Journal of Computational Mathematics, Vol.22, No.2, 2004, 188–199.

ON FINITE ELEMENT METHODS FOR INHOMOGENEOUS DIELECTRIC WAVEGUIDES [*1)]

Zhiming Chen
(*LSEC, Institute of Computational Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)
(*E-mail: zmchen@lsec.cc.ac.cn*)

Jian-hua Yuan
(*Institute of Computational Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)
(*E-mail: yuanjh@lsec.cc.ac.cn*)

Dedicated to Professor Zhong-ci Shi on the occasion of his 70th birthday

Abstract

We investigate the problem of computing electromagnetic guided waves in a closed, inhomogeneous, pillared three-dimensional waveguide at a given frequency. The problem is formulated as a generalized eigenvalue problem. By modifying the sesquilinear form associated with the eigenvalue problem, we provide a new convergence analysis for the finite element approximations. Numerical results are reported to illustrate the performance of the method.

Mathematics subject classification: 65N30, 35L15
Key words: Waveguide, Eigenvalue, Inf-Sup condition.

1. Introduction

We consider in this paper a closed waveguide defined by a right cylinder with cross section Ω, a bounded, Lipschitz, simply connected polyhedral domain in $\mathbf{R}^2$. The waveguide is filled with inhomogeneous media whose electromagnetic properties are described by the real-valued functions ε and μ. We assume the magnetic permeability $\mu = \mu_0$, the magnetic permeability in vacuum, and the dielectric permittivity ε is piecewise constant and has no variation along the waveguide. More precisely, let $\Omega_1 \subset \Omega$ be an open domain, $\Omega_2 = \Omega \backslash \bar{\Omega}_1$. We assume

$$\varepsilon(x) = \begin{cases} \varepsilon_1 \varepsilon_0 & \text{in} \quad \Omega_1, \\ \varepsilon_2 \varepsilon_0 & \text{in} \quad \Omega_2, \end{cases}$$

where ε_0 is the dielectric permittivity in vacuum.

The waveguide problem is to find solutions to Maxwell equations which are of the general form

$$\begin{cases} \mathcal{E}(x, x_3, t) = (\mathbf{E}(x), E_3(x)) e^{i(\omega t - \beta x_3)} \\ \mathcal{H}(x, x_3, t) = (\mathbf{H}(x), H_3(x)) e^{i(\omega t - \beta x_3)} \end{cases} \tag{1.1}$$

where $x \in \Omega$ and the x_3-axis is along the waveguide, $\omega > 0$ is the angular frequency of the guided wave, β is the constant of propagation, $\mathbf{E}$ and $\mathbf{H}$ are electric and magnetic field components in

* Received January 31, 2004.
1) This author was supported in part by China NSF under grant 10025102 and by China MOST under grant G1999032802.

the plane of the cross section, and E_3 and H_3 are electric and magnetic components along the waveguide.

With ansatz (1.1), the second order three dimensional Maxwell equations expressed in terms of electric field $(\mathbf{E}, E_3)$ reduce to the following two-dimensional equations (cf. e.g. [11]):

$$\nabla \times (\nabla \times \mathbf{E}) - i\beta \nabla E_3 - (\omega^2 \varepsilon_0 \mu_0)\varepsilon \mathbf{E} = -\beta^2 \mathbf{E} \quad \text{in} \quad \Omega, \tag{1.2}$$

$$\nabla \cdot (\varepsilon \mathbf{E}) - i\beta\varepsilon E_3 = 0 \quad \text{in} \quad \Omega. \tag{1.3}$$

For simplicity, perfect electric conductor boundary conditions are imposed

$$\mathbf{E} \times \mathbf{n} = 0, \quad E_3 = 0 \quad \text{ on } \partial\Omega, \tag{1.4}$$

where $\mathbf{n}$ is the unit outer normal to $\partial\Omega$.

Advances in various branches of photonics technologies have established the need for the development of numerical and approximate methods for the analysis of a wide range of waveguide structures that are not amenable to exact analytical studies [5]. The problem (1.2)-(1.3) is an eigenvalue problem. Either ω or β is assumed to be known, and the goal is to find all possible pairs which consist of the other missing constant β or ω and the corresponding field $(\mathbf{E}, E_3)$. The case with a given real-valued β has been extensively studied in the literature (see e.g. [8], [2] and the references therein). More physically relevant case with a given ω to find unknown β is recently studied in [11], in which the eigenvalue problem is studied under the assumption that the frequency ω does not belong to the spectrum of the variational eigenvalue problems associated with the curl-curl and div-grad operators. We remark that since the spectrum of these two operators are generally unknown, this assumption on ω cannot be verified in practical applications.

In this paper we are going to provide a new convergence analysis for the eigenvalue problem (1.2)-(1.3) which removes the restrictions on the frequency ω in [11]. This is achieved by modifying the sesquilinear form associated with the variational formulation of (1.2)-(1.3). The key technical difficulty is the proof of the inf-sup condition of the modified sesquilinear form which allows us to use the general framework for the approximation of the eigenvalue problems developed in [1]. We introduce a finite element method which uses the lowest order Nedelec edge element and standard conforming linear finite element to approximate $(\mathbf{E}, E_3)$, respectively. This choice of finite elements is shown in [11] to exclude spurious modes. Here again the discrete inf-sup condition is proved without any restrictions on the frequency ω and the mesh sizes. We also report several numerical experiments to illustrate the performance of the method studied in this paper.

2. The Continuous Problem

We begin with introducing the Hilbert space $\mathbb{X} = H_0(\text{curl};\Omega) \times H_0^1(\Omega)$ which is equipped with the norm

$$\|(\mathbf{V}, q)\|_{\mathbb{X}} = \|\, \mathbf{V} \,\|_{\text{curl},\Omega} + \|\, q \,\|_{H^1(\Omega)} \qquad \forall (\mathbf{V}, q) \in \mathbb{X}.$$

Here $\|\, \mathbf{V} \,\|_{\text{curl},\Omega} = (\|\, \nabla \times \mathbf{V} \,\|^2_{L^2(\Omega)} + \|\, \mathbf{V} \,\|^2_{L^2(\Omega)})^{1/2}$ is the norm of the space $H(\text{curl};\Omega)$ which is defined as the collection of all functions $\mathbf{V}$ in $L^2(\Omega)$ such that $\|\, \mathbf{V} \,\|_{\text{curl},\Omega} < \infty$. $H_0(\text{curl};\Omega)$ consists of functions $\mathbf{V}$ in $H(\text{curl};\Omega)$ whose tangential component $\mathbf{V} \times \mathbf{n}$ vanishes on the boundary $\partial\Omega$.

