Computation of Electromagnetic Waveguide Eigenmodes

Merle Backmeyer

December 2023

Contents

1 Introduction

A cylindrical waveguide, being z-invariant in geometry, is a well-studied structure in electromagnetic theory. Determining the eigenmodes of wave propagation in such waveguides has been extensively investigated for isotropic materials, including mode decomposition into TE/TM-waves [1]. However, this project seeks to extend this analysis to anisotropic materials, which exhibit varying electromagnetic properties in different directions, thus introducing added complexity. This research will expand the existing knowledge base by developing a comprehensive methodology for eigenmode determination in anisotropic waveguides. The boundary conditions for the waveguides under investigation will be assumed to be perfectly conducting, which is a common scenario in practical applications.

A time-harmonic field response is assumed and considering the z-invariance of the geometry,

$$\mathbf{E}(x, y, z, t) = \mathbf{E}(x, y) \exp(\mathrm{j}\omega t) \exp(-\gamma z) \tag{1}$$

with ω being the angular frequency and $\gamma = \alpha + \mathrm{j}\beta$ the propagation constant with attenuation (α) and phase constant (β).

The general problem, formulated using the electric field $\mathbf{E}$, reads

$$\nabla \times \left(\boldsymbol{\mu_r}^{-1} \nabla \times \mathbf{E}\right) - k_0^2 \boldsymbol{\varepsilon_r} \mathbf{E} = 0 \quad \text{on } \Omega \tag{2}$$

$$\mathbf{n} \times \mathbf{E} = 0 \quad \text{on } \delta\Omega \tag{3}$$

with $k_0 = \omega\sqrt{\mu_0 \varepsilon_0}$ [2]. Note that in the anisotropic case $\boldsymbol{\mu_r}$ and $\boldsymbol{\varepsilon_r}$ are tensors.

In the case of conductive losses, $\boldsymbol{\varepsilon_r}$ will be a complex number defined as

$$\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{r}} = \boldsymbol{\varepsilon_r} + \mathrm{j}\frac{\boldsymbol{\sigma}}{\omega\varepsilon_0}. \tag{4}$$

This project will be based on the Finite Element Method (FEM) to numerically solve the wave equation for anisotropic materials. This approach will be implemented within the open-source software SparseLizard (`https://www.sparselizard.org/`). In section 2, the weak formulation of the wave equation for anisotropic material conditions is derived from Maxwell's equations. Three different approaches are presented to validate the correctness of the solution. The weak formulation will be implemented in the C++ finite element library `SparseLizard`, which exploits `SLEPc` package as an eigenvalue solver. These two open-source libraries are presented in sections 3 and 4. Finally, in section 5 the complete numerical implementation is presented and in section 6 it is validated for many different scenarios and compared again results from papers, analytical formulas and other simulation software.

2 Derivation of Wave Equation for Anisotropic Materials

Maxwell's equations state the relationship between the electric field strength $\mathbf{E}$, the displacement field $\mathbf{D}$, the magnetic field strength $\mathbf{H}$, the magnetic flux $\mathbf{B}$, and the free charge density ϱ_f. They read

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{5}$$

$$\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t} \tag{6}$$

$$\boldsymbol{\nabla} \cdot \mathbf{D} = \varrho_f \tag{7}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0. \tag{8}$$

with the corresponding material laws

$$\mathbf{D} = \boldsymbol{\varepsilon}\mathbf{E} \tag{9}$$

$$\mathbf{B} = \boldsymbol{\mu}\mathbf{H} \tag{10}$$

$$\mathbf{J} = \boldsymbol{\sigma}\mathbf{E}, \tag{11}$$

with permittivity $\boldsymbol{\varepsilon}$, permeability $\boldsymbol{\mu}$, and conductivity $\boldsymbol{\sigma}$.

The general ansatz for the electric and magnetic field strength for a waveguide with z-invariant geometry and a time-harmonic excitation is

$$\mathbf{E}(x, y, z, t) = \tilde{\mathbf{E}}(x, y) \exp(j\omega t) \exp(-\gamma z) \tag{12}$$

$$\mathbf{H}(x, y, z, t) = \tilde{\mathbf{H}}(x, y) \exp(j\omega t) \exp(-\gamma z) \tag{13}$$

The aim is to obtain the propagation constant γ for a given frequency ω. To obtain the weak formulation of the eigenvalue problem with the eigenpair $(\gamma, \mathbf{E})$, the following three steps need to be performed:

- Replace z-derivatives with a multiplication with $-\gamma$
- Obtain the variational formulation
- Formulate the problem with only the unknown E and substitute H

Regardless of the order, they should result in the same formulation. For validation purposes, three different approaches are performed. There are three necessary steps to get to the eigenvalue problem:

- Derivation in terms of $\mathbf{E}$ only
- Substitution of z-derivatives
- Variational Formulation.

They can be carried out in any order leading to the same resulting formulation.
In the following, three different orders are presented:

1. Derivation in terms of $\mathbf{E}$ - Substitution of z-derivatives - Variational Formulation
2. Derivation in terms of $\mathbf{E}$ - Variational Formulation - Substitution of z-derivatives
3. Substitution of z-derivatives - Derivation in terms of $\mathbf{E}$ - Variational Formulation

First, the wave equation is derived in terms of E only for approach 1 and 2.
To derive the wave equation for E, we take the curl of (5), which results in

$$\boldsymbol{\nabla} \times \boldsymbol{\mu}^{-1} \boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial}{\partial t} \boldsymbol{\nabla} \times \mathbf{H} \tag{14}$$

and then by inserting (6), we obtain

$$\boldsymbol{\nabla} \times \boldsymbol{\mu}^{-1} \boldsymbol{\nabla} \times \mathbf{E} + \boldsymbol{\varepsilon} \frac{\partial^2}{\partial t^2} \mathbf{E} + \boldsymbol{\sigma} \frac{\partial \mathbf{E}}{\partial t} = 0 \tag{15}$$

Exploiting the fact that E is time-harmonic, we end up with

$$\boldsymbol{\nabla} \times \boldsymbol{\mu}^{-1} \boldsymbol{\nabla} \times \mathbf{E} - \boldsymbol{\varepsilon}\omega^2 \mathbf{E} + \mathrm{j}\omega\boldsymbol{\sigma}\mathbf{E} = 0 \tag{16}$$

or with $k_0 = \omega\sqrt{\mu_0 \varepsilon_0}$ we obtain

$$\boldsymbol{\nabla} \times \boldsymbol{\mu}_r^{-1} \boldsymbol{\nabla} \times \mathbf{E} - {k_0}^2 \boldsymbol{\varepsilon}_r \mathbf{E} + \mathrm{j}\omega\mu_0\boldsymbol{\sigma}\mathbf{E} = 0 \tag{17}$$

which in the case of $\boldsymbol{\sigma} = \mathbf{0}$ is equivalent to eq(93) from [2].

The walls are assumed to be perfectly conducting, so the general boundary value problem formulated using the electric field $\mathbf{E}$ reads

$$\nabla \times \left({\boldsymbol{\mu}_r}^{-1} \nabla \times \mathbf{E} \right) - k_o^2 \boldsymbol{\varepsilon}_r \mathbf{E} + \mathrm{j}\omega\mu_0\boldsymbol{\sigma}\mathbf{E} = 0 \quad \text{on } \Omega \tag{18}$$

$$\mathbf{n} \times \mathbf{E} = 0 \quad \text{on } \delta\Omega. \tag{19}$$

Note that in the anisotropic case $\boldsymbol{\mu}_r$ and $\boldsymbol{\varepsilon}_r$ are tensors.

2.1 Approach 1: Wave equation - Substitution of z-derivatives - Variational Formulation

First, we follow approach 1. Starting point is the wave equation, we rewrite it fully in terms of the component-wise derivatives in order to substitute the z-derivatives with $\frac{\partial}{\partial z} = -\gamma$.

For simplicity, let's work with the relative reluctivity $\boldsymbol{\nu}$, defined as the tensor:

$$\boldsymbol{\nu} = \boldsymbol{\mu_r^{-1}} = \begin{bmatrix} \nu_{11} & \nu_{12} & \nu_{13} \\ \nu_{21} & \nu_{22} & \nu_{23} \\ \nu_{31} & \nu_{32} & \nu_{33} \end{bmatrix} \tag{20}$$

Note that the subscript r is omitted here to increase readability. The three terms are treated one after another:

$$\underbrace{\boldsymbol{\nabla} \times \left({\boldsymbol{\mu_r}}^{-1} \, \boldsymbol{\nabla} \times \mathbf{E} \right)}_{[I]} - \underbrace{k_o^2 \boldsymbol{\varepsilon_r} \mathbf{E}}_{[II]} + \underbrace{\mathrm{j}\omega\mu_0 \boldsymbol{\sigma} \mathbf{E}}_{[III]} = 0 \tag{21}$$

First, term $[I]$ is analyzed and expanded:

$$\underbrace{\boldsymbol{\nabla} \times \left({\boldsymbol{\mu_r}}^{-1} \, \boldsymbol{\nabla} \times \mathbf{E} \right)}_{[I]} = \boldsymbol{\nabla} \times \left(\begin{bmatrix} \nu_{11} & \nu_{12} & \nu_{13} \\ \nu_{21} & \nu_{22} & \nu_{23} \\ \nu_{31} & \nu_{32} & \nu_{33} \end{bmatrix} \begin{bmatrix} \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \\ -\frac{\partial E_z}{\partial x} + \frac{\partial E_x}{\partial z} \\ \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \end{bmatrix} \right) \tag{22}$$

$$\begin{aligned} &= \begin{bmatrix} \nu_{31}(\frac{\partial^2 E_z}{\partial y^2} - \frac{\partial^2 E_y}{\partial y \partial z}) - \nu_{21}(\frac{\partial^2 E_z}{\partial y \partial z} - \frac{\partial^2 E_y}{\partial z^2}) \\ \nu_{11}(\frac{\partial^2 E_z}{\partial y \partial z} - \frac{\partial^2 E_y}{\partial z^2}) - \nu_{31}(\frac{\partial^2 E_z}{\partial x \partial y} - \frac{\partial^2 E_y}{\partial x \partial z}) \\ \nu_{21}(\frac{\partial^2 E_z}{\partial x \partial y} - \frac{\partial^2 E_y}{\partial x \partial z}) - \nu_{11}(\frac{\partial^2 E_z}{\partial y^2} - \frac{\partial^2 E_y}{\partial y \partial z}) \end{bmatrix} + \begin{bmatrix} \nu_{32}(-\frac{\partial^2 E_z}{\partial x \partial y} + \frac{\partial^2 E_x}{\partial y \partial z}) - \nu_{22}(-\frac{\partial^2 E_z}{\partial x \partial z} + \frac{\partial^2 E_x}{\partial z^2}) \\ \nu_{12}(-\frac{\partial^2 E_z}{\partial x \partial z} + \frac{\partial^2 E_x}{\partial z^2}) - \nu_{32}(-\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_x}{\partial x \partial z}) \\ \nu_{22}(-\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_x}{\partial x \partial z}) - \nu_{12}(-\frac{\partial^2 E_z}{\partial x \partial y} + \frac{\partial^2 E_x}{\partial y \partial z}) \end{bmatrix} \\ &+ \begin{bmatrix} \nu_{33}(\frac{\partial^2 E_y}{\partial x \partial y} - \frac{\partial^2 E_x}{\partial y^2}) - \nu_{23}(\frac{\partial^2 E_y}{\partial x \partial z} - \frac{\partial^2 E_x}{\partial y \partial z}) \\ \nu_{13}(\frac{\partial^2 E_y}{\partial x \partial z} - \frac{\partial^2 E_x}{\partial y \partial z}) - \nu_{33}(\frac{\partial^2 E_y}{\partial x^2} - \frac{\partial^2 E_x}{\partial x \partial y}) \\ \nu_{23}(\frac{\partial^2 E_y}{\partial x^2} - \frac{\partial^2 E_x}{\partial x \partial y}) - \nu_{13}(\frac{\partial^2 E_y}{\partial x \partial y} - \frac{\partial^2 E_x}{\partial y^2}) \end{bmatrix} \end{aligned} \tag{23}$$

