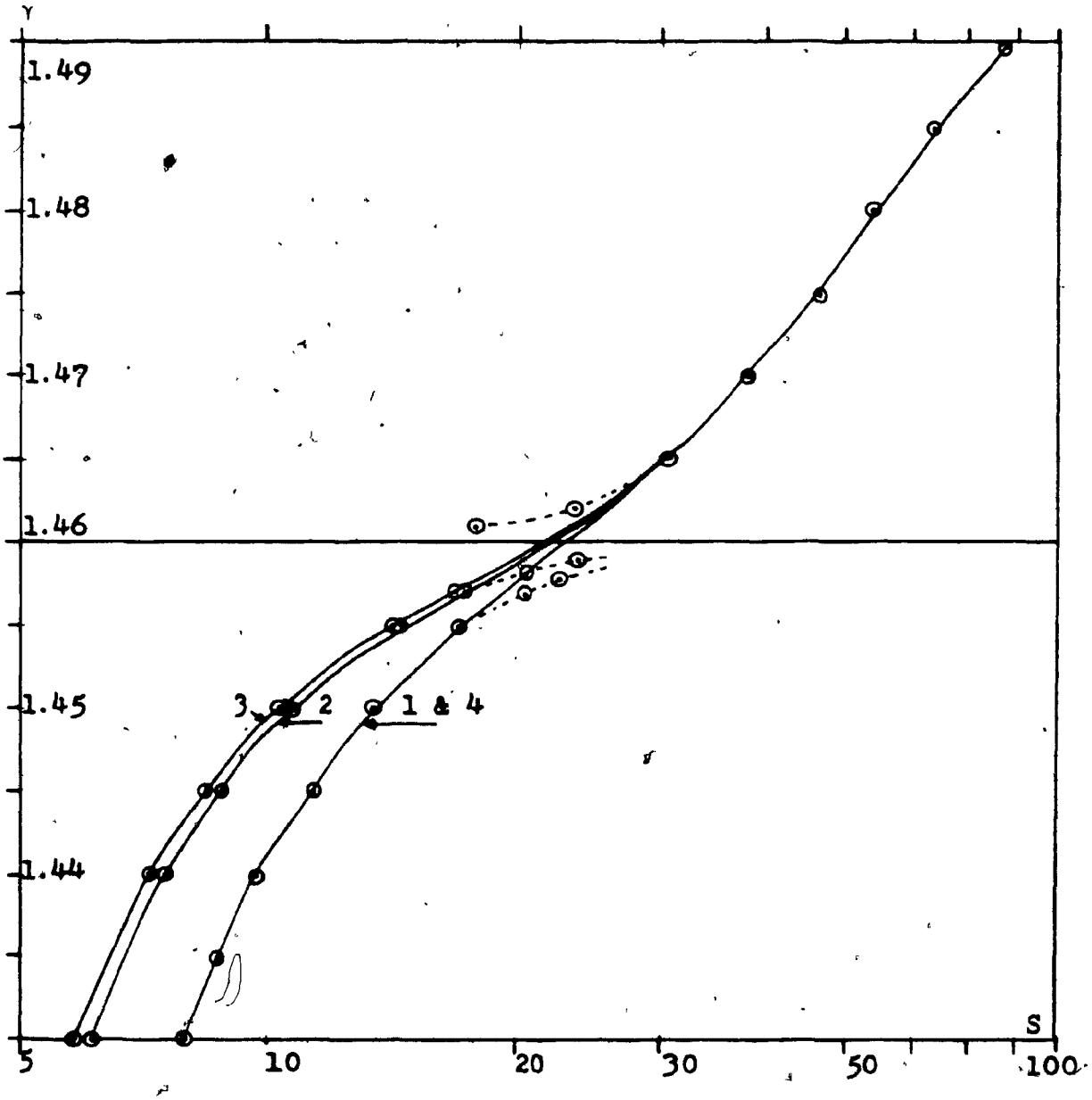


Fig. 26 : Coupler example 1.
Propagation characteristics of the first 4 modes.

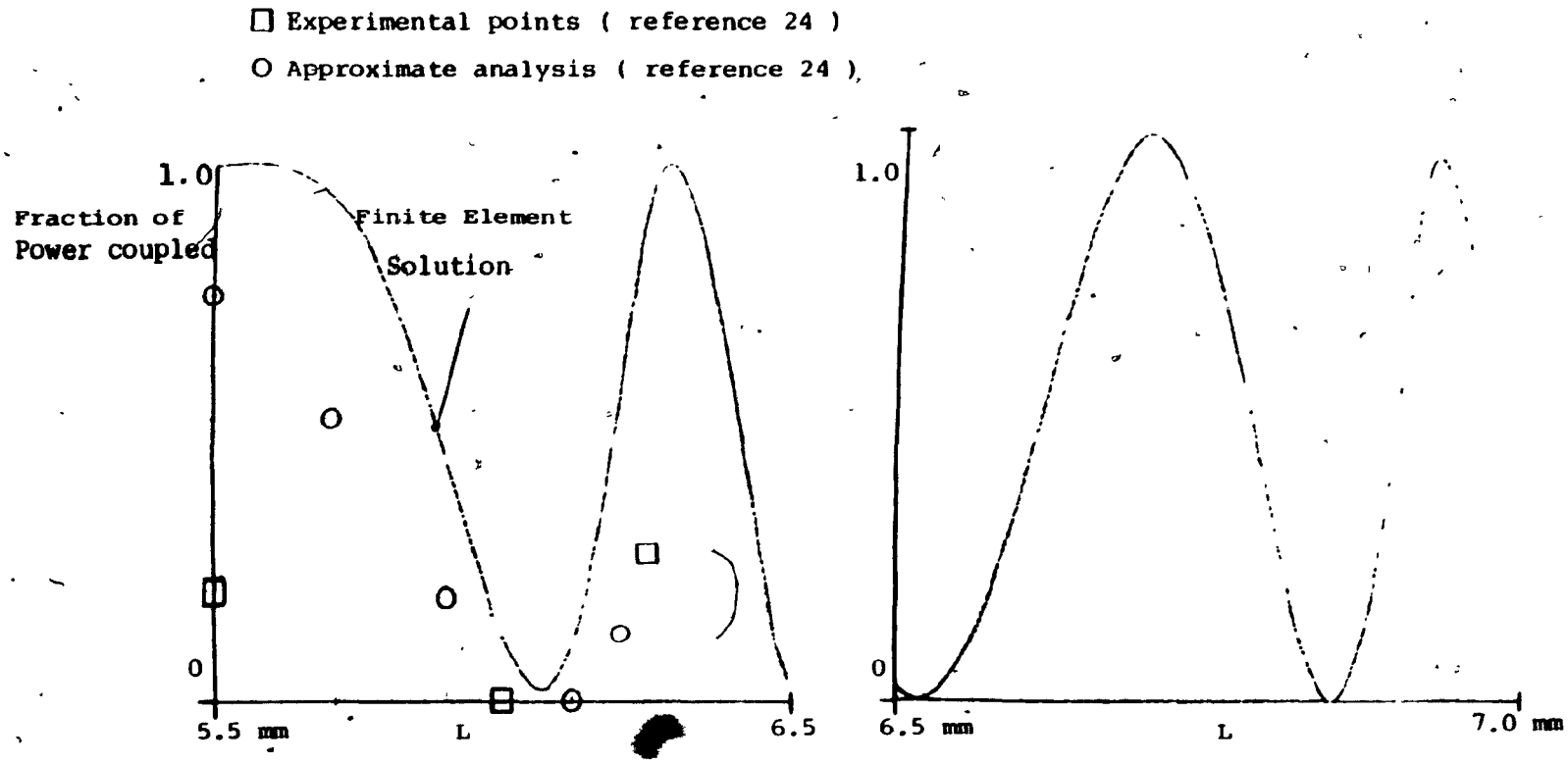


Fig. 27 : Coupled power versus coupler length L (example 1).

