

Penalty Function Improvement of Waveguide Solution by Finite Elements

B. M. AZIZUR RAHMAN MEMBER, IEEE, AND J. BRIAN DAVIES, MEMBER, IEEE

Abstract—The finite element method is a well-established method for the solution of a wide range of guided wave problems. One drawback associated with the powerful vector formulation is the appearance of spurious or nonphysical solutions. A penalty function method has been introduced to the finite element formulation, to reduce or eliminate spurious solutions. It also improves the quality of the physical field solutions. The method has been applied for the solution of metallic homogeneous and inhomogeneous guides, and integrated optics guides.

I. INTRODUCTION

THE FINITE ELEMENT method has become a powerful tool throughout engineering for its flexibility and versatility. This is a powerful technique for complicated electromagnetic, structural, heat, or fluid flow problems. There are a few different types of variational formulations which are available to solve electromagnetic wave propagation problems. A scalar variational formulation [1] which is simplest among them, is limited in scope, not being suitable (except as an approximation) for general inhomogeneous and general anisotropic problems. There are different types of vector variational formulations such as E_z/H_z , E , H , and $E + H$, and these are suitable for a wide range of practical complicated waveguides. However, these vector finite element solutions have been known to include nonphysical spurious solutions [2]–[12]. Variational formulations using transverse field components [13], [14], solved by the Rayleigh-Ritz method or method of moments, do not have spurious solutions, but these formulations are not valid for general anisotropic materials, the functionals are not self adjoint, and because of the additional differentiation involved with them, these are not very attractive for their implementation in a finite element method. The non-appearance of spurious solution in this formulation may be due to the divergence free basis functions.

Though the existence of spurious solutions has been quoted by many authors, little work has been done to reduce their number or to eliminate them. Konrad [5] suggested imposing boundary conditions rigorously to eliminate them, but that did not prove to be adequate [11], [12]. Mabaya *et al.* [6] have reduced the number of spurious solution for E_z/H_z formulation, by explicitly enforcing the continuity of the tangential components of the trans-

verse fields at the interfaces by using Lagrange multipliers. The disadvantage of that method was the greatly increased complexity of the program and of the increased numerical operations necessary to enforce those continuity conditions; also this E_z/H_z formulation is limited to a diagonal permittivity tensor. In this paper we will describe use of a penalty method to reduce or eliminate spurious solutions. A bonus of the penalty method, as reported in other applications [15], is found in the significantly improved smoothness of the resulting field.

II. SPURIOUS SOLUTIONS

The most serious difficulty in using a vector finite element analysis is the appearance of extraneous nonphysical or spurious modes [2]–[12]. Spurious solutions do not exist in a scalar formulation because the operator is positive definite, but with a vector finite element method the operator is no longer positive definite. In defining a vector field fully, we need to define both the curl and the divergence of the vector field at every point in space. In the conventional vector field finite element formulations, their Euler equations satisfy Helmholtz's equations but do not satisfy $\text{div } \mathbf{B} = 0$. This may cause the system to be excessively flexible, which in turn is believed [11] to be responsible for spurious modes.

Spurious solutions are found to spread all over the eigenvalue spectrum, some of them appearing below any true modes and some between the physical modes. As one increases the mesh refinement to improve the accuracy of the solution the number of spurious solutions also increases.

If one wants to compute a set of eigenmodes, it is difficult and quite cumbersome to distinguish between the spurious and the physical modes of the guide. Spurious modes can be identified by observing their dispersion curve [2], [4]. Another simple way to identify them is to observe their eigenvectors. In the case of a nonphysical mode the fields vary in an unreasonable, sometimes random fashion over the guide cross section, so that a plot of the field solution can also provide a good, if somewhat tedious test [3]. They can also be checked by observing convergence of the solution with mesh refinements. Because of the inconsistent spatial variations of the field for the spurious eigenmodes, if one calculates its $\text{div } \mathbf{H}$ it will be quite high

Manuscript received November 14, 1983; revised March 19, 1984.

The authors are with the Department of Electronic and Electrical Engineering, University College London, London WC1 7JE, England

compared to that of a physical eigenmode. This criteria has been successfully used [11], [12] to filter out most of the spurious solutions from a set of calculated eigenvectors.