We now replace $\frac{\partial}{\partial z} = -\gamma$.

$$\begin{aligned} &= \begin{bmatrix} \nu_{31}(\frac{\partial^2 E_z}{\partial y^2} + \gamma\frac{\partial E_y}{\partial y}) - \nu_{21}(-\gamma\frac{\partial E_z}{\partial y} - \gamma^2 E_y) \\ \nu_{11}(-\gamma\frac{\partial E_z}{\partial y} - \gamma^2 E_y) - \nu_{31}(\frac{\partial^2 E_z}{\partial x \partial y} + \gamma\frac{\partial E_y}{\partial x}) \\ \nu_{21}(\frac{\partial^2 E_z}{\partial x \partial y} + \gamma\frac{\partial E_y}{\partial x}) - \nu_{11}(\frac{\partial^2 E_z}{\partial y^2} + \gamma\frac{\partial E_y}{\partial y}) \end{bmatrix} + \begin{bmatrix} \nu_{32}(-\frac{\partial^2 E_z}{\partial x \partial y} - \gamma\frac{\partial E_x}{\partial y}) - \nu_{22}(\gamma\frac{\partial E_z}{\partial x} + \gamma^2 E_x) \\ \nu_{12}(\gamma\frac{\partial E_z}{\partial x} + \gamma^2 E_x) + \nu_{32}(\frac{\partial^2 E_z}{\partial x^2} + \gamma\frac{\partial E_x}{\partial x}) \\ \nu_{22}(-\frac{\partial^2 E_z}{\partial x^2} - \gamma\frac{\partial E_x}{\partial x}) + \nu_{12}(\frac{\partial^2 E_z}{\partial x \partial y} + \gamma\frac{\partial E_x}{\partial y}) \end{bmatrix} \\ &+ \begin{bmatrix} \nu_{33}(\frac{\partial^2 E_y}{\partial x \partial y} - \frac{\partial^2 E_x}{\partial y^2}) - \nu_{23}(-\gamma\frac{\partial E_y}{\partial x} + \gamma\frac{\partial E_x}{\partial y}) \\ \nu_{13}(-\gamma\frac{\partial E_y}{\partial x} + \gamma\frac{\partial E_x}{\partial y}) - \nu_{33}(\frac{\partial^2 E_y}{\partial x^2} - \frac{\partial^2 E_x}{\partial x \partial y}) \\ \nu_{23}(\frac{\partial^2 E_y}{\partial x^2} - \frac{\partial^2 E_x}{\partial x \partial y}) - \nu_{13}(\frac{\partial^2 E_y}{\partial x \partial y} - \frac{\partial^2 E_x}{\partial y^2}) \end{bmatrix} \end{aligned} \tag{24}$$

Now, the system will be brought together again and we introduce a splitting of the **E**-field into the transversal and longitudinal components, namely

$$\mathbf{E} = \mathbf{E_t} + \mathbf{E_z} = \begin{bmatrix} E_x \\ E_y \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ E_z \end{bmatrix}.$$

To bring it together, it makes sense to order the terms depending on their coefficients $(1, \gamma, \gamma^2)$ and sort the components of the **E**-field into transversal and longitudinal components.

$$\begin{aligned} &= \begin{bmatrix} \nu_{33}(\frac{\partial^2 E_y}{\partial x \partial y} - \frac{\partial^2 E_x}{\partial y^2}) \\ -\nu_{33}(\frac{\partial^2 E_y}{\partial x^2} - \frac{\partial^2 E_x}{\partial x \partial y}) \\ \nu_{23}(\frac{\partial^2 E_y}{\partial x^2} - \frac{\partial^2 E_x}{\partial x \partial y}) - \nu_{13}(\frac{\partial^2 E_y}{\partial x \partial y} - \frac{\partial^2 E_x}{\partial y^2}) \end{bmatrix} + \begin{bmatrix} \nu_{31}(\frac{\partial^2 E_z}{\partial y^2}) - \nu_{32}(\frac{\partial^2 E_z}{\partial x \partial y}) \\ -\nu_{31}(\frac{\partial^2 E_z}{\partial x \partial y}) + \nu_{32}(\frac{\partial^2 E_z}{\partial x^2}) \\ \nu_{21}(\frac{\partial^2 E_z}{\partial x \partial y}) - \nu_{11}(\frac{\partial^2 E_z}{\partial y^2}) + \nu_{22}(-\frac{\partial^2 E_z}{\partial x^2}) + \nu_{12}(\frac{\partial^2 E_z}{\partial x \partial y}) \end{bmatrix} \\ &+ \gamma \begin{bmatrix} \nu_{31}\frac{\partial E_y}{\partial y} - \nu_{32}\frac{\partial E_x}{\partial y} + \nu_{23}(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}) \\ -\nu_{31}\frac{\partial E_y}{\partial x} + \nu_{32}\frac{\partial E_x}{\partial x} + \nu_{13}(-\frac{\partial E_y}{\partial x} + \frac{\partial E_x}{\partial y}) \\ \nu_{21}\frac{\partial E_y}{\partial x} - \nu_{11}\frac{\partial E_y}{\partial y} - \nu_{22}\frac{\partial E_x}{\partial x} + \nu_{12}\frac{\partial E_x}{\partial y} \end{bmatrix} + \gamma \begin{bmatrix} \nu_{21}\frac{\partial E_z}{\partial y} - \nu_{22}\frac{\partial E_z}{\partial x} \\ -\nu_{11}\frac{\partial E_z}{\partial y} + \nu_{12}\frac{\partial E_z}{\partial x} \\ 0 \end{bmatrix} - \gamma^2 \begin{bmatrix} -\nu_{21}E_y + \nu_{22}E_x \\ \nu_{11}E_y - \nu_{12}E_x \\ 0 \end{bmatrix} \end{aligned} \tag{25}$$

The first two lines of the system can be put back together as

$$\begin{aligned} [\cdot]_{(xy)} = \boldsymbol{\nabla}_\mathrm{t} \times &\left(\begin{bmatrix} * & * & 0 \\ * & * & 0 \\ * & * & \nu_{33} \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t} \right) - \boldsymbol{\nabla}_\mathrm{t} \times \left(\begin{bmatrix} 0 & 0 & * \\ 0 & 0 & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t} \, E_z \right) - \gamma \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t} \, E_z \\ &- \gamma \, \boldsymbol{\nabla}_\mathrm{t} \times \left(\begin{bmatrix} 0 & 0 & * \\ 0 & 0 & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} \mathbf{E_t} \right) + \gamma \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & 0 \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t} - \gamma^2 \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t}, \end{aligned} \tag{26}$$

where the matrix entries denoted by $*$ can take an arbitrary number as they get multiplied by 0 in the calculation anyways.

The third line of the system results in

$$[\cdot]_{(z)} = \boldsymbol{\nabla}_\mathrm{t}\cdot\left(\begin{bmatrix}* & * & \nu_{23}\\ * & * & -\nu_{13}\\ * & * & *\end{bmatrix}\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{E_t}\right) - \boldsymbol{\nabla}_\mathrm{t}\cdot\left(\begin{bmatrix}\nu_{22} & -\nu_{21} & *\\ -\nu_{12} & \nu_{11} & *\\ * & * & *\end{bmatrix}\boldsymbol{\nabla}_\mathrm{t}\,E_z\right) - \gamma\,\boldsymbol{\nabla}_\mathrm{t}\cdot\left(\begin{bmatrix}\nu_{22} & -\nu_{21} & *\\ -\nu_{12} & \nu_{11} & *\\ * & * & *\end{bmatrix}\mathbf{E_t}\right) \tag{27}$$

Let's now take a look at part $[II]$ and $[III]$ of equation (21). They are straightforward as they just involve multiplications and no spatial derivatives. They result in

$$-\underbrace{k_0^2\boldsymbol{\varepsilon_r}\mathbf{E}}_{[II]} + \underbrace{j\omega\mu_0\boldsymbol{\sigma}\mathbf{E}}_{[III]} = -{k_0}^2\left(\begin{bmatrix}\varepsilon_{11} & \varepsilon_{12} & *\\ \varepsilon_{21} & \varepsilon_{22} & *\\ \varepsilon_{31} & \varepsilon_{32} & *\end{bmatrix}\mathbf{E_t} + \begin{bmatrix}* & * & \varepsilon_{13}\\ * & * & \varepsilon_{23}\\ * & * & \varepsilon_{33}\end{bmatrix}\mathbf{E_z}\right) + \mathrm{j}\omega\mu_0(\boldsymbol{\sigma}\mathbf{E_t} + \boldsymbol{\sigma}\mathbf{E_z}) \tag{28}$$

$$= -{k_0}^2\left(\boldsymbol{\varepsilon}\mathbf{E_t} + \boldsymbol{\varepsilon}\mathbf{E_z}\right) + \mathrm{j}\omega\mu_0(\boldsymbol{\sigma}\mathbf{E_t} + \boldsymbol{\sigma}\mathbf{E_z}) \tag{29}$$

The next step is to apply the variational formulation to the strong form, which is formed by equations (26), (27), and (29). We multiply the equations with a general test function $\mathbf{T}$:

$$\mathbf{T} = \mathbf{T_t} + \mathbf{T_z} = \begin{bmatrix}T_x\\ T_y\\ 0\end{bmatrix} + \begin{bmatrix}0\\ 0\\ T_z\end{bmatrix} \tag{30}$$

Then, we integrate over the whole domain Ω. Essentially, it reads:

$$\int_\Omega (\mathbf{T_t} + \mathbf{T_z})\cdot\left([\mathbf{I}]_{(\mathbf{xy})} + [I]_{(z)} + [\mathbf{II}] + [\mathbf{III}]\right)\,d\mathbf{s} = 0 \tag{31}$$

$$\Leftrightarrow \int_\Omega \mathbf{T_t}\cdot[\mathbf{I}]_{(\mathbf{xy})} + T_z\cdot[I]_{(z)} + (\mathbf{T_t} + \mathbf{T_z})\cdot([\mathbf{II}] + [\mathbf{III}])\,d\mathbf{s} = 0. \tag{32}$$

Let's go through the terms one by one and plug in equations (26), (27), and (29). We start with $\mathbf{T_t}\cdot[I]_{(xy)}$:

$$\begin{aligned}\mathbf{T_t}\cdot[I]_{(xy)} = \int_\Omega \mathbf{T_t}\cdot\Bigg(&\boldsymbol{\nabla}_\mathrm{t}\times\left(\begin{bmatrix}* & * & 0\\ * & * & 0\\ * & * & \nu_{33}\end{bmatrix}\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{E_t}\right) - \boldsymbol{\nabla}_\mathrm{t}\times\left(\begin{bmatrix}0 & 0 & *\\ 0 & 0 & *\\ \nu_{32} & -\nu_{31} & *\end{bmatrix}\boldsymbol{\nabla}_\mathrm{t}\,E_z\right)\\ &- \gamma\begin{bmatrix}\nu_{22} & -\nu_{21} & *\\ -\nu_{12} & \nu_{11} & *\\ * & * & *\end{bmatrix}\boldsymbol{\nabla}_\mathrm{t}\,E_z + \gamma\begin{bmatrix}* & * & \nu_{23}\\ * & * & -\nu_{13}\\ * & * & 0\end{bmatrix}\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{E_t} - \gamma\,\boldsymbol{\nabla}_\mathrm{t}\times\begin{bmatrix}0 & 0 & *\\ 0 & 0 & *\\ \nu_{32} & -\nu_{31} & *\end{bmatrix}\mathbf{E_t}\\ &- \gamma^2\begin{bmatrix}\nu_{22} & -\nu_{21} & *\\ -\nu_{12} & \nu_{11} & *\\ * & * & *\end{bmatrix}\mathbf{E_t}\Bigg)\,d\mathbf{s}.\end{aligned} \tag{33}$$