Set $E_3^{\text{new}} = -i\beta E_3$ in (1.3). To save the notation, E_3 will represent E_3^{new} for the reminder of this paper. Let $k_0^2 = \omega^2 \varepsilon_0 \mu_0$. For $\Lambda > 0$, by adding $\Lambda \mathbf{E}$ on both sides of (1.2), we can reformulate (1.2)-(1.3) with boundary condition (1.4) into the following variational form:

For a given $\omega > 0$, find all pairs $(\beta, (\mathbf{E}, E_3)) \in \mathbb{C} \times \mathbb{X}$, such that, for any $(\mathbf{V}, q) \in \mathbb{X}$,

$$\langle \nabla \times \mathbf{E}, \nabla \times \mathbf{V} \rangle + \langle (\Lambda - k_0^2 \varepsilon)\mathbf{E}, \mathbf{V} \rangle + \langle \nabla E_3, \mathbf{V} \rangle = (\Lambda - \beta^2)\langle \mathbf{E}, \mathbf{V} \rangle, \tag{2.1}$$

$$\langle \nabla q, \varepsilon \mathbf{E} \rangle - \langle \varepsilon E_3, q \rangle = 0. \tag{2.2}$$

Here $\langle \cdot, \cdot \rangle$ stands for the inner product of $[L^2(\Omega)]^2$ or $L^2(\Omega)$.

Now we introduce the sesquilinear forms $\mathbf{a} : \mathbb{X} \times \mathbb{X} \to \mathbb{C}$ and $\mathbf{b} : \mathbb{X} \times \mathbb{X} \to \mathbb{C}$ as follows: for any $(\mathbf{U}, p), (\mathbf{V}, q) \in \mathbb{X}$,

$$\begin{aligned} \mathbf{a}((\mathbf{U}, p), (\mathbf{V}, q)) &= \langle \nabla \times \mathbf{U}, \nabla \times \mathbf{V} \rangle + \langle (\Lambda - k_0^2 \varepsilon)\mathbf{U}, \mathbf{V} \rangle + \langle \nabla p, \mathbf{V} \rangle \\ &\quad - \langle \nabla q, \varepsilon \mathbf{U} \rangle + \langle \varepsilon p, q \rangle, \\ \mathbf{b}((\mathbf{U}, p), (\mathbf{V}, q)) &= \langle \mathbf{U}, \mathbf{V} \rangle. \end{aligned}$$

Then it is easy to see that (2.1)-(2.2) is equivalent to the following generalized eigenvalue problem:

For a given $\omega > 0$, find all pairs $(\lambda, (\mathbf{E}, E_3)) \in \mathbb{C} \times \mathbb{X}$, such that

$$\mathbf{a}((\mathbf{E}, E_3), (\mathbf{V}, q)) = \lambda \mathbf{b}((\mathbf{E}, E_3), (\mathbf{V}, q)) \quad \forall (\mathbf{V}, q) \in \mathbb{X}. \tag{2.3}$$

The addition of $\Lambda \mathbf{E}$ in (2.1) is solely for the ease of mathematical analysis. In particular, it does not affect the practical computations. The following theorem is the main result of this section.

Theorem 2.1. *Assume that the parameter Λ satisfies*

$$\max(k_0^2 \varepsilon_1, k_0^2 \varepsilon_2) < \Lambda \le k_0^2(\varepsilon_1 + \varepsilon_2), \tag{2.4}$$

then the sesquilinear form $\mathbf{a}(\cdot, \cdot)$ satisfies the following properties:

(i) *There exists a constant $C > 0$ such that*

$$|\mathbf{a}((\mathbf{U}, p), (\mathbf{V}, q))| \le C \| (\mathbf{U}, p) \|_{\mathbb{X}} \| (\mathbf{V}, q) \|_{\mathbb{X}}, \quad \forall (\mathbf{U}, p), (\mathbf{V}, q) \in \mathbb{X}. \tag{2.5}$$

(ii) *There exists a constant $\alpha_1 > 0$ such that, for any $(\mathbf{U}, p) \in \mathbb{X}$,*

$$\sup_{(\mathbf{V}, q) \in \mathbb{X}} \frac{\mathbf{a}((\mathbf{U}, p), (\mathbf{V}, q))}{\| (\mathbf{V}, q) \|_{\mathbb{X}}} \ge \alpha_1 \| (\mathbf{U}, p) \|_{\mathbb{X}}. \tag{2.6}$$

(iii) *There exists a constant $\alpha_2 > 0$ such that, for any $(\mathbf{V}, q) \in \mathbb{X}$,*

$$\sup_{(\mathbf{U}, p) \in \mathbb{X}} \frac{\mathbf{a}((\mathbf{U}, p), (\mathbf{V}, q))}{\| (\mathbf{U}, p) \|_{\mathbb{X}}} \ge \alpha_2 \| (\mathbf{V}, q) \|_{\mathbb{X}}. \tag{2.7}$$

Proof. Without loss of generality, we may assume $\varepsilon_1 > \varepsilon_2$. For brevity we set $\rho(x) = \Lambda - k_0^2 \varepsilon(x)$ and $\rho_i = \Lambda - k_0^2 \varepsilon_i$, $i = 1, 2$. By the assumption (2.4), $\rho_i > 0$, $i = 1, 2$.