Rather than calculate the spectrum and eigenvectors and then filter out, *a posteriori*, any spurious solution, the strategy of this new work is to apply the constraint *a priori*; spurious modes should then be removed rather than just recognized. We have added the least squares constraint of satisfying $\text{div} \mathbf{H}$ to the original functional J so that explicitly the Euler equations are the Helmholtz equation plus the vanishing of $\text{div} \mathbf{H}$.

III. PENALTY METHOD

The penalty method [16], [17] can impose a specific constraint on certain solution variables, and it has been used in structural engineering problems to impose specific boundary conditions [16]. This method has also been successfully used to eliminate spurious solutions in acoustic guide problems [15]. For our electromagnetic problem we add a functional whose Euler equation is the satisfaction of $\text{div} \mathbf{B} = 0$.

We have considered the penalty method, used with the full vector \mathbf{H} field formulation. This formulation has been chosen because of its accuracy near cutoff, ability to consider general lossless arbitrary tensor permittivity, and natural field continuity across the interelement boundaries [12]. This often quoted Berk's [18] formulation can be written as [5], [10]–[12]

$$\omega^2 = \frac{\int (\nabla \times \mathbf{H})^* \cdot \hat{\epsilon}^{-1} \cdot (\nabla \times \mathbf{H}) d\Omega}{\int \mathbf{H}^* \cdot \hat{\mu} \mathbf{H} d\Omega} \quad (1)$$

When we add the penalty constraint, the augmented functional can be written as

$$\omega^2 = \frac{\int (\nabla \times \mathbf{H})^* \cdot \hat{\epsilon}^{-1} (\nabla \times \mathbf{H}) d\Omega + (\alpha/\epsilon_0) \int (\nabla \cdot \mathbf{H})^* \cdot (\nabla \cdot \mathbf{H}) d\Omega}{\int \mathbf{H}^* \cdot \hat{\mu} \mathbf{H} d\Omega} \quad (2)$$

where α is the penalty number. This equation has the desired additional Euler equation $\text{div} \mathbf{B} = 0$, in precisely the manner that terms are added [18] to give desired boundary conditions. The constraint is imposed in a least square sense: the larger the value of penalty number, the heavier is the constraint imposed on the corresponding Euler equation. In using the penalty method, one important consideration is the choice of an appropriate penalty number; this will be discussed in the next section. This imposition of the constraint can be compared with the changing of natural boundary conditions of an Euler equation by introducing an additional appropriate integrand [18]. One advantage of using the penalty method is that it does not increase the matrix order, and the additional computational time is negligible.

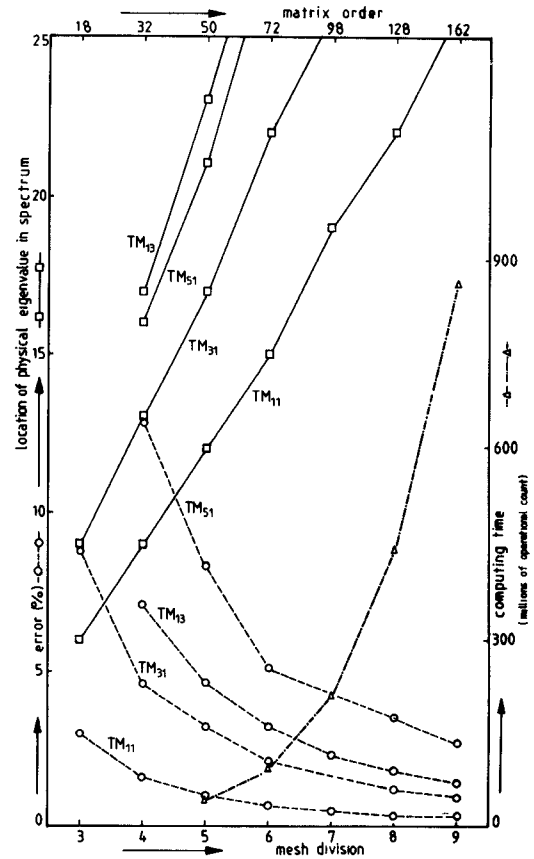


Fig. 1. Location of physical eigenvalues in the spectrum, percentage of cutoff frequency error, and computing time as function of mesh division and matrix order.

