

Waveguide Propagation Modes and Quadratic Eigenvalue Problems

*Prof. André Nicolet, Institut Fresnel, UMR CNRS 6133,
Université Paul Cézanne, Marseille, France.*

*Prof. Christophe Geuzaine, Mathematics Department,
Case Western Reserve University, Cleveland, OH, USA.*

Abstract This paper presents a direct approach to determine numerically the propagation modes in waveguides via a finite element method. Given a pulsation ω , a quadratic eigenvalue problem is solved to obtain the propagation constant β .

1 Propagation modes with finite elements

Our goal is to obtain numerically the propagation modes in an electromagnetic waveguide invariant along the z -axis—and therefore described by its cross-section in the xy -plane. We choose to formulate the problem in terms of the electric field \mathcal{E} with homogeneous boundary conditions, which corresponds to a waveguide with perfectly conducting metallic walls. (A formulation of the problem in terms of the magnetic field could be handled in a similar way.)

Choosing a time dependence in $e^{-i\omega t}$, and taking into account the invariance along the z -axis, we define the time-harmonic two-dimensional electric field \mathbf{E} such that

$$\mathcal{E}(x, y, z, t) = \Re e(\mathbf{E}(x, y) e^{-i(\omega t - \beta z)}), \quad (1)$$

where $\omega = k_0 c$ is the angular frequency and β is the propagating constant of the guided mode. Note that \mathbf{E} is a complex-valued field depending on two variables (coordinates x and y) but still with three components (along the three axes). The two-dimensional electric field is separated into a transverse component \mathbf{E}_t in the xy -plane and a longitudinal field E_ℓ along the z -axis of invariance so that $\mathbf{E} = \mathbf{E}_t + E_\ell \mathbf{e}^z$, with $\mathbf{E}_t \cdot \mathbf{e}^z = 0$.

Writing Maxwell's equations in terms of \mathbf{E} and denoting by k_0 the wave number $\omega\sqrt{\mu_0\varepsilon_0}$, one gets

$$\mathbf{curl}_\beta(\mu_r^{-1} \mathbf{curl}_\beta \mathbf{E}) = k_0^2 \varepsilon_r \mathbf{E}, \quad (2)$$

where the operator \mathbf{curl}_β is defined as:

$$\mathbf{curl}_\beta \mathbf{U}(x, y) = \mathbf{curl} \left(\mathbf{U}(x, y) e^{i\beta z} \right) e^{-i\beta z}.$$

Since the electric field satisfies a homogeneous Dirichlet boundary condition ($\mathbf{n} \times \mathbf{E} = 0$) on the boundary of the guide, the weak formulation of Eq. (2) writes

$$\mathcal{R}(\mathbf{E}, \mathbf{E}') = \int_{\Omega} \mu_r^{-1} \mathbf{curl}_\beta \mathbf{E} \cdot \overline{\mathbf{curl}_\beta \mathbf{E}'} dx dy - k_0^2 \int_{\Omega} \varepsilon_r \mathbf{E} \cdot \overline{\mathbf{E}'} dx dy = 0, \quad \forall \mathbf{E}' \in H(\mathbf{curl}_\beta, \Omega), \quad (3)$$

where the space $H(\mathbf{curl}_\beta, \Omega)$ of curl-conforming fields is defined as $H(\mathbf{curl}_\beta, \Omega) = \{\mathbf{v} \in [L^2(\Omega)]^3, \mathbf{curl}_\beta \mathbf{v} \in [L^2(\Omega)]^3\}$.

The discretization of this weak formulation is obtained via finite elements [1]. The cross-section of the guide is meshed with triangles and Whitney finite elements [2] are used, *i.e.*, edge elements for the transverse field and nodal elements for the longitudinal field:

$$\mathbf{E}_t = \sum_{j=1}^{\#\text{edges}} e_j^t \mathbf{w}_e^j(x, y) \quad \text{and} \quad E_\ell = \sum_{j=1}^{\#\text{nodes}} e_j^z w_n^j(x, y),$$

where e_j^t denotes the line integral of the transverse component \mathbf{E}_t on the edges, e_j^z denotes the line integral of the longitudinal component E_ℓ along one unit of length of the z -axis (which is equivalent to the nodal value), and \mathbf{w}_e^j and w_n^j are respectively the basis functions of Whitney 1-forms and 0-forms on triangles. On a triangle, if λ_i denotes the barycentric co-ordinate associated to the node i , $\mathbf{w}_e = \lambda_i \mathbf{grad} \lambda_j - \lambda_j \mathbf{grad} \lambda_i$ for the edge going from node i to node j and $w_n = \lambda_i$ for the node i .

