

Présentation d'un article :  
On shape optimization of optical waveguides using inverse problem  
techniques

Thomas Felici & Heinz W Engl

Alexandre VIEIRA

4 février 2015

## Table des matières

<b>1</b>	<b>Formulation du problème direct en 2D</b>	<b>2</b>
<b>2</b>	<b>Formulation du problème d'optimisation</b>	<b>5</b>
<b>3</b>	<b>Solution du problème direct</b>	<b>5</b>
3.1	Recherche d'une représentation locale . . . . .	5
3.2	Démonstration . . . . .	6
3.3	Équation d'évolution de l'amplitude . . . . .	8
3.4	Précisions sur les conditions aux bords . . . . .	8
<b>4</b>	<b>Solution numérique du problème d'optimisation</b>	<b>8</b>
4.1	Choix de la discrétisation . . . . .	8
4.2	Résultats et interprétation physique . . . . .	9
4.3	Un problème mal posé . . . . .	10

# Introduction

Les guides d'ondes optiques sont à la base de l'optoélectronique, ie les composants électroniques interagissant avec de la lumière, et de l'industrie des télécommunications. L'exemple le plus connu est bien évidemment la fibre optique, mais il existe également d'autres exemples tout aussi important, comme les composants optiques qui manipulent, filtrent et dispatchent des signaux entrants. Ces composants sont souvent de l'ordre du quart de nanomètre, ce qui représente bien sûr un avantage en poids et en volume, et qui permettent de remplir des fonctions que ne peuvent pas faire des objets plus classiques. Le développement de cette technologie repose de plus en plus sur des modèles mathématiques de paire avec des simulations numériques pour la prédiction du comportement de ces appareils, ainsi que la production d'un appareil 'optimal', ce qui repose bien souvent sur des modèles reposant sur des EDP.

Les adaptateurs de mode intégrés, aussi appelés "taper" en anglais, est un guide d'onde dont la section varie sur sa longueur, utilisés comme des sortes d'entonnoir. Il est bien connu que si le taper est assez long, la lumière sera transmise sans perte d'énergie. Mais plus la longueur est courte, plus le faisceau perd en énergie. Le but de ce papier était donc de trouver une formulation pour minimiser la perte d'énergie pour une longueur donnée.

Dans un premier temps, l'article met en place un premier problème direct pour la propagation d'une onde électromagnétique à travers un guide d'onde, puis il définit le problème d'optimisation. Comme on le verra, on peut ensuite dériver de ce modèle une formulation pour les modes, et établir des équations d'évolution pour l'excitation des modes dans le taper. Une problème d'optimisation en dimension finie basée sur une discrétisation sortira enfin de tout cela.

Des exemples numériques basés sur cette méthode montre que si la discrétisation devient trop fine, la convergence ralentit et la solution devient de plus en plus instable. Le problème vient du fait que le problème d'optimisation est mal posé. L'article présente donc une méthode pour maximiser l'énergie résultante basée sur un problème inverse non linéaire, ainsi que des résultats numériques corroborant leur idée.

## 1 Formulation du problème direct en 2D

Les équations homogènes de Maxwell dans un milieu continu avec la permittivité linéaire  $\varepsilon$  et la perméabilité magnétique  $\mu_0$  sont :

$$\begin{aligned}\nabla \wedge H &= \varepsilon \dot{E} \\ \nabla \wedge E &= -\mu_0 \dot{H} \\ \nabla \bullet (\varepsilon E) &= \nabla \bullet H = 0\end{aligned}\tag{1}$$

On se concentre uniquement sur le cas 2D. On assume donc que  $E$  et  $H$  sont indépendants de  $y$ . Si on assume une dépendance périodique en temps  $e^{-i\omega t}$  et qu'on note  $E = (E_x, E_y, E_z)$  et  $H = (H_x, H_y, H_z)$  les composantes spatiales de la solution, on a une solution avec  $H_y = E_x = E_z = 0$  - le champ TE (transverse electric) - vérifiant :

$$\begin{aligned}\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} &= -i\omega\varepsilon E_y \\ \frac{\partial E_y}{\partial z} &= -i\omega\mu_0 H_x \\ \frac{\partial E_y}{\partial x} &= i\omega\mu_0 H_z\end{aligned}\tag{2}$$

En redérivant les deux dernières équations de (2) par rapport à respectivement  $z$  et  $x$  et en réintroduisant le résultat dans la première équation, on obtient l'équation d'Helmholtz pour  $E_y$  :

$$\Delta E_y + k^2 n^2 E_y = 0 \text{ avec } \omega\sqrt{\mu_0\varepsilon_0} = k \text{ et } n^2 = \frac{\varepsilon}{\varepsilon_0}$$

où  $\varepsilon_0$  désigne la permittivité du vide.

On remarque également qu'on obtient une solution avec  $E_y = H_x = H_z = 0$  -le champ magnétique transverse ('TM') - donnant une équation légèrement modifiée :

$$-\Delta H_y + k^2 n^2 H_y = 0$$

Cependant, l'analyse ne se poursuivra qu'avec les modes du champ TE.

Nous allons maintenant formuler le problème de propagation de l'onde dans un guide d'onde générique  $\Omega$  en notant la frontière  $\Gamma \cup \Gamma_R \cup \Gamma_L$  (voir la figure 1).

À présent, pour plus de commodité,  $E_y$  sera noté  $U$ . Sur le bord  $\Gamma$ , nous prenons des conditions de Dirichlet ( $U = 0$ ).