Applying the vector identitiy (integration by parts)

$$\iiint_V \mathbf{A}\cdot(\boldsymbol{\nabla}\times\mathbf{B})dV = -\oiint_{\partial V}(\mathbf{A}\times\mathbf{B})\cdot d\mathbf{S} + \iiint_V(\boldsymbol{\nabla}\times\mathbf{A})\cdot\mathbf{B}dV \tag{34}$$

it results in

$$\begin{aligned}\mathbf{T_t}\cdot[I]_{(xy)} = \int_\Omega &(\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{T_t})\cdot\begin{bmatrix}* & * & *\\ * & * & *\\ * & * & \nu_{33}\end{bmatrix}(\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{E_t}) - (\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{T_t})\cdot\begin{bmatrix}* & * & *\\ * & * & *\\ \nu_{32} & -\nu_{31} & *\end{bmatrix}(\boldsymbol{\nabla}_\mathrm{t}\,E_z)\\ &- \gamma\mathbf{T_t}\cdot\begin{bmatrix}\nu_{22} & -\nu_{21} & *\\ -\nu_{12} & \nu_{11} & *\\ * & * & *\end{bmatrix}\boldsymbol{\nabla}_\mathrm{t}\,E_z + \gamma\mathbf{T_t}\cdot\begin{bmatrix}* & * & \nu_{23}\\ * & * & -\nu_{13}\\ * & * & *\end{bmatrix}(\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{E_t})\\ &- \gamma(\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{T_t})\cdot\begin{bmatrix}* & * & *\\ * & * & *\\ \nu_{32} & -\nu_{31} & *\end{bmatrix}\mathbf{E_t} - \gamma^2\mathbf{T_t}\cdot\begin{bmatrix}\nu_{22} & -\nu_{21} & *\\ -\nu_{12} & \nu_{11} & *\\ * & * & *\end{bmatrix}\mathbf{E_t}\,d\mathbf{s}\end{aligned} \tag{35}$$

Note that here the zeros in the matrix have been replaced with $*$, as the entry will be multiplied by 0 anyways due to the testfunction. The integration term over the boundary vanishes as we assume PEC-type boundary conditions everywhere. Next term is

$$T_z\cdot[I]_{(z)} = \int_\Omega T_z\cdot\boldsymbol{\nabla}_\mathrm{t}\cdot\left(\begin{bmatrix}* & * & \nu_{23}\\ * & * & -\nu_{13}\\ * & * & *\end{bmatrix}\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{E_t} - \begin{bmatrix}\nu_{22} & -\nu_{21} & *\\ -\nu_{12} & \nu_{11} & *\\ * & * & *\end{bmatrix}\boldsymbol{\nabla}_\mathrm{t}\,E_z - \gamma\begin{bmatrix}\nu_{22} & -\nu_{21} & *\\ -\nu_{12} & \nu_{11} & *\\ * & * & *\end{bmatrix}\mathbf{E_t}\right)d\mathbf{s} \tag{36}$$

With the vector identity

$$\iiint_V \psi\boldsymbol{\nabla}\cdot\mathbf{A}dV = \oiint_{\partial V}\psi\mathbf{A}\cdot d\mathbf{S} - \iiint_V\boldsymbol{\nabla}\psi\cdot\mathbf{A}dV \tag{37}$$

and vanishing boundary integral for the same reason as before, it results in

$$
\begin{aligned}
T_z \cdot [I]_{(z)} = \int_\Omega & -\boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t}) + \boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t}\, E_z \\
& + \gamma\, \boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t}\, d\mathbf{s}
\end{aligned} \tag{38}
$$

The terms $[II]$ and $[III]$ lead to

$$
\int_\Omega [II] + [III]\, d\mathbf{s} = \int_\Omega (\mathbf{T_t} + T_z) \cdot ({k_0}^2 \boldsymbol{\varepsilon_r} + \mathrm{j}\omega\mu_0\boldsymbol{\sigma})(\mathbf{E_t} + E_z)\, d\mathbf{s} \tag{39}
$$

So putting all together the variational formulation of the eigenvalue problem reads

$$
\begin{aligned}
& \int_\Omega (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & \nu_{33} \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t}) - (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t}\, E_z) \\
& - \gamma \mathbf{T_t} \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t}\, E_z + \gamma \mathbf{T_t} \cdot \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t}) - \gamma (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} \mathbf{E_t} \\
& - \gamma^2 \mathbf{T_t} \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t} - \boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t}) + \boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t}\, E_z \\
& + \gamma\, \boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t}\, d\mathbf{s} = \int_\Omega (\mathbf{T_t} + T_z) \cdot ({k_0}^2 \boldsymbol{\varepsilon_r} + \mathrm{j}\omega\mu_0\boldsymbol{\sigma})(\mathbf{E_t} + E_z)\, d\mathbf{s}
\end{aligned} \tag{40}
$$

2.2 Approach 2: Wave equation - Variational Formulation - Substitution of z-derivatives

The derivation can also be done starting with the weak form and perform the splitting

$$
\mathbf{E} = \begin{bmatrix} E_x \\ E_y \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ E_z \end{bmatrix} \tag{41}
$$

afterwards.

First, multiply with the testfunction

$$
\mathbf{T} = \begin{bmatrix} T_x \\ T_y \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ T_z \end{bmatrix} \tag{42}
$$

and integrate over the whole domain Ω

$$
\int_\Omega \mathbf{T} \cdot \left(\boldsymbol{\nabla} \times \left(\boldsymbol{\mu_r^{-1}} \boldsymbol{\nabla} \times \mathbf{E} \right) - ({k_0}^2 \boldsymbol{\varepsilon_r} - \mathrm{j}\omega\mu_0\boldsymbol{\sigma})\mathbf{E} \right) d\mathbf{s} = 0 \tag{43}
$$

Again with the vector identity (integration by parts) in equation 34, it can be rewritten to

$$
\int_\Omega \mathbf{T} \cdot \left(\boldsymbol{\nabla} \times \boldsymbol{\mu^{-1}} \boldsymbol{\nabla} \times \mathbf{E} - ({k_0}^2 \boldsymbol{\varepsilon_r} - \mathrm{j}\omega\mu_0\boldsymbol{\sigma})\mathbf{E} \right) d\mathbf{s} = 0 \tag{44}
$$

$$
\int_\Omega \mathbf{T} \cdot \left(\boldsymbol{\nabla} \times \boldsymbol{\mu^{-1}} \boldsymbol{\nabla} \times \mathbf{E} \right) d\mathbf{s} - \int_\Omega \mathbf{T} \cdot ({k_0}^2 \boldsymbol{\varepsilon_r} - \mathrm{j}\omega\mu_0\boldsymbol{\sigma})\mathbf{E}\, d\mathbf{s} = 0 \tag{45}
$$

$$
\int_\Omega (\boldsymbol{\nabla} \times \mathbf{T}) \cdot \left(\boldsymbol{\mu^{-1}} \boldsymbol{\nabla} \times \mathbf{E} \right) - \mathbf{T} \cdot ({k_0}^2 \boldsymbol{\varepsilon_r} - \mathrm{j}\omega\mu_0\boldsymbol{\sigma})\mathbf{E}\, d\mathbf{s} + \int_{\delta\Omega} \mathbf{T} \cdot (\hat{n} \boldsymbol{\nabla} \times \boldsymbol{\mu^{-1}} \boldsymbol{\nabla} \times \mathbf{E})\, d\mathbf{s} = 0 \tag{46}
$$

As we assume boundary conditions of PEC-type everywhere, the boundary integral evaluates to 0 and we are left with

$$
\int_\Omega (\boldsymbol{\nabla} \times \mathbf{T}) \cdot \left(\boldsymbol{\mu_r^{-1}} \boldsymbol{\nabla} \times \mathbf{E} \right) - \mathbf{T} \cdot ({k_0}^2 \boldsymbol{\varepsilon_r} - \mathrm{j}\omega\mu_0\boldsymbol{\sigma})\mathbf{E}\, d\mathbf{s} = 0 \tag{47}
$$

Now, we will plug in

$$
\mathbf{E} = \mathbf{E_t} + \mathbf{E_z} = \begin{bmatrix} E_x \\ E_y \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ E_z \end{bmatrix} \quad \text{and} \quad \mathbf{T} = \mathbf{T_t} + \mathbf{T_z} = \begin{bmatrix} T_x \\ T_y \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ T_z \end{bmatrix} \tag{48}
$$

and try to separate the equations. In cartesian coordinates, expanding the curl operations lead to

$$
\begin{aligned}
& \int_\Omega \left[\begin{pmatrix} \frac{\partial T_y}{\partial z} \\ -\frac{\partial T_x}{\partial z} \\ \frac{\partial T_y}{\partial x} - \frac{\partial T_x}{\partial y} \end{pmatrix} + \begin{pmatrix} \frac{\partial T_z}{\partial y} \\ -\frac{\partial T_z}{\partial x} \\ 0 \end{pmatrix} \right] \cdot \underbrace{\left(\begin{bmatrix} \nu_{11} \frac{\partial E_y}{\partial z} - \nu_{12} \frac{\partial E_x}{\partial z} + \nu_{13} \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \right) \\ \nu_{21} \frac{\partial E_y}{\partial z} - \nu_{22} \frac{\partial E_x}{\partial z} + \nu_{23} \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \right) \\ \nu_{31} \frac{\partial E_y}{\partial z} - \nu_{32} \frac{\partial E_x}{\partial z} + \nu_{33} \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \right) \end{bmatrix} + \begin{bmatrix} \nu_{11} \frac{\partial E_z}{\partial y} - \nu_{12} \frac{\partial E_z}{\partial x} \\ \nu_{21} \frac{\partial E_z}{\partial y} - \nu_{22} \frac{\partial E_z}{\partial x} \\ \nu_{31} \frac{\partial E_z}{\partial y} - \nu_{32} \frac{\partial E_z}{\partial x} \end{bmatrix} \right)}_{\boldsymbol{\nu} \boldsymbol{\nabla} \times (\mathbf{E_t} + \mathbf{E_z})} \\
& = \int_\Omega \mathbf{T} \cdot ({k_0}^2 \boldsymbol{\varepsilon_r} - \mathrm{j}\omega\mu_0\boldsymbol{\sigma})\mathbf{E}\, d\mathbf{s}
\end{aligned} \tag{49}
$$