The following transverse operators are defined for a scalar function $\varphi(x, y)$ and a transverse field $\mathbf{v} = v_x(x, y)\mathbf{e}^x + v_y(x, y)\mathbf{e}^y$:

$$\begin{aligned} \mathbf{grad}_t \varphi &= \frac{\partial \varphi}{\partial x} \mathbf{e}^x + \frac{\partial \varphi}{\partial y} \mathbf{e}^y \\ \mathbf{curl}_t \mathbf{v} &= \left(\frac{\partial v_x}{\partial y} - \frac{\partial v_y}{\partial x} \right) \mathbf{e}^z \end{aligned}$$

and are used to separate \mathbf{curl}_β into its transverse and longitudinal components:

$$\mathbf{curl}_\beta (\mathbf{v} + \varphi \mathbf{e}^z) = \mathbf{curl}_t \mathbf{v} + (\mathbf{grad}_t \varphi - i \beta \mathbf{v}) \times \mathbf{e}^z.$$

At this stage, the materials can still be supposed to be anisotropic so that ε and μ are tensorial—but they are requested not to mix longitudinal and

transverse components. The most general form with this property (called “z-anisotropy” in the following) is:

$$\underline{\underline{\varepsilon}} = \begin{pmatrix} \varepsilon_{xx} & \varepsilon_{xy} & 0 \\ \varepsilon_{yx} & \varepsilon_{yy} & 0 \\ 0 & 0 & \varepsilon_{zz} \end{pmatrix} = \underline{\underline{\varepsilon}}_{tt} \underline{\underline{\varepsilon}}_{zz} = \begin{pmatrix} \varepsilon_{xx} & \varepsilon_{xy} & 0 \\ \varepsilon_{yx} & \varepsilon_{yy} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \varepsilon_{zz} \end{pmatrix}, \quad (4)$$

so that $\underline{\underline{\varepsilon}}\mathbf{E} = \underline{\underline{\varepsilon}}_{tt}\mathbf{E}_t + \underline{\underline{\varepsilon}}_{zz}E_\ell\mathbf{e}^z$ and a similar form is supposed for the permeability $\underline{\underline{\mu}}$. Note that this property is conserved for the inverse tensors $\underline{\underline{\varepsilon}}^{-1}$ and $\underline{\underline{\mu}}^{-1}$. (In the sequel we remove the explicit notation for the tensorial nature of ε and μ to avoid heavy notations.)

Using the following matrix definitions [3, 1]:

$$\begin{aligned} A_{ij}^{tt} &= \int_{\Omega} \mu_r^{-1} \mathbf{curl}_t \mathbf{w}_e^i \cdot \mathbf{curl}_t \mathbf{w}_e^j dx dy \\ A_{ij}^{zz} &= \int_{\Omega} [\mu_r^{-1} (\mathbf{grad}_t w_n^i \times \mathbf{e}^z)] \cdot (\mathbf{grad}_t w_n^j \times \mathbf{e}^z) dx dy \\ A_{ij}^{tz} &= A_{ji}^{zt} = \int_{\Omega} [\mu_r^{-1} (\mathbf{w}_e^i \times \mathbf{e}^z)] \cdot (\mathbf{grad}_t w_n^j \times \mathbf{e}^z) dx dy \\ B_{ij}^{tt} &= \int_{\Omega} [\mu_r^{-1} (\mathbf{w}_e^i \times \mathbf{e}^z)] \cdot (\mathbf{w}_e^j \times \mathbf{e}^z) dx dy \\ C_{ij}^{tt} &= \int_{\Omega} \varepsilon_r \mathbf{w}_e^i \cdot \mathbf{w}_e^j dx dy & C_{ij}^{zz} &= \int_{\Omega} \varepsilon_r w_n^i w_n^j dx dy \end{aligned}$$

the Hermitian matrix system obtained from Eq.(3) for the propagation mode can be written in the form:

$$\begin{aligned} & \begin{pmatrix} A^{tt} + \beta^2 B^{tt} & i\beta A^{tz} \\ -i\beta A^{zt} & A^{zz} \end{pmatrix} \begin{pmatrix} \mathbf{e}_t \\ \mathbf{e}_\ell \end{pmatrix} \\ &= k_0^2 \begin{pmatrix} C^{tt} & 0 \\ 0 & C^{zz} \end{pmatrix} \begin{pmatrix} \mathbf{e}_t \\ \mathbf{e}_\ell \end{pmatrix}, \end{aligned} \quad (5)$$

where \mathbf{e}_t and \mathbf{e}_ℓ are the column arrays of transverse and longitudinal degrees of freedom respectively.

A first approach to the solution of (5) is to look for k_0 for a given β . Indeed, since the square of the wave number k_0^2 appears alone in (5), whereas the propagation constant β and its square β^2 are both involved, it seems natural to fix the value of β to obtain a generalized eigenvalue problem for k_0^2 . Nevertheless, from a physical point of view, it is often more natural to do the opposite—the frequency is usually given *a priori* and, due to the chromatic dispersion, ε does itself depend on k_0 . Moreover, the leaky modes [1] that play a fundamental

role in the study of microstructured optical fibres are traditionally described by a complex β whose imaginary part is related to losses of the mode at a given pulsation hence the importance to obtain β from k_0 .

In order to obtain an eigenvalue problem for β^2 , a second approach is to slightly modify (5) by dividing the longitudinal degrees of freedom by $i\beta$ [3, 4]. The new column vector of unknowns is

$$\begin{pmatrix} \mathbf{f}_t \\ \mathbf{f}_\ell \end{pmatrix} = \begin{pmatrix} \mathbf{e}_t \\ (i\beta)^{-1}\mathbf{e}_\ell \end{pmatrix}, \quad (6)$$

which leads to the following real and symmetric system of equations:

$$\begin{pmatrix} \mathbf{A}^{tt} + \beta^2\mathbf{B}^{tt} & -\beta^2\mathbf{A}^{tz} \\ -\beta^2\mathbf{A}^{zt} & \beta^2\mathbf{A}^{zz} \end{pmatrix} \begin{pmatrix} \mathbf{f}_t \\ \mathbf{f}_\ell \end{pmatrix} \\ = k_0^2 \begin{pmatrix} \mathbf{C}^{tt} & 0 \\ 0 & \beta^2\mathbf{C}^{zz} \end{pmatrix} \begin{pmatrix} \mathbf{f}_t \\ \mathbf{f}_\ell \end{pmatrix}, \quad (7)$$

which only involves terms in β^2 . This modification, however, breaks down when general anisotropic materials are considered (*i.e.* full anisotropic materials with possibly 9 non-zero coefficients instead of z -anisotropic materials), since in this case new terms appear in the matrices that prevent the reduction to a form without terms linear in β . They come from the fact that transverse and longitudinal components of the field are no more orthogonal after the action of the anisotropic constitutive laws.

This is a serious limitation for a range of modeling problems, in particular when it is necessary to take into account a change of coordinates by equivalent (inhomogeneous and anisotropic) material properties—as happens, for instance, when modeling twisted microstructured optical fibres [5, 1].

In this case of more general anisotropy, the new terms correspond to the addition of the following contribution to the matrix system:

$$\begin{pmatrix} i\beta\mathbf{D}^{tt} & \mathbf{F}^{zt} + k_0^2\mathbf{G}^{zt} \\ \mathbf{F}^{tz} + k_0^2\mathbf{G}^{tz} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{e}_t \\ \mathbf{e}_\ell \end{pmatrix},$$

where

$$\mathbf{D}_{ij}^{tt} = \int_{\Omega} -[\mu_r^{-1} \mathbf{curl}_t \mathbf{w}_e^i] \cdot (\mathbf{w}_e^j \times \mathbf{e}^z) + [\mu_r^{-1} (\mathbf{w}_e^i \times \mathbf{e}^z)] \cdot \mathbf{curl}_t \mathbf{w}_e^j \, dxdy$$

$$\mathbf{F}_{ij}^{tz} = \mathbf{F}_{ji}^{zt} = \int_{\Omega} [\mu_r^{-1} \mathbf{curl}_t \mathbf{w}_e^i] \cdot (\mathbf{grad}_t w_n^j \times \mathbf{e}^z) \, dxdy$$

$$\mathbf{G}_{ij}^{tz} = \mathbf{G}_{ji}^{zt} = \int_{\Omega} [\varepsilon_r \mathbf{w}_e^i] \cdot w_n^j \mathbf{e}^z \, dxdy$$

2 Quadratic eigenvalue problem

We thus propose a third approach: the idea is to transform the original quadratic problem into a generalized linear eigensystem [6] in β , with a technique similar to the one used to transform high order differential equations into first order equations (this is in fact the Fourier transform of this technique).