(2.5) is obvious. We now show the inf-sup condition (2.6). For any $(\mathbf{U}, p) \in \mathbb{X}$, define $\phi \in H_0^1(\Omega)$ as the weak solution of the problem

$$\langle \nabla \phi, \nabla \psi \rangle = -\left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right) \int_{\Omega_2} \mathbf{U}_2 \cdot \nabla \psi dx \quad \forall \psi \in H_0^1(\Omega), \tag{2.8}$$

where $\mathbf{U}_2 = \mathbf{U}|_{\Omega_2}$. By taking $\psi = \phi$ in (2.8) we know that

$$\| \nabla\phi \|_{L^2(\Omega)} \leq \left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right) \| \mathbf{U}_2 \|_{L^2(\Omega_2)}. \tag{2.9}$$

Now we set

$$\mathbf{V} = \mathbf{U} + \theta\nabla p + \nabla\phi, \quad q = \frac{1}{\varepsilon_1} p, \tag{2.10}$$

where $\theta = 2(\frac{\varepsilon_2}{\varepsilon_1})^2 \frac{1}{\rho_2}$. It is obvious that $(\mathbf{V}, q) \in \mathbb{X}$. We have

$$\begin{aligned}\mathbf{a}((\mathbf{U},p),(\mathbf{V},q)) &= \| \nabla\times\mathbf{U} \|^2_{L^2(\Omega)} + \langle \rho\mathbf{U}, \mathbf{U} + \theta\nabla p + \nabla\phi\rangle \\ &\quad + \langle \nabla p, \mathbf{U} + \theta\nabla p + \nabla\phi\rangle - \langle \frac{\varepsilon}{\varepsilon_1}\mathbf{U}, \nabla p\rangle + \langle \frac{\varepsilon}{\varepsilon_1} p, p\rangle.\end{aligned}$$

By (2.8), we know that

$$\langle \nabla p, \mathbf{U} + \nabla\phi\rangle - \langle \frac{\varepsilon}{\varepsilon_1}\mathbf{U}, \nabla p\rangle = 0.$$

Moreover, by Cauchy-Schwarz inequality and (2.9), we obtain

$$\begin{aligned}\theta|\langle \rho\mathbf{U}, \nabla p\rangle| &\leq \frac{1}{2}\theta\| \nabla p \|^2_{L^2(\Omega)} + \frac{1}{2}\theta\| \rho\mathbf{U} \|^2_{L^2(\Omega)}, \\ |\langle \rho\mathbf{U}, \nabla\phi\rangle| &\leq \left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right) \| \rho\mathbf{U} \|_{L^2(\Omega)} \| \mathbf{U}_2 \|_{L^2(\Omega_2)} \\ &= \rho_1 \left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right) \| \mathbf{U}_1 \|_{L^2(\Omega_1)} \| \mathbf{U}_2 \|_{L^2(\Omega_2)} + \rho_2 \left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right) \| \mathbf{U}_2 \|^2_{L^2(\Omega_2)}.\end{aligned}$$

Thus

$$\begin{aligned}\mathbf{a}((\mathbf{U},p),(\mathbf{V},q)) &= \| \nabla\times\mathbf{U} \|^2_{L^2(\Omega)} + \langle \rho\mathbf{U}, \mathbf{U}\rangle + \langle \frac{\varepsilon}{\varepsilon_1} p, p\rangle + \theta\| \nabla p \|^2_{L^2(\Omega)} \\ &\quad + \theta\langle \rho\mathbf{U}, \nabla p\rangle + \langle \rho\mathbf{U}, \nabla\phi\rangle \\ &\geq \| \nabla\times\mathbf{U} \|^2_{L^2(\Omega)} + \frac{1}{2}\theta\| \nabla p \|^2_{L^2(\Omega)} + \langle \frac{\varepsilon}{\varepsilon_1} p, p\rangle \\ &\quad + \rho_1\| \mathbf{U}_1 \|^2_{L^2(\Omega_1)} + \rho_2 \frac{\varepsilon_2}{\varepsilon_1}\| \mathbf{U}_2 \|^2_{L^2(\Omega_2)} \\ &\quad - \frac{1}{2}\theta\| \rho\mathbf{U} \|^2_{L^2(\Omega)} - \rho_1 \left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right) \| \mathbf{U}_1 \|_{L^2(\Omega_1)} \| \mathbf{U}_2 \|_{L^2(\Omega_2)}.\end{aligned}$$

It follows from $\varepsilon_2 < \varepsilon_1$ that $\rho_1 < \rho_2$ and

$$\rho_1\varepsilon_2 = \Lambda\varepsilon_2 - k_0^2\varepsilon_1\varepsilon_2 < \Lambda\varepsilon_1 - k_0^2\varepsilon_1\varepsilon_2 = \rho_2\varepsilon_1.$$

Moreover, since $\Lambda \leq k_0^2(\varepsilon_1 + \varepsilon_2)$ by (2.4), we have

$$\rho_2\varepsilon_2 = \Lambda\varepsilon_2 - k_0^2\varepsilon_2^2 = \rho_1\varepsilon_1 + (\varepsilon_2 - \varepsilon_1)[\Lambda - k_0^2(\varepsilon_1 + \varepsilon_2)] \geq \rho_1\varepsilon_1.$$

Then, since $\theta = 2(\frac{\varepsilon_2}{\varepsilon_1})^2 \frac{1}{\rho_2}$, we have

$$\begin{aligned}\rho_1 - \frac{1}{2}\theta\rho_1^2 &= \rho_1 - \left(\frac{\varepsilon_2}{\varepsilon_1}\right)^2 \frac{\rho_1^2}{\rho_2} \geq \rho_1 \left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right), \\ \rho_2 \left(\frac{\varepsilon_2}{\varepsilon_1}\right) - \frac{1}{2}\theta\rho_2^2 &= \rho_2 \left(\frac{\varepsilon_2}{\varepsilon_1}\right) \left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right) \geq \rho_1 \left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right).\end{aligned}$$

Thus

$$
\begin{aligned}
& \rho_1 \| \mathbf{U}_1 \|^2_{L^2(\Omega_1)} + \rho_2 \frac{\varepsilon_2}{\varepsilon_1} \| \mathbf{U}_2 \|^2_{L^2(\Omega_2)} \\
& -\frac{1}{2}\theta \| \rho \mathbf{U} \|^2_{L^2(\Omega)} - \rho_1 \left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right) \| \mathbf{U}_1 \|_{L^2(\Omega_1)} \| \mathbf{U}_2 \|_{L^2(\Omega_2)} \\
\geq \ & \frac{1}{2}\rho_1 \left(1 - \frac{\varepsilon_2}{\varepsilon_1}\right) \| \mathbf{U} \|^2_{L^2(\Omega)}.
\end{aligned}
$$

Therefore, $\mathbf{a}((\mathbf{U},p),(\mathbf{V},q)) \geq C_0 \| (\mathbf{U},p) \|^2_{\mathbb{X}}$, for some constant $C_0 > 0$. On the other hand, it is easy to see from (2.9) that there exists a constant $C > 0$ such that $\| (\mathbf{V},q) \|_{\mathbb{X}} \leq C \| (\mathbf{U},p) \|_{\mathbb{X}}$. This proves the desired result (2.6).

