



# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'université Toulouse III - Paul Sabatier  
Discipline ou spécialité : *Mathématiques Appliquées*

---

Présentée et soutenue par

**Sophie LAURENS**

Le 1 mars 2010

### ***Approximation de Haute Précision des Problèmes de Diffraction***

---

#### JURY

**Hélène BARUCQ**  
**Bruno DESPRES**  
**Philippe HELLUY**  
**Stéphane LANTERI**  
**Pierre-Alain MAZET**  
**Jean-Pierre RAYMOND**  
**François ROGIER**

*DR, Université de Pau*  
*PR, Université Paris VI*  
*PR, Université de Strasbourg*  
*DR, INRIA, Sophia Antipolis*  
*PAST, Université Toulouse III*  
*PR, Université Toulouse III*  
*DR, ONERA, Toulouse*

**Examineur**  
**Rapporteur**  
**Examineur**  
**Rapporteur**  
**Directeur de thèse**  
**Président du jury**  
**Examineur**

---

**Ecole doctorale :** *Mathématiques Informatique et Télécommunications de Toulouse*  
**Unité de recherche :** *Institut de Mathématiques de Toulouse - UMR 5219*  
**Directeur(s) de Thèse :** *Pierre-Alain MAZET*  
**Rapporteurs :** *Bruno DESPRES - Stéphane LANTERI*



# Remerciements

Il va de soi que mes remerciements vont en premier lieu à mon directeur de thèse, Pierre-Alain Mazet. Ce passionné à la culture mathématique sans limite m'aura transmis sa passion pour la recherche en aiguisant ma curiosité. La thèse qu'il m'a proposée a été un véritable défi que je ne regrette pas, et malgré les inévitables moments de déprime et les échecs numériques, je compte bien m'entêter et rester dans cette voie.

Merci à mes rapporteurs Bruno Desprès et Stéphane Lanteri d'avoir accepté de relire ce travail. Ils m'auront permis d'améliorer considérablement ce manuscrit grâce à leurs remarques toujours pertinentes. Merci également à tous les membres de mon jury qui se sont déplacés des quatre coins de la France pour juger cette thèse.

Je remercie également François Rogier, mon directeur d'unité, et Bernard Lécussan, directeur de département, de m'avoir accueillie au sein de leurs équipes. L'ONERA est un cadre très agréable pour faire ses premiers pas dans la recherche. Ayant été locataire du premier étage, je remercie en même temps que les chercheurs de mon équipe ceux du DPRS, nos partenaires de bureaux musicaux. En tant que doctorante, j'ai cotoyé nombre de stagiaires et de thésards de divers départements. Grâce à eux, les repas du midi ont été fort agréables. Tout ce beau monde va tellement me manquer ! En particulier, merci à mon coach technique, pour le thème beamer de ma soutenance qu'il m'a aidé à élaborer ainsi que tous les conseils sur les subtilités de LaTeX et emacs, mon coach perso, pour les remontants divers et variés (notamment l'apparition magique de Kinder Bueno sur mon bureau), et mon coach sportif, pour la découverte du jonglage. Et surtout, merci à celui avec qui j'ai partagé pendant plus de deux ans bien plus qu'un même bureau et certains déplacements, du CANUM au fin fond de la Vendée au SMAI de l'arrière pays niçois. Sa contribution à cette thèse est plus qu'importante.

Un grand merci aussi à « ma grosse ». Rencontrée en DEA, nous aurons partagé tellement de choses ensemble, des formations CIES aux galères administratives. Grâce à elle, j'ai découvert le monde du contrôle, et du coupage d'épsilon en quatre ! Merci à mes amis de Supaéro et plus généralement de Toulouse, avec une mention toute spéciale pour les filles (8 ans que ça dure !) pour tous les bons moments toulousains, parisiens, montpellierains, montalbanais, munichois passés ensemble, sans oublier tous ceux à venir, notamment à Bordeaux.

J'ai gardé l'essentiel pour la fin. Merci à mes proches qui n'ont eu de cesse de m'encourager et de me soutenir, en particulier yaa, et surtout, mes parents. Il est impossible de décrire ou d'expliquer tout ce qu'ils ont fait pour moi. Sans eux, ce manuscrit n'aurait jamais vu le jour. Merci.



# Table des matières

<b>Table des matières</b>	<b>5</b>
<b>1 Rappels sur les problèmes de propagation d’ondes écrits sous forme de systèmes de Friedrichs</b>	<b>5</b>
1.1 Enoncés des problèmes considérés . . . . .	5
1.1.1 Les Equations de Maxwell . . . . .	5
1.1.2 Les Equations des Ondes convectées . . . . .	6
1.2 Problèmes de diffraction considérés . . . . .	11
1.2.1 Cas instationnaire . . . . .	12
1.2.2 Cas harmonique . . . . .	13
1.3 Résultats généraux d’existence et d’unicité . . . . .	15
1.3.1 Rappels sur les systèmes de Friedrichs . . . . .	15
1.3.2 Cas de problèmes à coefficients constants avec $\Omega = \mathbb{R}^3$ . . . . .	22
1.3.3 Conclusions . . . . .	26
<b>2 Les Perfectly Matched Layers</b>	<b>27</b>
2.1 Introduction . . . . .	27
2.1.1 Historique . . . . .	27
2.1.2 Les formulations convexes classiques . . . . .	29
2.1.3 Le cas monodimensionnel . . . . .	33
2.1.4 Les PML de Bérenger revisités . . . . .	33
2.1.5 Pourquoi le changement de variable complexe semble la piste la plus naturelle ? . . . . .	35
2.2 Obtention d’un calcul différentiel intrinsèque sur la variété . . . . .	37
2.2.1 Rappels de géométrie différentielle . . . . .	37
2.2.2 Formalisme des PML . . . . .	40
2.2.3 Problème de Helmholtz sur $\mathbf{M}$ . . . . .	41
2.3 Résolution des Equations de Maxwell . . . . .	45
2.3.1 Cas des Equations de Maxwell avec un second membre non nul et $\mathbf{O} = \emptyset$ . . . . .	47
2.3.2 Cas des Equations de Maxwell sur $\mathbf{M}$ avec $\mathbf{O} \neq \emptyset$ et sans second membre (problème de scattering pour un obstacle parfaitement conducteur) . . . . .	48
2.4 Formulation harmonique . . . . .	50
2.4.1 Ecriture intrinsèque des équations . . . . .	50

2.4.2	Cas particulier des PML cartésiennes et cylindriques . . . . .	52
2.4.3	Existence et unicité de la solution au problème harmonique . . .	52
2.5	Formulations PML harmoniques . . . . .	60
2.5.1	Une première formulation des domaines étoilés . . . . .	60
2.5.2	Formulation hybride pour un domaine de forme $L$ . . . . .	63
2.5.3	Amélioration de cette formulation . . . . .	67
2.6	Relèvement instationnaire . . . . .	70
2.6.1	Formulation étoilée . . . . .	70
2.6.2	Formulation hybride . . . . .	75
2.6.3	Deuxième formulation hybride . . . . .	78
2.6.4	Conclusions et généralisation . . . . .	80
2.7	Equation des Ondes . . . . .	85
2.7.1	Formulation étoilée . . . . .	87
2.7.2	Formulation hybride . . . . .	90
2.7.3	Deuxième formulation hybride . . . . .	92
2.7.4	Généralisation . . . . .	93
2.8	Etude des formulations PML « naïves » . . . . .	96
2.9	Cas des Equations d'Euler Linéarisées Barotropes . . . . .	101
2.10	Simulations Numériques . . . . .	106
2.10.1	Relèvement des conditions aux limites . . . . .	106
2.10.2	Formulation variationnelle du problème . . . . .	109
2.10.3	Problème harmonique . . . . .	111
2.10.4	Problème instationnaire . . . . .	117
<b>3</b>	<b>Une nouvelle Approximation de Haute Précision</b>	<b>121</b>
3.1	Méthodes de Galerkin Discontinu : complexité et conditions CFL . . .	121
3.1.1	Calcul de la condition CFL . . . . .	122
3.1.2	Etude de la complexité d'un schéma de type Runge-Kutta couplé à une méthode de type Galerkin Discontinu . . . . .	135
3.1.3	Autres limites des méthodes de Galerkin Discontinu . . . . .	139
3.2	Les « Volumes Finis aux dérivées » . . . . .	140
3.2.1	Equation d'Advection monodimensionnelle . . . . .	142
3.2.2	Systèmes de Friedrichs à coefficients constants . . . . .	148
3.2.3	Généralisation de la méthode en bidimensionnel . . . . .	155
3.3	Simulations numériques . . . . .	159
	<b>Bibliographie</b>	<b>179</b>

# Introduction

Il est bien connu que la simulation numérique du problème de diffraction se heurte à deux difficultés : le domaine où l'on doit rechercher la solution est non borné dans l'espace ou dans l'espace temps (suivant que l'on recherche la solution dans le domaine fréquentiel ou temporel) et la rapidité des variations de la solution (d'autant plus importante que les fréquences des sources sont élevées). Bien que la mathématique concernant les principaux propagateurs soit parfaitement connue, ces deux points contribuent à la difficulté de leur simulation numérique.

Les méthodes intégrales réduisant d'une unité la dimension effective du domaine de résolution souffrent par ailleurs d'une inadaptation à des obstacles fortement hétérogènes et d'une implémentation délicate dans le domaine temporel. Pour pallier ces difficultés, deux voies, parmi d'autres, peuvent être considérées : borner artificiellement le domaine au plus près de l'obstacle par des conditions aux limites de type Absorbing Boundary Conditions ou un milieu fictif assurant une rapide décroissance du noyau de Schwartz de l'opérateur type Perfectly Matched Layers (PML), d'autre part les rapides oscillations de la solution peuvent être approximées avec plus d'efficacité par des schémas (structurés ou déstructurés) d'ordre « élevé ».

Ce travail s'inscrit dans le Projet de Recherche Fédérateur MAHPSO (Modèles d'Approximation de Haute Précision pour les Systèmes d'Ondes) de l'ONERA. Dans ce projet, on s'intéresse à la modélisation fine des phénomènes de propagation d'ondes pour des applications concernant la mécanique des fluides, la mécanique des structures, l'électromagnétisme et les plasmas. Il s'agit de réaliser un saut quantitatif dans la précision des schémas numériques de façon à permettre à l'ONERA de répondre avec succès aux nouveaux enjeux que posent la modélisation des phénomènes linéaires et non linéaires en propagation d'ondes. Ce projet concerne, d'une part, le domaine de l'approximation numérique des équations modélisant la propagation d'ondes et, d'autre part, le calcul de haute performance. Ce travail de thèse vient en appui du PRF MAHPSO sur les aspects théoriques et sur l'évaluation des méthodes. Elle a pour but de mener un travail de fond sur les aspects concernant les milieux fictifs absorbants de type PML et sur l'évaluation de fonctions de bases spécifiques pour l'amélioration de la précision des méthodes de Galerkin Discontinu.

## Contexte

Depuis plusieurs années on assiste à une utilisation de plus en plus importante des milieux fictifs absorbants pour réduire les domaines de calcul. Cette théorie, initiée par Bérenger, est une généralisation formelle du fait que la solution en monodimensionnel dans un domaine incluant l'obstacle n'est pas perturbée par l'introduction d'un opérateur de multiplication positif idoine dont le support est extérieur au domaine (ML de Bérenger). Pour généraliser cette technique en multidimensionnel, une idée naturelle (PML de Bérenger) a consisté alors à augmenter le nombre d'inconnues du problème de façon à ce que la partie spatiale de l'opérateur ne contienne, pour chaque équation scalaire qu'une seule direction de dérivation spatiale. Bien que les résultats numériques obtenus à partir de ce subterfuge se soient révélés pertinents, les fondements mathématiques de cette méthode sont pour le moins incertains : la partie principale de l'opérateur ainsi obtenu est faiblement hyperbolique et la perturbation par l'opérateur d'ordre zéro conduit à un problème  $L^2$  mal posé. Une autre interprétation suggérée par Chew et Wheedon dans le domaine fréquentiel du système ainsi obtenu conduit à remplacer les dérivations spatiales de l'opérateur par des combinaisons complexes de celles-ci. Le déterminant du symbole de l'opérateur de certaines équations (Maxwell par exemple) étant elliptique, et donc admettant un prolongement holomorphe en dehors de zéro, on pouvait inférer que la clé de la compréhension des domaines fictifs absorbants consistait à reconsidérer le noyau de Schwartz à partir d'un tel opérateur spatial. Sa décroissance pouvait être rendue aussi rapide que l'on voulait (sur exponentielle)... En l'état, et indépendamment de la forme de l'obstacle, le domaine de calcul demeurerait parallélépipédique. Un peu plus tard cette technique a été étendue à des domaines convexes en « complexifiant » les dérivations le long des normales de celui-ci.

Divers résultats d'existence d'unicité et de convergence des approximations (Colino, Lassas & al, Mazet & al) ont été obtenus dans le domaine fréquentiel, bien que le système ne soit pas de Friedrichs (les matrices le composant sont symétriques complexes et évidemment non hermitiennes). Une autre façon d'obtenir un relèvement temporel de ces systèmes écrits en fréquentiel est, de façon alternative à l'écriture originelle de Bérenger, en complexifiant la fréquence, d'opérer à une transformation de Laplace inverse par rapport au temps. On obtient ainsi des systèmes causaux de Friedrichs pouvant être lus, moyennant l'addition d'inconnues subsidiaires, comme des systèmes de Friedrichs identiques aux systèmes originaux à une perturbation d'ordre zéro près, auxquels on adjoint un système d'EDO. Ces systèmes sont évidemment bien posés (même si le semi groupe associé n'est pas forcément contractant et les conditions aux limites admissibles sont bien évidemment les mêmes que pour les systèmes originels). Cependant l'enveloppe convexe de l'obstacle peut être bien plus grande que celui-ci. Il reste donc dans le domaine de calcul effectif (directement corrélé au nombre d'inconnues) une grande partie qui peut sembler superflue. La question naturelle qui se pose alors est : peut-on au plus près d'un obstacle non convexe complexifier les dérivations de façon à obtenir l'effet de « damping » souhaité ? Quant aux méthodes connues de haut degré, elles souffrent, pour des schémas

explicites d'intégration pour la dérivée temporelle, d'un effondrement drastique du CFL. La question à laquelle nous allons tenter de répondre consiste à rechercher des approximations spatiales de haut degré dont les formulations explicites (Euler, Runge-Kutta) ne présenteraient pas le même inconvénient.

## Chapitre 1 : Rappels sur les problèmes de propagation d'ondes écrits sous forme de systèmes de Friedrichs

L'enjeu de cette thèse est donc l'étude des problèmes de diffraction. Plus exactement des équations de type Maxwell, Ondes et Ondes convectées. Ces systèmes différentiels sont gouvernés par des systèmes dit de Friedrichs, pour lesquels de nombreux résultats existent dans la littérature. Nous distinguerons deux types de problèmes : le problème instationnaire et le problème harmonique. Les résultats généraux d'existence et d'unicité sont spécifiques à chacun de ces problèmes, et dépendent de la nature des coefficients mis en jeu. Nous montrerons dans un premier temps que l'étude des problèmes à coefficients constants n'entraîne aucune perte de généralité ou d'information, puis nous étudierons cette classe d'équations sans condition aux limites. Nous nous attacherons donc dans cette partie à établir avec précision les hypothèses pour garantir l'existence et l'unicité de la solution aux problèmes étudiés.

## Chapitre 2 : Les Perfectly Matched Layers

Cette partie correspond à l'étude de ce type de conditions aux limites absorbantes. En partant de l'écriture par un changement de variable complexe des équations PML, une généralisation est proposée. Pour cela, nous procédons de manière similaire à Lassas et Somersalo [20] en définissant un milieu PML obtenu par exhaustion de  $\mathbb{R}^3$  sur une variété plongement de dimension 3 de  $\mathbb{C}^3$ , coïncidant avec  $\mathbb{R}^3$  sur l'intérieur du volume de calcul  $\Omega$  bordé par les PML. Une fois la théorie et le formalisme introduits, il convient de définir correctement le plongement précédent, et de vérifier ses propriétés. Au cours de cette thèse, plusieurs solutions ont été envisagées, et nous présenterons à chaque fois les inconvénients et les problèmes rencontrés, ainsi que les corrections apportées. Une écriture en système différentiel augmenté est ensuite proposée sous forme de décomposition en système de Friedrichs.

L'introduction du formalisme PML modifie les équations. Il convient donc de se demander si ces nouvelles équations admettent bien une unique solution, tant pour les problèmes instationnaires qu'harmoniques. A ce titre, les résultats du chapitre 1 seront donc largement exploités. Pour les problèmes harmoniques, les démonstrations d'existence et d'unicité nécessitent l'usage de famille d'opérateurs de type A présentant de bonnes propriétés et pour lesquels [18] fournit des théorèmes puissants. Pour les problèmes instationnaires, il est montré que le système d'équations pourra être mis, in fine, sous forme de systèmes de Friedrichs. Le problème des conditions aux limites sera également introduit. Les résultats d'existence, d'unicité (à un ensemble dénombrable de fréquences près (pouvant être vide)) et de décroissance

exponentielle dans les couches sont ainsi démontrés.

Une des applications sous-jacente à la réalisation de PML non convexes était de pouvoir entourer au plus près une forme complexe, comme par exemple un avion. Pouvant être comparé à une croix de Lorraine, la forme d'un avion peut être décomposée en parties semblables à un L. Nous nous sommes donc naturellement intéressés à un domaine de cette forme. Après avoir appliqué des formulations étoilées, nous avons souhaité utiliser au mieux la géométrie particulière de domaine, notamment sa décomposition en une partie convexe et une partie non convexe, et nous avons introduit une formulation hybride, mélangeant des PML cartésiennes classiques à une nouvelle formulation non convexe. La continuité du changement de variable est assurée à la jonction des deux types de PML. Des résultats numériques seront présentés, concernant deux types de problèmes : la propagation d'ondes électromagnétiques dans le vide ou la réflexion d'une onde plane incidente sur un scatterer parfait.

### Chapitre 3 : Une nouvelle Approximation de Haute Précision

Pour améliorer la précision des simulations numériques, une autre piste consistait à choisir une approximation spatiale de degré plus élevé que les méthodes de type Volumes Finis choisies pour les simulations précédentes. Celles de type Galerkin Discontinu ont été envisagées. Mais les formulations PML dégradent les conditions CFL à cause de perturbations sur les matrices de masses. Une étude des complexités et des conditions CFL pour ce type de méthode a été effectuée, et devant les résultats peu concluants mis en évidence, nous avons cherché une alternative. Le but étant d'écrire une nouvelle approximation où la solution est un polynôme de degré à choisir, mais dont la condition CFL est celle des Volumes Finis. L'idée majeure consiste à considérer que si une fonction est solution d'une Equation aux Dérivées Partielles de type Ondes, chacune de ses dérivées partielles l'est aussi. On obtient alors un système d'équations que nous approximerons avec une méthode de type Volumes Finis (où les fonctions tests seront les fonctions caractéristiques du maillage). La stabilité et la consistance de cette méthode sont prouvées, et nous présenterons les premières expériences numériques. Nous choisirons deux types de problèmes : en condition initiale (régularisation du Dirac) ou avec un second membre (onde incidente sur un obstacle parfait). A chaque fois, une comparaison avec une méthode de type Volumes Finis est présentée.

### Conventions et notations

Dans toute la suite de ce mémoire, les vecteurs seront dénotés en gras (ex :  $\mathbf{x} \in \mathbb{R}^3$ ), les matrices en gras et italique (ex :  $\mathbf{M}$ ) et les opérateurs en caligraphie (ex :  $\mathcal{A}$ ). Les conventions de notation d'Einstein seront régulièrement utilisées pour simplifier les écritures quand elles ne nuieront en rien à la lisibilité des formules. Les opérateurs et fonctions mathématiques usuelles seront également mises en valeur par une écriture en gras (ex : **grad**).

# Chapitre 1

## Rappels sur les problèmes de propagation d'ondes écrits sous forme de systèmes de Friedrichs

### 1.1 Enoncés des problèmes considérés

#### 1.1.1 Les Equations de Maxwell

Ces équations décrivent les phénomènes électromagnétiques dans les milieux continus à l'aide de quatre champs fonctionnels ou distributionnels définis de  $\mathbb{R}^+ \times \mathbb{R}^3$  dans  $\mathbb{R}^3$ . Par convention dans ce mémoire, les vecteurs seront notés en gras. Soit  $\mathbf{E}$  le champ électrique,  $\mathbf{H}$  le champ magnétique,  $\mathbf{B}$  l'induction électrique,  $\mathbf{D}$  l'induction magnétique,  $\mathbf{j}$  les courants électriques (le plus souvent distributionnels) et  $\rho$  les charges électriques. Au sens des distributions, les Equations de Maxwell s'écrivent :

$$\begin{cases} \mathbf{div} \mathbf{D} = \rho \\ \mathbf{div} \mathbf{B} = 0 \end{cases} \quad \begin{cases} \partial_t \mathbf{D} - \mathbf{rot} \mathbf{H} + \mathbf{j} = 0 \\ \partial_t \mathbf{B} + \mathbf{rot} \mathbf{E} = 0 \end{cases} \quad (1.1)$$

Ce système non fermé du point de vue de la relation entre les inductions et les champs et de celle entre les courants et les champs nous amène à énoncer deux types de lois phénoménologiques. Commençons par les lois inductions-champs. Dans le vide,

$$\mathbf{D} = \varepsilon_0 \mathbf{E} \quad \mathbf{B} = \mu_0 \mathbf{H},$$

où  $\varepsilon_0$  et  $\mu_0$  sont les permittivités électrique et magnétique du vide. On considère qu'un matériau est une perturbation du vide. Dans un matériau en statique, on privilégie, comme dans le vide, l'existence d'un potentiel électrique et l'existence d'un potentiel vecteur magnétique, soit l'existence d'une fonction  $V$  et d'une fonction vectorielle  $\mathbf{A}$  telles que  $\mathbf{E} = -\nabla V$  et  $\mathbf{B} = \mathbf{rot} \mathbf{A}$ . Ceci amène à introduire les vecteurs de polarisation électrique et magnétique  $\mathbf{P}_e$  et  $\mathbf{P}_m$ . Les lois constitutives s'écrivent alors :

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}_e \quad \mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{P}_m).$$

En l'absence de champs et de matière, on fait l'hypothèse que  $\mathbf{P}_e = \mathbf{P}_m = \mathbf{0}$ . On peut également faire l'hypothèse de localité statique :

$$\mathbf{P}_e = \varepsilon_0 f_E(\mathbf{E}) \quad \mathbf{P}_m = f_H(\mathbf{H})$$

en tout point. La linéarité statique est obtenue en confondant  $f_e$  et  $f_h$  avec leur dérivée en 0.

Soit  $\Omega$  un domaine de  $\mathbb{R}^3$ . Soit les variables adimensionnées  $\mathbf{E} := \sqrt{\varepsilon_0} \mathbf{E}$ ,  $\mathbf{H} := \sqrt{\mu_0} \mathbf{H}$  et  $t := c_0 t$  avec  $c_0 = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$ . Soit  $\psi = \begin{pmatrix} \mathbf{D} \\ \mathbf{B} \end{pmatrix}$  et  $\phi = \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}$ . L'opérateur associé aux Equations de Maxwell  $\begin{pmatrix} 0 & -\mathbf{rot} \\ \mathbf{rot} & 0 \end{pmatrix}$  est symétrique et peut être décomposé sous la forme  $\sum_i \mathbf{A}_i \partial_i$ , avec  $\mathbf{A}_i = \begin{pmatrix} 0 & -e_i \wedge \\ e_i \wedge & 0 \end{pmatrix}$  avec  $(e_i)_{i \in \llbracket 1,3 \rrbracket}$  base canonique de  $\mathbb{R}^3$  et  $e_i \wedge$  la matrice associée au produit vectoriel par  $e_i$ .

Les Equations de Maxwell dans le vide sur  $\mathbb{R}^+ \times \mathbb{R}^3$  s'écrivent :

$$\frac{\partial \phi(t, \mathbf{x})}{\partial t} + \sum_i \mathbf{A}_i \frac{\partial \phi(t, \mathbf{x})}{\partial \mathbf{x}_i} = f(t, \mathbf{x}).$$

Dans le cas d'un diélectrique, sur  $\mathbb{R}^+ \times \mathbb{R}^3$ , les Equations de Maxwell deviennent :

$$\mathbf{A}_0 \frac{\partial \phi(t, \mathbf{x})}{\partial t} + \sum_i \mathbf{A}_i \frac{\partial \phi(t, \mathbf{x})}{\partial \mathbf{x}_i} + \mathbf{B} \phi(t, \mathbf{x}) = f(t, \mathbf{x}).$$

Les conditions aux limites considérées sont de la forme :

$$\mathbf{n} \times E + \mathbf{n} \times \mathbf{Z}(\mathbf{n} \times H) = 0, \quad (1.2)$$

où  $\mathbf{Z}$  est appelée matrice d'impédance et telle que la matrice  $\mathbf{Z} + \mathbf{Z}^T$  soit positive ou nulle. Le cas où cette matrice est antisymétrique correspond à des conditions aux limites conservatives. Le cas particulier considéré dans cette thèse correspond à la condition d'un obstacle parfaitement conducteur, à savoir  $\mathbf{Z} = 0$ .

### 1.1.2 Les Equations des Ondes convectées

Nous examinerons donc ici la perturbation d'un écoulement porteur stationnaire subsonique régulier quelconque dans  $\mathbb{R}^3$  défini par :  $\mathbf{U}_0^i$  un champ de vitesse,  $\mathbf{P}_0$  un champ de pression,  $\rho_0$  une masse volumique,  $e_0$  une énergie interne,  $s_0$  une entropie, et vérifiant les équations d'Euler stationnaires suivantes :

$$\begin{cases} \partial_j (\rho_0 \mathbf{U}_0^j) & = 0 \\ \partial_j (\rho_0 \mathbf{U}_0^j \mathbf{U}_0^i + \mathbf{P}_0 \delta_{ij}) & = 0 \\ \partial_j \left( \left( \rho_0 \left( e_0 + \frac{\|\mathbf{U}_0\|_2^2}{2} \right) + \mathbf{P}_0 \right) \mathbf{U}_0^j \right) & = 0 \end{cases} \quad (1.3)$$

En l'absence de chocs (hypothèses supplémentaires de régularité de l'écoulement), la cinquième équation (bilan d'énergie) pourra être confondue avec le bilan d'entropie suivant ( $\sigma_0$  désignant l'entropie volumique de l'écoulement porteur) :  $\partial_j (\rho_0 \sigma_0 \mathbf{U}_0^j) = 0$ , la loi d'état du champ de pression s'écrivant alors indifféremment  $\mathbf{P}_0 = \mathbf{P}_0(\rho_0, e_0)$  ou  $\mathbf{P}_0 = \mathbf{P}_0(\rho_0, \sigma_0)$ . De plus, si le fluide porteur est barotrope, la loi d'état devient  $\mathbf{P}_0 = \mathbf{P}_0(\rho_0)$ , et seules les quatre premières équations de (1.3) peuvent être retenues.

Après perturbation au temps  $t = 0$ , l'écoulement est supposé vérifier les équations d'Euler instationnaires (complètes ou barotropes) au sens des distributions, soit le système de 5 équations sous la forme conservative suivante (on conservera les notations utilisées pour les variables entropiques de l'écoulement porteur sans l'indice 0 afin de stipuler leurs caractères désormais instationnaires) :

$$\begin{cases} \partial_t w + \partial_i f^i(w) & = g \\ w(0) & = w_0, \end{cases} \quad (1.4)$$

avec  $w \in \mathcal{C}_m^1(\mathbb{R}^4, \mathbb{R}^5)$  et les flux  $f^i \in (\mathbb{R}^5, \mathbb{R}^5)$ , fonctions « suffisamment régulières » données par :

$$w = \begin{pmatrix} \rho \\ \rho \mathbf{U}^i \\ \rho \left( e + \frac{\|\mathbf{U}\|_2^2}{2} \right) \end{pmatrix} \quad (\text{resp.}) \quad w = \begin{pmatrix} \rho \\ \rho \mathbf{U}^i \\ \rho \sigma \end{pmatrix}$$

$$f^i(w) = \begin{pmatrix} \rho \mathbf{U}^i \\ \left( \rho \left( e + \frac{\|\mathbf{U}\|_2^2}{2} \right) + \mathbf{P} \right) \mathbf{U}^i \end{pmatrix} \quad (\text{resp.}) \quad f^i(w) = \begin{pmatrix} \rho \mathbf{U}^i \\ \rho \mathbf{U}^i \mathbf{U}^j + \mathbf{P} \delta_{ij} \\ \rho \sigma \mathbf{U}^i \end{pmatrix},$$

dans le cas où l'on prend en compte la perturbation thermique (respectivement la perturbation entropique). Dans un premier temps, on se propose d'étudier le caractère bien posé du système (en espace libre) obtenu par une linéarisation entropique au premier ordre de (1.4).

(1.4) constitue un système hyperbolique non linéaire admettant une entropie de Lax [35]. Plus précisément, cela signifie que (1.4) admet une équation de conservation supplémentaire (en l'absence de chocs) de la forme :

$$\partial_t \mathbf{S}(w) + \partial_i \mathbf{S}^i(w) - g \nabla_w \mathbf{S}(w) = 0, \quad (1.5)$$

où  $\mathbf{S}(w)$  est strictement convexe en  $w$  et telle que le jacobien des flux  $f^i$  soit auto-adjoint par rapport à la métrique induite par le hessien de  $\mathbf{S}(w)$ . C'est une entropie (volumique) « physique » dans le cas général et dans le cas barotrope avec  $\mathbf{P} = \mathbf{P}(\rho)$ , on a :

$$\mathbf{S}(w) = \mathbf{H}(\rho) + \rho \frac{\|\mathbf{U}\|_2^2}{2}, \quad \mathbf{H}''(\rho) = \frac{\mathbf{P}'(\rho)}{\rho}.$$

On aura supposé bien entendu que  $\mathbf{P}'(\rho) = c^2(\rho)$  est positif.