Given a generalized quadratic eigenvalue problem of the form

$$-\beta^2 M \mathbf{u} + i\beta L \mathbf{u} + K \mathbf{u} = 0$$

where M, L , and K are $n \times n$ matrices and β and \mathbf{u} are the eigenvalue and eigenvector respectively, it suffices to take $\mathbf{w} = i\beta \mathbf{u}$ and to write the equivalent linear generalized eigenvalue problem involving $2n \times 2n$ matrices P and Q :

$$Q \begin{pmatrix} \mathbf{u} \\ \mathbf{w} \end{pmatrix} = i\beta P \begin{pmatrix} \mathbf{u} \\ \mathbf{w} \end{pmatrix} \quad (8)$$

with

$$P = \begin{pmatrix} L & M \\ I & 0 \end{pmatrix}$$

and

$$Q = \begin{pmatrix} -K & 0 \\ 0 & I \end{pmatrix}.$$

The expressions for the three matrices M, L , and K follow from (5) and (1):

$$M = \begin{pmatrix} -B^{tt} & 0 \\ 0 & 0 \end{pmatrix}, \quad L = \begin{pmatrix} D^{tt} & A^{tz} \\ -A^{zt} & 0 \end{pmatrix},$$

$$K = \begin{pmatrix} A^{tt} - k_0^2 C^{tt} & F^{zt} + k_0^2 G^{zt} \\ F^{tz} + k_0^2 G^{tz} & A^{zz} - k_0^2 C^{zz} \end{pmatrix}.$$

Note that the matrix $-\beta^2 M + i\beta L + K$ is Hermitian if the material properties correspond to real scalars or Hermitian tensors.

Iterative methods for large sparse eigensystems (such as the Implicitly Restarted Arnoldi Method (IRAM) used in Arpack [7]) require only to perform matrix-vector products. More exactly, if the smallest eigenvalue around a given value s is to be found, the finite rank operator $(Q - sP)^{-1}P$ is the only thing the algorithm (*e.g.*, the reverse iteration interface of Arpack) has to know. We emphasize the fact that it is an operator as it is a very bad idea to compute explicitly this matrix! This operator comes from the composition of three operations: first, sP is subtracted from both sides of the eigenvalue problem to

shift the spectrum by a value s , then the system is multiplied on the left by P^{-1} to transform the generalized eigenvalue problem into a regular one, and finally the matrix is inverted so that the numerical algorithm will provide the smallest eigenvalues instead of the largest ones. In practice, the fact that the two lower blocks of the matrix

$$Q - sP = \begin{pmatrix} -K - sL & -sM \\ -sI & I \end{pmatrix}$$

are made of identity matrices up to a scalar factor, commuting with any other matrix, makes it possible to give an explicit formula for the inversion of $Q - sP$. The block matrix determinant is the matrix $\Delta = -(s^2M + sL + K)$ and

$$(Q - sP)^{-1} = \begin{pmatrix} \Delta^{-1} & s\Delta^{-1}M \\ s\Delta^{-1} & -\Delta^{-1}K - s\Delta^{-1}L \end{pmatrix}.$$

The explicit expression for the requested operator is thus:

$$(Q - sP)^{-1}P \begin{pmatrix} \mathbf{u} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \Delta^{-1}(L\mathbf{u} + sM\mathbf{u} + M\mathbf{w}) \\ \Delta^{-1}(-K\mathbf{u} + sM\mathbf{w}) \end{pmatrix}.$$

It involves only two $n \times n$ system solutions with the same matrix Δ (Δ^{-1} is not explicitly computed but instead systems involving Δ are solved) and even if a little more time consuming, the algorithmic complexity of the method is the same as the one of the two previous methods.

