Acknowledgements:

First, I would like to thank Thomas Felici. He gave me a desk in its own office so that we could work together. I have been disturbing him all the time with questions about C++ or waveguide theory. He took the time to answer all of them and I owe him much of this work. His enthusiasm and kindness were the best support I could get.

I also would like to thank Mike Giles for his precious help concerning the adjoint method.

I enjoyed very much the time I spent in *PhotonDesign* and many thanks to everyone there.

Finally, I would like to thank Hilary Ockendon for her support and for organising this master I very much appreciated.

Contents

Introduction **4**

1 Modal Analysis and Computation of Optical Waveguides **6**

1.1 General Assumptions 7

1.2 Uniform and Homogeneous waveguides 8

1.3 Guided, Radiation and Evanescent modes 12

1.4 Eigenmode expansion in a uniform waveguide 16

1.5 Propagation in non z-uniform waveguides 23

1.6 *Fimmprop-3d* : the Maxwell's equations solver of *PhotonDesign* 25

2 Computation of the Derivative using Adjoint Method **27**

2.1 Introduction to the adjoint approach to optimisation 27

2.2 Computation of the derivative in the 2D case 30

2.2.1 Presentation of the 2D problem 30

2.2.2 Lagrangian Formulation 32

2.2.3 Adjoint Formulation 35

2.3 Computation of the derivative in the 3D case 38

2.3.1 Equations and boundary conditions 38

2.3.2 Computation of the adjoint operators 41

2.3.3 Boundary conditions for the adjoint problem 43

2.3.4 Computation of the derivative P_m for a shape deformation 44

2.4 Computation of the derivative combined with modal expansion 51

3 Numerical Evaluation **54**

3.1 General framework 54

3.2 Numerical computation of the derivative expression 55

3.3 Test cases 59

3.3.1 Test case 1 : uniform waveguide 59

3.3.2 Test case 2 60

Appendix **64**

1 Generalisation to any objective function 64

2 Test case 2 : calculation details 65

2.1 Modes normalisation 65

2.2 Computation of the direct problem 65

2.3 Computation of the adjoint problem 66

Conclusion **68**

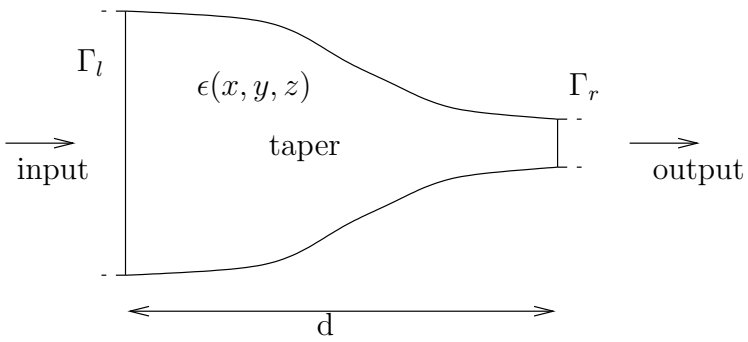
Further works 69

Bibliography **70**

Introduction

The photonics industry is a thriving industry which started twenty years ago when the optical fibre revolutionised telecommunications. Since then, in many places, optical fibres have replaced copper wires transmitting signals at remarkable speeds. Nevertheless at different stages, signals need to be processed and rooted. This task, if done by electronic means, slows down the traffic and undermines the optical fibre performances. However, in recent years, optical components have been developed which can carry out some of those processing. Designing new components offering more functions is the challenge of the photonics industry. As in any field which involves design, efficient numerical simulation is very much needed. The software *Fimmprop-3d* developped by *PhotonDesign* is a comprehensive tool which solves Maxwell's equations for most of these components.

Tapered waveguides are commonly used to channel guided light from an input waveguide into an output waveguide with different cross-section (typically between two optical fibre of different width). If the distance between the sections is sufficently long, a slowly varying taper can work in the adiabatic regime and the power loss can be minimized in this way (cf [5]). However in optical components length is a constraint which has to be taken into account and a taper need to be optimised in order to fit with the dimension of the component. The optimisation problem is then : Given two section profiles Γ_l and Γ_r located at a distance d one from the other, what is the shape of the taper which maximises the transmitted power ?



In pratice we are interested in maximising the power remaining in one mode, P_m (m is the mode number). The shape of a taper is in fact determined by the electric permittivity distribution $\epsilon(x, y, z)$ and we want to find the optimal $\epsilon_{\max}$:

$$P_m(\epsilon_{\max}) = \max_{\epsilon} P_m(\epsilon)$$

Optimal design methods involving the solution of an adjoint system of equations are more and more efficently used in computational fluid dynamics particularly for aeronautical applications. The aim of this dissertation is to apply these methods to electromagnetism in the taper case and find an adjoint formulation for the derivative $\partial P_m/\partial\epsilon$. The adjoint method can lead to a huge gain in computation time. Implementing it should improve a lot any optimisation routine which requires at each step the computation of the derivative. This is the aim of the computational part of this industrial project.

The Maxwell's equations solver *Fimmprop-3d* is completely based on waveguide theory and modal expansion. The first chapter provides the necessary background to understand where the modal expansions come from and how they can be used to implement a solver such as *Fimmprop-3d* . In the second chapter, we derive an adjoint formulation for the taper case from which we deduce the derivative $\partial P_m/\partial\epsilon$. The last chapter is dedicated to the numerical implementation of this derivative and its utilisation in *Fimmprop-3d* within an optimisation module.

Chapter 1

Modal Analysis and Computation of Optical Waveguides

"There is nothing so practical as a good theory"

Will Harvey [1]

A waveguide is a device which carries energy from one point to another. The most simple waveguide is the uniform (invariant in the direction of propagation) and homogeneous waveguide which will be studied in section 1.2. In this case, each mode carries a certain amount of energy without interfering with the others. However for many common waveguides, energy is radiated on the sides. The section 1.3 explains with the simplified case of a planarly layered medium how radiation is integrated in eigenmode expansion and how the numerical results have to be interpreted. The section 1.4 gives some insight on the eigenmode expansion in an arbitrary uniform waveguide. In section 1.5, we assemble together bits of uniform waveguides and then we are able to model any particular waveguide. These two sections provide the theoretical background to understand how *Fimmprop-3d* , the Maxwell's equations solver of *PhotonDesign* works.

[1] I found this quotation in P. Bienstman's thesis [1] which gives very clear explanations about the utilisation of eigenmode expansions for waveguides

1.1 General Assumptions

Our starting point is Maxwell's equations.

$$\begin{cases} \nabla \times E + \dfrac{\partial B}{\partial t} &= 0 \\ \nabla \times H - \dfrac{\partial D}{\partial t} &= J \\ \nabla . D &= \rho \\ \nabla . B &= 0 \end{cases}$$

where

E	electric field
B	magnetic field
D	electric flux density
H	magnetic density
J	current density
ρ	charge density

Since there are 12 unknowns for 8 scalar equations, complementary relations are needed to solve Maxwell's equations. These remaining relations are the constitutive relations which are related to the material properties. Since will only consider isotropic dielectric medium, we have

$$\begin{aligned} D &= \epsilon E \\ B &= \mu H \end{aligned}$$

where

ϵ	electric permitivity
μ	magnetic permeability

In the remain of this report, we will always assume μ to be constant and equal to μ_0, the magnetic permeability in free space. We also consider a medium without source of charge or current and Maxwell's Equations reduce to

$$\begin{cases} \nabla \times E + \mu \dfrac{\partial H}{\partial t} &= 0 \\ \nabla \times H - \epsilon \dfrac{\partial E}{\partial t} &= 0 \\ \nabla . \epsilon E &= 0 \\ \nabla . \mu H &= 0 \, . \end{cases}$$

For periodic phenomena, we introduce time dependancy of the form

$$\begin{aligned} E(\mathbf{x}, t) &= E(\mathbf{x}) e^{-i\omega t} \\ H(\mathbf{x}, t) &= H(\mathbf{x}) e^{-i\omega t} \end{aligned}$$

and we end up with the following equations which are the form of Maxwell's equations we will use in the next sections.

$$\begin{cases} \nabla \times E - i\omega\mu H &= 0 \\ \nabla \times H + i\omega\epsilon E &= 0 \\ \nabla . \epsilon E &= 0 \\ \nabla . \mu H &= 0 \, . \end{cases}$$

1.2 Uniform and Homogeneous waveguides

We consider in this section a uniform and homogeneous waveguide. By uniform waveguide, we mean that the waveguide is invariant in the direction of propagation z. An homogeneous waveguide is a waveguide with a constant electric permitivity ϵ.

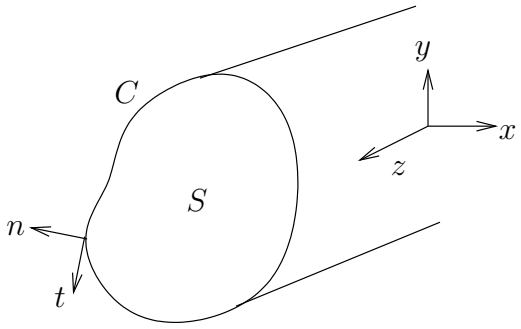


Figure 1.1: uniform homogeneous waveguide

The boundary conditions on C (cf Figure (1.1)) are

$$E \times n = 0$$

which is equivalent to

$$E_t = E_z = 0$$

We denote E_s and E_z the tangential and perpendicular components of E with respect to the cross-section S and $E_s = E_t\hat{t} + E_n\hat{n}$.

These boundary conditions correspond to a perfectly conducting material located on C such as metal and are called *electric wall* or *hard wall* conditions.

Since the medium is homogeneous, ϵ is constant and Maxwell's equations give us

$$\nabla \times \nabla \times E - \epsilon\mu\omega^2 E = 0 \tag{1.1}$$

and since $\nabla \times \nabla \times E = \nabla\nabla.E - \nabla^2 E$ and $\nabla.E = 0$ (from Maxwell's equations), (1.1) can be rewritten as a vector wave equation

$$\nabla^2 E + k^2 E = 0 \tag{1.2}$$

where $k^2 = \epsilon\mu\omega^2$

We will first focuss on E_z, the $\hat{z}$ component of E. It satisfies the scalar wave equation

$$\nabla^2 E_z + k^2 E_z = 0 \tag{1.3}$$

$$E_z = 0 \text{ on } C \tag{1.4}$$

If we now define the eigenfunctions ϕ_m as follows

$$\begin{aligned} \nabla^2\phi_m + \nu_m^2\phi_m = 0 \text{ in } S \\ \phi_m = 0 \text{ on } C \end{aligned}$$

where ν_m^2 is the eigenvalue associated to ϕ_m, the Sturm-Liouville Theorem tells us then that the ϕ_m form an orthogonal basis of the space of the continuous functions which vanish on C. Hence, for each z, $E_z(x, y, z)$ can be expanded in term of $\phi_m(x, y)$ and so, there exist a sequence $a_m(z)$ such that

$$E_z(x, y, z) = \sum_{m\geq 1} a_m(z)\phi_m(x, y) \tag{1.5}$$

Plugging (1.5) into (1.3), we get

$$\sum_{m\geq 1} \left(\frac{d^2}{dz^2} a_m(z) + (k^2 - \nu_m^2)a_m(z) \right) \phi_m = 0$$

and, by using the orthogonality of the ϕ_m (i.e. $\int_S \phi_m.\phi_n \; dxdy = \delta_{m,n}$), we end up with

$$a_m''(z) + (k^2 - \nu_m^2)a_m(z) = 0 \quad \text{for } \forall m \tag{1.6}$$

finally, after solving (1.6) for all m, we get the following decomposition:

$$E_z = \sum_{m\geq 1} (A_m e^{i\lambda_m z} + B_m e^{-i\lambda_m z})\phi_m$$

where

$$\lambda_m^2 = k^2 - \nu_m^2 \tag{1.7}$$

Since H satisfies the same scalar wave equation (1.3) as E, it is sensible to decompose H_z in the same way as we did for E_z but we first have to find out what the boundary conditions for H_z are.

On C, $E_t = E_z = 0$ therefore we have:

$$\nabla \times E \,.\, n = 0$$

Since $H = \frac{1}{i\omega\mu}\nabla \times E$, it implies that

$$H_n = 0 \text{ on } C \tag{1.8}$$

Furthermore, since on C, $\nabla \times H \,.\, s = -i\omega\epsilon E_t = 0$, we have

$$\frac{\partial H_z}{\partial n} - \frac{\partial H_n}{\partial z} = \nabla \times H \,.\, s = 0$$

and finally, after using (1.8), we end up with the following boundary condition for H_z:

$$\frac{\partial H_z}{\partial n} = 0 \tag{1.9}$$

Hence, we have to introduce another set of eigenfunctions ψ_m satisfying (1.9) on the boundary.

$$\nabla^2 \psi_m + \mu_m^2 \psi_m = 0 \text{ on } S$$
$$\frac{\partial \psi_m}{\partial n} = 0 \text{ on } C$$

and we can derive the following expansion for H_z

$$H_z = \sum_{m \geq 1} (C_m e^{i\kappa_m z} + D_m e^{-i\kappa_m z}) \psi_m$$

where

$$\kappa_m^2 = k^2 - \mu_m^2 \tag{1.10}$$

It is now possible to exhibit two sets of modes for the waveguide.

- The first set is induced by setting

  $$E_z(x, y, z) = \phi_m(x, y) e^{\pm i\lambda_m z} \text{ and } H_z(x, y, z) = 0$$

  The tangential components are then retrieved by using Maxwell's equations (please refer to Jones' book [8] for more details about this calculation). Since the axial component of H is zero, these modes are called transverse magnetic (T.M.) modes.

- The second set is induced by setting

  $$E_z(x, y, z) = 0 \text{ and } H_z(x, y, z) = \psi_m(x, y) e^{\pm i\kappa_m z}$$

  these modes are naturally called transverse electric (T.E.) modes.

One can show (the proof is in [8]) that in the case of a simply connected domain S, the T.E. and T.M. modes that we have just described form a complete basis ; any electro-magnetic field in a homogeneous and uniform waveguide can be split into a transverse magnetic and transverse electric part and expanded as follows:

$$E = \sum_{m \geq 1} A_m e^{i\lambda_m z} E_m^M + B_m e^{-i\lambda_m z} E_{-m}^M + C_m e^{i\kappa_m z} E_m^E + D_m e^{-i\kappa_m z} E_{-m}^E$$

and a similar expansion for H. $(e^{i\pm\lambda_m z}E^M_{\pm m})$ and$(e^{i\pm\kappa_m z}E^E_{\pm m})$ are respectively the T.M. and T.E. modes.

The propagating constant defined in (1.7) λ_m (or in (1.10) κ_m) of a mode can be imaginary if $k^2 < \nu_m^2$ and the contribution of such a mode becomes

$$A_m e^{-|\lambda_m|z}E_m + B_m e^{|\lambda_m|z}E_m$$

For a semi-infinite waveguide $z > 0$, this is only acceptable if $B_m = 0$. Since this mode then decays exponentially, it is said to be cut off and we can define from (1.7) a cutoff wavelength

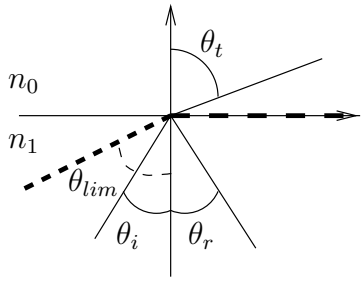
$$\frac{2\pi}{\lambda_{\text{cutoff}}} = \nu_1$$

where ν_1 is the smallest of the ν_m. Any entering signal with a wavelength higher then λ_{cutoff} is not guided by the waveguide but cut off.

This kind of waveguide has been historically the first to be studied. Hollow tubes in metal are extensively used as antennas in radiotelecommunication. Nevertheless, the technology that has emerged during the past twenty years with the optical fibre relies on a different principle. The guidance of the electromagnetic field in an optical fibre doesn't come from the boundaries as for the hollow tube (hard wall boundary conditions) but from a sharp gradient in the refractive index distribution. It is easy to get a rough idea of how an optical fibre works by simply using the Snell's laws (which are a first order approximation of Maxwell's equations). Snell's laws tell us that a planar wave at the boundary between two homogeneous media of different refractive index splits into a reflected and transmitted waves with angles obeying the following equality:

$\sin\theta_t = \frac{n_1}{n_0}\sin\theta_i$
$\theta_r = \theta_i$

In this example,
$n_1 > n_0$ and $\sin\theta_{lim} = \frac{n_0}{n_1}$



If $n_1 > n_0$, there exists a limit angle θ_{lim} for which there is not any transmitted wave. If we now consider an optical fibre with a core and a cladding of refractive index n_1 and n_0 ($n_1 > n_0$) as depicted in Figure (1.2), a wave entering with an angle bigger than θ_{lim} will never escape the core and therefore is guided. On the other hand, a wave entering with an angle smaller than θ_{lim} will emit a transmitted wave each time it hits the core boundary

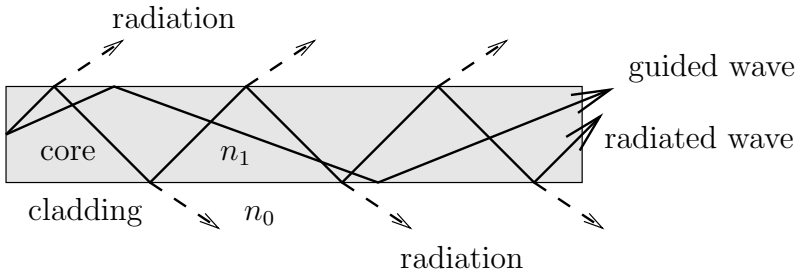


Figure 1.2: guided and radiation modes in optical fibre

and so loses power. At a sufficient distance, nothing remains of this wave; it has been completely radiated.

As far as modal expansions are concerned, this drives us to take into consideration radiation modes and add them to the guided modes in order to form a complete set of modes.

$$E = E_{\text{guided}} + E_{\text{radiation}}$$

The problem with radiation modes is that they correspond to a continuum set of eigenvalues (see [11]) which makes them very tricky to handle numerically. A way to overcome this problem is to introduce artificial boundary conditions such as hard walls far from the region of interest for the computation. The aim of the next section is to explain how the modes are modified by such boundaries and how the results we then get have to be interpreted.

1.3 Guided, Radiation and Evanescent modes

In this section, we will consider a planarly layered medium as depicted in Figure (1.3). This is the simplest case of inhomogeneous medium but nevertheless it gives us the main features of the modes and allows us to classify them in three important and physically meaningful categories : Guided, radiation and evanescent modes.

The medium we are considering is constituted of N layers invariant in the y and z direction. Each layer is characterised by its thickness and its constant electric permitivity ϵ_i ($\mu = \mu_0$ everywhere). ϵ is then a piece-wise constant function of x. At $x = x_{min}$ and $x = x_{max}$ we have hard wall boundary conditions. We assume that the fields are invariant in the $\hat{y}$ direction and therefore we deal with a 2D problem.