Now we turn to the proof of (2.7). For any $(\mathbf{V},q) \in \mathbb{X}$, let $\varphi \in H^1_0(\Omega)$ be the weak solution of the problem

$$
\langle k_0^2 \varepsilon \nabla\varphi, \nabla\psi \rangle = \langle \rho \mathbf{V}, \nabla\psi \rangle \quad \forall \psi \in H^1_0(\Omega). \tag{2.11}
$$

It is easy to see that

$$
\| \nabla\varphi \|_{L^2(\Omega)} \leq \frac{\rho_2}{k_0^2 \varepsilon_2} \| \mathbf{V} \|_{L^2(\Omega)}. \tag{2.12}
$$

We set

$$
\mathbf{U} = k_0^2 \mathbf{V} - \zeta \nabla q + k_0^2 \nabla\varphi, \quad p = \Lambda q,
$$

where $\zeta = k_0^2 \varepsilon_2 / \rho_2$. It is clear that $(\mathbf{U},p) \in \mathbb{X}$, and we have

$$
\begin{aligned}
\mathbf{a}((\mathbf{U},p),(\mathbf{V},q)) \ = \ & \| \nabla \times \mathbf{V} \|^2_{L^2(\Omega)} + \langle \rho (k_0^2 \mathbf{V} - \zeta \nabla q + k_0^2 \nabla\varphi), \mathbf{V} \rangle + \langle \Lambda \nabla q, \mathbf{V} \rangle \\
& - \langle \nabla q, \varepsilon (k_0^2 \mathbf{V} - \zeta \nabla q + k_0^2 \nabla\varphi) \rangle + \Lambda \langle \varepsilon q, q \rangle.
\end{aligned}
$$

By (2.11) we have $\langle \rho \nabla\varphi, \mathbf{V} \rangle = k_0^2 \langle \varepsilon \nabla\varphi, \nabla\varphi \rangle \geq 0$, and

$$
\langle \Lambda \nabla q, \mathbf{V} \rangle - \langle \nabla q, \varepsilon (k_0^2 \mathbf{V} + k_0^2 \nabla\varphi) \rangle = \langle (\Lambda - k_0^2 \varepsilon - \rho) \mathbf{V}, \nabla q \rangle = 0.
$$

Thus

$$
\begin{aligned}
\mathbf{a}((\mathbf{U},p),(\mathbf{V},q)) \ \geq \ & \| \nabla \times \mathbf{V} \|^2_{L^2(\Omega)} + k_0^2 \langle \rho \mathbf{V}, \mathbf{V} \rangle - \zeta \langle \rho \nabla q, \mathbf{V} \rangle \\
& + \zeta \langle \varepsilon \nabla q, \nabla q \rangle + \Lambda \langle \varepsilon q, q \rangle.
\end{aligned}
$$

By Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
\zeta |\langle \rho \nabla q, \mathbf{V} \rangle| \ & \leq \ \frac{1}{2} \zeta \langle \varepsilon \nabla q, \nabla q \rangle + \frac{1}{2} \langle \rho^2 \zeta \varepsilon^{-1} \mathbf{V}, \mathbf{V} \rangle \\
& = \ \frac{1}{2} \zeta \langle \varepsilon \nabla q, \nabla q \rangle + \frac{1}{2} k_0^2 \langle \frac{\rho}{\varepsilon} \frac{\varepsilon_2}{\rho_2} \rho \mathbf{V}, \mathbf{V} \rangle \\
& \leq \ \frac{1}{2} \zeta \langle \varepsilon \nabla q, \nabla q \rangle + \frac{1}{2} k_0^2 \langle \rho \mathbf{V}, \mathbf{V} \rangle,
\end{aligned}
$$

which implies that, for some constant $C_1 > 0$, $\mathbf{a}((\mathbf{U},p),(\mathbf{V},q)) \geq C_1 \| (\mathbf{V},q) \|^2_{\mathbb{X}}$. On the other hand, it is easy to see from (2.12) that $\| (\mathbf{U},p) \|_{\mathbb{X}} \leq C \| (\mathbf{V},q) \|_{\mathbb{X}}$. This proves (2.7) and completes the proof of the theorem.

It follows from Theorem 2.1 that there exists a unique bounded operator $T \ \mathbb{X} \to \mathbb{X}$ satisfying

$$
\mathbf{a}(T(\mathbf{U},p),(\mathbf{V},q)) = \mathbf{b}((\mathbf{U},p),(\mathbf{V},q)) \quad \forall (\mathbf{V},q) \in \mathbb{X}, \forall (\mathbf{U},p) \in \mathbb{X}. \tag{2.13}
$$

Let $\mathbb{H} = [L^2(\Omega)]^2 \times L^2(\Omega)$. It is clear that T is also well-defined on $\mathbb{H}$. By using the compactness embedding of $H_0(\text{curl};\Omega) \cap H(\text{div}_\varepsilon;\Omega)$ in $[L^2(\Omega)]^2$ in [12], it is shown in [11, Theorem 4] that $T : \mathbb{H} \to \mathbb{H}$ is compact.

Because of the inf-sup condition proved in Theorem 2.1, there is no nontrivial fields $(\mathbf{E}, E_3)$ satisfying (2.3) for $\lambda = 0$. Thus $(\lambda, (\mathbf{E}, E_3))$ solves (2.3) if and only if $(\lambda^{-1}, (\mathbf{E}, E_3))$ is an eigenpair of the compact operator $T : \mathbb{H} \to \mathbb{H}$. Hence the spectral properties of the compact operator T provide the information of the spectral properties of the variationally posed eigenvalue problem (2.3).