IV. RESULTS

We have used the penalty method to eliminate or reduce spurious solutions to various problems, the first being homogeneous and then inhomogeneous-loaded metal waveguide where we know the exact solutions. We then continue to some integrated optic structures.

A. Hollow Rectangular Metal Guide

We have chosen this first problem because of its simple closed form solutions. Applying the finite element method [1], the accuracy of the solution depends on the mesh refinement and order of the shape functions. For higher accuracy we need finer mesh divisions or higher order shape functions, but increasing either the number of mesh divisions or the order of shape functions increases the number of spurious solutions. The effect of mesh has been shown in Fig. 1 for a rectangular guide considering only one-quarter of the guide (using two symmetric magnetic wall planes). Using first degree shape functions, the abscissa denotes the number of mesh divisions in each coordinate (and the corresponding matrix order); the right-hand

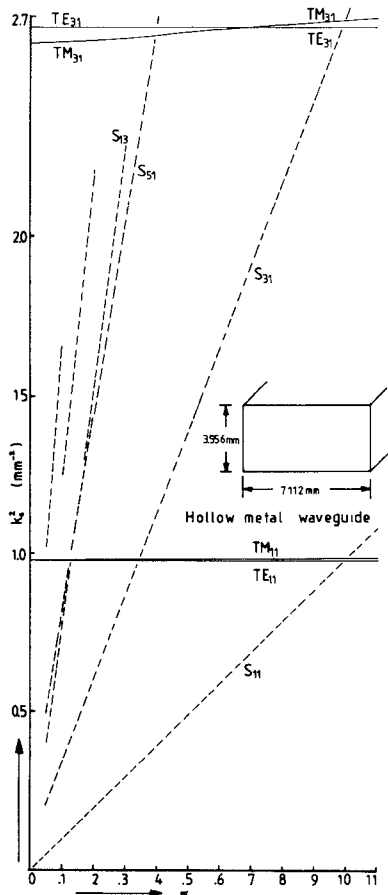


Fig. 2. Variation of k_c^2 with penalty term for hollow rectangular waveguide, finite element division 7×7 .

ordinate gives the computing time (in millions of operational counts) and the left hand gives, for the first four TM modes, the resulting cutoff frequency error and its location (order) in the computed spectrum.

Fig. 2 shows the variation of computed eigenvalue (k_c^2) for TE and TM modes at cutoff with the penalty parameter (for mesh division = 7). All the spurious solutions in this part of the spectrum come from the origin, varying almost linearly with α and having quite large but different slopes. Eigenvalues for true TM modes increase very slowly with penalty parameter, but for TE mode do not vary with α (at cutoff). In Fig. 3 we can observe the locations in the spectrum of the TE_{11} and TM_{11} modes, as α is increased from zero, where the locations are 19 and 20, respectively. At α equal to 0.5, the TE_{11} and TM_{11} mode now appear at 2 and 3 eigenlocation, respectively; i.e., there is only one spurious mode before the first two physical modes, instead of 18 when no penalty term has been added. Similarly we can see at $\alpha=1.1$ there is no spurious solution before the first two eigenvalues. We can also observe that there are 7 spurious modes between TM_{11} and next physical mode TM_{31} for α zero, whereas at $\alpha=1.1$ there is only one spurious mode among the first four physical modes. Looking at Figs. 2 and 3, we see how one can eliminate spurious modes among any number of dominant physical modes. The error associated with the physical modes are quite low, particularly for the lower order modes. For TE modes at