One can then derive the following equations for E_y and H_y (refer to [3] for

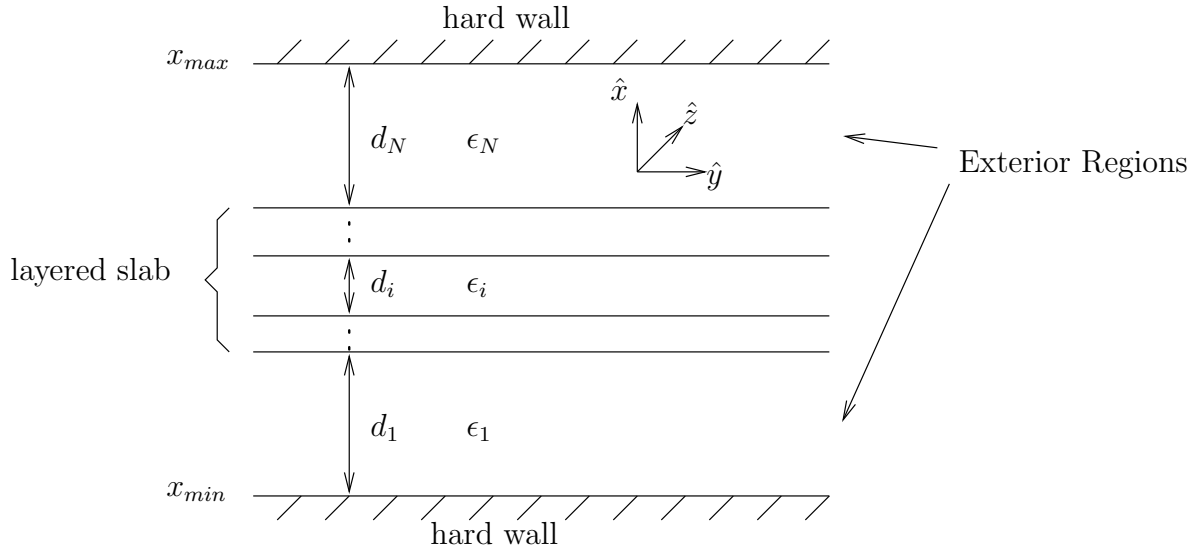


Figure 1.3: Planarly layered medium

more details about the calculations):

$$\left[\mu\frac{\partial}{\partial x}\mu^{-1}\frac{\partial}{\partial x} + k^2 + \frac{\partial^2}{\partial z^2}\right] E_y(x) = 0$$
$$\left[\epsilon\frac{d}{dx}\epsilon^{-1}\frac{d}{dx} + k^2 + \frac{\partial^2}{\partial z^2}\right] H_y(x) = 0$$

where $k = \epsilon\mu\omega^2$.

One can show that any electromagnetic field can be decomposed into transverse electric components ($H_y = 0$ which implies $E_x = E_z = 0$) and transverse magnetic components (E_y=0 which implies $H_x = H_z = 0$) components and therefore we are led to solve the following eigenvalue problem for $p = \epsilon$ and $p = \mu$:

$$\mathcal{L}\phi = k_z^2\phi \tag{1.11}$$

where k_z^2 is the eigenvalue and $\mathcal{L}$ the differential operator defined as follows:

$$\mathcal{L}\phi(x) = \left[p\frac{d}{dx}p^{-1}\frac{d}{dx} + k^2\right]\phi(x) \tag{1.12}$$

with homogeneous Dirichlet or Neuman boundary conditions.

The scalar product defined in (1.13) makes $\mathcal{L}$ self-adjoint.

$$< u, v >= \int_{x_{min}}^{x_{max}} p^{-1}(x)u(x)v(x)\ dx \tag{1.13}$$

Indeed, by integrating by parts and using the Dirichlet or Neuman boundary conditions satisfied by u and v, we have

$$\begin{aligned} < \mathcal{L}u, v > &= \int_{x_{min}}^{x_{max}} \frac{d}{dx}[p^{-1}(x)\frac{d}{dx}u(x)]\ v(x)\ dx + \int_{x_{min}}^{x_{max}} k^2 p^{-1}(x)u(x)v(x)\ dx \\ &= - < \frac{du}{dx}, \frac{dv}{dx} > + < ku, kv > \\ &= < u, \mathcal{L}v > \end{aligned}$$

The Dirichlet or Neuman boundary conditions implies that the eigenvalues of (1.11) form a discrete set and since $\mathcal{L}$ is self-adjoint, the eigenfunctions form an orthonormal basis with respect to the scalar product (1.13) of the continuous functions satisfying the same boundary conditions.

In the simple case of a planarly layered medium, we can find explicitly the eigenfunctions. In each layer, ϵ is constant and we have to solve

$$\frac{d^2\phi}{dx^2} + (k_i^2 - k_z^2)\phi = 0 \tag{1.14}$$

where $k_i = k = \epsilon_i \mu \omega^2$.

ϕ_i, the solution in the ith region, can be then written as

$$\phi_i = A_i e^{-ik_{ix}x} + B_i e^{ik_{ix}x} \tag{1.15}$$

where

$$k_{ix}^2 = k_i^2 - k_z^2$$

The value of A_i and B_i are found by forcing the continuity of ϕ and $\frac{1}{p}\frac{d\phi}{dx}$. We end up with the right number of equations to calculate the eigenvalues k_z^2. We do not pursue this calculation here and the results can be found in Chew's book [3]. We rather focuss on a qualitative analysis of the eigenvalues k_z^2 and the corresponding eigenvectors.

k_z^2 is always real because it is eigenvalue of a self-adjoint operator. Hence, k_z can be either real or purely imaginary depending on the sign of k_z^2 and two cases have to be considered.

1. If $k_z^2 < 0$, k_z is purely imaginary and the corresponding mode is said to be evanescent. The mode is in fact cut off as described in the previous section and decays as
$$e^{ik_z z} = e^{-|k_z|z}$$

2. If $k_z^2 > 0$, k_z is real and the corresponding mode is guided. However two cases show up depending on the values of the propagating constants, k_{1x} and k_{Nx}, in the exterior regions (note that we are now talking about propagation in the $\hat{x}$ direction.)

   (a) k_{1x} and k_{Nx} are purely imaginary. It occurs when $|k_1| < |k_z|$ and $|k_N| < |k_z|$. The mode behaves in region 1 and N as a combination of real exponential (cf (1.15)) and one can show that, after normalization, only a small fraction of the energy travels in the exterior regions. The most important part of the mode energy remains in the layered slab. Hence, the mode is effectively guided by the slab. Furthermore, since the mode slowly vanishes at the walls, it will not be significantly modified as we move the walls away.

   (b) k_{1x} or k_{Nx} or both are real. The mode is still guided (k_z is real) and will propagate along the direction $\hat{z}$ but most of the power lies in the exterior regions. The mode is in fact guided by the hard walls on the boundaries and for this reason it will be very sensitive to a shift of the boundaries. As it carries energy outside the slab, this exterior mode corresponds to a radiation mode.

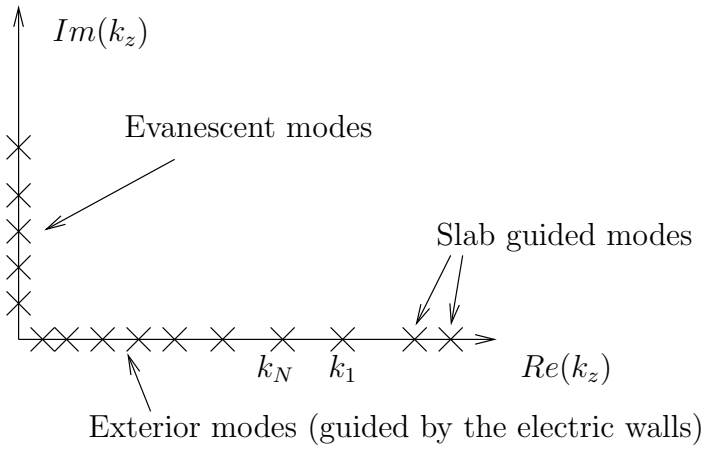


Figure 1.4: distribution of k_z

In practice, we are interested in an open region where the walls we have introduced for computation convenience are replaced by more realistic boundary conditions. A first improvement is to let x_{min} and x_{max} tend to infinity. There is no more guidance by the walls and the eigenvalues described above and in Figure (1.4) are modified in different ways. The slab guided modes

are almost not changed because they are not sensitive to the location of the boundaries. The corresponding eigenvalues still belong to a discrete set. However, as the walls move away, the number of exterior modes increases because there is "more room" for them in the exterior regions 1 and N and therefore there are more and more radiation and evanescent modes. In the limit case where the walls are infinitely far, the eigenvalues corresponding to those modes form a continuous set. The linear operator $\mathcal{L}$ (or rather $\mathcal{L}^{-1}$) is not compact any more and has a continuous spectrum. The evanescent and radiation modes come from this continuous spectrum.

The hard wall boundary conditions are "artificial" boundary conditions. Sommerfield boundary conditions (1.16) would have been more appropriate since they really model an open region with no reflection at the boundaries.

$$\begin{pmatrix} E \\ H \end{pmatrix} e^{i\omega t} \sim \frac{e^{i(\omega t - k|x|)}}{|x|} \quad \text{as } |x| \to \infty \tag{1.16}$$

However they also lead to a continuous set of radiation modes which is difficult to handle numerically. When we introduce hard wall boundary conditions far away from the region of interest (the layered slab), we tend to approximate the Sommerfield boundary conditions and the results we get in the vicinity of the layered slab are acceptable. For example, the ray 3 in Figure (1.5) does not interfere with the region of interest (colorred in grey in the figure) but the ray 2 does. Nevertheless, even the limit case where the walls are at infinity is not equivalent to Sommerfield boundary conditions because the ray 1 in Figure (1.5) which is in the perpendicular direction of propagation is reflected in one case and not in the other. Hence, we have to be very careful as we come to interpret computational results.

1.4 Eigenmode expansion in a uniform waveguide

We consider a uniform (invariant in the z direction) waveguide but with an arbitrary electric permitivity $\epsilon(x, y)$ within the cross-section. Maxwell's equations naturally lead us to the following equations for E:

$$\mu\nabla \times (\frac{1}{\mu}\nabla \times E) - k^2 E = 0 \tag{1.17}$$

$$\nabla.(\epsilon E) = 0 \tag{1.18}$$

The boundary conditions are again electric wall boundary conditions i.e.

$$E \times n = 0 \tag{1.19}$$

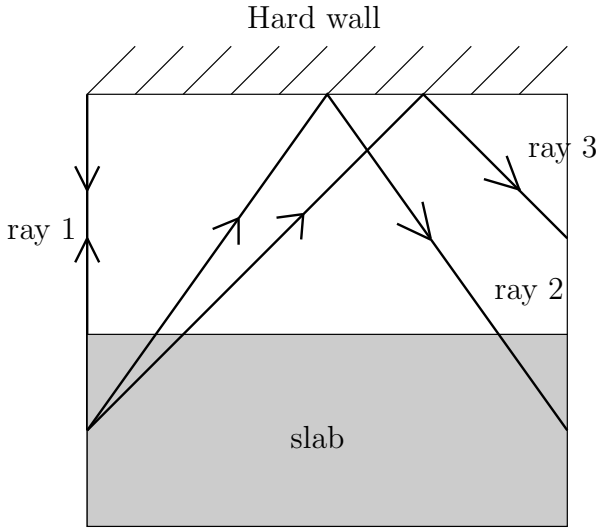


Figure 1.5: Hard wall boundary conditions induce reflections which do not occur in a open waveguide

We split E into its tangential and perpendicular parts (with respect to the cross-section S), E_s and E_z.

$$\begin{aligned} E_s &= (E.\hat{x})\hat{x} + (E.\hat{y})\hat{y} \\ E_z &= E.\hat{z} \end{aligned}$$

We also introduce the differential operator ∇_s which only acts on the $\hat{x}$ and $\hat{z}$ components.

$$\nabla_s = \begin{pmatrix} \frac{\partial .}{\partial x} \\ \frac{\partial .}{\partial y} \\ 0 \end{pmatrix}$$

Thus, we have $\nabla = \nabla_s + \hat{z}\frac{\partial .}{\partial z}$

Since the waveguide is z invariant, by equation (1.18), E_z can be expressed in function of the tangential component E_s:

$$\frac{\partial E_z}{\partial z} = \frac{1}{\epsilon}\nabla_s.\epsilon E_s \tag{1.20}$$

Plugging this into (1.17), one can get rid of E_z:

$$\mu\nabla_s \times (\mu^{-1}\nabla_s \times E_s) - \nabla_s(\epsilon^{-1}\nabla_s.\epsilon E_s) - (k^2 + \frac{\partial^2}{\partial z^2})E_s = 0 \tag{1.21}$$

This equation naturally leads us to the following eigenvalue problem in the cross-section S

$$\mu\nabla_s \times (\mu^{-1}\nabla_s \times E_s) - \nabla_s(\epsilon^{-1}\nabla_s.\epsilon E_s) - (k^2 - \beta^2)E_s = 0 \tag{1.22}$$

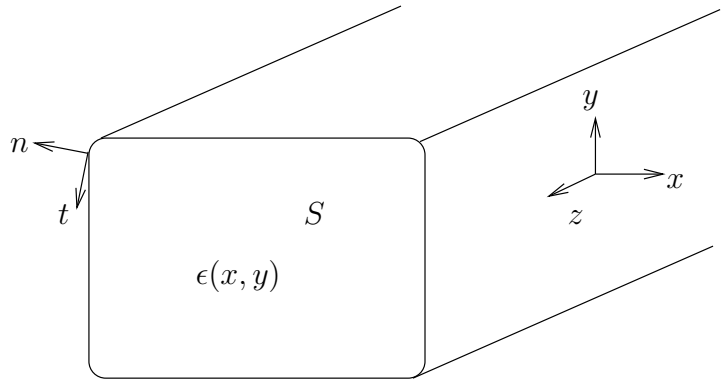


Figure 1.6: Uniform waveguide

Once (1.22) is solved, E_z can be easily retrieved by equation (1.20):

$$E_z = \frac{1}{i\epsilon\beta}\nabla_s.\epsilon E_s \tag{1.23}$$

Following Chew in [3], we multiply equation (1.22) by $\hat{z}\times$ and define the linear operators $\mathcal{L}$ and $\mathcal{B}$ as follows:

$$\begin{aligned}\mathcal{L}E_s &= \mu\hat{z}\times\nabla_s\times(\mu^{-1}\nabla_s\times E_s) - \hat{z}\times\nabla_s(\epsilon^{-1}\nabla_s.\epsilon E_s) - k^2\hat{z}\times E_s\\ \mathcal{B}E_s &= \hat{z}\times E_s\end{aligned}$$

The eigenvalue problem (1.22) becomes

$$\mathcal{L}E_s + \beta^2\mathcal{B}E_s = 0 \tag{1.24}$$

We would like to prove that the set of all the solutions E of (1.24) and (1.23) is a complete basis in which any general solution of our initial problem (1.17) and (1.18) can be expanded. In this purpose, we are going to compute the adjoint $\mathcal{L}^*$ of $\mathcal{L}$ hoping that $\mathcal{L}$ is self-adjoint. Indeed, if $\mathcal{L}$ is self-adjoint, the existence of an orthonormal basis is straightforward (as in the previous sections). A natural inner product in the cross-section S is

$$< E_s, A_s >= \int_S E_s.A_s\ dxdy$$

$\mathcal{L}^*$ is then defined as follows

$$< A_s, \mathcal{L}E_s >=< \mathcal{L}^*A_s, E_s >$$

for any A_s, E_s such that $E_s\times n = 0$ and A_s satisfies the adjoint boundary conditions (which have to be determined).

The derivation of the adjoint would require several additional pages and the reader is asked to refer to the article [2] for the calculation. Finally, we get:

$$\mathcal{L}^* A_s = -\epsilon \hat{z} \times \nabla_s \times (\epsilon^{-1} \nabla_s \times A_s) + \hat{z} \times \nabla_s (\mu^{-1} \nabla_s . \mu A_s) + k^2 \hat{z} \times A_s \quad (1.25)$$

and the adjoint boundary conditions are

$$n \times [\hat{z} \times A_s] = 0$$
$$[\nabla_s \times A_s] \times n = 0$$

Hence, $\mathcal{L}$ is not in general self-adjoint but $\mathcal{L}^*$ has many similarities with $\mathcal{L}$ and in fact if we interchange the role of ϵ and $-\mu$ in the expression of $\mathcal{L}^*$ (equation (1.25)), we get $\mathcal{L}$. But the transformation

$$\begin{array}{ccc} \epsilon & \longleftrightarrow & -\mu \\ E & \longleftrightarrow & H \end{array}$$

is also a transformation which let invariant Maxwell's equations and which is known as the duality principle (refer to [8] for more details)

The duality principle applied to (1.22) implies that H satisfies:

$$\epsilon \hat{z} \times \nabla_s \times (\epsilon^{-1} \nabla_s \times H_s) - \hat{z} \times \nabla_s (\mu^{-1} \nabla_s . \mu H_s) - (k^2 - \beta^2) \hat{z} \times H_s = 0$$

and hence, since one can easily prove that $\mathcal{B}^* = -\mathcal{B}$,

$$\mathcal{L}^* H_s + \beta^2 \mathcal{B}^* H_s = 0 \quad (1.26)$$

When E is an eigenmode with eigenvalue β^2, H is solution of the adjoint equation (1.26). However, we have to be careful with the boundary conditions. The adjoint conditions have to be fulfilled and we have to check that

$$n \times [\hat{z} \times H_s] = 0 \text{ and } [\nabla_s \times H_s] \times n = 0 \text{ on } \partial S$$

But a quick look at section 1.2, equations (1.8) and (1.9) tells us that these boundary conditions on H are naturally induced by the hard wall boundary conditions:

$$n \times [\hat{z} \times H_s] = H_n = 0 \; ; \; [\nabla_s \times H_s] \times n = [\nabla \times H \, . \, \hat{z}] \; t = 0$$

The way how H is related to the adjoint problem looks miraculous. H exactly plays the role of the adjoint variable of E. There must be some physical interpretation of this result but I did not find it in the literature I read.

Finally, the eigenmode (E, H, β^2) satisfies

$$\mathcal{L}E_s + \beta^2 \mathcal{B}E_s = 0 \tag{1.27}$$
$$\mathcal{L}^* H_s + \beta^2 \mathcal{B}^* H_s = 0 \tag{1.28}$$

The computation of the adjoint was worth doing because we can now use theorems of spectral theory to prove the existence of eigenmodes and eigenmode decomposition (refer for example to [10]). Since $\mathcal{B}^*\mathcal{B} = Id$, (1.27) implies :

$$\mathcal{D}E_s + \beta^2 E_s = 0 \tag{1.29}$$

where

$$D = \mathcal{B}^*\mathcal{L} \tag{1.30}$$

As in the previous sections, the hard wall boundary conditions imply that the β^2 solutions of (1.29) form a discrete set (This result I am not able to prove but which is commonly assumed may be related to the compactness of the operator $\mathcal{D}^{-1}$).

If, due to possible symetries, $\mathcal{D}$ is self-adjoint ($D^* = D$), β^2 is real and an eigenmode expansion is possible : Any E_s satisfying $E_s \times n = 0$ on ∂S can be expanded as a linear combination of the eigenmodes $E_{s,n}$ i.e.

$$E_s = \sum_{n \geq 1} a_n E_{s,n}$$

Several waveguide symetries give a self-adjoint operator $\mathcal{D}$. In section 1.2, ϵ was constant and we ended up with a self-adjoint operator. The case of a cylindrical symmetry (as in an optical fibre) also leads to a self-adjoint problem (see [3] for more details). For the planarly layered medium (section 1.3) the configuration was invariant in the $\hat{z}$ direction and a new scalar product made the equations self-adjoint. This last case may suggest that we should introduce an other inner product which would make $\mathcal{D}$ self-adjoint.