On a alors :  $\mathbf{S}^i(w) = (\mathbf{S} + \rho)\mathbf{U}^i$ , et l'équation de conservation supplémentaire a physiquement le sens d'une équation d'énergie. Rappelons un théorème dû à S.K Godunov [13] :

**Théorème 1.1**

*Le système (1.4) est symétrisable si et seulement si il admet une entropie.*

Par application du théorème (1.1), le changement de variables bijectif (dans le domaine admissible pour  $w$ )  $\Phi = \nabla_w(S(w))$  permet d'obtenir une forme symétrique de (1.4) :

$$\partial_t(\nabla_\Phi \mathbf{S}^*(\Phi)) + \partial_i(\nabla_\Phi \mathbf{S}^{*i}(\Phi)) = g, \tag{1.6}$$

avec :  $\mathbf{S}^*(\Phi) = \Phi \cdot w(\Phi) - \mathbf{S}(w(\Phi))$ ,  $\mathbf{S}^{*i}(\Phi) = \Phi \cdot f^i(w(\Phi)) - \mathbf{S}^i(w(\Phi))$ . On suppose que (1.6) est bien posée, donc que sa solution est continue par rapport aux données du problème. En posant  $g = \varepsilon h$ , et  $\Phi = \Phi_0 + \varepsilon \varphi$  ( $\Phi_0$  variables entropiques associées à l'écoulement porteur), on obtient formellement en faisant un développement limité en  $\varepsilon$  au premier ordre de la formulation faible de (1.6), une expression linéarisée du problème :

$$\partial_t(\mathbf{H}_{\mathbf{S}^*}(\Phi_0)\varphi) + \partial_i(\mathbf{H}_{\mathbf{S}^{*i}}(\Phi_0)\varphi) = h,$$

où  $\mathbf{H}_{\mathbf{S}^*}(\Phi_0)$  sont les Hessiens symétriques définis positifs de  $\mathbf{S}^*(\Phi)$  et  $\mathbf{H}_{\mathbf{S}^{*i}}(\Phi_0)$  symétriques de  $\mathbf{S}^{*i}(\Phi)$ , calculés en  $\Phi_0$ .

En notant  $\mathbf{A}^0(x) = \mathbf{H}_{\mathbf{S}^*}(\Phi_0)$ , et  $\mathbf{A}^i(x) = \mathbf{H}_{\mathbf{S}^{*i}}(\Phi_0)$  la matrice symétrique définie positive et les matrices symétriques, (1.1.2) constitue un système de Friedrichs :

$$\partial_t(\mathbf{A}^0\varphi) + \partial_i(\mathbf{A}^i\varphi) = h. \tag{1.7}$$

Le cas qui nous intéressera dans cette thèse est le cas des Equations d'Euler linéarisées Barotropes. On considère donc un écoulement porteur vérifiant (1.3) de loi d'état en  $\mathbf{P}_0$  quelconque. L'hypothèse de barotropie de la perturbation consiste à supposer que l'influence des variations entropiques est négligeable devant celle des variations de la masse volumique, autrement dit que :

$$\rho = \rho_0 + \varepsilon \rho' + o(\varepsilon) \quad \Rightarrow \quad \mathbf{P} = \mathbf{P}_0 + \varepsilon c_0^2 \rho' + o(\varepsilon), \tag{1.8}$$

où  $c_0 = c_0(x)$  désigne la vitesse du son dans le milieu. On pose les changements de variables suivants :  $\mathbf{U}^i = \mathbf{U}_0^i + \varepsilon u^i$ ,  $\rho = \frac{c_0 \rho'}{\rho_0}$  et  $\varphi = (u^1, u^2, u^3, \rho)^T$  qui est un adimensionnement de la variable entropique associée à l'entropie de Lax du système non linéaire barotrope isotherme. Dès lors, après développement en  $\varepsilon$  à l'ordre 1 des équations générales d'Euler instationnaires (1.4), on obtient que la variable  $\varphi$  est solution de :

$$\partial_t \varphi + \mathbf{A}^i \partial_i \varphi + \mathbf{B} \varphi = f, \tag{1.9}$$

$$\begin{aligned} \text{avec } \mathbf{A}^i \partial_i &= (\mathbf{U}_0^i \partial_i) \mathbf{I}_4 + c_0 \begin{pmatrix} 0 & \nabla \\ \nabla^T & 0 \end{pmatrix} \\ \mathbf{B} &= \begin{pmatrix} [\partial_i \mathbf{U}_0^j]_{i,j} & \frac{1}{\rho_0} \left( \nabla (c_0 \rho_0) - \frac{1}{c_0} \nabla \mathbf{P}_0 \right) \\ \frac{c_0}{\rho_0} \nabla^T \rho_0 & \frac{-1}{c_0} (\mathbf{U}_0^i \partial_i c_0) \end{pmatrix} \\ f &= \frac{1}{\rho_0} (g^1 - \mathbf{U}_0^1 g^0, \dots, g^d - \mathbf{U}_0^d g^d, c_0 g^0)^T. \end{aligned}$$

**Remarques** : la partie diagonale de la partie principale de l'opérateur est un opérateur de convection tandis que la partie non diagonale est l'écriture sous la forme d'un système d'ordre 1 de l'équation des ondes.

On obtient ces équations en utilisant la formule de dérivation d'un produit, donc elles n'ont de sens distributionnel que si l'écoulement porteur est régulier. De même, dans le cas d'un écoulement porteur uniforme en dehors d'un compact  $\bar{\Omega}$ , la matrice  $\mathbf{B}$  s'annule, les vitesses  $\mathbf{U}_0$  et  $c_0$  sont constantes. Le cas particulier où  $\mathbf{U}_0 = 0$  (cas sans convection) correspond à l'équation des ondes classique à l'ordre 1.

### Conditions aux Limites

Dans ce paragraphe sont passés en revue les différents problèmes selon le type d'écoulement. On pourra se reporter à [38] pour davantage de détails. On représente classiquement l'ensemble des conditions aux limites, noté  $\mathcal{N}(\mathbf{x})$  par le noyau d'une matrice  $\mathbf{M} \in \mathcal{M}_3(\mathbb{C})$ . Donc  $\mathcal{N}(\mathbf{x}) = \ker(\mathbf{M}(\mathbf{x}))$ . Soit  $\mathcal{A}$  l'opérateur spatial non fermé défini par  $\mathbf{A}^i \partial_i + \mathbf{B}$ . Son domaine  $\mathcal{D}(\mathcal{A})$  est  $C^1(\bar{\Omega})^m$  avec des conditions aux limites maximales positives (hypothèse indispensable pour assurer l'existence et l'unicité du problème instationnaire posé avec des conditions aux limites)

$$\forall \varphi \in \mathcal{D}(\mathcal{A}), (\mathbf{A}^i n_i \varphi, \varphi)_{\mathbb{C}^3} \geq 0 \quad \forall \psi \in \mathcal{D}(\mathcal{A}^\#), (\mathbf{A}^i n_i \psi, \psi)_{\mathbb{C}^3} \leq 0, \quad (1.10)$$

où  $\mathcal{A}^\# = -\partial_i(\mathbf{A}^i \cdot) + \mathbf{B}^T$  est l'adjoint formel non fermé de l'opérateur spatial. On exprime les flux  $\mathbf{A}^i n_i$  dérivant de (1.9) sous la forme :

$$\mathbf{A}^i n_i = \mathbf{U}_0^i n_i \mathbf{I}_3 + c_0 \begin{pmatrix} 0 & \mathbf{n} \\ \mathbf{n}^T & 0 \end{pmatrix} \quad (1.11)$$

où  $\mathbf{U}_0^i n_i$  désigne la vitesse normale sortante de l'écoulement porteur. La condition aux limites dépend donc de cette vitesse, soit encore du type de l'écoulement.

#### ► L'écoulement porteur glissant

Pour  $\mathbf{U}_0^i n_i = 0$ , l'écoulement porteur est dit glissant. Dans ce cas, les flux  $\mathbf{A}^i n_i$  correspondent aux flux de l'équation des ondes écrite à l'ordre 1. Deux types de conditions d'obstacles apparaissent alors.

Si  $\forall \varphi \neq 0 \in \ker(\mathbf{M})$ , alors  $(\mathbf{A}^i n_i \varphi, \varphi)_{\mathbb{C}^3} = 0$ . C'est en particulier le cas pour la condition de Dirichlet pour l'équation des ondes ou condition d'obstacle mou qui consiste à imposer  $\rho = 0$  et pour la condition de Neumann ou condition d'obstacle dur à savoir  $u^i n_i = 0$ .

Si  $\forall \varphi \neq 0 \in \ker(\mathbf{M})$ , alors  $(\mathbf{A}^i n_i \varphi, \varphi)_{\mathbb{C}^3} > 0$ . On impose alors la condition aux limites homogène  $\forall \mathbf{x} \in \Omega$ ,  $\rho = \alpha(\mathbf{x}) \mathbf{u} \cdot \mathbf{n}$ , avec  $\mathbf{n}$  normale sortante à  $\Omega$  sur  $\partial\Omega$ , et la condition suivante sur l'impédance :

$$\alpha(\mathbf{x}) \in \mathbb{C} \quad \text{et} \quad \mathbf{Re}(\alpha) \geq 0.$$

La condition d'impédance peut être écrite en fonction de  $\beta$ , absorption acoustique de module inférieur à 1 définie par :

$$\beta(\mathbf{x}) = \frac{\alpha(\mathbf{x}) - 1}{\alpha(\mathbf{x}) + 1},$$

ce qui se traduit sous la forme classique :

$$(\beta - 1)\rho + (\beta + 1)\mathbf{u} \cdot \mathbf{n} = 0. \quad (1.12)$$

#### ► L'écoulement porteur sortant

Pour  $\mathbf{U}_0^i n_i > 0$ , l'écoulement porteur est dit sortant. Dans ce cas, les conditions aux limites intéressantes correspondent à des conditions transparentes à l'ordre 1 (complètement non réfléchissantes pour les ondes normales). Elles correspondent à des coefficients gelés, supposés localement constants.

Dans ce cas, l'expression de la matrice  $\mathbf{M}$  est donnée par :

$$\mathbf{M} = -(\mathbf{A}^i n_i)^- = -\frac{1}{2} (\mathbf{U}_0^i n_i - c_0) \begin{pmatrix} \mathbf{n} \otimes \mathbf{n} & -\mathbf{n} \\ -\mathbf{n}^T & 1 \end{pmatrix}. \quad (1.13)$$

Mais la condition  $\mathbf{M}\varphi = 0$  implique  $\rho = u^i n_i = \mathbf{u} \cdot \mathbf{n}$ , donc dans ce cas, il ne subsiste qu'une seule condition aux limites correspondant à  $\alpha = (1, 0)$  ou  $\beta = (0, 0)$ .

#### ► L'écoulement porteur entrant

Pour  $\mathbf{U}_0^i n_i < 0$ , l'écoulement porteur est dit entrant. Les conditions aux limites sont les mêmes que dans le cas précédent, avec cette fois-ci :

$$(\mathbf{A}^i n_i)^+ = \frac{1}{2} (\mathbf{U}_0^i n_i + c_0) \begin{pmatrix} \mathbf{n} \otimes \mathbf{n} & \mathbf{n} \\ \mathbf{n}^T & 1 \end{pmatrix}$$

et  $\mathbf{M} = -(\mathbf{A}^i n_i)^- = (\mathbf{A}^i n_i)^+ - \mathbf{A}^i n_i$ , donc

$$\mathbf{M} = \begin{pmatrix} -\mathbf{U}_0^i n_i \mathbf{I}_3 + \frac{1}{2} (\mathbf{U}_0^i n_i + c_0) \mathbf{n} \otimes \mathbf{n} & \frac{1}{2} (\mathbf{U}_0^i n_i - c_0) \mathbf{n} \\ \frac{1}{2} (\mathbf{U}_0^i n_i - c_0) \mathbf{n}^T & -\frac{1}{2} (\mathbf{U}_0^i n_i - c_0) \end{pmatrix}.$$

La condition non réfléchissante  $\mathcal{M}\varphi = 0$  se formule dans ce cas de la façon suivante :

$$\begin{cases} \rho = \mathbf{u} \cdot \mathbf{n} \\ \mathbf{U}_0^i n_i \mathbf{u} - \frac{1}{2} (\mathbf{U}_0^i n_i + c_0) (\mathbf{u} \cdot \mathbf{n}) \mathbf{n} + \frac{1}{2} (c_0 - \mathbf{U}_0^i n_i) \rho \mathbf{n} = 0. \end{cases}$$

**Remarque** : ces relations ne sont pas indépendantes. La première relation se déduit des trois autres en multipliant scalairement par le vecteur  $\mathbf{n}$ . Cette condition non réfléchissante approchée ne comporte donc que deux relations indépendantes.

## 1.2 Problèmes de diffraction considérés

Les systèmes linéaires d'évolution provenant d'équations de conservation au premier ordre (non-linéaires) sont symétrisables si et seulement si le système admet une entropie de Lax, systèmes pour lesquels on a des chances d'obtenir l'unicité de la solution au sens des distributions. Ce dernier cas est celui des systèmes de Friedrichs, qui représentent la forme la plus naturelle pour des problèmes de théorie des champs (équations de Maxwell, de Dirac, ...).

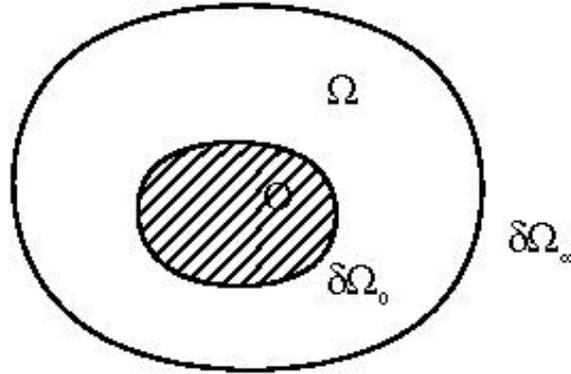


FIG. 1.1: Description des domaines en présence

Soit  $\mathbf{O}$  un obstacle borné, régulier non nécessairement connexe pouvant être vide. Soit  $\Omega$  un domaine connexe, non nécessairement simplement connexe représentant le domaine spatial. Soit  $\delta\Omega_0$  la frontière entre  $\Omega$  et  $\mathbf{O}$  et  $\delta\Omega_\infty$  sa frontière extérieure. Elles sont toutes deux  $C^1$ , et  $\delta\Omega_\infty$  peut éventuellement être rejetée à l'infini dans le cas d'un domaine  $\Omega$  non borné. Dans toute la suite du chapitre, nous nommerons  $n$  la dimension d'espace,  $n = 2$  ou  $3$ .

### 1.2.1 Cas instationnaire

Soit  $\mathbf{x} \in \mathbb{R}^n$  et  $t \in \mathbb{R}^+$ . Soit  $m$  le nombre d'équations. Les problèmes instationnaires étudiés sont de la forme :

$$\begin{cases} \mathbf{A}_0(\mathbf{x}) \frac{\partial \phi(\mathbf{x}, t)}{\partial t} + \sum_i \mathbf{A}_i(\mathbf{x}) \frac{\partial \phi(\mathbf{x}, t)}{\partial \mathbf{x}_i} + \mathbf{B}(\mathbf{x}) \phi(\mathbf{x}, t) = F(\mathbf{x}, t) \\ \phi(\mathbf{x}, 0) = 0 \quad \phi(\cdot, t)|_{\delta\Omega_0} \in \mathcal{N}_0(\mathbf{x}) \quad \phi(\cdot, t)|_{\delta\Omega_\infty} \in \mathcal{N}_\infty(\mathbf{x}). \end{cases} \quad (1.14)$$

où les matrices  $\mathbf{A}_i \in M_{m \times m}(\mathbb{R})$  sont symétriques et assez régulières (par exemple lipschitziennes),  $\mathbf{A}_0 \in M_{m \times m}(\mathbb{R})$  est uniformément définie positive à coefficients dans  $L^\infty(\mathbb{R})$ , et  $\mathbf{B} \in M_{m \times m}(\mathbb{R})$  à coefficients dans  $L^\infty(\mathbb{R})$ .

$$\begin{aligned} \text{supp}(\mathbf{A}_0(\mathbf{x}) - \mathbf{I}_m) &\subset \Omega \cup \delta\mathbf{O} & \text{supp}F(\cdot, t) &\subset \Omega \cup \delta\mathbf{O} \\ \text{supp}(\mathbf{A}_i(\mathbf{x}) - \mathbf{A}_i^c) &\subset \Omega \cup \delta\mathbf{O} & \text{supp}\mathbf{B} &\subset \Omega \cup \delta\mathbf{O} \end{aligned} \quad (1.15)$$

ce qui signifie que les matrices  $\mathbf{A}_i^c$  deviennent des matrices à coefficients constants en dehors du domaine  $\Omega \cup \delta\mathbf{O}$ .

**Remarques :** dans le cas de conditions initiales ou de conditions aux limites inhomogènes, on suppose que l'on peut les relever dans le domaine de l'opérateur. La solution  $\phi$  recherchée appartient à l'espace  $L^2([0, T] \times \Omega)^m$ . Par la suite, on notera  $Q = [0, T] \times \Omega$  le domaine spatio-temporel.

**Proposition 1.2**

*On pourra toujours supposer  $\Omega$  borné (mais grand) en choisissant sa frontière  $\delta\Omega_\infty$ .*

Soit un problème à coefficients constants. Soit  $\Gamma_c$  le **cône d'évolution**, à savoir l'ensemble des  $\lambda \mathbf{n}$  avec  $\lambda \geq 0$  et  $\mathbf{n} \in S^n$  tel que  $\forall \xi \in \mathbb{R}^m, \quad \xi^T (\mathbf{I}_m \mathbf{n}_t + \mathbf{A}_c^i \mathbf{n}_i) \xi \geq 0$ . C'est un cône convexe. Le **cône de propagation**, noté  $\Gamma_p$ , est défini comme le polaire du cône d'évolution  $\Gamma_c$  dans  $\mathbb{R}^{n+1}$ . C'est aussi le plus petit cône convexe contenant le support de la solution élémentaire.

**Corollaire 1.3 (Principe de Huygens [16])**

*La solution autour de  $(\mathbf{x}, t)$  ne dépend que de l'intersection du support de  $f$  avec le cône  $\tau_{(\mathbf{x}, t)}(\Gamma_p)$ , image du cône de propagation par une symétrie d'hyperplan  $t = 0$ , puis une translation de  $(\mathbf{x}, t)$ . Si cette intersection est vide, alors la solution est nulle dans un voisinage de  $(\mathbf{x}, t)$ .*

Le principe de Huygens nous assure donc que la solution est nulle à l'extérieur de  $\Omega \cup \overline{\mathbf{O}} + \Gamma_p$ . Donc pour tout  $T$ , on peut toujours choisir la frontière  $\delta\Omega_\infty$  telle que  $\phi(\cdot, t)|_{\delta\Omega_\infty} = 0$  et  $\forall \mathcal{N}_\infty(\mathbf{x}), \{0\} \in \mathcal{N}_\infty(\mathbf{x})$  : toutes les conditions homogènes seront donc correctes.

Définissons le problème à coefficients constants dérivé de (1.14) :

$$\begin{cases} \frac{\partial \phi(\mathbf{x}, t)}{\partial t} + \sum_i \mathbf{A}_i^c \frac{\partial \phi(\mathbf{x}, t)}{\partial \mathbf{x}_i} = G(\mathbf{x}, t) \\ \phi(\mathbf{x}, 0) = 0 \quad \phi(\cdot, t)|_{\delta\Omega_0} \in \mathcal{N}_0(\mathbf{x}) \quad \phi(\cdot, t)|_{\delta\Omega_\infty} \in \mathcal{N}_\infty(\mathbf{x}) \end{cases} \quad (1.16)$$

**Proposition 1.4**

Soit un domaine  $\Omega$  infini. Si le problème à coefficients variables (1.14) admet une unique solution  $\bar{\phi}$  et si le problème à coefficients constants (1.16) admet également une solution unique, alors (1.14) peut être remplacé par (1.16) avec

$$G(\mathbf{x}, t) = F(\mathbf{x}, t) + (\mathbf{I}_m - \mathbf{A}_0(\mathbf{x}))\bar{\phi}(\mathbf{x}, t) + (\mathbf{A}_i^c - \mathbf{A}_i(\mathbf{x}))\frac{\partial \bar{\phi}(\mathbf{x}, t)}{\partial \mathbf{x}_i} - \mathbf{B}(\mathbf{x})\bar{\phi}(\mathbf{x}, t). \quad (1.17)$$

**Démonstration** : soit  $\bar{\phi}$  une solution du problème à coefficients variables (1.14). Montrons qu'il existe une unique solution au problème à coefficients constants (1.16) telle que les restrictions de  $\phi$  et de cette solution au domaine  $\Omega$  soient égales. On souhaite donc établir que l'opérateur à coefficients variables est le perturbé d'un opérateur à coefficients constants, et ainsi bénéficier d'une équivalence entre la résolution des problèmes à coefficients variables, difficiles, et ceux à coefficients constants. Soit  $(\mathbf{x}, t) \in \mathbb{R}^3 \times \mathbb{R}^+$ .

$$\begin{aligned} \frac{\partial \bar{\phi}(\mathbf{x}, t)}{\partial t} + \sum_i \mathbf{A}_i^c \frac{\partial \bar{\phi}(\mathbf{x}, t)}{\partial \mathbf{x}_i} &= (\mathbf{I}_m - \mathbf{A}_0(\mathbf{x}) + \mathbf{A}_0(\mathbf{x}))\frac{\partial \bar{\phi}(\mathbf{x}, t)}{\partial t} \\ &\quad + \sum_i (\mathbf{A}_i^c - \mathbf{A}_i(\mathbf{x}) + \mathbf{A}_i(\mathbf{x}))\frac{\partial \bar{\phi}(\mathbf{x}, t)}{\partial \mathbf{x}_i} \\ &\quad + (\mathbf{B}(\mathbf{x}) - \mathbf{B}(\mathbf{x}))\bar{\phi}(\mathbf{x}, t) \\ &= (\mathbf{I}_m - \mathbf{A}_0(\mathbf{x}))\frac{\partial \bar{\phi}(\mathbf{x}, t)}{\partial t} + \sum_i (\mathbf{A}_i^c - \mathbf{A}_i(\mathbf{x}))\frac{\partial \bar{\phi}(\mathbf{x}, t)}{\partial \mathbf{x}_i} \\ &\quad - \mathbf{B}(\mathbf{x})\bar{\phi}(\mathbf{x}, t) + F(\mathbf{x}, t). \end{aligned}$$

L'existence et l'unicité de la solution  $\bar{\phi}$  de (1.14) permet de remplacer l'équation appliquée en  $\bar{\phi}$  par le second membre  $F \in L^2(\mathbb{R}^m \times \mathbb{R}^+)$ . En posant  $G$  la fonction (1.17), on obtient bien que  $\bar{\phi}$  est solution du problème à coefficients constants. Grâce aux propriétés sur les supports des matrices en jeu,  $G$  est à support dans  $\Omega \cup \delta\mathbf{O}$  et  $G \in L^2(\mathbb{R}^m \times \mathbb{R}^+)$ .  $\square$

Cette proposition est motivée par l'étude faite au chapitre 2 d'une formulation PML du problème à coefficients constants (ou gelés). Ainsi, il suffit d'étudier le problème à coefficients constants et ses propriétés à l'extérieur du domaine  $\Omega$ .

**1.2.2 Cas harmonique**

Il existe trois définitions possibles pour les problèmes harmoniques, qui ne sont, en général, pas équivalentes.

**Définition 1.1** (Amplitude Limite)

Cette définition du problème harmonique passe par la transformée de Fourier d'une éventuelle solution asymptotique ( $T$  et  $t$  tendant vers l'infini), périodique en temps, de période  $\frac{2\pi}{\omega}$  avec  $F = f e^{i\omega t}$  avec  $f \in L^2(\Omega)^m$

**Définition 1.2** (Absorption Limite)

Il s'agit de la limite quand  $\varepsilon$  tend vers 0 dans  $L^2_{loc}(\Omega)^m$  de l'éventuelle solution  $L^2(\Omega)^m$  de (1.18). Le problème harmonique s'écrit alors :

$$\begin{cases} (i\omega + \varepsilon) \mathbf{A}_0(\mathbf{x})\phi(\mathbf{x}) + \sum_i \mathbf{A}_i(\mathbf{x}) \frac{\partial \phi(\mathbf{x})}{\partial \mathbf{x}_i} + \mathbf{B}(\mathbf{x})\phi(\mathbf{x}) = f(\mathbf{x}) \\ \phi|_{\delta\Omega_0} \in \mathcal{N}_0(\mathbf{x}). \end{cases} \quad (1.18)$$

**Remarque** : si  $\mathbf{B}(\mathbf{x}) + \mathbf{B}(\mathbf{x})^T - \partial_i(\mathbf{A}^i(\mathbf{x}))$  n'est pas une matrice positive, le système (1.18) n'a pas forcément une solution  $L^2$  mais il se peut pourtant que pour  $\varepsilon$  tendant vers 0 il y ait une solution  $L^2_{loc}$  (instabilité convective).

**Définition 1.3** (Condition de Radiation)

Quand  $F = f e^{i\omega t}$  et que le domaine  $\Omega$  est infini, il s'agit de la définition d'une solution instationnaire dans  $\mathbb{R} \times \Omega$  de la forme  $\phi e^{i\omega t}$ , où  $\phi \in L^2_{loc}(\Omega)^m$ , avec adjonction d'une condition de décroissance de  $\phi$  à l'infini dans  $\mathbb{R}^n$ .

Dans la suite de ce chapitre, nous considérerons la définition par absorption limite (1.18) en prenant le cas limite  $\varepsilon = 0$ .

**Remarques** : la propriété (1.2) n'est alors plus vraie. Il n'est dès lors plus possible de se ramener à un domaine borné : l'apport d'un second membre, et donc d'énergie, de façon périodique ne permet pas de rester à énergie bornée. En revanche, la propriété (1.4) de pouvoir se ramener à l'étude des problèmes à coefficients admet l'équivalent suivant.

**Proposition 1.5**

Soit un domaine  $\Omega$  infini. Soit le problème à coefficients constants harmonique

$$\begin{cases} (i\omega + \varepsilon) \phi(\mathbf{x}) + \sum_i \mathbf{A}_i^c \frac{\partial \phi(\mathbf{x})}{\partial \mathbf{x}_i} = G(\mathbf{x}) \in L^2_{cpt}(\Omega)^m \\ \phi|_{\delta\Omega_0} \in \mathcal{N}_0(\mathbf{x}). \end{cases} \quad (1.19)$$

Si le problème à coefficients variables (1.18) admet une unique solution  $\bar{\phi}$  et si le problème à coefficients constants (1.19) admet également une solution unique, alors (1.18) peut être remplacé par (1.19) avec

$$G(\mathbf{x}) = F(\mathbf{x}) + (i\omega + \varepsilon) (\mathbf{I}_m - \mathbf{A}_0(\mathbf{x}))\bar{\phi}(\mathbf{x}) + (\mathbf{A}_i^c - \mathbf{A}_i(\mathbf{x})) \frac{\partial \bar{\phi}(\mathbf{x})}{\partial \mathbf{x}_i} - \mathbf{B}(\mathbf{x})\bar{\phi}(\mathbf{x}).$$

**Remarque** : l'étude du problème harmonique est particulièrement délicate. Dans le cas  $\varepsilon > 0$  il existe toujours une solution dans  $L^2$  à condition que le second membre soit  $L^2$  à support compact fixé. Cependant pour le problème  $\varepsilon$  tend vers zéro, même dans le cas des coefficients constants il peut y avoir des problèmes si on ne sait pas évaluer l'inverse du noyau de Green.

## 1.3 Résultats généraux d'existence et d'unicité

### 1.3.1 Rappels sur les systèmes de Friedrichs

#### Généralités

Notons par la suite  $Q = \Omega$  dans le cadre de problèmes harmoniques ou bien  $\Omega \times [0, T]$  pour des problèmes instationnaires. On cherche à résoudre de façon faible ou forte le problème  $\mathcal{A}u = f$ , où  $f \in L^2(\overline{Q})^m$  et  $\mathcal{A}$  un opérateur non borné, fermable, défini par  $\mathbf{A}_0 \partial_t + \sum_i \mathbf{A}_i \partial_i + \mathbf{B}$ . Son domaine  $\mathcal{D}(\mathcal{A})$  est  $C^1(\overline{Q})^m$  avec la trace sur  $\partial Q$  dans  $\mathcal{N}$  un sous espace vectoriel.

On note  $\mathcal{A}^\sharp$  son adjoint formel  $-\mathbf{A}_0 \partial_t - \sum_i \partial_i(\mathbf{A}_i, \cdot) + \mathbf{B}^T$  défini par la relation  $\forall (u, v) \in \mathbf{H}^2$ ,  $\langle \mathcal{A}u, v \rangle = \langle u, \mathcal{A}^\sharp v \rangle$ . Son domaine  $\mathcal{D}(\mathcal{A}^\sharp)$  est  $C^1(\overline{Q})^m$  avec la trace sur  $\partial Q$  dans  $(\mathbf{A}^i \mathbf{n}_i(\mathcal{N}))^\perp$ . C'est le plus « grand » adjoint formel qui annule les termes de bords dans les intégrations par parties. On notera  $\overline{\mathcal{A}}$  la plus petite extension fermée de  $\mathcal{A}$ .

**Définition 1.4** (Solution faible)

$u \in L^2(\overline{Q})^m$  est appelé solution faible si  $\forall v \in \mathcal{D}(\mathcal{A}^\sharp)$  (donc  $v|_{\partial Q} \in (\mathbf{A}^i \mathbf{n}_i(\mathcal{N}))^\perp$ ) on ait :  $\langle u, \mathcal{A}^\sharp v \rangle = \langle f, v \rangle$ .