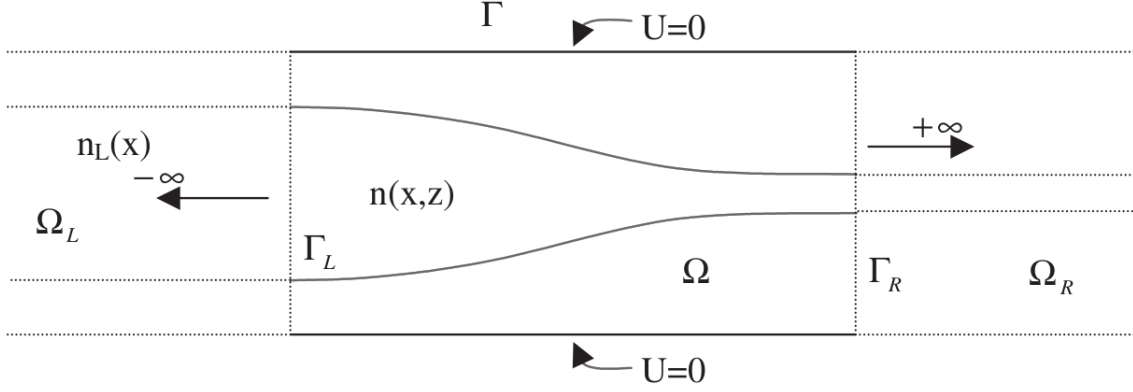


FIGURE 1 – Profil du taper

Physiquement, cela signifie que le champ est réfléchi vers  $\Omega$ . En pratique, cela n'est pas un problème si on cherche des solutions de champs guidés, car ils sont par définition limités au guide d'onde et décroissent exponentiellement en dehors (ce qui sera discuté plus en détail ultérieurement). On réécrit donc les équations :

$$\Delta U + n^2 U = 0 \text{ pour } (x, z) \in \Omega \quad (3)$$

$$U|_{\Gamma} = 0 \quad (4)$$

où on a posé  $k = 1$ . Cela n'est rien d'autre qu'une normalisation de l'équation : on pose simplement  $x' = kx$  et  $z' = kz$ . Ainsi :

$$(\partial_x^2 + \partial_z^2)U(x', z') = k^2(\partial_{x'}^2 + \partial_{z'}^2)U(x', z') = k^2 \Delta U(x', z')$$

D'où :

$$k^2 \Delta U + n^2 k^2 U = 0$$

On obtient l'équation précédente en simplifiant par  $k^2$ .

Nous avons besoin à présent de conditions à  $\pm\infty$ . Nous prenons comme hypothèse que nous avons un champ entrant venant de  $-\infty$ . Ce champ  $U_I$  satisfait également (3), (4) dans  $\Omega_L$ . Nous devons retrouver le fait que le champ diffusé  $U - U_I$  voyage vers l'extérieur des deux côtés du guide d'onde. Nous allons faire cela en terme de modes propres dans  $\Omega_L$  et  $\Omega_R$ , ce qui amène donc à considérer un problème spectral dans ces régions.

Dans  $\Omega_L$ , on prend une dépendance périodique par rapport à  $z$  :  $U(x, z) = \tilde{U}(x)e^{i\beta z}$ . En réinjectant cela dans (3), on obtient :

$$\frac{d^2 \tilde{U}}{dx^2} + (n^2 - \beta^2) \tilde{U} = 0 \text{ pour un } z \text{ fixé avec } \tilde{U}(x_{min}) = \tilde{U}(x_{max}) = 0 \quad (5)$$

Les conditions de Dirichlet (4) nous assure que nous avons en ensemble de valeurs propres discrètes  $\{(\tilde{U}_k, \pm\beta_k) : k \in \mathbb{N}^*\}$ . On a donc :

$$U(x, z) = \sum_{k=1}^{\infty} (C_k e^{i\beta_k z} + C_{-k} e^{-i\beta_k z}) \tilde{U}_k(x) \text{ dans } \Omega_L \quad (6)$$

Notons que le raisonnement tient toujours dans  $\Omega_R$  et donne un résultat similaire.

Dans (6), on pose  $\beta_k > 0$  si  $\beta_k$  est réel,  $\beta_k = i|\beta_k|$  si  $\beta_k$  est imaginaire. Ainsi, les coefficients  $C_k$  et  $C_{-k}$  représentent respectivement le voyage vers la droite et vers la gauche de l'onde.

De plus, on remarque que l'opérateur utilisé dans (3) est auto-adjoint pour le produit scalaire  $\langle a, b \rangle = \int_{x_{min}}^{x_{max}} a(x)b(x)dx$ .

En effet :

$$\begin{aligned}
\langle (\Delta + n^2 Id)u, v \rangle &= \langle \Delta u, v \rangle + \langle n^2 u, v \rangle \\
&= \int_{x_{min}}^{x_{max}} \Delta u(x) v(x) dx + \langle u, n^2 v \rangle \\
&= - \int_{x_{min}}^{x_{max}} \nabla u(x) \cdot \nabla v(x) dx + \underbrace{[\nabla u(x) \cdot n v(x)]_{x_{min}}^{x_{max}}}_{=0 \text{ dû au conditions de Dirichlet}} + \langle u, n^2 v \rangle \\
&= \int_{x_{min}}^{x_{max}} u(x) \Delta v(x) dx - \underbrace{[\nabla v(x) \cdot n u(x)]_{x_{min}}^{x_{max}}}_{=0} + \langle u, n^2 v \rangle \\
&= \langle u, (\Delta + n^2 Id)v \rangle
\end{aligned}$$

Cela assure ainsi que les modes propres sont orthogonaux relativement à ce produit scalaire. De plus, si on normalise les normes :

$$\langle \tilde{U}_k, \tilde{U}_k \rangle = \frac{1}{\beta_k} \quad (7)$$

la puissance transmise dans chaque mode propre  $k$  est donné par  $|C_k|^2$ .