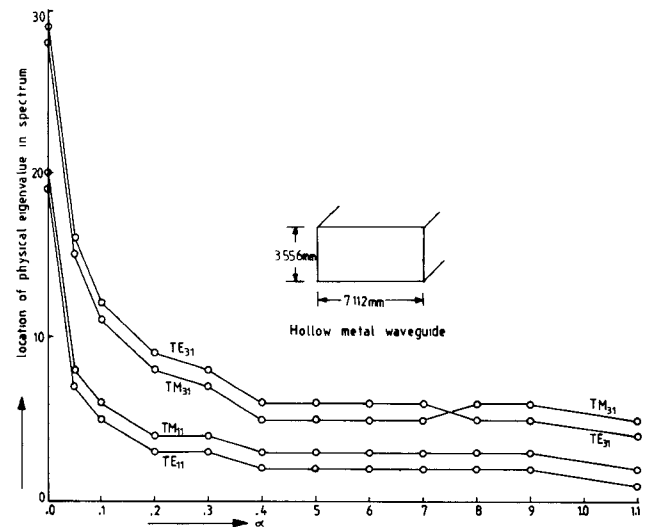


Fig. 3. Reduction of spurious solutions with penalty term for hollow rectangular waveguide, mesh division 7×7 .

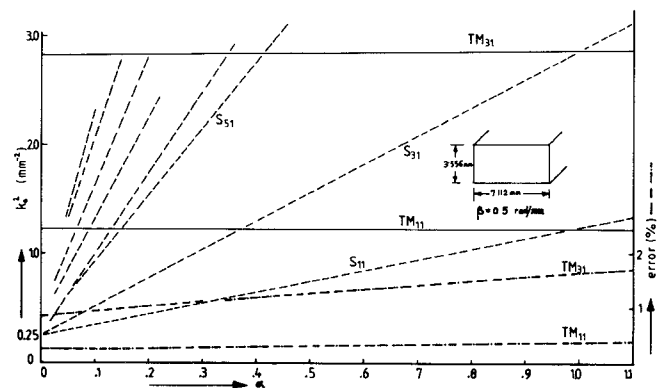


Fig. 4. Variation of k_0^2 with penalty term for hollow rectangular waveguide, $\beta = 0.5$ rad/mm, mesh division 9×9 .

cutoff there is exceptionally no addition of error from the penalty term as the additional term of (2) is identically zero in this formulation at cutoff. This is because the TE mode at cutoff has a longitudinal component only of H field, to give a divergence-free field.

Fig. 4 shows the variation of k_0^2 with penalty parameter α . Here we calculate the TM modes above cutoff, at a fixed propagation constant = 0.5 rad/mm and using a mesh division of 9. Here again, as in Fig. 2, k_0^2 varies very slowly with α for physical modes but rapidly for spurious modes. It is of interest to examine the field solution for these spurious modes, and for the above problem with α zero, the magnetic fields are shown of the physical TM_{11} mode in Fig. 5 and of the spurious ' S_{11} ' mode in Fig. 6. In these figures we have considered only one-quarter of the guide using two-fold symmetries with the two magnetic walls. For the TM_{11} mode, the H_x and H_y spatial variations and their vector combinations are consistent, but not for the S_{11} mode. Specifically $\delta H_x / \delta x$ and $\delta H_y / \delta y$ have opposite signs throughout the region for the TM_{11} mode but the same sign for the S_{11} . The latter solution is therefore incapable of satisfying $\text{div } \mathbf{B} = 0$, and so is not physically admissible.

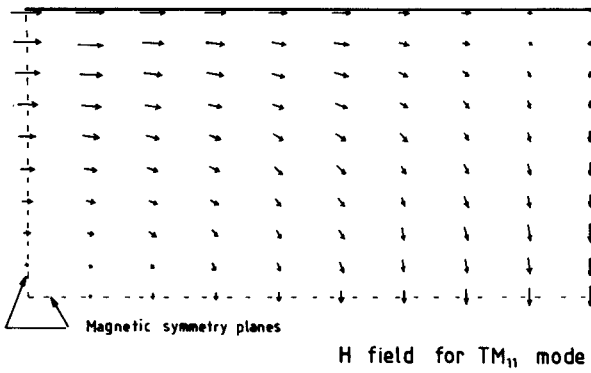


Fig. 5. H field vector for TM_{11} mode, in one-quarter of the hollow guide (considering two magnetic symmetry planes), mesh division 9×9 .