**Définition 1.5** (Solution forte)

$u \in L^2(\overline{Q})^m$  est appelé solution forte si  $\overline{\mathcal{A}}u = f \in L^2$  avec  $u|_{\partial Q} \in \mathcal{N}$ .

**Remarque** : une solution forte est au sens des distributions et non au sens  $C^1$ .

**Proposition 1.6**

On dira que « faible équivaut à fort » (ie toute solution du problème faible est solution du problème fort, et inversement) si  $\overline{\mathcal{A}^\sharp} = \overline{\mathcal{A}}^* = \mathcal{A}^*$ .

En effet, une solution forte est toujours solution faible par continuité du produit scalaire. Par contre, faible équivaut à fort exige que  $(\mathcal{A}^\sharp)^* = \overline{\mathcal{A}}$  et donc que les domaines de l'opérateur et de l'opérateur adjoint formel soient assez équilibrés.

**Théorème 1.7**

Dans le cas où « faible équivaut à fort », on a équivalence entre les propositions suivantes :

1. Il existe une unique solution au problème  $\overline{\mathcal{A}}u = f$ , avec  $f \in L^2$  avec  $\overline{\mathcal{A}}^{-1}$  continue.
2. Il existe une unique solution au problème  $\overline{\mathcal{A}^\sharp}u = f$ , avec  $f \in L^2$  avec  $(\overline{\mathcal{A}^\sharp})^{-1}$  continue.
3. Les opérateurs  $\mathcal{A}$  et  $\mathcal{A}^\sharp$  sont coercifs.

**Démonstration** : la coercivité de  $\mathcal{A}$  est équivalente à la coercivité de  $\overline{\mathcal{A}}$ . De même pour  $\mathcal{A}^\sharp$  et  $\overline{\mathcal{A}^\sharp}$ . Comme  $\mathcal{A}$  est un opérateur fermé dans  $\mathbf{H}$  un espace de Hilbert, on a équivalence entre les propriétés  $\mathcal{A}$  surjectif et  $\mathcal{A}^*$  coercif ou encore  $\mathbf{Im}(\mathcal{A}^*)$  fermé et  $\ker(\mathcal{A}^*) = \{0\}$  (théorème de surjectivité). De même  $\mathcal{A}^*$  surjectif équivaut à  $\mathcal{A}$

coercif ou encore  $\mathbf{Im}(\mathcal{A})$  fermé et  $\mathbf{ker}(\mathcal{A}) = \{0\}$  (théorème de surjectivité dual). Les opérateurs  $\mathcal{A}$  et  $\mathcal{A}^\sharp$  sont coercifs si, respectivement,

$$\forall u \in \mathcal{D}(\mathcal{A}), |\mathcal{A}u| \geq \alpha|u| \quad \forall v \in \mathcal{D}(\mathcal{A}^\sharp), |\mathcal{A}^\sharp v| \geq \beta|v|. \quad (1.20)$$

Si  $\mathcal{A}$  est injectif, alors  $\mathcal{A}^{-1}$  existe mais si  $\mathcal{A}$  coercif, en posant  $v = \mathcal{A}u$ , on peut réécrire l'inégalité ci dessus sous la forme  $|v| \geq \alpha|\mathcal{A}^{-1}v|$ , ce qui donne la continuité de  $\overline{\mathcal{A}^{-1}}$ . Donc  $\mathcal{A}$  et  $\mathcal{A}^*$  coercifs assurent la continuité de la solution par rapport au second membre.

Donc, pour établir l'existence et l'unicité de la solution, il suffit de montrer la coercivité de  $\mathcal{A}$  et  $\mathcal{A}^\sharp$  pour des fonctions aussi régulières que souhaité. Si l'on prend une restriction  $\overline{\mathcal{A}}$  (resp.  $\overline{\mathcal{A}^\sharp}$ ) quelconque de  $\overline{\mathcal{A}}$  (resp.  $\overline{\mathcal{A}^\sharp}$ ) à un sous-espace dense dans  $\mathbf{H}$ , et si l'on a la propriété « faible équivaut à fort », alors une condition nécessaire et suffisante pour que les problèmes  $\overline{\mathcal{A}}u = f$  et  $\overline{\mathcal{A}^\sharp}v = g$  admettent chacun une unique solution est que  $\mathcal{A}$  et  $\mathcal{A}^\sharp$  soient coercifs par rapport à la norme  $L^2$ . Dans la pratique, on cherche des inégalité du type  $\langle \mathcal{A}u, u \rangle_{L^2} \geq \alpha|u|_{L^2}^2$  ou  $\langle \mathcal{A}^\sharp v, v \rangle_{L^2} \geq \beta|v|_{L^2}^2$  qui impliquent bien les conditions de coercivité sur la norme des opérateurs (1.20).

### Cas instationnaire

Rappel :  $Q = \Omega \times [0, T]$  le domaine spatio-temporel. Soit  $\mathbf{H} = L^2(Q)^m$ . L'opérateur considéré ici est  $\mathcal{A} = \mathbf{A}_0(\mathbf{x})\partial_t + \sum_i \mathbf{A}_i(\mathbf{x})\partial_{x_i} + \mathbf{B}(\mathbf{x})$  et son domaine  $\mathcal{D}(\mathcal{A})$  est l'ensemble des  $u \in C_0^1(\overline{\Omega} \times [0, T])$  si le domaine  $\Omega$  est borné,  $u \in C^1(\overline{\Omega} \times [0, T])$  dans le cas inverse, tels que  $u(\mathbf{x}, 0) = 0$  et  $u|_{\partial\Omega \times [0, T]} \in \mathcal{N}(\mathbf{x})$ .

On rappelle que l'adjoint formel  $\mathcal{A}^\sharp$  de  $\mathcal{A}$  est défini par la relation  $-\mathbf{A}_0(\mathbf{x})\partial_t - \partial_{x_i}(\mathbf{A}_i(\mathbf{x})\cdot) + \mathbf{B}^*(\mathbf{x})$  et son domaine est l'ensemble des  $u \in C_0^1(\overline{\Omega} \times [0, T])$  si le domaine  $\Omega$  est borné, ou  $C^1(\overline{\Omega} \times [0, T])$  dans le cas inverse, tels que  $u(\mathbf{x}, T) = 0$  et  $u|_{\partial\Omega \times [0, T]} \in \mathbf{A}^i \mathbf{n}_i(\mathcal{N}(\mathbf{x}))^\perp$ .

De plus, on suppose que le domaine spatial répond à certaines hypothèses. Soit  $\Omega$  un domaine, borné ou non, de  $\mathbb{R}^n$ , régulier à frontière bornée  $C^1$  et localement d'un seul côté de la frontière. Soit  $\mathcal{N}$  une application à valeurs dans l'ensemble des sous espaces vectoriels de  $\mathbb{C}^{n+1}$  définissant des conditions aux limites homogènes, lipschitziennes et contenant le noyau de  $\mathbf{A}^i \mathbf{n}_i$  sur les composantes connexes de  $\Omega$  où  $\mathbf{A}^i \mathbf{n}_i$  de rang constant et  $C^1$  sur la composante connexe où  $\mathbf{A}^i \mathbf{n}_i$  saute de rang sur une variété immergée différentiable de codimension 2 notée  $\partial\Omega_g$  et  $C^1$  sur le complémentaire de  $\partial\Omega_g$  dans  $\partial\Omega$ .

Dernière hypothèse :  $\forall (\mathbf{x}, t) \in \partial\Omega_g \times [0, T]$ , il existe un ouvert  $\mathcal{V}$  de  $\partial\Omega$  avec  $\mathbf{x} \in \mathcal{V}$  tel que  $\mathcal{V}$  est difféomorphe à une boule  $B$  de  $\mathbb{R}^d$  et  $\partial\Omega_g$  à un équateur de cette boule, et tel qu'il existe deux sous espaces vectoriels  $\mathcal{N}_{small}$  et  $\mathcal{N}_{big}$  de  $\mathbb{C}^n$  tels que  $\mathbf{ker}(\mathbf{A}^i \mathbf{n}_i) \subset \mathcal{N}_{small}$ ,  $\mathcal{N}_{small} \subset \mathcal{N}_{big}$ ,  $\mathcal{N} = \mathcal{N}_{big}$  d'un côté de  $\mathcal{V} \setminus \partial\Omega_g$  et  $\mathcal{N} = \mathcal{N}_{small}$  de l'autre côté.

**Théorème 1.8** (Théorème de Rauch pour les problèmes instationnaires)  
 $\mathcal{N} = \{0\}$  sur  $\Omega \times \{0\}$  et  $\mathcal{N} = \mathbb{R}^m$  sur  $\Omega \times \{T\}$ . Soit  $\mathbf{n}$  la normale unitaire sortante à  $\partial\Omega$ . La propriété « faible équivaut à fort » est vraie dans les cas suivants :

- $\mathbf{A}^i \mathbf{n}_i$  est de rang constant sur toute composante connexe de  $\partial\Omega \times [0, T]$  et la condition initiale est nulle.
- $\mathbf{A}^i \mathbf{n}_i$  change de rang sur  $\Sigma$ , variété de classe  $C^1$  de codimension 1 immergée dans  $\partial\Omega$ . Pour  $\mathbf{x}_0 \in \Sigma$ , il faut que  $\Sigma$  partage  $\partial\Omega$  en deux espaces tels que  $\mathbf{A}^i \mathbf{n}_i$  soit de rang constant de part et d'autre, et que  $\mathcal{N}_{big}(\mathbf{x}_0) \supset \mathcal{N}_{small}(\mathbf{x}_0)$  localement sur  $\Sigma$  et la condition initiale est nulle.

**Remarque :** Soit  $\mathbf{H}_A = \{u \in \mathbf{H}, \mathcal{A}u \in \mathbf{H}\}$ . Alors  $C_0^1(\overline{Q})$  ou  $C_0^1(\overline{\Omega})$  dense  $\mathbf{H}_A$ . Donc, si  $u$  vérifie la condition aux limites souhaitée, on peut définir sa trace modulo  $\mathcal{N}$ . On a alors que  $\mathbf{A}^i \mathbf{n}_i u|_{\partial Q}$  (resp.  $\mathbf{A}^i \mathbf{n}_i u|_{\partial\Omega}$ ) est défini dans  $H^{-\frac{1}{2}}(\partial Q)$  (resp.  $H^{-\frac{1}{2}}(\partial\Omega)$ ).

On rajoute deux hypothèses supplémentaires. La première, notée (H1),  $\mathcal{N}$  est maximal dans le cône de positivité de  $\mathbf{A}^i \mathbf{n}_i$ . La seconde hypothèse est (H2) : il existe  $\alpha \geq 0$  tel que  $\mathbf{B} + \mathbf{B}^* - \partial_i \mathbf{A}^i \geq \alpha \mathbf{I}$ . Cette dernière condition est équivalente à  $\mathbf{B} + \mathbf{B}^* - \partial_i \mathbf{A}^i \geq 0$ .

Si l'on pose  $\varphi = \varphi' e^{\alpha t}$ , avec  $\alpha > 0$ , et  $\mathbf{K} = \mathbf{B} + \mathbf{B}^T - \partial_i \mathbf{A}^i$  l'équation est transformée en une équation similaire mais avec une matrice  $\mathbf{K}' = \mathbf{K} + 2\alpha \mathbf{I}_{n+1}$ , qui en choisissant  $\alpha$  de manière convenable sera toujours définie positive. Ainsi si les coefficients des  $\mathbf{A}^i$  et  $\mathbf{B}$  sont par exemple dans  $W^{1,\infty}$ , on pourra supposer que  $\mathbf{K} > 0$ . Ceci est à la base du choix des conditions aux limites rendant (1.9) bien posé en domaine temporel borné (et uniquement dans ce cas, car le changement d'inconnue  $\varphi$  devient  $\varphi' e^{\alpha t}$  est non homéomorphe sur  $L^2([0, T] \times \Omega)$ ). Le fait de modifier cette matrice  $\mathbf{B}$  ne modifie pas les conditions aux limites décrites par (1.10).

**Remarque :** la condition initiale est la seule condition aux limites maximales monotones sur  $\overline{\Omega} \times \{0\}$ .

### Théorème 1.9

Il existe une unique solution au problème (1.14) avec l'hypothèse (H1).

**Démonstration :** la démonstration s'appuie très fortement sur le théorème de Rauch (1.8). Ce dernier garantit l'obtention de la propriété « faible équivaut à fort ». Ainsi, pour démontrer l'existence et l'unicité de la solution, il ne reste qu'à vérifier la coercivité des opérateurs  $\mathcal{A}$  et  $\mathcal{A}^\sharp$ . Rappelons que  $\forall i \in \llbracket 1, n \rrbracket$ ,  $\mathbf{A}_i = \mathbf{A}_i^*$  (cas réel ou complexe) et  $\mathbf{A}_0$  définie positive (H3).

La notation  $\langle \mathbf{A}^i \mathbf{n}_i u, u \rangle$  fera référence à l'intégrale sur  $\partial\Omega \times [0, T]$  dans le cas de problèmes instationnaires et  $\partial\Omega$  dans le cas harmonique, plus l'ajout des intégrales de bords sur  $\Omega \times \{0\}$  (qui sera nulle vu les hypothèses considérées) et  $\Omega \times \{T\}$ .

**Coercivité de  $\mathcal{A}$**  : Soit  $u \in \mathbf{H}$  tel que  $u|_{\partial Q} \in \mathcal{N}$ . On sait que

$$\langle \mathbf{B}u, u \rangle_{\mathbb{R}^m} = \frac{1}{2} \langle (\mathbf{B} + \mathbf{B}^*)u, u \rangle_{\mathbb{R}^m} \quad (1.21)$$

Dans le cas réel, on a :

$$\begin{aligned} \langle \mathcal{A}u, u \rangle_{L^2} &= \int_Q \langle \mathbf{A}^i \partial_i u + \mathbf{B}u, u \rangle \\ &= \int_Q \langle \partial_i (\mathbf{A}^i u), u \rangle + \int_Q \langle -(\partial_i \mathbf{A}^i)u + \mathbf{B}u, u \rangle \\ \stackrel{\text{(Green)}}{=} &= \int_Q \langle -\mathbf{A}^i u, \partial_i u \rangle + \langle \mathbf{A}^i \mathbf{n}_i u, u \rangle + \int_Q \langle -(\partial_i \mathbf{A}^i)u + \mathbf{B}u, u \rangle \\ \stackrel{\text{(H3)}}{=} &= \int_Q \langle -u, \mathbf{A}_i \partial_i u \rangle + \langle \mathbf{A}^i \mathbf{n}_i u, u \rangle + \int_Q \langle -(\partial_i \mathbf{A}^i)u + \mathbf{B}u, u \rangle \\ &= \frac{1}{2} \langle \mathbf{A}^i \mathbf{n}_i u, u \rangle + \int_Q \langle -(\frac{1}{2} \partial_i \mathbf{A}^i)u + \mathbf{B}u, u \rangle \\ \stackrel{\text{(1.21)}}{=} &= \frac{1}{2} \langle \mathbf{A}^i \mathbf{n}_i u, u \rangle + \frac{1}{2} \langle (\mathbf{B} + \mathbf{B}^* - \partial_i \mathbf{A}^i)u, u \rangle_{L^2} \stackrel{\text{(H1)-(H2)}}{\geq} \frac{1}{2} \alpha |u|_{L^2}^2. \end{aligned}$$

En effet  $\langle \mathbf{A}^i \mathbf{n}_i u, u \rangle \geq 0$  de façon évidente sur  $\partial\Omega \times [0, T]$  puisque  $\mathcal{N}$  est dans le cône de positivité de  $\mathbf{A}^i \mathbf{n}_i$  et  $u \in \mathcal{N}$ . Pour le problème instationnaire, sur  $\Omega \times \{0\}$ ,  $\mathcal{N} = \{0\}$  donc  $\langle \mathbf{A}^i \mathbf{n}_i u, u \rangle = 0$ . Et sur  $\Omega \times \{T\}$ ,  $\mathcal{N} = \mathbb{R}^m$  donc  $\langle \mathbf{A}^i \mathbf{n}_i u, u \rangle = \langle \mathbf{A}_0 u, u \rangle \geq 0$  puisque  $\mathbf{A}_0$  est définie positive.

Dans le cas complexe, la démonstration est analogue en remplaçant le produit scalaire réel par la partie réelle du produit hermitien. Dans tous les cas, l'opérateur  $\mathcal{A}$  est bien coercif.

**Coercivité de  $\mathcal{A}^\sharp$**  : Soit  $v \in \mathbf{H}$  tel que  $v|_{\partial Q} \in [\mathbf{A}^i \mathbf{n}_i(\mathcal{N})]^\perp$ , on a :

$$\langle \mathcal{A}^\sharp v, v \rangle_{L^2} = -\frac{1}{2} \langle \mathbf{A}^i \mathbf{n}_i v, v \rangle_{L^2(\partial Q)} + \langle (\mathbf{B} + \mathbf{B}^* - \partial_i \mathbf{A}^i)v, v \rangle_{L^2} \geq \frac{1}{2} \alpha |v|_{L^2}^2.$$

Pour montrer la coercivité de l'adjoint et le calcul des termes de bord, on sait que la maximale monotonie de  $\mathcal{N}$  implique la négativité de  $\langle \mathcal{A}^\sharp v, v \rangle_{L^2}$ . Sur  $\Omega \times \{T\}$  ce terme est nul et sur  $\Omega \times \{0\}$ , il est égal à  $-\langle \mathbf{A}_0 v, v \rangle$ . Donc l'opérateur  $\mathcal{A}^\sharp$  est coercif. Comme « faible équivaut à fort », le théorème (1.8) donne le résultat proposé.  $\square$

## Cas harmonique

**Théorème 1.10** (Théorème de Rauch pour les problèmes harmoniques)

La propriété « faible équivaut à fort » est vraie dans les cas suivants :

- $\mathbf{A}^i \mathbf{n}_i$  est de rang constant sur toute composante connexe de  $\partial\Omega$ .
- $\mathbf{A}^i \mathbf{n}_i$  change de rang sur  $\Sigma$ , variété de classe  $C^1$  de codimension 1 immergée dans  $\partial\Omega$ . Pour  $\mathbf{x}_0 \in \Sigma$ , il faut que  $\Sigma$  partage  $\partial\Omega$  en deux espaces tels que  $\mathbf{A}^i \mathbf{n}_i$  soit de rang constant de part et d'autre, et que  $\mathcal{N}_{\text{big}}(\mathbf{x}_0) \supset \mathcal{N}_{\text{small}}(\mathbf{x}_0)$  localement sur  $\Sigma$ .

**Remarque :** dans le cas harmonique, l'hypothèse (H2) ne peut être modifiée comme dans le cas instationnaire précédemment détaillé.

**Théorème 1.11**

Si  $\mathbf{H} = L^2_{\mathbf{A}_0}(\Omega)$ , espace où le produit scalaire est défini par  $\langle u, u \rangle = \langle \mathbf{A}_0 u, u \rangle_{L^2(\Omega)}$ , alors les hypothèses (H1) et (H2) impliquent que l'opérateur  $\mathbf{A}_0^{-1}(\overline{\mathcal{A}})$  est maximal monotone.

**Démonstration :** si l'opérateur est fermé, alors son adjoint est maximal monotone. En effet, la maximale monotonie est équivalente à l'existence et l'unicité des solutions pour les problèmes définis par les opérateurs  $\mathcal{I} + \overline{\mathcal{A}}$  et  $\mathcal{I} + \overline{\mathcal{A}}^\sharp$  simultanément. Or si l'opérateur  $\mathcal{A}$  est monotone, alors  $\overline{\mathcal{A}}$  aussi. La propriété « fort équivaut à faible » donne l'égalité  $\overline{\mathcal{A}}^\sharp = \mathcal{A}^*$ . Les hypothèses (Hyp1) et (Hyp2) impliquent alors que l'opérateur  $\mathcal{A}^\sharp$  est monotone, donc  $\alpha\mathcal{I} + \overline{\mathcal{A}}$  et  $\alpha\mathcal{I} + \mathcal{A}^*$  sont coercifs, avec  $\mathbf{Re}(\alpha) > 0$ . Ils sont donc simultanément inversibles, et on a donc la maximale monotonie en passant à la fermeture.  $\square$

**Conséquences :** il existe une unique solution au problème (1.18) mais la norme  $L^2(\Omega)^m$  de la solution est majorée par  $\frac{1}{\varepsilon} \|f\|$ . On ne pourra donc pas conclure quand  $\varepsilon$  tendra vers 0.

Soit  $Q = \Omega$  le domaine spatial. Soit  $\mathbf{H} = L^2_{\mathbf{A}_0}(\Omega)$ . L'opérateur considéré est  $(i\omega + \varepsilon) \mathbf{A}_0(\mathbf{x}) + \sum_i \mathbf{A}_i(\mathbf{x}) \partial_{\mathbf{x}_i} + \mathbf{B}(\mathbf{x})$  et son domaine  $\mathcal{D}(\mathcal{A})$  est l'ensemble des  $u$  appartenant à  $C^1_0(\overline{\Omega})$  si le domaine  $\Omega$  est borné,  $u \in C^1(\overline{\Omega})$  dans le cas inverse, tels que  $u|_{\partial\Omega} \in \mathcal{N}(\mathbf{x})$ . Afin de simplifier les notations, nous adopterons la convention que  $(i\omega + \varepsilon)\mathcal{I} + \mathcal{A} := \mathbf{A}_0(\mathbf{x})^{-1}(\overline{\mathcal{A}}) = (i\omega + \varepsilon)\mathbf{I} + \mathbf{A}_0(\mathbf{x})^{-1}(\sum_i \mathbf{A}_i(\mathbf{x})\partial_{\mathbf{x}_i} + \mathbf{B}(\mathbf{x}))$ .

**Les études en « amplitude limite »**

Un problème harmonique est un problème obtenu en régime périodique établi, qui se définit comme la limite asymptotique en temps d'un problème transitoire entreteu. Nous allons étudier le problème harmonique associé au système instationnaire (1.14) dans le cas où les coefficients des opérateurs  $\mathbf{A}^i$  et  $\mathbf{B}$  ne dépendent pas du temps. La démonstration de l'existence et l'unicité d'une solution au problème où l'on dissymétrise l'espace et le temps découle naturellement du théorème de Hille-Yosida (dans ce cas,  $-\mathcal{A}$  est le générateur d'un semi-groupe) :

**Théorème 1.12**

Si  $f \in C^1(\mathbb{R}^+, (L^2_{\mathbf{A}_0}(\Omega)))$  et si  $\mathbf{B} + \mathbf{B}^* - \partial_i \mathbf{A}^i \geq 0$ , alors le système (1.14) admet une solution unique dans  $C^1(\mathbb{R}^+, (L^2_{\mathbf{A}_0}(\Omega))) \cap C^0(\mathbb{R}^+, \mathcal{D}(\mathcal{A}))$  donnée par la formule de Duhamel :

$$\varphi(t) = e^{-\mathcal{A}t} \varphi_0 + \int_0^t e^{-\mathcal{A}(t-s)} f(s) ds, \quad (1.22)$$

où  $e^{-\mathcal{A}t}$ ,  $t \geq 0$ , est un semi-groupe sur l'espace de Hilbert  $\mathbf{H}$ .

Une manière naturelle de définir un problème harmonique associé à (1.22) est

d'examiner sa « réponse forcée » asymptotique en temps, c'est à dire en remplaçant  $f(t)$  par  $Y(t)g(\mathbf{x}) \mathbf{exp}(i\omega t)$ , où  $Y$  est la fonction de Heaviside de  $\mathbb{R}^+$  et  $g \in \mathbf{H}$ .

Un calcul simple utilisant la formule de Duhamel et les propriétés des semi-groupes montre que si  $i\omega$  est dans l'ensemble résolvant de  $-\mathcal{A}$  (ouvert connexe de  $\mathbb{C}$ ), alors :

$$\varphi(\cdot, t) = \mathbf{exp}(-\mathcal{A}t) (\varphi_0 - (i\omega\mathcal{I} + \mathcal{A})^{-1}g) + \mathbf{exp}(i\omega t) (i\omega\mathcal{I} + \mathcal{A})^{-1}g. \quad (1.23)$$

Le premier terme est appelé terme transitoire. Il correspond à une solution de (1.22) homogène ( $f(t) = 0$ ) avec  $\varphi_0 - (i\omega\mathcal{I} + \mathcal{A})^{-1}g$  comme condition initiale (qui est bien dans  $\mathcal{D}(\mathcal{A})$ ) et la solution forcée  $\mathbf{exp}(i\omega t) (i\omega\mathcal{I} + \mathcal{A})^{-1}g$  périodique en temps. Le cas le plus simple (et le plus rare) où l'on peut définir la solution harmonique comme solution forcée asymptotique en temps de la solution de (1.22) correspond donc au cas où les deux hypothèses suivantes sont vérifiées :

- $i\omega$  est dans le domaine résolvant de  $-\mathcal{A}$ . Ceci sera le cas pour tout  $\omega$  hormis un ensemble discret, par exemple pour  $\Omega$  borné,  $\mathbf{B} + \mathbf{B}^* - \partial_i \mathbf{A}^i \geq 0$  et des conditions aux limites adéquates.
- Le terme transitoire tend vers 0, par exemple si l'opérateur  $-\mathcal{A}$  est strictement dissipatif avec  $\mathbf{B} + \mathbf{B}^* - \partial_i \mathbf{A}^i > 0$  ou des conditions aux limites suffisamment dissipatives.

Dans ce cas favorable, le premier terme tendra exponentiellement vers 0 dans  $H$  et  $\varphi(\cdot, t) \simeq \mathbf{exp}(i\omega t) (i\omega\mathcal{I} + \mathcal{A})^{-1}g$ , pour  $t$  assez grand.

Remarquons que même si  $-\mathcal{A}$  est dissipatif (non strictement),  $\mathcal{A}$  peut être par exemple maximal conservatif, alors  $i\mathcal{A}$  est un opérateur auto-adjoint (corollaire du théorème de Stone) et le spectre de  $\mathcal{A}$  est tout entier sur l'axe des imaginaires purs. Dans le cas où  $\Omega$  est non borné, il peut même être confondu avec l'axe des imaginaires purs, et alors il n'existe pas de  $\omega$  tel que  $i\omega$  soit dans la résolvante de  $-\mathcal{A}$ .

Il s'agit de donner dans un espace convenable un sens à la disparition de la solution transitoire, même quand la formule (1.23) ne peut être écrite faute de possibilité de définition dans  $\mathbf{H}$  de l'inverse de  $i\omega + \mathcal{A}$ . Le principe est la décomposition du spectre de  $i\mathcal{A}$ , quand  $\mathcal{A}$  est maximal conservatif [39], ou plus généralement quand  $\mathcal{A}$  est maximal monotone c'est à dire  $(\mathcal{I} + \mathcal{A})^{-1} \in L(\mathbf{H})$  et  $(\mathcal{A}\varphi, \varphi) \geq 0$ . On a alors le théorème dit théorème de Rage [30], en notant  $\mathbf{H}_b^\perp$  l'orthogonal dans  $\mathbf{H}$  des sous-espaces propres associés à une valeur propre imaginaire pure :

**Théorème 1.13**

Soit  $\mathcal{C} \in L(\mathbf{H})$ . Si  $\mathcal{C}(\mathcal{A} + \mathcal{I})^{-1}$  est compact, alors :

$$\forall f \in \mathbf{H}_b^\perp, \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \|\mathcal{C} \mathbf{exp}(-\mathcal{A}t) f\|_{\mathbf{H}}^2 dt = 0.$$

Ce théorème permet de « contrôler », si  $\mathcal{C}$  est l'opérateur de restriction à un compact spatial dans  $\mathbf{H} = (L^2(\Omega))^{m+1}$ , l'énergie locale produite par la condition initiale, et de définir une solution du problème harmonique dans  $(L_{loc}^2(\Omega))^{m+1}$ .

**Remarque :** l'hypothèse de compacité de  $\mathcal{C}(\mathcal{I} + \mathcal{A})^{-1}$  sera vérifiée si les conditions aux limites sur  $\mathcal{A}$  permettent d'avoir une estimation de la solution dans  $(H_{loc}^1(\Omega))^{m+1}$ . On dit alors que les conditions aux limites sur  $\Omega$  sont coercives et que  $\Omega$  a la propriété de compacité locale.

Lorsque  $\Omega$  est non borné, des difficultés spécifiques se greffent. La forme du bord  $\partial\Omega$  a une grande importance.  $\mathbf{O}$  doit être non « trapping », c'est-à-dire suffisamment convexe pour ne pas piéger l'énergie de la condition initiale et ceci peut être élucidé uniquement en fréquence infinie (non piégeage des rayons).

Outre l'hypothèse de monotonie sur  $\mathcal{A}$ , à savoir  $\mathbf{B} + \mathbf{B}^* - \partial_i \mathbf{A}^i \geq 0$ , une hypothèse de compacité est essentielle pour pouvoir appliquer le théorème de Rage :

$$\exists C > 0, \forall f \in \mathcal{D}(\mathcal{A}) \cap (\ker(\mathcal{A}))^\perp, \quad \sum_j |\partial_{x_j} f| \leq C (\|f\|_{\mathbf{H}} + \|\mathcal{A}(f)\|_{\mathbf{H}}), \quad (1.24)$$

où  $C$  est une constante indépendante de  $f$ . Sous ces hypothèses, on contrôle le terme transitoire de (1.23) et :

$$\forall f \in \mathbf{H}_b^\perp, \quad \lim_{t \rightarrow +\infty} \inf \| \exp(-\mathcal{A}t) f \|_{(L_{loc}^2(\Omega))^{m+1}} = 0. \quad (1.25)$$

### Les études en « absorption limite »

On peut définir plus largement le problème harmonique, c'est-à-dire l'inversion de  $i\omega + \mathcal{A}$  en lui conférant un sens « causal ». Plus précisément, cela signifie, qu'après avoir défini arbitrairement un sens de propagation des ondes, on éliminera dans la solution les ondes provenant de l'infini.