Étant donné cette normalisation, une onde sortante  $U_L$  (vers  $-\infty$ ) est donnée en posant  $C_k = 0$  pour tout  $k > 0$ . Ainsi :

$$\begin{aligned}
\beta_k \langle U_L, \tilde{U}_k \rangle &= \beta_k \sum_{j=1}^{\infty} C_{-j} e^{-i\beta_j z} \langle \tilde{U}_j, \tilde{U}_k \rangle \\
&= \beta_k C_{-k} e^{-i\beta_k z} \frac{1}{\beta_k} \\
&= C_{-k} e^{-i\beta_k z}
\end{aligned}$$

De même, une onde entrante  $U_I$  est donnée en fixant  $C_{-k} = 0$  pour tout  $k > 0$  et  $C_k$  est donné par

$$C_k e^{i\beta_k z} = \beta_k \langle U_I, \tilde{U}_k \rangle$$

Enfin, on différencie (6) pour  $U_L$  et  $U_I$  pour obtenir les conditions au bord sur les côtés. On obtient ainsi :

$$\frac{\partial U_L}{\partial z} = \sum_{k=1}^{\infty} -i\beta_k C_{-k} e^{-i\beta_k z} \tilde{U}_k(x)$$

$$\frac{\partial U_L}{\partial z} = -i \sum_{k=1}^{\infty} \beta_k^2 \langle U_L, \tilde{U}_k \rangle \tilde{U}_k(x) \text{ dans } \Omega_L - \text{condition pour les ondes partant vers } -\infty \quad (8)$$

De même :

$$\frac{\partial U_I}{\partial z} = -i \sum_{k=1}^{\infty} \beta_k^2 \langle U_I, \tilde{U}_k \rangle \tilde{U}_k(x) \text{ dans } \Omega_L - \text{condition pour les ondes venant de } -\infty \quad (9)$$

Puisque  $U = U_I + U_L$ , en remplaçant  $U_L$  dans (8) et en utilisant (9), on obtient :

$$\frac{\partial U}{\partial z} = -i \sum_{k=1}^{\infty} \beta_k^{(L)2} \langle U, \tilde{U}_k^{(L)} \rangle \tilde{U}_k^{(L)} + 2i \sum_{k=1}^{\infty} \beta_k^{(L)2} \langle U_I, \tilde{U}_k^{(L)} \rangle \tilde{U}_k^{(L)}$$

En surmontant le tout de l'exposant  $(L)$  pour indiquer que le tout vient de la gauche, ie de  $-\infty$ .

De même, pour les ondes se propageant vers  $+\infty$ , et en supposant qu'il n'y a aucune onde venant de là, on obtient :

$$\frac{\partial U}{\partial z} = i \sum_{k=1}^{\infty} \beta_k^{(R)2} \langle U, \tilde{U}_k^{(R)} \rangle \tilde{U}_k^{(R)} \text{ dans } \Omega_R - \text{condition pour les ondes allant vers } +\infty$$

Le problème complet est donc : trouver  $U(x, z)$  tel que :

$$\left\{ \begin{array}{ll} \Delta U + n^2 U &= 0 & \text{pour } (x, z) \in \Omega \\ U|_{\Gamma} &= 0 & \text{(murs réfléchissants)} \\ \frac{\partial U}{\partial z} + i \sum_{k=1}^{\infty} \beta_k^{(L)2} \langle U, \tilde{U}_k^{(L)} \rangle \tilde{U}_k^{(L)} &= 2i \sum_{k=1}^{\infty} \beta_k^{(L)2} \langle U_I, \tilde{U}_k^{(L)} \rangle \tilde{U}_k^{(L)} & \text{sur } \Gamma_L \\ \frac{\partial U}{\partial z} - i \sum_{k=1}^{\infty} \beta_k^{(R)2} \langle U, \tilde{U}_k^{(R)} \rangle \tilde{U}_k^{(R)} &= 0 & \text{sur } \Gamma_R \end{array} \right. \quad (10)$$

où  $\{(\tilde{U}_k^{(L)}, \pm\beta_k^{(L)}) : k \in \mathbb{N}^*\}$ ,  $\{(\tilde{U}_k^{(R)}, \pm\beta_k^{(R)}) : k \in \mathbb{N}^*\}$  sont les modes propres dans  $\Omega_L$ ,  $\Omega_R$  respectivement.

## 2 Formulation du problème d'optimisation

En général, nous nous intéressons à la forme optimale du taper (ou la distribution des indices de refraction) qui dans un sens maximise la puissance transférée entre l'entrée et la sortie du guide d'onde.

Souvent, pour des raisons pratiques, on suppose que l'onde en entrée est une excitation de l'élément propre fondamental (indiqué donc par 1), et on s'intéresse à la puissance restante dans l'élément propre fondamental  $\tilde{U}_1^{(R)}$  de l'onde de sortie (on assume bien évidemment que les deux ondes admettent au moins un mode guidé). Dans la décomposition spectral (6), l'énergie restante correspond au coefficient  $|C_1|^2$ . Cela est donné par :

$$P(n^2) = \beta_1^2 |\langle U, \tilde{U}_1^{(R)} \rangle|^2 = \beta_1^2 \left| \int_{x \in \Gamma_R} U(x, z_R) \tilde{U}_1^{(R)}(x) dx \right|^2 \quad (11)$$

Intuitivement (et cela se vérifie, comme montré dans [2]), la puissance transférée est maximale (ie  $P = 1$ ) quand la longueur du taper tend vers l'infini. Il est donc évident qu'on doit imposer une contrainte supplémentaire sur la longueur finie du taper.

## 3 Solution du problème direct

### 3.1 Recherche d'une représentation locale

À cause des conditions aux bords que nous avons imposé, le système (10) est en général résolu en utilisant une décomposition spectrale.

On définit  $L_t(U) = \frac{\partial^2 U}{\partial x^2} + n^2(x, z)U$  l'opérateur 1D de Helmholtz.