Initial diameter of the fibers = 100 μ = $2R_{in}$

Laser beam wavelength = 0.6328 μ = λ

Flame width = 1.0 mm = Δz

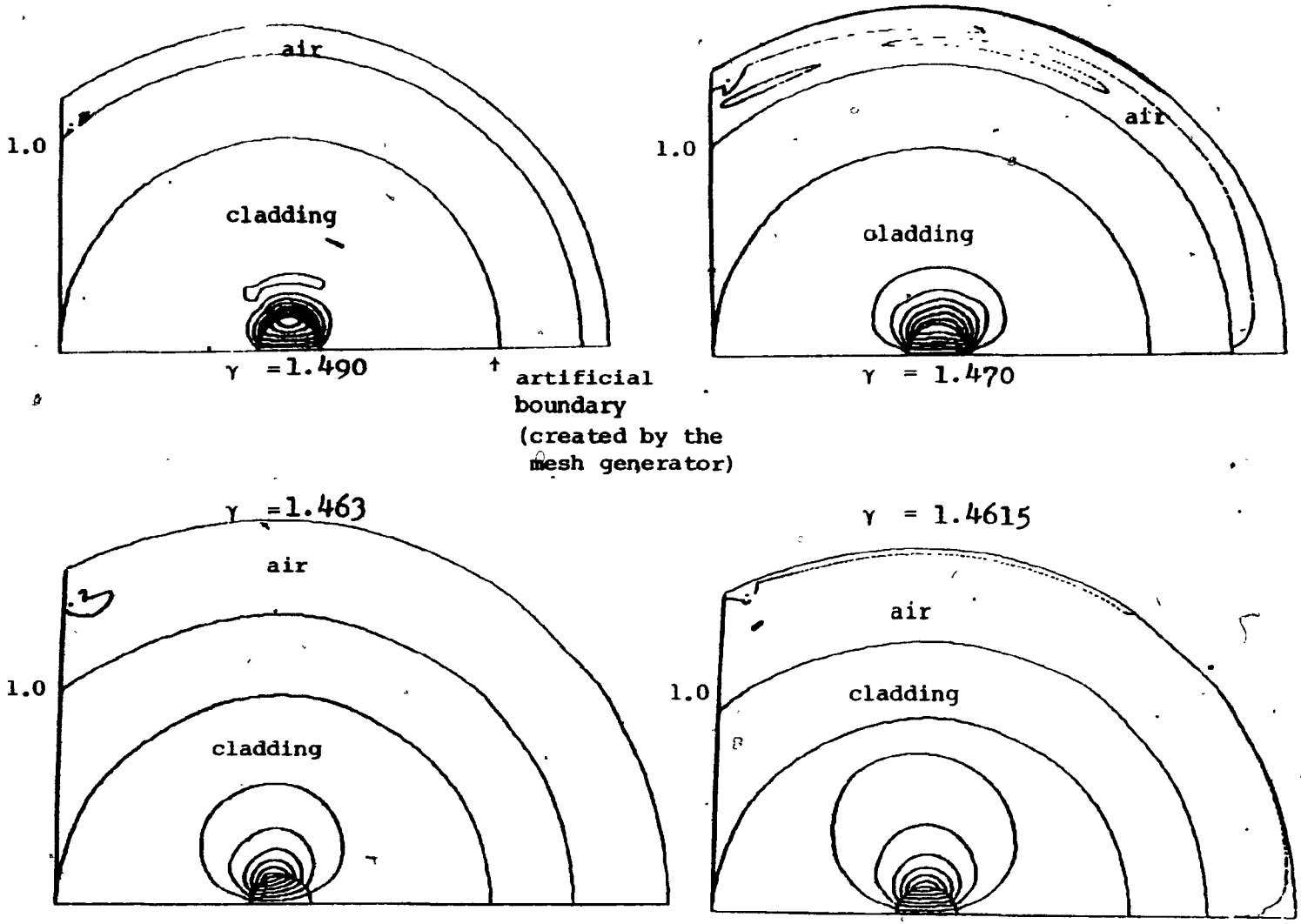
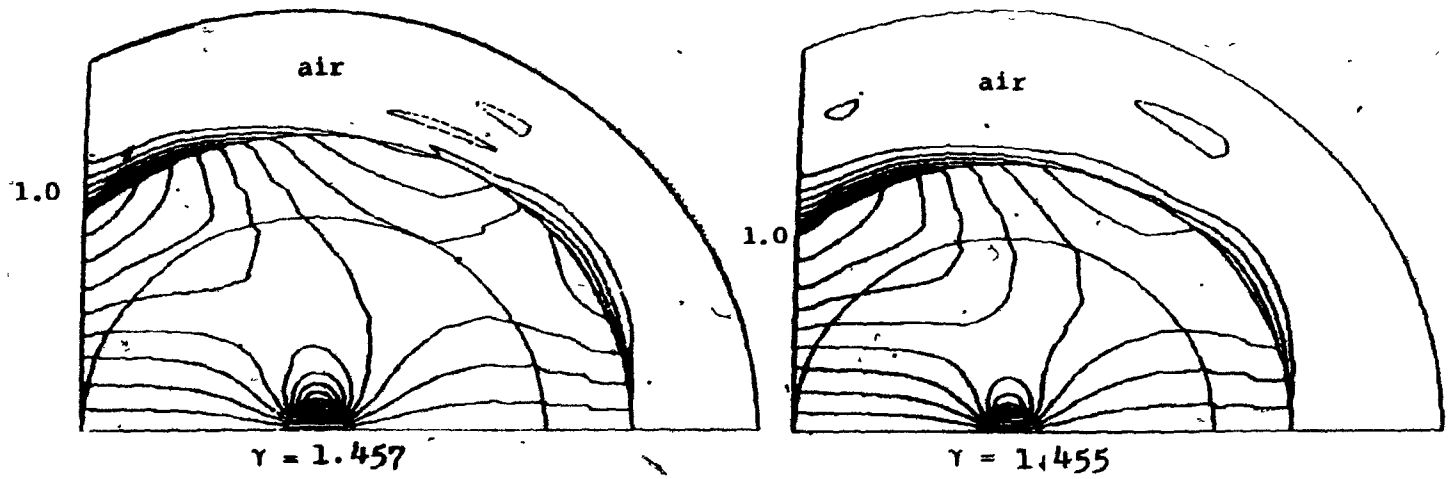


Fig. 28 : Coupler, Ez field, linear scale, boundary conditions code = 2, fundamental mode.



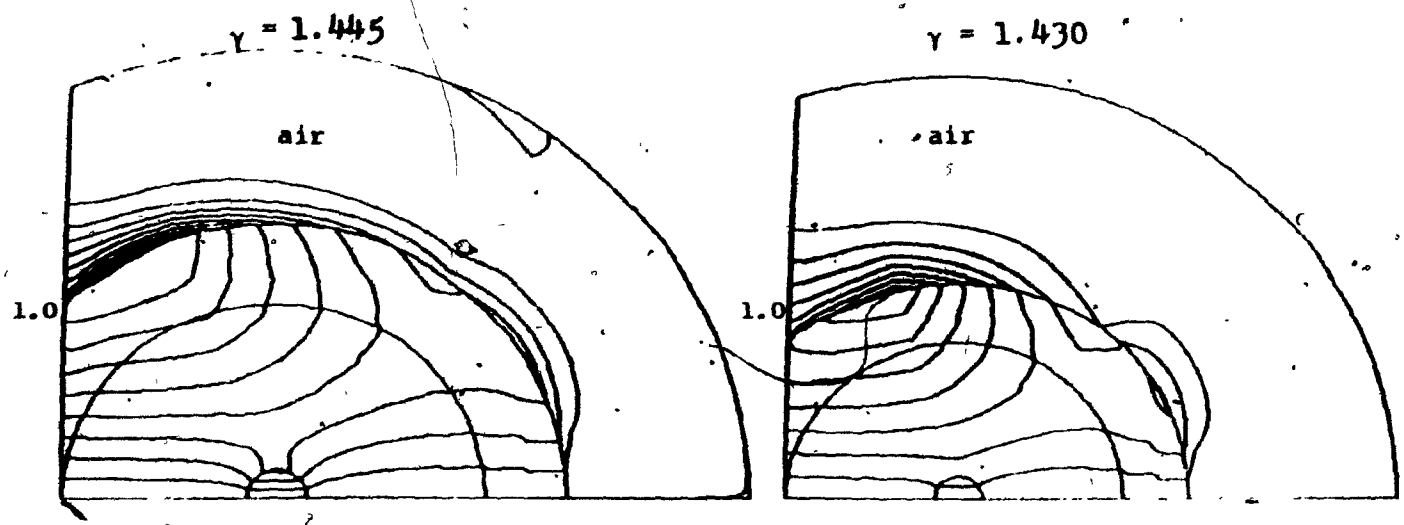


Fig. 29 : Coupler, Ez field, linear scale, boundary conditions code = 2, fundamental mode.

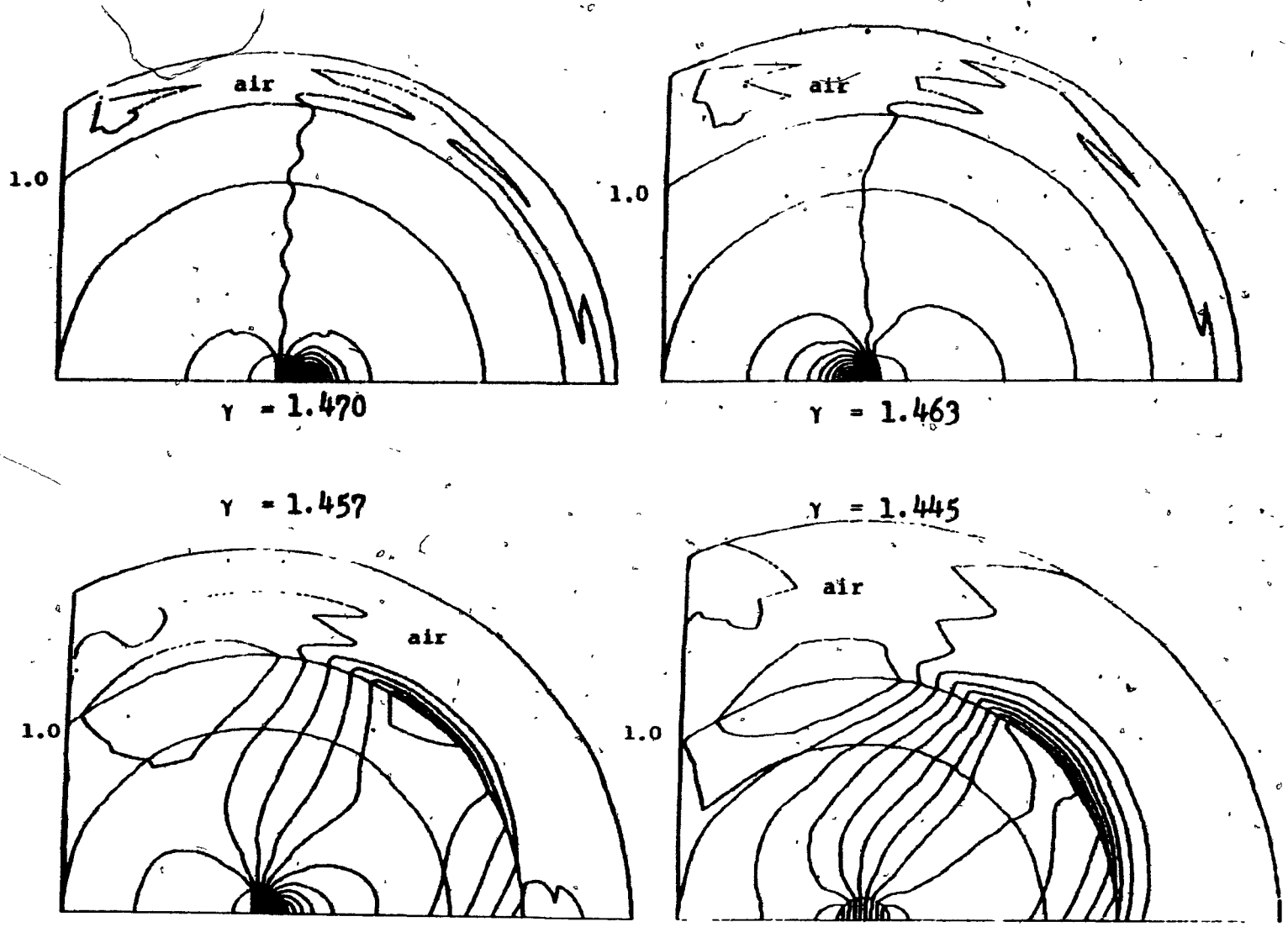


Fig. 30 : Coupler, Hz field, linear scale, boundary conditions code = 2, fundamental mode.

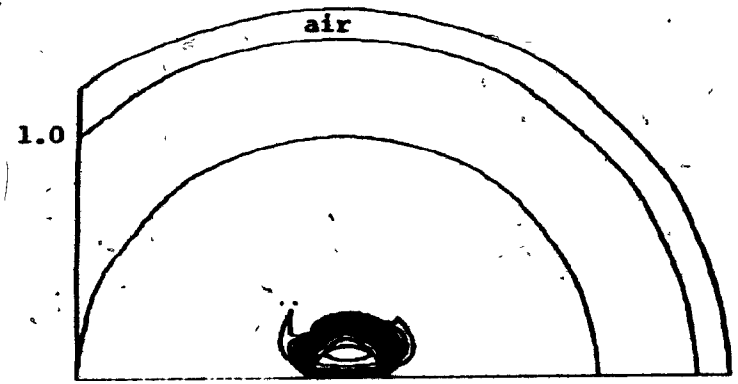


Fig 31 : Coupler, Ez field, log. scale,
γ = 1.490,
boundary conditions code = 2 ,
fundamental mode.