Pour cela, on cherche à donner un sens dans un espace convenable au problème :

$$((i\omega + \varepsilon)\mathcal{I} + \mathcal{A})\varphi = h, \quad \text{quand } \varepsilon \rightarrow 0^+. \quad (1.26)$$

**Remarque :** On aurait pu choisir  $\varepsilon \rightarrow 0^-$  en définissant autrement le sens de propagation, autrement dit le sens direct du repère dans  $\mathbb{R}_t \times \mathbb{R}_x$ .

Si  $\mathcal{A}$  est maximal monotone, (1.26) a toujours une solution unique dans  $\mathbf{H}$ , mais  $\|((i\omega + \varepsilon)\mathcal{I} + \mathcal{A})^{-1}\|_{\mathcal{L}(\mathbf{H})} \leq \frac{1}{\varepsilon}$  explose quand  $\varepsilon \rightarrow 0$ . Il convient alors pour plus de clarté de distinguer deux cas suivant la nature du domaine  $\Omega$ .

#### ► Cas d'un domaine $\Omega$ borné

Si la résolvante de  $-\mathcal{A}$  est compacte (en pratique si l'on vérifie au moins l'hypothèse (1.24)), le spectre de  $-\mathcal{A}$  est ponctuel discret et on est ramené à une alternative de Fredholm. Autrement dit, il existe  $\sigma_p$  discret dans  $\mathbb{R}$  tel que : si  $\omega \notin \sigma_p$ ,  $(i\omega + \mathcal{A})\varphi = f$  a une solution unique dans  $\mathbf{H}$  continue par rapport à  $f$ , et si  $\omega \in \sigma_p$ ,  $i\omega + \mathcal{A}$  est non injectif. Alors  $i\omega$  est alors valeur propre de  $-\mathcal{A}$  et suivant  $f$ ,

$(i\omega + \mathcal{A})\varphi = f$  n'a pas de solution ou une infinité. On dit que  $\omega$  est une pulsation de résonance de  $\Omega$  pour  $\mathcal{A}$ .

► **Cas d'un domaine  $\Omega$  non borné**

Pour les problèmes de diffraction avec obstacle, le spectre de  $-\mathcal{A}$  est souvent continu. Cependant, si l'on suppose que l'opérateur  $\mathcal{A}$  est à résolvante compacte et que  $\mathcal{A}$  est maximal monotone, on pourra montrer que la solution de (1.26) tend généralement vers une limite dans  $(L^2_{loc}(\Omega))^{m+1}$  si  $\Omega$  a la propriété de compacité locale décrite ci-dessus.

**Remarques :** on peut aussi définir des espaces pondérés à l'infini ou bien, lorsque  $\mathcal{A}$  est à coefficients constants en dehors d'un compact, on ajoute à l'équation  $(i\omega + \mathcal{A})\varphi = f$  une condition à l'infini, dite condition de radiation. Que  $\Omega$  soit borné ou non, il existe des solutions du problème d'absorption limite qui ne sont pas solutions du problème d'amplitude limite. Cela revient pour  $\Omega$  borné à « oublier » le terme transitoire de (1.23) et, pour  $\Omega$  non borné, à oublier l'énergie piégée dans un obstacle trapping. Dans tous les cas, l'hypothèse de monotonie est essentielle aux démonstrations. Dans le cas des équations de Maxwell, cette hypothèse est réalisée avec une inégalité du type de (1.24) associée à une décomposition de Hodge.

**1.3.2 Cas de problèmes à coefficients constants avec  $\Omega = \mathbb{R}^3$**

Dans cette section, nous allons considérer la résolution des problèmes sous forme algébrique. Pour résoudre un système linéaire sur un module dont les coefficients de la matrice sont dans une algèbre commutative unitaire incluse dans le module, il suffit de vérifier que le déterminant est un inversible de l'algèbre pour garantir l'existence et l'unicité de la solution.

Les algèbres et modules sont différents selon les problèmes étudiés. Dans le cas instationnaire, l'algèbre sera celle des distributions hyperboliques et le module celui des distributions paraboliques. Dans le cas harmonique, pour  $\mathbf{p} = i\omega + \varepsilon$ , avec  $\varepsilon$  positif, l'algèbre sera l'algèbre de convolution  $\mathcal{O}'_c$  et le module  $\mathcal{S}'$  celui des distributions tempérées. Les diverses définitions sont rappelées dans la section suivante mais on constate que n'interviennent que des algèbres régulières, garantissant ainsi un calcul formel.

**Systèmes symétriques**

**Théorème 1.14**

*Il existe une unique solution au problème instationnaire à coefficients constants (1.16).*

**Définition 1.6** (Noyau de Schwartz – Noyau de Green)

*Le noyau de Schwartz  $\mathcal{K}$  est une distribution à deux variables telle que dans l'espace  $\mathcal{D}'(\Omega \times \Omega)$ ,  $\forall (x, y) \in \Omega \times \Omega$  on ait que  $\langle \langle \mathcal{K}(x, y), \phi(y) \rangle, \varphi(x) \rangle$  définit une*

distribution, dont on saura définir un prolongement par densité. Un opérateur peut donc être défini par une distribution à deux variables. Pour un système à coefficients constants, on parle aussi de noyau de Green convolutif. Dans le cas de coefficients non constants on parle de noyau de Green diadique.

Le noyau de Green matriciel ou résolvante est la matrice de convolution, à savoir « l'inverse » de la matrice de l'opérateur.

Soit  $\Gamma_c$  le cône d'évolution et  $\Gamma_p$  le cône de propagation, polaire de  $\Gamma_c$  dans  $\mathbb{R} + \times \mathbb{R}^n$ .

**Définition 1.7** (Distributions hyperboliques et paraboliques)

[19] Une distribution hyperbolique est une distribution à support inclus dans  $\Gamma_p$ . Une distribution parabolique ou causale est une distribution telle que  $\exists t_0 > 0$  tel que son support soit inclus dans  $[t_0, +\infty[ \times \mathbb{R}^n$ . On peut toujours prendre  $t_0 = 0$  sans perte de généralité par translation temporelle.

Les distributions hyperboliques forment une algèbre unifère, l'élément neutre étant le Dirac  $\delta$ , et régulière ( $A * F = A * G \Rightarrow F = G$  lorsque  $A \neq 0$ ) pour la convolution  $*$ . On notera  $\mathcal{D}'_H(\Gamma_p)$  cette algèbre.

Toute distribution hyperbolique est parabolique.

De plus, les distributions paraboliques (ou causales) forment un module sur cette algèbre et  $\mathcal{D}'_H(\Gamma_p)$  est incluse dans ce module..

Définissons le système de convolution associé au problème instationnaire par :

$$(\mathbf{A}_0\delta_0 + \mathbf{A}_i\delta_i + \mathbf{B}) * X = f. \quad (1.27)$$

Au sens des distributions, dériver dans une direction revient à convoler par la dérivée partielle du Dirac dans la même direction. Il vient le théorème d'évolution suivant.

**Théorème 1.15**

Le problème instationnaire admet une solution unique quelle que soit la distribution parabolique  $f$ .

De plus, le déterminant du système de convolution  $\mathbf{A}_0\delta_0 + \mathbf{A}_i\delta_i + \mathbf{B}$  est un inversible de  $\mathcal{D}'_H(\Gamma_p)$  et  $\Gamma_p$  est l'enveloppe convexe du support de l'inverse du déterminant [16].

**Remarque :** la matrice  $\mathbf{B}$  n'influe en rien dans le résultat précédent [16].

Ceci revient en fait à inverser la matrice  $\mathbf{A} = \mathbf{A}_0\delta_0 + \mathbf{A}_i\delta_i + \mathbf{B}$ , dont les éléments appartiennent à  $\mathcal{D}'_H(\Gamma_p)$ .

**Remarque :** le cône  $\Gamma_p$  est toujours pointé dans  $t \geq 0$  et le déterminant est un polynôme fortement hyperbolique.

**Théorème 1.16**

Si le déterminant est un inversible de l'algèbre, la solution élémentaire n'est autre que :

$$[\mathbf{E}] = \frac{1}{\det(\mathbf{A})} \mathbf{Com}\mathbf{A}^T.$$

A ce moment là, la solution du problème (1.27) où  $f$  est une distribution parabolique, existe et est unique :

$$X = [\mathbf{E}] * f \quad \text{et} \quad \text{supp}[\mathbf{E}] \subset \Gamma_p.$$

**Remarque** : ce théorème implique les principes de Huygens : dans le cas du principe de Huygens fort, on utilise la propriété que le support de la convolution est inclus dans la somme du support de  $f$  et de  $[\mathbf{E}]$ , alors que le principe de Huygens faible s'appuie sur l'égalité entre l'enveloppe convexe du support de la convolution et la somme des enveloppes convexes des supports de  $f$  et de  $[\mathbf{E}]$ . Rappelons que le support de  $[\mathbf{E}]$  est la frontière d'un cône convexe en dimension impaire.

Rappelons que  $\mathcal{S}(\mathbb{R}^n)$  est le sous ensemble de  $\mathcal{E}(\mathbb{R}^n)$  des fonctions indéfiniment différentiables à décroissance rapide ainsi que toutes ses dérivées partielles.

**Définition 1.8 (Algèbre des multiplicateurs)**

On appelle algèbre des multiplicateurs le sous ensemble de  $\mathcal{E}(\mathbb{R}^n)$  des fonctions à croissance lente que l'on note  $\mathcal{O}_M(\mathbb{R}^n)$ . C'est une sous algèbre unifère de  $\mathcal{E}(\mathbb{R}^n)$  et  $\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^n)$  sont des modules unitaires sur  $\mathcal{O}_M(\mathbb{R}^n)$ .

**Définition 1.9 (Algèbre des convoleurs)**

On appelle convoleur une distribution tempérée qui envoie continûment  $\mathcal{S}(\mathbb{R}^n)$  dans  $\mathcal{S}(\mathbb{R}^n)$ . On appelle également ce type de distribution tempérée une distribution tempérée à décroissance rapide et l'on note  $\mathcal{O}'_C(\mathbb{R}^n)$  le sous espace vectoriel de  $\mathcal{S}'(\mathbb{R}^n)$  les contenant. C'est une sous algèbre associative, commutative et unifère de  $\mathcal{E}(\mathbb{R}^n)$  sur lesquels  $\mathcal{S}(\mathbb{R}^n)$  et  $\mathcal{S}'(\mathbb{R}^n)$  sont des modules unitaires.  $\mathcal{O}'_C(\mathbb{R}^n)$  contient les distributions à support compact et les fonctions localement intégrables à décroissance rapide.

**Théorème 1.17 ([19])**

Dans  $\mathcal{S}(\mathbb{R}^n)$  ou  $\mathcal{S}'(\mathbb{R}^n)$  (quand elle existe), la transformée de Fourier échange les produits convolutif et multiplicatif. Elle échange les algèbres (convolutive)  $\mathcal{O}'_C(\mathbb{R}^n)$  et (multiplicative)  $\mathcal{O}_M(\mathbb{R}^n)$ .

De plus,  $\forall T \in \mathcal{S}'(\mathbb{R}^n)$  une distribution tempérée,  $\forall m \in \mathbb{N}$ , il existe une famille finie de fonctions continues  $f_\beta$  avec  $\beta \in \mathbb{N}^n$  et  $|\beta| \leq k(m)$  telle que  $T = \sum_{|\beta| \leq k(m)} D^\beta f_\beta$

avec  $(1 + |M|^2)^m f_\beta \in L^\infty(\mathbb{R}^n)$ ,  $|\beta| \leq k(m)$ .

Ce théorème nous indique que tout distribution à support compact est un convoleur pour  $\mathcal{S}'(\mathbb{R}^n)$ .

**Théorème 1.18 (Paley-Wiener-Schwartz [19])**

Soit  $b \in \mathbb{R}_+$ . Les deux assertions suivantes sont équivalentes pour une fonction  $f$  définie sur  $\mathbb{R}^n$  :

- $f$  est l'image par Fourier d'une distribution  $T \in \mathcal{E}'(\mathbb{R}^n)$  à support contenu dans la boule compacte  $\mathcal{B}(0, b)$
- $f$  est prolongeable par une fonction  $\tilde{f}$  holomorphe sur  $\mathbb{C}^n$  possédant la propriété

suivante :  $\exists m \in \mathbb{N}, C > 0$  tels que  $\forall \xi \in \mathbb{C}^n$  on ait :

$$|\tilde{f}(\xi)| \leq C(1 + |\xi|^2)^{\frac{m}{2}} e^{2\pi b |\operatorname{Im}(\xi)|}$$

Un opérateur est inversible si la matrice associée est un inversible de l'algèbre. L'inverse d'une matrice  $\mathbf{M}$  est donnée par la relation  $\mathbf{M}^{-1} = \frac{1}{\det(\mathbf{M})} \operatorname{Com}(\mathbf{M})^T$ . On en déduit qu'une condition suffisante pour que le problème soit bien posé est que le système avec second membre ait une unique solution dans  $\mathcal{S}'(\mathbb{R}^n)$  et que le déterminant soit un inversible de l'algèbre  $\mathcal{O}'_c(\mathbb{R}^n)$ . On peut en tirer un théorème plus général encore :

**Théorème 1.19**

Si le second membre  $f$  est une distribution tempérée sur  $\mathbb{R}^n$ , à savoir  $f \in \mathcal{S}'(\mathbb{R}^n)$ , pour  $\operatorname{Re}(\mathbf{p}) > 0$ , avec  $\mathbf{p}$  la variable de Laplace, le problème harmonique admet une unique solution dans  $\mathcal{S}'(\mathbb{R}^n)$ .

**Remarque :** si  $\det(\mathbf{A}_0\delta_0 + \mathbf{A}_i\delta_i + \mathbf{B}) \in \mathcal{O}'_c$ , alors la matrice  $\mathbf{B}$  est correctement signée et il n'est pas besoin de considérer davantage d'hypothèses sur elle.

**Théorème 1.20** (Lojasiewicz [34])

Pour toute équation aux dérivées partielles à coefficients constants ( $\mathcal{A}$  est un polynôme de dérivation) il existe au moins une solution élémentaire tempérée et au moins une solution tempérée toutes les fois où le second membre est temporel.

Il existe donc toujours une solution élémentaire pour le déterminant. Comme la convolution par une comatrice à support ponctuel est toujours définie, la solution est toujours tempérée. Cela ne garantit cependant pas l'unicité. Le déterminant peut avoir une infinité d'inverses. Comme  $\mathcal{O}'_c(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$  on pourra utiliser le théorème d'inversion.

### Systèmes non symétriques

Supposons maintenant que les matrices qui interviennent dans le système (1.16) ne soient plus symétriques. Ce paragraphe est motivé par l'étude des équations d'Euler barotropes, non symétriques. Rappelons que la condition de forte hyperbolicité garantit la  $L^2$  stabilité.

**Théorème 1.21** (Kreiss)

Le système (1.16) est  $L^2$  stable ssi  $\mathbf{A}_c^i \xi_i, \xi \in \mathcal{S}^{n-1}$  est autoadjointe par rapport à  $\mathbf{H}(\xi)$  matrice uniformément définie positive, ie indépendamment de  $\xi$ .

**Corollaire 1.22**

Le système est  $L^2$  stable ssi  $\mathbf{A}_c^i \xi_i, \xi \in \mathcal{S}^{n-1}$  a des valeurs propres réelles et que la matrice  $\mathbf{A}_c^i \xi_i$  soit uniformément diagonalisable pour tout  $\xi$ .

**Remarque :** si  $\mathbf{A}_c^i \xi_i$  a des valeurs propres réelles (système hyperbolique éventuellement faiblement hyperbolique) alors le problème (1.16) est bien posé si la partie imaginaire des valeurs propres de  $\mathbf{A}_c^i \xi_i + \mathbf{B}$  est bornée pour tout  $\xi$ .

### 1.3.3 Conclusions

Nous avons donc établi dans cette partie les principaux résultats d'existence et d'unicité pour les problèmes instationnaires et harmoniques étudiés dans cette thèse. La dernière catégorie pose néanmoins quelques problèmes. En effet, pour les équations de Maxwell ou des ondes convectées à coefficients constants, il y a existence et unicité de la solution mais pas nécessairement dans le cas des coefficients variables pour le problème harmonique.

Cependant, la proposition (1.4) permet de ramener l'étude du problème à coefficients variables à celui à coefficients constants. Aussi dans le chapitre suivant serons nous amenés à travailler sur des systèmes PML à coefficients constants, assurés de ne pas perdre en généralité. L'étude sera d'ailleurs menée dans le cas harmonique via l'étude des noyaux de Green. Il n'y a en effet, à notre connaissance, aucune théorie générale pour les problèmes harmoniques à coefficients variables. Dans toute la suite de ce mémoire, il sera supposé que le problème harmonique est bien posé (même en coefficients variables).

Par ailleurs, ce chapitre nous permet d'énoncer le théorème suivant.

**Théorème 1.23**

Soit  $n = 2$  ou  $3$  la dimension d'espace et un problème à coefficients constants de la forme :

$$\mathbf{K}(\mathbf{p}, \mathbf{x})\phi(\mathbf{x}) + \sum_{i=1}^n \mathbf{A}_i^c \frac{\partial \phi(\mathbf{x})}{\partial x_i} = f(\mathbf{x}).$$

Sous les hypothèses suivantes

-  $\mathbf{K}(\mathbf{p}, \mathbf{x}) \in L^\infty$  en  $\mathbf{x}$

-  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  fraction rationnelle en  $\mathbf{p} : \mathbf{K}_{ij} = \frac{P_{ij}(\mathbf{p}, \mathbf{x})}{Q_{ij}(\mathbf{p}, \mathbf{x})}$  avec  $\deg P \leq \deg Q + 1$

-  $\lim_{\mathbf{p} \rightarrow 0} \frac{1}{\mathbf{p}} \mathbf{K}(\mathbf{p}, \mathbf{x}) = \widetilde{\mathbf{A}}_0(\mathbf{x})$  uniformément définie positive,

le système correspondant en temps par transformation de Laplace inverse ( $\mathbf{p} \approx \partial_t$ ) avec des conditions aux limites maximales monotones est bien posé dans  $L^2(\Omega \times [0, T])^m$  où  $\Omega^\infty$  est borné ou égal à  $\mathbb{R}^n$ .

Il peut de plus être écrit in fine sous la forme d'un système de Friedrichs avec ajout d'Equations Différentielles Ordinaires

$$\begin{cases} \widetilde{\mathbf{A}}_0(\mathbf{x}) \frac{\partial \phi(\mathbf{x}, t)}{\partial t} + \mathbf{B}(\mathbf{x})\phi(\mathbf{x}, t) + \mathbf{R}(\mathbf{x})v(\mathbf{x}, t) + \sum_{i=1}^n \mathbf{A}_i^c \frac{\partial \phi(\mathbf{x}, t)}{\partial x_i} = f(\mathbf{x}, t) \\ \frac{\partial v(\mathbf{x}, t)}{\partial t} + \mathcal{E}(\mathbf{x})v(\mathbf{x}, t) + \mathbf{S}(\mathbf{x})\phi(\mathbf{x}, t) = 0 \text{ avec } \mathbf{B}, \mathbf{R}, \mathcal{E}, \mathbf{S} \text{ dans } L^\infty. \end{cases}$$

# Chapitre 2

## Les Perfectly Matched Layers

### 2.1 Introduction

#### 2.1.1 Historique

Jean-Pierre Bérenger [3] a introduit en 1994 ce milieu fictif absorbant qu'est le milieu PML. Considérant le système instationnaire de Maxwell en 2D pour un obstacle parfaitement conducteur, dont les inconnues sont les composantes  $E_x, E_y, H_z$  des champs électrique et magnétique (mode transverse électrique) suivant dans  $\mathbb{R}^2/\mathcal{O}$  :

$$\begin{cases} \frac{\partial}{\partial t} H_z = \frac{\partial}{\partial y} E_x - \frac{\partial}{\partial x} E_y \\ \frac{\partial}{\partial t} E_x = -\frac{\partial}{\partial x} H_z \\ \frac{\partial}{\partial t} E_y = \frac{\partial}{\partial y} H_z \end{cases} \quad (2.1)$$

où  $\mathcal{O}$  contient le support de l'obstacle. Il a alors ajouté une équation au système en découplant l'inconnue  $H_z$  en deux sous composantes  $H_{zx}$  et  $H_{zy}$ , puis en introduisant des facteurs absorbants  $\sigma_x(x)$  et  $\sigma_y(y)$  vérifiant avec  $\mathcal{O} \subset ]-a, a[ \times ]-b, b[$  :

$$\begin{cases} \sigma_x(x) = 0 \text{ sur } ]-a, a[ \\ \sigma_x(x) > 0 \text{ sur } \mathbb{R}/]-a, a[ \\ \sigma_y(y) = 0 \text{ sur } ]-b, b[ \\ \sigma_y(y) > 0 \text{ sur } \mathbb{R}/]-b, b[ \end{cases}$$

Le système instationnaire qu'il propose est dans  $\mathbb{R}^2/\mathcal{O}$  :

$$\begin{cases} \frac{\partial}{\partial t} H_{zx} + \sigma_y H_{zy} = \frac{\partial}{\partial y} E_x \\ \frac{\partial}{\partial t} H_{zy} + \sigma_x H_{zx} = -\frac{\partial}{\partial x} E_y \\ \frac{\partial}{\partial t} E_x + \sigma_y E_x = \frac{\partial}{\partial y} H_z \\ \frac{\partial}{\partial t} E_y + \sigma_x E_y = -\frac{\partial}{\partial x} H_z \end{cases} \quad (2.2)$$

Lorsque  $\sigma_x = \sigma_y = 0$ , les deux premières équations de (2.2) sont regroupables pour n'en former plus qu'une, et l'on retrouve ainsi le système de Maxwell originel. De plus, si Jean-Pierre Bérenger a dans un premier temps considéré la diffraction des ondes électromagnétiques par un obstacle parfaitement conducteur, le cas du diélectrique parfait peut également être traité par cette méthode. Il suffit dans tous les cas de disjointre le support de l'obstacle du milieu PML ( $\sigma_x \neq 0$  ou  $\sigma_y \neq 0$ ).

L'opérateur associé au système (2.2) est faiblement hyperbolique mais pas fortement hyperbolique, contrairement à celui du système de Maxwell associé. Le symbole complet de l'opérateur ne vérifie pas les conditions d'Hadamard Petrovsky et donc le problème est mal posé en unicité. Il convient donc de modifier les équations de ce système tout en conservant les propriétés de ce milieu fictif.

Dès 1994, W.C. Chew et W.H. Weedon [4] ont proposé à l'aide d'une transformation formelle par Laplace des équations et d'un regroupement judicieux de celles ci, une interprétation du milieu PML pour les équations de Maxwell en 3D en termes d'un changement de coordonnées complexes. Ils démontrent en outre une décroissance exponentielle de l'amplitude des ondes planes dans le milieu fictif.

De cette interprétation, F. Collino et P. Monk [5] ont démontré l'existence et l'unicité du problème des PML harmoniques en coordonnées polaires pour l'équation de Helmholtz scalaire 2D en milieu non borné, et en milieu borné et pour toute fréquence, excepté un ensemble discret et localement fini, pouvant être vide. Les tests numériques qu'ils ont effectués à l'aide d'une méthode d'éléments finis ont confirmé l'intérêt de cette approche. Une extension au problème écrit en coordonnées cartésiennes est présentée dans [5] et pour les équations de Maxwell dans [27].

Dans [21], l'auteur travaille sur les équations de Helmholtz pour des domaines de forme convexe. Il établit ainsi l'existence et l'unicité de la solution pour toute fréquence en harmonique assez loin du domaine. Actuellement, les PML sont largement répandues pour borner les domaines de calcul quand ceux ci ont des géométries cartésiennes ou convexes. Le but de cette thèse est d'introduire le formalisme pour des géométries non convexes, ou plus exactement, le formalisme le plus général possible.

Nous travaillerons avec des Equations admettant des formulations intrinsèques, comme les Equations de Maxwell, ou des ondes. En s'appuyant sur les travaux de [20], nous définirons un changement de coordonnées complexes qui sera le plongement dans une sous variété réelle de  $\mathbb{C}^n$  à fibré tangent et cotangent complexes,  $n$  étant la dimension d'espace. A l'aide de la métrique induite par ce changement, nous expliciterons la transformée de Hodge qui permet d'aboutir à la formulation des nouvelles Equations dans le milieu PML.

Commençons tout d'abord par écrire des formulations cartésiennes et cylindriques, qui nous serviront de solutions de références pour les expériences numériques. De plus, le formalisme étant général, nous vérifierons qu'il correspond bien

aux cas convexes classiques précédemment cités.

## 2.1.2 Les formulations convexes classiques

### Les PML cartésiennes

Soit un domaine d'étude cartésien  $\mathbf{D} = [-a, a] \times [-b, b]$ . Soit  $\mathbf{p} = i\omega + \varepsilon$  avec  $\varepsilon \geq 0$ . Soit les équations de Maxwell transformées formelles de Laplace en temps :

$$\begin{cases} \mathbf{p} E - \nabla \times H = F \\ \mathbf{p} H + \nabla \times E = G \end{cases} \Leftrightarrow \mathbf{p} \begin{pmatrix} E \\ H \end{pmatrix} + \begin{pmatrix} 0 & -\frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} & 0 & -\frac{\partial}{\partial x} \\ -\frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \end{pmatrix} \begin{pmatrix} E \\ H \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix}, \quad (2.3)$$

où  $(F, G)$  sont à support compact dans le domaine  $\mathbf{D}$ . On introduit le changement de variable complexe suivant :

$$\tilde{x} = x + \frac{1}{\mathbf{p}} \int \sigma_x(\tau) d\tau, \quad \tilde{y} = y + \frac{1}{\mathbf{p}} \int \sigma_y(\tau) d\tau, \quad \tilde{z} = z + \frac{1}{\mathbf{p}} \int \sigma_z(\tau) d\tau,$$

où les fonctions  $\sigma_x$ ,  $\sigma_y$  et  $\sigma_z$  vérifient les mêmes propriétés que précédemment (on pourra se reporter à la figure (2.1) pour un exemple bidimensionnel).

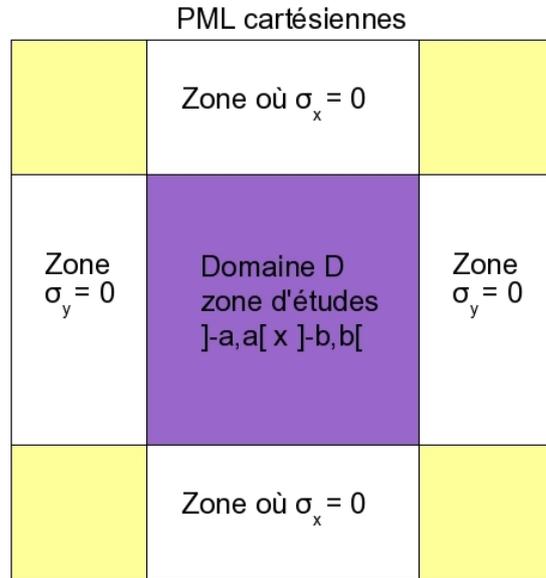


FIG. 2.1: Exemple de domaines et propriétés des fonctions  $\sigma$

On constate que  $\tilde{x}$  (resp.  $\tilde{y}$ ,  $\tilde{z}$ ) ne dépend que de  $x$  (resp.  $y$ ,  $z$ ). Ce cas est simple car le jacobien de la transformation est une matrice diagonale. Les opérateurs de dérivations spatiales  $\frac{\partial}{\partial x}$ ,  $\frac{\partial}{\partial y}$  et  $\frac{\partial}{\partial z}$  présents dans les termes  $\nabla \times$  sont alors remplacés

par des opérateurs de dérivations à coefficients complexes  $\frac{\partial}{\partial \tilde{x}}$ ,  $\frac{\partial}{\partial \tilde{y}}$  et  $\frac{\partial}{\partial \tilde{z}}$ .