Soit  $\Omega_z$  la section de la région au point  $z$ . On définit la base locale  $\{(U_k, \beta_k); k \geq 1\}$  à chaque position  $z$  par :

$$\begin{aligned} L_t(U_k) &= \beta_k^2 U_k \quad \text{dans } \Omega_z \\ U_k|_{\partial\Omega_z} &= 0 \end{aligned} \quad (12)$$

Cela est analogue à (5), donc nous parlerons des solutions  $U_k$  comme les fonctions propres locales. Le caractère auto-adjoint de l'opérateur  $L_t$  (comme on l'a prouvé précédemment) assure que les  $(U_k)$  forment une base orthogonale pour les fonctions  $f \in \mathcal{C}(\Omega_z)$  avec  $f|_{\partial\Omega_z} = 0$ . Ainsi, toute fonction  $U$  définie dans une région  $\Omega$  avec  $U|_{\partial\Omega_z} = 0$  peut être exprimé comme l'unique décomposition  $U = \sum_{k=1}^{\infty} c_k(z)U_k$  où les coefficients (uniques)  $c_k$  dépendent exclusivement de  $z$ .

On s'intéresse maintenant à une formulation qui donne une représentation locale du champ électromagnétique en terme de propagation vers  $-\infty$  ou  $+\infty$  des modes du champ EM. On choisit donc le développement de notre fonction  $U$  solution de (10) de la forme :

$$\begin{aligned} U &= \sum_{k=1}^{\infty} (a_k + a_{-k})U_k \\ \frac{\partial U}{\partial z} &= \sum_{k=1}^{\infty} (a_k - a_{-k})i\beta_k U_k \end{aligned} \quad (13)$$

Ce choix de coefficients nous assure que les  $a_k, a_{-k}$  correspondent aux champs se propageant vers  $+\infty$  ou  $-\infty$ . On réutilise la normalisation (7) pour les fonctions propres  $U_k$  définies par (12) :

$$\int_{\Omega_z} U_k^2 ds = \frac{1}{\beta_k} \quad (14)$$

Cela assure que la puissance contenue dans le  $k^{ieme}$  mode de l'onde se propageant de chaque côté est obtenu par  $|a_k|^2$  ou  $|a_{-k}|^2$  (voir [2] pour plus de détail).

On remarque en particulier que l'énergie contenue dans l'état fondamental à la sortie du guide d'onde est maintenant

$$P(n^2) = |a_1|^2 \quad (15)$$

En référence à l'expression (11).

En utilisant le fait que la dérivée par rapport à  $z$  de  $U$  en (13) doit être identique à la seconde ligne de (13), et en substituant (13) dans (3), on en déduit que les  $a_k$  doivent vérifier le système d'EDO suivant :

$$\dot{a}_k(z) - i\beta_k a_k(z) = \sum_{j \neq k, 0} r_{kj}(z) a_j(z), \quad k \neq 0 \quad (16)$$

avec  $\beta_{-k} = -\beta_k$  et

$$r_{kj}(z) = \frac{\int_{\Omega_z} \frac{\partial n^2}{\partial z} U_k U_j ds}{2(\beta_k - \beta_j)}$$

pour tout  $j \neq k, j, k \neq 0$

### 3.2 Démonstration

Précisons tout d'abord que nous allons noter pour plus de commodité  $U_{-k} = -U_k$ . Pour démontrer (16), on différencie (13) et on l'identifie à la seconde ligne, on obtient que  $a_k$  vérifie :

$$\sum_{k \neq 0} \dot{a}_k U_k + \sum_{k \neq 0} a_k \frac{\partial U_k}{\partial z} = \sum_{k \neq 0} a_k i \beta_k U_k$$

Par orthogonalité des  $U_k$ , on projette cette équation sur chaque  $U_i$ , et cela mène à :

$$\dot{a}_k + \dot{a}_{-k} = (a_k - a_{-k}) i \beta_k - \sum_{j \neq 0} a_j p_{jk} \quad \forall k \neq 0 \quad (17)$$

où  $p_{jk} = \beta_{|k|} \int_{\Omega_z} \frac{\partial U_j}{\partial z} U_k dx$  ne sont pas moins que les coefficients du développement de  $\frac{\partial U_j}{\partial z}$  dans la base des  $(U_k)$ . La première expression de (3) peut être réécrite

$$\frac{\partial^2 U}{\partial z^2} + L_t(U) = 0$$

Et donc, en utilisant (13) :

$$\frac{\partial}{\partial z} \left( \sum_{k \neq 0} a_k i \beta_k U_k \right) + L_t \left( \sum_{k \neq 0} a_k U_k \right) = 0$$

comme  $L_t$  est linéaire :

$$\begin{aligned} \Rightarrow & \quad \frac{\partial}{\partial z} \left( \sum_{k \neq 0} a_k i \beta_k U_k \right) + \sum_{k \neq 0} a_k L_t(U_k) = 0 \\ \Rightarrow & \quad \sum_{k \neq 0} \frac{d}{dz} (a_k i \beta_k) U_k + \sum_{k \neq 0} a_k i \beta_k \frac{\partial U_k}{\partial z} + \sum_{k \neq 0} a_k \beta_k^2 U_k = 0 \end{aligned}$$

En projetant sur la base des  $(U_k)$  :

$$\begin{aligned} \Rightarrow & \quad \frac{d}{dz} (a_k i \beta_k + a_{-k} i \beta_{-k}) \frac{1}{\beta_{|k|}} + \sum_{j \neq 0} a_j i \frac{\beta_j}{\beta_{|k|}} p_{jk} + (a_k + a_{-k}) \frac{\beta_k^2}{\beta_{|k|}} = 0 \quad \forall k > 0 \\ \Rightarrow & \quad \frac{d}{dz} (a_k i \beta_k + a_{-k} i \beta_{-k}) + \sum_{j \neq 0} a_j i \beta_j p_{jk} + (a_k + a_{-k}) \beta_k^2 = 0 \\ \Rightarrow & \quad \frac{d}{dz} [i \beta_k (a_k - a_{-k})] + \sum_{j \neq 0} a_j i \beta_j p_{jk} + (a_k + a_{-k}) \beta_k^2 = 0 \\ \Rightarrow & \quad i \dot{\beta}_k (a_k - a_{-k}) + i \beta_k (\dot{a}_k - \dot{a}_{-k}) + \sum_{j \neq 0} a_j i \beta_j p_{jk} + (a_k + a_{-k}) \beta_k^2 = 0 \end{aligned}$$