Now we briefly recall the spectral property of compact operators. Further details can be found in [1], for example. The resolvent set of a compact operator T, $\rho(T)$, is the set of complex numbers z such that $(z-T)$ has a bounded inverse operator in $\mathbb{H}$. The spectrum of T is the set $\sigma(T) = \mathbb{C}\backslash\rho(T)$. Classical spectral theory of compact operators implies that $\sigma(T)$ is countable with no nonzero limit points. Nonzero numbers in $\sigma(T)$ are eigenvalues.

Let $\mu \in \sigma(T)$ be nonzero. The ascent α of $\mu - T$ is the smallest number such that $N((\mu - T)^\alpha) = N((\mu - T)^{\alpha+1})$, where N denotes the null space. $N((\mu - T)^\alpha)$ is finite dimensional and its dimension is called the algebraic multiplicity of μ. The vectors in $N((\mu - T)^\alpha)$ is called generalized eigenvectors of T corresponding to μ. The vectors in $N(\mu - T)$ is the eigenvectors of T corresponding to μ. The geometric multiplicity of μ is the dimension of $N(\mu - T)$ which is less than the algebraic multiplicity.

3. The Discrete Problem

Let $\mathcal{M}_h$ be a shape regular triangulation of Ω. For any $K \in \mathcal{M}_h$, we denote h_K its diameter, and $h = \max_{K\in\mathcal{M}_h} h_K$. Let $Q_h \subset H_0^1(\Omega)$ be the standard conforming linear finite element space, and $\mathbf{W}_h \subset H_0(\text{curl};\Omega)$ be the finite element space of the lowest order $H(\text{curl};\Omega)$ conforming edge element

$$\mathbf{W}_h = \{\mathbf{V}_h \in H_0(\text{curl};\Omega) \quad : \quad \mathbf{V}_h|_K = (a_K - c_K x_2, b_K + c_K x_1)^T, \\ \text{where } a_K, b_K, c_K \in \mathbf{R}, \quad K \in \mathcal{M}_h\}.$$

Denote $\mathbb{X}_h = \mathbf{W}_h \times Q_h$. Then we introduce the following finite element approximation of the variationally posed eigenvalue problem (2.3) as follows:

For a given frequency $\omega > 0$, find all pairs $(\lambda_h, (\mathbf{E}_h, E_{3h})) \in \mathbb{C} \times \mathbb{X}_h$, such that

$$\mathbf{a}((\mathbf{E}_h, E_{3h}), (\mathbf{V}_h, q_h)) = \lambda_h \mathbf{b}((\mathbf{E}_h, E_{3h}), (\mathbf{V}_h, q_h)) \qquad \forall (\mathbf{V}_h, q_h) \in \mathbb{X}_h. \tag{3.1}$$

Notice that $\nabla Q_h \subset \mathbf{W}_h$. The discrete inf-sup conditions formulated in the following theorem can be proved exactly as in Theorem 2.1.

Theorem 3.1. *Assume that the parameter* Λ *satisfies*

$$\max(k_0^2\varepsilon_1, k_0^2\varepsilon_2) < \Lambda \le k_0^2(\varepsilon_1 + \varepsilon_2),$$

then the sesquilinear form $\mathbf{a}(\cdot,\cdot)$ *satisfies the following properties:*

(i) *There exists a constant* $\hat{\alpha}_1 > 0$ *independent of* h *such that, for any* $(\mathbf{U}_h, p_h) \in \mathbb{X}_h$,

$$\sup_{(\mathbf{V}_h, q_h)\in\mathbb{X}} \frac{\mathbf{a}((\mathbf{U}_h, p_h), (\mathbf{V}_h, q_h))}{\|(\mathbf{V}_h, q_h)\|_{\mathbb{X}}} \ge \hat{\alpha}_1 \|(\mathbf{U}_h, p_h)\|_{\mathbb{X}}. \tag{3.2}$$

(ii) *There exists a constant* $\hat{\alpha}_2 > 0$ *independent of* h *such that, for any* $(\mathbf{V}_h, q_h) \in \mathbb{X}$,

$$\sup_{(\mathbf{U}_h, p_h)\in\mathbb{X}} \frac{\mathbf{a}((\mathbf{U}_h, p_h), (\mathbf{V}_h, q_h))}{\|(\mathbf{U}_h, p_h)\|_{\mathbb{X}}} \ge \hat{\alpha}_2 \|(\mathbf{V}_h, q_h)\|_{\mathbb{X}}. \tag{3.3}$$

Theorem 3.1 implies that there exists a unique bounded operator $T_h : \mathbb{X}_h \to \mathbb{X}_h$ satisfying

$$\mathbf{a}(T_h(\mathbf{U}_h, p_h), (\mathbf{V}_h, q_h)) = \mathbf{b}((\mathbf{U}_h, p_h), (\mathbf{V}_h, q_h)) \quad \forall (\mathbf{V}_h, q_h) \in \mathbb{X}_h. \tag{3.4}$$

We also have that $(\lambda_h, (\mathbf{E}_h, E_{3h}))$ solves (3.1) if and only if $(\lambda_h^{-1}, (\mathbf{E}_h, E_{3h}))$ is an eigenpair of the operator $T_h : \mathbb{H}_h \to \mathbb{H}_h$.

The operator T_h can be written as $T_h = P_h T$, where P_h is the projection from $\mathbb{X}$ to $\mathbb{X}_h$ defined by

$$\mathbf{a}(P_h(\mathbf{U}, p), (\mathbf{V}_h, q_h)) = \mathbf{a}((\mathbf{U}, p), (\mathbf{V}_h, q_h)) \quad \forall (\mathbf{V}_h, q_h) \in \mathbb{X}_h, \ \ (\mathbf{U}, p) \in \mathbb{X}.$$

From Theorem 3.1 we easily deduce that $P_h \to I$ pointwise as $h \to 0$. Since $T : \mathbb{H} \to \mathbb{H}$ is compact, $T_h \to T$ in the operator norm on $\mathbb{H}$. Let λ be an eigenvalue of (2.3) with algebraic multiplicity m, by which we mean that λ^{-1} is an eigenvalue of T with algebraic multiplicity m. Let α be the ascent of $\lambda^{-1} - T$. Since $T_h \to T$ in norm, m eigenvalues $\lambda_1(h), \cdots, \lambda_m(h)$ of (3.1) will converge to λ. The $\lambda_j(h)$ are counted according to the algebraic multiplicities of $\lambda_j(h)^{-1}$ as eigenvalues of T_h.