Soit :

$$s_x = 1 + \frac{1}{\mathbf{p}}\sigma_x, \quad s_y = 1 + \frac{1}{\mathbf{p}}\sigma_y, \quad s_z = 1 + \frac{1}{\mathbf{p}}\sigma_z.$$

Alors :

$$\frac{\partial}{\partial \tilde{x}} = \frac{1}{s_x} \frac{\partial}{\partial x}, \quad \frac{\partial}{\partial \tilde{y}} = \frac{1}{s_y} \frac{\partial}{\partial y}, \quad \frac{\partial}{\partial \tilde{z}} = \frac{1}{s_z} \frac{\partial}{\partial z}.$$

Le système (2.3) est remplacé par :  $\begin{cases} \mathbf{p} \tilde{E} - \widetilde{\nabla} \times \tilde{H} = F \\ \mathbf{p} \tilde{H} + \widetilde{\nabla} \times \tilde{E} = G, \end{cases}$  où

$$\widetilde{\nabla} \times = \begin{pmatrix} 0 & -\frac{1}{s_z} \frac{\partial}{\partial z} & \frac{1}{s_y} \frac{\partial}{\partial y} \\ \frac{1}{s_x} \frac{\partial}{\partial x} & 0 & -\frac{1}{s_z} \frac{\partial}{\partial z} \\ -\frac{1}{s_y} \frac{\partial}{\partial y} & \frac{1}{s_x} \frac{\partial}{\partial x} & 0 \end{pmatrix}$$

Il faut ensuite factoriser cette matrice pour obtenir la formulation PML [10].  
Posons :

$$\mathbf{M} = \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{pmatrix} \quad \mathbf{N} = \frac{1}{s_x s_y s_z} \mathbf{M} \quad \tilde{E} = \mathbf{M} \tilde{E} \quad \tilde{H} = \mathbf{M} \tilde{H}.$$

D'où les équations de Maxwell dans les couches PML :

$$\begin{cases} \mathbf{p} \mathbf{N}^{-1} \mathbf{M}^{-1} \tilde{E} - \nabla \times \tilde{H} = F \\ \mathbf{p} \mathbf{N}^{-1} \mathbf{M}^{-1} \tilde{H} + \nabla \times \tilde{E} = G. \end{cases} \quad (2.4)$$

Cette écriture peut être interprétée comme la perturbation d'un opérateur différentiel par un opérateur d'ordre 0 et permet d'être implémentée sur des schémas de type Volumes Finis par exemple. Le système peut alors se relever en temps, une fois la matrice  $\mathbf{N}^{-1} \mathbf{M}^{-1}$  décomposée en éléments simples.

$$\begin{aligned} \mathbf{p} \mathbf{N}^{-1} \mathbf{M}^{-1} &= \begin{pmatrix} \frac{s_y s_z}{s_x} & 0 & 0 \\ 0 & \frac{s_x s_z}{s_y} & 0 \\ 0 & 0 & \frac{s_x s_y}{s_z} \end{pmatrix} \\ &= \begin{pmatrix} \frac{(\mathbf{p} + \sigma_y)(\mathbf{p} + \sigma_z)}{(\mathbf{p} + \sigma_x)} & 0 & 0 \\ 0 & \frac{(\mathbf{p} + \sigma_x)(\mathbf{p} + \sigma_z)}{(\mathbf{p} + \sigma_y)} & 0 \\ 0 & 0 & \frac{(\mathbf{p} + \sigma_x)(\mathbf{p} + \sigma_y)}{(\mathbf{p} + \sigma_z)} \end{pmatrix} \\ &= \mathbf{p} \mathbf{I} + \mathbf{A} + \mathbf{R}_1 (\mathbf{p} \mathbf{I} + \mathbf{C})^{-1} \mathbf{R}_2, \end{aligned}$$

avec :

$$\begin{aligned}\mathbf{A} &= \mathbf{diag}(\sigma_z + \sigma_y - \sigma_x, \sigma_x + \sigma_z - \sigma_y, \sigma_x + \sigma_y - \sigma_z) \\ \mathbf{C} &= \mathbf{diag}(\sigma_x, \sigma_y, \sigma_z) \\ \mathbf{R}_1 &= \mathbf{diag}(\sigma_x - \sigma_y, \sigma_z - \sigma_y, \sigma_y - \sigma_z) \\ \mathbf{R}_2 &= \mathbf{diag}(\sigma_x - \sigma_z, \sigma_x - \sigma_y, \sigma_x - \sigma_z).\end{aligned}$$

Le système (2.4) peut alors être relevé en temps sous la forme :

$$\begin{cases} \partial_t E + \mathbf{A}E + \mathbf{R}_1 D - \nabla \times H = 0 \\ \partial_t H + \mathbf{A}H + \mathbf{R}_1 B + \nabla \times E = 0, \end{cases}$$

où  $(B, D)$  sont solutions des Equations aux Dérivées Ordinaires suivantes :

$$\begin{cases} \partial_t D + \mathbf{C}D - \mathbf{R}_2 E = 0 \\ \partial_t B + \mathbf{C}B - \mathbf{R}_2 H = 0. \end{cases}$$

On obtient alors un système dit de Friedrichs, facilement implémentable et dont on démontrera dans la partie suivante l'existence et l'unicité de solutions.

### Les PML cylindriques

Le principe reste le même.  $\theta$  restant invariant, le changement de variable complexe se résume à :

$$\tilde{\rho} = \rho + \frac{1}{\mathbf{p}} \tilde{\sigma}(\rho, \theta) = \rho + \frac{1}{\mathbf{p}} \int_{\rho_0}^{\rho} \sigma(\tau) d\tau,$$

où  $\rho_0 = \partial\Omega$  ne dépend pas de  $\theta$  et  $\sigma$  respecte toujours les mêmes propriétés. Posons  $\beta = 1 + \frac{1}{\mathbf{p}} \sigma$  et  $\tilde{\beta} = 1 + \frac{1}{\mathbf{p}} \tilde{\sigma}$ . Notons  $\widetilde{\nabla \times}$  l'opérateur  $\nabla \times$  exprimé en coordonnées polaires. Alors :

$$\widetilde{\nabla \times} = \begin{pmatrix} 0 & 0 & \partial_\rho \\ 0 & 0 & \frac{1}{\rho} \partial_\theta \\ \partial_\rho & \frac{1}{\rho} \partial_\theta & 0 \end{pmatrix}.$$

La formulation PML cylindrique des équations de Maxwell (2.3) devient alors :

$$\begin{cases} \mathbf{pM} \tilde{\tilde{E}} - \widetilde{\nabla \times} \tilde{\tilde{H}} = 0 \\ \mathbf{pM} \tilde{\tilde{H}} + \widetilde{\nabla \times} \tilde{\tilde{E}} = 0. \end{cases} \quad \text{avec } \mathbf{pM} = \mathbf{p} \begin{pmatrix} \beta & 0 & 0 \\ \tilde{\beta} & \tilde{\beta} & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \beta \tilde{\beta} \end{pmatrix}$$

On obtient une décomposition de la matrice  $M$  en éléments simples comme précédemment sous la forme :  $\mathbf{pM} = \mathbf{pI} + \mathbf{A} + \mathbf{R}_1 (\mathbf{pI} + \mathbf{C})^{-1} \mathbf{R}_2$ , avec :

$$\begin{aligned} \mathbf{A} &= \mathbf{diag}(\sigma - \tilde{\sigma}, \tilde{\sigma} - \sigma, \sigma + \tilde{\sigma}) \\ \mathbf{C} &= \mathbf{diag}(\tilde{\sigma}, \sigma, 0) \\ \mathbf{R}_1 &= \mathbf{diag}(\tilde{\sigma} - \sigma, \sigma - \tilde{\sigma}, \sigma) \\ \mathbf{R}_2 &= \mathbf{diag}(\tilde{\sigma}, \sigma, \tilde{\sigma}). \end{aligned}$$

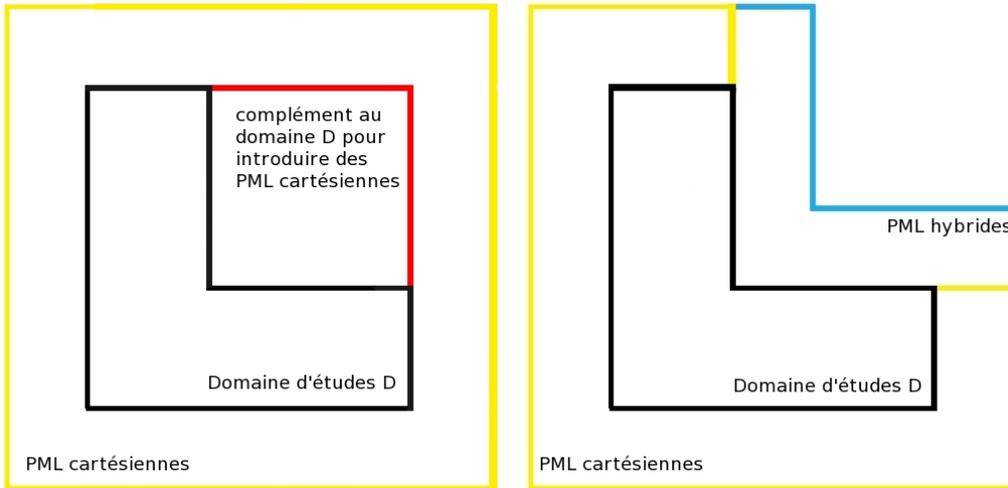
Ce système peut alors être relevé en temps sous la forme d'un système de Friedrichs, comme pour les PML cartésiennes :

$$\begin{cases} \partial_t E + \mathbf{A}E + \mathbf{R}_1 D - \nabla \times H = 0 \\ \partial_t H + \mathbf{A}H + \mathbf{R}_1 B + \nabla \times E = 0, \end{cases}$$

où  $(B, D)$  sont solutions des Equations aux Dérivées Ordinaires suivantes :

$$\begin{cases} \partial_t D + \mathbf{C}D - \mathbf{R}_2 E = 0 \\ \partial_t B + \mathbf{C}B - \mathbf{R}_2 H = 0. \end{cases}$$

**Remarque** : ces deux méthodes de PML sont largement utilisées. Cependant, dans le cas d'un domaine de forme L, utiliser des PML non convexes apportera un gain significatif. Bien évidemment ce gain serait encore plus important en 3D.



Quand le nombre d'inconnues tend vers l'infini, le gain tend lui aussi vers l'infini : en bidimensionnel en  $\mathcal{O}(N)$ , où  $N$  représente le nombre d'inconnues, et en tridimensionnel en  $\mathcal{O}(N^2)$ . Les domaines présentés sur la figure précédente seront d'ailleurs ceux utilisés pour une comparaison numérique (voir section Numérique).

### 2.1.3 Le cas monodimensionnel

Soit le problème défini sur  $] \alpha, \beta[ \times \mathbb{R}^+$  suivant :

$$\frac{\partial \phi(x, t)}{\partial t} + \mathbf{A} \frac{\partial \phi(x, t)}{\partial x} = f(x, t), \quad (2.5)$$

avec  $f$  à support dans l'intervalle  $] \alpha, \beta[ \times \mathbb{R}^+$  et les conditions aux limites pour tout  $t \in \mathbb{R}^+$  :

$$\mathbf{A}^+ \phi(\alpha, t) = 0 \quad \mathbf{A}^- \phi(\beta, t) = 0.$$

En résolvant ce problème sur  $] \alpha, \beta[$  avec les conditions aux limites précédentes, on trouve la restriction de  $\phi$  sur  $] \alpha, \beta[$ . Le noyau de Schwartz est très régulier, il s'agit d'une fonction constante  $L_{loc}^1$  à l'intérieur du cône de propagation, et nulle ailleurs. La localisation est donc facile et devient algébrique.

#### Définition 2.1

Sur un intervalle de type  $] \alpha, \beta[$ , on a le problème dit *Matched Layers (ML)* de Bérenger suivant :

$$\frac{\partial \tilde{\phi}(x, t)}{\partial t} + \sigma(x) \tilde{\phi}(x, t) + \mathbf{A} \frac{\partial \tilde{\phi}(x, t)}{\partial x} = f(x, t).$$

La fonction  $\sigma$  est nulle sur  $] \alpha, \beta[$  et strictement positive sinon.

Les solutions de (2.1) et (2.5) coïncident sur  $] \alpha, \beta[$ . L'avantage de ce problème est qu'une perturbation en  $x_0 \notin ] \alpha, \beta[$  revient sur  $] \alpha, \beta[$  pondérée par  $\exp\left(-\int_{\beta}^{x_0} \sigma(x) dx\right)$ .

On obtient alors une décroissance exponentielle de l'erreur dans le retour et donc une très bonne approximation du problème.

**Remarque :** cette propriété est encore vraie pour des problèmes de la forme :

$$\frac{\partial \phi(\mathbf{x}, t)}{\partial t} + \sum_{i=1}^n \mathbf{A}_i \frac{\partial \phi(\mathbf{x}, t)}{\partial x_i} = f(\mathbf{x}, t),$$

si  $\forall (i, j) \in \llbracket 1, n \rrbracket$ ,  $[\mathbf{A}^i, \mathbf{A}^j] = 0$ , autrement dit si les  $\mathbf{A}^i$  sont sur une même base de vecteurs propres et qu'il est alors possible de se ramener à un ensemble de problèmes scalaires monodimensionnels.

### 2.1.4 Les PML de Bérenger revisités

L'idée énoncée précédemment serait de doubler les inconnues afin de se ramener à des problèmes monodimensionnels. Considérons le problème suivant sur  $\mathbb{R}^2 \times \mathbb{R}^+$  :

$$\frac{\partial \phi(x, y, t)}{\partial t} + \mathbf{A}_1 \frac{\partial \phi(x, y, t)}{\partial x} + \mathbf{A}_2 \frac{\partial \phi(x, y, t)}{\partial y} = 0. \quad (2.6)$$

On décompose  $\phi(x, y, t) = \phi_1(x, y, t) + \phi_2(x, y, t)$ , où  $\phi_1$  et  $\phi_2$  sont solutions de (2.6). On a alors le système suivant :

$$\begin{cases} \frac{\partial \phi_1(x, y, t)}{\partial t} + \mathbf{A}_1 \frac{\partial \phi_1(x, y, t)}{\partial x} + \mathbf{A}_1 \frac{\partial \phi_2(x, y, t)}{\partial x} = 0 \\ \frac{\partial \phi_2(x, y, t)}{\partial t} + \mathbf{A}_2 \frac{\partial \phi_1(x, y, t)}{\partial y} + \mathbf{A}_2 \frac{\partial \phi_2(x, y, t)}{\partial y} = 0 \end{cases} \quad (2.7)$$

**Remarque** : il sera possible d'économiser des dédoublements en tenant compte des zéros des matrices  $\mathbf{A}_i$  en présence.

### Propriété 2.1

Le problème (2.7) est hyperbolique.

**Démonstration** : soit  $\mu(\xi, \eta)$  une valeur propre de la matrice  $\mathbf{M} = \begin{pmatrix} \mathbf{A}_1 \xi & \mathbf{A}_1 \xi \\ \mathbf{A}_2 \eta & \mathbf{A}_2 \eta \end{pmatrix}$ .

Alors  $\det(\mathbf{M} - \mu(\xi, \eta)\mathbf{I}) = \det \begin{pmatrix} \mathbf{A}_1 \xi - \mu(\xi, \eta)\mathbf{I} & \mathbf{A}_1 \xi \\ \mathbf{A}_2 \eta & \mathbf{A}_2 \eta - \mu(\xi, \eta)\mathbf{I} \end{pmatrix}$ . En sommant les lignes, on obtient :

$$\det(\mathbf{M} - \mu(\xi, \eta)\mathbf{I}) = \det \begin{pmatrix} \mathbf{A}_1 \xi - \mu(\xi, \eta)\mathbf{I} & \mathbf{A}_1 \xi \\ \mathbf{A}_1 \xi + \mathbf{A}_2 \eta - \mu(\xi, \eta)\mathbf{I} & \mathbf{A}_1 \xi + \mathbf{A}_2 \eta - \mu(\xi, \eta)\mathbf{I} \end{pmatrix}.$$

Soit  $\lambda(\xi, \eta)$  une valeur propre de la matrice  $\mathbf{A}_1 \xi + \mathbf{A}_2 \eta$ . Alors,  $\mu(\xi, \eta) = \lambda(\xi, \eta)$ . Toutes les valeurs propres du symbole sont donc réelles. Cependant, il va exister une perturbation d'ordre 0 rendant le problème mal posé : le système est en réalité faiblement hyperbolique.

**Lemme 2.2** (Dans [17])

L'uniforme diagonalisabilité de  $a$  (endomorphisme) par rapport à  $N$  (norme) est équivalente à la forte hyperbolicité par rapport à  $N$  de l'opérateur différentiel  $a(D)$ .

On pourra montrer en général que la matrice  $\mathbf{M}$  n'est pas uniformément diagonalisable. Par exemple, dans le cas scalaire, soit  $(a, b)$  et  $\xi = \eta = 1$ . Donc  $\mathbf{M} = \begin{pmatrix} a & a \\ b & b \end{pmatrix}$ . Le déterminant de cette matrice est nul, donc elle est semblable à une matrice triangulaire, ce qui prouve la non uniforme diagonalisabilité.

Pour les Equations de Maxwell, les matrices  $\mathbf{A}_i$  sont données (au signe près) par  $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$  et  $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$ . Ces matrices ont les mêmes valeurs propres (0, 1 et -1). Ce sont donc les valeurs propres de la somme et, a fortiori, de  $\mathbf{M}$ , qui n'est donc pas uniformément diagonalisable.  $\square$

Bien que ce problème sous forme ML puisse être mal posé, l'écriture est quand même tentée, de la même façon que pour le problème monodimensionnel, en introduisant des fonctions  $\sigma_1$  et  $\sigma_2$  strictement positives en dehors du domaine, et nulles à l'intérieur. Le problème (2.7) devient alors :

$$\begin{cases} \frac{\partial \phi_1(x, y, t)}{\partial t} + \sigma_1(x)\phi_1(x, y, t) + \mathbf{A}_1 \frac{\partial \phi_1(x, y, t)}{\partial x} + \mathbf{A}_1 \frac{\partial \phi_2(x, y, t)}{\partial x} = 0 \\ \frac{\partial \phi_2(x, y, t)}{\partial t} + \sigma_2(y)\phi_2(x, y, t) + \mathbf{A}_2 \frac{\partial \phi_1(x, y, t)}{\partial y} + \mathbf{A}_2 \frac{\partial \phi_2(x, y, t)}{\partial y} = 0 \end{cases}$$

**Propriété 2.3** (Dans [4])

Pour des fonctions  $\sigma_1$  et  $\sigma_2$  constantes par morceaux, et pour des systèmes de Maxwell en bidimensionnel ou tridimensionnel (ainsi que pour l'équation des ondes en bidimensionnel), les ondes planes solutions du problème dédoublé ne sont ni diffractées, ni réfléchies sur le support de  $\sigma_1 \cup \sigma_2$  et leur amplitude décroît en  $\exp\left(-\int_{\text{trajet}} \sigma_i\right)$

### 2.1.5 Pourquoi le changement de variable complexe semble la piste la plus naturelle ?

W.C. Chew et W.H. Weedon [4] introduisent pour la première fois la notion de changement de variable complexe. Cette piste fournit la théorie la plus acceptable en harmonique et en instationnaire de la justification de l'idée de Bérenger de passer d'une équation ML à  $n$  équations PML par dédoublement des inconnues.

Soit  $\mathbf{p} = i\omega + \varepsilon \in \mathbb{C}$ . Pour les problèmes harmoniques, nous remplacerons l'opérateur  $\partial_t$  par  $\mathbf{p}$  en faisant tendre  $\varepsilon$  vers 0. Ecrivons maintenant les PML en dédoublant les inconnues. On introduit le changement de variables cartésiens de la même façon que dans (2.1.2) à savoir :

$$\tilde{x} = x + \frac{1}{\mathbf{p}} \int \sigma_1(\tau) d\tau, \quad \tilde{y} = y + \frac{1}{\mathbf{p}} \int \sigma_2(\tau) d\tau.$$

On pose  $s_1 = 1 + \frac{1}{\mathbf{p}}\sigma_1$  et  $s_2 = 1 + \frac{1}{\mathbf{p}}\sigma_2$ . Alors  $\frac{\partial}{\partial \tilde{x}} = \frac{1}{s_1} \frac{\partial}{\partial x}$ , et  $\frac{\partial}{\partial \tilde{y}} = \frac{1}{s_2} \frac{\partial}{\partial y}$ . Le système d'équations devient alors dans les couches PML :

$$\mathbf{p} \mathbf{M} \begin{pmatrix} \phi_1(x, y, t) \\ \phi_2(x, y, t) \end{pmatrix} + \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_1 \\ \mathbf{A}_2 & \mathbf{A}_2 \end{pmatrix} \begin{pmatrix} \phi_1(x, y, t) \\ \phi_2(x, y, t) \end{pmatrix} = 0, \quad (2.8)$$

$$\text{avec } \mathbf{p} \mathbf{M} = \begin{pmatrix} \frac{\mathbf{p} + \sigma_2}{\mathbf{p} + \sigma_1} & 0 & 0 \\ 0 & \frac{\mathbf{p} + \sigma_1}{\mathbf{p} + \sigma_2} & 0 \\ 0 & 0 & \frac{(\mathbf{p} + \sigma_1)(\mathbf{p} + \sigma_2)}{\mathbf{p}} \end{pmatrix}.$$

L'idée de Bérenger est de faire intervenir la structure de l'opérateur  $\mathbf{p} + \mathbf{A}^i \partial_i$  à travers celle d'un opérateur scalaire  $\det(\mathbf{p}\mathbf{I} + \mathbf{A}^i \partial_i)$  qui est peut être elliptique dans le cas par exemple des équations de Maxwell ou des Ondes. Cette remarque

permettra la généralisation à d'autres systèmes de Friedrichs. Elle permet de comprendre une structure dispersive équivalente, et de généraliser à d'autres systèmes de coordonnées, de milieux homogènes (à pertes, dispersif, ...) avec des équations approximatives.

Dans le cas d'opérateurs elliptiques, le noyau de Schwartz est holomorphe et n'a qu'une singularité en 0. Soit la base naturelle du fibré tangent  $\frac{\partial M}{\partial \tilde{\mathbf{x}}_i}$ . Si les  $\tilde{\mathbf{x}}_i$  définissent un plongement de la variété de  $\tilde{\mathbf{x}}(\mathbf{x})$ , alors les  $\partial_{\tilde{\mathbf{x}}_i}$  sont des repères naturels de l'espace tangent et les crochets de Lie commutent. La base naturelle est un repère qui ne dépend pas de la carte choisie et dans lequel la composition de dérivée est encore une dérivée. La commutativité de la composition des dérivées n'est vraie que si l'on a une forme différentielle. Dans le cas d'un changement de coordonnées polaires, il existe deux bases fréquemment utilisées. La base orthonormée n'est pas la base naturelle (division par  $r$  sur la deuxième composante).

#### Propriété 2.4

Dans notre cas, les  $\partial_{\tilde{\mathbf{x}}_i}$  sont une base dans le fibré tangent complexe d'une variété réelle  $C^\infty$  de  $\mathbb{C}^n$ . On a donc la nullité des crochets de Lie, à savoir la propriété :

$$\forall (i, j) \in \llbracket 1, n \rrbracket, \quad [\partial_{\tilde{\mathbf{x}}_i}, \partial_{\tilde{\mathbf{x}}_j}] = 0.$$

**Démonstration** : Soit  $\partial_{\tilde{\mathbf{x}}_i} = b_i^k(\mathbf{x})\partial_{\mathbf{x}^k}$  avec pour tout  $\mathbf{x}$ ,  $b_i^k(\mathbf{x}) \in \mathcal{GL}(\mathbb{C}^n)$  alors  $[\partial_{\tilde{\mathbf{x}}_i}, \partial_{\tilde{\mathbf{x}}_j}] = \gamma^r \partial_{\mathbf{x}^r}$  avec  $\gamma^r = b_i^k \partial_{\mathbf{x}^k}(b_j^r) - b_j^k \partial_{\mathbf{x}^k}(b_i^r)$ . En posant  $a_k^j b_i^k = \delta_i^j$  on obtient que  $a_k^j \gamma^r = b_i^k \partial_{\mathbf{x}^k}(b_j^r) - b_j^k \partial_{\mathbf{x}^k}(b_i^r) = 0$ , ce qui implique la nullité de  $\gamma^r$ .  $\square$

Cette condition est très générale mais nécessite de travailler avec les bases adaptées. En effet, pour garantir l'invariance du changement, il faut de plus avoir la nullité des crochets de Lie pour les dérivées covariantes. La condition précédente peut s'écrire sous la forme  $[\nabla_{\partial_{\tilde{\mathbf{x}}_i}}, \nabla_{\partial_{\tilde{\mathbf{x}}_j}}] = 0$ . Pour avoir une écriture intrinsèque et un calcul commutatif, il faut que la propriété de nullité soit vraie pour tout vecteur (et notamment les  $\partial_{\mathbf{x}_i}$ ). D'où la condition supplémentaire :

$$\forall (u, v) \in \llbracket 1, n \rrbracket, \quad [\nabla_u, \nabla_v] = 0.$$

Mais avec ce changement de variable complexe, est-il toujours possible d'avoir une formulation dispersive ? Obtenir cette dernière signifie qu'il existe une matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  telle que l'opérateur s'écrive  $\mathbf{K}(\mathbf{p}, \mathbf{x})\mathbf{I} + \mathbf{A}^i \partial_i$  et présente quatre intérêts.

Le premier est que ces systèmes sont des systèmes de Friedrichs, candidats pour les approximations habituelles et sur lesquels nous disposons de nombreux résultats (cf. premier chapitre).

Le deuxième est lié à la forme de la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$ . Si il s'agit d'une fraction rationnelle (ce qui sera notre cas), alors elle pourra se décomposer en éléments

simples et permettra, grâce à la transformation de Laplace inverse, de pouvoir relever simplement en temps le système à l'aide d'Equations Différentielles Ordinaires.

L'immersion transforme les matrices  $\mathbf{A}^i \partial_i$  en matrices  $\widetilde{\mathbf{A}}^i$ . Dans l'exemple précédent, on a vu que  $\widetilde{\mathbf{A}}^i = s_i \mathbf{A}^i$ . Comme les  $s_i \notin \mathbb{R}$ , les matrices  $\widetilde{\mathbf{A}}^i$  ne peuvent être hermitiennes puisqu'elles sont symétriques complexes.

Enfin, dernier intérêt, tout ce qui est écrit est vrai pour tout immersion. Il ne sera donc plus utile de se restreindre à des immersions de la forme  $\widetilde{\mathbf{x}}(\mathbf{x})$  et  $\widetilde{\mathbf{x}}(\mathbf{y})$ . Il n'y aura dès lors plus de limitation au cas cartésien.

Si l'on dispose d'une formulation intrinsèque des équations, la formulation dispersive existera toujours. En revanche, l'équation d'Euler avec une convection monotone [31] ne permet pas de l'obtenir.

## 2.2 Obtention d'un calcul différentiel intrinsèque sur la variété

On sait que les équations de Maxwell ou des ondes admettent une formulation dispersive pour des PML cartésiennes. Nous allons généraliser les systèmes PML en considérant que le domaine  $\Omega$  est la trace sur  $\mathbb{R}^3$ , éventuellement à un difféomorphisme de  $\mathbb{R}^3$  près, d'une variété complexe  $\mathbf{M}$  de  $\mathbb{C}^3$  et que le système d'équations s'écrit de façon intrinsèque sur cette variété.

Les changements de variables complexes précédemment introduits correspondent alors à des exemples d'équations de telles variétés décrites par un difféomorphisme  $\phi^{-1} : \mathbb{R}^3 \rightarrow \mathbf{M} \subset \mathbb{C}^3$  vérifiant  $\phi^{-1}(\mathbf{x}) = \mathbf{x}$  pour  $\mathbf{x} \in \Omega \subset \mathbf{M} \cap \mathbb{R}^3$ . Dans [20], *Lassas et al.* ont alors proposé de définir des variétés pseudo-Riemanniennes complexes absorbantes généralisant ces propriétés. On entend par variété pseudo-Riemannienne complexe une variété réelle de dimension 3, à fibrés tangents et cotangents complexes, munie d'une métrique  $g(.,.)$  complexe symétrique telle que son déterminant  $g = \sqrt{\det(g(.,.))}$  puisse toujours être défini avec la même détermination. L'objectif est de décrire géométriquement la complexification des PML.

### 2.2.1 Rappels de géométrie différentielle

On va travailler sur une variété  $\mathbf{M}$  réelle à fibré tangent complexe noté  $TM$  dotée d'une métrique riemannienne  $g_{jl}$ . Soit  $\mathbf{G} = (g_{ij})$  la matrice associée à la métrique  $g(.,.)$ . L'espace du fibré tangent associé  $T_{\mathbf{x}}M$ , pour tout  $\mathbf{x}$  appartenant à  $M$ , consiste en la somme de  $U + iV$  avec  $U$  et  $V$  des vecteurs réels tangent à  $\mathbf{x}$ . Le fibré cotangent, noté  $T_{\mathbf{x}}^*M$ , est le dual complexe de  $T_{\mathbf{x}}M$ . Les bases respectives sur ces espaces sont les dérivées partielles  $\frac{\partial}{\partial \mathbf{x}_1}, \dots, \frac{\partial}{\partial \mathbf{x}_n}$  pour  $T_{\mathbf{x}}M$  et les différentielles extérieures notées  $d\mathbf{x}_1, \dots, d\mathbf{x}_n$  pour  $T_{\mathbf{x}}^*M$ . Un élément de  $T_{\mathbf{x}}M$  s'écrira donc comme une combinaison

linéaire complexe sur cette base, soit  $\sum_j X_j(\mathbf{x}) \frac{\partial}{\partial \mathbf{x}_j}$ , et on nommera ces éléments vecteur et vecteur tangent, alors que l'on parlera de covecteur et de vecteur cotangent pour des éléments de  $T_x^*M$ , de la forme  $\sum_j X_j(\mathbf{x}) d\mathbf{x}_j$ , avec  $X_j(\mathbf{x}) \in \mathbb{C}$ .