Avec (17), on obtient donc :

$$\begin{aligned} \dot{a}_k + \dot{a}_{-k} &= (a_k - a_{-k}) i \beta_k - \sum_{j \neq 0} a_j p_{jk} & \forall k \neq 0 \\ \dot{a}_k - \dot{a}_{-k} &= (a_k + a_{-k}) i \beta_k - (a_k - a_{-k}) \frac{\dot{\beta}_k}{\beta_k} - \sum_{j \neq 0} a_j \frac{\beta_j}{\beta_k} p_{jk} & \forall k \neq 0 \end{aligned} \quad (18)$$

On somme et on fait la différence des deux équations de (18), et on obtient :

$$\begin{aligned} \dot{a}_k - i \beta_k a_k &= -\frac{1}{2} (a_k - a_{-k}) \frac{\dot{\beta}_k}{\beta_k} - \frac{1}{2 \beta_k} \sum_{j \neq 0} p_{jk} a_j (\beta_k + \beta_j) \\ \dot{a}_{-k} + i \beta_k a_k &= -\frac{1}{2} (a_{-k} - a_k) \frac{\dot{\beta}_k}{\beta_k} - \frac{1}{2 \beta_k} \sum_{j \neq 0} p_{jk} a_j (\beta_k - \beta_j) \end{aligned} \quad (19)$$

On peut réécrire le tout en une seule équation :

$$\dot{a}_k - i \beta_k a_k = -\frac{\dot{\beta}_k}{2 \beta_k} (a_k - a_{-k}) - \frac{1}{2 \beta_k} \sum_{j \neq 0} p_{jk} a_j (\beta_k + \beta_j) \quad (20)$$

On va maintenant exprimer plus précisément  $\dot{\beta}$  et  $p_{jk}$ . On rappelle que par définition :

$$\frac{\partial U_k}{\partial z} = \sum_{j=1}^{\infty} p_{kj} U_j$$

En différentiant en utilisant (12), on obtient :

$$\begin{aligned} L_t \left( \frac{\partial U_k}{\partial z} \right) &= \frac{\partial^2}{\partial x^2} \left( \frac{\partial U_k}{\partial z} \right) + n^2 \frac{\partial U_k}{\partial z} \\ \frac{\partial}{\partial z} L_t (U_k) &= \frac{\partial^2}{\partial x^2} \left( \frac{\partial U_k}{\partial z} \right) + n^2 \frac{\partial U_k}{\partial z} + \frac{\partial n^2}{\partial z} U_k \\ \Rightarrow L_t \left( \frac{\partial U_k}{\partial z} \right) &= \frac{\partial}{\partial z} L_t (U_k) - \frac{\partial n^2}{\partial z} U_k = \frac{\partial}{\partial z} (\beta_k^2 U_k) - \frac{\partial n^2}{\partial z} U_k \\ \Rightarrow L_t \left( \frac{\partial U_k}{\partial z} \right) &= \beta_k^2 \frac{\partial U_k}{\partial z} + \frac{\partial}{\partial z} (\beta_k^2 - n^2) U_k \end{aligned}$$

On a donc, en reprenant la décomposition spectrale de  $\frac{\partial U_k}{\partial z}$  et par linéarité de  $L_t$  :

$$\begin{aligned} L_t \left( \sum_{j=1}^{\infty} p_{kj} U_j \right) &= \beta_k^2 \sum_{j=1}^{\infty} p_{kj} U_j + \frac{\partial}{\partial z} (\beta_k^2 - n^2) U_k \\ \Rightarrow \sum_{j=1}^{\infty} p_{kj} (\beta_j^2 - \beta_k^2) U_j &= \frac{\partial}{\partial z} (\beta_k^2 - n^2) U_k \\ \Rightarrow p_{kj} (\beta_j^2 - \beta_k^2) &= \beta_j \int_{\Omega_z} \frac{\partial}{\partial z} (\beta_k^2 - n^2) U_k U_j dx \end{aligned}$$

en projetant sur  $U_k$ . Si  $j \neq k$ , en utilisant l'orthogonalité des fonctions propres, on a :

$$p_{kj} = \frac{\beta_j}{(\beta_k^2 - \beta_j^2)} \int_{\Omega_z} \frac{\partial}{\partial z} (\beta_k^2 - n^2) U_k U_j dx \quad (21)$$

Si  $j = k$ , le membre de gauche s'annule, et on a donc une condition sur les  $\beta_k$  :

$$\frac{\partial \beta_k}{\partial z} = \frac{1}{2\beta_k} \int_{\Omega_z} \frac{\partial n^2}{\partial z} U_k^2 dx \quad (22)$$

$p_{kk}$  est déterminé par la normalisation des  $U_k$  :

$$p_{kk} = \beta_k \int_{\Omega_z} \frac{\partial U_k}{\partial z} U_k dx = \frac{\beta_k}{2} \int_{\Omega_z} \frac{\partial U_k^2}{\partial z} dx = \frac{\beta_k}{2} \frac{\partial}{\partial z} \int_{\Omega_z} U_k^2 dx = \frac{\beta_k}{2} \frac{\partial \beta_k^{-1}}{\partial z} = -\frac{1}{2\beta_k} \frac{\partial \beta_k}{\partial z}$$