Let

$$\begin{aligned} \mathbb{M} = \mathbb{M}(\lambda) &= \{(\mathbf{U}, p) : (\mathbf{U}, p) \text{ is a generalized eigenvector of (2.3)} \\ &\quad\ \text{corresponding to } \lambda, \|(\mathbf{U}, p)\|_{\mathbb{X}} = 1.\} \\ \mathbb{M}^* = \mathbb{M}^*(\lambda) &= \{(\mathbf{V}, q) : (\mathbf{V}, q) \text{ is a generalized adjoint eigenvector} \\ &\quad\ \text{of (2.3) corresponding to } \lambda, \|(\mathbf{V}, q)\|_{\mathbb{X}} = 1.\} \end{aligned}$$

and define

$$\begin{aligned} \varepsilon_h &= \varepsilon_h(\lambda) = \sup_{(\mathbf{U}, p) \in \mathbb{M}} \inf_{(\mathbf{F}_h, g_h) \in \mathbb{X}_h} \|(\mathbf{U}, p) - (\mathbf{F}_h, g_h)\|_{\mathbb{X}} \\ \varepsilon_h^* &= \varepsilon_h^*(\lambda) = \sup_{(\mathbf{V}, q) \in \mathbb{M}^*} \inf_{(\mathbf{F}_h, g_h) \in \mathbb{X}_h} \|(\mathbf{V}, q) - (\mathbf{F}_h, g_h)\|_{\mathbb{X}}. \end{aligned}$$

The following theorem which extends the result in [1] is proved in [11].

Theorem 3.2. *There is a constant C such that*

$$|\lambda - \lambda_j(h)|^\alpha \le C \varepsilon_h \varepsilon_h^*, \quad j = 1, \cdots, m.$$

This theorem shows that the rate of convergence depends on the ascent α and the interpolation error for the eigenspaces, which in turn depends on the regularity of the corresponding eigenfunctions.

Now we consider a special consequence of Theorem 3.2. Let $\Omega_1 \subset\subset \Omega$ and the interface Γ is smooth (say C^2). We assume the mesh $\mathcal{M}_h$ of Ω is so constructed that the domain Ω_1 is approximated by a domain Ω_1^h with a polygonal boundary Γ_h whose vertexes all lie on the interface Γ. Let Ω_2^h stand for the domain with $\partial\Omega$ and Γ_h as the exterior and interior boundaries. In addition, we assume each $K \in \mathcal{M}_h$ is either in Ω_1^h or in Ω_2^h, and has at most two vertexes lying on Γ_h. It is proved in [4, Lemma 2.1] that for any $v \in Y = H^1(\Omega) \cap H^2(\Omega_1) \cap H^2(\Omega_2)$, the following error estimate holds

$$\| v - I_h v \|_{L^2(\Omega)} + h \| \nabla (v - I_h v) \|_{L^2(\Omega)} \le C h^2 |\log h|^{1/2} \| v \|_Y, \tag{3.5}$$

where $I_h : C(\bar{\Omega}) \to Q_h$ is the standard linear finite element nodal interpolant.

Similar to the argument in [4, Lemma 2.1], we can show that

$$\| \mathbf{U} - \Pi_h \mathbf{U} \|_{\mathrm{curl}, \Omega} \le C h |\log h|^{1/2} \| \mathbf{U} \|_Y, \tag{3.6}$$

where $\Pi_h : H^1(\Omega) \cap H_0(\text{curl};\Omega) \to \mathbf{W}_h$ is the canonical interpolant defined through the relation

$$\int_e \Pi_h \mathbf{U} \cdot \tau \, ds = \int_e \mathbf{U} \cdot \tau \, ds,$$

for any edges e of the mesh $\mathcal{M}_h$, where τ is the unit tangential vector along e.

The following result is a direct consequence of Theorem 3.2 and the estimates (3.5)-(3.6).

Corollary 3.3. *Let $\lambda \in \mathbb{C}$ be an eigenvalue of* (2.3) *with algebraic multiplicity m and ascent α. Assume the corresponding generalized eigenvectors $(\mathbf{E}, E_3)$ and adjoint eigenvectors $(\mathbf{F}, F_3)$ have the regularity properties: $\mathbf{E}, \mathbf{F} \in [H^2(\Omega_1) \cap H^2(\Omega_2)]^2$ and $E_3, F_3 \in H^2(\Omega_1) \cap H^2(\Omega_2)$. Let $\lambda_1(h), \cdots, \lambda_m(h)$ be the eigenvalues of* (3.1) *that converge to λ as $h \to 0$. Then there exists a constant C independent of h such that*

$$|\lambda - \lambda_j(h)|^\alpha \le Ch^2 |\log h|, \quad j = 1, 2, \cdots, m.$$

Things become more complicated, however, when the interface Γ is not smooth. In this situation, it is well-known that the eigenfunctions usually display singularities which deteriorate the finite element convergence if uniform mesh refinements are used. We will show one such situation in next section. One possible way to overcome this difficulty is to use adaptive mesh refinements based on a posteriori error estimation. This is a topic of our current research.

4. Numerical Experiments

In this section we report several numerical examples to illustrate the performance of the method studied in this paper. In the computations we used the PDE toolbox of MATLAB. The stiffness matrix in (3.1) is assembled by choosing $\Lambda = 0$. The discrete algebraic eigenvalue problems are solved by shifted inverse iteration algorithm with shift κ.

Example 1. This example is taken from [7] which concerns the simplest semi-filled rectangular waveguide. Let $\Omega = [0, a] \times [0, \frac{1}{2}a]$ with $a = 2.00\mu m$. Set $\Omega_1 = [0, a] \times [0, \frac{1}{4}a]$ and $\Omega_2 = [0, a] \times [\frac{1}{4}a, \frac{1}{2}a]$ (see Figure 4.1). We take $\varepsilon_1 = 4$ and $\varepsilon_2 = 1$.