Une connexion linéaire :  $TM \times TM \rightarrow TM : (\mathbf{u}, \mathbf{v}) \rightarrow \nabla_{\mathbf{u}}\mathbf{v}$  est une application bilinéaire vérifiant, avec  $f$  une fonction quelconque, les axiomes

$$\begin{cases} \nabla_{\mathbf{u}}(\mathbf{v} + \mathbf{w}) &= \nabla_{\mathbf{u}}\mathbf{v} + \nabla_{\mathbf{u}}\mathbf{w} \\ \nabla_{f\mathbf{u}}\mathbf{v} &= f\nabla_{\mathbf{u}}\mathbf{v} \\ \nabla_{\mathbf{u}}(f\mathbf{v}) &= \mathbf{u}(f)\mathbf{v} + f\nabla_{\mathbf{u}}\mathbf{v} \end{cases} \quad (2.9) \quad (\text{r\`egle de Leibnitz}).$$

Si  $g(\mathbf{u}, \mathbf{v})$  est la forme bilinéaire sur  $TM \times TM$  équivalente à l'isomorphisme sur  $T^*M$ , la connexion de Levi-Civita est la seule connexion linéaire qui conserve la métrique  $g$ , c'est-à-dire telle que

$$\mathbf{u}(g(\mathbf{v}, \mathbf{w})) = g(\nabla_{\mathbf{u}}\mathbf{v}, \mathbf{w}) + g(\mathbf{v}, \nabla_{\mathbf{u}}\mathbf{w}).$$

Elle est donnée par

$$2g(\nabla_{\mathbf{u}}\mathbf{v}, \mathbf{w}) = \mathbf{u}g(\mathbf{v}, \mathbf{w}) + \mathbf{v}g(\mathbf{w}, \mathbf{u}) - \mathbf{w}g(\mathbf{u}, \mathbf{v}) + g([\mathbf{u}, \mathbf{v}], \mathbf{w}) - g([\mathbf{u}, \mathbf{w}], \mathbf{v}) - g([\mathbf{v}, \mathbf{w}], \mathbf{u}),$$

où  $[\mathbf{u}, \mathbf{v}]$  désigne le crochet de Lie de  $\mathbf{u}$  et  $\mathbf{v}$ . Remarquons que si la composition  $\mathbf{u}\mathbf{ov}$  n'est pas une dérivation a priori,  $\mathbf{u}\mathbf{ov} - \mathbf{v}\mathbf{ou}$  en est bien une. Sur un système de coordonnées  $\partial_{\mathbf{x}_i}$  de  $T_xM$  on a pour tout tenseur  $t_{l_1 \dots l_q}^{j_1 \dots j_p}$  avec  $(p, q)$  quelconques, en posant  $\nabla_i = \nabla_{\partial_{\mathbf{x}_i}}$  la dérivée covariante selon  $\mathbf{x}_i$  est

$$\nabla_i t_{l_1 \dots l_q}^{j_1 \dots j_p} = \partial_i t_{l_1 \dots l_q}^{j_1 \dots j_p} + \sum_{\nu=1}^p \Gamma_{h\nu, i}^{j\nu} t_{l_1 \dots l_q}^{j_1 \dots h\nu \dots j_p} - \sum_{\nu=1}^q \Gamma_{l\nu, i}^{h\nu} t_{l_1 \dots h\nu \dots l_q}^{j_1 \dots j_p},$$

où  $\Gamma_{hi}^j$  sont les symboles de Christoffel :  $\Gamma_{hi}^j = \frac{1}{2}g^{jk}(\partial_i g_{kh} + \partial_h g_{ik} - \partial_h g_{ki})$ ,  $g^{jk}$  désignant les composantes de l'inverse de la métrique  $g(\cdot, \cdot)$ . On aura donc sur une base donnée :  $\nabla_i(\partial_{\mathbf{x}_h}) = \Gamma_{ih}^j \partial_{\mathbf{x}_j}$ . La connexion  $\nabla$  est alors sans torsion si

1.  $\nabla_{\mathbf{u}}\mathbf{v} - \nabla_{\mathbf{v}}\mathbf{u} - [\mathbf{u}, \mathbf{v}] = 0$  (d'où une expression simple des crochets de Lie)
2.  $\nabla_i g = 0$ .

La courbure complexe s'exprime par,

$$R(\mathbf{u}, \mathbf{v})\mathbf{w} = \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\mathbf{w} - \nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\mathbf{w} - \nabla_{[\mathbf{u}, \mathbf{v}]} \mathbf{w}.$$

Sur une base de  $TM$ , comme  $[\partial_{\mathbf{x}_i}, \partial_{\mathbf{x}_j}] = 0$ , la nullité de la courbure exprime la commutation des dérivées covariantes. L'expression sur une base  $\partial_{\sigma} \otimes d\mathbf{x}_{\mu} \otimes d\mathbf{x}_{\nu} \otimes d\mathbf{x}_{\rho}$  du tenseur (1, 3) ainsi défini est

$$R_{\mu\nu\rho}^{\sigma} = \frac{\partial \Gamma_{\mu\nu}^{\sigma}}{\partial \mathbf{x}_{\rho}} - \frac{\partial \Gamma_{\mu\rho}^{\sigma}}{\partial \mathbf{x}_{\nu}} + \Gamma_{\mu\nu}^{\lambda} \Gamma_{\rho\lambda}^{\sigma} - \Gamma_{\mu\rho}^{\lambda} \Gamma_{\nu\lambda}^{\sigma}.$$

On voit ainsi que si  $g(\cdot, \cdot)$ , ses dérivées partielles et son inverse dépendent holomorphiquement d'un paramètre  $s \in \mathbb{C}$ , les composantes de  $R$  sont à leur tour holomorphes en  $s$ .

**Définition 2.2** (Transformée de Hodge)

Pour deux  $p$ -formes  $\eta$  et  $\tau$ , on note le produit scalaire  $\eta \lrcorner \tau = \eta_J \tau^J$ , où  $J$  représente l'ensemble des  $p$  indices ordonnés, et on définit la transformée de Hodge  $\star$  par la relation :

$$\eta \wedge \star \tau = \tau \wedge \star \eta = (\eta \lrcorner \tau) \sqrt{|g|} dx_1 \wedge \cdots \wedge dx_n.$$

De plus, si  $n$  est la dimension de la variété,

$$\star^{-1} \tau = (-1)^{p(n-p)} \tau.$$

Notons  $g_c(\cdot, \cdot)$  la métrique euclidienne canonique complexe dans  $\mathbb{C}^3$  qui sera considérée comme la sous variété de dimension 3 de  $\mathbb{R}^6$  à fibrés tangent et cotangent complexes. Elle sera exprimée en coordonnées cartésiennes, soit  $\forall (\mathbf{u}, \mathbf{v}) \in \mathbb{C}^3$ ,

$$g_c(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^3 \mathbf{u}_j \mathbf{v}_j.$$

A partir du changement de variable :  $\mathbf{x} \mapsto \tilde{\mathbf{x}}(\mathbf{x})$  nous définissons alors une métrique complexe sur  $\mathbb{R}^3$ , notée  $g_{\mathbf{x}} : T_{\mathbf{x}} M\mathbb{R}^3 \times T_{\mathbf{x}} M\mathbb{R}^3 \rightarrow \mathbb{C}^3$ , par :

$$g_{\mathbf{x}}(\mathbf{u}, \mathbf{v}) = g_c(d\tilde{\mathbf{x}}(\mathbf{u}), d\tilde{\mathbf{x}}(\mathbf{v})),$$

où  $d\tilde{\mathbf{x}}$  est la différentielle de  $\tilde{\mathbf{x}}$ .

Dans le repère naturel, il y a toujours commutation des vecteurs tangents. Les dérivées covariantes ne commutent que si la variété est plate (tenseur de Riemann nul), ce qui permettra d'écrire le déterminant et le symbole de l'opérateur de façon intrinsèque. Une variété plate définit ainsi une structure de parallélisme sur la variété qui permettra de généraliser la notion de convolution.

**Définition 2.3** (Variétés Pseudo-Riemanniennes Absorbantes de [20])

Une variété Pseudo-Riemannienne  $(\mathbf{M}, g)$  est dite absorbante si :

1. la variété  $(\mathbf{M}, g)$  est plate et  $\mathbf{M}$  est difféomorphe à  $\mathbb{R}^3$  (on note  $\phi$  ce difféomorphisme),
2. il existe  $\Omega \subset \mathbf{M}$  un ouvert relativement compact où la métrique est euclidienne réelle, et est donnée par  $g = \phi^* g^e$ , où  $\phi^*$  est le pull-back selon  $\phi$  et  $g^e$  la métrique euclidienne,
3.  $\forall \mathbf{v} \in T_{\mathbf{x}}^{\mathbb{R}}, \mathbf{v} \neq 0, \quad g_{\mathbf{x}}(\mathbf{v}, \mathbf{v}) \neq 0$ ,
4. l'immersion  $\tilde{\mathbf{x}} : \mathbf{M} \rightarrow \mathbb{C}^3$  vérifie les deux propriétés suivantes :

$$(\tilde{\mathbf{x}}(\mathbf{x}_1) - \tilde{\mathbf{x}}(\mathbf{x}_2))^2 = 0 \text{ si et seulement si } \mathbf{x}_1 = \mathbf{x}_2, \quad (2.10)$$

$$\exists (c_0, c_1) > 0, \quad \mathbf{Re} \left( \frac{\tilde{\mathbf{x}}(\mathbf{x})^2}{|\phi(\mathbf{x})|^2} \right) \leq -2c_0 \text{ pour } |\phi(\mathbf{x})| > c_1. \quad (2.11)$$

Dans [20], l'hypothèse 4 lui permet de garantir l'existence et l'unicité de la solution pour toute fréquence en bornant le domaine « assez loin ». Or nous cherchons à obtenir rapidement de la décroissance pour réduire le domaine de calcul le plus près possible du domaine d'étude. Aussi, nous introduirons une définition (2.4).

**Définition 2.4**

$(\mathbf{M}, g)$  est une variété Pseudo-Riemannienne « absorbante », c'est à dire que

1. la variété  $(\mathbf{M}, g)$  est plate et  $\mathbf{M}$  est difféomorphe à  $\mathbb{R}^3$  (on note  $\phi$  ce difféomorphisme),
2. il existe  $\Omega \subset \mathbf{M}$  un ouvert relativement compact où la métrique est euclidienne réelle, et est donnée par  $g = \phi^* g^e$ , où  $\phi^*$  est le pull-back selon  $\phi$  et  $g^e$  la métrique euclidienne,
3.  $\forall \mathbf{v} \in T_{\mathbf{x}}^{\mathbb{R}}, \mathbf{v} \neq 0, \quad g_{\mathbf{x}}(\mathbf{v}, \mathbf{v}) \neq 0,$
4. l'immersion  $\tilde{\mathbf{x}} : \mathbf{M} \rightarrow \mathbb{C}^3$  garantie que la partie imaginaire de  $\tilde{\mathbf{x}}(\mathbf{x}_1) - \tilde{\mathbf{x}}(\mathbf{x}_2)$  soit signée et de signe constant.

La définition (2.3) indique alors que les tenseurs  $\bigotimes_{k=1}^p d\tilde{\mathbf{x}}_{i_k}$   $\bigotimes_{r=1}^q d\tilde{\mathbf{x}}_{i_r}$  sont parallèles et leurs dérivées covariantes sont nulles. En particulier, la platitude de la variété va permettre de trouver un noyau de Schwartz à l'opérateur de Maxwell analogue à la convolution par le noyau de l'équation des ondes et la matrice des codérivations permettant de trouver le noyau des équations de Maxwell sur  $\mathbb{R}^3$ . L'hypothèse sur la platitude de la variété assure ici la commutativité des dérivations et des compositions par convolution.

**2.2.2 Formalisme des PML**

Soit une variété réelle à valeurs dans  $\mathbb{C}^3$  décrite par le changement de variable suivant :

$$\tilde{\mathbf{x}} = \varphi(\mathbf{x}) + \frac{1}{\mathbf{p}} f(\varphi(\mathbf{x})). \quad (2.12)$$

Rappelons les notations introduites dans le chapitre 1 (1.1) :  $\mathbf{O}$  un obstacle borné, régulier non nécessairement connexe pouvant être vide,  $\Omega$  un domaine connexe, non nécessairement simplement connexe représentant le domaine spatial,  $\delta\Omega_0$  la frontière entre  $\Omega$  et  $\mathbf{O}$ ,  $\delta\Omega_{\infty}$  sa frontière extérieure. Elles sont toutes deux  $C^1$ , et  $\delta\Omega_{\infty}$  peut éventuellement être rejetée à l'infini dans le cas d'un domaine  $\Omega$  non borné.

**Hypothèses 2.5**

Le couple  $(\varphi, f)$  appartient à  $C^1(\mathbb{R}^3, \mathbb{R}^3)$ .  $\forall x \in \Omega$ , on a  $\varphi(\mathbf{x}) = \mathbf{x}$  et  $f(\mathbf{x}) = 0$ . De plus,  $|\varphi(\mathbf{x})| = \mathcal{O}_{|\mathbf{x}| \rightarrow \infty}(|\mathbf{x}|)$  et  $f$  est strictement croissante avec une croissance linéaire à l'infini. Enfin, on suppose qu'il existe  $S$  une fonction convexe telle que  $f = \mathbf{grad} S$ .

**Remarque :** Nous justifierons plus loin ces hypothèses. On pourrait se contenter de la régularité  $C_{mex}^1(\mathbb{R}^3, \mathbb{R}^3) \cap C^0(\mathbb{R}^3, \mathbb{R}^3)$  pour les couples  $(\varphi, f)$  si les raccordements sont lipschitziens. En effet, le fait que les fonctions soit continues par morceaux implique la présence d'arêtes, et donc la variété n'est pas différentiable. Dans le cas où ces arêtes sont lipschitziennes, les raccordements n'impliquent pas de sauts et il sera possible d'écrire les différentielles extérieures faibles grâce au théorème du partitionnement de l'unité.

L'intérêt d'une telle formulation réside dans le fait que la complexification de chaque composante se fait alors à partir d'une fonction donnée sur tout  $\mathbb{R}^3$  et ne dépendant plus de la simple coordonnée choisie. En outre, ce changement de variable ne préjuge pas a priori de la forme donnée aux PML ni de la direction d'absorption de celles-ci. Le choix de ces fonctions influe sur la régularité de la variété.

Comme  $f|_{\Omega} = 0_{\Omega}$  et  $\varphi|_{\Omega} = \mathbf{id}_{\Omega}$ , le système d'équations est inchangé à l'intérieur du domaine d'étude. Si ces fonctions sont de classe  $C^k$ , alors la variété  $\mathbf{M}$  définie par le changement de variables (2.12) est bien une sous-variété de classe  $C^k$  de dimension 3 dans le  $\mathbb{R}$ -espace vectoriel  $\mathbb{C}^3$  de dimension 6.

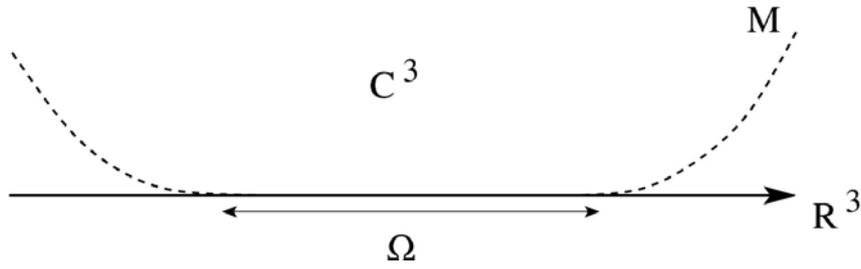


FIG. 2.2: Représentation du plongement

### 2.2.3 Problème de Helmholtz sur $\mathbf{M}$

On s'intéresse d'abord à la résolution de l'équation de Helmholtz sur la variété  $\mathbf{M}$  qui satisfait aux axiomes de la définition (2.4). L'opérateur de Hodge  $\star$  associé correspondant à la métrique  $g$  d'une variété absorbante  $(\mathbf{M}, g)$  est défini par l'identité :

$$\begin{aligned} U \wedge \star V &= g(U, V) d\tilde{\mathbf{x}}_1 \wedge d\tilde{\mathbf{x}}_2 \wedge d\tilde{\mathbf{x}}_3 \\ \mathbf{dvol}_g &= g dx_1 \wedge dx_2 \wedge dx_3 = d\tilde{\mathbf{x}}_1 \wedge d\tilde{\mathbf{x}}_2 \wedge d\tilde{\mathbf{x}}_3. \end{aligned}$$

Il est alors possible de définir le Laplacien généralisé  $\Delta_g^r$  pour les  $r$ -formes sur  $\mathbf{M}$  par la généralisation à la métrique  $g$  de la relation :

$$\Delta_g^r = (-1)^r (\star \mathbf{d} \star \mathbf{d} - \mathbf{d} \star \mathbf{d} \star).$$

Si la variété  $\mathbf{M}$  est plate,  $[\nabla_i, \nabla_j] = 0$ , et les composantes d'une 1-forme quelconque sont exprimées à l'aide de la métrique  $g$ . Alors  $\Delta_g^0$  correspond à l'opposé de l'opérateur de Laplace-Beltrami  $\Delta : \varphi \mapsto g^{ij} \nabla_i \nabla_j \varphi$ .

On s'intéresse maintenant à la résolution de l'équation de Helmholtz. Soit  $\mathbf{x}, \mathbf{y}$  appartenant à la variété  $\mathbf{M}$ , notons  $\tilde{\mathbf{x}} = \tilde{\mathbf{x}}(\mathbf{x})$ ,  $\tilde{\mathbf{y}} = \tilde{\mathbf{x}}(\mathbf{y})$  et :

$$\{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\} = \left( \sum_{j=1}^3 (\tilde{\mathbf{x}}_j - \tilde{\mathbf{y}}_j)^2 \right)^{1/2}.$$

Remarquons que la condition (2.10) correspond à la détermination d'une racine carrée complexe. En effet, d'après la dernière condition de la définition (2.3), la partie imaginaire de  $\{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\}$  pour tout  $(\mathbf{x}, \mathbf{y}) \in \Omega$ . Le choix  $\{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\} \geq 0$  correspondant à une condition de radiation d'onde sortante, et  $\{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\} \leq 0$  à une onde entrante. La condition (2.10) est celle des ondes sortantes. Le signe de la racine carrée complexe  $\{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\}$  pour  $(\tilde{\mathbf{x}}(\mathbf{x}), \tilde{\mathbf{x}}(\mathbf{y}))$  peut alors être déterminé sur la variété simplement connectée  $\mathbf{M} \times \mathbf{M}$ . Dans ce cas, la variété  $\mathbf{M}$  est appelée **variété absorbante sortante**.

Soit un courant  $\delta_{\mathbf{y}}$  et  $\mathbf{p} = \varepsilon + i\omega \in \mathbb{C}$  la variable de Laplace, le problème de Helmholtz sur la variété  $\mathbf{M}$  est

$$\forall \mathbf{y} \in \mathbf{M}, \quad (\Delta_g - \mathbf{p}^2)\Phi(., \mathbf{y}) = -\delta_{\mathbf{y}}. \quad (2.13)$$

Le delta Dirac  $\delta_{\mathbf{y}}$  ci dessus sera interprété par rapport à la métrique  $g(., .)$  : si  $\psi$  est une 0-forme  $C^\infty$  sur la variété  $\mathbf{M}$ , alors  $\int_{\mathbf{M}} \psi(\mathbf{x})\delta_{\mathbf{y}}(\mathbf{x})d\text{vol}_g(\mathbf{x}) = \psi(\mathbf{y})$ . La recherche du noyau de Green de cette équation donnera ainsi la solution élémentaire de (2.13) pour des 1-formes, composantes par composantes.

### ***Théorème 2.6***

*Soit  $(\mathbf{M}, g)$  une variété absorbante sortante, alors la fonction :*

$$G_H(\mathbf{x}, \mathbf{y}) = \frac{\exp(-i\omega \{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\})}{4\pi \{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\}} \quad (2.14)$$

*est la solution fondamentale de l'opérateur de Helmholtz décrit sur  $\mathbf{M}$  et satisfait l'équation (2.13).*

*De plus,  $G_H$  vérifie la relation asymptotique :*

$$|G_H(\mathbf{x}, \mathbf{y})| \leq C_{\mathbf{y}} e^{-c_0\omega|\phi(\mathbf{x})|} \text{ pour } |\phi(\mathbf{x})| > c_2, \quad (2.15)$$

*où  $c_0 > 0$ ,  $c_2 = c_2(\mathbf{y}) > 0$  et  $\phi : \mathbf{M} \rightarrow \mathbb{R}^3$  est le difféomorphisme donné par la définition (2.3).*

Le théorème (2.6) décrit la solution élémentaire (2.14) de la partie principale des équations de Maxwell écrites sur la variété complexe  $\mathbf{M}$ , soit de l'équation de Helmholtz (2.13) avec le Laplacien  $\Delta_g$  sur la variété. La relation (2.15) correspond à la condition de décroissance à l'infini sur la variété  $\mathbf{M}$ .

Soit  $n = 2$  ou  $3$  la dimension d'espace. Etudions l'« injectivité » du changement de variables. On parle ici d'injectivité au sens de la condition du plongement de la variété dans  $\mathbb{C}^n$ . Le plongement est une immersion injective isométrique par rapport à la métrique pseudo euclidienne de  $\mathbb{C}^n$ . Le noyau de Green des Equations de Maxwell dépend fortement du terme  $\{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\}$  : il ne doit pas s'annuler au dénominateur et doit avoir une partie imaginaire signée positive au numérateur afin de garantir la décroissance exponentielle souhaitée. Notons  $[\tilde{\mathbf{x}}]^2 = \{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\}$ ,  $[f] = [f(\varphi(\mathbf{x})) - f(\varphi(\mathbf{y}))]$  et

$[\varphi] = [\varphi(\mathbf{x}) - \varphi(\mathbf{y})]$ . En considérant le cas limite  $\mathbf{p} = i\omega$ , on obtient que

$$[\tilde{\mathbf{x}}]^2 = [\varphi]^2 - \frac{1}{\omega^2} [f]^2 - \frac{2i}{\omega} [\varphi] \cdot [f]. \quad (2.16)$$

Il est nécessaire de calculer les parties réelle et imaginaire de  $[\tilde{\mathbf{x}}]^2$  : il faut vérifier qu'elles ne s'annulent pas simultanément et que la partie imaginaire est de signe constant positif. Une condition pour garantir la non nullité simultanée des parties réelle et imaginaire se déduit de la relation précédente :  $[f] = 0$ . Dans le cas contraire, il est toujours possible d'annuler la partie réelle.

$$[f] = \int_0^1 (1-t) \left( \frac{Df}{D\varphi} \right) (\varphi_1 + t(\varphi_2 - \varphi_1)) dt.$$

Notons  $\mathbf{M}(\varphi_1, \varphi_2)$  la matrice précédente. Une condition pour annuler  $[f]$  est donc que  $(\mathbf{M}[\varphi], [\varphi]) = 0$ , ce qui implique que  $[\varphi]$  est un vecteur propre de la matrice  $\mathbf{M} + \mathbf{M}^T$  associé à une valeur propre nulle. En effet, les valeurs propres strictement positives et négatives de  $\mathbf{M} + \mathbf{M}^T$  ne garantissent pas l'injectivité. Une condition suffisante est donc que la matrice  $\mathbf{M}$  soit positive ou nulle. Si la matrice  $\mathbf{M}$  est strictement positive, alors la partie imaginaire de  $[\tilde{\mathbf{x}}]^2$  ne s'annule jamais. Ce cas est impossible : le contre exemple le plus trivial est celui des formulations PML classiques. Dans le cas des PML cartésiennes,  $\varphi = \mathbf{id}$  et  $\left( \frac{Df}{D\varphi} \right) = (\delta_{ij} \sigma_i)_{i,j \in \llbracket 1, n \rrbracket}$ .

Pour avoir une détermination continue de la racine, il faut que la partie imaginaire soit de signe constant. Une condition suffisante est donc que la matrice  $\mathbf{M}$  soit symétrique. Si tel est le cas, alors la matrice  $\left( \frac{Df}{D\varphi} \right)$  est symétrique, et donc il existe une fonction  $S$  convexe telle que  $f = \underset{\varphi}{\mathbf{grad}} S$ . Ceci justifie les hypothèses (2.5) faites précédemment.

### Propriété 2.7

*Les décroissances locale et asymptotique du noyau de Green sont indépendantes de la fréquence  $\omega$ .*

**Démonstration** : nous allons déterminer explicitement le terme  $\{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\}$  la racine carrée. Soit  $\alpha$  la partie réelle et  $\beta$  la partie imaginaire. En reprenant l'équation (2.16), on obtient par égalité des parties imaginaire et réelle le système

$$\begin{cases} \alpha^2 - \beta^2 = [\varphi]^2 - \frac{1}{\omega^2} [f]^2 \\ 2\alpha\beta = -\frac{2}{\omega} [\varphi] \cdot [f]. \end{cases}$$

En combinant ces deux équations, on est ramené à la recherche des racines du polynôme

$$\beta^4 + \left( [\varphi]^2 - \frac{1}{\omega^2} [f]^2 \right)^2 \beta^2 - \left( \frac{[\varphi] \cdot [f]}{\omega} \right)^2 = 0.$$

D'où

$$2\beta^2 = \frac{1}{\omega^2} [f]^2 - [\varphi]^2 + \sqrt{\left([\varphi]^2 - \frac{1}{\omega^2} [f]^2\right)^2 + 4 \left(\frac{[\varphi] \cdot [f]}{\omega}\right)^2}.$$

Si  $[\varphi] \cdot [f] = 0$  alors nécessairement  $[f] = 0$  (condition d'injectivité précédente). A ce moment là,  $2\beta^2 = -[\varphi]^2 + \sqrt{([\varphi]^2)^2} = 0$ , il n'y a donc pas de décroissance du noyau de Green. Cela définit les lignes d'exhaustion du domaine PML. La décroissance du milieu PML se fait dans la direction orthogonale à ces dernières. Par exemple, dans le cas des PML cartésiennes, dans la zone où le coefficient d'absorption en  $x$  est nul, les directions sans décroissance sont horizontales, et la décroissance dans le milieu se fait verticalement (cf figure (2.1)).

Si  $\frac{[f]}{\omega} < [\varphi]$ . On obtient alors localement une estimation de la partie imaginaire  $\beta^2 \leq \frac{1}{\omega^2} [f]^2$  d'où  $|G_H| \leq \frac{\exp(-\omega \sqrt{[\varphi] \cdot [f]})}{4\pi | \{[\varphi]\} |}$ . Mais dans ce cas,  $\omega \geq [f] / [\varphi]$  donc

$$|G_H| \leq \frac{\exp(-\sqrt{[f] / [\varphi]})}{4\pi | \{[\varphi]\} |}.$$

Dans le cas contraire, si  $[\varphi] < \frac{[f]}{\omega}$  la matrice  $\mathbf{M}$  est définie positive. On peut développer le calcul par rapport aux valeurs propres de cette matrice. En se plaçant sur la bonne base, il est possible d'orthogonaliser, c'est à dire écrire  $[f]$  comme  $[\varphi]$  plus quelque chose d'orthogonal au saut de  $\varphi$ . Mais alors  $\beta = \frac{[f]}{\omega}$ . Or le jacobien de  $f$  devient constant à l'infini, donc on peut majorer par  $[\varphi]$ . Soit  $k$  la plus petite valeur propre de la matrice  $\mathbf{M}$ , on a  $|G_H| \leq \frac{\exp(-k [\varphi])}{4\pi | \{[\varphi]\} |}$ . De plus, comme  $\varphi$  est équivalent à  $\mathbf{x}$  en l'infini, on obtient que

$$|G_H| \leq \frac{\exp(-k [\mathbf{x}])}{4\pi | \{[\varphi]\} |}.$$

Dans tous les cas, la décroissance locale du noyau de Green est assurée indépendamment de la fréquence. Asymptotiquement, comme  $\exists c > 0$ , telle que pour toute fréquence  $\omega$  on ait  $\mathbf{Re}(-i\omega \{[\tilde{\mathbf{x}}]\}) \leq -c |\mathbf{x}|$  quand  $|\mathbf{x}|$  tend vers l'infini, on a donc la décroissance exponentielle de chaque composante de  $G_H$ .  $\boxtimes$

**Remarque** : l'obtention de cette propriété est liée à la forme du changement de variable choisi  $\varphi(\mathbf{x}) + \frac{1}{\mathbf{p}} f(\varphi(\mathbf{x}))$ . Remplacer  $\frac{1}{\mathbf{p}}$  par une fonction homographe ne permettrait pas de conclure.

Pour la suite du mémoire, nous considérerons donc les couples admissibles  $(\varphi, f)$  appartenant à  $C^1(\mathbb{R}^3, \mathbb{R}^3)$  tels que  $\forall \mathbf{x} \in \Omega$ ,

$$\varphi(\mathbf{x}) = \mathbf{x} \quad \text{et} \quad f(\mathbf{x}) = 0.$$

Le  $\mathbb{R}^3$  difféomorphisme  $\varphi$  est tel que  $|\varphi(\mathbf{x})| = \mathcal{O}(|\mathbf{x}|)$ . La fonction  $f$  doit être strictement croissante, sa croissance doit être linéaire à l'infini et il existe  $S$  une fonction convexe telle que  $f$  soit le gradient de  $S$  par rapport à  $\varphi$ . Sous ces hypothèses, le plongement effectué est isométrique, permet de garantir l'injectivité du changement de variable complexe et la décroissance, tant locale qu'asymptotique du noyau de Green des Equations, ce qui garantit les propriétés du milieu PML.

## 2.3 Résolution des Equations de Maxwell

La résolution des équations de Maxwell sur la variété  $\mathbf{M}$  est étudiée afin de garantir dans un premier temps l'existence et l'unicité de la solution sur tout  $\mathbf{M}$ . Dans un second temps, le problème posé sur la variété  $\mathbf{M}$  en domaine non-borné sera restreint à un ensemble borné décrivant le problème PML en domaine borné avec condition limite.

Si les hypothèses du théorème de Rauch sont vérifiées [32], ce dernier garantit que toute solution forte est solution faible, et que les solutions fortes sont définies régulières  $C^\infty$  avec des seconds membres de même régularité, si le système admet une solution forte régulière, alors il admet une solution faible. Ainsi, les champs électrique et magnétique  $e$  et  $h$  seront uniquement considérés comme des 1-formes régulières. Etant donné une métrique  $g(.,.)$ , il y a une correspondance univoque bien connue entre les champs de vecteurs et les 1-formes. Soit  $j$  et  $m$  des 2-formes représentant les seconds membres à coefficients  $C^\infty$  à support compact. Les équations de Maxwell harmoniques dans le vide peuvent être écrites sous la forme

$$\begin{cases} \mathbf{p} \star e - \mathbf{d} h = -j \\ \mathbf{p} \star h + \mathbf{d} e = -m, \end{cases} \quad (2.17)$$

avec  $\mathbf{p} = \varepsilon + ik \in \mathbb{C}$ ,  $k = \omega/c = \omega\sqrt{\varepsilon_0\mu_0}$ ,  $\star$  l'opérateur de Hodge qui transforme des 1-formes en 2-formes défini avec la métrique Euclidienne. Les champs sont adimensionnés de la façon suivante :  $E \rightarrow \sqrt{\varepsilon_0}E$  et  $H \rightarrow \sqrt{\mu_0}H$  pour obtenir des notations symétriques. L'opérateur de Maxwell  $\mathcal{M}$  est défini par

$$\mathcal{M} = \begin{pmatrix} 0 & \mathbf{d} \\ -\mathbf{d} & 0 \end{pmatrix} = -\mathcal{M}^T,$$

et  $\star(e, h)$  est la transformation de Hodge du doublet  $(e, h)$  à savoir  $(\star e, \star h)$ .

$$(\mathbf{p} \star + \mathcal{M}) \begin{pmatrix} e \\ h \end{pmatrix} = \begin{pmatrix} -j \\ -m \end{pmatrix}. \quad (2.18)$$

### **Lemme 2.8**

Soit  $\varphi$  une 0-forme et  $A$  une 1-forme avec  $\nabla A = 0$ . Alors  $\Delta^1(\varphi A) = (\Delta^0 \varphi) A$ .