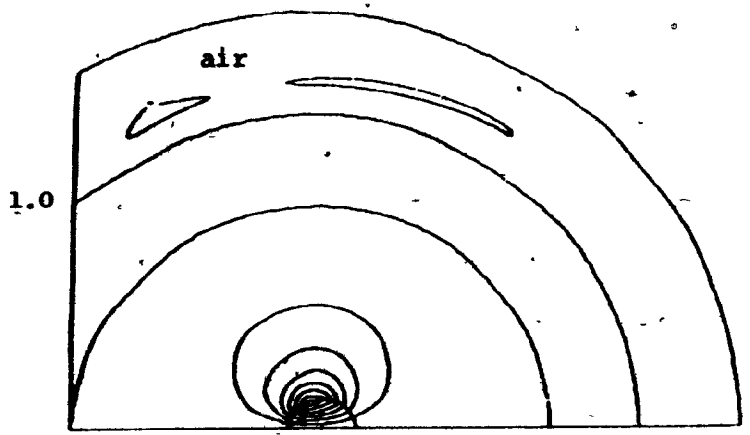


Fig. 32 : Coupler, Ez field, linear scale,
γ = 1.463,
boundary conditions code = 4,
fundamental mode.

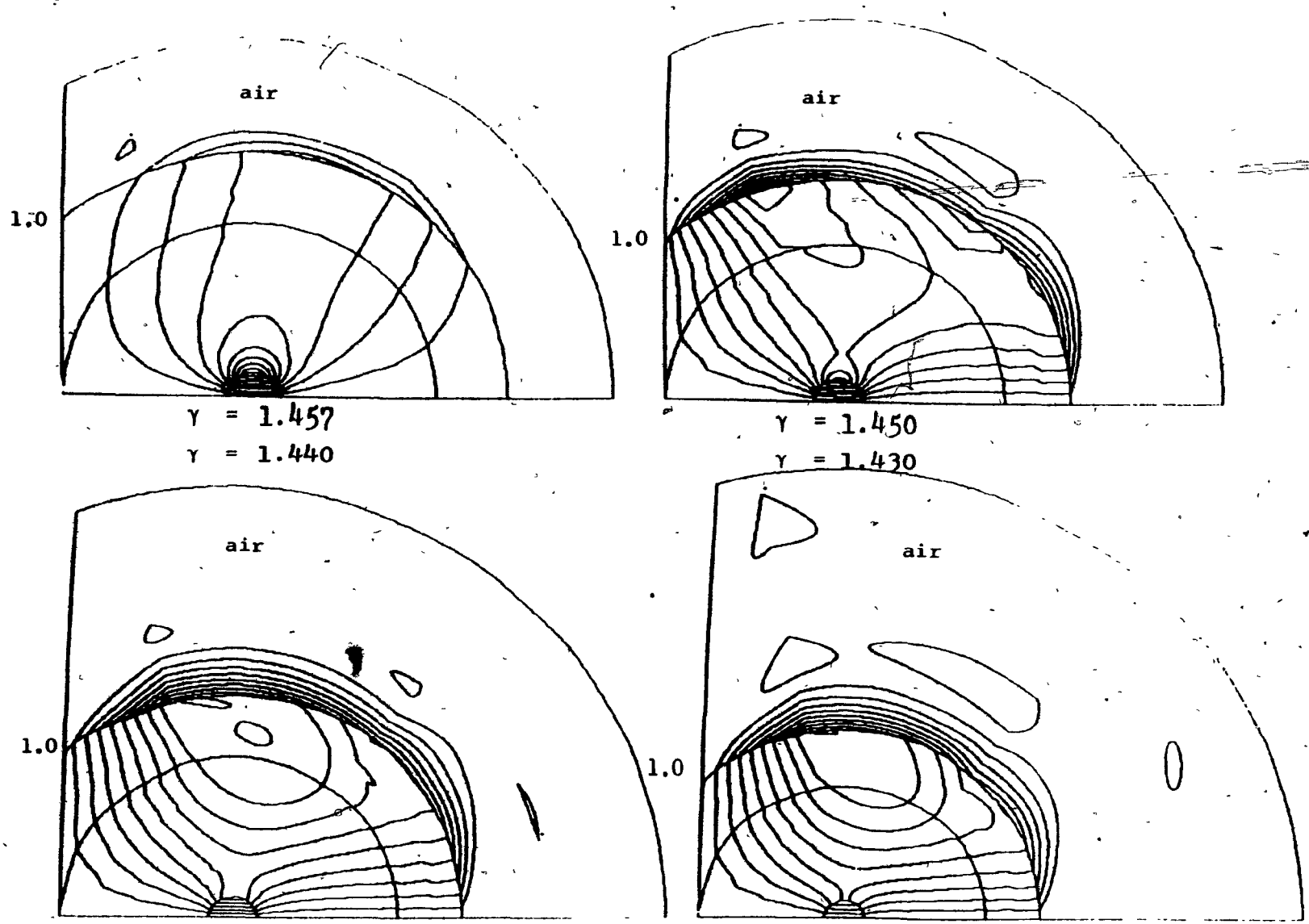


Fig. 33 : Coupler, Ez field, linear scale, boundary conditions code = 4, fundamental mode.

γ	S KBC=1	S KBC=2	S KBC=3	S KBC=4
1.4645	138.06	138.10	138.06	-
1.4630	105.69	105.70	105.69	-
1.4620	89.12	89.11	89.11	-
1.4570	24.367	20.316	20.061	-
1.4550	17.598	15.768	15.471	
1.4500	14.275	11.065	10.746	-
1.4450	11.757	9.022	8.709	11.803
1.4400	10.126	7.747	7.428	10.173
1.430	8.189	6.224	5.900	8.234

TABLE 15 : Example 2
FEM Solution for the first modes

Note : Because of the degeneracy between modes "KBC=1" and "KBC=4", some values have not been computed.

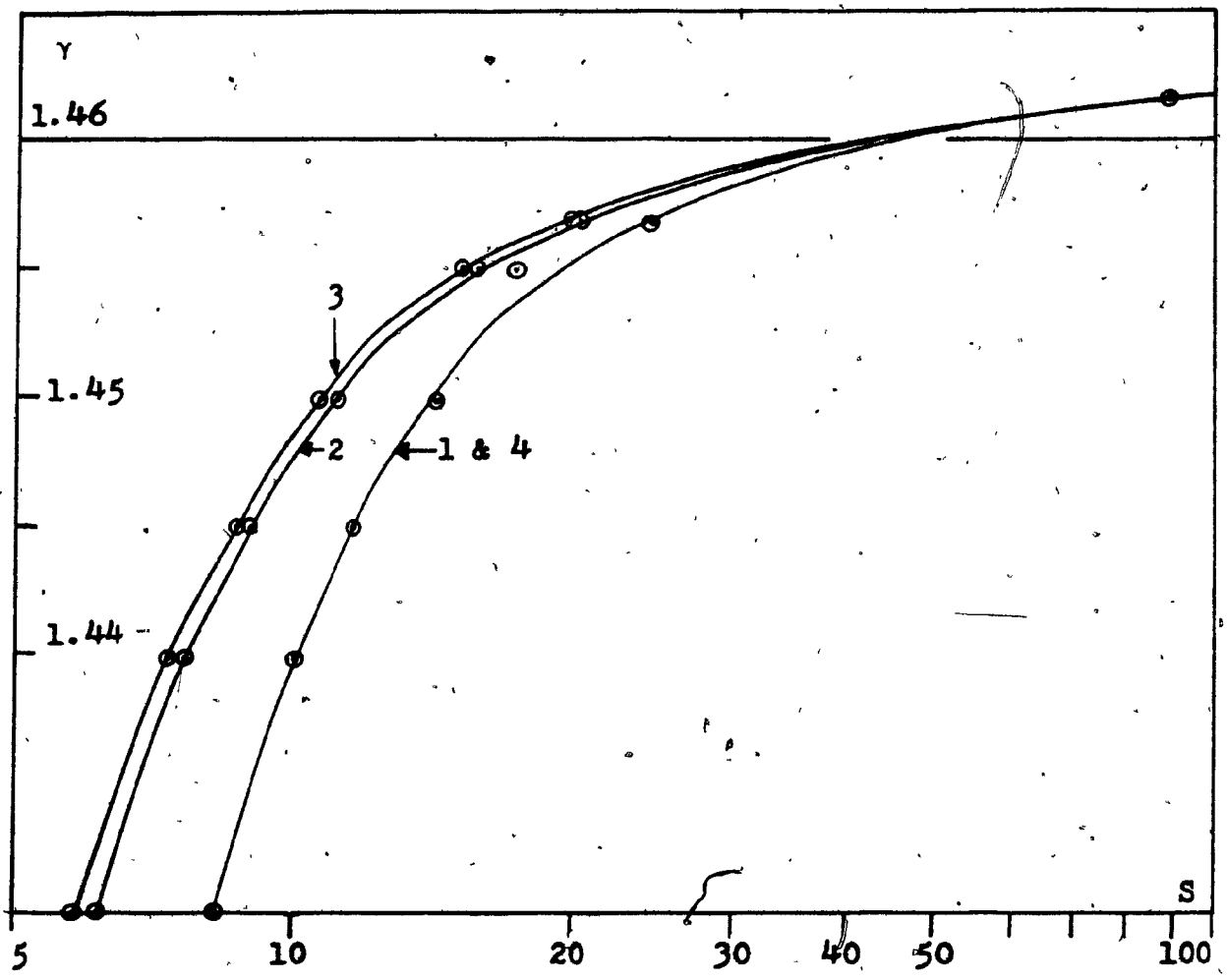


Fig. 34 : Coupler example 2.
Propagation characteristics of the first 4 modes.

γ	KBC=1 S	KBC=2 S	KBC=3 S	KBC=4 S
1.463	78.8	79.1	79.1	-
1.457	3.95	2.74	2.61	-
1.455	0.38	1.35	1.26	-
1.450	0.89	0.45	0.39	-
1.445	0.61	0.28	0.26	0.62
1.440	0.40	0.19	0.17	0.41
1.430	0.23	0.11	0.10	0.24

TABLE 16 : Examples 1 and 2

Difference between the normalized frequency versus γ .

5.4.3 Comparison with experimental data

Such a comparison is difficult because of the great difficulty in obtaining reliable, accurate and experimentally repeatable data.
The coupler analyzed in the example 2 has been studied in Ref. [24]. Unfortunately, its real shape and the exact index values are not known with a great accuracy. Even when the fiber's data are good, the building process which involves a high temperature manipulation may change the parameters in an unpredictable way.
A better model will necessitate detailed analysis of the built coupler, with accurate measurements of the indices and the dimensions.
But the previous analysis has shown the general behavior of these structures and particularly the triggering effect of the cladding cutoff. It has also given the right order of magnitude for the coupling data as is shown on Fig. 27.

5.4.4 Accuracy of numerical results

As has been shown previously, the discontinuity which occurs when γ is equal to one of the material indices must be avoided. Therefore the points close to the cladding cutoff should be selected with care.

The use of double precision in that particular case would probably extend the range of validity of the method by an appreciable margin at the additional cost of memory size and double precision arithmetic. The contour plots distortion is a good indication of the accuracy of the eigenvalues.

On both examples, we can see some degradation close to the discontinuity which is also partially caused by the increasingly complicated nature of these curves. In those cases, a discretisation with more elements in the angular distribution would improve the final results. Unfortunately, such discretisation increases computational expense.