As in approach 1 (2.1, eq. (24)-(29)) , we will replace the z-derivative with $\frac{\partial}{\partial z} = -\gamma$ and perform the multiplication between the testfunction T (i.e. the curl of the splitted notation) and E. In the end, we will bring the terms together into a notation with vector operators and material tensors again.We will take a seperate look at the terms corresponding to $\mathbf{E_t}$ and E_z. Starting with the $\mathbf{T_t}$-part, we end up with

$$\left[\begin{pmatrix} \frac{\partial T_y}{\partial z} \\ -\frac{\partial T_x}{\partial z} \\ \frac{\partial T_y}{\partial x} - \frac{\partial T_x}{\partial y} \end{pmatrix}\right] \cdot \boldsymbol{\nu} \left(\boldsymbol{\nabla} \times \mathbf{E_t} + \boldsymbol{\nabla} \times \mathbf{E_z}\right) \tag{50}$$

$$\begin{aligned}
&= \nu_{33} \left(\frac{\partial E_x}{\partial y}\frac{\partial T_x}{\partial y} - \frac{\partial E_x}{\partial y}\frac{\partial T_y}{\partial x} - \frac{\partial E_y}{\partial x}\frac{\partial T_x}{\partial y} + \frac{\partial E_y}{\partial x}\frac{\partial T_y}{\partial x} \right) \\
&\quad + \nu_{31} \left(\frac{\partial E_z}{\partial y}\frac{\partial T_y}{\partial x} - \frac{\partial E_z}{\partial y}\frac{\partial T_x}{\partial y} \right) + \nu_{32} \left(\frac{\partial E_z}{\partial x}\frac{\partial T_x}{\partial y} - \frac{\partial E_z}{\partial x}\frac{\partial T_y}{\partial x} \right) \\
&\quad + \gamma \left(-\nu_{11} E_y \frac{\partial T_z}{\partial y} + \nu_{12} E_x \frac{\partial T_z}{\partial y} + \nu_{21} E_y \frac{\partial T_z}{\partial x} - \nu_{22} E_x \frac{\partial T_z}{\partial x} \right) \\
&\quad + \gamma \left(\nu_{13} \left(T_y \frac{\partial E_y}{\partial x} - T_y \frac{\partial E_x}{\partial y} \right) + \nu_{23} \left(T_x \frac{\partial E_y}{\partial x} - T_x \frac{\partial E_x}{\partial y} \right) \right) \\
&\quad + \gamma \left(\nu_{31} \left(E_y \frac{\partial T_y}{\partial x} - E_y \frac{\partial T_x}{\partial y} \right) + \nu_{32} \left(E_x \frac{\partial T_y}{\partial x} - E_x \frac{\partial T_x}{\partial y} \right) \right) \\
&\quad + \gamma^2 \left(\nu_{11} E_y T_y - \nu_{12} E_x T_y - \nu_{21} E_y T_x + \nu_{22} E_x T_x \right)
\end{aligned} \tag{51}$$

$$\begin{aligned}
&= (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & \nu_{33} \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t}) - (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} E_z) \\
&\quad - \gamma \mathbf{T_t} \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t} E_z + \gamma \mathbf{T_t} \cdot \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t}) \\
&\quad - \gamma (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} \mathbf{E_t} - \gamma^2 \mathbf{T_t} \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t}
\end{aligned} \tag{52}$$

The T_z-part can be summarized as

$$\left[\begin{pmatrix} \frac{\partial T_z}{\partial y} \\ -\frac{\partial T_z}{\partial x} \\ 0 \end{pmatrix}\right] \cdot \boldsymbol{\nu} \left(\boldsymbol{\nabla} \times \mathbf{E_t} + \boldsymbol{\nabla} \times \mathbf{E_z}\right) \tag{53}$$

$$\begin{aligned}
&= \left(\nu_{13} \frac{\partial E_x}{\partial x}\frac{\partial T_z}{\partial y} - \nu_{13} \frac{\partial E_x}{\partial y}\frac{\partial T_z}{\partial y} - \nu_{23} \frac{\partial E_y}{\partial x}\frac{\partial T_z}{\partial x} + \nu_{23} \frac{\partial E_x}{\partial y}\frac{\partial T_z}{\partial x} \right) \\
&\quad + \left(\nu_{11} \frac{\partial E_z}{\partial y}\frac{\partial T_z}{\partial y} - \nu_{12} \frac{\partial E_z}{\partial x}\frac{\partial T_z}{\partial y} - \nu_{21} \frac{\partial E_z}{\partial y}\frac{\partial T_z}{\partial x} + \nu_{22} \frac{\partial E_z}{\partial x}\frac{\partial T_z}{\partial x} \right) \\
&\quad + \gamma \left(\nu_{11} E_y \frac{\partial T_z}{\partial y} - \nu_{12} E_x \frac{\partial T_z}{\partial y} - \nu_{21} E_y \frac{\partial T_z}{\partial x} + E_x \nu_{22} \frac{\partial T_z}{\partial x} \right)
\end{aligned} \tag{54}$$

$$= \boldsymbol{\nabla}_\mathrm{t} T_z \cdot \begin{bmatrix} * & * & -\nu_{23} \\ * & * & \nu_{13} \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t} + \boldsymbol{\nabla}_\mathrm{t} T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t} E_z + \gamma \boldsymbol{\nabla}_\mathrm{t} T_z \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t} \tag{55}$$

Adding eq. (55) and (52) up, we end up with the final weak formulation in the splitted notation ($\mathbf{E_t}$ and E_z) and with the z-derivatives replaced with their constant equivalent value γ, namely

$$\begin{aligned}
&\int_\Omega (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & \nu_{33} \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t}) - (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} E_z) \\
&- \gamma \mathbf{T_t} \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t} E_z + \gamma \mathbf{T_t} \cdot \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t}) - \gamma (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} \mathbf{E_t} \\
&- \gamma^2 \mathbf{T_t} \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t} - \boldsymbol{\nabla}_\mathrm{t} T_z \cdot \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t} \times \mathbf{E_t}) + \boldsymbol{\nabla}_\mathrm{t} T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t} E_z \\
&+ \gamma \boldsymbol{\nabla}_\mathrm{t} T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t} \, d\mathbf{s} = \int_\Omega (\mathbf{T_t} + \mathbf{T_z}) \cdot (k_0{}^2 \boldsymbol{\varepsilon_r} + \mathrm{j}\omega\mu_0\boldsymbol{\sigma})(\mathbf{E_t} + \mathbf{E_z}) \, d\mathbf{s}
\end{aligned} \tag{56}$$

This is exactly equivalent to equation (40) obtained from approach 1 in subsection 2.1. That is expected and validates the correctness of the above calculations.

For further validation, we will assume an isotropic medium with a scalar $\nu = 1/\mu$, and $\alpha = 0$, so $\gamma = \beta$ and compare it to the known formula for that case in [2]. Reducing it to isotropic media, the formulation simplifies to

$$\int_\Omega \frac{1}{\mu}\left((\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{T_t})\cdot(\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{E_t}) - \mathrm{j}\beta\mathbf{T_t}\cdot\boldsymbol{\nabla}E_z + \beta^2\mathbf{T_t}\cdot\mathbf{E_t} + \mathrm{j}\beta\,\boldsymbol{\nabla}_\mathrm{t}\,T_z\cdot\mathbf{E_t} + \boldsymbol{\nabla}_\mathrm{t}\,T_z\cdot\boldsymbol{\nabla}_\mathrm{t}\,E_z\right)\,d\mathbf{s} \\ = \int_\Omega {k_0}^2\varepsilon_r(\mathbf{E_t}\cdot\mathbf{T_t} + E_zT_z)\,d\mathbf{s} \tag{57}$$

If we perform the same field scaling as in the literature ([2]), i.e. $\mathbf{E_t} = \frac{\mathbf{E_t}}{\beta}$ and $E_z = -\frac{E_z}{j}$ (also $\mathbf{T_t} = \frac{\mathbf{T_t}}{\beta}$ and $T_z = -\frac{T_z}{j}$), we get

$$\int_\Omega \frac{1}{\mu}\left(\frac{1}{\beta^2}(\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{T_t})\cdot(\boldsymbol{\nabla}_\mathrm{t}\times\mathbf{E_t}) + \mathbf{T_t}\cdot\boldsymbol{\nabla}_\mathrm{t}\,E_z + \mathbf{T_t}\cdot\mathbf{E_t} - \boldsymbol{\nabla}_\mathrm{t}\,T_z\cdot\mathbf{E_t} - \boldsymbol{\nabla}_\mathrm{t}\,T_z\cdot\boldsymbol{\nabla}_\mathrm{t}\,E_z\right)\,d\mathbf{s} \\ = \int_\Omega {k_0}^2\varepsilon_r(\frac{1}{\beta^2}\mathbf{E_t}\cdot\mathbf{T_t} - E_zT_z)\,d\mathbf{s} \tag{58}$$

which is equivalent to the linear combination $\left[\frac{1}{\beta^2}\text{eq.(108)} - \beta^2\cdot\text{eq.(109)}\right]$ from [2].

2.3 Approach 3: Substitution of z- derivatives - Derivation in terms of E - Variational Formuatlion

Another approach is to start directly from Maxwell's equations and substitute the z-derivatives immediately before deriving the wave equation. Equation (5) can be expanded to

$$\boldsymbol{\nu}\boldsymbol{\nabla}\times\mathbf{E} = \boldsymbol{\nu}\begin{pmatrix}\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\\ \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\\ \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\end{pmatrix} = \boldsymbol{\nu}\begin{pmatrix}\frac{\partial E_z}{\partial y} + j\beta E_y\\ -j\beta E_x - \frac{\partial E_z}{\partial x}\\ \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\end{pmatrix} = \boldsymbol{\nu}\begin{pmatrix}\hat{E}_1\\ \hat{E}_2\\ \hat{E}_3\end{pmatrix} = -j\omega\begin{pmatrix}H_x\\ H_y\\ H_z\end{pmatrix} \tag{59}$$

and written linewise,

$$\nu_{11}\hat{E}_1 + \nu_{12}\hat{E}_2 + \nu_{13}\hat{E}_3 = -\mathrm{j}\omega H_x \tag{60a}$$
$$\nu_{21}\hat{E}_1 + \nu_{22}\hat{E}_2 + \nu_{23}\hat{E}_3 = -\mathrm{j}\omega H_y \tag{60b}$$
$$\nu_{31}\hat{E}_1 + \nu_{32}\hat{E}_2 + \nu_{33}\hat{E}_3 = -\mathrm{j}\omega H_z \tag{60c}$$

Similarly, equation (6) expands to

$$\boldsymbol{\nabla}\times\mathbf{H} = \begin{pmatrix}\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\\ \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}\\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\end{pmatrix} = \begin{pmatrix}\frac{\partial H_z}{\partial y} + \gamma H_y\\ -\gamma H_x - \frac{\partial H_z}{\partial x}\\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\end{pmatrix} = (\mathrm{j}\omega\boldsymbol{\varepsilon} + \boldsymbol{\sigma})\mathbf{E} \tag{61}$$

and linewise

$$\frac{\partial H_z}{\partial y} + \gamma H_y = [\mathrm{j}\omega\boldsymbol{\varepsilon} + \boldsymbol{\sigma}\mathbf{E}]_x \tag{62a}$$
$$-\gamma H_x - \frac{\partial H_z}{\partial x} = [\mathrm{j}\omega\boldsymbol{\varepsilon} + \boldsymbol{\sigma}\mathbf{E}]_y \tag{62b}$$
$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = [\mathrm{j}\omega\boldsymbol{\varepsilon} + \boldsymbol{\sigma}\mathbf{E}]_z \tag{62c}$$

By a smart choice of linear combinations of the different lines and their derivatives of (60a) to (60b), i.e.

$$\partial_y[(60\mathrm{c})] + \gamma(60\mathrm{b}) = -\mathrm{j}\omega(62\mathrm{a}) \tag{63}$$
$$-\gamma(60\mathrm{c}) - \partial_z[(60\mathrm{a})] = -\mathrm{j}\omega(62\mathrm{b}) \tag{64}$$
$$-\partial_y[(60\mathrm{b})] - \partial_x[(60\mathrm{a})] - \mathrm{j}\omega(62\mathrm{c}) \tag{65}$$

we can replace the unknown H of the right-hand-side by the right-hand-side of (61).