Donc :

$$p_{kk} = -\frac{\dot{\beta}_k}{2\beta_k} \quad (23)$$

À présent, dans (20), on remplace tous les  $p$  en utilisant (21) et (23) et on remplace le terme  $\dot{\beta}$  en utilisant (22). Le terme  $j = k$  dans la somme se simplifie avec le premier terme dans le membre de droite, et on peut réécrire (20) comme :

$$\begin{aligned} \forall k \neq 0, \dot{a}_k - i\beta_k a_k &= -\frac{\dot{\beta}_k}{2\beta_k} (a_k - a_{-k}) - \frac{1}{2\beta_k} \sum_{j \neq 0} p_{jk} a_j (\beta_k + \beta_j) \\ &= -\frac{\dot{\beta}_k}{2\beta_k} (a_k - a_{-k}) + \frac{1}{2\beta_k} a_k 2\beta_k \frac{\dot{\beta}_k}{2\beta_k} - \frac{1}{2\beta_k} \sum_{j \neq 0, k} p_{jk} a_j (\beta_k + \beta_j) \\ &= \frac{\dot{\beta}_k}{2\beta_k} a_{-k} - \frac{1}{2\beta_k} \sum_{j \neq 0, k} p_{jk} a_j (\beta_k + \beta_j) \\ &= \frac{a_{-k}}{4\beta_k} \int_{\Omega_z} \frac{\partial n^2}{\partial z} U_k^2 dx - \frac{1}{2\beta_k} \sum_{j \neq 0, k} \frac{\beta_k}{\beta_j^2 - \beta_k^2} (\beta_k + \beta_j) a_j \int_{\Omega_z} \frac{\partial n^2}{\partial z} U_k U_j dx \\ &= \frac{a_{-k}}{4\beta_k} \int_{\Omega_z} \frac{\partial n^2}{\partial z} U_k^2 dx + \sum_{j \neq 0, k} a_j \frac{1}{2(\beta_k - \beta_j)} \int_{\Omega_z} \frac{\partial n^2}{\partial z} U_k U_j dx \end{aligned}$$

Finalement, on inclut les termes les termes  $a_{-k}$  comme le terme  $-k$  de la somme et nous arrivons à l'équation (16).

### 3.3 Équation d'évolution de l'amplitude

Pour compléter l'équation (16), on a également des conditions au début ( $z = 0$ ) et à la fin du guide d'onde ( $z = z_R$ ), qui se traduisent par :

$$a_k(0) = A_k; \quad a_{-k}(z_R) = 0 \quad \forall k > 0 \quad (24)$$

où les  $A_k$  sont les coefficients correspondant au champ d'entrée  $U_I$ .

L'équation (16) peut se réécrire, on notant  $\gamma_k(z) = e^{-i \int \beta_k(z) dz} a_k(z)$  :

$$\dot{\gamma}_k(z) = \sum_{j \neq 0, k} r_{kj}(z) e^{i \int_0^z (\beta_j(z) - \beta_k(z)) dz} \gamma_j(z) \quad (25)$$

C'est l'équation d'évolution des amplitudes  $\gamma_k(z)$ . Notons que :

1. Le membre de droit ne dépend pas de  $\gamma_k$ , ce qui signifie que l'équation décrit le couplage entre les différents modes
2. Dans un guide d'onde de section constante,  $r_{kj} = 0$ , donc  $\gamma_k(z)$  est constant, comme attendu.
3. Une borne supérieure pour le couplage est donné par  $r_{kj} \leq \frac{\int_{\Omega_z} \left| \frac{\partial n^2}{\partial z} \right| ds}{2|\beta_k - \beta_j|}$ , et donc le couplage tend à être plus fort avec les voisins immédiats, pour lesquels  $|\beta_k - \beta_j|$  est petit.

### 3.4 Précisions sur les conditions aux bords

Les conditions aux bords (4) sont en un sens "artificielles". Si nous utilisions plutôt des conditions plus vraisemblables, les modes locaux seraient divisés en un ensemble discret de modes guidés et en une continuité de modes dû à la radiation. Les équations (13) auraient alors une composante intégrale, ce qui est difficile à analyser analytiquement et numériquement.

C'est pour cela que des "murs" sont utilisés ici. Comme on l'a vu, on obtient ainsi un ensemble discret de modes guidés, qui ne sont pratiquement pas affectés par cette approximation puisqu'ils décroissent exponentiellement vers 0 à l'extérieur du taper, et un ensemble discret de modes dû à la radiation. Bien évidemment, toute radiation du taper sera réfléchi par le bord et pourrait se coupler avec les modes guidés dans le taper. Cela affecterait donc la puissance que nous cherchons justement à optimiser ! Cependant, cet effet sera négligeable si le ration entre la distance entre le bord et le taper et la longueur totale du taper est assez grande pour empêcher le champ réfléchi d'atteindre le cœur du taper avant sa fin.

## 4 Solution numérique du problème d'optimisation

### 4.1 Choix de la discrétisation

Le problème posé précédent a été discrétisé de la manière suivante : la paramétrisation du profil du taper a été divisé en  $N$  sous-sections équidistantes. L'indice de réfraction dans chaque sous-section est contrôlé par un nombre fini de paramètres. Les paramètres à optimiser ici sont les hauteurs du guide d'onde  $\{w_1, \dots, w_N\}$  (voir figure 2).

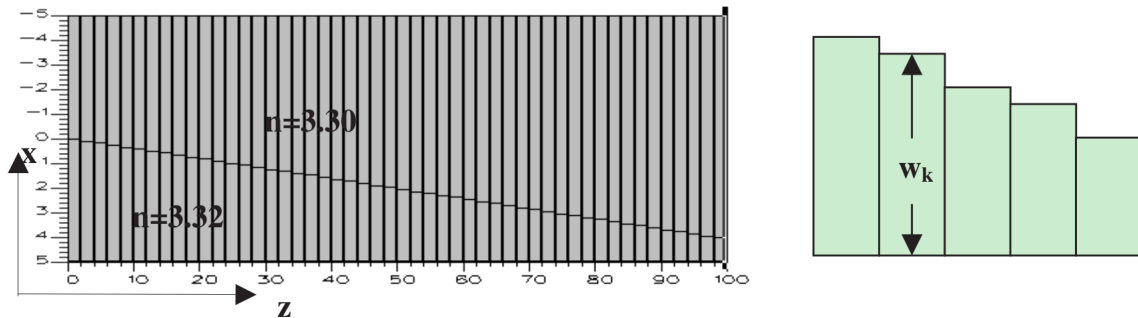


FIGURE 2 – Profil du taper : discrétisation pour le problème d'optimisation



Cette discrétisation impose une régularité très faible pour le profil du taper. On pourrait en effet obtenir des solution avec un profil totalement discontinu. Le champ est calculé grâce à l'analyse faite sur les modes locaux. Les guides d'onde en entrée et en sortie sont choisis tels qu'ils aient au moins un mode guidé. Les dérivées requises sont calculées via un schmé de type différences finies, avec  $N = 13$  ou 48 et une solution initiale qui est une droite comme montrée figure 2.