**Démonstration** : rappelons tout d'abord les transformées de Hodge des vecteurs de bases du fibré tangent à la variété  $\mathbf{M}$  :

$$\begin{cases} \star d\tilde{x}_1 = d\tilde{x}_2 \wedge d\tilde{x}_3 \\ \star d\tilde{x}_2 = d\tilde{x}_3 \wedge d\tilde{x}_1 \\ \star d\tilde{x}_3 = d\tilde{x}_1 \wedge d\tilde{x}_2 \end{cases}$$

Soit  $A$  une 1-forme alors  $A$  est combinaison linéaire des  $d\tilde{x}_i$ . Supposons pour les calculs (sans perdre en généralité) que  $A = d\tilde{x}_1$ .

$$\Delta^1(\varphi A) = -(\star \mathbf{d} \star \mathbf{d} - \mathbf{d} \star \mathbf{d} \star)(\varphi d\tilde{x}_1).$$

Décomposons l'opérateur. Soit  $\nabla_i := \nabla_{\partial \tilde{x}_i}$ .

$$\begin{aligned} \mathbf{d}(\varphi d\tilde{x}_1) &= \nabla_k \varphi d\tilde{x}_k \wedge d\tilde{x}_1 = \nabla_2 \varphi d\tilde{x}_2 \wedge d\tilde{x}_1 + \nabla_3 \varphi d\tilde{x}_3 \wedge d\tilde{x}_1 \\ \star \mathbf{d}(\varphi d\tilde{x}_1) &= -\nabla_2 \varphi d\tilde{x}_3 + \nabla_3 \varphi d\tilde{x}_2 \\ \mathbf{d} \star \mathbf{d}(\varphi d\tilde{x}_1) &= -\nabla_{12} \varphi d\tilde{x}_1 \wedge d\tilde{x}_3 - \nabla_{22} \varphi d\tilde{x}_2 \wedge d\tilde{x}_3 \\ &\quad + \nabla_{13} \varphi d\tilde{x}_1 \wedge d\tilde{x}_2 + \nabla_{33} \varphi d\tilde{x}_3 \wedge d\tilde{x}_2 \\ \star \mathbf{d} \star \mathbf{d}(\varphi d\tilde{x}_1) &= \nabla_{12} \varphi d\tilde{x}_2 - \nabla_{22} \varphi d\tilde{x}_1 + \nabla_{13} \varphi d\tilde{x}_3 - \nabla_{33} \varphi d\tilde{x}_1 \\ \star(\varphi d\tilde{x}_1) &= \varphi d\tilde{x}_2 \wedge d\tilde{x}_3 \\ \mathbf{d} \star(\varphi d\tilde{x}_1) &= \nabla_1 \varphi d\tilde{x}_1 \wedge d\tilde{x}_2 \wedge d\tilde{x}_3 \\ \star \mathbf{d} \star(\varphi d\tilde{x}_1) &= \nabla_1 \varphi \\ \mathbf{d} \star \mathbf{d} \star(\varphi d\tilde{x}_1) &= \nabla_{11} \varphi d\tilde{x}_1 + \nabla_{12} \varphi d\tilde{x}_2 + \nabla_{13} \varphi d\tilde{x}_3. \end{aligned}$$

Donc

$$\begin{aligned} (\star \mathbf{d} \star \mathbf{d} - \mathbf{d} \star \mathbf{d} \star)(\varphi d\tilde{x}_1) &= (-\partial_{11} \varphi - \partial_{22} \varphi - \partial_{33} \varphi) d\tilde{x}_1 \\ &\quad + (\partial_{12} \varphi - \partial_{12} \varphi) d\tilde{x}_2 + (\partial_{13} \varphi - \partial_{13} \varphi) d\tilde{x}_3 \\ &= -(\Delta^0 \varphi) d\tilde{x}_1 \end{aligned}$$

On obtient des résultats similaires avec  $A = d\tilde{x}_2$  et  $A = d\tilde{x}_3$ . Le lemme s'obtient alors en passant à la combinaison linéaire des  $d\tilde{x}_i$ .  $\square$

Soit deux 2-formes  $A$  et  $B$  telles que  $\nabla A = \nabla B = 0$ . Définissons l'application  $G_{A,B}(\mathbf{x}, \mathbf{y})$  à partir du noyau de Green  $G_H(\mathbf{x}, \mathbf{y})$  de l'équation de Helmholtz sur  $\mathbf{M}$

$$G_{A,B}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} \mathbf{p} - \mathbf{p}^{-1} \mathbf{d} \star \mathbf{d} \star & -\star \mathbf{d} \\ \star \mathbf{d} & \mathbf{p} - \mathbf{p}^{-1} \mathbf{d} \star \mathbf{d} \star \end{pmatrix} \begin{pmatrix} G_H(\mathbf{x}, \mathbf{y}) A \\ G_H(\mathbf{x}, \mathbf{y}) B \end{pmatrix}.$$

Comme sur des variétés euclidiennes, sur la variété  $\mathbf{M}$ , l'application  $G_{A,B}(x, y)$  vérifie alors les propriétés du noyau de Green pour les équations de Maxwell.

### **Théorème 2.9**

Soit  $(\mathbf{M}, g)$  une variété absorbante sortante,  $G_{A,B}$  vérifie :

$$\forall \mathbf{y} \in \mathbf{M}, \quad (\mathbf{p} + \star \mathcal{M}^T) G_{A,B}(\cdot, \mathbf{y}) = \begin{pmatrix} A \delta_{\mathbf{y}} \\ B \delta_{\mathbf{y}} \end{pmatrix}. \quad (2.19)$$

**Démonstration** : comme  $(\mathbf{p} + \star \mathcal{M}^T) = \begin{pmatrix} \mathbf{p} & \star \mathbf{d} \\ -\star \mathbf{d} & \mathbf{p} \end{pmatrix}$ , il vient

$$\begin{aligned} & (\mathbf{p} + \star \mathcal{M}^T) G_{A,B}(\cdot, \mathbf{y}) \\ &= \begin{pmatrix} \mathbf{p} & \star \mathbf{d} \\ -\star \mathbf{d} & \mathbf{p} \end{pmatrix} \begin{pmatrix} \mathbf{p} - \mathbf{p}^{-1} \mathbf{d} \star \mathbf{d} \star & -\star \mathbf{d} \\ \star \mathbf{d} & \mathbf{p} - \mathbf{p}^{-1} \mathbf{d} \star \mathbf{d} \star \end{pmatrix} \begin{pmatrix} G_H(\cdot, \mathbf{y}) A \\ G_H(\cdot, \mathbf{y}) B \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{p}^2 - \mathbf{d} \star \mathbf{d} \star + \star \mathbf{d} \star \mathbf{d} & 0 \\ 0 & \mathbf{p}^2 - \mathbf{d} \star \mathbf{d} \star + \star \mathbf{d} \star \mathbf{d} \end{pmatrix} \begin{pmatrix} G_H(\cdot, \mathbf{y}) A \\ G_H(\cdot, \mathbf{y}) B \end{pmatrix} \\ &= \begin{pmatrix} (\mathbf{p}^2 - \Delta_g) G_H(\cdot, \mathbf{y}) A \\ (\mathbf{p}^2 - \Delta_g) G_H(\cdot, \mathbf{y}) B \end{pmatrix} = \begin{pmatrix} A \delta_{\mathbf{y}} \\ B \delta_{\mathbf{y}} \end{pmatrix} \quad (\text{lemme (2.8)}) \end{aligned}$$

puisque  $G_H(\cdot, \mathbf{y})$  est le noyau de Green de l'équation de Helmholtz sur la variété  $\mathbf{M}$  et donc la solution élémentaire de (2.13).  $\square$

### 2.3.1 Cas des Equations de Maxwell avec un second membre non nul et $\mathbf{O} = \emptyset$

**Théorème 2.10**

Soit  $(\mathbf{M}, g)$  une variété absorbante sortante. Alors le problème (2.17) en l'absence d'objet diffractant admet une unique solution

$$\begin{cases} e_i = \int_{\mathbf{M}} G_{d\tilde{\mathbf{x}}_i, 0} \wedge \begin{pmatrix} -j \\ -m \end{pmatrix} \\ h_i = \int_{\mathbf{M}} G_{0, d\tilde{\mathbf{x}}_i} \wedge \begin{pmatrix} -j \\ -m \end{pmatrix}. \end{cases}$$

**Démonstration** : Soit  $X, Y$  des 1-formes. On note  $\star \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \star X \\ \star Y \end{pmatrix}$ . Soit  $P, Q$  des 1-formes, on rappelle que le produit extérieur  $\begin{pmatrix} X \\ Y \end{pmatrix} \wedge \begin{pmatrix} P \\ Q \end{pmatrix} = X \wedge P + Y \wedge Q$ . De plus,  $P^* \wedge Q = -P \wedge Q^*$ .

Considérons la formulation faible des équations de Maxwell. Soit  $A$  et  $B$  des 1-formes,

$$(\mathbf{p} \star + \mathcal{M}) \begin{pmatrix} e \\ h \end{pmatrix} \wedge \varphi(A, B) = \begin{pmatrix} e \\ h \end{pmatrix} \wedge \left( -\mathbf{p} \star \varphi(A, B) + \mathcal{M} \varphi(A, B) \right)$$

En appliquant la transformée de Hodge à l'équation (2.19), on obtient

$$(\mathbf{p} \star + \mathcal{M}^T) G_{A,B}(\cdot, \mathbf{y}) = \begin{pmatrix} \star A \delta_{\mathbf{y}} \\ \star B \delta_{\mathbf{y}} \end{pmatrix}.$$

D'où

$$\begin{aligned} e \wedge \star A + h \wedge \star B &= \left\{ \int_{\mathbf{M}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathbf{p} \star + \mathcal{M}^T) G_{A,B}(\cdot, \mathbf{y}) \right\} \mathbf{dvol}_g(\mathbf{y}) \\ &= \left\{ \int_{\mathbf{M}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathbf{p} \star G_{A,B}(\cdot, \mathbf{y})) \right\} \mathbf{dvol}_g(\mathbf{y}) \\ &\quad + \left\{ \int_{\mathbf{M}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathcal{M}^T G_{A,B}(\cdot, \mathbf{y})) \right\} \mathbf{dvol}_g(\mathbf{y}) \end{aligned}$$

Par commutation de la transformée de Hodge,  $\int_{\mathbf{M}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathbf{p} \star G_{A,B}(\cdot, \mathbf{y}))$  devient  $\int_{\mathbf{M}} G_{A,B}(\cdot, \mathbf{y}) \wedge \left( \mathbf{p} \star \begin{pmatrix} e \\ h \end{pmatrix} \right)$ .

Soit  $\omega$  et  $\beta$  des 1-formes. On rappelle que  $\mathbf{d}(\omega \wedge \beta) = \mathbf{d}\omega \wedge \beta - \omega \wedge \mathbf{d}\beta$ . Alors,

$$0 = \int_{\mathbf{M}} \mathbf{d}(\omega \wedge \beta) = \int_{\mathbf{M}} \mathbf{d}\omega \wedge \beta - \int_{\mathbf{M}} \omega \wedge \mathbf{d}\beta \text{ donc } \int_{\mathbf{M}} \mathbf{d}\omega \wedge \beta = \int_{\mathbf{M}} \omega \wedge \mathbf{d}\beta.$$

Grâce aux propriétés du noyau de Green,

$$\int_{\mathbf{M}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathcal{M}^T G_{A,B}(\cdot, \mathbf{y})) = \int_{\mathbf{M}} \mathcal{M}^T \begin{pmatrix} e \\ h \end{pmatrix} \wedge G_{A,B}(\cdot, \mathbf{y}) = \int_{\mathbf{M}} G_{A,B}(\cdot, \mathbf{y}) \wedge \mathcal{M} \begin{pmatrix} e \\ h \end{pmatrix}.$$

Donc

$$\begin{aligned} e \wedge \star A + h \wedge \star B &= \left\{ \int_{\mathbf{M}} G_{A,B}(\cdot, \mathbf{y}) \wedge (\mathbf{p} \star + \mathcal{M}) \begin{pmatrix} e \\ h \end{pmatrix} \right\} \mathbf{dvol}_g(\mathbf{y}) \\ &= \left\{ \int_{\mathbf{M}} G_{A,B}(\cdot, \mathbf{y}) \wedge \begin{pmatrix} -j \\ -m \end{pmatrix} \right\} \mathbf{dvol}_g(\mathbf{y}). \end{aligned} \quad \text{équation (2.18)}$$

Etant donné que  $e = \sum e_i d\tilde{\mathbf{x}}_i$  et  $h = \sum h_i d\tilde{\mathbf{x}}_i$  et que la formule précédente a été établie pour des 1-formes  $A$  et  $B$  quelconques, posons respectivement  $A = d\tilde{\mathbf{x}}_i, B = 0$  et  $A = 0, B = d\tilde{\mathbf{x}}_i$  pour trouver les composantes  $e_i$  et  $h_i$  des champs électrique et magnétique.  $\square$

### 2.3.2 Cas des Equations de Maxwell sur $\mathbf{M}$ avec $\mathbf{O} \neq \emptyset$ et sans second membre (problème de scattering pour un obstacle parfaitement conducteur)

#### **Théorème 2.11**

Soit  $(\mathbf{M}, g)$  une variété absorbante sortante et  $G_{A,B}(\cdot, \mathbf{y}) = \begin{pmatrix} G_1 \\ G_2 \end{pmatrix}$ , où  $A$  et  $B$  sont des 1-formes. On a alors la formulation type Stratton-Chu suivante, avec  $(e, h)$  les solutions du problème (2.17),

$$e \wedge \star A + h \wedge \star B = \left\{ \int_{\partial\Omega_\infty} e \wedge G_2 - h \wedge G_1 \right\} \mathbf{dvol}_g(\mathbf{y}).$$

**Démonstration** : soit  $A$  et  $B$  des 1-formes,

$$\begin{aligned} e \wedge \star A + h \wedge \star B &= \left\{ \int_{\mathbb{R}^3/\mathbf{0}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathbf{p} \star + \mathcal{M}^T) G_{A,B}(\cdot, \mathbf{y}) \right\} \mathbf{dvol}_g(\mathbf{y}) \\ &= \left\{ \int_{\mathbb{R}^3/\mathbf{0}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathbf{p} \star G_{A,B}(\cdot, \mathbf{y})) \right\} \mathbf{dvol}_g(\mathbf{y}) \\ &\quad + \left\{ \int_{\mathbb{R}^3/\mathbf{0}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathcal{M}^T G_{A,B}(\cdot, \mathbf{y})) \right\} \mathbf{dvol}_g(\mathbf{y}) \end{aligned}$$

$$\text{Or } \int_{\mathbb{R}^3/\mathbf{0}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathcal{M}^T G_{A,B}(\cdot, \mathbf{y})) = \int_{\mathbb{R}^3/\mathbf{0}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge \left( \begin{pmatrix} 0 & \mathbf{d} \\ -\mathbf{d} & 0 \end{pmatrix} G_{A,B}(\cdot, \mathbf{y}) \right).$$

$$\text{De même, } \int_{\mathbb{R}^3/\mathbf{0}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathbf{p} \star G_{A,B}(\cdot, \mathbf{y})) \text{ becomes } \int_{\mathbb{R}^3/\mathbf{0}} G_{A,B}(\cdot, \mathbf{y}) \wedge \left( \mathbf{p} \star \begin{pmatrix} e \\ h \end{pmatrix} \right).$$

Comme le second membre est nul, l'équation (2.18) implique  $\mathbf{p} \star \begin{pmatrix} e \\ h \end{pmatrix} = -\mathcal{M} \begin{pmatrix} e \\ h \end{pmatrix}$ .

Mais  $-\mathcal{M} = \mathcal{M}^T$ , donc

$$\int_{\mathbb{R}^3/\mathbf{0}} \begin{pmatrix} e \\ h \end{pmatrix} \wedge (\mathbf{p} \star G_{A,B}(\cdot, \mathbf{y})) = \int_{\mathbb{R}^3/\mathbf{0}} G_{A,B}(\cdot, \mathbf{y}) \wedge \begin{pmatrix} 0 & \mathbf{d} \\ -\mathbf{d} & 0 \end{pmatrix} \begin{pmatrix} e \\ h \end{pmatrix}$$

Posons  $G_{A,B}(\cdot, \mathbf{y}) = \begin{pmatrix} G_1 \\ G_2 \end{pmatrix}$ . Alors :

$$\begin{aligned} e \wedge \star A + h \wedge \star B &= \left\{ \int_{\mathbb{R}^3/\mathbf{0}} \begin{pmatrix} e \wedge \mathbf{d} G_2 + G_2 \wedge \mathbf{d} e \\ -h \wedge \mathbf{d} G_1 + G_1 \wedge (-\mathbf{d} h) \end{pmatrix} \right\} \mathbf{dvol}_g(\mathbf{y}) \\ &= \left\{ \int_{\mathbb{R}^3/\mathbf{0}} \begin{pmatrix} e \wedge \mathbf{d} G_2 - \mathbf{d} e \wedge G_2 \\ -h \wedge \mathbf{d} G_1 + \mathbf{d} h \wedge G_1 \end{pmatrix} \right\} \mathbf{dvol}_g(\mathbf{y}) \\ &= \left\{ \int_{\mathbb{R}^3/\mathbf{0}} \begin{pmatrix} \mathbf{d}(e \wedge G_2) \\ -\mathbf{d}(h \wedge G_1) \end{pmatrix} \right\} \mathbf{dvol}_g(\mathbf{y}). \end{aligned}$$

En appliquant la formule de Stokes, il vient

$$e \wedge \star A + h \wedge \star B = \left\{ \int_{\partial\Omega_\infty} e \wedge G_2 - h \wedge G_1 \right\} \mathbf{dvol}_g(\mathbf{y}).$$

**Remarque** : ce théorème est très important. D'une part parce qu'il implique le théorème suivant, qui donne la solution pour un problème de scattering pour les équations de Maxwell, d'autre part parce que cette formule permet d'estimer l'erreur commise en bornant le domaine. Pour tout l'espace, on a exactement la restriction de la solution qui coïncide avec la solution « réelle ». Si le problème est bien posé, on aura correspondance entre le problème inhomogène et la trace de la solution. Dans le cas d'une solution homogène, cette formule donnera donc l'« erreur ». De plus, si les erreurs commises sur le bord de la frontière artificielle sont petites, étant donné la décroissance exponentielle du noyau de Green, nous sommes assurés de ne pas avoir besoin de choisir une frontière artificielle loin du domaine d'étude  $\mathbf{D}$  : les traces de

la solution PML seront petites et seront encore exponentiellement atténuées lors de leur retour dans le domaine  $\mathbf{D}$ .

De plus cette formule permet de superposer les solutions : pour un obstacle diffractant parfait, la solution des Equations de Maxwell sera fonction de la valeur de la trace des champs sur la frontière de cet objet.

**Théorème 2.12**

Soit  $(\mathbf{M}, g)$  une variété absorbante sortante. Alors le problème (2.17) admet une unique solution

$$\begin{cases} E = \left\{ \int_{\partial\Omega_0} (-\eta \wedge (G_{d\tilde{\mathbf{x}}_j, 0})_2 + \gamma \wedge (G_{d\tilde{\mathbf{x}}_j, 0})_1) \right\} d\tilde{\mathbf{x}}_j \\ H = \left\{ \int_{\partial\Omega_0} (-\eta \wedge (G_{0, d\tilde{\mathbf{x}}_j})_2 + \gamma \wedge (G_{0, d\tilde{\mathbf{x}}_j})_1) \right\} d\tilde{\mathbf{x}}_j. \end{cases}$$

où  $\eta$  et  $\gamma$  sont des courants suffisamment réguliers.

## 2.4 Formulation harmonique

Dans cette section, nous allons montrer que le problème PML peut être écrit sous la forme suivante pour les Equations de Maxwell et des Ondes :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) \begin{pmatrix} e \\ h \end{pmatrix} + \begin{pmatrix} 0 & -\nabla \times \\ \nabla \times & 0 \end{pmatrix} \begin{pmatrix} e \\ h \end{pmatrix} = \begin{pmatrix} -j \\ -m \end{pmatrix}$$

où  $(j, m)$  sont des seconds membres réguliers. Grâce à cette écriture, nous établissons un théorème d'existence et d'unicité pour la formulation harmonique de ces équations.

**Théorème 2.13**

Sous les hypothèses suivantes :

- $\mathbf{K}(\mathbf{p}, \mathbf{x})$   $L^\infty$  dans  $\mathbf{x}$
- $\mathbf{K}(\mathbf{p}, \mathbf{x})$  holomorphe en  $\mathbf{p} \in \mathbf{H} \supset i\mathbb{R}_*^+$
- $\exists \mathbf{p}_0 \in \mathbf{H}$  tel que  $\mathbf{Re}(\mathbf{K}(\mathbf{p}_0, \mathbf{x}))$  est coercive
- $\exists \mathbf{p}_1 \in \mathbf{H}$  tel que  $\mathbf{K}(\mathbf{p}_1, \mathbf{x}) + \mathbf{A}_i^c \partial_i$  soit à résolvante compacte,

alors, pour  $\mathbf{p} \in \mathbf{H}/\mathbf{S}$  où  $\mathbf{S}$  est un ensemble localement fini, pouvant être nul,  $\Omega^\infty$  borné :  $(\mathbf{K}(\mathbf{p}, \mathbf{x}) + \mathbf{A}_i^c \partial_i)^{-1} \in \mathcal{B}(L^2(\Omega)^m)$  avec des conditions aux limites maximales monotones.

### 2.4.1 Ecriture intrinsèque des équations

Ecrivons une formulation PML non convexes à partir des équations de Maxwell intrinsèques sur des variétés pseudo-Riemanniennes absorbantes.

Notations : les coordonnées contravariantes d'un point  $\mathbf{x}$  quelconque de l'espace seront données par  $(x_1, x_2, x_3)$ , les vecteurs de base du fibré cotangent par  $dx_i$  et on

désignera par  $d\tilde{x}_i$  la 2-forme élémentaire donnée par le produit extérieur  $dx_j \wedge dx_k$  avec  $j \neq i, k \neq i$  et  $j < k$ .

Soient  $E$  et  $H$  des 1-formes. Soit  $(\mathbf{M}, g)$  une variété absorbante de métrique  $g$ . Soit  $\star$  la transformée de Hodge. Les équations de Maxwell peuvent s'écrire en terme d'égalités de 2-formes,

$$\begin{cases} \mathbf{p} \star E - \mathbf{d} H = 0, \\ \mathbf{p} \star H + \mathbf{d} E = 0. \end{cases}$$

En coordonnées cartésiennes, on peut dire que  $\mathbf{rot}$  en tant que matrice est la transformée de Hodge d'une 2-forme en une 1-forme, donc on obtient en utilisant les variables  $\mathbf{E}, \mathbf{H}, \mathbf{rotH}, \mathbf{rotE}$  des égalités de 1-formes.

$$\begin{cases} \mathbf{pE} - \mathbf{rotH} = 0, \\ \mathbf{pH} + \mathbf{rotE} = 0. \end{cases}$$

Soit  $\mathbf{G}$  la matrice de la métrique  $(g_{ij})$  et  $g = \mathbf{det}(\mathbf{G})$ . Par la transformée de Hodge  $\star$  sur  $\mathbf{M}$ , une 1-forme  $E = \sum e_i dx_i$  se transforme en une 2-forme dont la composante en  $d\tilde{x}_i$  est donnée par  $\sqrt{g}g^{ij}e_j$ , où  $g^{ij}$  désignent les composantes de l'inverse de  $g$ ,  $\sqrt{g}$  est la racine carrée complexe du déterminant de  $g$ . On a ainsi :  $\star E = \sqrt{g}\mathbf{G}^{-1}e$ . D'où l'écriture intrinsèque des équations de Maxwell

$$\begin{cases} \mathbf{p}\sqrt{g}\mathbf{G}^{-1}E - \mathbf{d} H = 0, \\ \mathbf{p}\sqrt{g}\mathbf{G}^{-1}H + \mathbf{d} E = 0. \end{cases}$$

L'objectif est de restreindre les solutions de l'équation de Maxwell à un domaine  $\Omega$  de  $\mathbb{R}^3$ . Nous allons dans un premier temps ramener la représentation des équations de Maxwell dans un autre système de coordonnées éventuellement plus propice à l'introduction des PML.

Rappelons que l'écriture des PML revient à faire un stretching de coordonnées en travaillant avec la variété réelle à valeurs dans  $\mathbb{C}^3$  décrite par le changement de variables  $\tilde{\mathbf{x}} = \varphi(\mathbf{x}) + \frac{1}{\mathbf{p}}f(\varphi(\mathbf{x}))$ , où le couple  $(\varphi, f)$  doit respecter les hypothèses (2.5). Soit  $\mathbf{J}(\mathbf{p}, \mathbf{x})$  la matrice jacobienne de cette transformation. Il est donc maintenant possible d'écrire la métrique  $g$  du changement de variable à l'aide de cette matrice,

$$\mathbf{G}(\mathbf{p}, \mathbf{x}) = \mathbf{J}(\mathbf{p}, \mathbf{x})^T \mathbf{J}(\mathbf{p}, \mathbf{x}).$$

La transformation de Hodge associée étant

$$\star = \sqrt{g}\mathbf{G}^{-1}(\mathbf{p}, \mathbf{x}) = \mathbf{det}(\mathbf{J}(\mathbf{p}, \mathbf{x}))\mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} (\mathbf{J}(\mathbf{p}, \mathbf{x})^T)^{-1},$$

on en déduit la matrice de la formulation PML dispersive

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{p}\star = \mathbf{pdet}(\mathbf{J}(\mathbf{p}, \mathbf{x}))\mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} (\mathbf{J}(\mathbf{p}, \mathbf{x})^T)^{-1}. \quad (2.20)$$

### 2.4.2 Cas particulier des PML cartésiennes et cylindriques

Le but de ce formalisme est d'être le plus général possible. Aussi convient-il de vérifier que les PML « classiques », à savoir les PML cartésiennes et cylindriques présentées dans (2.1.2), s'inscrivent dans ce cadre.

Définissons les ensembles emboîtés  $\Omega_\alpha$  tels que  $S(\varphi) = \alpha$  sur  $\partial\Omega_\alpha$ , où  $S$  est la fonction convexe dont  $f$  est le gradient. Ces ensembles définissent une exhaustion de domaines bornés avec  $\mathbf{D} \subset \Omega_0$ .

Pour les PML cartésiennes, le changement de variable complexe est donné par

$$\tilde{\mu} = \mu + \frac{1}{\mathbf{p}} \int_0^\mu \sigma_\mu(\tau) d\tau.$$

Le domaine peut s'écrire sous la forme  $\mathbf{D} = \prod_i [\alpha_i, \beta_i]$ . Il est déjà convexe, et vu le changement de variable, en posant  $\varphi(\mathbf{x}) = \mathbf{x}$  et  $f(\mathbf{x}) = \left( \left( \int_0^{x_i} \sigma_{x_i}(\tau) d\tau \right) \cdot e_i \right)_{i=1, \dots, n}$  on retrouve bien le formalisme énoncé. Vu les régularités des fonctions  $\sigma_\mu$  en présence, la fonction  $f$  respecte bien les conditions (2.5). Pour les PML cylindriques, le changement de variable complexe est donné par

$$\tilde{\rho} = \rho + \frac{1}{\mathbf{p}} \int_{\rho_0}^\rho \sigma(\tau) d\tau.$$

Pour les PML convexes, le domaine  $\mathbf{D}$  est donc confondu avec  $\Omega_0$ . En posant  $\phi(\mathbf{x}) = \mathbf{x}$  et  $f(\mathbf{x}) = \partial_i h(d(\mathbf{x}, \partial\Omega))$  avec  $h$  une fonction croissante convexe asymptotiquement linéaire, les hypothèses (2.5) sont encore respectées.

**Remarque fondamentale** : les ensembles  $\Omega_\alpha$  ne sont pas nécessairement convexes mais  $\varphi(\Omega_\alpha)$  l'est. On retrouve donc la notion de convexité, mais à un difféomorphisme de  $\mathbb{R}^3$  près. Cependant plus  $\Omega_0$  sera proche de  $\mathbf{D}$  plus le jacobien de la transformation  $\left( \frac{D\varphi}{D\mathbf{x}} \right)$  sera grand...

### 2.4.3 Existence et unicité de la solution au problème harmonique

Le domaine de la famille d'opérateurs  $(\mathbf{K}(\mathbf{p}, \mathbf{x}) + \mathbf{A})_{\mathbf{p} \in \mathbf{D}_0}$  est holomorphe et son domaine est indépendant de la fréquence  $\omega$ . Résoudre une Equation aux Dérivées Partielles revenant à inverser un opérateur, il semble naturel de regarder plus en détail les familles holomorphes d'opérateurs à domaines dont le domaine est indépendant de  $z$ . Un de ces types de famille est appelé opérateur de type A. Le but est de démontrer l'existence et l'unicité de solutions pour les Equations de Maxwell en harmonique. La majeure partie des résultats énoncés dans ce chapitre sont tirés du livre de T. Kato [18].