We end up with the same expansion as with approach 1 (see eq.(2.1)), and therefore the next steps will be the same and we again end up with the same weak formulation.

3 Introduction to SparseLizard

For a numerical solution of this weak formulation, it will be discretized in the context of the Finite Element Method (FEM). This will be done using the open-source C++ finite element library `SparseLizard` (`https://www.sparselizard.org/`). This package was mainly developed by A.Halbach. It provides the necessary elements to provide the weak formulation, construct the Galerkin problem and solve it. In the following, the necessary objects will be briefly presented.

In the beginning of each simulation, a `mesh` object will be created that holds the finite element mesh of the geometry. Physical regions can be defined to assign material properties to specific domains. The material definition can be done using the `parameter` object, it can hold different expression objects on different geometric regions. The expressions can be scalar, or matrices.

The unknows are introduced by the `field` object. The type of shape functions ("h1" for nodal and "hcurl" for Nedelec's edge shape functions) and the interpolation order can be set. The method `setconstraint` allows to set Dirichlet boundary conditions.

To input the weak formulation, a `formulation` object is created. For example,

```
formulation mode;
mode += integral(vol, dof(v*tf (v) );
mode += integral(vol, curl(dof(v))*grad(tf (v) ));
```

creates a `formulation` object, adds two terms that are assembled for unknowns (dof) and test functions (tf) defined on region 'vol' and for an integration on all elements in region 'vol' as well. All terms are added together and their sum equals zero. The use of time derivatives on the dofs assigns the term to the correct matrix, i.e. terms containing `dtdt(dof(v))` are assigned to the mass matrix $\mathbf{M}$, `dt(dof(v))` to the damping matrix $\mathbf{C}$ and all terms without time derivative to the stiffness matrix $\mathbf{K}$. The matrices can be assembled and accessed by

```
mode.generate();
mat M = mode.M();
mat C = mode.C();
mat K = mode.K();
```

where the `mat` object holds a sparse algebraic matrix. The `eigenvalue` object in `SparseLizard` allows to solve classical, generalized and polynomial eigenvalue problems. It can be either created with two or three matrix arguments for a linear or quadratic EVP, respectively. The computation is done by `SLEPc`, a scalable library for eigenvalue problem computation, which will be introduced in the next chapter. It is important to note that only the scalar version of all solver librarys (`SLEPc`, `PETSc` and `MUMPS`) is included. Therefore lossy eigenmode problems need to be transformed to consist of purely real system matrices. It will later be discussed in detail. This and further information on `SparseLizard` can be found in the documentation [3].

4 Introduction to SLEPc Library

For solving the eigenvalue problem efficiently, the Scalable Library for Eigenvalue Problem Computations (SLEPc) (`https://slepc.upv.es/`) developed by researchers from Universitat Politècnica de València (Spain) is used. SLEPc is an open-source library designed to address large-scale sparse eigenvalue problems on parallel computing architectures. It extends the functionality of PETSc and is applicable to linear eigenvalue problems, both in standard and generalized formulations, supporting real or complex arithmetic. Additionally, SLEPc can also tackle nonlinear eigenvalue problems, whether polynomial or general in nature. It is based on the PETSc data structures and it employs the MPI standard for message-passing communication. This chapter introduces the main capabilities used in this project.

4.1 Main Solver Classes

4.1.1 Eigenvalue Problem Solver (EPS)

The EPS class supports both standard eigenvalue problems ($\mathbf{Ax} = \lambda\mathbf{x}$) and generalized eigenvalue problems ($\mathbf{Ax} = \lambda\mathbf{Bx}$). The problemtype can be provided in order to use a specific solving strategy, it can be either hermitian, general hermitian, non-hermitian or general non-hermitian. In chapter 5, it is specified which type applies to which material configuration. The type of eigenvalue solver is set to the Krylov-Schur method, that works for symmetric and non-symmetric problems. It typically outperforms Arnoldi in terms of speed and effectiveness. Additionally, both methods exhibit similar levels of robustness [4]. Nevertheless, for hermitian problems it might be faster to use Lanczos methods, which is specifically designed for them, because SLEPc does not exploit the symmetry in its Krylov-Schur implementation. Other options than can be given to the solver are the number of eigenvalues to compute (`EPSSetDimensions`), the tolerance and maximum number of iterations (`EPSSetTolerances`). In the example cases that will be shown in section 6 a tolerance of 1e−6 and maximum 100 iterations were enough to match the results given by papers or other simulation software.

SLEPC provides the option to apply spectral transformations to the problem which involves transforming the original eigenvalue problem to a new one, where eigenvalues are relocated while eigenvectors remain unchanged. This transformation serves various purposes, such as computing internal eigenvalues within a specified range or proximity to a given target value. Additionally, these transformations can enhance convergence by optimizing the distribution of eigenvalues, particularly in cases where difficult distributions hinder convergence properties [5].

In the realm of eigenvalue problem-solving, particularly with algorithms such as Lanczos and Arnoldi, the challenge often lies in slow convergence, especially when faced with closely spaced eigenvalues.To address this challenge, the shift-and-invert spectral transformation emerges as a powerful strategy. By applying this transformation to the original operator $\mathbf{A}$, specifically in the form of $(\mathbf{A} - \sigma\mathbf{I})^{-1}$, where σ is a chosen shift, the eigenvalues are remapped in a way that not only allows computation of eigenvalues beyond the spectrum boundaries but also significantly accelerates convergence for eigenvalues closest to the chosen shift [5]. The given reasons makes it a strategic choice for eigenmode problems of waveguides, as they result in clustered eigenvalues close to a target.

As we usually want to perform a frequency shift, it makes sense that we provide a target that holds for all frequencies. So we provide the effective refractive index $\eta_{\text{eff}} = \beta/k_0$ and then derive our actual eigenvalue target β from it.

4.1.2 Polynomial Eigenvalue Problem (PEP)

The PEP class supports general polynomial EVPs. It can be applied to general, hermitian and hermitian hyperbolic problems. Among others, it provides a solver `PEPLINEAR` which applies a linerization to the problem and then solves it using the before presented EPS solver class. If the problem is given as hermitian, it performs a linearization that maintains this beneficial property, i.e. for the quadratic eigenvalue problem

$$(\lambda^2\mathbf{M} + \lambda\mathbf{C} + \mathbf{K})\mathbf{x} = 0 \tag{66}$$

the resulting matrix pencil

$$\begin{bmatrix} \beta K & \alpha K \\ \alpha K & \alpha C - \beta M \end{bmatrix} - \lambda \begin{bmatrix} \alpha K - \beta C & -\beta M \\ -\beta M & -\alpha M \end{bmatrix} \tag{67}$$

is symmetric, although indefinite [5].

Despite linearization results in a doubling of the unknowns, it enables the use of linear solvers which are known to be faster and more robust.

All other options, such as dimension, target and tolerances, can be set in the same way as in the EPS class.

5 Numerical Implementation

A general script to compute the eigenmodes of a waveguide on its two-dimensional cross-section will have the following structure.

1. Mesh import and region assignment
2. Field definition (type and order)
3. Set perfect conductor boundary condition on the whole boundary
4. Material properties definition
5. Weak Formulation Construction
6. Galerkin matrices assembly
7. Eigenvalue solver initialization
8. Compute solution of EVP
9. Postprocessing

There are two fields to define, first the transverse E-field **Et** which is discretized using Nedelec's edge elements (in the solver it is a 3x1 array with the z-component equal to zero) and then the longitudinal E-field E_l, which is a scalar unknown discretitized by nodal shape functions. For proper calculation, the unknows need to be put into a 3x1 array in the form

$$\begin{bmatrix} 0 \\ 0 \\ \mathtt{dof}(E_l) \end{bmatrix}. \tag{68}$$

In Appendix A there is an example script for the fully anisotropic case.

The eigenvalue solver is compiled in the scalar version, so only accepts real-valued matrices as input. It is nevertheless capable to return complex eigenvalues and -vectors.

There are several options to convert an EVP consisting of complex matrices into a one consisting of purely real matrices.

- Scaling the fields in a way that the resulting formulation is purely real
- Redefining the eigenvalue
- Expanding the $n \times n$-size problem into an $2n \times 2n$-size (in the case of n unknowns)

Not all methods are always applicable and in the following I will go through the possible cases to illustrate the different methods.

5.1 Lossy case ($\alpha \neq 0$)

In the lossy case, we apply the scaling $\mathbf{E_t} = \frac{\mathbf{E_t}}{\gamma}$ and $\frac{\mathbf{T_t}}{\gamma}$. Then equation (56) results in

$$\begin{aligned}
&\int_\Omega \frac{1}{\gamma^2}(\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & \nu_{33} \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{E_t}) - \frac{1}{\gamma}(\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t}\, E_z) \\
&- \mathbf{T_t} \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t}\, E_z - \frac{1}{\gamma}\mathbf{T_t} \cdot \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{E_t}) + \frac{1}{\gamma}(\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} \mathbf{E_t} \\
&- \mathbf{T_t} \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t} - \frac{1}{\gamma}\boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{E_t}) + \boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t}\, E_z \\
&+ \boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t}\, d\mathbf{s} = \int_\Omega (\frac{1}{\gamma}\mathbf{T_t} + \mathbf{T_z}) \cdot (k_0{}^2 \boldsymbol{\varepsilon_r} + \mathrm{j}\omega\mu_0\boldsymbol{\sigma})(\frac{1}{\gamma}\mathbf{E_t} + \mathbf{E_z})\, d\mathbf{s}
\end{aligned} \tag{69}$$

and multiplying by γ^2 leads to

$$\begin{aligned}
&\int_\Omega (\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & \nu_{33} \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{E_t}) - \gamma(\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t}\, E_z) \\
&- \gamma^2\mathbf{T_t} \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t}\, E_z - \gamma\mathbf{T_t} \cdot \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{E_t}) \\
&+ \gamma(\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{T_t}) \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ \nu_{32} & -\nu_{31} & * \end{bmatrix} \mathbf{E_t} - \gamma^2\mathbf{T_t} \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t} \\
&- \gamma\boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} * & * & \nu_{23} \\ * & * & -\nu_{13} \\ * & * & * \end{bmatrix} (\boldsymbol{\nabla}_\mathrm{t}\times \mathbf{E_t}) + \gamma^2\boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \boldsymbol{\nabla}_\mathrm{t}\, E_z \\
&+ \gamma^2\, \boldsymbol{\nabla}_\mathrm{t}\, T_z \cdot \begin{bmatrix} \nu_{22} & -\nu_{21} & * \\ -\nu_{12} & \nu_{11} & * \\ * & * & * \end{bmatrix} \mathbf{E_t}\, d\mathbf{s} = \int_\Omega (\mathbf{T_t} + \gamma\mathbf{T_z}) \cdot (k_0{}^2 \boldsymbol{\varepsilon_r} + \mathrm{j}\omega\mu_0\boldsymbol{\sigma})(\mathbf{E_t} + \gamma\mathbf{E_z})\, d\mathbf{s}
\end{aligned} \tag{70}$$

This can be written into the general problem

$$\left(\gamma^2\mathbf{M} + \gamma\mathbf{C} + \mathbf{K}\right)\mathbf{x} = 0. \tag{71}$$

It is important to note, that in the case of losses due to a complex permittivity ε_r or the presence of conductive materials with σ, we always end up having complex matrices even in the isotropic case. The original problem of matrix size $n \times n$ (in the case of n degrees of freedom) has to be expanded to size $2n \times n$ to convert it into a purely real problem. It is illustrated in the following in the case of real μ and the non-diagonal material entries related to propagation direction z being purely real (or zero) for ε_r and purely imaginary (or zero) for σ. In that case the damping matrix $\mathbf{C}$ is purely real. However, it can be easily extended later. Then we can rewrite (71) into

$$\left(\gamma^2 \begin{bmatrix} \mathrm{Re}\{\mathbf{M}\} & -\mathrm{Im}\{\mathbf{M}\} \\ \mathrm{Im}\{\mathbf{M}\} & \mathrm{Re}\{\mathbf{M}\} \end{bmatrix} + \gamma \begin{bmatrix} \mathbf{C} & 0 \\ 0 & \mathbf{C} \end{bmatrix} + \begin{bmatrix} \mathrm{Re}\{\mathbf{K}\} & -\mathrm{Im}\{\mathbf{K}\} \\ \mathrm{Im}\{\mathbf{K}\} & \mathrm{Re}\{\mathbf{K}\} \end{bmatrix}\right) \begin{bmatrix} \mathrm{Re}\{\mathbf{x}\} \\ \mathrm{Im}\{\mathbf{x}\} \end{bmatrix} = \mathbf{0}, \tag{72}$$

which consists of purely real matrices of size $2n \times 2n$. Note that in this case the symmetry of the stiffness matrix $\mathbf{K}$ is lost.