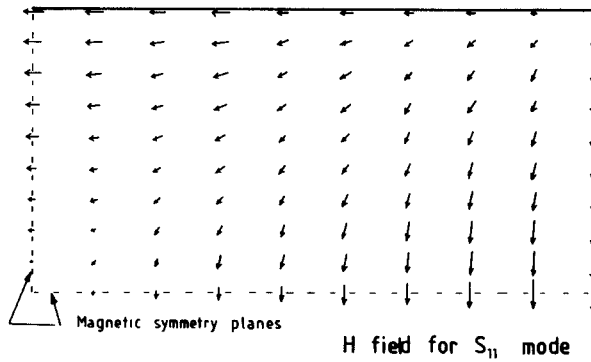


Fig. 6. H field vector for spurious ' S_{11} ' mode, in one-quarter of the hollow guide (considering two magnetic symmetry walls), mesh division 9×9 .

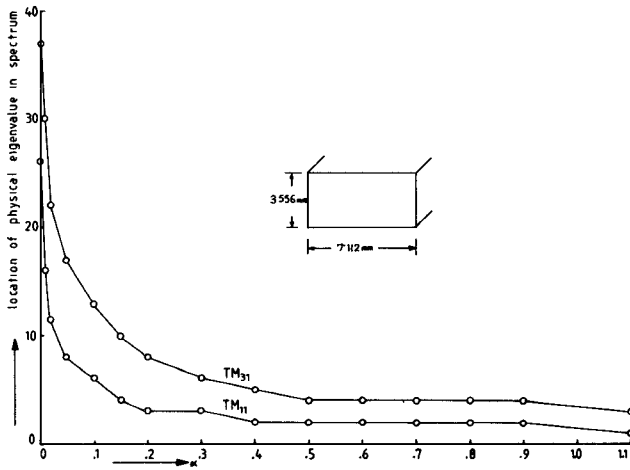


Fig. 7. Reduction of spurious solution with penalty term, for hollow rectangular waveguide, mesh division 9×9 .

Fig. 7 shows the reduction of spurious solutions with α for this same mesh refinement of Figs. 4–6. Fig. 4 also shows the error introduced by this penalty method. At $\alpha = 0$ the percentage of error was 0.27 percent for the TM_{11} mode (due to mesh divisions, word length, etc.). At $\alpha = 0.5$ the increase in error due to this penalty method was only 0.05 percent and the error increases by about 0.13 percent at $\alpha = 1.1$ for the same TM_{11} mode. Fig. 8 shows the variation of H_x and H_y field along the x direction (for H_x field at the center of the guide, for H_y field at the top edge

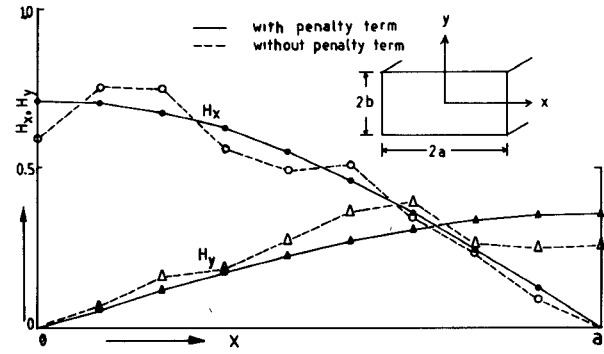


Fig. 8. H_x and H_y field variation with x direction, mesh division 9×9 .

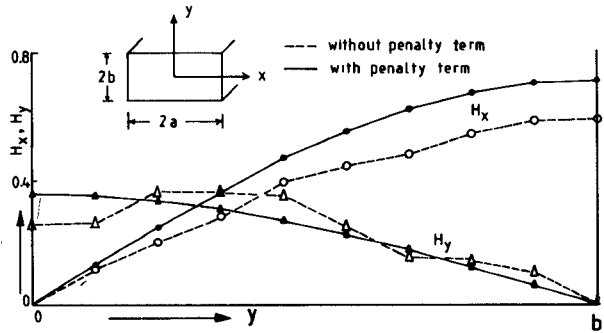


Fig. 9. H_x and H_y field variation with y direction, mesh division 9×9 .