However, the appearance of complex modes in a closed, lossless and uniform waveguide which has been observed in practice (see the article of Rozzi [9]) shows us that, in general, $\mathcal{D}$ is not self-adjoint for any scalar product. Nevertheless, eigenmodes expansion with complex eigenvalues β^2 seem still valid (as said in the reference book of Collin [4]) and are often used by physicis. It could by justified by proving that $\mathcal{D}$ is normal ($\mathcal{D}^*\mathcal{D} = \mathcal{D}\mathcal{D}^*$) but this question is beyond the scope of this dissertation. However, these complex modes have very particular properties (they do not carry time-averaged energy, they can only appear in conjugated pairs) and neglecting

them in the modal expansions should not affect much the accuracy of the solution. That's why we will from now on assume that β^2 is real.

Equations (1.27) and (1.28) are very useful to derive orthogonality relations between the eigenmodes. Let (E_m, H_m, β_m), (E_n, H_n, β_n) be two eigenmodes which satisfy (1.27) and (1.28).

$$\mathcal{L}E_{s,m} + \beta_m^2 \mathcal{B}E_{s,m} = 0$$

implies, after multiplying it by $< ., H_{s,n} >$ that

$$< \mathcal{L}E_{s,m}, H_{s,n} > + \beta_m^2 < \mathcal{B}E_{s,m}, H_{s,n} >= 0$$

and then

$$< E_{s,m}, \mathcal{L}^* H_{s,n} > + \beta_m^2 < \mathcal{B}E_{s,m}, H_{s,n} >= 0 \tag{1.31}$$

After using (1.28), (1.31) becomes

$$< E_{s,m}, -\beta_n^2 \mathcal{B}^* H_{s,n} > + \beta_m^2 < \mathcal{B}E_{s,m}, H_{s,n} >= 0$$

which is equivalent to

$$(\beta_m^2 - \beta_n^2) < \mathcal{B}E_{s,m}, H_{s,n} >= 0$$

Hence, if $\beta_m^2 \neq \beta_n^2$, we get

$$< \mathcal{B}E_{s,m}, H_{s,n} >= \int_S E_{s,m} \times H_{s,n}.\hat{z}\ dxdy = 0$$

or, after normalisation of the modes,

$$\int_S E_{s,m} \times H_{s,n}.\hat{z}\ dxdy = \delta_{m,n} \tag{1.32}$$

In a similar way as we did in section 1.2, once the eigenmodes have been calculated, one can prove that any solution (E, H) of the initial problem (1.17) and (1.18) can be expanded in the form

$$E(x,y,z) = \sum_{n \geq 1} a_n e^{i\beta_n z} e_n + a_{-n} e^{-i\beta_n z} e_{-n}$$
$$H(x,y,z) = \sum_{n \geq 1} a_n e^{i\beta_n z} h_n + a_{-n} e^{-i\beta_n z} h_{-n}$$

where $\begin{cases} e_n = e_{n,s} + e_{n,z}\hat{z} \\ e_{-n} = e_{n,s} - e_{n,z}\hat{z} \end{cases}$ with $\begin{cases} e_{n,s} = E_{n,s} \\ e_{n,z} = \frac{1}{i\epsilon\beta_n} \nabla_s.(\epsilon E_{n,s}) \end{cases}$ from (1.23)

$$\text{and} \qquad h_n = \tfrac{1}{i\omega\mu}\nabla \times e_n,\ h_{-n} = \tfrac{1}{i\omega\mu}\nabla \times e_{-n}$$

The orthogonality relationships between the modes are the following:

$$\int_{\Gamma_l} e_i \times h_j.\hat{z}\ ds = \delta_{i,j} \text{ for any } i,j > 0 \tag{1.33}$$

One can easily see that in fact, these integrals only depend on the tangential part of the vectors, i.e.:

$$\int_{\Gamma_l} e_i \times h_j.\hat{z}\ ds = \int_{\Gamma_l} e_{t,i} \times h_{t,j}.\hat{z}\ ds \tag{1.34}$$

where $e_{t,i}$ and $h_{t,i}$ are respectively the tangential part of e_i and h_i.

Then we can retrieve the orthogonal relations for any i and j (not only the positive ones) by using the following identities that link forward and backward propagating modes:

$$\left\{ \begin{array}{rcl} e_{-j} & = & e_{t,j} - e_{zj}\hat{z} \\ h_{-j} & = & -h_{t,j} + h_{zj}\hat{z} \end{array} \right.$$

where $e_{t,j}$ and $h_{t,j}$ are the tangential parts of e_j and h_j while e_{zj} and h_{zj} are the normal components of e_j and h_j.

These orthogonality relations can be considered as the main result of this section and will be extensively used in the next sections.

This expansion is possible within a $\hat{z}$-uniform waveguide but in the next section we will see how such expansions can still be very useful for a more general type of waveguide.

In our optimisation problem, we have to compute the power remaining in the mth mode P_m. As we are going to see, the normalization (1.32) induces a very simple relation between P_m and the coefficients a_n of the modal expansion.

The average flow of energy P crossing the surface S is given by

$$P = \frac{1}{2} Re \int_S E \times \overline{H}\,.\,ds$$

$\overline{H}$ is the conjugate of H and $E \times \overline{H}$ is called the Poynting vector. Hence, by using the orthogonality of the modes

$$\begin{array}{rcl} P & = & \frac{1}{2} Re \int_S \left(\sum_{n\geq 1} a_n e^{i\beta_n z} e_n + a_{-n} e^{-i\beta_n z} e_{-n} \right) \times \left(\sum_{n\geq 1} \overline{a_n} e^{-i\beta_n z} h_n + \overline{a}_{-n} e^{i\beta_n z} h_{-n} \right) . z\ ds \\ & = & \frac{1}{2} \sum_{n\geq 1} Re \left(|a_n|^2 - a_n \overline{a}_{-n} e^{2i\beta_n z} + a_{-n}\overline{a}_n e^{-2i\beta_n z} - |a_{-n}|^2 \right) \end{array}$$

and, since $a_n\overline{a}_{-n}e^{2i\beta_n z} - a_{-n}\overline{a}_n e^{-2i\beta_n z}$ is purely imaginary,

$$P = \frac{1}{2}\sum_{n\geq 1} |a_n|^2 - |a_{-n}|^2 \tag{1.35}$$

The minus sign in front of $|a_{-n}|^2$ indicates that the power flow is in the negative $\hat{z}$ direction and it is coherent with the fact that the mode (e_{-n}, h_{-n}) propagates in the same direction ($-\beta_n < 0$ by definition).

If the modes are normalized as in (1.32) then $P_n = |a_n|^2/2$ and $P_{-n} = |a_{-n}^2|/2$ represent the power carried by the n^{th} mode in the $\hat{z}$ positive and negative directions.

1.5 Propagation in non z-uniform waveguides

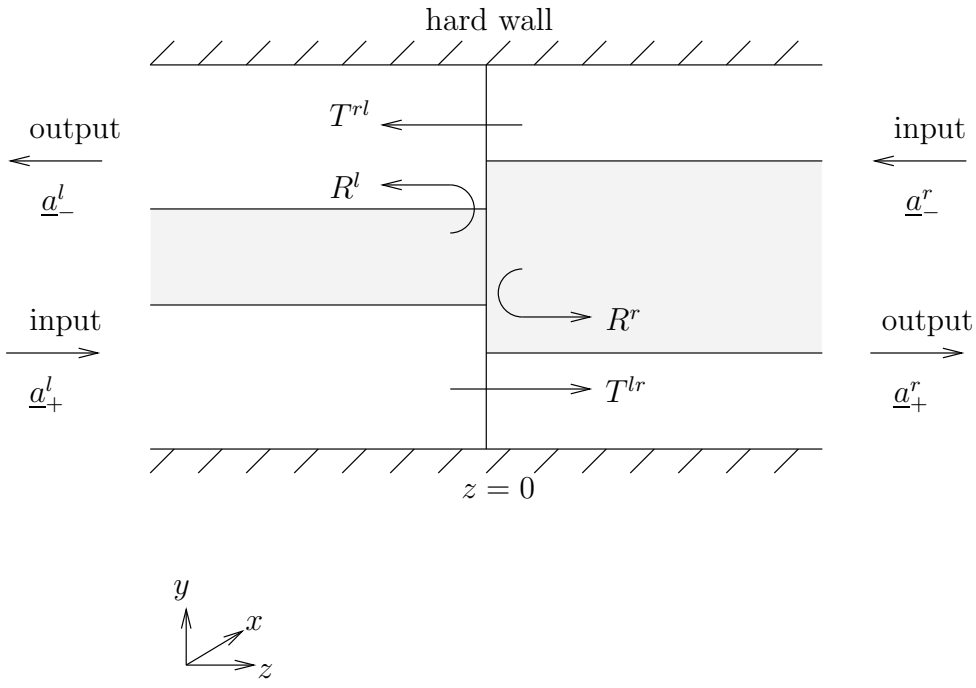


Figure 1.7: junction between two uniform waveguides

Let's consider two uniform waveguides with different cross-section put side by side as shown in Figure (1.7). On both sides, since the waveguides are z-invariant, the electromagnetic field can be expanded in an linear combination

of the eigenmodes and we have on the left-hand side

$$\begin{aligned} E &= \sum_{n\geq 1} a_n^l\ e^{i\beta_n^l z}\ e_n^l\ +\ a_{-n}^l\ e^{-i\beta_n^l z}\ e_{-n}^l \\ H &= \sum_{n\geq 1} a_n^l\ e^{i\beta_n^l z}\ h_n^l\ +\ a_{-n}^l\ e^{-i\beta_n^l z}\ h_{-n}^l \end{aligned}$$

and on the right-hand side

$$\begin{aligned} E &= \sum_{n\geq 1} a_n^r\ e^{i\beta_n^r z}\ e_n^r\ +\ a_{-n}^r\ e^{-i\beta_n^r z}\ e_{-n}^r \\ H &= \sum_{n\geq 1} a_n^r\ e^{i\beta_n^r z}\ h_n^r\ +\ a_{-n}^r\ e^{-i\beta_n^r z}\ h_{-n}^r \end{aligned}$$

At a discontinuity, the tangential components of E and H are continuous. For the sake of simplicity, we assume that the junction between the two waveguides is at $z = 0$. Hence,

$$\sum_{n\geq 1} a_n^l e_{n,s}^l + a_{-n}^l e_{-n,s}^l = \sum_{n\geq 1} a_n^r e_{n,s}^r + a_{-n}^r e_{-n,s}^r \tag{1.36}$$

$$\sum_{n\geq 1} a_n^l h_{n,s}^l + a_{-n}^l h_{-n,s}^l = \sum_{n\geq 1} a_n^r h_{n,s}^r + a_{-n}^r h_{-n,s}^r \tag{1.37}$$

After multiplying (1.36) by $< ., h_{n,s}^r >$ and (1.37) by $< e_{n,s}^r, . >$ and making use of the orthogonality relations, we get:

$$\sum_{n\geq 1} a_n^l < e_{n,s}^l, h_{n,s}^r > + a_{-n}^l < e_{-n,s}^l, h_{n,s}^r >= a_n^r + a_{-n}^r \tag{1.38}$$

$$\sum_{n\geq 1} a_n^l < e_{n,s}^r, h_{n,s}^l > + a_{-n}^l < e_{n,s}^r, h_{-n,s}^l >= a_n^r - a_{-n}^r \tag{1.39}$$

Assuming that the field is known on the left hand-side (i.e. a_n^l and a_{-n}^l are known), the equations above allow us to determine the field on the right-hand side.

In practice, the coefficients which are given are the input coefficients (a_n^l and a_{-n}^r) and the output coefficients (a_{-n}^l and a_n^r) are looked for. Some more algebra is needed to rewrite (1.38) and (1.39) as outputs in function of inputs:

$$\begin{aligned} \underline{a}_+^r &= T^{lr}\underline{a}_+^l + R^r\underline{a}_-^r \\ \underline{a}_-^l &= T^{rl}\underline{a}_-^r + R^l\underline{a}_+^l \end{aligned}$$

where $\underline{a}_+^l = (a_1^l, a_2^l, ...)$ and $\underline{a}_-^r = (a_{-1}^r, a_{-2}^r, ...)$ are the input vectors while $\underline{a}_+^r = (a_{-1}^r, a_2^r, ...)$ and $\underline{a}_-^l = (a_{-1}^l, a_{-2}^l, ...)$ are the outputs vectors. T^{lr} and T^{rl} are transmission operators from left to right and right to left. R^l and R^r are reflection operators (see Figure (1.7)).

In a uniform waveguide, eigenmodes constitute a basis and the coefficients $a_{\pm n}e^{\pm i\beta_n z}$ are the projection of (E, H) into each component $e_{\pm n}$. This basis is relevant as long as we are in one given uniform waveguide. When we consider another uniform waveguide, we have to change the basis and this is what the transmission and reflection operators allow us to do in projecting one basis into the other.

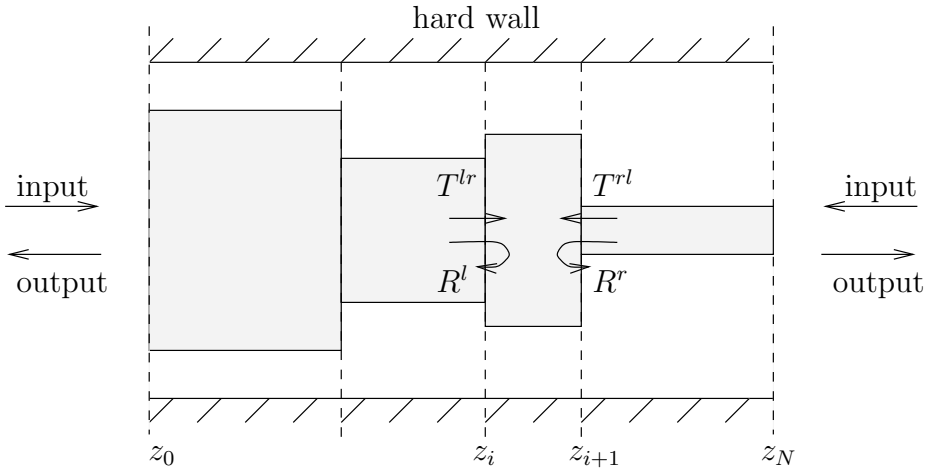


Figure 1.8: assembling uniform waveguides

Let's now consider a succession of z-uniform waveguides as shown in Figure (1.8). At any junction ($z = z_i$), it is possible to derive tranmission and reflection operators as we have just done. Between two junctions, coefficients are only modified by a phase shift of $\pm\beta_n z$. Hence, we are able to compute for any given input the resulting output of such a waveguide. This method can be numerically implemented and the result is *Fimmprop-3d* of *PhotonDesign* .

1.6 *Fimmprop-3d* : the Maxwell's equations solver of *PhotonDesign*

Fimmprop-3d is a computational software designed for the photonic industry. It uses an eigenmode expansion algorithm, where the optical properties of a structure are fully characterised by the local modes of the structure, and coupling matrices between different local modes.

The modal decomposition constitutes the discretisation of Maxwell's equations which is always needed when one wants to solve a problem numerically.

The advantage of this approach compared with others (like for example a finite difference method) is that the coefficients we get by this process are closely related to the modes of the structure and therefore have a strong physical content. It yields fast computation times and a better understanding of the physical phenomena involved.

Fimmprop-3d follows the procedure described in the previous section. Any optical device is reduced to a succession of $\hat{z}$-uniform waveguides. The modes of each waveguide is then computed. Of course, only a finite number of modes are considered. Nevertheless since we know that the power flow is the same in any coss-section S, it can be used as a test for accuracy in order to evaluate the number of modes needed.

Chapter 2

Computation of the Derivative using Adjoint Method

2.1 Introduction to the adjoint approach to optimisation

The principle of the adjoint approach is to formulate an adjoint equation which, once it has been solved, allows us to take into account the constraints coming from the original equation without having to solve it. In order to make this statement clearer, let's introduce some notation and see how one can solve an optimisation problem using the adjoint method. For more explanation, the reader is asked to refer to the article of Mike Giles [7] from which the rest of this section is highly inspired.

Let's consider an objective function $P(U, \epsilon)$ which is the function we want to optimise. ϵ is an optimisation variable and U is a physical quantity (in our case, the electric field). U is solution of a physical problem which depends on ϵ and is implicitly determined by a relation of the form $N(U(\epsilon), \epsilon) = 0$. Our optimisation problem is then:

$$\max_{\epsilon} P(U(\epsilon), \epsilon)$$

where U is given by

$$N(U(\epsilon), \epsilon) = 0 \tag{2.1}$$

Solving (2.1) requires computation (we have to solve the direct problem) and the aim of the adjoint method is to avoid that. After differentiating P,

we get:

$$\frac{dP}{d\epsilon} = \frac{\partial P}{\partial U}.\frac{dU}{d\epsilon} + \frac{\partial P}{\partial \epsilon} \tag{2.2}$$

and since U satisfies equation (2.1),

$$\frac{\partial N}{\partial U}.\frac{dU}{d\epsilon} + \frac{\partial N}{\partial \epsilon} = 0 \tag{2.3}$$

By defining

$$u = \frac{dU}{d\epsilon} \ ; \ L = \frac{\partial N}{\partial U} \ ; \ f = -\frac{\partial N}{\partial \epsilon}$$

equations (2.2) and (2.3) become

$$\frac{dP}{d\epsilon} = \frac{\partial P}{\partial U}.u + \frac{\partial P}{\partial \epsilon} \tag{2.4}$$

$$Lu = f \tag{2.5}$$

$\partial P/\partial U$ in (2.4) is a linear form applied to u. If we are in a Hilbert space H, the Riesz theorem tells us that there is a bijection between H and the linear forms on H and so, we can find g in H such that:

$$\frac{\partial P}{\partial U}.u = \langle g, u \rangle$$

where $\langle ., . \rangle$ denotes the scalar product of H. Hence, equation (2.4) can be rewritten as

$$\frac{dP}{d\epsilon} = \langle g, u \rangle + \frac{\partial P}{\partial \epsilon}$$

If we now solve the adjoint equation

$$L^* v = g \tag{2.6}$$

where L^* is the adjoint operator of L (i.e. $\langle Lv, u \rangle = \langle v, L^* u \rangle$), we get

$$\begin{aligned} \frac{dP}{d\epsilon} &= \langle L^* v, u \rangle + \frac{\partial P}{\partial \epsilon} \\ &= \langle v, Lu \rangle + \frac{\partial P}{\partial \epsilon} \end{aligned}$$

and finally,

$$\frac{dP}{d\epsilon} = \langle v, f \rangle + \frac{\partial P}{\partial \epsilon} \tag{2.7}$$

Note that u doesn't appear in the expression (2.7) and, as we said earlier, this is what makes the adjoint method interesting : there is no need to solve

the original equation for u in order to get the derivative. However, we have to compute v from equation (2.6) and we expect this equation to be as difficult to solve as the original one. It may seem that the method doesn't bring any simplification but when one has to compute $dP/d\epsilon$ for many different source terms f then the method becomes very interesting : the adjoint equation has to be solved once and we get $dP/d\epsilon$ for any f by (2.7).

We can relate this result to the computation of a Green function. The Green function G of an operator L comes from a particular choice of g as a Dirac function.

$$L^*G = \delta(x - x')$$

and we have

$$\begin{aligned} u(x) &= \langle \delta(x - x'), u(x') \rangle \\ &= \langle L^*G, u \rangle \\ &= \langle G, f \rangle \end{aligned}$$

There is then the same trade-off to do : If the original equation $Lu = f$ has to be solved once, it is probably better to solve it directly but if it has to be solved for many f, it may be worth to compute the Green function.