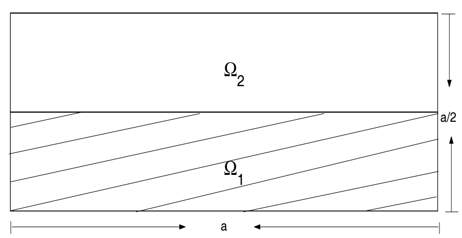

Figure 4.1: Semi-filled rectangular dielectric waveguide

Let $k_0 = \omega\sqrt{\epsilon_0\mu_0} = 1.25$ and $\kappa = 0$. The wave analysis in [6] indicates that the minimum wave number $\beta^2 \simeq 0.42044$. Table 4.1 shows the computed eigenvalues λ_k and the corresponding error $|\lambda_k - \beta^2|$ on successively uniformly refined meshes $\mathcal{M}_k$ with N_k nodes.

Figure 4.2 shows clearly that the meshes and the associated numerical complexity are quasi-optimal: $|\lambda_k - \beta^2| \approx CN_k^{-1}$ is valid asymptotically for different choices of $k_0 = 1.25, 1.5$, and 2.5. The performance of the quasi-optimal method is indicated by the dotted line of slope -1.

Example 2. We consider step-index circular waveguide.This example is taken form [9] and concerns dielectric waveguide. Let $\Omega = [0, a] \times [0, a]$ be a square with $a = 20.00\mu m$. Let $\Omega_1 \subset \Omega$

Table 4.1: Semi-filled rectangular dielectric waveguide: The level of mesh refinements k, the number of nodes N_k, the computed minimum eigenvalue λ_k, and the error $|\lambda_k - \beta^2|$ when $k_0 = 1.25$.

k	N_k	λ_k	$\lvert\lambda_k - \beta^2\rvert$
1	13	0.47398020	0.05354020
2	41	0.43877230	0.01833230
3	145	0.42634740	0.00590740
4	545	0.42209659	0.00165659
5	2113	0.42087439	0.00043439
6	8321	0.42055350	0.00011350
7	33025	0.42047198	0.00003198

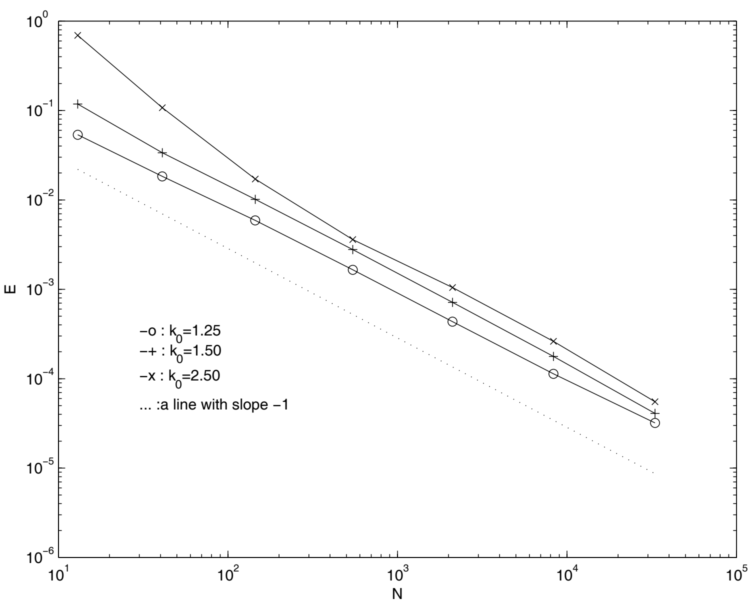

Figure 4.2: Semi-filled rectangular dielectric waveguide: Performance of the error in terms of the number of nodes of the meshes. The quasi-optimal decay is indicated by the dotted line of slope -1.

be a circle with the radius $r = 4.50\mu m$, and $\Omega_2 = \Omega \backslash \overline{\Omega}_1$ (see Figure 4.3). Let $\varepsilon_1 = 11.66$ and $\varepsilon_2 = 10.02$.

Let $k_0 = 4.833219$ and $\kappa = 270$. The analytical solution for the fundamental mode is $\beta \approx 3.4130933k_0$. We still observe the asymptotically quasi-optimal decay of the error in terms of the mesh complexity in Figure 4.4.

Example 3. This example also is taken form [9] but concerns the dielectric rib waveguide. Let $\Omega = [0, a] \times [0, h]$ with $a = 4.00\mu m$ and $h = 3.00\mu m$. The rib waveguide is sketched in Figure 4.5, where $a_1 = 2.00\mu m$, $h_1 = 1.00\mu m$, $h_2 = 1.10\mu m$, and $h_3 = 1.80\mu m$. Let $\varepsilon_1 = 11.1556, \varepsilon_2 = 11.8336$, and $\varepsilon_3 = 1$.

Let $k_0 = 4.053668$ and $\kappa = 185$. The analytical solution is $\beta \approx 3.388687k_0$. Figure 4.6 indicates that $|\lambda_h - \beta^2| \approx CN_k^{-1/2}$ which is not quasi-optimal. This can be explained by the

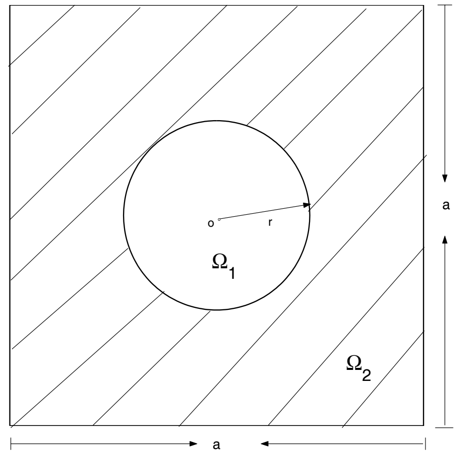

Figure 4.3: Step-index circular dielectric waveguide: Performance of the error in terms of the number of nodes of the meshes. The quasi-optimal decay is indicated by the dotted line of slope −1.