No attempt was made to compare isoparametric solutions with conventional finite element solutions.

5.4.5 Equivalence with Snyder coupled-mode theory

From the reference [25], the total field ($\vec{E}$ and $\vec{H}$) in a coupler may be written as a combination of z-modulated modes :

$$\vec{E} = \sum_p a_p(z)\ \tilde{e}_p(x,y) \qquad (5.64)$$

$$\vec{H} = \sum_p a_p(z)\ \tilde{h}_p(x,y) \qquad (5.65)$$

Where :

- $\vec{E}$ and $\vec{H}$ are the total field.
- a_p is a modulation factor.
- $\tilde{e}_p$ and $\tilde{h}_p$ are the modal fields of each separate fiber.
- p is the mode index.

The coefficients a_p satisfy the following coupled equations :

$$\frac{da_p^j(z)}{dz} + \frac{i}{} \beta_p^j\ a_p^j = \frac{i}{} \sum_{s \neq j} a_p^s\ C_{pp}^{js} \qquad (5.66)$$

Where :

- β_p^j is the propagation constant of mode "p", fiber "j"

- C_{pp}^{js} is the coupling coefficient between the mode p of two fibers j and s.

- $\pm i$ is $\sqrt{-1}$.

If we restrict this analysis to two identical monomode fibers i and j, we obtain :

$$\beta_p^i = \beta_p^j$$

and

$$\frac{da_p^i}{dz} + \pm i \beta_p^i a_p^i = \pm i a_p^j C_{pp}^{ij} \tag{5.67}$$

$$\frac{da_p^j}{dz} + \pm i \beta_p^j a_p^j = \pm i a_p^i C_{pp}^{ji} \tag{5.68}$$

As it is shown in reference [25], the solution of these differential equations leads to a ratio of transfer power $\bar{P}_S = P_S/Po$ which can be expressed as :

$$\bar{P}_S = \cos^2 (C_{pp}^{ij} z) \tag{5.69}$$

when C_{pp}^{ij} is constant, or :

$$\bar{P}_S = \cos^2 \int_{-\infty}^{z} C_{pp}^{ij} (\xi) \, d\xi \tag{5.70}$$

This expression corresponds to equation (5.63) for each polarization. If we call p_1 the first one and p_2 the other, we obtain :

$$\bar{P}_S = \cos^2 \int_{-\infty}^{z} C_{p1,p1}^{ij} \, d\xi + \cos^2 \int_{-\infty}^{z} C_{p2,p2}^{ij} \, d\xi \tag{5.71}$$

Comparing this with the results given in section 5.3.5, we obtain :

$$C_{p1,p1}^{ij} = \frac{\beta_1 - \beta_3}{2} \quad \text{and} \quad C_{p2,p2}^{ij} = \frac{\beta_2 - \beta_4}{2} \tag{5.72}$$

Therefore, we can write that the coupling coefficient defined by Synder in his theory is an elegant approximation of the difference between the propagation constants of quasi-degenerate modes.

For simple structures such as quasi-isolated optical fibers, this

approximation is a quick method to obtain the results without solving the complete problem as it has been shown in reference [24], by ignoring the core.

But it does not show the effect of the cladding cutoff as our model does, and it does not predict the difference between the coupling ratios of the different polarizations. These two points, in addition to the achievement of contour plots had improved our understanding of the phenomenon and raised new questions which will probably find some answers in a near future with other experimental work.

6. PROGRAMS

6.1 Structure

The solution of a problem is found by executing several programs in a sequence. The file manipulation is done through a procedure file which can be adapted to particular needs or constraints.
The detailed system flow chart is shown in Figs. 35 to 37. There are four major and five secondary programs for the processing of the results.
The maximum number of elements is restricted by the computer central memory size available for the eigenvalue solver.

6.2 The Pre-processors

6.2.1 Input program "DWIN"

This is an interactive input program which builds the input file "DWIN.DAT" from the answers to the questions asked of the user. This file is read by the other programs which discretise certain geometries and give the FEM solutions. This method is fast, efficient and suppresses input format errors which often occur when the user edits his own input file directly.

6.2.2 The automatic mesh generator "ELF"

It is fed by the input file "DWIN.DAT". For each value of γ (the normalized propagation constant), and with the information collected from "DWIN", it generates sets of two data files :
-"ELEMn.DAT" which contains the characteristics of the elements
-"NODEn.DAT" which contains the characteristics of the nodes.
n is a label for each γ, from 1 to 16.
The mesh generator contains two main subroutines :
-"ELGN1F' which generates the mesh for a circular or elliptical multilayer fiber.

-"ELGN2F" which generates the mesh for a circular or elliptical multilayer coupler. Fig. 5, 11, 13 and 23 show typical meshes.
Both subroutines use a radial and angular division of the region to be discretised. Table 18 gives the number of elements, nodes and free variables versus the number of angular and radial subdivisions.
An automatic mesh generator has been preferred to an interactive mesh generator (in which the user sets the mesh in a more direct approach) because of the special geometry under study.
"ELGN1F" has been used extensively to test the programs with known examples (see section 4). "ELGN2F" has been used for the analysis of the coupler (see section 5).
Any other type of mesh generator may be substituted for "DWIN" and "ELF" as long as their output formats are compatible with the other programs.

6.3 The Solvers

The solver is made of two distinct programs : the global matrices generator ("DW1P") and the eigenvalue solver (DW2PX, DW2PXX).

6.3.1 The Global Matrices Generator "DW1P"

This program executes the coding which is mathematically described in chapter 3. It is made of 4 main and 4 secondary subroutines, in addition to the main program "DWMAIP". This version keeps all the arrays in central memory and requires more than 64 Kbytes.
These subroutines are :
-"COPYX" : copies the input data files "ELEM.DAT" and "NODE.DAT" in central memory arrays. These two files are chosen among the "ELEMn.DAT" and "NODEn.DAT" files previously generated by the program "ELF".
-"CALPAR" : It has be shown easily in section 3 that many quantities are independent of the element and may be computed only once, at the beginning of the program. This subroutine does these calculations and stores the results in an array called "XVA".
-"FPINIP" does some preliminary processing before the generation of

the global matrices.

-"FONPRP" executes the most important part of the program. Using the "ABEL" and "JACOB" subroutines, it generates the global matrices $[A_G]$ and $[B_G]$ which are stored in a compressed format in the "DWABMX.DAT" data file. Only the non-zero terms are kept, with their row and column numbers.

The other subroutines are :

-"ABEL" which computes the contribution of each pair of nodes to the global matrices using the results in section 3.

-"JACOB" which computes the Jacobian matrix and some quantities related to it for each element.

-"FN" computes the interpolation functions and their partial derivatives in the reference element.

-"LOCATE" returns the linear storage mode location index of one element of a symmetric matrix.

"FPINIP" and "FONPRP" subroutines are modified versions of two earlier versions respectively called "FPINIT" and "FONPRE". Other versions of this program have been developed using temporary disk data storage, which may be run on machines with small central memory capacity.

6.3.2 The Eigenvalue Solvers DW2PX, DW2PXX

These are two full matrix eigenvalue solvers.

They reduce the generalized eigenvalue problem to a standard problem, tridiagonalize the matrix and use a Sturm sequence search to find the eigenvalues. Eigenvectors are computed by inverse iteration. No use is made of the sparsity of the original matrices, and little time is saved by requesting only a few eigenpairs, as most of the computational effort is involved in the tridiagonalization.

Because of the central memory size limitation of the minicomputer being used, three temporary disk files (shown on the system flow chart Fig. 36) are needed to perform the computations.

The eigenvalue solver is by far the most time-consuming and expensive part of the total process. Its large memory size requirements

had limited our problem size to approximatively 62 quadratic elements.
An optimum discretisation had to be found to reduce costly iterations.
An efficient, fast and reliable sparse matrix eigenvalue solver for general systems (not only for positive definite matrices) will make FEM analysis of waveguides much more attractive.

6.4 The Post Processors

6.4.1 Computation of the transverse field components with "DW5P" program.

This program carries out the computations described in section 3.5, using four input files and five subroutines.
The input files are :
-"DWIN.DAT", "ELEM.DAT", "NODE.DAT", (The input files for the solver "DW1P") and "DWEzHz.DAT" which is a file copied from the output of "DW2PX" or "DW2PXX".
The subroutines are :
-"CALTAR" which computes the "V_{jk}" functions previously defined in section 3.5 at each node of the reference element.
-"COPYTR" which reads the input files.
-"XDERIV" which computes the transverse field components ϕ_x, ϕ_y, ψ_x and ψ_y at each node, in each element.
-"DFN" which contains the derivatives of the interpolation functions needed by "XDERIV".
The output and the averaging process are done in the main program "TRMAIN".
Three output files are generated :
-"DWExEy.DAT" contains ψ_x and ψ_y.
-"DWHxHY.DAT" contains ϕ_x and ϕ_y.
-"DWSz.DAT" contains the z-component of the Poynting vector $\vec{S}$. All these files are compatible with the graphic program "DW6P" (see next section) if the graphic code in "DWIN.DAT" file is on.

6.4.2 The graphic program "DW6P"

6.4.2.1 Theory

This program draws the equipotential lines in a cross section of the waveguide for any component of the field. The plot is done point by point with a double sweep in each element executed from the reference triangle using the following sequence :

-Normalise the nodal values of the field component to be plotted.

-Compute the real size of each element to determine the number of points in the sweep process.

-Plot the boundaries of the waveguide, and if wanted, the discretisation.

-Execute the sweep, element by element.

This is done knowing that the field in each element can be written as :

$$\phi_i (\xi,\eta) = \sum_j N_j(\xi,\eta) \phi_{ij} \qquad (6.1)$$

Where the N_j's are the interpolation functions and the ϕ_{ij} are the nodal values obtained from "DW2PX" or "DW2PXX".

The equipotential line "k" will be the set of points which satisfies the following equation :

$$\phi_k = \sum_j N_j \phi_{ij} \qquad (6.2)$$

where ϕ_k is a value generated by the program from data entered by the user during execution.

The equipotential line is computed in the reference element first and then mapped into the real element using the transformation T_i previously defined in section 3. Using the interpolation functions from the table 1, we can write :

$$\phi_k = (2\phi_1^i - 4\phi_2^i + 2\phi_3^i)\xi^2 + (2\phi_1^i + 2\phi_5^i - 4\phi_6^i)\eta^2$$
$$+ (4\phi_1^i - 4\phi_2^i + 4\phi_4^i - 4\phi_6^i)\xi\eta + (-3\phi_1^i + 4\phi_2^i - \phi_3^i)\xi \qquad (6.3)$$
$$+ (-3\phi_1^i - \phi_5^i + 4\phi_6^i)\eta + \phi_1^i$$

The sweep process consists of finding the set of points (ξ ,η) which satisfies the above equation for a given set of ϕ_k values. One ξ-sweep (given ξ, solve for η) and one η-sweep (given η, solve for ξ) are done for an optimum coverage. Equation (6.3) is solved using standard second order polynomial analysis.