5.2 Lossless case (α=0)

In the lossless case, so $\sigma = 0$ and ε is purely real, there is no attenuation ($\alpha = 0$). That means that all matrices $\mathbf{M}, \mathbf{C}$ and $\mathbf{K}$ are purely real. In the following we insert $\gamma = j\beta$ into equation 70 and end up with the general problem

$$(-\beta^2\mathbf{M} + \mathrm{j}\beta\mathbf{C} + \mathbf{K})\mathbf{x} = 0 \tag{73}$$

Note, that the matrix $\mathbf{C}$ reduces to a zero matrix, if the material property entries $_{13}$ and $_{23}$ of μ and ε are equal to zero, which is the case for isotropic, orthotropic or only in xy-plane anisotropic materials. Then the problem reduces to

$$\left(-\beta^2\mathbf{M} + \mathbf{K}\right)x = 0 \tag{74}$$

and with the substitution $\lambda = -\beta^2$, we end up with a general linear eigenvalue problem (EVP) with symmetric matrices (in the case of symmetric material matrices)

$$\left(\lambda\mathbf{M} + \mathbf{K}\right)\mathbf{x} = 0. \tag{75}$$

In the case of fully anisotropic material properties (so at least one of the entries $_{13}$, $_{23}$, $_{31}$ or $_{32}$ is non-zero), we need to consider the full quadratic EVP. There is no possiblity to tell slepc that the matrix $\mathbf{C}$ has to be multiplied with the imaginary number using SparseLizard. Therefore we apply the substitution $\lambda^* = j\beta$ and end up with

$$\left(\lambda^{*2}(\mathbf{M}) + \lambda^*\mathbf{C} + \mathbf{K}\right)\mathbf{x} = 0. \tag{76}$$

Then the real parts of the returned eigenvalues correspond to the phase constants β and we can easily target the desired modes.

So in both cases for symmetric material tensors, we can specify the problem type as hermitian in `SLEPc`. In the full anisotropic case PEP solver is used, in all other lossless cases the EPS solver can be used.

As a summary, table 1 summarizes when to use which problem size and solver in the case of symmetric material properties.

The returned eigenvector holds the scaled electric field distribution in the waveguide. The magnetic field can be computed with

$$\mathbf{H} = \mathrm{j}\frac{1}{\omega}\boldsymbol{\mu}^{-1}\boldsymbol{\nabla} \times \mathbf{E} \tag{77}$$

, after the electric field has been scaled back. With the splitting that results in

$$\mathbf{H_t} = \frac{1}{\omega}\boldsymbol{\mu}^{-1}\left(\mathrm{j}\beta\mathbf{E_t} \times \vec{n} + \mathrm{j}\,\boldsymbol{\nabla}_\mathrm{t}\, E_z\right) \tag{78}$$

$$H_z = \mathrm{j}\frac{1}{\omega}\boldsymbol{\mu}^{-1}\,\boldsymbol{\nabla}_\mathrm{t} \times E_z. \tag{79}$$

$*_{13} = *_{23} = *_{31} = *_{32} = 0$	$\sigma = 0$, real ε and μ	Problem Size	Solver
x	x	$n \times n$	EPS (Hermitian)
x		$2n \times 2n$	EPS
	x	$n \times n$	PEP (Hermitian)
		$2n \times 2n$	PEP

Table 1: Summary about solver type for each setting

6 Validation examples

6.1 Isotropic Material without Losses

For isotropic materials without losses, the attenuation constant α is zero and with $\lambda = -\beta^2$ we end up with a general linear EVP, so EPS solver can be used. To obtain the phase constant, we take the square root of

the returned eigenvalues, so negative eigenvalues result in a purely real phase constant (evanescent modes) and positive ones in undamped propagating modes (purely imaginary phase constant). The formulation has been validated on a frequently used test case [6–9]. It is a simple rectangular inhomogenous waveguide of size $2a \times a$ bounded by perfectly conducting walls. It is partially filled with a dielectric material whose relative permittivity and permeability are, $\varepsilon_r = 2.25$ and $\mu_r = 1.0$ respectively. The other half is filled with vacuum. The setup is shown in Figure 1 and the results in Figure 2, where the colored lines corresponds to the first four modes given by [9] and the black dotted values are obtained using the presented formulation in SparseLizard. We can see perfect agreement.

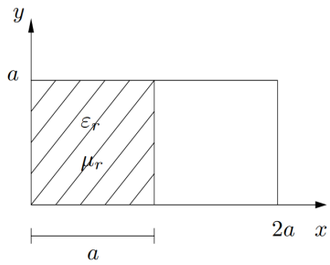

Figure 1: Geometry of the dielectric waveguide in the isotropic, non-lossy case

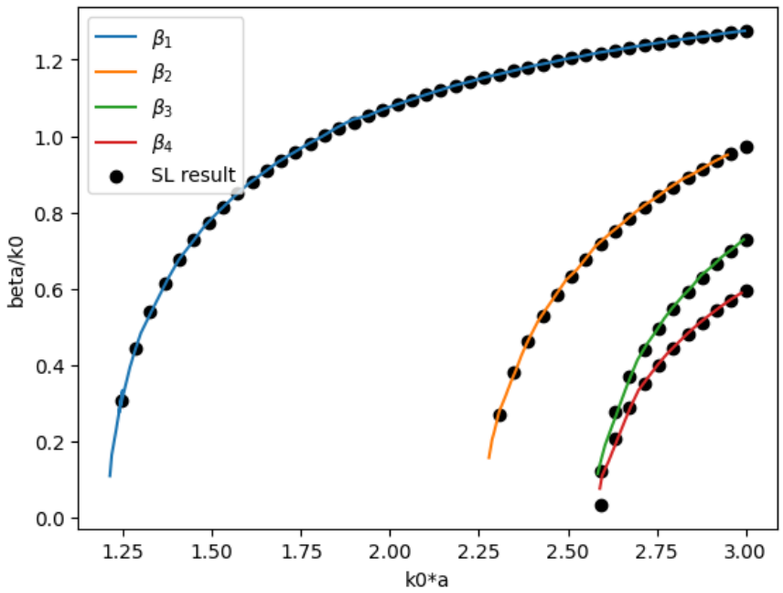

Figure 2: Dispersion characteristics for a $2a \times a$ inhomogeneous rectangular waveguide

6.2 Isotropic Materials with Losses

The expanded $2n \times 2n$ problem, which is needed if we have losses and still want to use a purely real setup, is validated for dielectric and also for conductive losses.

6.2.1 Dielectric Losses

For a rectangular metallic waveguide filled with homogeneous, isotropic, and lossy dielectric, the analytical solution for the propagation constant γ of a $TE_{m,n}$-mode is know as

$$\gamma = \alpha + j\beta = k_0 \sqrt{\left(\frac{m\pi}{k_0 a}\right)^2 + \left(\frac{n\pi}{k_0 b}\right)^2 - \epsilon' + j\epsilon''}, \tag{80}$$

where a and b are the geometrical dimensions and ε' and ε'' correspond to the real and imaginary part of the relative permittiviy respectively [10]. In this example, $a = 1\,\text{mm}$, $b = 0.5\,\text{mm}$ and the dielectric has a relative permittivity $\varepsilon_r = 1.5 - j1.5$. The results for the propagation constants are shown in Figure 3 and for the attenuation constants in Figure 4.

6.2.2 Conductive Losses

In principle, it is the same procedure as for dielectric losses, if we define $\varepsilon'' = \frac{1}{\omega\varepsilon_0}\sigma$. The geometry is shown in Figure 5 with $a = 3\,\text{mm}$, $b = 2\,\text{mm}$ and the electric conductivities $\sigma_1 = 58 \times 10^6\,\text{S}$ in the (thick) wall and $\sigma_2 = 10\,\text{S}$ in the core. The relative permeability and permittivity are 1.0 in all media.

The corrrectness of the formulation is shown by validating it against the well-established simulation tool `Ansys HFSS` (`https://www.ansys.com/products/electronics/ansys-hfss`).

The validation results are shown in Figure 6 and Figure 7.

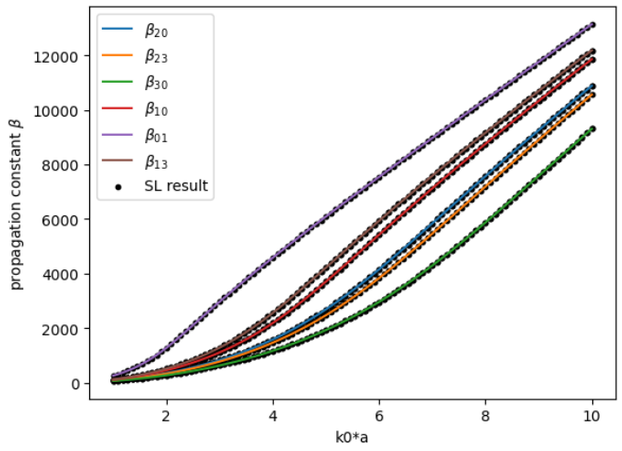

Figure 3: Propagation constants of rectangular waveguide with dielectric losses validated against analytical formula

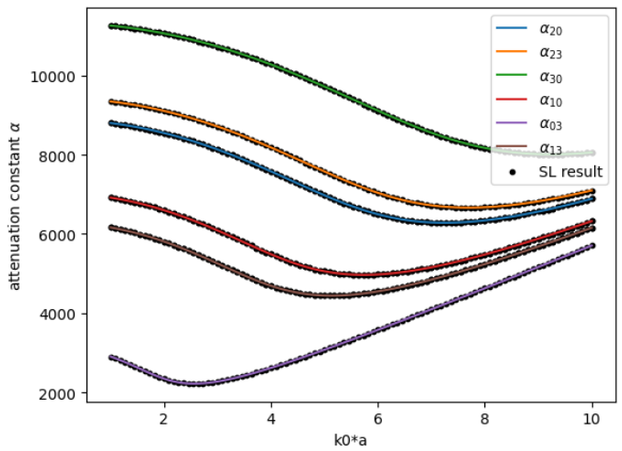

Figure 4: Attenuation constants of rectangular waveguide with dielectric losses validated against analytical formula