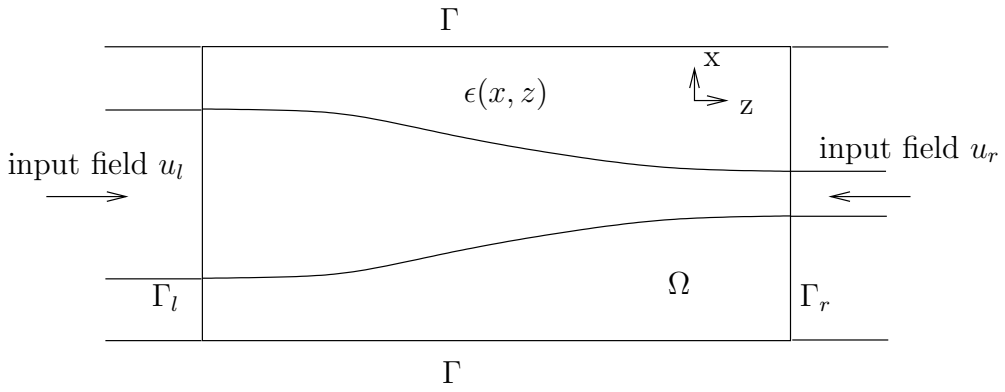
So far we have implicitely assumed that the original problem has homogeneous boundary conditions and the objective function consists only of an inner product over the whole domain and not a boundary integral. In the taper case, the objective function (the power P_m remaining in the mth mode) is computed on the boundary and we have to consider the general form of the adjoint identity:

$$\langle v, Lu \rangle + \langle C^*v, Bu \rangle_{\partial\Omega} = \langle L^*v, u \rangle + \langle B^*v, Cu \rangle_{\partial\Omega}$$

for all functions u,v where $\langle ., . \rangle_{\partial\Omega}$ is an inner-product on the boundaries. B and C are both boundary operators given in the definition of the original problem. We will explain in the next section in the 2D case of a taper how this identity is used to derive the derivative of P_m.

2.2 Computation of the derivative in the 2D case

2.2.1 Presentation of the 2D problem



In 2D, we assume that $\partial/\partial y = 0$. Maxwell's equations can be reduced to two independent scalar equations (see [8]) for the T.E. and T.M. field. The T.E. field is obtained by setting

$$H_y = E_x = E_z = 0$$

then we have

$$\begin{cases} \dfrac{\partial H_x}{\partial z} - \dfrac{\partial H_z}{\partial x} &= -i\omega\epsilon E_y \\ \dfrac{\partial E_y}{\partial z} &= -i\omega\mu H_x \\ \dfrac{\partial E_y}{\partial x} &= i\omega\mu H_z \end{cases}$$

which leads to the Helmotz equation for E_y

$$\nabla^2 E_y + \epsilon\mu\omega^2 E_y = 0$$

In this section, we will only consider the T.E. field. The T.M. field is obtained by setting $E_y = H_x = H_z = 0$ and leads to a slightly different equation. On Γ, the boundary conditions are those of a reflective wall i.e.

$$E_y = 0$$

After renaming E_y by u, we end up with the following equations:

$$\begin{aligned} \nabla^2 u + \epsilon\mu\omega^2 u &= 0 \\ u_{|\Gamma} &= 0 \end{aligned}$$

On the left and right-hand side of Ω where the waveguide is uniform, it is possible to expand the field u in function of the eigenmodes of the waveguide.

$$u(x,z) = \sum_{k=1}^{\infty} \left(c_k^l e^{i\beta_k^l z} + c_{-k}^l e^{-i\beta_k^l z} \right) u_k^l(x) \quad \text{on the left-hand side} \tag{2.8}$$

$$u(x,z) = \sum_{k=1}^{\infty} \left(c_k^r e^{i\beta_k^r z} + c_{-k}^r e^{-i\beta_k^r z} \right) u_k^r(x) \quad \text{on the right-hand side} \tag{2.9}$$

The modes u_k^l are orthonormal with respect to the inner product (2.10)

$$\langle u, v \rangle_{\Gamma_l} = \int_{\Gamma_l} u(x) v(x) \, dx \tag{2.10}$$

Hence,

$$\left\langle u_i^l, u_j^l \right\rangle_{\Gamma_l} = \delta_{i,j} \tag{2.11}$$

On the right-hand side, the modes are also orthonormal and we have:

$$\left\langle u_i^r, u_j^r \right\rangle_{\Gamma_l} = \int_{\Gamma_r} u_i^r(x) u_j^r(x) \, dx = \delta_{i,j}$$

The input field u_I^l on the left-hand side Γ_l is given by the coefficients c_k^l ($k > 0$) which correspond to z positive direction of propagation ($\beta_k^l > 0$).

$$u_I^l(x,z) = \sum_{k=1}^{\infty} c_k^l e^{i\beta_k^l z} u_k^l(x) \tag{2.12}$$

We are now going to work out what the boundary conditions on Γ_l are. From (2.8), after using the orthogonality relations (2.11), we get:

$$\left\langle u, u_k^l \right\rangle_{\Gamma_l} = c_k^l e^{i\beta_k^l z} + c_{-k}^l e^{-i\beta_k^l z} \tag{2.13}$$

and from (2.12),

$$\left\langle u_I^l, u_k^l \right\rangle_{\Gamma_l} = c_k^l e^{i\beta_k^l z} \tag{2.14}$$

Hence,

$$\begin{aligned}
\frac{\partial u}{\partial z} &= \sum_{k=1}^{\infty} \left(i\beta_k^l c_k^l e^{i\beta_k^l z} - i\beta_k^l c_{-k}^l e^{-i\beta_k^l z} \right) u_k^l \\
&= \sum_{k=1}^{\infty} \left(i\beta_k^l \left\langle u_I^l, u_k^l \right\rangle_{\Gamma_l} - i\beta_k^l \left(\left\langle u, u_k^l \right\rangle_{\Gamma_l} - \left\langle u_I^l, u_k^l \right\rangle_{\Gamma_l} \right) \right) u_k^l \quad \text{from (2.12) and (2.14)} \\
&= \sum_{k=1}^{\infty} \left(-i\beta_k^l \left\langle u, u_k^l \right\rangle_{\Gamma_l} + 2i\beta_k^l \left\langle u_I^l, u_k^l \right\rangle_{\Gamma_l} \right) u_k^l
\end{aligned}$$

A similar expression can be found for the right-hand side. We introduce the operators B^l and B^r defined as follows

$$B^{(s)}(u) = \sum_{k=1}^{\infty} \beta_k^{(s)} \left\langle u_k^{(s)}, u \right\rangle_{\Gamma_{(s)}} u_k^{(s)} \quad \text{with } (s) = (l) \text{ or } (r)$$

and the boundary conditions can be written as

$$\begin{cases} \partial_n u - iB^l(u) &= -2iB^l(u_l) \quad \text{on } \Gamma_l \\ \partial_n u - iB^r(u) &= -2iB^r(u_r) \quad \text{on } \Gamma_r \end{cases}$$

Finally, we end up with u being solution of

$$\nabla^2 u + \epsilon\mu\omega^2 u = 0 \tag{2.15}$$

$$\begin{cases} u = 0 & \text{on } \Gamma \\ \partial_n u - iB^l(u) &= -2iB^l(u_l) & \text{on } \Gamma_l, \\ \partial_n u - iB^r(u) &= -2iB^r(u_r) & \text{on } \Gamma_r, \end{cases}$$

where $\partial_n u$ is the normal derivative of u.

In the taper problem, given the input fields, we are interested in the power P_m remaining in the mth mode. It can be shown that:

$$P_m = \beta_m^r |\, \langle u, u_m^r \rangle_{\Gamma_r} \,|^2$$

In the next sections, we derive two methods to evaluate the derivative of P_m with respect to ϵ.

2.2.2 Lagrangian Formulation

In the Lagrangian formulation, we consider the complex Lagrangian :

$$L(u, \lambda, \epsilon) = P_m(u, \epsilon) + Re\left(\int_\Omega \lambda(\nabla^2 u + \epsilon\mu\omega^2 u)\, dv \right) \tag{2.16}$$

The lagrangian variable λ is a field over Ω that we are going to determine. When the potential u satisfies the equation (2.15), it is completely determined by the electric permittivity distribution ϵ which remains as our only independent variable and we have from (2.16):

$$L(u(\epsilon), \lambda, \epsilon) = P_m(u(\epsilon), \epsilon)$$

because $\nabla^2 u + \epsilon\mu\omega^2 u = 0$ for any solution u. hence

$$\frac{\partial L(u(\epsilon), \lambda, \epsilon)}{\partial \epsilon} = \frac{dP_m}{d\epsilon}.$$

ϵ does not appear explicitly in P_m ; it only appears implicitly through u and therefore the derivative $dP_m/d\epsilon$ can not be computed directly. We use the freedom we have on λ to overcome this difficulty.

Let's consider a small perturbation $\delta\epsilon$ which induces a perturbation δu of the field u. From (2.16), we get:

$$\begin{aligned}\delta L &= \frac{\partial P_m}{\partial u}\delta u + Re\left(\int_\Omega (\nabla^2 u + \epsilon\mu\omega^2 u)\delta\lambda\, dv\right) \\ &\quad + Re\left(\int_\Omega \lambda(\nabla^2\delta u + \epsilon\mu\omega^2\delta u)\, dv\right) + \omega^2\mu Re\left(\int_\Omega \lambda u\delta\epsilon\, dv\right)\end{aligned}$$

For any solution u, the term $Re\left(\int_\Omega (\nabla^2 u + \epsilon\mu\omega^2 u)\delta\lambda\right)$ vanishes.

We use the freedom we have on λ to get rid of the terms involving δu. So, we want to impose:

$$\frac{\partial P_m}{\partial u}\delta u + Re\left(\int_\Omega \lambda(\nabla^2\delta u + \epsilon\mu\omega^2\delta u)\, dv\right) = 0 \tag{2.17}$$

Let's first compute $\frac{\partial P_m}{\partial u}\delta u$,

$$P_m(u) = \beta_m^r |\left\langle u, u_m^r\right\rangle_{\Gamma_r}|^2$$

Hence,

$$\frac{\partial P_m}{\partial u}\delta u = 2\beta_m^r Re\left(\overline{\left\langle u, u_m^r\right\rangle}_{\Gamma_r}\left\langle u_m^r, \delta u\right\rangle_{\Gamma_r}\right) \tag{2.18}$$

By using Green's theorem, we get :

$$\int_\Omega \lambda(\nabla^2\delta u + \epsilon\mu\omega^2\delta u)\, dv = \int_\Omega (\nabla^2\lambda + \epsilon\mu\omega^2\lambda)\delta u\, dv + \int_{\partial\Omega} \lambda\,\partial_n\delta u\, ds - \int_{\partial\Omega}\partial_n\lambda\,\delta u\, ds \tag{2.19}$$

Since $u + \delta u$ is solution of the initial problem (2.15) where ϵ is replaced by $\epsilon + \delta\epsilon$, δu satisfies the following homogeneous boundary conditions:

$$\left\{\begin{array}{rcll} \delta u & = & 0 & \text{on } \Gamma \\ \partial_n\delta u & = & iB^l(\delta u) & \text{on } \Gamma_l, \\ \partial_n\delta u & = & iB^r(\delta u) & \text{on } \Gamma_r, \end{array}\right.$$

and the first line integral above can be rewritten:

$$\int_{\partial\Omega} \lambda\, \partial_n \delta u\, ds = \int_{\Gamma} \lambda\, \partial_n \delta u\, ds - \left\langle \lambda, iB^l(\delta u) \right\rangle_{\Gamma_l} + \left\langle \lambda, iB^r(\delta u) \right\rangle_{\Gamma_r} \tag{2.20}$$

where we use the scalar product defined in (2.10) but since we have

$$\begin{aligned} \left\langle \lambda, iB^l(\delta u) \right\rangle_{\Gamma_l} &= \left\langle \lambda, i \textstyle\sum_{k\geq 1} \beta_k^l \left\langle u_k^l, \delta u \right\rangle_{\Gamma_l} u_k^l \right\rangle_{\Gamma_l} \\ &= \textstyle\sum_{k\geq 1} i\beta_k^l \left\langle u_k^l, \lambda \right\rangle_{\Gamma_l} \left\langle u_k^l, \delta u \right\rangle_{\Gamma_l} \\ &= i \left\langle \textstyle\sum_{k\geq 1} \beta_k^l \left\langle u_k^l, \lambda \right\rangle_{\Gamma_l} u_k^l, \delta u \right\rangle_{\Gamma_l} \\ &= \left\langle iB^l(\lambda), \delta u \right\rangle_{\Gamma_l} \end{aligned}$$

and a similar expression for $\langle \lambda, iB^r(\delta u)\rangle_{\Gamma_r}$, (2.20) is equivalent to

$$\int_{\partial\Omega} \lambda \partial_n \delta u\, ds = \int_{\Gamma} \lambda \partial_n \delta u\, ds - \left\langle iB^l(\lambda), \delta u \right\rangle_{\Gamma_l} + \left\langle iB^r(\lambda), \delta u \right\rangle_{\Gamma_r} \tag{2.21}$$

If we set :

$$\nabla^2 \lambda + \epsilon\mu\omega^2 \lambda = 0 \text{ on } \Omega$$
$$\lambda = 0 \text{ on } \Gamma,$$

we get rid of the integral over Ω in (2.19) and of the integral over Γ in (2.21). Equation (2.19) then becomes:

$$\int_{\Omega} \lambda(\nabla^2 \delta u + \epsilon\mu\omega^2 \delta u)\, dv = \left\langle \partial_n \lambda - iB^l(\lambda), \delta u \right\rangle_{\Gamma_l} - \left\langle \partial_n \lambda - iB^r(\lambda), \delta u \right\rangle_{\Gamma_r} \tag{2.22}$$

Hence, the Lagrangian conditions we want to impose and which are given by (2.17) become:

$$Re\left[\left\langle \partial_n \lambda - iB^l(\lambda),\ \delta u \right\rangle_{\Gamma_l} - \left\langle \partial_n \lambda - iB^r(\lambda) - 2\beta_m^r \left\langle \overline{u, u_m^r} \right\rangle_{\Gamma_r} u_m^r,\ \delta u \right\rangle_{\Gamma_r} \right] = 0 \tag{2.23}$$

Since we want this equality to be satisfied for any δu, we have to impose

$$\begin{aligned} \partial_n \lambda - iB^l(\lambda) &= 0 \text{ on } \Gamma_l \\ \partial_n \lambda - iB^r(\lambda) &= 2\beta_m^r \left\langle \overline{u, u_m^r} \right\rangle_{\Gamma_r} u_m^r \text{ on } \Gamma_r \end{aligned}$$

And, finally, the lagrangian variable λ satisfies the following equations:

$$\nabla^2 \lambda + \epsilon\mu\omega^2 \lambda = 0 \ \text{ on } \Omega$$

$$\begin{cases} \lambda &= 0 & \text{on } \Gamma \\ \partial_n \lambda &= iB^l(\lambda) & \text{on } \Gamma_l, \\ \partial_n \lambda &= iB^r(\lambda) + 2\beta_m^r \left\langle \overline{u, u_m^r} \right\rangle_{\Gamma_r} u_m^r & \text{on } \Gamma_r, \end{cases}$$

These equations have the same form of the original equation ; only the inhomogeneous terms in the boundary conditions are different. Once λ has been computed, the derivative of P_m can be expressed as

$$\delta L = \frac{dP_m}{d\epsilon}\delta\epsilon = \omega^2 \mu Re \left[\int_\Omega \lambda u \delta\epsilon \; dv \right]$$

2.2.3 Adjoint Formulation

In this section we follow the adjoint method derived in Mike Giles' article [7]. In the same article, it is shown how the adjoint and the Lagrangian methods are closely related. Nevertheless the adjoint method provides a clearer theoritical framework which makes it more generic.

Let $\delta\epsilon$ be again a small perturbation which induces δu. The initial set of equations (2.15) implies that :

$$L(\delta u) \equiv \nabla^2 \delta u + \epsilon\mu\omega^2 \delta u = -\omega^2 \mu \delta\epsilon \, u \tag{2.24}$$

$$\begin{aligned} \delta u &= 0 \text{ on } \Gamma \\ \partial_n \delta u &= iB(\delta u) \text{ on } \Gamma_s \end{aligned}$$

where

$$\Gamma_s = \Gamma_l \cup \Gamma_r \;\text{ and } \begin{cases} B(v) &= B^l(v) \text{ on } \Gamma_l \\ &= B^r(v) \text{ on } \Gamma_r \end{cases}$$

We are now led to find the adjoint operators of L, A and C such that

$$\langle w, Lw' \rangle_\Omega + \langle C^* w, Aw' \rangle_{\partial\Omega} = \langle L^* w, w' \rangle_\Omega + \langle A^* w, Cw' \rangle_{\partial\Omega} \quad \forall (w, w')$$

A and C are both boundary operators resulting from the definition of our problem and we have :

$$A(w) = -\partial_n w + iB(w) \text{ on } \Gamma_s$$

so $A(\delta u) = 0$ for any perturbation δu.

$C = C^* = I$ because our linearised objective function doesn't involve any derivative of u.

Calculations exactly similar to those done in the previous section give us that for any w and w' such that $w_{|\Gamma} = w'_{|\Gamma} = 0$,

$$\langle w', Lw\rangle_\Omega + \langle w', -\partial_n w + iBw\rangle_{\Gamma_s} = \langle Lw', w\rangle_\Omega + \langle -\partial_n w' + iBw', w\rangle_{\Gamma_s} \quad (2.25)$$

Hence, since the adjoint operators are defined by:

$$\langle w', Lw\rangle_\Omega + \langle w', Aw\rangle_{\Gamma_s} = \langle L^* w', w\rangle_\Omega + \langle A^* w', w\rangle_{\Gamma_s} \ , \quad (2.26)$$

we deduce from (2.25):

$$L^* = L \quad \text{and} \quad A^* = A$$

The linearised function g of section 2.1 is given by (2.18):

$$g = 2\beta_m^r \langle \overline{u, u_m^r}\rangle_{\Gamma_r} u_m^r,$$

which implies that

$$\frac{\partial P_m}{\partial u}\delta u = Re\, \langle g, \delta u\rangle_{\Gamma_r}$$

Let's now impose:

$$L^* v = 0 \ \text{ and } \ A^* v = g \quad (2.27)$$

Then, for any pertubation δu, if we apply (2.26) with $w = \delta u$ and $w' = v$, we get:

$$\langle g, u\rangle_{\Gamma_s} = \langle v, L(\delta u)\rangle_\Omega$$

and, since $L(\delta u) = -\omega^2 \mu \delta\epsilon\, u$,

$$\frac{dP_m}{d\epsilon}\delta\epsilon = Re\, \langle g, u\rangle_{\Gamma_r} = -\omega^2 \mu Re\left(\int_\Omega v u \delta\epsilon \, d\mathrm{v}\right) \quad (2.28)$$

Furthermore, the conditions (2.27) are equivalent to:

$$\nabla^2 v + \epsilon\mu\omega^2 v = 0 \ \text{ on } \Omega$$

$$\begin{cases} v &= 0 & \text{on } \Gamma \\ \partial_n v &= iB^l(v) & \text{on } \Gamma_l, \\ \partial_n v &= iB^r(v) - 2\beta_m^r \langle \overline{u, u_m^r}\rangle_{\Gamma_r} u_m^r & \text{on } \Gamma_r, \end{cases}$$

which is again a well posed problem of the same form of the original problem.

If we want to find exactly the same equations as in the Lagrangian case, we have to set $-\partial_n v + iBv = -g$ in (2.27).