This basic method is improved greatly using two subroutines "ELSIZE" and "EXTREM" which have the following purpose :

-"ELSIZE" computes the relative size of each element with respect to the total area using the following relation :

$$Si = \iint_{Si} dx\, dy = \iint_{\text{Reference triangle}} |J_i|\, d\xi\, d\eta \qquad (6.4)$$

In that computation, the accuracy is not critical, and we use a simple numerical integration with three points :

$$Si = \sum_{R=1}^{3} W_R \, | J_i(\xi_R, \eta_R) | \qquad (6.5)$$

But we have seen previously in section [3] that :

$$| J_i | = \sum_j \sum_k -V_{jk} X_{ij} Y_{ij} \qquad (6.6)$$

Then, we obtain :

$$Si = \sum_R \sum_j \sum_k W_R V_{jk} (\xi_R, \eta_R) X_{ij} Y_{ij} \qquad (6.7)$$

This computation is done for each element and the number of sweeps is proportional to the square root of the ratio $S_i / \Sigma S_i$.

The density of points in the sweep may be changed during program execution by the user.

-"EXTREM" subroutine computes the extremum of the field function inside an element "i" knowing the nodal values ϕ_{ij}. This problem is reduced to the determination of a maximum and a minimum of a 2-

dimensional quadratic equation with the following constraints :

$$0 \leq \xi \leq 1 \qquad 0 \leq \eta \leq 1-\xi \tag{6.8}$$

This standard problem is solved in appendix 4. Once the extremum of ϕ_i in the element "i" are known, it is not necessary to look for the equipotential lines ϕ_k such that $\phi_k > \text{Max}(\phi_i)$ or $\phi_k < \text{Min}(\phi_i)$. This feature increases the plotting speed by a factor of 2, 3 or even greater in some cases, when the field strength is concentrated in small regions.

6.4.2.2 Implementation

This program is run with the "PLOTCAD" graphic routines implemented on a PERQ minicomputer, but may be easily adapted to any fortran compatible graphic system.

Three input files are needed :

"ELEM.DAT" and "NODE.DAT", the input files to the solver "DW1P", and "DWGR.DAT" generated by "DW2PX","DW2PXX", or "DW5P" with a compatible graphic format.

The subroutines are :

-"COPYGR" copies the input files into the corresponding arrays.

-"NORMF" normalises the field.

-"ELSIZE" computes the relative size of each element.

-"XSCALE" executes any linear scale-transformation required by the user (translation, scale change, rotation).

-"ELSHOW" shows the detail of the discretisation and plots the axes through "POINT" and "LINE" subroutines.

-"ELEQUI" solves the quadratic equations and plots the equipotential lines through "POINT" subroutine.

-"EXTREM" computes the local extremum of the field in each element.

-"REFXY" maps any point from the reference triangle to the real element using T_i transformation (see section [3]).

-"DFN" computes the partial derivatives of the N_j interpolation functions

The subroutines from "PLOTCAD" library are :

-"MOVABS" which moves the cursor on the screen.
-"DRWABS" which draws one line on the sceen.
-"ERASE" which erases the screen.
-"INITT" which initialises the graphic system at the beginning.

"GRMAIN" is the main program which requires some additional interactive input data. A log or linear scale may be selected, with up to 50 equipotential lines to be drawn. The output goes on a screen with 512 x 512 addressable points. All the contours shown in this thesis have been plotted with this program on a "PERQ" mincomputer high resolution black and white screen.
The point-by-point process is of course slower than a line-by-line process, but it gives smooth contours and a simple fault-free alternative for quadratic problems.

6.4.3 The coupler analyser "DW8P(2)"

This program does several important computations for the analysis of quasi-degenerate modes.

6.4.3.1 The $\Delta\gamma$ Characteristics and Error Analysis

The main program "DBETAS" implements the method described in sections 4.3.2 and 4.3.3.

6.4.3.2 Computation of the Coupling Coefficients

The computation process described in section 5.3.5 is carried out in subroutine "CPL", and the process described in section 5.3.6 in subroutine "CPL2".

PROGRAM NAME	DESCRIPTION
DWIN	Input program (Implemented on the PERQ minicomputer).
ELF	Automatic mesh generator for a circular or elliptic fiber or coupler (Implemented on the PERQ).
DW1P	Global matrices generator (Implemented on the PERQ).
DW2PX	Full matrix eigenvalue solver.
DW2PXX	Identical to "DW2PX", but using unformatted temporary disc files (Implemented on the PERQ).
DW3	Dielectric rod double-precision analytical solver (Implemented on the PDP 11/23 minicomputer).
DW5P	Transverse components program (Implemented on the PERQ).
DW6P	Graphic program (Implemented on the PERQ).
DW6X	Identical to "DW6P", but implemented on the PDP 11/23.
DW8P	Computes the $\Delta\gamma$ characteristics, the coupling coefficients and the coupled power for two modes interaction (Implemented on the PERQ).
DW8P2	Identical to "DW8P", but can solve 4 modes interaction problems.

TABLE 17 : List of programs

M		N = 17	19	21	23	25	27	29	31	33
5	NEL	30	34	38	42	46	50	54	58	62
	NN	79	89	99	109	119	129	139	149	159
	N_{max}	117	133	149	165	181	197	213	229	245
7	NEL	45	51	57	63	69	75			
	NN	110	124	138	152	166	180			
	N_{max}	175	199	223	247	271	295			
9	NEL	60	68	76						
	NN	141	159	177						
	N_{max}	233	265	297						
		255 Limit	298 Limit							

NEL = Number of elements.

NN = Number of nodes.

N_{max} = Maximum dimension of the eigenvalue problem.

M = Number of angular divisions.

N = Number of radial divisions.

TABLE 18 : Dimension of the discretised problem vs the radial and angular division.

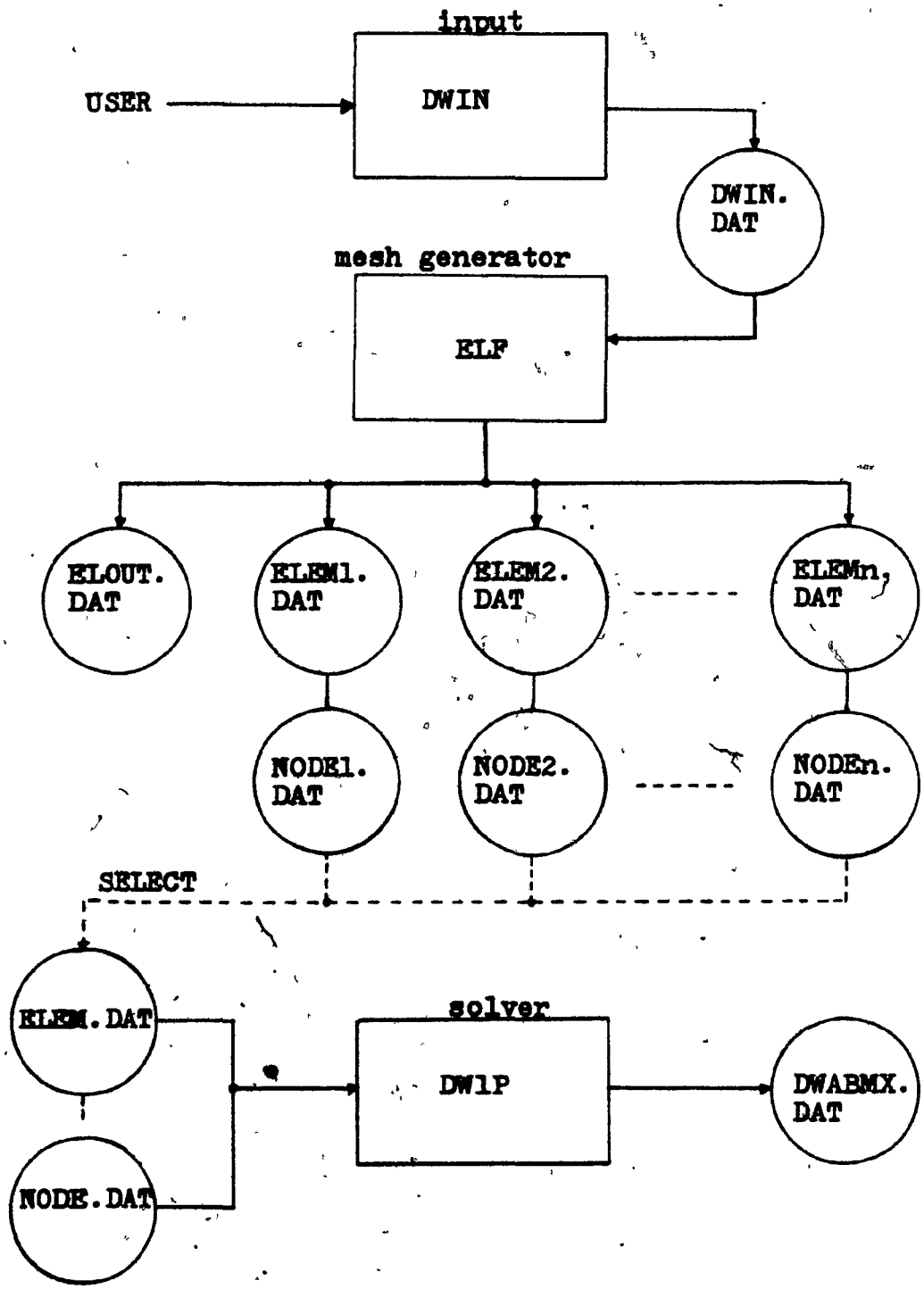


Fig. 35 : Flow chart (1)