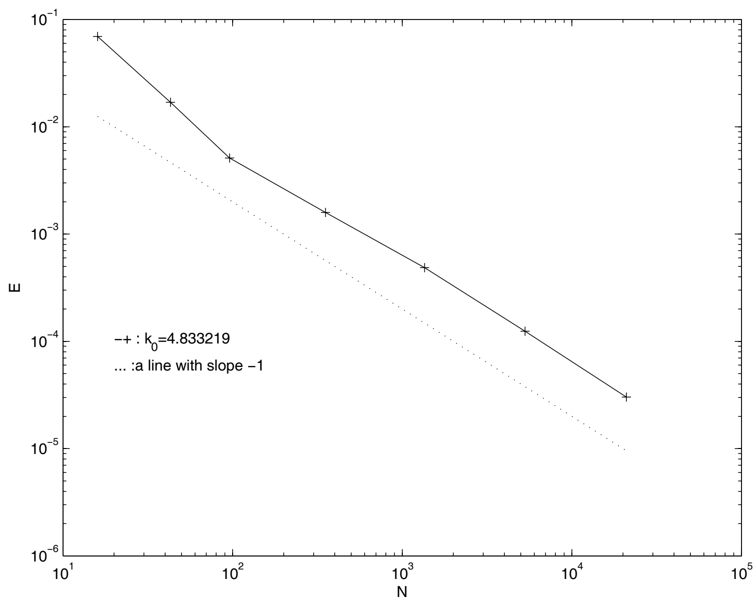

Figure 4.4: Step-index circular dielectric waveguide: Performance of the error in terms of the number of nodes of the meshes. The quasi-optimal decay is indicated by the dotted line of slope −1.

fact that that for the rib waveguide, the eigenfunctions have singularities which deteriorate the finite element convergence. Adaptive finite element methods based on a posteriori error estimates are known to be successful in resolving this difficulty [3]. We will report progress in this direction in a future work.

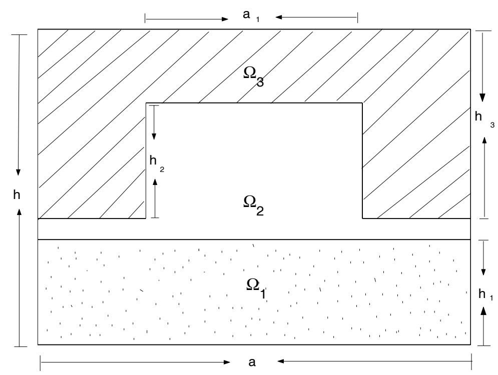

Figure 4.5: Rib dielectric waveguide

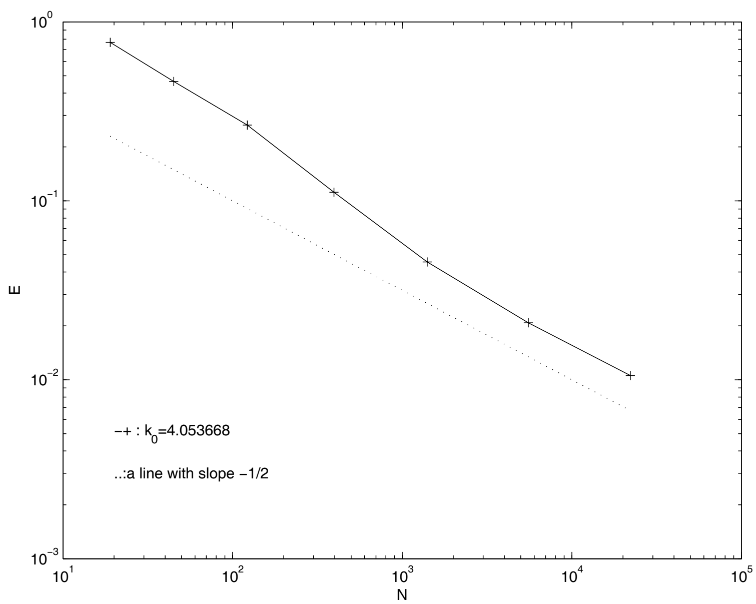

Figure 4.6: Rib dielectric waveguide: Performance of the error in terms of the number of nodes of the meshes. A dotted line of slope $-1/2$ is depicted.

References

[1] I.Babuška and J.Osborn, Eigenvalue Problems, in Handbook of Numerical Analysis, Elsevier-North Holland Amsterdam, 1991.

[2] A.Bermúdez and D.G.Padreira, Mathematical analysis of a finite element method without spurious for computation of dielectric waveguides, *Numer. Math.*, **61** (1992), 39-57.

[3] Z. Chen and S. Dai, On the efficiency of adaptive finite element methods for elliptic problems with discontinuous coefficients, *SIAM J. Sci. Comput.*, **24** (2002), 443-462.

[4] Z. Chen and J. Zou, Finite element methods and their convergence for elliptic and parabolic interface problems, *Numer. Math.*, **79** (1998), 175-202.

[5] K.S. Chiang, Review of numerical and approximate methods for the modal analysis of general optical dielectric waveguide, Optical and Quantum Electronics, **26** (1994), 113-134.

[6] R.F.Harrington, Time-Harmonic Electromagnetic Fields, New York: McGraw-Hill, 1961.

[7] J.M. Jin, The Finite Element Methods in Electromagnetics New York: John Wiley & Sons, 1993.

[8] P.Joly, C.Poirier, J.E.Roberts, and P.Trouve, A New Nonconforming Finite Element Method for the Computation of Electromagnetic Guided Waves I: Mathematical Analysis, *SIAM J. Numer. Anal.*, **33**:4 (1996), 1494-1525.

[9] P.Lüsse, P.Stuwe, and H.-G.Unger. Analysis of Vectorial Mode Fields in Optical Waveguides by a New Finite Diffrernce Method, *J. Lightwave Technology*, **12**:3 (1994), 487-493.

[10] J.C.Nedelec, Mixed finite elements in $\mathbb{R}^3$, *Numer. Math.*, **35** (1980), 315-341.

[11] L.Vardapetyan and L.Demkowicz, Full-wave analysis of dielectric waveguides at a given frequency, *Math. Comp.*, **72**:241 (2002), 105-129.

[12] C.Weber, A local compactness theorem for Maxwell's equations, *Math. Meth. Appl. Sci.*, **2** (1980), 12-25.