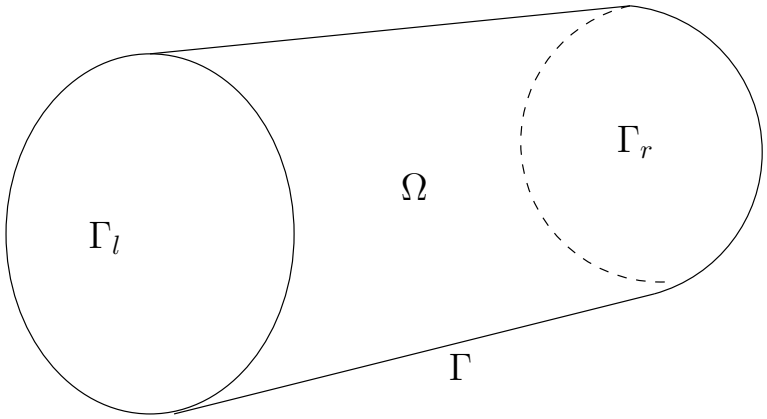
Before leaving this section and starting the 3d case, I would like to compare the adjoint method with a classical finite difference method and see what advantages we get.

Finite difference method	Adjoint method
• Computation of $u(\epsilon)$	• Computation of $u(\epsilon)$
• Computation of $u(\epsilon + \delta\epsilon)$	• Computation of $v(\epsilon)$
• $dPm/d\epsilon$ is given by:	• $dPm/d\epsilon$ is given by:
$\frac{dPm}{d\epsilon}\delta\epsilon = u(\epsilon + \delta\epsilon) - u(\epsilon)$	$\frac{dPm}{d\epsilon}\delta\epsilon = -\omega^2\mu Re\int_\Omega vu\delta\epsilon \ d\mathrm{v}$

The computations of $v(\epsilon)$ and $u(\epsilon + \delta\epsilon)$ take the same time and if we only consider one perturbation $\delta\epsilon$, the two methods are equivalent. However if we have more than one perturbation $\delta\epsilon$, the computation of $v(\epsilon)$ is done once and then $dP_m/d\epsilon$ is retrieved by (2.28). Let's say we have N parameters with respect to which we want to compute the derivative. With the finite different method we have to solve $N + 1$ times Maxwell's equation whereas with the adjoint method we have to solve them twice. The gain in computation time is then obvious.

2.3 Computation of the derivative in the 3D case

2.3.1 Equations and boundary conditions



In free space, Maxwell's equations for periodic phenomena in the absence of sources and current are

$$\begin{array}{rclcrcll} \nabla \times E - i\omega\mu H & = & 0, & & \nabla.\epsilon E & = & 0 & \\ \nabla \times H + i\omega\epsilon E & = & 0, & & \nabla.H & = & 0 & (\mu \text{ is constant}) \end{array}$$

From the equations of the first column, it can be deduced:

$$\nabla \times \nabla \times E - \omega^2 \mu\epsilon E = 0 \tag{2.29}$$

Since $\omega \neq 0$, once E is known, H can be simply retrieved by

$$H = \frac{1}{i\omega\mu} \nabla \times E \tag{2.30}$$

Equation (2.29) also suffices to retrieve the condition $\nabla.\epsilon E = 0$, since, for any vector A, $\nabla.\nabla \times A = 0$.

On Γ, we impose wall conditions. Hence, we have

$$n \times E = 0, \text{ on } \Gamma \tag{2.31}$$

which, in other words, means that the tangential component of E vanishes.

On $\Gamma_s = \Gamma_l \cup \Gamma_r$, one wonders which form of the boundary conditions we will need. To get a hint, we start the calculation of the adjoint operators and see what happens.

By multiplying by F and integrating over Ω, we get from (2.29) that, for any vector field F,

$$\int_\Omega \left(\nabla \times \nabla \times E - \omega^2 \epsilon \mu E\right) . F \; dv = 0$$

By using Stokes formulae, the first bit of the left-hand term can be rewritten as:

$$\int_\Omega \nabla \times \nabla \times E . F \; dv = \int_\Omega E . \nabla \times \nabla \times F \; dv + \int_{\partial\Omega} n . E \times \nabla \times F \; ds - \int_{\partial\Omega} n . F \times \nabla \times E \; ds \quad (2.32)$$

The surface integrals at the boundaries involve terms with the curl of E and F. On Γ, these integrals vanish because we require E and F to satisfy the boundary condition (2.31). On Γ_s, we would like to express $\nabla \times E$ as a linear function of E, just as we did for $\partial_n u$ in the two dimensional case.

On Γ_l, the electromagnetic field can be expanded in terms of the eigenmodes of the uniform waveguide of cross section Γ_l :

$$E = \sum_{k \geq 1} \alpha_k e^{i\beta_k^l z} e_k^l + \alpha_{-k} e^{-i\beta_k^l z} e_{-k}^l \quad (2.33)$$

$$H = \sum_{k \geq 1} \alpha_k e^{i\beta_k^l z} h_k^l + \alpha_{-k} e^{-i\beta_k^l z} h_{-k}^l \quad (2.34)$$

The inputs fields E_I^l and H_I^l are known and given by the forward propagating coefficient α_k, $k \geq 0$.

$$E_I^l = \sum_{k \geq 1} \alpha_k e^{i\beta_k^l z} e_k^l \quad (2.35)$$

$$H_I^l = \sum_{k \geq 1} \alpha_k e^{i\beta_k^l z} h_k^l \quad (2.36)$$

To handle these equations, we are going to use the following inner product which induces orthogonal relationships between the eigenmodes (cf section 1.4)

$$\langle u, v \rangle_{\Gamma_l} = \int_{\Gamma_l} u \times v . \hat{z} \; ds$$

By using the orthogonality of the modes, (2.33) and (2.35) give us:

$$\left\langle E - E_I^l, h_k^l \right\rangle_{\Gamma_l} = \alpha_{-k} e^{-i\beta_k^l z}$$

Hence, (2.34) becomes:

$$\begin{aligned} H &= H_I + \sum_{k\geq 1} \left\langle E - E_I^l, h_k^l \right\rangle_{\Gamma_l} h_{-k}^l \\ &= \sum_{k\geq 1} \left\langle E_I^l, h_k^l \right\rangle_{\Gamma_l} h_k^l - \sum_{k\geq 1} \left\langle E_I^l, h_k^l \right\rangle_{\Gamma_l} h_{-k}^l + \sum_{k\geq 1} \left\langle E, h_k^l \right\rangle_{\Gamma_l} h_{-k}^l \\ &= B^{l-}(E) + (B^{l+} - B^{l-})(E_I^l) \end{aligned}$$

where $B^{l\pm}$ is a linear operator defined as follows:

$$B^{l\pm}(E) = \sum_{k\geq 1} \left\langle E, h_k^l \right\rangle_{\Gamma_l} h_{\pm k}^l \tag{2.37}$$

On the right-hand side Γ_r, the input field is given by the backward propagating coefficients α_k, $k \leq 0$. We can compute H in the same way as before and we get:

$$H = B^{r+}(E) + (B^{r-} - B^{r+})(E_I^r) \text{ on } \Gamma_r$$

where B_I^r is defined is as in (2.37) except the integration is done over Γ_r and, of course, the eigenmodes are those of Γ_r.

Since $\nabla \times E = i\omega\mu H$, we have:

$$\nabla \times E = i\omega\mu B^{l-}(E) + i\omega\mu(B^{l+} - B^{l-})(E_I^l) \text{ on } \Gamma_l \tag{2.38}$$

$$\nabla \times E = i\omega\mu B^{r+}(E) + i\omega\mu(B^{r-} - B^{r+})(E_I^r) \text{ on } \Gamma_r \tag{2.39}$$

Finally, the electric field E is the solution of :

$$\nabla \times \nabla \times E - \omega^2 \epsilon\mu E = 0$$

$$\begin{cases} E \times n = 0 \text{ on } \Gamma \\ \nabla \times E = i\omega\mu B^{l-}(E) + i\omega\mu(B^{l+} - B^{l-})(E_I^l) \text{ on } \Gamma_l \\ \nabla \times E = i\omega\mu B^{r+}(E) + i\omega\mu(B^{r-} - B^{r+})(E_I^r) \text{ on } \Gamma_r \end{cases}$$

If we define L and A as follows:

$$\begin{cases} L(E) &= \nabla \times \nabla \times E - \omega^2 \epsilon\mu E \\ A(E) &= \nabla \times E - i\omega\mu B(E) \end{cases}$$

where, for convenience, we have defined B as $B = B^{l-}$ on Γ_l and $B = B^{r+}$ on Γ_r,we have that E satisfies

$$L(E) = 0$$

$$\begin{cases} E \times n &= 0 \text{ on } \Gamma \\ A(E) &= i\omega\mu(B^{l+} - B^{l-})(E_I^l) \text{ on } \Gamma_l \\ &\quad\; i\omega\mu(B^{r-} - B^{r+})(E_I^r) \text{ on } \Gamma_r \end{cases}$$

We are interested in the effect of a small perturbation $\delta\epsilon$ of the electric permitivity which induces a perturbation δE of the electric field. The set of equations above gives us that $\delta\epsilon$ and δE have to satisfy

$$\nabla \times \nabla \times \delta E - \omega^2 \epsilon\mu \, \delta E = \omega^2 \delta\epsilon\mu E \tag{2.40}$$

$$\begin{cases} \delta E \times n = 0 \text{ on } \Gamma \\ A(\delta E) = \nabla \times \delta E - i\omega\mu B(\delta E) = 0 \text{ on } \Gamma_s \end{cases}$$

2.3.2 Computation of the adjoint operators

For any field F satisfying $n \times F = 0$ on Γ, since the terms $n.E \times \nabla \times F = (n \times E).\nabla \times F$ and $n.F \times \nabla \times E = (n \times F).\nabla \times E$ vanish on Γ, (2.32) becomes:

$$\begin{aligned} \int_\Omega \nabla \times \nabla \times E.\, F \, dv &= \int_\Omega E.\nabla \times \nabla \times F \, dv + \\ &\quad \int_{\Gamma_s} n.E \times \nabla \times F \, ds - \int_{\Gamma_s} n.F \times \nabla \times E \, ds \end{aligned}$$

which can be finally written as

$$\int_\Omega \nabla \times \nabla \times E.F \, dv + \int_{\Gamma_s} n.F \times \nabla \times E \, ds = \int_\Omega E.\nabla \times \nabla \times F \, dv + \int_{\Gamma_s} n.E \times \nabla \times F \, ds \tag{2.41}$$

To make things clearer, let's have a look at the point we want to get to. The adjoint formulation consists of finding the adjoint operators L^* and A^* such that for any E and F satisfying $E \times n = F \times n = 0$ on Γ,

$$\langle L(E), F\rangle_\Omega + \langle A(E), F\rangle_{\Gamma_s} = \langle E, L^*(F)\rangle_\Omega + \langle E, A^*(F)\rangle_{\Gamma_s} \tag{2.42}$$

We will explain very soon how we will use L^* and A^* to find a simple expression of the derivative of P_m, the power remaining in the mth mode but first, let's compute them.

For this purpose we want to evaluate:

$$\int_{\Gamma_s} n.F \times (i\omega\mu B(E)) = -\left\langle F, i\omega\mu B^{l-}(E)\right\rangle_{\Gamma_l} + \left\langle F, i\omega\mu B^{r+}(E)\right\rangle_{\Gamma_r}$$

Let's first consider the integral over Γ_l, we have :

$$\begin{aligned} \left\langle F, i\omega\mu B^{l-}(E)\right\rangle_{\Gamma_l} &= i\omega\mu \left\langle F, \sum_{k\geq 1} \langle E, h_k^l\rangle_{\Gamma_l} h_{-k}^l \right\rangle_{\Gamma_l} \\ &= i\omega\mu \sum_{k\geq 1} \langle E, h_k^l\rangle_{\Gamma_l} \langle F, h_{-k}^l\rangle_{\Gamma_l} \end{aligned} \tag{2.43}$$

Since

$$\left\langle E, h_k^l \right\rangle_{\Gamma_l} = \int_{\Gamma_l} E \times h_k^l . \hat{z} ds = \int_{\Gamma_l} E_t \times h_{t,k} . \hat{z} ds = \int_{\Gamma_l} E_t \times -h_{t,-k} . \hat{z} ds = - \left\langle E, h_{-k}^l \right\rangle_{\Gamma_l}$$

and similarly

$$\left\langle F, h_{-k}^l \right\rangle_{\Gamma_l} = - \left\langle F, h_k^l \right\rangle_{\Gamma_l},$$

(2.43) becomes

$$\begin{aligned} \left\langle F, i\omega\mu B^{l-}(E) \right\rangle_{\Gamma_l} &= i\omega\mu \sum_{k\geq 1} \left\langle E, h_{-k}^l \right\rangle_{\Gamma_l} \left\langle F, h_k^l \right\rangle_{\Gamma_l} \\ &= i\omega\mu \left\langle E, \sum_{k\geq 1} \left\langle F, h_k^l \right\rangle_{\Gamma_l} h_{-k}^l \right\rangle_{\Gamma_l} \\ &= \left\langle E, i\omega\mu B^{l-}(F) \right\rangle_{\Gamma_l} \end{aligned} \tag{2.44}$$

Similarly, on the right-hand side Γ_r,

$$\begin{aligned} \left\langle F, i\omega\mu B^{r+}(E) \right\rangle_{\Gamma_r} &= i\omega\mu \left\langle F, \sum_{k\geq 1} \langle E, h_k^r \rangle_{\Gamma_r} h_k^r \right\rangle_{\Gamma_r} \\ &= i\omega\mu \left\langle E, \sum_{k\geq 1} \langle F, h_k^r \rangle_{\Gamma_r} h_k^r \right\rangle_{\Gamma_r} \\ &= i\omega\mu \left\langle E, i\omega\mu B^{r+}(F) \right\rangle_{\Gamma_r} \end{aligned} \tag{2.45}$$

Gathering (2.44) and (2.45), we get

$$\langle F, i\omega\mu B(E) \rangle_{\Gamma_s} = \langle E, i\omega\mu B(F) \rangle_{\Gamma_s}$$

which shows that B is self-adjoint with respect to the inner product $\langle ., . \rangle_{\Gamma_s}$.

After subtracting $\langle F, i\omega\mu B(E) \rangle_{\Gamma_s}$ on both sides, (2.41) becomes:

$$\begin{aligned} \langle \nabla \times \nabla \times E, F \rangle_\Omega &+ \langle F, \nabla \times E \rangle_{\Gamma_s} - \langle F, i\omega\mu B(E) \rangle_{\Gamma_s} \\ &= \langle E, \nabla \times \nabla \times F \rangle_\Omega + \langle E, \nabla \times F \rangle_{\Gamma_s} - \langle F, i\omega\mu B(E) \rangle_{\Gamma_s} \\ &= \langle E, \nabla \times \nabla \times F \rangle_\Omega + \langle E, \nabla \times F \rangle_{\Gamma_s} - \langle E, i\omega\mu B(F) \rangle_{\Gamma_s} \end{aligned}$$

and finally,

$$\langle L(E), F \rangle_\Omega + \langle A(E), F \rangle_{\Gamma_s} = \langle E, \nabla \times \nabla \times F \rangle_\Omega + \langle E, \nabla \times F - i\omega\mu B(F) \rangle_{\Gamma_s}$$

and the adjoint operators L^* and A^* are

$$L^*(F) = L(F) = \nabla \times \nabla \times F - \omega^2 \epsilon\mu F \tag{2.46}$$

$$A^*(F) = A(F) = \nabla \times F - i\omega\mu B(F) \tag{2.47}$$

2.3.3 Boundary conditions for the adjoint problem

We want to maximise P_m, the power remaining in the mth mode on the right-hand side, i.e.

$$P_m = |\langle E, h_m^r \rangle_{\Gamma_r}|^2 = \left\langle \overline{E, h_m^r} \right\rangle_{\Gamma_r} \langle E, h_m^r \rangle_{\Gamma_r} .$$

We have:

$$\begin{aligned} \frac{\partial P_m}{\partial \epsilon}\delta\epsilon &= 2Re \left\langle \overline{E, h_m^r} \right\rangle_{\Gamma_r} \langle \delta E , h_m^r \rangle_{\Gamma_r} \\ &= Re \left\langle \delta E , 2 \left\langle \overline{E, h_m^r} \right\rangle_{\Gamma_r} h_m^r \right\rangle_{\Gamma_r} \\ &= Re \langle \delta E , g \rangle_{\Gamma_r} \end{aligned}$$

where

$$g = 2 \left\langle \overline{E, h_m^r} \right\rangle_{\Gamma_r} h_m^r$$

If we impose that F satisfies

$$\begin{aligned} L^*(F) &= 0 \\ A^*(F) &= \nabla \times F - i\omega\mu B(F) = g \end{aligned}$$

then plugging it into (2.42) with $E = \delta E$ and since $A(\delta E) = 0$, we get:

$$\langle L(\delta E), F \rangle_\Omega = \langle \delta E , g \rangle_{\Gamma_r}$$

Hence,

$$\frac{\partial P_m}{\partial \epsilon}\delta\epsilon = Re \langle L(\delta E), F \rangle_\Omega$$

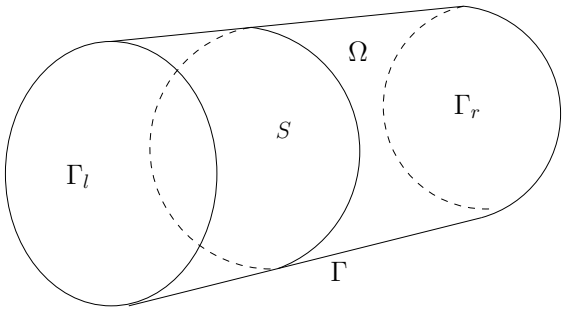
and finally, since $L(\delta E) = \omega^2 \mu \delta\epsilon \, E$ (from equation (2.40)),

$$\boxed{\frac{\partial P_m}{\partial \epsilon}\delta\epsilon = \omega^2 \mu Re \int_\Omega \delta\epsilon E.F dv} \tag{2.48}$$

This is one of the main results of this dissertation . The derivative $\partial P_m/\partial\epsilon$ is expressed in function of the adjoint variable F whose computation doesn't require the computation of the perturbed field. Moreover F is the solution of equations which have the same form of the original problem and we can use without modification the Maxwell's equation solver *Fimmprop-3d* to solve them. We just have to set the right boundary conditions.

2.3.4 Computation of the derivative P_m for a shape deformation

The previous result is valid under the assumption that the distribution of electric permitivity ϵ is continuous and then the field E and F are continuous too. We are now going to study the case where there is a discontinuity of the electric permittivity on a fixed surface S.