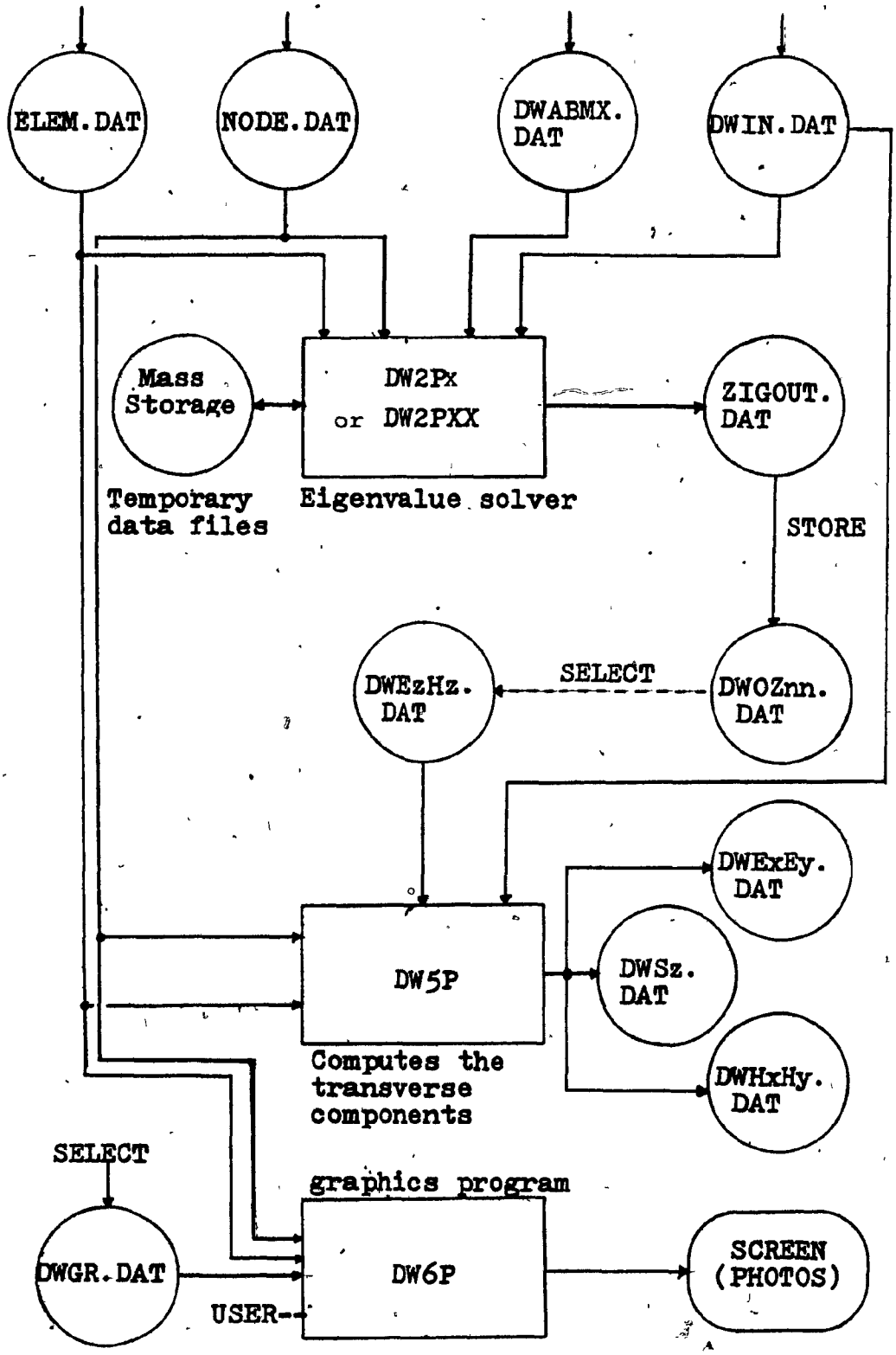


Fig. 36 : Flow chart (2)

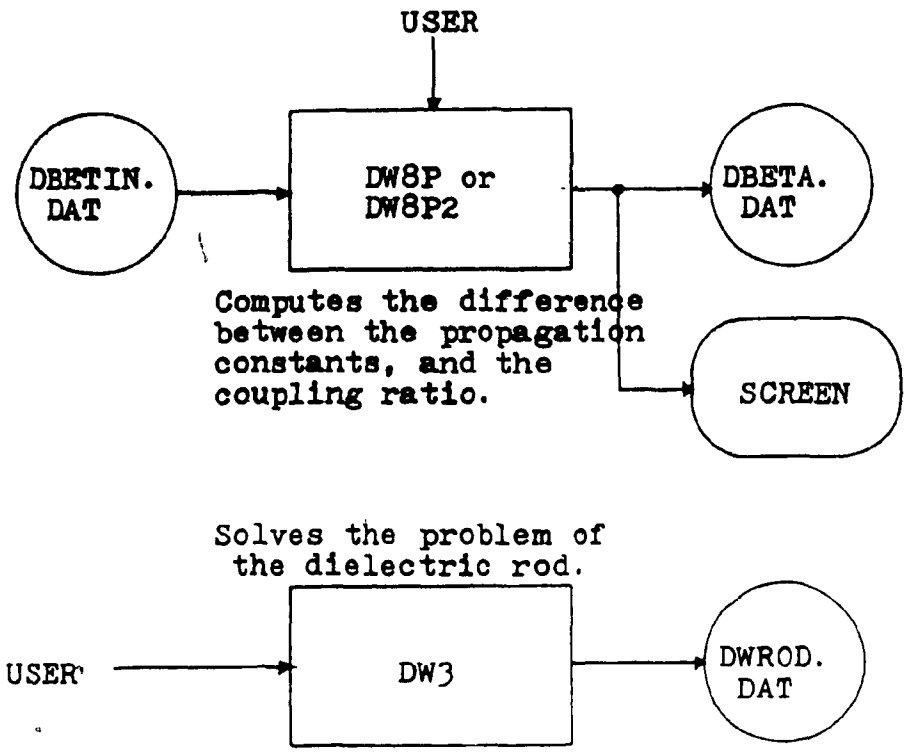


Fig. 37 : Other programs

7. CONCLUSION

We have demonstrated the ability of a general 2-dimensional Finite Element software package to analyse dielectric waveguides using quadratic, isoparametric elements.
Future developments of this method would include the use of a more efficient eigenvalue solver.

The coupler analysis had shown the mechanism involved in the energy transfer process with more details than ever before, even with the discontinuity due to the cladding cutoff.

The method is not limited by the shape of the guide and shows the real complexity of the modes interaction in general non symmetric structures.
The discretisation is made with a relatively small number of elements and presents a very small level of contour distortion.

Appendix 1 : The continuity conditions of the radial field components.

From the equations (3.19) and (3.20), we can write :

$$\vec{\psi}_t^i \cdot \vec{n} = A\tau_i \{ \gamma^2 \nabla_t \psi_i \cdot \vec{n} - (\vec{e}_z \times \nabla_t \phi_i) \cdot \vec{n} \}$$

$$\vec{\phi}_t^i \cdot \vec{n} = A\tau_i \gamma \{ \nabla_t \phi_i \cdot \vec{n} + (\frac{\varepsilon_i}{\varepsilon_0}) (\vec{e}_z \times \nabla_t \psi_i) \cdot \vec{n} \}$$

Equations (3.23) and (3.26) give :

$$\vec{e}_z (\nabla_t \phi_i \cdot \vec{n}) = (\frac{1}{A\tau_i} \vec{\psi}_t^i - \gamma^2 \nabla_t \psi_i) \times \vec{n}$$

$$(\frac{\varepsilon_i}{\varepsilon_0}) \vec{e}_z (\nabla_t \psi_i \cdot \vec{n}) = (- \frac{1}{A\gamma\tau_i} \vec{\phi}_t^i + \nabla_t \phi_i) \times \vec{n}$$

Using the relation $(\vec{e}_z \times \vec{f}) \cdot \vec{n} = \vec{e}_z \cdot (\vec{f} \times \vec{n})$ for any vector field $\vec{f}$ in x-y plane, we obtain :

$$(\frac{\varepsilon_i}{\varepsilon_0}) (\vec{\psi}_t^i \cdot \vec{n}) = \gamma \vec{e}_z \cdot \{ (-\vec{\phi}_t^i \times \vec{n}) + A\tau_i (\gamma^2 - \frac{\varepsilon_i}{\varepsilon_0}) (\nabla_t \phi_i \times \vec{n}) \}$$

$$\vec{\phi}_t^i \cdot \vec{n} = \gamma \vec{e}_z \cdot \{ (+\vec{\psi}_t^i \times \vec{n}) + A\tau_i (\frac{\varepsilon_i}{\varepsilon_0} - \gamma^2) (\nabla_t \psi_i \times \vec{n}) \}$$

But $\tau_i = \dfrac{\gamma^2 - 1}{\gamma^2 - \dfrac{\varepsilon_i}{\varepsilon_0}}$ and the constant A is independent of "i".

Therefore :

$$(\frac{\varepsilon_i}{\varepsilon_0}) (\vec{\psi}_t^i \cdot \vec{n}) = \gamma \vec{e}_z \cdot \{ -(\vec{\phi}_t^i \times \vec{n}) + A(\gamma^2 - 1) (\nabla_t \phi_i \times \vec{n}) \}$$

$$(\vec{\phi}_t^i \cdot \vec{n}) = \gamma \vec{e}_z \cdot \{ +(\vec{\psi}_t^i \times \vec{n}) - A(\gamma^2 - 1) (\nabla_t \psi_i \times \vec{n}) \}$$

$\vec{\phi}_t^i \times \vec{n}$ and $\vec{\psi}_t^i \times \vec{n}$ are continuous (see section 3.1.2).

$\nabla_t \phi_i \times \vec{n}$ and $\nabla_t \psi_i \times \vec{n}$ are also continuous because ϕ_i and ψ_i are continuous.

Therefore, $(\frac{\varepsilon_i}{\varepsilon_0}) \vec{\psi}_t^i \cdot \vec{n}$ and $\vec{\phi}_t^i \cdot \vec{n}$ are continuous.

The integral formulation of the continuity can be written :

$$\sum_i \int_{C_i} (\vec{\phi}_t^{\,i} \cdot \vec{n}) \, dl_1 = \sum_i \int_{C_i} \gamma \vec{e}_z \cdot (\vec{\psi}_t^{\,1} \times \vec{n}) \, dl_1 - \sum_i \int_{C_1} \gamma \vec{e}_z \cdot A(\gamma^2 - 1) (\nabla_t \psi_1 \times \vec{n}) \, dl_i$$

where "i" is the element number, C_1 its contour and dl_1 an infinitely small part of C_1.
Using the result (3.25), the divergence theorem and the relation $\vec{e}_z \cdot (\nabla_t \psi_1 \times \vec{n}) = (\vec{e}_z \times \nabla_t \psi_i) \cdot \vec{n}$, this integral becomes :

$$\sum_1 \int_{S_1} \nabla_t \cdot (\vec{e}_z \times \nabla_t \psi_1) \, dS_1$$

But we know from vector calculus that :

$$\nabla_t \cdot (\vec{e}_z \cdot \nabla_t \psi_1) = \nabla_t \psi_i \cdot (\nabla_t \times \vec{e}_z) - \vec{e}_z \cdot (\nabla_t \times \nabla_t \psi_1) = 0$$

Therefore :

$$\sum_1 \int_{C_1} (\vec{\phi}_t^{\,1} \cdot \vec{n}) \, dl_i = 0$$

Using the same arguments, we find :

$$\sum_i \int_{C_1} \left(\frac{\varepsilon_i}{\varepsilon_0} \right) (\vec{\psi}_t^{\,i} \cdot \vec{n}) \, dl_i = 0$$

Appendix 2 : Details of the discretisation process.

By definition, we have :

$$[J_i] = \begin{bmatrix} \frac{\partial X_i}{\partial \xi} & \frac{\partial Y_i}{\partial \xi} \\ \frac{\partial X_i}{\partial \eta} & \frac{\partial Y_i}{\partial \eta} \end{bmatrix}$$

with : $X_i = \sum_{j=1}^{6} \bar{N}_j X_{ij}$ and $Y_i = \sum_{j=1}^{6} \bar{N}_j Y_{ij}$,

where $\bar{N}_j$ is one of the geometrical transformation functions, X_{ij} and Y_{ij} are the coordinates of the nodes "j" of the element "i".