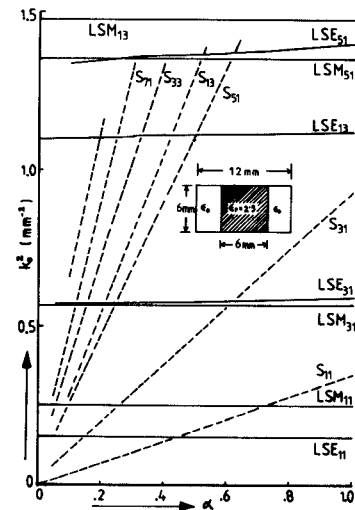


Fig. 10. Variation of k_c^2 with penalty term for inhomogeneously loaded metal guide, mesh division 8×8 .

of the guide) for $\alpha = 0$ and $\alpha = 0.5$, while Fig. 9 shows the variation along the y direction for $\alpha = 0$ and $\alpha = 0.5$. These figures showed dramatic improvements in smoothness of eigenvectors when applying the penalty functional—a feature noted in other applications of the penalty function [15].

B. Inhomogeneously Loaded Metal Guide

A rectangular inhomogeneous waveguide problem has been solved using the vector H field finite element formulation. This problem involves a material discontinuity which gives rise to different wavenumbers of k_x and k_y in the

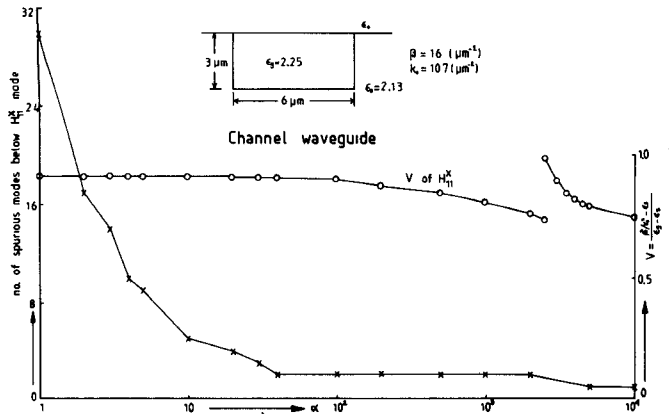


Fig. 13. Reduction of spurious solution and variation of V for channel waveguide with penalty term, total mesh division 18×19 .

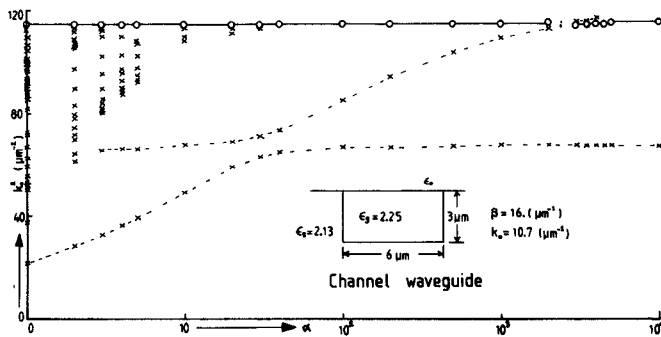


Fig. 14. Eigenvalues for channel waveguide with penalty term, total mesh division 18×19 .

index = 2.25, or $V=1.0$), and Fig. 15 shows the reduction of spurious solutions in that range. We know that with zero penalty parameter, most of the spurious solutions appear at the beginning of the spectrum, which is outside the above mentioned range.

Again, as illustrated in Figs. 8 and 9, we found a major benefit in using the penalty method was the improvement of eigenvectors. Above a small penalty parameter, all the physical propagating eigenmodes were clearly identified for the given symmetry plane, whereas without the penalty parameter many of them were unrecognizable. The dispersion characteristics are shown in Fig. 16 for the complete set of propagating eigenmodes for this channel waveguide.

V. COMPUTATIONAL REMARKS

A vector \mathbf{H} field finite element formulation has been used for all these solutions. For channel waveguide, infinite elements [12] have been added to orthodox finite element representation to extend the problem's physical domain to infinity. For smaller order eigenvalue problems we have used the NAG [20] F02AEF routine (Householder reduction and QL algorithm), but for larger order problems (order more than 200) we have used an efficient sparse routine, being a modification of an earlier complex version [11], but now exploiting the real symmetric properties of the matrices. For a matrix order of 1120 this routine takes less than 2.9 percent of the storage and only about 0.17 percent of the computational time to calculate any 5 eigen-