The jump conditions for E and F at the interface can be derived from Maxwell equations (refer for example to Jones [8]) and are the following:

$$[E \times n]^+_- = 0, \qquad [\epsilon E]^+_- = 0$$
$$[H \times n]^+_- = 0, \qquad [\mu H]^+_- = 0$$

Since $H = \frac{1}{i\omega\mu}\nabla \times E$ and μ is constant, the last line above becomes:

$$[n \times \nabla \times E]^+_- = 0, \qquad [\nabla \times E]^+_- = 0$$

We can do again all the integrations we have done before splitting them in two parts, one for the inside of S the other for the outside. We end up with the same expression as (2.32) except that we have an additional integral over S:

$$\begin{aligned}\int_\Omega \nabla \times \nabla \times E.F \; dv \;\; = \;\; & \int_\Omega E.\nabla \times \nabla \times F \; dv + \\ & \int_{\partial\Omega} n.(E \times \nabla \times F - F \times \nabla \times E) \; ds + \\ & \int_S n.\,[E \times \nabla \times F - F \times \nabla \times E]^+_- \; ds\end{aligned}$$

$[\,]^+_-$ denotes the jump between the outside and the inside of S. The last integral in the right-hand side becomes :

$$\begin{aligned}\int_S n.\,[E\times\nabla\times F - F\times\nabla\times E]^+_-\ ds &= \int_S [\nabla\times F.n\times E]^+_- - [F.\nabla\times E\times n]^+_-\ ds \\ &= \int_S [\nabla\times F]^+_-.n\times E - [F]^+_-.\nabla\times E\times n\ ds\end{aligned}$$

because $[n\times E]^+_- = [\nabla\times E\times n]^+_- = 0$ and finally,

$$\int_S n.\,[E\times\nabla\times F - F\times\nabla\times E]^+_-\ ds = \int_S E.\,[\nabla\times F\times n]^+_- - \nabla\times E.\,[n\times F]^+_-\ ds$$

Hence, the jump conditions we want to impose on F in order for (2.32) to be still valid even if we have discontinuities are :

$$[\nabla\times F\times n]^+_- = 0 \text{ and } [n\times F]^+_- = 0 \tag{2.49}$$

However, we have seen in the section 2.3.3 that F satisfy

$$L(F) = \nabla\times\nabla\times F - \omega^2\mu\epsilon F = 0 \tag{2.50}$$

which makes it equivalent to an electric field. As any electric field, it will then automatically fulfill the jump conditions (2.49): they are already implied by Maxwell's equations through (2.50).

We can then conclude that equation (2.32) is still valid even when we have a surface of discontinuity.

We are interested in cases where the position of S is one of our parameter with respect to which we want to compute the derivative. The software discretises any profile in a succession of z-invariant cross-sections. Following this approach, we assume that the element of which we are going to change the shape is entirely determined by any one of its cross-section. For the sake of simplicity we will consider only two regions of constant electric permitivity ϵ_1 and ϵ_2 as depicted in Figure (2.1).

The position of a point belonging to S_α is given by the curvilinear abscissa l and the function $\mathbf{r}(l,\alpha)$:

$$\begin{pmatrix} x \\ y \end{pmatrix} \in S_\alpha \Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{r}(l,\alpha).$$

At the first order, for a small perturbation $d\alpha = \alpha_1 - \alpha_0$, we have:

$$\mathbf{r}(l,\alpha_1) = \mathbf{r}(l,\alpha_0) + d\alpha\ \mathbf{u}(l)$$

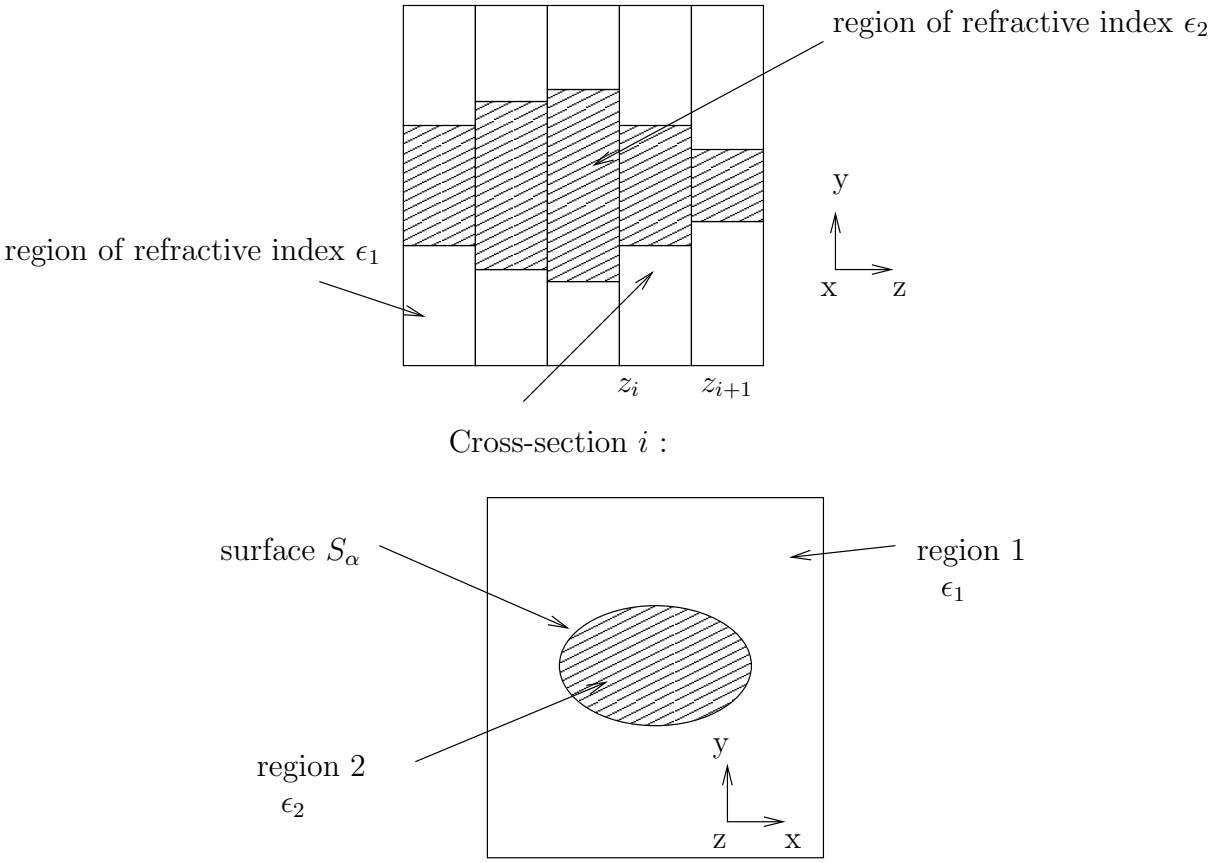


Figure 2.1: Description of the profile

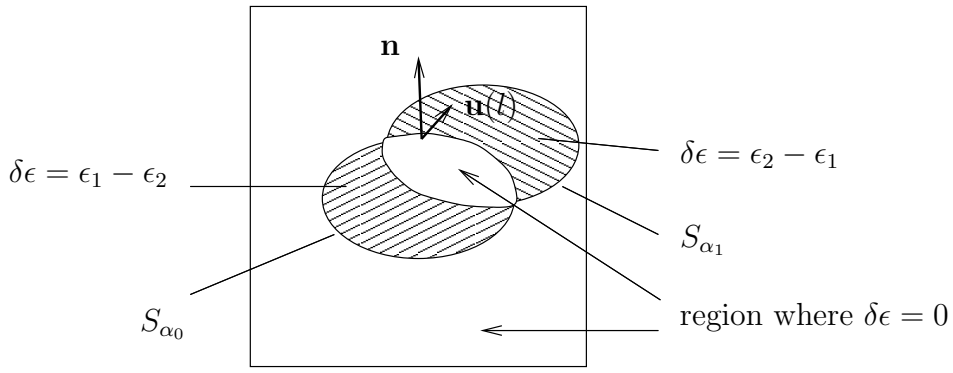


Figure 2.2: map of $\delta\epsilon$

where $\mathbf{u}(l)$ depends on the way we have defined our shape transformation (cf Figure (2.2)).

We have just shown that even when there is a discontinuity, equation (2.32) is still valid. Then, the same calculations as in the continuous case can be carried out and we still have:

$$\delta P_m = Re\,\langle\,\delta E\,, g\rangle_{\Gamma_r} = Re\,\langle L(\,\delta E\,), F\rangle_{\Omega}$$

However, we have to be very careful in the expression of $L(\,\delta E\,)$ because neither $\delta\epsilon$ nor δE are of a small order in $d\alpha$. Therefore, we have to consider the second order term $\delta\epsilon\,\delta E$ in the expression of $L(\,\delta E\,)$, i.e

$$L(\,\delta E\,) = \omega^2\delta\epsilon\mu E + \omega^2\delta\epsilon\mu\,\delta E$$

Hence,

$$\langle L(\,\delta E\,), F\rangle_{\Omega} = \omega^2\mu\int_{z_i}^{z_{i+1}} dz\int_{C_z}(\delta\epsilon E.F + \delta\epsilon\,\delta E\,.F)\;ds$$

where $C_z = \Omega\cap\{z\}$.

In order to evaluate the integrand above, we have to consider two cases depending on whether it is the region 2 or the region 1 which expands.

- $u.n \geq 0$, the region 2 expands.

$$(\delta\epsilon E.F + \delta\epsilon\,\delta E\,.F)\;ds = (\epsilon_2 - \epsilon_1)(E_1.F_1 + \,\delta E\,.F_1)u.n\;dld\alpha \tag{2.51}$$

E_1, F_1 and E_2, F_2 are the values of E and F on S_{α_0} respectively in the region 1 and region 2 sides.

We remind here that in the adjoint equation (2.48) the values of E and F that appears on the right-hand side are those computed before the deformation. This is the goal of the adjoint formulation: E and F are computed once and then one can deduce the result for any deformation. Equation (2.51) involves δE and so implies the computation of the field after deformation but it is in fact not necessary since at the first order in $d\alpha$ we have:

$$\delta E\; = E_2 - E_1 + o(d\alpha)$$

in the region where $\delta\epsilon = \epsilon_2 - \epsilon_1$.

This comes from the physical argument that E is continuous with respect to a small deformation $d\alpha$. Hence,

$$(\delta\epsilon E.F + \delta\epsilon\,\delta E\,.F)ds = (\epsilon_2 - \epsilon_1)E_2.F_1u.n\;dl\;d\alpha \tag{2.52}$$

• $u.n \leq 0$, the region 1 expands. We have:

$$\begin{aligned}(\delta\epsilon E.F + \delta\epsilon\,\delta E.F)ds &= (\epsilon_2 - \epsilon_1)(E_2.F_2 + (E_1 - E_2).F_2)u.n\ dl\ d\alpha \\ &= (\epsilon_2 - \epsilon_1)E_1.F_2(u.n)\ dl\ d\alpha \qquad (2.53)\end{aligned}$$

We get two different expressions for the integrand depending on which region expands. But, the derivative has to be continuous and it shouldn't matter whatever the expanding region is. We are going to prove that (2.52) and (2.53) are in fact equivalent.

The jump conditions on S_{α_0} are

$$[E_t]^+_- = [F_t]^+_- = 0 \text{ and } [\epsilon E_n]^+_- = [\epsilon F_n]^+_- = 0$$

where E_t, F_t and E_n, F_n are respectively the tangential and normal part of E, F with respect to S_0. We have

$$(\epsilon_2 - \epsilon_1)E_2.F_1 = (\epsilon_2 - \epsilon_1)(E_{2t}.F_{1t} + E_{2n}.F_{1n})$$

Since the tangential components are continuous,

$$(\epsilon_2 - \epsilon_1)(E_{2t}.F_{1t}) = (\epsilon_2 - \epsilon_1)(E_{1t}.F_{2t}).$$

As far as the normal component is concerned,

$$\epsilon_1\epsilon_2 E_{2n}F_{1n} = \epsilon_2 E_{2n}\epsilon_1 F_{1n} = \epsilon_1 E_{1n}\epsilon_2 F_{2n} = \epsilon_1\epsilon_2 E_{1n}F_{2n}$$

implies that

$$E_{2n}F_{1n} = E_{1n}F_{2n} \qquad (2.54)$$

and so,

$$(\epsilon_2 - \epsilon_1)E_{2n}F_{1n} = (\epsilon_2 - \epsilon_1)E_{1n}F_{2n} \qquad (2.55)$$

Hence,

$$(\epsilon_2 - \epsilon_1)E_2.F_1 = (\epsilon_2 - \epsilon_1)E_1.F_2$$

and (2.52) and (2.53) are equivalent.

Since

$$\begin{aligned}E_{1n}F_{2n} &= \pm\sqrt{(E_{1n}F_{2n})^2} \\ &= \pm\sqrt{E_{1n}F_{1n}E_{2n}F_{2n}} \text{ after using (2.54)}\end{aligned}$$

$E_{1n}F_{2n}$ that showed up in (2.55) can be seen as a geometrical mean of E_nF_n between the two sides of the discontinuity.

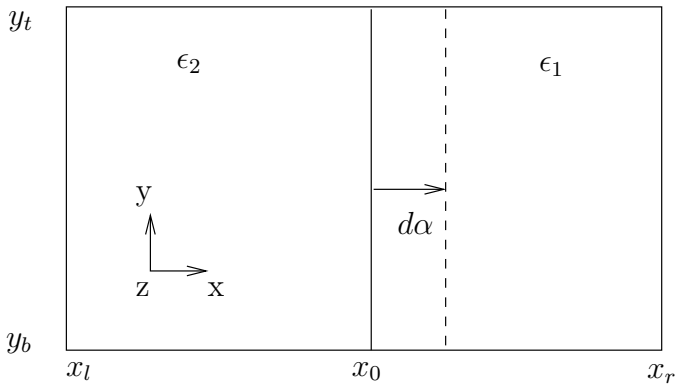


Figure 2.3: simple case of a rectangular shape

Finally, the derivative of P_m is given by

$$\frac{dP_m}{d\alpha} = Re \int_{z_i}^{z_{i+1}} dz \int_{C_z} (\epsilon_2 - \epsilon_1) E_1.F_2 \ (u.n) \ dl$$

Let's apply this result to the simple case of a rectangular shape moving from left to right of a distance $d\alpha$ (cf Figure (2.3))

We get :

$$\frac{dP_m}{d\alpha} = Re \int_{z_i}^{z_{i+1}} dz \int_{y_b}^{y_t} (\epsilon_2 - \epsilon_1) E_1.F_2 \ dy \tag{2.56}$$

Alternative derivation of dP_m for a rectangular shape

When dealing with physical problems, the proper way to handle a discontinuity is to consider it as a limit of a continuous configuration. That's what we are going to do now for the simple case of the rectangular shape and of course we hope to find the same result for the derivative.

We introduce the continuous function $H_\nu(x)$ which tend to $H(x)$ as ν tends to 0. H is a step function which gives the final discontinuous profile (cf Figure (2.4)). We let $\epsilon = H_\nu(x)$

Since we have smoothened the problem, there is no discontinuity and we can use equation 2.48.

$$\delta P_m = Re \int_{z_i}^{z_{i+1}} dz \int_{y_b}^{y_t} \int_{x_l}^{x_r} \delta\epsilon E.F \ dy \ dx$$

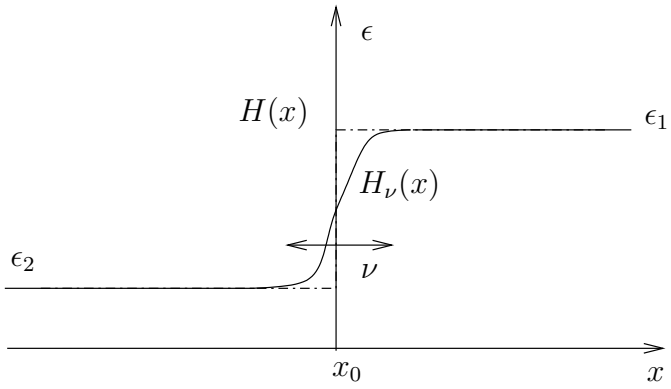


Figure 2.4: $H_\nu(x)$ and $H(x)$

Since $\delta\epsilon = H_\nu(x - d\alpha) - H_\nu(x)$, at the first order we have:

$$\delta\epsilon = -H'_\nu(x)d\alpha$$

and

$$\int_{y_b}^{y_t}\int_{x_l}^{x_r} \delta\epsilon E.F\ dy\ dx = -\int_{y_b}^{y_t} dy\ \int_{x_0-\frac{\nu}{2}}^{x_0+\frac{\nu}{2}} H'_\nu(x)E.F\ dx\ d\alpha$$

We again decompose $E.F$ into its tangential and normal components.

$$\int_{x_0-\frac{\nu}{2}}^{x_0+\frac{\nu}{2}} H'_\nu(x)E_t.F_t\ dx = E_t.F_t \int_{x_0-\frac{\nu}{2}}^{x_0+\frac{\nu}{2}} H'_\nu(x)\ dx\ + o(\nu)$$

because the fact that E_t and F_t are continuous across a discontinuity implies that E_t and F_t tend to a constant in a small region around x_0.

Hence,

$$\int_{x_0-\frac{\nu}{2}}^{x_0+\frac{\nu}{2}} H'_\nu(x)E_t.F_t\ dx = -E_{t1}.F_{t2}(\epsilon_2 - \epsilon_1)\ + o(\nu)$$

For the normal part, it is $K = H_\nu(x)^2 E_n F_n$ that will be constant in a small region around x_0. Thus we have :

$$\begin{aligned}
\int_{x_0-\frac{\nu}{2}}^{x_0+\frac{\nu}{2}} H'_\nu(x)E_n.F_n\ dx &= \int_{x_0-\frac{\nu}{2}}^{x_0+\frac{\nu}{2}} \frac{K}{H_\nu(x)^2} H'_\nu(x)\ dx \\
&= K\left[-\frac{1}{H_\nu(x)}\right]_{x_0-\frac{\nu}{2}}^{x_0+\frac{\nu}{2}} + o(\nu) \\
&= K\left[\frac{1}{\epsilon_2} - \frac{1}{\epsilon_1}\right] + o(\nu)
\end{aligned}$$

Finally, as ν tends to 0,

$$\int_{y_b}^{y_t}\int_{x_l}^{x_r} \delta\epsilon E.F\ dy\ dx \rightarrow \int_{y_b}^{y_t}\left((\epsilon_2-\epsilon_1)E_{t1}.F_{t1}+\epsilon_1^2 E_{1n}F_{1n}\left[\frac{1}{\epsilon_1}-\frac{1}{\epsilon_2}\right]\right)\ d\alpha dy$$

But the second term in the right-hand side can be rewritten as

$$\begin{aligned}\epsilon_1^2 E_{1n}F_{1n}\left[\frac{1}{\epsilon_1}-\frac{1}{\epsilon_2}\right] &= (\epsilon_2-\epsilon_1)\frac{\epsilon_1 E_{1n}F_{1n}}{\epsilon_2}\\ &= (\epsilon_2-\epsilon_1)E_{1n}F_{2n}\end{aligned}$$

and we retrieve the same result as before.

2.4 Computation of the derivative combined with modal expansion

Fimmprop-3d works entirely with modal expansions. This is the way Maxwell's equations are discretised and we have to reformulate our results in order to fit to the software.