Then, we can write :

$$J^i_{11} = \frac{\partial X_i}{\partial \xi} = \sum_{j=1}^{6} \frac{\partial \bar{N}_j}{\partial \xi} X_{ij}$$

$$J^i_{12} = \frac{\partial Y_i}{\partial \xi} = \sum_{j=1}^{6} \frac{\partial \bar{N}_j}{\partial \xi} Y_{ij}$$

$$J^i_{21} = \frac{\partial X_i}{\partial \eta} = \sum_{j=1}^{6} \frac{\partial \bar{N}_j}{\partial \eta} X_{ij}$$

$$J^i_{22} = \frac{\partial Y_i}{\partial \eta} = \sum_{j=1}^{6} \frac{\partial \bar{N}_j}{\partial \eta} Y_{ij}$$

Now, we compute $\nabla\phi_i$ in terms of the known parameters ($\nabla\psi_i$ is equivalent). We obtain :

$$\nabla\phi_i = \left[J_i^{-1} \right] \begin{Bmatrix} \frac{\partial \phi_i}{\partial \xi} \\ \frac{\partial \phi_i}{\partial \eta} \end{Bmatrix} = \frac{1}{|J_i|} \begin{bmatrix} J^i_{22} & -J^i_{12} \\ -J^i_{21} & J^i_{11} \end{bmatrix} \begin{Bmatrix} \frac{\partial \phi_i}{\partial \xi} \\ \frac{\partial \phi_i}{\partial \eta} \end{Bmatrix}$$

$\frac{\partial \phi_i}{\partial \xi}$ and $\frac{\partial \phi_i}{\partial \eta}$ can be written in terms of the nodal values ϕ_{ij} :

$$\phi_i = \sum_{j=1}^{6} N_j \phi_{ij} \quad \rightarrow \quad \frac{\partial \phi_i}{\partial \xi} = \sum_{j=1}^{6} \frac{\partial N_j}{\partial \xi} \phi_{ij} \quad \text{and}$$

$$\frac{\partial \phi_i}{\partial \eta} = \sum_{j=1}^{6} \frac{\partial N_j}{\partial \eta} \phi_{ij}$$

Putting these results back in the expression of $\nabla \phi_i$, we can write all the quantities needed in the variational form. We obtain :

$$|\nabla \phi_i|^2 = \frac{1}{|J_i|^2} \sum_{j=1}^{6} \sum_{k=1}^{6} \{ (J_{22}^{i2}+J_{21}^{i2}) \frac{\partial N_j}{\partial \xi} \frac{\partial N_k}{\partial \xi} +$$

$$(J_{12}^{i2}+J_{11}^{i2}) \frac{\partial N_j}{\partial \eta} \frac{\partial N_k}{\partial \eta} - (J_{12}^{i} J_{22}^{i}+J_{21}^{i} J_{11}^{i}) (\frac{\partial N_j}{\partial \xi} \frac{\partial N_k}{\partial \eta} + \frac{\partial N_k}{\partial \xi} \frac{\partial N_j}{\partial \eta}) \} \phi_{ij} \phi_{ik}$$

$|\nabla \psi_i|^2$ is similar, replacing $\phi_{ij}\phi_{ik}$ by $\psi_{ij}\psi_{ik}$. The coupling term $\vec{e}_z \cdot (\nabla \psi_i \times \nabla \phi_i)$ becomes :

$$\frac{1}{|J_i|} \{ \sum_{j=1}^{6} \sum_{k=1}^{6} (\frac{\partial N_j}{\partial \xi} \frac{\partial N_k}{\partial \eta} - \frac{\partial N_k}{\partial \xi} \frac{\partial N_j}{\partial \eta}) \psi_{ij} \phi_{ik} \}$$

and the other terms are :

$$|\phi_i|^2 = \sum_{j=1}^{6} \sum_{k=1}^{6} N_j N_k \phi_{ij} \phi_{ik}$$

$$|\psi_i|^2 = \sum_{j=1}^{6} \sum_{k=1}^{6} N_j N_k \psi_{ij} \psi_{ik}$$

Finally, the total integral may be written :

$$\delta \sum_{i=1}^{N} \int_0^1 d\xi \int_0^{1-\xi} d\eta \sum_{j=1}^{6} \sum_{k=1}^{6} \{ \tau_i U_{jk}^{i} \phi_{ij} \phi_{ik} + \gamma^2 \tau_i (\frac{\varepsilon_i}{\varepsilon_0}) U_{jk}^{i} \psi_{ij} \psi_{ik} +$$

$$\gamma^2 \tau_i V_{jk} \psi_{ij} \phi_{ik} + \gamma^2 \tau_i V_{jk} \phi_{ik} \psi_{ij} + k_0^2 W_{jk}^{i} \phi_{ij} \phi_{ik} + k_0^2 \gamma^2 (\frac{\varepsilon_i}{\varepsilon_0}) W_{jk}^{i} \psi_{ij} \psi_{ik} \} = 0$$

where U^i_{jk}, V_{jk} and W^i_{jk} are defined in page 21.

Using numerical integration on the triangle, we obtain :

$$\delta \sum_{i=1}^{N} \tau_i \sum_{j=1}^{6}\sum_{k=1}^{6} \{ \sum_{R=1}^{NI} w_R U^i_{jk}(\xi_R,\eta_R) \} \phi_{ij}\phi_{ik} +$$

$$\gamma^2 \tau_i(\tfrac{\varepsilon_i}{\varepsilon_0}) \sum_{j=1}^{6}\sum_{k=1}^{6} \{ \sum_{R=1}^{NI} w_R U^i_{jk}(\xi_R,\eta_R) \} \psi_{ij}\psi_{ik} +$$

$$\gamma^2 \tau_i \sum_{j=1}^{6}\sum_{k=1}^{6} \{ \sum_{R=1}^{NI} w_R V_{jk}(\xi_R,\eta_R) \} \psi_{ij}\phi_{ik} +$$

$$\gamma^2 \tau_i \sum_{j=1}^{6}\sum_{k=1}^{6} \{ \sum_{R=1}^{NI} w_R V_{jk}(\xi_R,\eta_R) \} \phi_{ik}\psi_{ij} +$$

$$k_0^2 \sum_{j=1}^{6}\sum_{k=1}^{6} \{ \sum_{R=1}^{NI} w_R W^i_{jk}(\xi_R,\eta_R) \} \phi_{ij}\phi_{ik} +$$

$$k_0^2\gamma^2(\tfrac{\varepsilon_i}{\varepsilon_0}) \sum_{j=1}^{6}\sum_{k=1}^{6} \{ \sum_{R=1}^{NI} w_R W^i_{jk}(\xi_R,\eta_R) \} \psi_{ij}\psi_{ik} = 0$$

The decomposition :

$$2\gamma^2\tau_i \sum_{j}\sum_{k}\sum_{R} w_R V_{jk}\, \psi_{ij}\phi_{ik} \qquad \text{in :}$$

$$\gamma^2\tau_i \sum_{j}\sum_{k}\sum_{R} \ldots \psi_{ij}\phi_{ik} + \gamma^2\tau_i \sum_{j}\sum_{k}\sum_{R} \ldots \phi_{ik}\psi_{ij}$$

is necessary to obtain a matrix formulation which is fully SYMMETRIC. We can easily prove from matrix algebra that the above formulation in terms of double summations is exactly equivalent to the matrix equations given in page 20.

Appendix 3 : Analytical solution of the dielectric rod problem.

The dielectric rod is an infinitely long circular dielectric waveguide (see Fig. 5) such that ε_1 is greater then ε_2.

We have, by definition :

$$n_1^2 = \frac{\varepsilon_1}{\varepsilon_0} \quad \text{and} \quad n_2^2 = \frac{\varepsilon_2}{\varepsilon_0} .$$

ε_0 is the free space permittivity.

Solving the Helmholtz equation with the separation of variables method (ref. 3), we obtain the z-components of the field in the inner and outer media.

The interior solution ($r < a$) is :

$$E_{z1} = A_n^1 J_n(k_1 r) \cos(n\phi) e^{j(\omega t - \beta z)}$$

$$H_{z1} = B_n^1 J_n(k_1 r) \sin(n\phi) e^{j(\omega t - \beta z)}$$

with $k_1^2 = (\frac{\omega}{c})^2 (n_1^2 - \gamma^2)$;

The exterior solution ($r > a$) is :

$$E_{z2} = A_n^2 K_n(k_2 r) \cos(n\phi) e^{j(\omega t - \beta z)}$$

$$H_{z2} = B_n^2 K_n(k_2 r) \sin(n\phi) e^{j(\omega t - \beta z)}$$

with $k_2^2 = (\frac{\omega}{c})^2 (\gamma^2 - n_2^2)$;

and : n = label of the mode,

$$\gamma = \frac{\beta c}{\omega}$$

The ϕ-components and r-components may be deduced from the z-components using equations (2.12-2.15) and the usual rectangular to cylindrical coordinates transformations.

Only the ϕ-components are needed to match the tangential field components at the interface between the two media, when $r = a$.

We have :

$$E_{z1}(a) = E_{z2}(a) \quad ; \; E_{\phi 1}(a) = E_{\phi 2}(a) \quad \text{and}$$

$$H_{z1}(a) = H_{z2}(a) \quad ; \; H_{\phi 1}(a) = H_{\phi 2}(a)$$

where :

$$E_{\phi m} = \frac{j}{k_m^2} \left\{ \left(\frac{\omega\gamma}{c} \right) \frac{1}{r} \frac{\partial E_{zm}}{\partial \phi} - \omega\mu_0 \frac{\partial H_{zm}}{\partial r} \right\}$$

$$H_{\phi m} = \frac{j}{k_m^2} \left\{ \omega\bar{\epsilon}_m \frac{\partial E_{zm}}{\partial r} + \left(\frac{\omega\gamma}{c} \right) \frac{1}{r} \frac{\partial H_{zm}}{\partial r} \right\}$$

with : $m = 1$ for the interior solution, 2 for the exterior solution.

μ_0 is the free space permeability.

$\bar{\epsilon}_m$ is the relative permittivity of the medium "m".

The above equations lead to a matrix equation.

Using some straight forward algebraic manipulations, we obtain :

$$\begin{bmatrix} n\gamma \left(\frac{1}{S_1^2} + \frac{1}{S_2^2} \right) & \frac{K_n'(S_2)}{S_2 K_n(S_2)} + \frac{J_n'(S_1)}{S_1 J_n(S_1)} \\ \frac{n_2^2 K_n'(S_2)}{S_2 K_n(S_2)} + \frac{n_1^2 J_n'(S_1)}{S_1 J_n(S_1)} & n\gamma \left(\frac{1}{S_1^2} + \frac{1}{S_2^2} \right) \end{bmatrix} \begin{bmatrix} A_n^1 \\ Z_0 B_n^1 \end{bmatrix} = 0$$

with : $S = \left(\frac{\omega a}{c} \right)$,

$S_1 = S\sqrt{n_1^2 - \gamma^2}$ and $S_2 = S\sqrt{\gamma^2 - n_2^2}$,

$Z_0 = \eta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$

J_n' and K_n' are the derivative of the Bessel functions J_n and K_n.