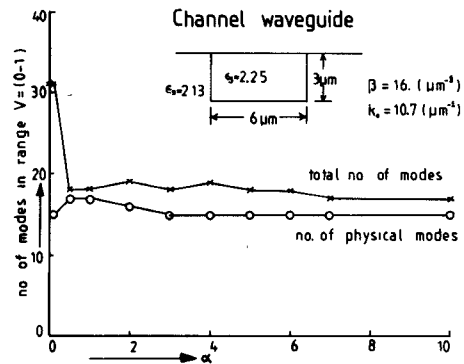


Fig. 15. Total eigenmodes and recognized eigenmodes for channel waveguide in the range $V=0$ to $V=1$, with penalty term, total mesh division 18×19 .

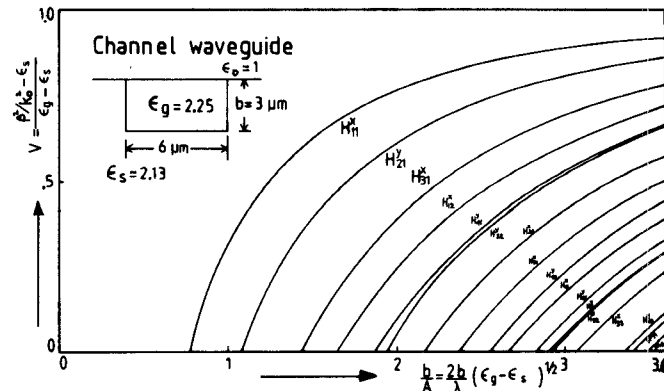


Fig. 13 shows the same for channel guide. Fig. 16 represents a set of dispersion characteristics for integrated optic channel waveguide. It was possible to recognize all the physical eigenmodes for this guide (with magnetic wall symmetry) when using the penalty parameter. Without the penalty many of the higher order modes were not recognizable properly for the same mesh refinement.

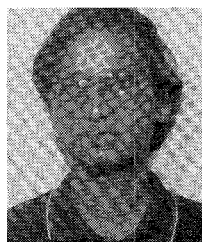
Since completing this work, the authors have come across similar studies in elastic waves of coupled fluid/solid systems [21]. Briefly, a penalty term is added for $|\text{curl } \mathbf{v}|^2$ over the fluid like our $|\text{div } \mathbf{H}|^2$ term of (2). No reference is made to any improvement of the field solutions, but interestingly, their results for eigenvalue dependence on penalty function have precisely the form of our Figs. 2, 4, and 10. Like our \mathbf{H} field plot of a spurious mode in Fig. 6, they give [21] sketches of spurious "circulation modes."

A reviewer has pointed out to us a recent conference paper [22] which describes use of the penalty method to eliminate spurious modes in three-dimensional electromagnetic cavity problems.