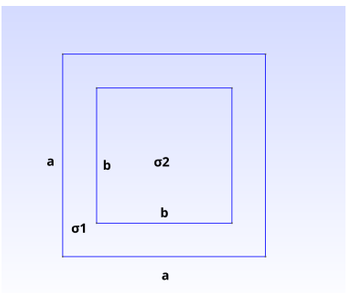

Figure 5: Geometry of the rectangular waveguide with conductive losses

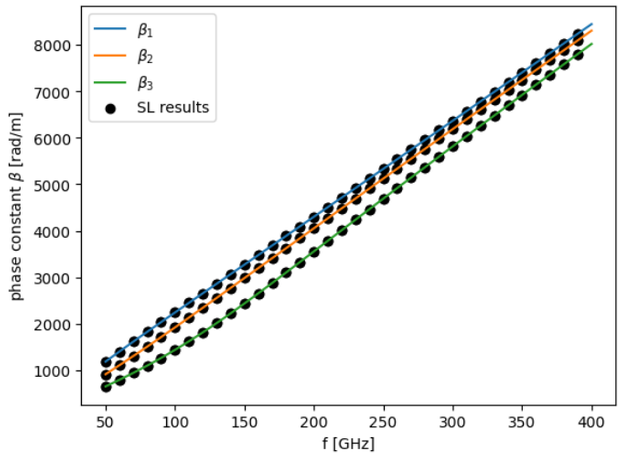

Figure 6: Propagation constants of a inhomogeneous rectangular waveguide with conductive losses

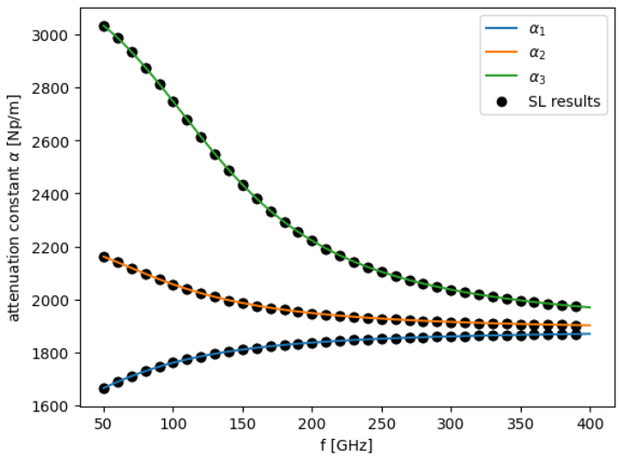

Figure 7: Propagation constants of a inhomogeneous rectangular waveguide with conductive losses

6.3 Orthotropic Materials

As mentioned above, the scaled formulation will also result in purely real valued matrices for orthotropic materials.

6.3.1 Orthotropic dielectric

To validate the case of orthotropic dielectric materials, an anisotropic, lossless dielectric buried waveguide of rectangular cross section of height t, width $w = 2t$, core permittivity $\varepsilon_{xx} = \varepsilon_{yy} = 2.31$ and cladding permittivity $\varepsilon_2 = 2.05$ The first four modes are shown in Figure 8 and agree very well with those obtained by Ohtaka [10, 11].

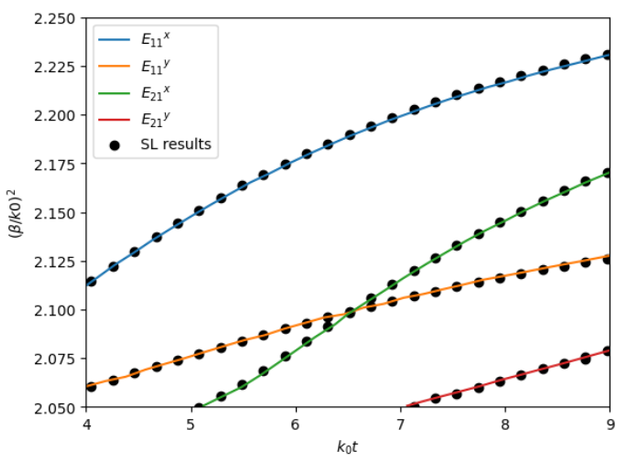

Figure 8: Dispersion characteristics of the four lowest modes in an orthotropic rectangular dielectric waveguide

6.3.2 Lossy Orthotropic dielectric

As a next step in validation, the implementation was applied to a structure consisting of a lossy orthotropic medium. The dispersion characteristics in the slow wave region for the ${E^1}_{yy}$-mode of a lossy anisotropic image waveguide shown in Figure 9 are compared. The imaginary part of the relative permittivity ε''_{yy} is chosen as a parameter. The Figures 10 and 11 show perfect agreement with the results from the doctoral thesis of Lu [12].

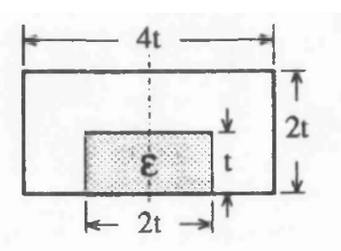

Figure 9: Geometry of Orthotropic dielectric lossy inhomogenous rectangular waveguide

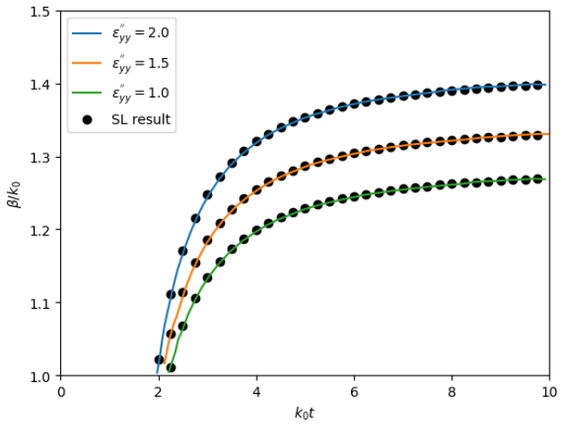

Figure 10: Orthotropic dielectric lossy inhomogenous rectangular waveguide (validated against Paper)

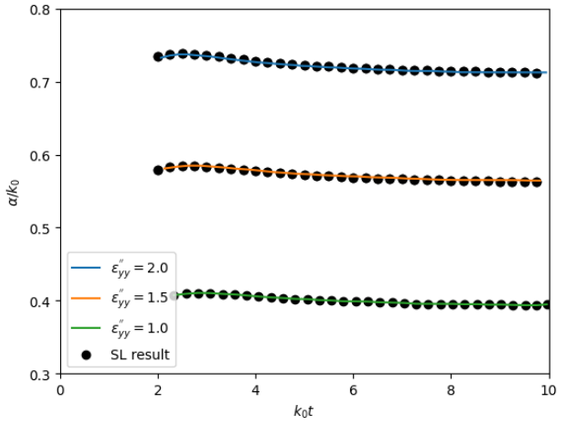

Figure 11: Orthotropic dielectric lossy inhomogenous rectangular waveguide (validated against Paper)

6.3.3 Orthotropic permittivity and permeability

This example shows that the implemented formulation works also in the case that both, the relative permittivity and permeability are orthotropic. The propagation constant of a rectangular waveguide of size $15\,\text{mm} \times 10\,\text{mm}$ filled with a homogenouse medium with

$$\boldsymbol{\varepsilon_r} = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\mu_r} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{81}$$

was computed with SparseLizard and with `Ansys HFSS`. The phase constant β obtained from both tools are plotted in 12.

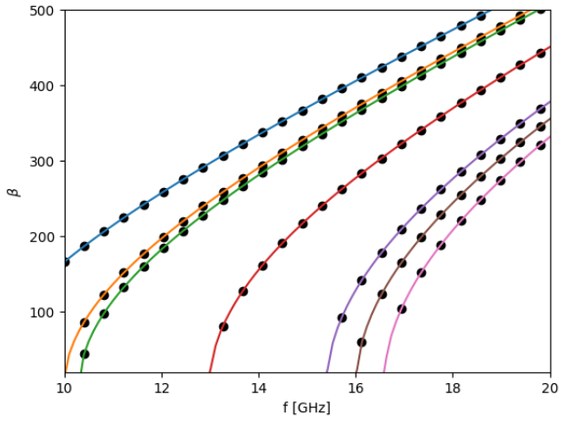

Figure 12: Rectangular waveguide with orthotropic permittivity and permeablity

6.4 Anisotropic Materials in x-y-plane

As already stated, for material properties of the form

$$\begin{bmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{82}$$

the scaling also results in a real valued formulation and matrix $\mathbf{C}$ having only zero entries. This can again be reduced to a linear EVP of size $n \times n$.

6.4.1 Anisotropic permittivity

We consider a square anisotropic waveguide with a uniaxial material surrounded by an isotropic material of relative permittivity 2.05, where the ordinary and extraordinary refractive indices of the core are $\sqrt{2.31}$ and $\sqrt{2.19}$, respectively. Figure 13 shows the comparison in the case that the optical lies at an angle $\theta = 45°$ to the xy-plane. That results in a permittivity tensor of

$$\boldsymbol{\varepsilon_r} = \begin{bmatrix} 2.25 & -0.06 & 0 \\ -0.06 & 2.25 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{83}$$

in the core.

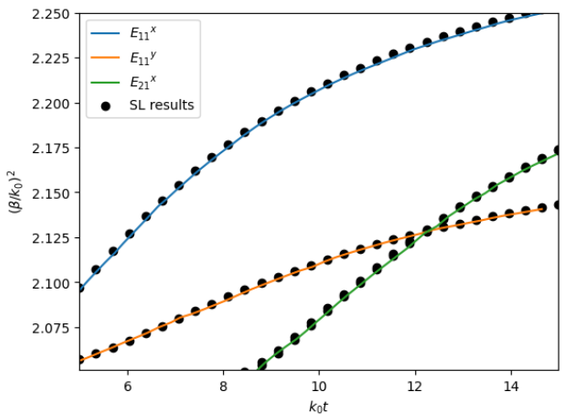

Figure 13: Dispersion characteristics of an anisotropic square waveguide whose optic axis lies in the xy-plane at an angle $\theta = 45°$ from the x axis.

6.4.2 Anisotropic permeability

For validation purposes, the geometry shown in Figure 14 was simulated and validated against the eigenvalues obtained by `Ansys HFSS`. The permittivity in the loaded ferrite is $\varepsilon_r = 10$ and the permeability is assumed to be complex with

$$\boldsymbol{\mu_r} = \begin{bmatrix} 1 & 1i & 0 \\ 1i & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{84}$$

so in this case an expansion of the problem to size $2n \times 2n$ is needed following the explanation in Section 5.1, but this time for the mass matrix $\mathbf{M}$.

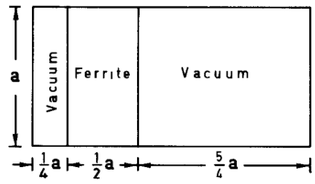

Figure 14: The cross section of a ferrite-loaded waveguide.

The phase constants of the first four modes are presented in Figure 15, where again the colored lines show the simulation results obtained from `Ansys HFSS` and the black dotted ones with the presented formulation. The agreement is perfect and it is important to note that the two bumps in the orange and green line are non-physical and due to the fact that the solution in `Ansys HFSS` could not fully converge for that parameters.

6.5 Fully Anisotropic Materials

As I could not find an example where eigenvalues where computed for the fully anisotropic case, the approach for validation is to take first orthotropic matrices and then push them to fully anisotropic ones by increasing the entries on the off-diagonal terms. The studies have been performed on a homogeneous rectangular waveguide of dimensions 15 mm × 10 mm.