This system has a solution only if its determinant is zero, which leads to the following equation in γ and S :

$$\{ n\gamma(\frac{1}{S_1^2} + \frac{1}{S_2^2}) \}^2 - \{ n(\frac{1}{S_1^2} + \frac{1}{S_2^2}) - \frac{K_{n+1}(S_2)}{S_2 K_n(S_2)} - \frac{J_{n+1}(S_1)}{S_1 J_n(S_1)} \} \cdot$$

$$\{ n(\frac{n_2^2}{S_2^2} + \frac{n_1^2}{S_1^2}) - \frac{n_2^2 K_{n+1}(S_2)}{S_2 K_n(S_2)} - \frac{n_1^2 J_{n+1}(S_1)}{S_1 J_n(S_1)} \} = 0$$

The solutions of this transcendental equation give the propagation characteristics of the rod. For a given value of A_n^1 (or B_n^1), the other three coefficients A_n^2 , B_n^2 , B_n^1 (or A_n^1) can be computed.

The arbitrary coefficient simply describes the power (or the field strength) which travels inside the waveguide.

Appendix 4 : A simple method to find the extremum inside the reference element.

As we have seen in section 6.4.2, the field in the reference element can always be written :

$$\phi(\xi,\eta) = A\xi^2 + B\eta\xi + C\eta^2 + D\xi + E\eta + F$$

where A,B,C,D,E and F are computed from the nodal values of ϕ.

Because $\nabla\phi$ is linear in ξ and η, the extremum of ϕ over the element must be located at one of the following points:

1) The stationary point of the quadratic IF IT BELONGS to the reference element. In that case, we have :

$$\nabla\phi\ (\xi_0,\eta_0) = 0 \qquad \xi_0 = \frac{2CD - BE}{B^2 - 4AC}$$

$$\eta_0 = \frac{2AE - BD}{B^2 - 4AC}$$

and : $B^2 - 4AC \neq 0$, $0 \leq \xi_0 \leq 1$, $0 \leq \eta_0 \leq 1-\xi_0$

2) The stationary point of the quadratic on the boundary $\eta = 0$.

In that case, we obtain :

$$\xi_1 = -\frac{D}{2A} \quad , 0 \leq \xi_1 \leq 1 \quad \text{and } A \neq 0$$

3) The stationary point of the quadratic on the boundary $\xi = 0$.

In that case, we obtain :

$$\eta_2 = -\frac{E}{2C} \quad , 0 \leq \eta_2 \leq 1 \quad \text{and } C \neq 0$$

4) The stationary point of the quadratic on the boundary $\xi+\eta = 1$.

In that case, we obtain :

$$\xi_3 = -\{\frac{B - 2C + D - E}{2(A - B + C)}\} \quad , \eta_3 = 1 - \xi_3 \quad \text{and :}$$

$A - B + C \neq 0$, $0 \leq \xi_3 \leq 1$.

5) A vertex of the reference element.

The maximum (minimum) of the function (field) is the maximum (minimum) of the field values at these 7 points.

The subroutine "EXTREM" (see section 6.4.2) simply computes the field intensity at these points and returns to the main program the maximium (minimum) of these.

Appendix 5 : A variational formulation which removes the discontinuity at τ_i.

From ref. 3 , we can write :

$$\nabla_t \psi_z = j \left(\frac{\omega}{c}\right) \left(-\gamma \vec{\psi}_t + \frac{1}{\gamma} \vec{e}_z x \vec{\phi}_t \right)$$

$$\nabla_t \phi_z = - j \left(\frac{\omega}{c}\right) \left(\gamma \vec{\phi}_t + \bar{\varepsilon}_i \gamma \vec{e}_z x \vec{\psi}_t \right)$$

where : ψ_z , ϕ_z = z-components of the field

$\vec{\psi}_t$, $\vec{\phi}_t$ = transverse field,

$\bar{\varepsilon}_i$ = relative permittivity in the region of interest.

In cartesian coordinates, we have :

$$\vec{\psi}_t = \vec{x}\psi_x + \vec{y}\psi_y \quad \text{and} \quad \vec{\phi}_t = \vec{x}\phi_x + \vec{y}\phi_y .$$

Therefore, we can write :

$$\nabla_t \psi_z = j \left(\frac{\omega}{c}\right) \{ \vec{x}\left(-\gamma\psi_x - \frac{1}{\gamma}\phi_y \right) + \vec{y}\left(-\gamma\psi_y + \frac{1}{\gamma}\phi_x \right) \}$$

$$\nabla_t \phi_z = - j \left(\frac{\omega}{c}\right) \{ \vec{x}\left(\gamma\phi_x - \bar{\varepsilon}_i\gamma\psi_y \right) + \vec{y}\left(\gamma\phi_y + \bar{\varepsilon}_i\gamma\psi_x \right) \}$$

With some mathematical manipulations, we easily find :

$$\tau_i \{\gamma^2 \bar{\varepsilon}_i |\nabla_t \psi_z|^2 + |\nabla_t \phi_z|^2 + 2\gamma^2 \vec{e}_z . (\nabla\psi_z \times \nabla\phi_z) \} =$$

$$\tau_i (\bar{\varepsilon}_i - \gamma^2) \left(\frac{\omega}{c}\right)^2 \{\phi_x^2 + \bar{\varepsilon}_i\gamma^2\psi_y^2 + \phi_y^2 + \bar{\varepsilon}_i\gamma^2\psi_x^2 + 2\gamma^2(\phi_y\psi_x - \phi_x\psi_y) \}$$

But $\tau_i = \dfrac{\gamma^2 - 1}{\gamma^2 - \bar{\varepsilon}_i}$,

and this expression becomes :

$$(1 - \gamma^2) \left(\frac{\omega}{c}\right)^2 \{ |\vec{\phi}_t|^2 + |\vec{\psi}_t|^2 \bar{\varepsilon}_i\gamma^2 + 2\gamma^2 \vec{e}_z . (\vec{\psi}_t \times \vec{\phi}_t) \}$$

The discontinuity at τ_i is suppressed.

REFERENCES

1. Thomas G. Giallorenzi, Joseph A. Bucaro, Anthony Dandridge, G.H. Sigel, JR., James H. Cole, Scott C. Rashleigh, Richard G. Priest
"Optical Fiber Sensor Technology"
IEEE Transaction on Microwave and Techniques (MTT)
Vol. MTT-30, No 4, April 1982, pp. 473-480.

2. Reference [1], p. 500

3. J.D. Jackson
"Classical Electrodynamics", 2nd Edition
John Wiley and Sons, New York, 1975, p.340

4. G.L. Yip, Y.H. Amhew
"Propagation Characteristics of Radially Inhomogeneous Optical Fibers"
Electronics Letters, 21st February 1974, Vol. 10, No 4

5. H.H. Yao, G.L. Yip
"Numerical Study of Inhomogeneous Optical Waveguide Problems using Predictor-Corrector Method"
Electronics Letters, 19th March 1981, Vol. 17, No 6, pp.229-230

6. Gouri Dhatt, Gilbert Touzot
"Une Presentation de la Methode des Elements Finis"
Les Presses de l'Universite Laval, Quebec
Maloine S.A. Editeur, Paris, 1981

7. O.C. Zienkiewicz
"The Finite Element Method for Engineering Sciences"
McGraw Hill, New York, 3rd Edition, 1977

8. K.J. Bathe and E.L. Wilson
"Numerical Methods in Finite Element Analysis"
Prentice Hall, 1976

9. P. Silvester and P. Rafinejad
"Curvilinear Finite Elements for Two Dimensional Saturable Magnetic Fields"
IEEE meeting, New York, Jan 27-Feb 1 1974, Paper T74 214-3

10. S. Ahmed, P. Daly
"Finite Element Method For Inhomogeneous Waveguides"
Proc. IEE, Vol. 116, No 10, October 1969, pp. 1661-1664

11. Zoltan J. Csendes, P. Silvester
"Numerical Solution of Dielectric Loaded Waveguide : I-Finite-Element Analysis"
IEEE Vol MTT-48, No 12, December 1970, pp 1124-1131

12. C. Yeh, S.B. Dong and W. Oliver
"Arbitrarily Shaped Inhomogeneous Optical Fibers or Integrated Optical Waveguides"
Journal of Applied Physics, Vol. 46, No 5, May 1975, p. 2126

13. C. Yeh, K. Ha, S.B. Dong, W.P. Brown
"Single-mode Optical Waveguides"
Applied Optics, Vol. 18, No 10, 15 May 1979, pp. 1490-1495

14. N.Mabaya, P.E. Lagasse, P.Vandenbulcke
"Finite Element Analysis of Optical Waveguides"
IEEE Vol. MTT-29, No 6, June 1981, pp 600-605

15. Joseph Katz
"Novel Solution of 2-D Waveguides Using the Finite Element Method"
Applied Optics, Vol 21, No 15, 1 August 1982, pp 2747-2750

16. B.M. Azizur Rahman and J. Brian Davies
"Finite Element Analysis of Optical and Microwave Waveguide Problems"
IEEE Vol. MTT-32, No 1, January 1984, pp 20-28

17. Murray R. Spiegel
"Formules et Tables de Mathematiques"
Serie Schaum, McGraw Hill, New York, Paris, 1974

18. Michael K. Barnoski
"Fundamentals of Optical Fiber Communications"
Academic Press, 2nd Edition, 1981, p. 16

19. Ahmew Y.H.
"Propagation Characteristics of the Self Focusing Fibre Waveguide"
M.Eng. Thesis
Department of Electrical Engineering, Mc Gill University
Montreal, Canada

20. M.S. Sodha and A.K. Ghatak
"Inhomogeneous Optical Waveguides"
Indian Institute of Technology, New Delhi, India
Plenum Press, New York and London, 1971

21. S.R. Rengarajan, J.E. Lewis
"Single Mode Propagation in Multi-layer Elliptical Fiber Waveguide"
Radio Science, Vol. 16, No 4, July-August 1981, pp. 541-547

22. J. Martucci
"Propagation and Loss Characteristics of Cladded Optical Fibers"
M.Eng. Thesis
Department of Electrical Engineering, Mc Gill University,
Montreal, Canada, 1973

23. Masatoshi Ikeuchi, Hideo Sawami, and Hiroshi Niki
"Analysis of Open-Type Dielectric Waveguides by the Finite-Element Iterative Method"
IEEE Vol. MTT-29, No 3, March 1981, pp. 234-239

24. Jacques Bures, Suzanne Lacroix et Jean Lapierre
"Analyse d'un Coupleur Bidirectionnel a Fibres Optiques monomodes fusionnées"
Applied Optics, Vol. 22, June 15 1983, pp. 1918-1922

25. Allan W. Snyder
"Coupled Mode Theory for Optical Fibers"
Journal of The Optical Society of America, Vol. 62, No 11, November 1972, pp. 1267-1277

26. Mitsuo Hano
"Finite Element Analysis of Dielectric-Loaded Waveguides"
IEEE Vol. MTT-32, No 10, October 1984, pp. 1275-1279