In section 2.3.1, we derived the boundary conditions in a waveguide as a function of the input fields on the left and right-hand sides (equations (2.38) and (2.39)). The input fields are given by their modal expansion and so are completely determined by the coefficients α_k^l and α_{-k}^r where:

$$\begin{array}{lll} E_I^l = \sum_{k\geq 1}\alpha_k^l e_k^l & \text{and} & \alpha_k^l = \left\langle E_I^l, h_k^l\right\rangle_{\Gamma_l} \\ E_I^r = \sum_{k\geq 1}\alpha_{-k}^r e_{-k}^r & \text{and} & \alpha_{-k}^r = \left\langle E_I^r, h_k^r\right\rangle_{\Gamma_r}\end{array}$$

Hence, equations (2.38) and (2.39) become :

$$\begin{aligned}\nabla\times E - i\omega\mu B^{l-}(E) &= i\omega\mu\sum_{k\geq 1}\alpha_k^l(h_k^l - h_{-k}^l) \ \text{ on } \ \Gamma_l\\ \nabla\times E - i\omega\mu B^{r+}(E) &= i\omega\mu\sum_{k\geq 1}\alpha_{-k}^r(h_{-k}^r - h_k^r) \ \text{ on } \ \Gamma_r\end{aligned}$$

The boundary conditions for the adjoint problem have to be rewritten in this form : We have to find which input the field g in equation (2.57) corresponds to.

$$\nabla\times E - i\omega\mu B^{r+}(E) = g \ \text{ on the left-hand side} \tag{2.57}$$

First, we require g to satisfy

$$\left\langle g,\, \delta E\right\rangle_{\Gamma_r} = 2\left\langle\overline{E, h_m^r}\right\rangle_{\Gamma_r}\left\langle h_m^r,\, \delta E\right\rangle_{\Gamma_r} \tag{2.58}$$

This condition is necessary if we want to end up with the right expression for the derivative (equation (2.48)). We proposed $g = 2\left\langle \overline{E, h_m^r} \right\rangle_{\Gamma_r} h_m^r$ which obviously fullfills equation (2.58) but then this choice for g doesn't give us the form of an input and leads to an ill-posed problem. However, we know that the inner product $\langle ., . \rangle_{\Gamma_r}$ is not definite and in particular only depends on the tangential components of the vectors it is applied to (see equation (1.34)). Hence, since $h_{-m,s}^r = -h_{m,s}^r$,

$$g = -\left\langle \overline{E, h_m^r} \right\rangle_{\Gamma_r} (h_{-m}^r - h_m^r) \tag{2.59}$$

fulfills (2.58) and can be identified to an input on the right-hand side with the following coefficients:

$$\alpha_{-m}^r = \frac{i}{\omega\mu} \left\langle \overline{E, h_m^r} \right\rangle_{\Gamma_r} \quad \text{and } \alpha_{-k}^r = 0 \quad \forall k \neq m$$

Once the boundary conditions have been set, *Fimmprop-3d* computes the adjoint field F and we then have to compute the integrale:

$$\delta P_m = \omega^2 \mu \int_\Omega \delta\epsilon \, E.F \tag{2.60}$$

Here again, *Fimmprop-3d* gives us the modal expansion of the fields E and F i.e. the coefficients $a_n(z)$ and $b_n(z)$ in

$$E = \sum_{n \neq 0} a_n(z) e_n(z)$$
$$F = \sum_{n \neq 0} b_n(z) e_n(z)$$

where $\{e_n(z)\}_{n \neq 0}$ are the modes in the cross-section $\Omega \cap \{z\}$.

In a uniform waveguide, the modes do not couple one with the others and we have the following relations between the coefficients:

$$a_n(z_1) = a_n(z_0) e^{i\beta_n (z_1 - z_0)}$$

where β_n is the propagating constant of the mode n. As we will see later, the way we parametrise the taper only lead to integrate equation (2.60) over a z-uniform element. In this case, the modes are the same within the element and the fields E and F have the following expansions:

$$E = \sum_{n \neq 0} a_n^l e^{i\beta_n z} e_n$$
$$F = \sum_{n \neq 0} b_n^l e^{i\beta_n z} e_n$$

where a_n^l and b_n^l are respectively equal to $a_n(z)$ and $b_n(z)$ at the beginning of the element. Hence, integrated over an element K, (2.60) becomes:

$$\begin{aligned}
\delta P_m &= Re\left[\omega^2\mu \int_K \sum_{p,q\neq 0} a_p^l b_q^l \, e^{i(\beta_p+\beta_q)z}\delta\epsilon(x,y)e_p(x,y).e_q(x,y)\; dxdydz\right] \\
&= Re\Bigg[\omega^2\mu \sum_{\beta_p+\beta_q\neq 0} a_p^l b_q^l \, \frac{e^{i(\beta_p+\beta_q)d}-1}{i(\beta_p+\beta_q)} \int_S \delta\epsilon(x,y)e_p(x,y).e_q(x,y)\; dxdy \\
&\quad + \omega^2\mu \sum_{\beta_p+\beta_q=0} a_p^l b_q^l \, d \int_S \delta\epsilon(x,y)e_p(x,y).e_q(x,y)\; dxdy\Bigg]
\end{aligned}$$

where d is the length of the element and S its cross-section. Finally, after introducing a_n^r and b_n^r, the value of the coefficients $a_n(z)$ and $b_n(z)$ at the end of the element, we get:

$$\begin{aligned}
\delta P_m = Re\Bigg[&\omega^2\mu \sum_{\beta_p+\beta_q\neq 0} \frac{a_p^r b_q^r - a_p^l b_q^l}{i(\beta_p+\beta_q)} \int_S \delta\epsilon\, e_p.e_q\; dxdy \\
&+ \omega^2\mu \sum_{\beta_p+\beta_q=0} a_p^l b_q^l \, d \int_S \delta\epsilon\, e_p.e_q\; dxdy\Bigg] \qquad (2.61)
\end{aligned}$$

The coefficients a_p^l,b_q^l,a_p^r,b_q^l and the modes are computed by *Fimmprop-3d* and now, given the modes and the modified distribution $\delta\epsilon$, we have to implement the computation of the integrale $\int_S \delta\epsilon\, e_p.e_q\; dxdy$. It was part of my numerical work to do it and the next chapter explains how we proceed.

Chapter 3

Numerical Evaluation

3.1 General framework

It has been part of my collaboration with *PhotonDesign* to develop an optimisation module in C++ using adjoint methods to compute the derivative. We tried to make this module as generic as possible in order to handle a wide range of problems.

Fimmprop-3d is used via a client-server interface. Basically, we send commands that set the boundary conditions (the input fields), launch the computation and retrieve the results. All the computation parameters are set within *Fimmprop-3d* and this allows us to use the whole efficiency of the software.

The user of the optimisation module first has to create a *Fimmprop-3d* project which includes a taper to be optimised. The taper is sliced into N $\hat{z}$-uniform elements (the number of elements N is given by the user, see Figure (3.1)). Then, the optimization parameters $(p_1, p_2, ..., p_q)$ have to be selected. There is a large variety of parameters that can be chosen, as many as the *Fimmprop-3d* section builder proposes. For example, it can be the radius of a round shape of constant refractive index or a parameter in a refractive index distribution law etc...

In a taper, the power remaining in the first mode is often the quantity one wants to optimise but we give the user the possibility to define a wider range of objective functions (see appendix).

Two optimisation routines are available : a quasi-newton method based on recursive approximation of the hessian matrix or an inverse problem method

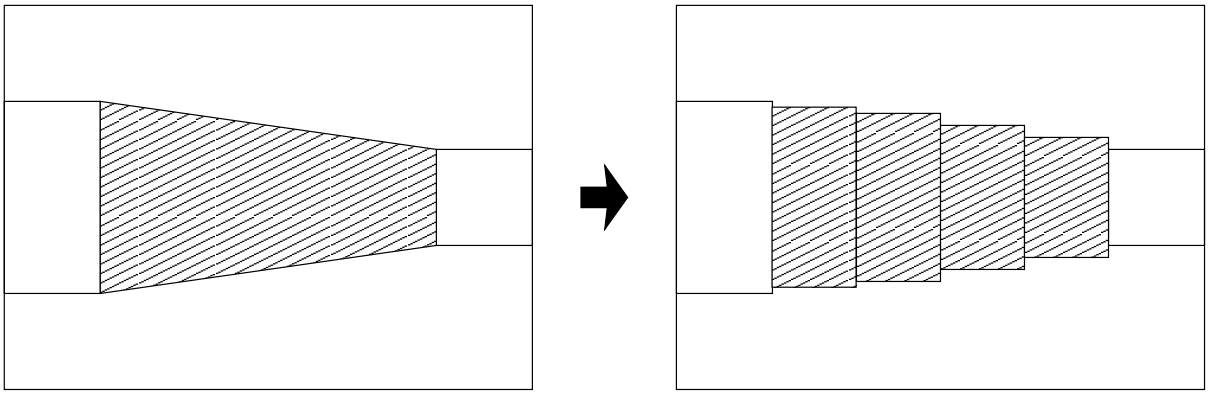


Figure 3.1: The taper is discretised in $N = 4$ elements

which takes into account the ill-possedness of the problem (refer to Thomas Felici's article [6]). Both of them are iterative methods and require at each step the derivative of P_m.

3.2 Numerical computation of the derivative expression

Each parameter p_i acts separately on every $\hat{z}$ invariant element and that's why we are led, as shown in the previous chapter, to compute integrals of the type:

$$\int_S \frac{\partial \epsilon}{\partial p_i}(x,y) e_p(x,y).e_q(x,y)\ dxdy \tag{3.1}$$

where K is a $\hat{z}$-uniform element whose modes are $\{e_p\}_{p\neq 0}$. δP_m will then be retrieved by equation (2.61) as shown in the previous chapter.

We first have to evaluate $\frac{\partial \epsilon}{\partial p_i}(x,y)$ for every optimisation parameter p_i. These parameters can be chosen among a wide range of variables and it is hard to derive an analytic expression for all the derivatives $\partial\epsilon/\partial p_i$. Instead, since it is easy with *Fimmprop-3d* to build a shape for a given set of parameters, we create a slightly modified section $S(p_i + \delta p_i)$ then by comparing it with the initial section $S(p_i)$, we construct a virtual section $\tilde{S}_i$

$$\tilde{S}_i := S(p_i + \delta p_i) - S(p_i)$$

which gives us a map of the modified electric permittivity $\delta\epsilon$. In fact, we are using a kind of finite difference scheme to compute the derivative $\frac{\partial \epsilon}{\partial p_i}$.

Two cases show up depending whether the parameter p_i modifies the geometry of the cross-section or its refractive index. If the geometry is modified, we have to consider the integrals of type *II* derived in section 2.3.4 for a shape deformation (C denotes the surface which is deformed) otherwise we compute a surface integral of type *I*.

$$\int_S \delta\epsilon \; e_p.e_q \; dxdy \qquad \text{type } I$$
$$\int_C \delta\alpha[\,\epsilon\,]^+_- e_p^- e_q^+ (u.n) \; dl \qquad \text{type } II$$

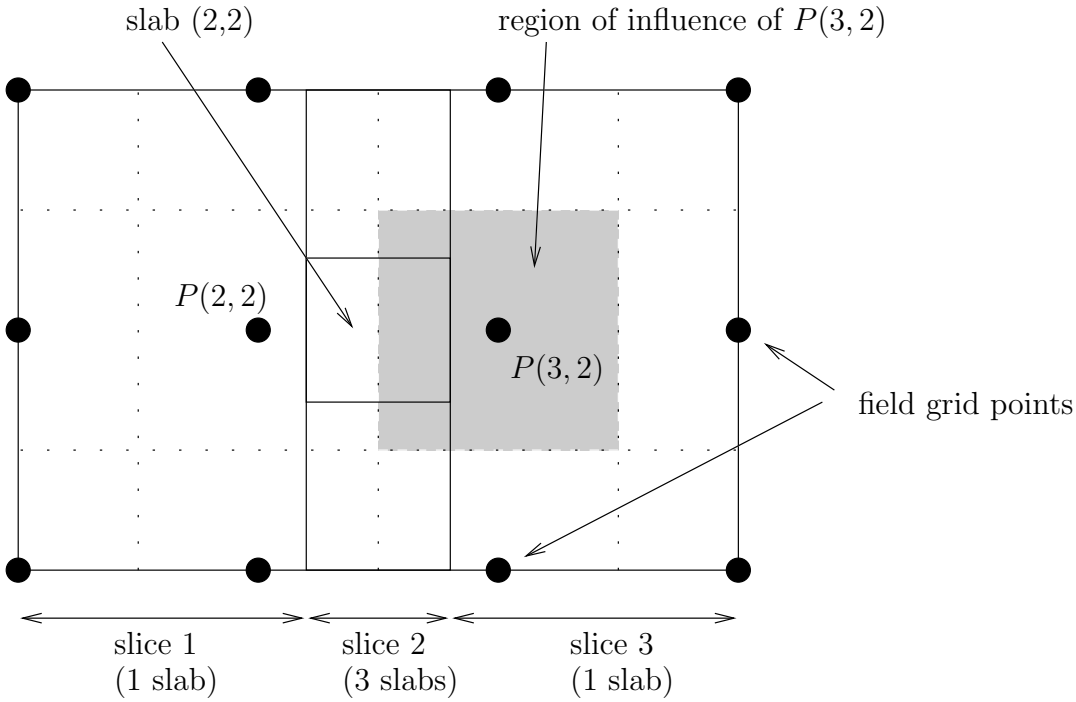


Figure 3.2: discrete waveguide

Fimmprop-3d reduces any waveguides to discrete waveguides comprising of slices and slabs of constant refractive index. The problem of such waveguides is that they induce a lot of discontinuities in the refractive index distribution which is piecewise constant and so the eigenmodes have also many discontinuities. The integrations become then particularly tricky. On the other hand, the advantage of this decomposition is that in the case of integrals of type *II* we only have to consider vertical and horizontal lines of discontinuity because a discrete waveguide is only determined by such lines.

The integration (3.1) is done successively over every slab in every slice. In this way, we integrate successively over domains where the fields are continuous because the refractive index is constant. $\delta\epsilon$ is given by $\tilde{S}_i$. The modes e_p are retrieved via *Fimmprop-3d* on a grid of the section (see Figure (3.2)).

The fields need to be evaluated everywhere in the domain and not only at the grid points. In this purpose, for every point of the grid, we define a *region of influence* (in dotted lines in Figure (3.2)), a rectangle centred around a field grid point. In a region of influence the fields are the best approximated by the value at the center point of this region because it is the closest. Of course, this procedure makes sense only if the field is continuous or if we take into account the discontinuities via jump conditions. For example, let consider the case of a square section of constant electric permitivity ϵ as depicted in Figure (3.3). *Fimmprop-3d* produces the equivalent discrete waveguide of Figure (3.2) and the square becomes the slab indexed (2,2). This slab does not have any field grid point in it but it belongs to the regions of influence of points $P(2,2)$ and $P(3,2)$. If we take the value at those points as approximation of the fields in the slab, we do a big error because the fields are discontinuous. However, instead of ignoring these regions (which should be small if the grid is fine enough), we use the jump conditions $[E_t] = [\epsilon E_n] = 0$ to retrieve the fields across a discontinuity (see Figure 3.4). However, this procedure

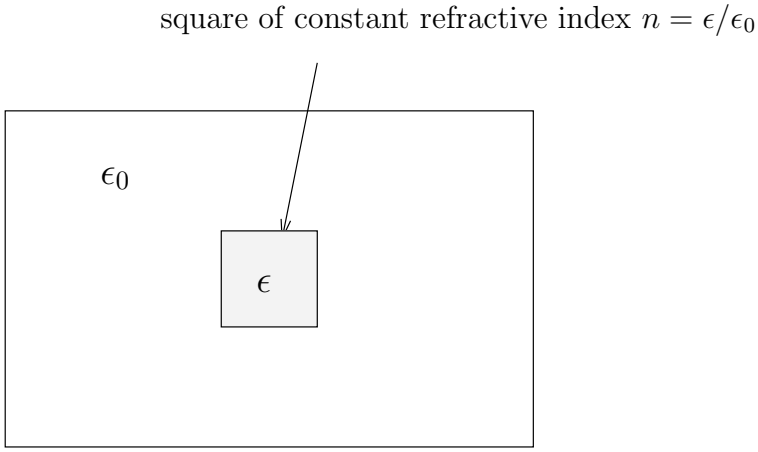


Figure 3.3: this section is transformed into the discrete waveguide of Figure (3.2)

fails when it is applied to corners. In electromagnetism, edges always induce scattering and so, if we want to approximate the value of a field E inside a region with a corner A (see Figure (3.5)) with a value given outside A, the jump condtions do not suffice. We have to take a closer look at the corner

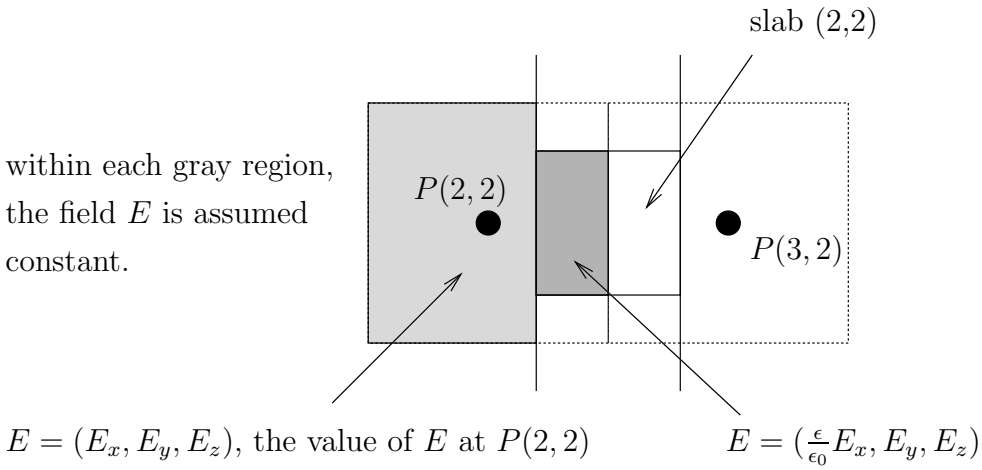


Figure 3.4: jump conditions are used to retrieve the field across discontinuity

region and probably an asymptotic expansion in this region would give us the right extrapolation relations. This work hasn't be done yet but we can still expect good results if we use a fine grid. The price has then to be paid in computation time.

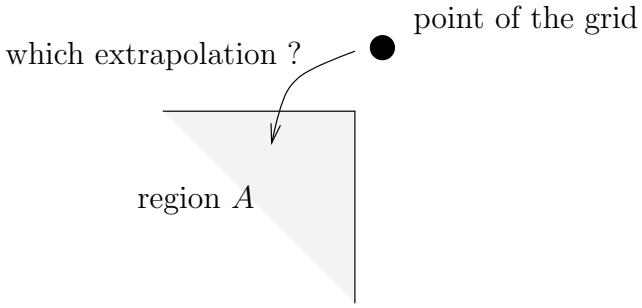


Figure 3.5: In a corner, new extrapolation relations have to be found

3.3 Test cases

3.3.1 Test case 1 : uniform waveguide

This case is the simplest : On both sides are the same cross-sections. A uniform waveguide between these two sections carries the first mode from one side to the other without exciting the others and so all the power remains in this mode. We expect then the derivative δP_1 to vanish whatever the variation $\delta\epsilon$ is because we have reached a maximum of transmitting power in the first mode.

Following the calculations derived in section 2.4, we set

$$a_1^l = 1 \quad a_k^l = 0, \ \forall\, k > 1$$

which means that only the first mode is excited on the left-hand side. Since the waveguide is uniform the coefficient $a_n(z)$ on the right-hand side are simply given by a phase shift:

$$a_1^r = e^{i\beta_1 d} \quad a_k^r = 0, \ \forall\, k > 1$$

where d is the length of the waveguide.

The input field for the adjoint problem is on the right-hand side :

$$b_{-1}^r = \frac{i}{\omega\mu} \langle E, h_m^r \rangle_{\Gamma_r} \quad b_k^r = 0, \ \forall\, k < -1$$

hence

$$b_{-1}^r = \frac{i}{\omega\mu} \overline{a_1^r} = \frac{i}{\omega\mu} e^{-i\beta_1 d}$$

Again, since the waveguide is uniform, b_{-1}^l is retrieved by a phase shift:

$$b_{-1}^l = \frac{i}{\omega\mu}$$

Plugging these coefficients into equation (2.61), we are left with

$$\begin{aligned} \delta P_1 &= Re\left[\omega^2 \mu a_1^l b_{-1}^l d \int_S \delta\epsilon \, e_1 . e_{-1} \, dxdy\right] \\ &= Re\left[i\omega d \int_S \delta\epsilon \, e_1 . e_{-1} \, dxdy\right] \\ &= 0 \end{aligned}$$

for any deformation $\delta\epsilon$, as expected.