3 A numerical example

As a simple numerical example, a metallic rectangular waveguide ($\varepsilon_r = 1, \mu_r = 1$, width = $2m$, height = $1m$) and a propagation constant $\beta = 2m^{-1}$ are considered. The first three modes correspond to k_0^2 equal to $\pi^2/4 + 4 = 6.467401100m^{-2}$, $\pi^2 + 4 = 13.86960440m^{-2}$, and $5\pi^2/4 + 4 = 16.33700550m^{-2}$ respectively.

The corresponding finite element model uses a uniform mesh with 2596 triangles, leading to a system of equations with 5047 complex unknowns. We used Arpack [7] to solve the eigensystems iteratively and prescribed a relative error tolerance of 10^{-8} on the eigenvalues and Krylov subspaces of dimension 20.

Given $\beta = 2m^{-1}$, the first method gives a correct answer for the values of k_0^2 (6.467378559331, 13.86950286971, and 16.35633078641) in 32 Arpack steps.

Note that the discrepancy with the analytic result is due to the finite element discretization and not to rounding errors in matrix computations, as we we have successfully verified by using a finer mesh.

The two methods allowing the computation of β from given k_0 (numerically obtained above) are now considered. The first method (that recovers a linear problem in β^2) gives $\beta = 2$ with 13 figures for the three modes using 31, 25 and 30 Arpack steps respectively. The method directly solving the quadratic eigenvalue problem gives $\beta = 2.000000000000$, $\beta = 2.000000000003$, and $\beta = 1.999999999996$ using 58, 52, and 78 Arpack steps respectively. (Note that the non-Hermitian routine from Arpack is used and that all the computed imaginary parts are negligible as expected, *i.e.*, of the order of the prescribed accuracy in Arpack.)

Although the simple example presented here can be computed by the three methods, an application where our new method is necessary can be found in [8].

4 Conclusion

We have presented three methods to solve the quadratic eigenvalue problem that arises in the modeling of electromagnetic waveguides. All three algorithms are based on the same discretization of the problem and use the same elementary matrices—only with a different organization of the computations. Even if the direct computation of the quadratic eigenvalue problem is a little bit more time consuming, the three methods have the same algorithmic complexity and are stable and accurate. The main advantage of the new method lies in its generality: it allows the computation of β from a given k_0 , which is important to cope with the chromatic dispersion of optical materials, and it can take into account general anisotropic materials—a crucial advantage for the modeling of twisted microstructured optical fibres [1, 5, 8].

References

- [1] F. Zolla, G. Renversez, A. Nicolet, B. Khulmey, S. Guenneau, and D. Felbacq, *Foundations of Photonic Crystal Fibres*, Imperial College Press, London, 2005.
- [2] A. Bossavit, *Whitney forms: a class of finite elements for three-dimensional computations in electromagnetism*, IEE Proceedings 135, Pt. A (1988), no. 8, 493–499.
- [3] A. F. Peterson, D. R. Wilton, *Curl-conforming mixed-order edge elements for discretizing the 2D and 3D vector Helmholtz equation*, in *Finite Element*

-
- Software for Microwave Engineering*, P. P. Silvester, G. Pelosi, and T. Itoh, Eds. New York: Wiley, pp. 101-125, 1996.
- [4] J.-F. Lee, D.-K. Sun, Z. J. Csendes, *Full-wave analysis of dielectric waveguides using tangential vector finite elements*, IEEE Trans. Microwave Theory Tech, MTT-39 (1991), no. 8, pp. 1262-1271.
- [5] A. Nicolet, S. Guenneau, F. Zolla, *Modelling of twisted optical waveguides with edge elements*, Eur. Phys. J. Appl. Phys. (28) pp. 153-157, 2004.
- [6] F. Tisseur, K. Meerbergen, *The Quadratic Eigenvalue Problem*, SIAM Review, Vol. 43 (2001), No. 2, pp. 235-286.
- [7] R. B. Lehoucq, D. C. Sorensen, C. Yang, ARPACK Users' Guide: *Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*, SIAM, Philadelphia, 1998.
- [8] A. Nicolet, F. Zolla, *Finite Element Analysis of Helicoidal Waveguides*, IEE Proc., Sci. Meas. Technol., *this issue*.