6.5.1 Permittivity

First that has been done for the relative permittivity tensor, while the permeability was homogenously 1.0. The permittivity matrix was varied with

$$\boldsymbol{\varepsilon_{orth}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad \boldsymbol{\varepsilon_1} = \begin{bmatrix} 2 & 0.5 & 0.1 \\ 0.5 & 4 & 0.2 \\ 0.1 & 0.2 & 1 \end{bmatrix}, \quad \boldsymbol{\varepsilon_2} = \begin{bmatrix} 2 & 0.7 & 0.3 \\ 0.7 & 4 & 0.3 \\ 0.3 & 0.3 & 1 \end{bmatrix} \tag{85}$$

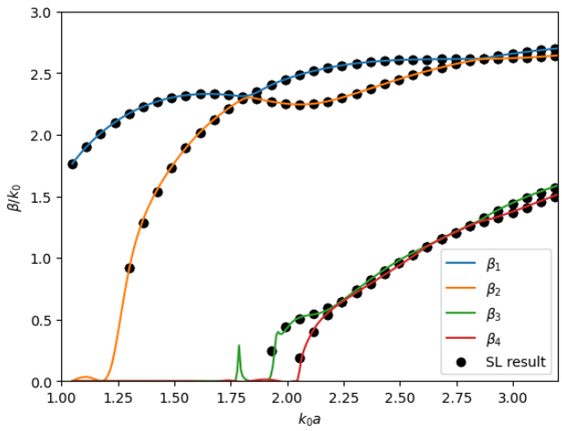

Figure 15: Dispersion characteristics of an waveguide with anisotropic permeability (validated against `Ansys HFSS`)

In Figure 16 it can be seen that the bigger the off-diagonal entries of the material tensor, the more the eigenvalues differ from the ones corresponding to the orthotropic problem. In fact, some modes even start to be lossy for $\boldsymbol{\varepsilon_2}$.

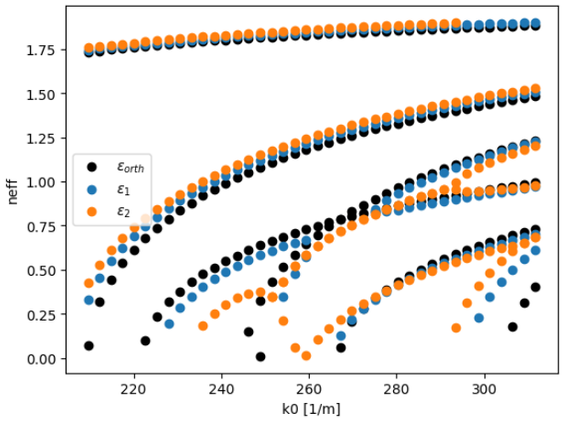

Figure 16: Dispersion characteristics of an waveguide with fully anisotropic permeability

6.5.2 Permeability

Secondly, the contribution from a fully anisotropic permeability was analyzed while the relative permittivity was kept isotropic at 1.0. The three following material tensors

$$\boldsymbol{\mu_{\text{orth}}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{\mu_1} = \begin{bmatrix} 2 & 0.1 & 0.2 \\ 0.1 & 4 & 0.1 \\ 0.2 & 0.1 & 1 \end{bmatrix}, \quad \boldsymbol{\mu_2} = \begin{bmatrix} 2 & 0.5 & 0.3 \\ 0.5 & 4 & 0.2 \\ 0.3 & 0.2 & 1 \end{bmatrix} \tag{86}$$

have been studied. The result for the first two propagating modes are displayed in Figure 17. We can observe that for off-diagonal terms approaching zero, it approaches the solution for orthotropic permeability, a case that has been validated. That gives a good argument to believe that the formulation has been implemented correctly.

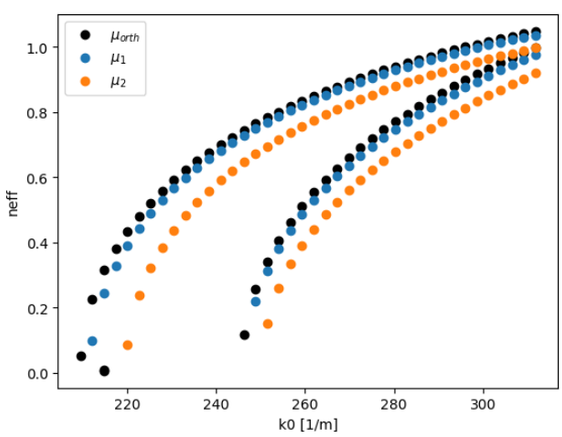

Figure 17: Dispersion characteristics of an waveguide with fully anisotropic permeability

6.5.3 Permittivity and Permeability fully anisotropy

The same procedure was also applied having both material tensor first orthotropic ($\boldsymbol{\varepsilon_{orth}}, \boldsymbol{\mu_{orth}}$) and then anisotropic ($\boldsymbol{\varepsilon_1}, \boldsymbol{\mu_1}$) . The electric field for one mode at 12 GHz is normalized in a way that a power of 1 W is fed into the waveguide is shown in Figure 18 and the influence of the off-diagonal terms in the material tensor can clearly be observed.

(a) orthotropic $\boldsymbol{\mu}$ and $\boldsymbol{\varepsilon}$

(b) fully anistropic $\boldsymbol{\mu}$ and $\boldsymbol{\varepsilon}$

Figure 18: E-Field of a supported mode in a rectangular waveguide (normalized to 1 W)

References

[1] J. D. Jackson, "Classical electrodynamics," 1999.

[2] C. Reddy, *Finite element method for eigenvalue problems in electromagnetics*. NASA, Langley Research Center, 1994, vol. 3485.

[3] A. Halbach, *Sparselizard documentation*, Jan. 2023, accessed: December 7, 2023. [Online]. Available: https://www.sparselizard.org/DOCUMENTATION.pdf

[4] V. Hernández, J. E. Román, A. Tomás, and V. Vidal, "Krylov-schur methods in slepc," *Universitat Politecnica de Valencia, Tech. Rep. STR-7*, 2007.

[5] J. E. Roman, C. Campos, L. Dalcin, E. Romero, and A. Tomás, *SLEPc Users Manual*, Sep. 2023, accessed: December 7, 2023. [Online]. Available: https://slepc.upv.es/documentation/slepc.pdf

[6] T. Angkaew, M. Matsuhara, and N. Kumagai, "Finite-element analysis of waveguide modes: A novel approach that eliminates spurious modes," *IEEE transactions on microwave theory and techniques*, vol. 35, no. 2, pp. 117–123, 1987.

[7] C. Johansson, "Numerical methods for ports in closed waveguides," Ph.D. dissertation, Numerisk analys och datalogi, 2003.

[8] J. A. Svedin, "A numerically efficient finite-element formulation for the general waveguide problem without spurious modes," *IEEE transactions on microwave theory and techniques*, vol. 37, no. 11, pp. 1708–1715, 1989.

[9] A. Zdunek and W. Rachowicz, "Inhomogeneous lossy waveguide mode analysis," *Computers & Mathematics with Applications*, vol. 75, no. 3, pp. 798–808, 2018.

[10] Y. Lu and F. A. Fernandez, "An efficient finite element solution of inhomogeneous anisotropic and lossy dielectric waveguides," *IEEE transactions on microwave theory and techniques*, vol. 41, no. 6, pp. 1215–1223, 1993.

[11] M. Ohtaka, "Analysis of the guided modes in the anisotropic dielectric rectangular waveguides," *Trans. Inst. Electron. Commun. Eng. Japan*, vol. 64, pp. 674–681, 1981.

[12] Y. Lu, *The finite element solution of inhomogeneous anisotropic and lossy dielectric waveguides*. University of London, University College London (United Kingdom), 1991.

A Example Code

Algorithm 1: "Example script"

```
#include "sparselizard.h"

using namespace sl;

int main(void)
{
// 1. Mesh import and region assignment
    int wg = 5, bound = 13;

    mesh mymesh;
    mymesh.selectskin(bound, all);
    mymesh.load("RectangularWG.msh");


// 2. Field definition
//Edge shape functions 'hcurl' for the tranverse electric field Et.
// Nodal shape functions 'h1' for the longitudinal electric field El.

    field Et("hcurl"), El("h1");
    Et.setorder(all, 2);
    El.setorder(all, 3);

// 3. Set Perfect conductor boundary condition:
    El.setconstraint(bound);
    Et.setconstraint(bound);

// 4. Material properties definition

    parameter epsr, nur;
    expression epsr_an(3,3,{2.0,0.7,0.3,
                            0.7,4.0,0.3,
                            0.3,0.3,2.0});
    expression nur_an(3,3,{0.25916, 0.0597, -0.14925,
                            0.0597, 0.53731, 0.03392,
                            -0.14925,0.03392, 1.05156});
    epsr|wg = epsr_an;
    nur|wg = nur_an;

// Define operating frequency
   double f0 = 12e9, c = 299792458, lambda = c/f0, k0 = 2.0*getpi()/lambda;

// 5. Weak Formulation Construction

    // Operators grad() and curl() in the transverse plane:
    expression gradtfEl = grad(tf(El)).resize(3,1);
    expression graddofEl = grad(dof(El)).resize(3,1);
    expression dtdtgraddofEl = dtdt(grad(dof(El))).resize(3,1);
    expression dtdtdofEl(3,1,{0,0, dtdt(dof(El))});
    expression dtgraddofEl = dt(grad(dof(El))).resize(3,1);

    formulation mode;

    //matrix M:
    mode += integral(wg, nur * (dtdt(dof(Et)) - dtdtgraddofEl)
            * tf(Et));
    mode += integral(wg,  nur * (dtdt(dof(Et)) - dtdtgraddofEl)
            * gradtfEl);

    //matrix C:
      mode += integral(wg, -k0*k0*epsr*dt(dof(Et))*tfEl
              - k0*k0*epsr*dtdofEl * tf(Et));
      mode += integral(all, -nur * dtgraddofEl * curl(tf(Et))
                - nur * dt(dof(Et))*curl(tf(Et))
                + nur * curl(dt(dof(Et)))*tf(Et)
                - nur*curl(dt(dof(Et)))*gradtfEl);

    //matrix K:
      mode += integral(wg, +nur * curl(dof(Et)) * curl(tf(Et))
```

```
            - k0*k0*epsr * dof(Et) * tf(Et)
            + k0*k0*epsr * dtdtdofEl * tfEl);

// 6. Galerkin matrices assembly

      mode.generate();

      // Get the stiffness matrix K, damping matrix C and mass matrix M:
      mat K = mode.K();
      mat M = mode.M();
      mat C = mode.C();

// 7. Create the object to solve the eigenvalue problem:

      eigenvalue eig(K, C, M);

      // Compute the 10 eigenvalues closest to the target magnitude.
      double neff_target = 1.0, bt = k0*neff_target;

      eig.compute(10, -bt);
      eig.printeigenvalues();

// 8. Postprocessing
    // Get all eigenvectors and eigenvalues found:
      std::vector<vec> realeigenvectors = eig.geteigenvectorrealpart();
      std::vector<double> realeigenvalues = eig.geteigenvaluerealpart();
      std::vector<vec> imageigenvectors = eig.geteigenvectorimaginarypart();
      std::vector<double> imageigenvalues = eig.geteigenvalueimaginarypart();

      // Loop on all eigenvalues found:
      int index = 1;
      for (int i = 0; i < myrealeigenvalues.size(); i++)
      {
        // Transfer the data from the ith eigenvector to fields Et and El:
        Et.setdata(wg, realeigenvectors[i]);
        El.setdata(wg, realeigenvectors[i]);


        // Compute the propagation constant and the effect. refract. index:

        double beta = -realeigenvalues[i];
        double neffc = beta/k0;

        double alpha =imageigenvalues[i];
        double neffa = alpha/k0;

        // Display mode information:
        std::cout << "Mode " << index << ": real = " << btc << " rad/m,
        imag = " << atc << " rad/m "<< std::endl;

        index++;

       }
}
```