REFERENCES

- [1] P. Silvester, "A general high-order finite-element waveguide analysis program," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 204-210, Apr. 1969.
- [2] D. G. Corr and J. B. Davies, "Computer analysis of the fundamental and higher order modes in single and coupled microstrip," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-20, pp. 669-678, Oct. 1972.
- [3] Z. J. Csendes and P. Silvester, "Numerical solution of dielectric loaded waveguide: I- Finite-element analysis," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, pp. 1124-1131, Dec. 1970.
- [4] M. Ikeuchi, H. Sawami, and H. Niki, "Analysis of open-type dielectric waveguides by the finite-element iterative method," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-29, pp. 234-239, Mar. 1981.
- [5] A. Konrad, "Vector variational formulation of electromagnetic fields in anisotropic media," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-24, pp. 553-559, Sept. 1976.
- [6] N. Mabaya, P. E. Lagasse, and P. Vandenbulcke, "Finite element analysis of optical waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-29, pp. 600-605, June 1981.
- [7] P. Vandenbulcke and P. E. Lagasse, "Eigenmode analysis of anisotropic optical fibres or integrated optical waveguides," *Electron. Lett.*, vol. 12, no. 5, pp. 120-122, 1976.
- [8] S. Ahmed and P. Daly, "Finite element methods for inhomogeneous waveguides," *Proc. Inst. Elec. Eng.*, vol. 116, pp. 1661-1664, Oct. 1969.
- [9] C. Yeh, K. Ha, S. B. Dong, and W. P. Brown, "Single-mode optical waveguides," *Appl. Opt.*, vol. 18, pp. 1490-1504, May 1979.
- [10] T. S. Bird, "Propagation and radiation characteristics of rib waveguide," *Electron. Lett.*, vol. 13, p. 401-403, July 7, 1977.
- [11] J. B. Davies, F. A. Fernandez, and G. Y. Philippou, "Finite element analysis of all modes in cavities with circular symmetry," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-30, pp. 1975-1980, Nov. 1982.
- [12] B. M. A. Rahman and J. B. Davies, "Finite element analysis of optical and microwave waveguide problems," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-32, pp. 20-28, Jan. 1984.
- [13] C. G. Williams and G. K. Cambrell, "Numerical solution of surface waveguide modes using transverse field components," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-22, pp. 329-330, Mar. 1974.
- [14] M. Ohtaka, M. Matsuhara, and N. Kumagai, "Analysis of the guided modes in slab-coupled waveguides using a variational method," *IEEE J. Quantum Electron.*, vol. QE-12, pp. 378-382, July 1976.
- [15] J. R. Winkler and J. B. Davies, "Elimination of spurious modes in finite element analysis," to be published in *J. Comp. Phys.*
- [16] K. Bathe, *Finite Element Procedure in Engineering Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1982.
- [17] O. C. Zienkiewicz, *The Finite Element Method*, 3rd ed. London, England: McGraw-Hill, 1977.
- [18] A. D. Berk, "Variational principles for electromagnetic resonators and waveguides," *IRE Trans. Antennas Propagat.*, vol. AP-4, pp. 104-111, Apr. 1956.
- [19] H.-G. Unger, *Planar Optical Waveguides and Fibres*. Oxford, England: Clarendon, 1977.
- [20] NAG Fortran Library, Numerical Algorithms Groups Ltd., Oxford, England.
- [21] M. A. Hamdi, Y. Ousset, and G. Verchery, "A displacement method for the analysis of vibrations of coupled fluid-structure systems," *Int. J. Num. Meth. Eng.*, vol. 13, pp. 139-150, 1978.
- [22] M. Hara, T. Wada, T. Fukasawa, and F. Kikuchi, "Three-dimensional analysis of RF electromagnetic field by finite element method," in *COMPUMAG 83*, Paper E07.

+



B. M. Azizur Rahman (S'80—M'83) was born in Bangladesh in 1953. He received the B.Sc. and M.Sc. degrees in electrical engineering from Bangladesh University of Engineering and Technology (BUET), Dacca, in 1976 and 1979, respectively. He received the Ph.D. degree in electronic engineering from University College London London, England, in 1982.

From 1976 to 1979, he worked at the Electrical Engineering Department, BUET, as a Lecturer.

In 1983, he joined the Electrical and Electronic Engineering Department of University College London, as an Associate Research Assistant. He worked mainly on the application of finite element technique to the solution of microwave and optical waveguide problems. His current interest is also in the solution of microstrip discontinuity problems, using the least square boundary residual method.

+



J. Brian Davies (M'73) was born in Liverpool, England, in 1932. He received a degree in mathematics in 1955 from Jesus College, Cambridge, England. He received that M.Sc. degree in mathematics, the Ph.D. degree in mathematical physics, and the D.Sc. (Eng.) degree in engineering in 1957, 1960, and 1980, respectively, from the University of London, England.

From 1955 to 1963, he worked at the Mullard Research Laboratories, Salford, Surrey, England, except for two years spent at University College London. In 1963 he joined the Staff of the Department of Electrical Engineering, University of Sheffield, England. Since 1967, he has been on the Staff at University College London, where he is a Reader in Electrical Engineering. From 1971 to 1972, he was a Visiting Scientist at the National Bureau of Standards, Boulder, CO. In 1983, he was a Visitor at the Department of Engineering Science, University of Oxford, England. His research interests have been with problems of electromagnetic theory, especially those requiring computer methods of solution. But recently, his interests have extended into the field problems of acoustic microscopy and of surface acoustic waves.

Dr. Davies is a Member of the Institute of Electrical Engineers, London, England.