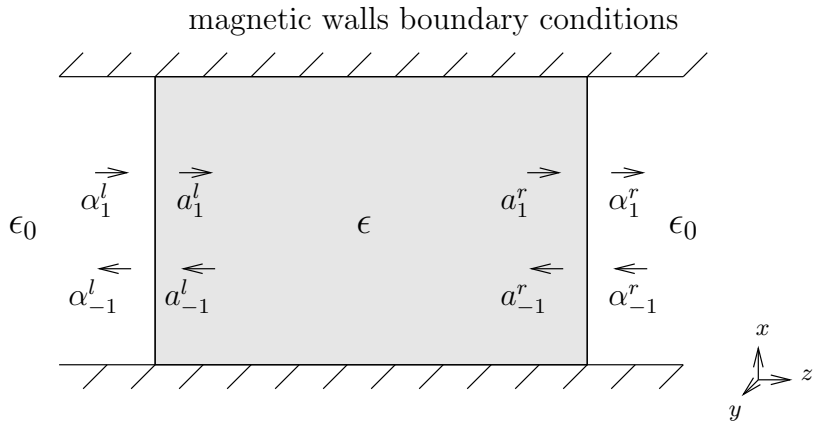


Figure 3.6: Test case 2 : the refractive index in the waveguide is constant and different from the sides

3.3.2 Test case 2

In this test-case, we will consider a uniform 2d waveguide with a constant electric permitivity ϵ. On the sides, the cross-sections are the same but the electric permitivity is different and equals to ϵ_0 (see Figure (3.6)). In 2d, Maxwell's equations reduce to

$$\nabla^2 u + k^2 u = 0$$

where $u = E_y$, $E_x = E_z = 0$ and $k = \epsilon\mu\omega^2$. k is constant.

We replace the electric walls boundary conditions by magnetic walls boundary conditions i.e.:

$$\frac{\partial u}{\partial x} = 0 \text{ on the boundaries}$$

One can prove that there still exists an eigen expansion for the field u and the modes satisfy the equations

$$\begin{aligned} &\frac{\partial^2 u}{\partial x^2} + (k^2 - \beta^2)u = 0 \\ &\frac{\partial u}{\partial x} = 0 \text{ on the boundaries} \end{aligned} \tag{3.2}$$

$\beta_1 = k$ and u_1 constant is a solution of these equations and therefore is a mode.

The modes are orthogonal :

$$\int_S u_i.u_j \; dx = 0 \;\; \forall i \neq j$$

The modes u_l on the left-hand side of the waveguide satisfy similar equations:

$$\begin{aligned} &\frac{\partial^2 u_l}{\partial x^2} + (k_0^2 - \beta_l^2)u_l = 0 \\ &\frac{\partial u_l}{\partial x} = 0 \quad \text{on the boundaries} \end{aligned} \tag{3.3}$$

where $k_0 = \epsilon_0 \omega \mu$.

(3.3) is equivalent to

$$\nabla^2 u + [k^2 - (k^2 + \beta_l^2 - k_0^2)]u = 0$$

Hence u_l is solution of of (3.2) for $\beta^2 = k^2 - k_0^2 + \beta_l^2$ and the modes on the left-hand side (and on the righ-hand side) are the same as in the waveguide but with different propagating constants β and β_l ($= \beta_r$). It implies that the modes on the left-hand side ($\{u_i^l\}_{i\geq 1}$) and the modes in the waveguide ($\{u_i\}_{i\geq 1}$) are orthogonal:

$$\int_S u_i^l . u_j \; dx = 0 \quad \forall i \neq j$$

After imposing the continuity of the tangential components of E and H across the section S_0 (see figure (3.6)) and making use of the previous orthogonality relations as we did in section (1.5), we end up with

$$\alpha_1^l u_1^l + \alpha_{-1}^l u_1^l = a_1^l u_1 + a_{-1}^l u_1 \tag{3.4}$$

$$\alpha_1 \beta_1^l u_1^l - \alpha_{-1}^l \beta_1^l u_1^l = a_1^l \beta_1 u_1 - a_{-1}^l \beta_1 u_1 \tag{3.5}$$

where α_1^l and α_{-1}^l are the coefficients of the modal decomposition on the left-hand side.

u_1^l and u_1 are constants which are determined by normalization and we have (see appendix):

$$u_1 = \sqrt{\frac{\omega\mu}{k_0 n w}} \quad \text{and} \quad u_1^l = \sqrt{\frac{\omega\mu}{k_0 n^l w}}$$

(3.4) and (3.5) then become

$$(\alpha_1^l + \alpha_{-1}^l)\sqrt{1/n^l} = (a_1^l + a_{-1}^l)\sqrt{1/n} \tag{3.6}$$

$$(\alpha_1^l - \alpha_{-1}^l)\sqrt{n^l} = (a_1^l - a_{-1}^l)\sqrt{n} \tag{3.7}$$

where n and n^l are respectively the refractive index in the waveguide and on the left-hand side. Since $\epsilon = \epsilon_0$ on the left-hand side, $n^l = 1$. A similar relation holds an the right-hand side (with the coefficients α_1^r and α_{-1}^r) and allow us to solve completely the problem.

We set the inputs:

$$\begin{array}{ll} \alpha_1^l &= 1 : \text{the first mode is excited} \\ \alpha_{-1}^r &= 0 : \text{no input on the right-hand side} \end{array}$$

and the outputs are:

$$\begin{array}{ll} \alpha_{-1}^l &= r : \text{a part of the signal is reflected} \\ \alpha_1^r &= t : \text{the rest of it is transmitted on the} \\ & \quad \text{right-hand side} \end{array}$$

r and t are unknown and are determined by equations (3.6) and (3.7):

$$t = \frac{4ne^{ik_0nd}}{(n+1)^2 - (n-1)^2 e^{2ik_0nd}}$$

(see appendix for details).

The remaining power in the first mode P_1 and its derivative are given by:

$$\begin{aligned} P_1 &= |t|^2 \\ \frac{dP_1}{dn} &= 2Re\left[\bar{t}\frac{dt}{dn}\right] \end{aligned} \tag{3.8}$$

with

$$\frac{dt}{dn} = \frac{4e^{ik+0nd}\left(1 - n^2 + ik_0nd(n+1)^2\right) + 4e^{3ik_ond}\left(n^2 - 1 + ik_0nd(n-1)^2\right)}{\left[(n+1)^2 - (n-1)^2 e^{2ik_0nd}\right]^2}$$

Our goal is now to check that we end up with the same expression by using the adjoint method. The input field on the the right-hand side is $\alpha_{-1}^r = \frac{i\bar{t}}{\omega\mu}$ and we set

$$\begin{array}{ll} \alpha_{-1}^r = 2\dfrac{i\bar{t}}{\omega\mu} &: \text{excitation of the first mode on the} \\ & \quad \text{righ-hand side} \\ \alpha_1^l = 0 &: \text{no input on the left-hand side} \end{array}$$

The outputs are:

$$\begin{array}{ll} \alpha_1^r &= \tilde{r} : \text{a part of the signal is reflected} \\ \alpha_{-1}^l &= \tilde{t} : \text{the rest of it is transmitted on the} \\ & \quad \text{left-hand side} \end{array}$$

$\tilde{t}$ and $\tilde{r}$ are again unknown which have to be determined. With the coefficient b_1^l,b_{-1}^l,b_1^r and b_{-1}^r of the adjoint equation, they are computed in the appendix. From (2.61), we get:

$$\begin{aligned}
\frac{dP_1}{dn} &= \frac{dP_1}{d\epsilon}\frac{d\epsilon}{dn} \\
&= 2n\epsilon_0\omega^2\mu Re\left[\frac{a_1^r b_1^r - a_1^l b_1^l}{2i\beta_1} - \frac{a_{-1}^r b_{-1}^r - a_{-1}^l b_{-1}^l}{2i\beta_1} + (a_1^l b_{-1}^l + a_{-1}^l b_1^l)w\right] \\
&= 2Re\left[\overline{t}\frac{dt}{dn}\right] \quad \text{as computed in equation (3.8) (see appendix for the details)}
\end{aligned}$$

The two methods happily lead to the same results.

This test case is not of great interest in itself but nevertheless it allowed us to correct a lot of programming errors and it also shows how rapidly analytic expressions for the derivative become very complex.

APPENDIX

1 Generalisation to any objective function

We can generalise the computation of the derivative to any objective function f. We have:

$$\delta f = \nabla f.(\delta a_1, \delta a_2, ..., \delta a_p)$$

We consider here only p modes. The coefficients a_i are the modal decomposition cofficients of the output field on the right-hand side.

Since, for any $i = 1, ..., p$,

$$a_i = \langle E, h_i^r \rangle_{\Gamma_r}$$

we have

$$\delta a_i = \langle\, \delta E\,, h_i^r \rangle_{\Gamma_r}$$

and so,

$$\begin{aligned} \delta f &= \nabla f.\left(\langle\, \delta E\,, h_1^r \rangle_{\Gamma_r}\,, ..., \langle\, \delta E\,, h_p^r \rangle_{\Gamma_r}\right) \\ &= \left\langle \delta E\,, \sum_{i=1}^{p} \frac{\partial f}{\partial a_i} h_i^r \right\rangle_{\Gamma_r} \end{aligned}$$

Hence, we deduce g:

$$g = \sum_{i=1}^{p} \frac{\partial f}{\partial a_i} \frac{h_i^r - h_{-i}^r}{2}$$

and the adjoint problem input coefficients on the right-hand side are

$$\alpha_{-i}^r = \frac{i}{2\omega\mu} \partial f a_i$$

Once the adjoint field has been found, δf is given by

$$\delta f = \int_\Omega \delta\epsilon\, E.F\; dv$$

2 Test case 2 : calculation details

2.1 Modes normalisation

The modes are normalized such that each mode satisfies:

$$\int_S E \times H.\hat{z}\ ds = 1 \tag{A-1}$$

Since we are dealing with a 2d case, we have $E = (0, u, 0)e^{i\beta z}$ and so,

$$H = \frac{1}{i\omega\mu}\nabla \times E = \frac{1}{i\omega\mu}\begin{pmatrix} -i\beta u \\ 0 \\ \frac{\partial u}{\partial x} \end{pmatrix} e^{i\beta z}$$

since u_1, the first mode, is constant, we get

$$\int_S E_1 \times H_1.\hat{z}\ ds = \frac{\beta u_1^2 w}{\omega\mu}$$

hence, from (A-1),

$$u_1 = \sqrt{\frac{\omega\mu}{\beta w}} \tag{A-2}$$

2.2 Computation of the direct problem

On the left-hand side,

$$\begin{aligned} 1 + r &= (a_1^l + a_{-1}^l)\sqrt{1/n} \qquad (n_1^l = 1 \text{ because } \epsilon^l = \epsilon_0) \\ 1 - r &= (a_1^l - a_{-1}^l)\sqrt{n} \end{aligned}$$

implies

$$2\sqrt{n} = a_1^l(n+1) + a_{-1}^l(1-n) \tag{A-3}$$

On the righ-hand side,

$$\begin{aligned} (a_1^r + a_{-1}^r)\sqrt{1/n} &= t \\ (a_1^r - a_{-1}^r)\sqrt{n} &= t \end{aligned} \tag{A-4}$$

implies

$$a_1^r(n-1) - a_{-1}^r(n+1) = 0 \tag{A-5}$$

By using the identities

$$\left\{ \begin{array}{lcll} a_1^r & = & a_1^l e^{ik_0 nd} & \\ a_{-1}^r & = & a_{-1}^l e^{-ik_0 nd} & k_0 n = k = \beta_1 \end{array} \right.$$

(A-3) and (A-5) give us

$$\begin{aligned} a_1^l &= \frac{2\sqrt{n}(n+1)}{A} \\ a_{-1}^l &= \frac{2\sqrt{n}e^{2ik_0 nd}(n-1)}{A} \end{aligned}$$

where $A = (n+1)^2 - (n-1)^2 e^{2ik_0 nd}$ and from (A-4),

$$t = \frac{4ne^{ik_0 nd}}{A}$$

2.3 Computation of the adjoint problem

On the right-hand side,

$$\begin{aligned} \tilde{r} + \frac{i\overline{t}}{\omega\mu} &= (b_1^r + b_{-1}^r)\sqrt{1/n} \\ \tilde{r} - \frac{i\overline{t}}{\omega\mu} &= (b_1^r - b_{-1}^r)\sqrt{n} \end{aligned}$$

implies

$$\frac{2i\overline{t}}{\omega\mu} = b_1^r(1-n) + b_{-1}^r(1+n) \qquad \text{(A-6)}$$

On the right-hand side,

$$\begin{aligned} \tilde{t} &= (b_1^l + b_{-1}^l)\sqrt{1/n} \\ -\tilde{t} &= (b_1^l - b_{-1}^l)\sqrt{n} \end{aligned}$$

implies

$$b_1^l(1+n) + b_{-1}^l(1-n) = 0 \qquad \text{(A-7)}$$

By using the identities

$$\left\{ \begin{array}{lcl} b_1^r & = & b_1^l e^{ik_0 nd} \\ b_{-1}^r & = & b_{-1}^l e^{-ik_0 nd} \end{array} \right.$$

(A-6) and (A-7) give us

$$\begin{aligned} b_1^l &= \frac{2i\bar{t}\sqrt{n}(n-1)e^{ik_0nd}}{\omega\mu A} \\ b_{-1}^l &= \frac{2i\bar{t}\sqrt{n}(n+1)e^{ik_0nd}}{\omega\mu A} \end{aligned}$$

We have:

$$\begin{aligned} \frac{a_1^r b_1^r - a_1^l b_1^l}{2i\beta_1} &= \frac{4in\bar{t}}{\omega\mu 2ik_o nA^2}(n^2-1)e^{ik_0nd}(e^{2ik_0nd}-1) \\ \frac{a_{-1}^l b_{-1}^l - a_{-1}^r b_{-1}^r}{2i\beta_1} &= \frac{4in\bar{t}}{\omega\mu 2ik_0 nA^2}(n^2-1)e^{3ik_0nd}(1-e^{-2ik_0nd}) \\ a_1^l b_{-1}^l d &= \frac{4in\bar{t}}{\omega\mu A^2}(n+1)^2 e^{ik_0nd}d \\ a_{-1}^l b_1^l d &= \frac{4in\bar{t}}{\omega\mu A^2}(n-1)^2 e^{3ik_0nd}d \end{aligned}$$

Hence, plugging these results into equation (2.61), we get:

$$\begin{aligned} \frac{dP_1}{dn} &= \frac{dP_1}{d\epsilon}\frac{d\epsilon}{dn} \\ &= 2n\epsilon_0\omega^2\mu Re\left[\frac{a_1^r b_1^r - a_1^l b_1^l}{2i\beta_1} + \frac{a_{-1}^l b_{-1}^l - a_{-1}^r b_{-1}^r}{2i\beta_1} + (a_1^l b_{-1}^l + a_{-1}^l b_1^l)d\right] u_1^2 w \\ &= 2\frac{u_1^2 w\epsilon_0 nw}{k_0} Re\Bigg[\frac{4\bar{t}}{A^2}\big((n-1)^2(e^{3ik_0nd}-e^{ik_0nd}) + \\ &\qquad ik_0nd(n+1)^2e^{ik_0nd} + ik_0nd(n-1)^2e^{4ik_0nd}\big)\Bigg] \end{aligned}$$

From (A-2), we get

$$\frac{u_1^2 w\epsilon_o n\omega}{k_0} = \frac{\omega^2\mu\epsilon_0 n}{\beta_1 k_0} = \frac{k_0^2 n}{k_0^2 n} = 1$$

and we end up with

$$\frac{dP_1}{dn} = 2Re\left[\frac{4\bar{t}}{A^2}\big((n-1)^2(e^{3ik_0nd}-e^{ik_0nd}) + ik_0nd(n+1)^2e^{ik_0nd} + ik_0nd(n-1)^2e^{4ik_0nd}\big)\right]$$

Conclusion

The adjoint method provides a general and solid framework to solve efficently a lot of optimisation problems. In this dissertation, we have applied it successfully to electromagnetism in the case of a taper. We have derived an adjoint formulation of Maxwell's equations in a waveguide from which we deduced an expression of the derivative of P_m, the power remaining in the mth mode.

$$\delta P_m = \omega^2 \mu Re \int_\Omega \delta\epsilon \, E.F \; dv \tag{C-1}$$

The electric field E satisfies Maxwell's equations while the adjoint field F satisfies the adjoint equations. In the taper case we have the remarkable property that the adjoint equations are the same as the original ones. F is solution of Maxwell's equations in the waveguide for specific boundary conditions which depend on the function we want to optimise, in this case P_m.

Since we want to optimise a shape, the expression (C-1) which has been derived for continuous fields can't be applied. That's why derived another expression for a shape deformation which takes into account the field's discontinuities at the shape boundary S:

$$\delta P_m = \int_S \delta\alpha \; [\epsilon]_-^+ \; E_+.F_- \, (u.n) \; ds \tag{C-2}$$

An important part of this project consisted of writing an optimisation module integrated in *Fimmprop-3d* , the Maxwell's equation solver of *PhotonDesign* . We implemented the computation of the derivatives (C-1) and (C-2) in order to replace the finite difference method currently in use by an adjoint method which should improve significantly the computation time. We get satisfying results for the two test cases described in the last chapter but not for more complex configurations.

Further works

The very next thing to do is naturally to make the code work!

We could also improve the evaluation of the integrals in (C-1) and (C-2) in taking more into account the singularities of the fields in particular regions such as corners.

This study focussed on the computation of the derivative which is an important part of the optimisation process but the optimisation scheme is also of course very important. We only considered step taper which are discretised in z uniform sections. It should be very interesting to consider smoother discretisations where the shape varies continuously.

To conclude, I would like to say that I very much enjoyed working on this industrial project. It was a very challenging project and I learnt a lot in many fields. I am now looking forward to doing further works to complete it.

Bibliography

[1] Peter Bienstman. *Rigorous and efficient modelling of wavelength sacel photonic components*. University of Gent, 2001.

[2] Weng Cho. Chew. A variational analysis of anisotropic, inhomogeneous dielectric waveguides. *IEEE transactions on microwave theory and tech.*, **37**:661–668, 1989.

[3] Weng Cho. Chew. *Waves and fields in inhomogeneous media*. IEEE Press series on electromagnetic waves, 1995.

[4] Robert E. Collin. *Field theory of guided waves*. IEEE, Oxford University Press, 1991.

[5] Thomas Felici. On propagation through long step tapers. *Optical and quantum Electronics*, **33**:399–411, 2001.

[6] Thomas Felici. On shape optimization of optical waveguides using inverse problem techniques. *Inverse Problems*, **17**:1141–1162, 2001.

[7] Mike Giles. An introduction to the adjoint approach in design. *Flow, Turbulence and Combustion*, **65**:393–415, 2000.

[8] D.S. Jones. *Acoustic and electromagnetic waves*. Clarendon Press, 1986.

[9] Tullio Rozzi. General constraints on the propagation of complex waves in closed lossless isotropic waveguides. *IEEE transactions on microwave theory and tech.*, **46**:512–516, 1998.

[10] Walter Rudin. *Functional analysis*. McGraw-Hill, 1991.

[11] A Snyder and J.D. Love. *Optical waveguide theory*. Chapman and Hall, 1983.