## 4.2 Résultats et interprétation physique

Les résultats sont assez étonnants : en effet, on pense en général que le profil devrait plutôt ressembler à une courbe assez lisse. Cependant, comme on peut le voir figure 3, le résultat n'est pas vraiment ce à quoi on s'attendait. Ces résultats ont été obtenus en 2900 itérations.

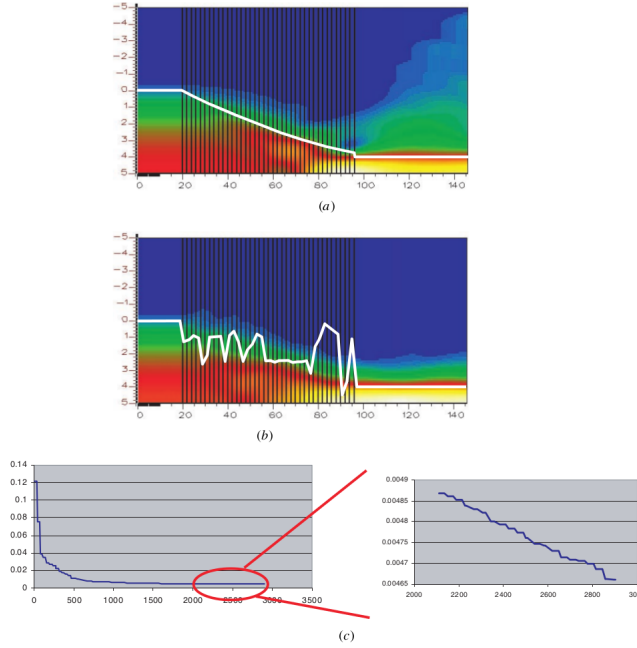


FIGURE 3 – Profil du taper : résultat avec  $N = 48$ . (a) Forme initial du taper (b) Forme optimale (c) Perte d'énergie en fonction du nombre d'itérations.

On remarque dans la figure 4 que le mécanisme d'optimisation repose surtout sur une résonance entre les modes. On remarque que le profil est fait de petits pics qui se répètent tous les  $10\mu m$ . Cela fait que l'énergie reste presque

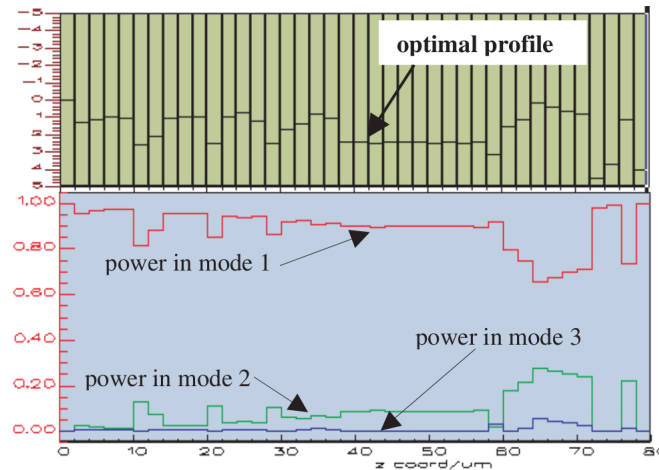


FIGURE 4 – L'énergie est préservée grâce à la résonance avec les autres modes

entièrement dans les deux premiers modes. Après un passage assez plat, le saut final juste avant la fin réintroduit toute la puissance du mode 2 dans le mode 1. La perte d'énergie totale est de seulement 0.5% en  $80\mu\text{m}$ , tandis qu'un taper plus classique aurait besoin, pour avoir la même efficacité, d'être au moins 5 fois plus long. On ne sait pas, bien sûr, si c'est la solution optimale, mais cette perte d'énergie faible est bien plus importante. Cependant, ces résultats soulèvent une autre problématique : le problème est mal posé.

### 4.3 Un problème mal posé

Les résultats précédents montrent une certaine instabilité au niveau de l'optimisation du profil. Pire encore, cette instabilité semble croître encore avec le nombre de sous-sections  $N$ .

Cela s'explique par le fait qu'on peut toujours trouver une variation arbitrairement grande de la solution qui donnerait toujours un champ  $U$  arbitraire assez proche. La solution optimale est donc instable par rapport aux variations de l'indice de réfraction du profil. On peut voir cela ainsi : supposons que nous avons une solution avec  $n^2$  de notre problème d'optimisation, et appliquons une perturbation  $\delta n^2$  à  $n^2$ . Le champ résultant  $U$  subira également une perturbation  $\delta U$  qui est donné au premier ordre par (10) :

$$\begin{aligned} \Delta \delta U + n^2 \delta U &= -U \delta n^2 & (x, z) \in \Omega \\ \delta U|_{\Gamma} &= 0 \\ \frac{\partial \delta U}{\partial z} + i \sum_{k=1}^{\infty} \beta_k^{(L)} \langle \delta U, \tilde{U}_k^{(L)} \rangle \tilde{U}_k^{(L)} &= 0 & \text{sur } \Gamma_L \\ \frac{\partial \delta U}{\partial z} - i \sum_{k=1}^{\infty} \beta_k^{(R)} \langle \delta U, \tilde{U}_k^{(R)} \rangle \tilde{U}_k^{(R)} &= 0 & \text{sur } \Gamma_R \end{aligned} \quad (26)$$

Il est souvent utile de réexprimer (26) sous forme intégrale : on introduit donc la fonction de Green  $G(r, r')$  avec  $r = (x, z)$  et  $r' = (x', z')$ .

Fonction de Green : Si on note  $\mathfrak{D}$  un opérateur différentiel linéaire, on note  $G(x)$  toute solution de l'équation :

$$\mathfrak{D}G(x) = \delta(x)$$

où  $\delta(x)$  note ici la distribution de Dirac.

Pour en revenir à notre problème, la fonction de Green  $G(r, r')$  est exprimée par :

$$\begin{aligned} \Delta G(r, r') + n(r)^2 G(r, r') &= \delta(r - r') & r, r' \in \Omega \\ G(r, r')|_{r \in \Gamma} &= 0 & r \in \Omega \\ \frac{\partial G(r, r')}{\partial z} + i \sum_{k=1}^{\infty} \beta_k^{(L)} \langle G, \tilde{U}_k^{(L)} \rangle \tilde{U}_k^{(L)} &= 0 & \text{sur } \Gamma_L \\ \frac{\partial G(r, r')}{\partial z} - i \sum_{k=1}^{\infty} \beta_k^{(R)} \langle G, \tilde{U}_k^{(R)} \rangle \tilde{U}_k^{(R)} &= 0 & \text{sur } \Gamma_R \end{aligned} \quad (27)$$

Le champ  $\delta U$  est ainsi donné par l'expression :

$$G \Delta \delta U - \delta U \Delta G \equiv \nabla \cdot (G \nabla \delta U - \delta U \nabla G)$$

[... Passage incompris...]

On a donc :

$$\delta U(r) = - \int_{r' \in \Omega} G(r, r') U(r') \delta n^2(r') dv := K(\delta n^2) \quad (28)$$

C'est une relation intégrale reliant  $\delta U$  et  $\delta n^2$ . Si on considère  $\delta U$  donné, (28) peut être vu comme une équation intégrale (de Fredholm) du premier type pour  $\delta n^2$  avec l'opérateur intégral :

$$K : \delta n^2 \mapsto - \int_{z \in \Omega} G(r, r') U(r') \delta n^2(r') dv$$

qui est, en tant qu'opérateur à noyau, compact dans l'espace des fonctions continues  $\mathcal{C}^0(\Omega)$ .

Selon [1], les équations intégrales du premier ordre à opérateur compact débouchent sur des problèmes mal posés, dans le sens que pour tout  $\delta U$  admissible correspondant à une solution  $\delta n^2$ , pour tout  $\varepsilon$  petit, il existe une autre fonction  $\delta n_1^2$  assez éloignée de  $\delta n^2$  (mais typiquement, différent de  $\delta n^2$  uniquement sur un petit ensemble) tel que  $\|K(\delta n_1^2) - \delta U\| < \varepsilon$

On en vient donc aux remarques suivantes :

1. Puisque les fonctions objectifs et les contraintes sur les deux problèmes d'optimisation précédents dépendent explicitement et continuellement du champ  $U$ , et non de l'indice de réfraction  $n^2$ , de petites variations dans le champ impliquent de petites variations dans les contraintes et les fonctions objectifs. Donc, si  $n^2$  est une solution optimale, alors il existe  $n_1^2$  aussi éloigné de  $n^2$  que l'on veut qui minimise la fonction objectif et qui satisfait également les contraintes à une précision arbitraire.
2. Tout schéma d'optimisation qui utilise d'une quelconque manière la linéarisation précédente pour trouver une direction locale de descente rencontrera des instabilités si on n'ajoute aucune présomption sur l'espace dans lequel on recherche la solution, c'est-à-dire si on cherche une solution optimale  $n^2$  dans tout  $\mathcal{C}^0(\Omega)$ .
3. D'un point de vue physique, ce caractère mal-posé du problème vient du fait que la puissance transmise n'est pas sensible à de grands changements de l'indice de réfraction sur un petit intervalle (comme un pic). En effet, l'onde  $U$  "ne voit pas" les variations dans l'indice de réfraction qui sont plus petites que la longueur d'onde de  $U$ .
4. On pourrait gagner en stabilité en limitant notre recherche à un espace de fonction plus petit. Par exemple, on pourrait chercher la solution optimale  $n^2$  dans un sous-ensemble borné de  $\mathcal{C}^1(\Omega)$ , ce qui éviterait les pics sus-mentionnés. L'équation (28) deviendrait alors une fonction  $K : \mathcal{C}^1(\Omega) \rightarrow K(\mathcal{C}^1(\Omega))$  qui aurait un inverse (s'il existe) dans  $\mathcal{C}^0(\Omega)$  dû à la compacité de  $\mathcal{C}^0(\Omega)$  dans  $\mathcal{C}^1(\Omega)$ . Si cet inverse n'existe pas, on peut toujours considérer l'inverse généralisé et calculer de manière stable en troncant les valeurs singulières (voir [1]). Ces considérations seront la base de l'algorithme alternatif décrit juste après.

Pour terminer cette discussion, nous allons discuter le rôle que joue le fait que ce problème soit mal posé dans ce problème.

On considère le problème d'optimisation qui est de trouver une (et pas la) distribution des indices de réfraction remplissant un certain nombre de critères. Dans un tel problème, l'incertitude n'est pas importante, ni même la stabilité de la solution optimale, puisqu'on souhaite principalement remplir les critères d'optimalité. A contrario, on pourrait aussi considérer le problème inverse, qui est d'identifier la distribution des indices de réfraction provenant d'une certaine mesure. Ce problème inverse est un problème d'identification de paramètre pour une équation aux dérivées partielles et donc mal-posé dans le sens que pour aucune norme disons raisonnable, sa solution (si elle est unique) ne sera pas stable au regard des perturbations sur les données (voir [1]). Mathématiquement, les deux problèmes sont équivalents - seuls la manière de voir le problème est différente.

## Références

- [1] H.W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Mathematics and Its Applications. Springer, 1996.
- [2] A.W. Snyder and J. Love. *Optical Waveguide Theory*. Science paperbacks. Springer, 1983.