**Résultats principaux sur les opérateurs de type A**

Soit  $(\mathbf{X}, \|\cdot\|)$  un espace de Banach et  $(\mathcal{T}, \mathcal{D}(\mathcal{T}))$  un opérateur non borné. Nous noterons aussi  $\|\mathcal{T}\|$  la norme de l'opérateur  $\mathcal{T}$  (lorsqu'elle existe). Si le domaine de l'opérateur  $\mathcal{T}$  est fermé, l'opérateur est borné à savoir  $\mathcal{T} \in \mathcal{B}(\mathcal{D}(\mathcal{T}), \mathbf{X})$  (théorème du graphe fermé). Il est alors possible de définir la norme de  $\mathcal{T}$  par 
$$\|\mathcal{T}\| = \sup_{u \in \mathcal{D}(\mathcal{T}) \setminus \{0\}} \frac{\|\mathcal{T}u\|}{\|u\|}.$$

**Définition 2.5**

On appelle résolvante de l'opérateur  $(\mathcal{T}, \mathcal{D}(\mathcal{T}))$  l'opérateur  $\mathcal{R}(z, \mathcal{T}) = (\mathcal{T} - z)^{-1}$ . Lorsque la résolvante existe, on dit que  $z$  appartient à l'ensemble résolvant de  $\mathcal{T}$ . L'ensemble résolvant de  $\mathcal{T}$  est noté  $\mathbf{P}(\mathcal{T})$ . L'ensemble  $\Sigma(\mathcal{T}) = \mathbb{C} \cap \mathbf{P}(\mathcal{T})^c$  est appelé spectre de  $\mathcal{T}$ .

**Remarque :** pour tout  $z$  appartenant à  $\mathbf{P}(\mathcal{T})$  l'opérateur  $\mathcal{R}(z, \mathcal{T})$  admet  $\mathbf{X}$  pour domaine et  $\mathcal{D}(\mathcal{T})$  pour image. Le théorème du graphe fermé nous assure alors que :  $\forall z \in \mathbf{P}(\mathcal{T}), \mathcal{R}(z, \mathcal{T}) \in \mathcal{B}(\mathbf{X}, \mathcal{D}(\mathcal{T}))$ .

Donc  $\mathbf{P}(\mathcal{T})$  est un ouvert de  $\mathbb{C}$  et la résolvante (en tant que fonction de  $z$ ) ne peut s'étendre en une fonction analytique au voisinage de la frontière de  $\mathbf{P}(\mathcal{T})$ .

L'alternative de Fredholm analytique donne un critère d'inversion d'une famille holomorphe d'opérateurs. Soit  $(\mathbf{X}, \|\cdot\|)$  un espace de Banach. Considérons une famille d'opérateurs dépendant de manière holomorphe d'un paramètre. Soit  $\mathbf{D}_0$  le domaine d'holomorphie de l'opérateur.

Soit  $\mathbf{X}, \mathbf{Y}$  deux espaces de Banach. On note  $C(\mathbf{X}, \mathbf{Y})$  des opérateurs fermés de  $\mathbf{X}$  dans  $\mathbf{Y}$ .

Rappelons la définition de l'holomorphie pour des opérateurs bornés. On dit qu'une famille d'opérateurs  $\mathcal{T}(x)$  est holomorphe sur  $\mathbf{D}_0$  un domaine de  $\mathbb{C}$  (on dira aussi holomorphe-bornée) si elle est différentiable en norme pour tout  $x$  appartenant à  $\mathbf{D}_0$ . Lorsque  $\mathcal{T}(x) \in C(\mathbf{X}, \mathbf{Y})$  on a la définition suivante :

**Définition 2.6**

Une famille d'opérateurs  $\mathcal{T}(x) \in C(\mathbf{X}, \mathbf{Y})$  définie au voisinage de  $x = 0$  est holomorphe en  $x = 0$  (au sens généralisé) s'il existe un troisième espace de Banach  $\mathbf{Z}$  et deux familles d'opérateurs  $\mathcal{U}(x) \in B(\mathbf{X}, \mathbf{Y})$  et  $\mathcal{V}(x) \in B(\mathbf{Z}, \mathbf{Y})$  qui sont holomorphes-bornées en  $x = 0$  telles que  $\mathcal{U}(x)$  envoie injectivement  $\mathbf{Z}$  dans  $\mathcal{D}(\mathcal{T}(x))$  et  $\mathcal{T}(x)\mathcal{U}(x) = \mathcal{V}(x)$ .

$\mathcal{T}(x)$  est holomorphe (au sens généralisé) sur un domaine  $\mathbf{D}_0$  de  $\mathbb{C}$  si  $\mathcal{T}(x)$  est holomorphe pour tout  $x$  appartenant à  $\mathbf{D}_0$ .

**Théorème 2.14** (Alternative de Fredholm analytique)

Soit  $\mathcal{T}(z) \in C(\mathbf{X})$  une famille holomorphe d'opérateurs compacts pour tout  $z$  appartenant à  $\mathbf{D}_0$ . On a alors l'alternative suivante :

- $(\mathcal{I} - \mathcal{T}(z))^{-1}$  n'existe pour aucun  $z$  appartenant à  $\mathbf{D}_0$ ,
- $(\mathcal{I} - \mathcal{T}(z))^{-1}$  existe pour tout  $z$  appartenant à  $\mathbf{D}_0 \setminus \mathbf{S}$  où  $\mathbf{S}$  est un sous ensemble discret et localement fini éventuellement vide de  $\mathbf{D}_0$ .

La proposition suivante permet de faire le lien entre le caractère holomorphe d'une famille d'opérateurs et son inverse.

**Proposition 2.15**

Soit  $\mathcal{T}(z)$  une famille holomorphe d'opérateurs. On suppose que  $\mathcal{T}(z_0)$  est inversible pour  $z_0$  fixé dans  $\mathbf{D}_0$ . Alors, pour  $z$  dans un voisinage de  $z_0$ , l'opérateur  $\mathcal{T}(z)$  est inversible d'inverse holomorphe sur ce même voisinage.

**Démonstration :** Un calcul montre que pour tout  $z$  appartenant à  $\mathbf{D}_0$ , on a  $\mathcal{T}(z) = \mathcal{T}(z_0)\{1 - \mathcal{T}(z_0)^{-1}(\mathcal{T}(z_0) - \mathcal{T}(z))\}$ . Par continuité de  $\mathcal{T}(z)$  par rapport à  $z$ , il existe  $r > 0$  tel que pour tout  $z \in \mathbf{B}(z_0, r)$ , on ait :  $\|\mathcal{T}(z) - \mathcal{T}(z_0)\| < \frac{1}{\|\mathcal{T}(z_0)\|}$ . Dans ce cas, la série de Neumann de  $1 - \mathcal{T}(z_0)^{-1}(\mathcal{T}(z_0) - \mathcal{T}(z))$  est convergente pour tout  $z \in \mathbf{B}(z_0, r)$ . Ainsi l'opérateur  $\mathcal{T}(z)$  est inversible pour tout  $z \in \mathbf{B}(z_0, r)$ .

Pour montrer que cet inverse est holomorphe, on calcule la quantité

$$\frac{1}{h}\{\mathcal{T}(z+h)^{-1} - \mathcal{T}(z)^{-1}\} = \frac{1}{h}\{\mathcal{T}(z+h)^{-1}(\mathcal{T}(z+h) - \mathcal{T}(z))\mathcal{T}(z)^{-1}\}.$$

En passant à la limite quand  $h \rightarrow 0$ , on obtient :  $\frac{\partial \mathcal{T}(z)^{-1}}{\partial z} = \mathcal{T}(z)^{-1} \frac{\partial \mathcal{T}(z)}{\partial z} \mathcal{T}(z)^{-1}$ .

⊠

Dans cette section  $\mathbf{D}_0$  représente le domaine d'holomorphie mais il peut être remplacé par tout autre domaine contenu strictement dans le domaine d'holomorphie. Nous utiliserons plutôt une autre version de l'alternative de Fredholm analytique liée au caractère compact de la résolvante et non plus à la compacité de l'opérateur.

**Théorème 2.16**

Soit  $\mathcal{T}(z) \in C(\mathbf{X})$  une famille holomorphe d'opérateurs pour  $z \in \mathbf{D}_0$ . On suppose que la résolvante de  $\mathcal{T}(z)$  est compacte pour tout  $z$ . On a alors l'alternative suivante :

- $\mathcal{T}(z)^{-1}$  n'existe pour aucun  $z$  appartenant à  $\mathbf{D}_0$ ,
- $\mathcal{T}(z)^{-1}$  existe pour tout  $z$  appartenant à  $\mathbf{D}_0 \setminus \mathbf{S}$  où  $\mathbf{S}$  est un sous ensemble discret et localement fini éventuellement vide de  $\mathbf{D}_0$ .

L'opérateur à inverser pour résoudre les équations de Maxwell harmoniques est un opérateur non borné perturbé par un opérateur de multiplication holomorphe et borné. De plus le domaine de l'opérateur perturbé est indépendant de la fréquence.

Nous allons donc nous intéresser à une classe d'opérateurs holomorphes dont le domaine ne dépend pas de  $z$  : les opérateurs de type A.

Soient  $M$  et  $N$  deux variétés linéaires fermées d'un espace de Banach  $\mathbf{Z}$ . Soit

$\mathcal{S}_M$  la sphère unité de  $M$  et les définitions d'écart suivantes :

$$\delta(M, N) = \sup_{u \in \mathcal{S}_M} \mathbf{dist}(u, N) \qquad \bar{\delta}(M, N) = \mathbf{max}(\delta(M, N), \delta(N, M)).$$

Soient  $\mathcal{T}, \mathcal{S} \in C(\mathbf{X}, \mathbf{Y})$ . Leurs graphes  $G(\mathcal{T}), G(\mathcal{S})$  sont des variétés linéaires fermées de l'espace produit  $\mathbf{X} \times \mathbf{Y}$ . On définit l'écart entre  $\mathcal{T}$  et  $\mathcal{S}$  par

$$\delta(\mathcal{T}, \mathcal{S}) = \delta(G(\mathcal{T}), G(\mathcal{S})) \qquad \bar{\delta}(\mathcal{T}, \mathcal{S}) = \bar{\delta}(G(\mathcal{T}), G(\mathcal{S})).$$

**Définition 2.7**

Soit  $(\mathcal{T}_n)$  une suite d'éléments de  $C(\mathbf{X}, \mathbf{Y})$ . On dit que  $\mathcal{T}_n$  converge vers  $\mathcal{T}$  (au sens généralisé) si  $\bar{\delta}(\mathcal{T}_n, \mathcal{T}) \rightarrow 0$ .

Le théorème suivant permet de faire le lien entre la convergence en norme des opérateurs bornés et la convergence généralisée.

**Théorème 2.17**

Soient  $\mathcal{T}, \mathcal{T}_n \in C(\mathbf{X}, \mathbf{Y})$ . Si  $\mathcal{T} \in \mathcal{B}(\mathbf{X}, \mathbf{Y})$ ,  $\mathcal{T}_n$  converge vers  $\mathcal{T}$  au sens généralisé si et seulement si  $\mathcal{T}_n \in \mathcal{B}(\mathbf{X}, \mathbf{Y})$  pour des  $n$  assez grands et  $\|\mathcal{T}_n - \mathcal{T}\| \rightarrow 0$ .

Ce théorème montre que la convergence en norme implique la convergence généralisée.

**Théorème 2.18**

Soient  $\mathcal{T}_n, \mathcal{T} \in C(\mathbf{X})$ . On suppose que  $\mathcal{T}_n \rightarrow \mathcal{T}$  au sens généralisé. Si tous les  $\mathcal{T}_n$  admettent une résolvante compacte et si  $\mathcal{T}$  admet un ensemble résolvant non vide alors  $\mathcal{T}$  a une résolvante compacte.

Ce résultat sera utile par la suite pour démontrer un résultat fort concernant les opérateurs de type A.

**Définition 2.8**

Soit  $\mathbf{D}_0$  un domaine de  $\mathbb{C}$ . On dit qu'une famille holomorphe d'opérateurs  $\mathcal{T}(z) \in C(\mathbf{X}, \mathbf{Y})$  est de type (A) lorsque :

- $\mathcal{D}(\mathcal{T}(z)) = \mathbf{D}$ , le domaine de  $\mathcal{T}(z)$  ne dépend pas de  $z$ .
- $\forall z \in \mathbf{D}_0, \forall u \in \mathcal{D}, z \in \mathbf{D}_0 \rightarrow \mathcal{T}(z)u \in \mathbf{Y}$  est holomorphe.

**Proposition 2.19**

Une famille holomorphe de type A est aussi holomorphe au sens généralisé de la définition 2.6.

**Démonstration** : soit  $z \in \mathbf{D}_0$ . On peut supposer sans perte de généralité que  $z = 0$ . Comme  $\mathcal{T} = \mathcal{T}(0)$  est un opérateur fermé,  $\mathbf{D} = \mathcal{D}(\mathcal{T})$  est un espace de Banach pour la norme  $\|u\|_{\mathbf{D}} = \|u\| + \|\mathcal{T}u\|$ . On notera cet espace  $\mathbf{Z}$ . Soit  $U$  l'injection de  $\mathbf{Z}$  dans  $\mathbf{X}$  qui est bornée car  $\|u\| \leq \|u\|_{\mathbf{D}}$ . Pour tout  $z$  l'opérateur  $\mathcal{T}(z)$  peut être vu comme un opérateur de  $\mathbf{Z}$  dans  $\mathbf{Y}$  que l'on va noter  $\mathcal{V}(z)$ . Comme  $\mathcal{T}(z)$  est fermé et que  $\|u\| \leq \|u\|_{\mathbf{D}}$  l'opérateur  $\mathcal{V}(z)$  est fermé. Dans la mesure où  $\mathcal{V}(z)$  est définie sur  $\mathbf{Z}$ , le théorème du graphe fermé implique que  $\mathcal{V}(z)$  appartient à  $\mathcal{B}(\mathbf{Z}, \mathbf{Y})$ . Comme  $\mathcal{V}(z)u = \mathcal{T}(z)u$  et que  $\mathcal{T}(z)u$  est holomorphe pour tout  $u \in \mathbf{Z}$

il vient que  $\mathcal{V}(z)$  est holomorphe-borné. Pour conclure on remarque que  $U$  envoie injectivement  $\mathbf{Z}$  dans  $\mathbf{D}$  et  $\mathcal{T}(z)U = \mathcal{V}(z)$  ce qui prouve que  $\mathcal{T}(z)$  est holomorphe au sens généralisé.  $\square$

Le résultat principal de cette section permet de ramener l'étude de la résolvante d'une famille d'opérateurs de type A à l'étude de la résolvante de l'opérateur  $\mathcal{T}(z_0)$  pour un  $z_0$  fixé dans  $\mathbf{D}_0$  :

**Théorème 2.20**

Soit  $\mathcal{T}(z) \in C(\mathbf{X})$  une famille holomorphe d'opérateurs de type A sur un domaine  $\mathbf{D}_0$  de  $\mathbb{C}$ . On suppose que l'ensemble résolvant de  $\mathcal{T}(z)$  est non vide  $\forall z \in \mathbf{D}_0$  et qu'il existe un point  $z_0$  de  $\mathbf{D}_0$  tel que la résolvante de  $\mathcal{T}(z_0)$  est compacte. Alors, pour tout  $z$  appartenant à  $\mathbf{D}_0$ , la résolvante de  $\mathcal{T}(z)$  est compacte.

**Remarque :** ce théorème est plus fort que (2.16). C'est son équivalent pour les opérateurs de type A. La preuve de ce théorème utilise les résultats ci dessous.

**Proposition 2.21**

Soit  $\mathcal{T}(z)$  une famille holomorphe de type A. Alors pour tout  $\varepsilon > 0$ , il existe  $\delta > 0$  tel que pour tout  $|z_1 - z_2| < \delta$ , on ait  $\| \mathcal{T}(z_1) - \mathcal{T}(z_2) \| \leq \varepsilon (\| u \| + \| \mathcal{T}(z)u \|)$ .

**Théorème 2.22**

Soit  $\mathcal{T}$  et  $\mathcal{A}$  deux opérateurs non bornés de  $\mathbf{X}$  dans  $\mathbf{X}$  vérifiant les hypothèses suivantes :  $\mathcal{T}$  est fermé,  $\mathcal{D}(\mathcal{T}) \subset \mathcal{D}(\mathcal{A})$  et il existe deux constantes positives  $a$  et  $b$  telles que  $\| \mathcal{A}u \| \leq a \| u \| + b \| \mathcal{T}u \|^2$ .

Soit  $\mathcal{R}(x, \mathcal{T}) = (\mathcal{T} - x)^{-1}$ . On suppose qu'il existe un point  $z$  appartenant à l'ensemble résolvant de  $\mathcal{T}$  tel que  $a \| \mathcal{R}(x, \mathcal{T}) \| + b \| \mathcal{T} \mathcal{R}(x, \mathcal{T}) \| < 1$ .

Alors  $\mathcal{S} = \mathcal{T} + \mathcal{A}$  est fermé et

$$\| \mathcal{R}(x, \mathcal{S}) \| \leq \| \mathcal{R}(x, \mathcal{T}) \| (1 - a \| \mathcal{R}(x, \mathcal{T}) \| - b \| \mathcal{T} \mathcal{R}(x, \mathcal{T}) \|^2)^{-1}.$$

En particulier, si  $\mathcal{T}$  admet une résolvante compacte,  $\mathcal{S}$  admet elle aussi une résolvante compacte.

**Démonstration** Supposons que l'on dispose d'une famille d'opérateurs de type A dont la résolvante est compacte en un point de  $\mathbf{D}_0$ . Alors le théorème (2.20) entraîne que la résolvante de  $\mathcal{T}(z)$  est compacte pour tout  $z$  appartenant à  $\mathbf{D}_0$ . Par conséquent, par l'alternative de Fredholm analytique, l'opérateur  $\mathcal{T}(z)$  est soit singulier sur  $\mathbf{D}_0$  soit inversible sur  $\mathbf{D}_0$  hormis un sous ensemble discret et localement fini éventuellement vide de  $\mathbf{D}_0$ .

**Corollaire 2.23**

Soit les opérateurs  $\mathbf{K}$  borné et  $\mathbf{A}$  maximal monotone. Si pour tout  $z$  appartenant au domaine, il existe un  $\lambda$  assez grand tel que l'opérateur  $\mathbf{I} + \mathbf{K}$  soit borné coercif et monotone, alors  $\mathbf{I} + \mathbf{K} + \mathbf{A}$  est inversible.

**Etude des Equations de Maxwell en présence d'éléments PML**

Les équations de Maxwell harmoniques dans les milieux PML peuvent s'écrire sous la forme :

$$(\mathbf{K}(\mathbf{p}, \mathbf{x}) + \mathbf{A}(\mathbf{x}))u(\mathbf{x}) = f(\mathbf{x}) \tag{2.21}$$

avec  $\mathbf{K} : (\mathbf{p}, \mathbf{x}) \in \mathbf{D}_0 \times \Omega \longrightarrow \mathbf{Hom}(\mathbb{C}^6)$  une fonction holomorphe en  $\mathbf{p} \in \mathbf{D}_0$  et appartenant à  $L^\infty(\Omega)$  en  $\mathbf{x}$  et  $\mathbf{A} = \begin{pmatrix} 0 & -\nabla \times \\ \nabla \times & 0 \end{pmatrix}$ . On suppose de plus que  $\mathbf{D}_0$  est un domaine de  $\mathbb{C}$  contenant un segment du type  $[i\omega_{min}, i\omega_{max}]$ , que  $f \in L^2(\Omega)$  et que  $\mathbf{A}$  est l'opérateur de Maxwell dont le domaine  $\mathcal{D}(\mathbf{A})$  est indépendant de  $\mathbf{p} \in \mathbb{C}$ .

Résoudre cette équation revient à inverser l'opérateur  $(\mathbf{K}(\mathbf{p}, \cdot) + \mathbf{A})$  de domaine  $\mathcal{D}(\mathbf{A})$  au point  $\mathbf{p} = i\omega$ . Par hypothèses, l'opérateur  $(\mathbf{K}(\mathbf{p}, \cdot) + \mathbf{A}, \mathcal{D}(\mathbf{A}))$  est holomorphe en  $\mathbf{p} \in \mathbf{D}_0$  et admet un domaine indépendant de  $\mathbf{p}$  (et donc de la fréquence) : c'est donc un opérateur de type A.

**Théorème 2.24** (Rellich)  
*L'injection de  $H^1(\Omega)$  dans  $L^2(\Omega)$  est compacte .*

Pour tout  $\mathbf{p}$  appartenant à  $\mathbf{D}_0$ ,  $\mathbf{K}(\mathbf{p}, \cdot) + \mathbf{A}$  admet  $\mathcal{D}(\mathbf{A})$  pour domaine et est holomorphe en  $\mathbf{p}$  par hypothèses sur la matrice  $\mathbf{K}(\mathbf{p}, \cdot)$ . C'est donc un opérateur de type A dont le domaine d'holomorphie est  $\mathbf{D}_0$ . De plus, l'opérateur  $(\mathbf{A}, \mathcal{D}(\mathbf{A}))$  est maximal monotone et admet un ensemble résolvant non vide. Comme  $\mathbf{K}(\mathbf{p}, \cdot)$  est  $L^\infty(\Omega)$ , l'opérateur  $(\mathbf{K}(\mathbf{p}, \cdot) + \mathbf{A}, \mathcal{D}(\mathbf{A}))$  admet un ensemble résolvant non vide pour tout  $\mathbf{p}$  appartenant à  $\mathbf{D}_0$ . Par le théorème (2.20), étudier la compacité de la résolvante de l'opérateur  $(\mathbf{K}(\mathbf{p}, \cdot) + \mathbf{A}, \mathcal{D}(\mathbf{A}))$  revient à l'étudier en un point de  $\mathbf{D}_0$ .

**Théorème 2.25**  
*Soit  $Q = \begin{pmatrix} \mathbf{div} & 0 \\ 0 & \mathbf{div} \end{pmatrix}$ . Pour tout  $u$  dans  $\mathcal{D}(\mathbf{A}) \cap (H(\mathbf{div}, \Omega)^2)$ , on a :*

$$\| u \|_{(H^1(\Omega))^6} \leq C \{ \| u \|_{(L^2(\Omega))^6} + \| \mathbf{A}u \|_{(L^2(\Omega))^6} + \| Qu \|_{(L^2(\Omega))^6} \} .$$

**Définition 2.9**  
*On notera  $P_0 : L^2(\Omega)^3 \longrightarrow H(\mathbf{div}_0, \Omega)$  et  $P_\perp : L^2(\Omega)^3 \longrightarrow (\mathbf{grad}(H_0^1(\Omega)))^2$  les projections associées à la décomposition de Hodge et, pour  $g \in L^2(\Omega)^3$ ,  $g_0 = P_0g$ ,  $g_\perp = P_\perp g$ .  
 L'espace  $H(\mathbf{div}_0, \Omega)^2$  est stable par application de l'opérateur  $\mathbf{A}$ .*

**Lemme 2.26**  
*L'opérateur  $P_\perp \mathbf{K}(\mathbf{p}, \cdot) P_\perp$  est inversible sur  $(\mathbf{grad}(H_0^1(\Omega)))^2$  pour tout  $\mathbf{p}$  appartenant à  $\mathbf{D}_0$  hormis un sous-ensemble fini (eventuellement vide) de  $\mathbf{D}_0$ , noté  $\mathbf{S}$ . De plus, son inverse est holomorphe sur  $\mathbf{D}_0 \setminus \mathbf{S}$  .*

**Démonstration** : soit  $h \in H^{-1}(\Omega)^2$ . Inverser  $P_\perp \mathbf{K}(\mathbf{p}) P_\perp$  sur  $(\mathbf{grad}(H_0^1(\Omega)))^2$  revient à trouver  $\omega$  dans  $H_0^1(\Omega)^2$  tel que  $\forall v \in H_0^1(\Omega)^2$  on ait

$$\int_\Omega \langle \mathbf{K}(\mathbf{p}, \cdot) \nabla \omega, \overline{\nabla v} \rangle dx = \langle h, v \rangle_{H^{-1}(\Omega)^2 \times H_0^1(\Omega)^2}, \tag{2.22}$$

avec  $\nabla\omega = \begin{pmatrix} \nabla\omega_1 \\ \nabla\omega_2 \end{pmatrix}$ .

Soient  $a(\omega, v) = \int_{\Omega} \langle \mathbf{K}(\mathbf{p}, \cdot) \nabla\omega, \overline{\nabla v} \rangle dx$ , et  $L(v) = \langle h, v \rangle_{H^{-1}(\Omega)^2 \times H_0^1(\Omega)^2}$ . Les applications  $a(\cdot, \cdot)$  et  $L(\cdot)$  sont respectivement bilinéaire et linéaire. De plus, une inégalité de Cauchy-Schwartz montre que  $a(\cdot, \cdot)$  est continue sur  $H_0^1(\Omega)^2 \times H_0^1(\Omega)^2$  et que  $L(\cdot)$  est continue sur  $H_0^1(\Omega)^2$ . On remarque de plus que le problème précédent revient en fait à trouver  $\omega$  dans  $H_0^1(\Omega)^2$  tel que  $\forall v \in H_0^1(\Omega)^2$ ,

$$a(\omega, v) = L(v). \quad (2.23)$$

En utilisant le théorème de Lax-Milgram en  $\mathbf{p} = \mathbf{p}_0$ , on obtient que cette équation est bien posée. En utilisant la coercivité de  $\mathbf{K}(\mathbf{p}_0, \cdot)$  ainsi que la continuité de  $L(\cdot)$  sur  $H_0^1(\Omega)^2$ , on obtient que la solution vérifie  $\beta \| \omega \|_{H_0^1(\Omega)^2} \leq \| h \|_{H^{-1}(\Omega)^2}$ , avec  $\beta$  une constante strictement positive.

De cette dernière inégalité, on déduit que la résolvante de  $P_{\perp} \mathbf{K}(\mathbf{p}_0, \cdot) P_{\perp}$  est compacte car l'injection  $H_0^1(\Omega)$  dans  $H^{-1}(\Omega)$  est compacte. Alors, par le théorème (2.20), la résolvante de  $P_{\perp} \mathbf{K}(\mathbf{p}, \cdot) P_{\perp}$  est compacte pour tout  $\mathbf{p}$  dans  $\mathbf{D}_0$ . Par conséquent, cet opérateur est soit singulier sur  $\mathbf{D}_0$  soit inversible pour tout  $\mathbf{p}$  appartenant à  $\mathbf{D}_0$  hormis un sous-ensemble discret localement fini et éventuellement vide de  $\mathbf{D}_0$  noté  $\mathbf{S}$ . Pour lever cette alternative, on remarque que le problème (2.23) est bien posé pour  $\mathbf{p} = \mathbf{p}_0$ .

Montrons maintenant que l'opérateur  $(P_{\perp} \mathbf{K}(\mathbf{p}, \cdot) P_{\perp})^{-1}$  est holomorphe sur  $\mathbf{D}_0 \setminus \mathbf{S}$ . En appliquant le théorème du graphe fermé au fermé  $\mathbf{grad}(H_0^1(\Omega))^2$  de  $L^2(\Omega)^6$ , on obtient que  $(P_{\perp} \mathbf{K}(\mathbf{p}, \cdot) P_{\perp})^{-1} \in \mathcal{L}((\mathbf{grad}(H_0^1(\Omega)))^2)$ . Comme  $P_{\perp} \mathbf{K}(\mathbf{p}, \cdot) P_{\perp}$  est holomorphe sur  $\mathbf{D}_0$ ,  $(P_{\perp} \mathbf{K}(\mathbf{p}, \cdot) P_{\perp})^{-1}$  est holomorphe sur  $\mathbf{D}_0 \setminus \mathbf{S}$ .  $\square$

**Théorème 2.27** ([18])

*On suppose qu'il existe  $\mathbf{p}_0 \in \mathbf{D}_0$  tel que  $\mathbf{K}(\mathbf{p}_0, \cdot)$  soit coercif. Alors l'équation (2.21) est bien posée pour toute fréquence réelle hormis pour un ensemble discret localement fini et éventuellement vide de  $\mathbb{R}$ .*

**Démonstration** : en projetant l'équation (2.22) selon la décomposition de Hodge le problème est désormais de trouver  $u = u_0 + u_{\perp} \in \mathcal{D}(\mathbf{A})$  tels que

$$\begin{cases} P_0 \mathbf{K}(\mathbf{p}, \cdot) u + \mathbf{A} u_0 = f_0, \\ P_{\perp} \mathbf{K}(\mathbf{p}, \cdot) u = f_{\perp}. \end{cases}$$

Comme  $u = P_0 u + P_{\perp} u$ , l'équation précédente devient :

$$\begin{cases} P_0 \mathbf{K}(\mathbf{p}, \cdot) P_0 u + P_0 \mathbf{K}(\mathbf{p}, \cdot) P_{\perp} u + \mathbf{A} u = f_0, \\ P_{\perp} \mathbf{K}(\mathbf{p}, \cdot) P_{\perp} u + P_{\perp} \mathbf{K}(\mathbf{p}, \cdot) P_0 u = f_{\perp}. \end{cases}$$

En utilisant le lemme (2.26), si l'on note  $\tilde{\mathbf{A}}$  l'opérateur défini par la restriction de  $\mathbf{A}$  à l'ensemble  $(H(\mathbf{div}_0, \Omega))^2$  alors  $\mathcal{D}(\tilde{\mathbf{A}}) = \mathcal{D}(\mathbf{A}) \cap (H(\mathbf{div}_0, \Omega))^2$  et le problème précédent revient à inverser dans  $(H(\mathbf{div}_0, \Omega))^2$  l'opérateur fermé

$$\mathcal{R}(\mathbf{p}) := P_0 \mathbf{K}_0(\mathbf{p}, \cdot) P_0 - P_0 \mathbf{K}_0(\mathbf{p}, \cdot) P_\perp (P_\perp \mathbf{K}_0(\mathbf{p}, \cdot) P_\perp)^{-1} P_\perp \mathbf{K}_0(\mathbf{p}, \cdot) P_0 + P_0 \mathbf{K}_1(\mathbf{p}, \cdot) \tilde{\mathbf{A}}$$

L'opérateur  $\mathcal{R}(\mathbf{p})$  est holomorphe sur  $\mathbf{D}_0 \setminus \mathbf{S} = \mathbf{D}_1$ . En appliquant les théorèmes (2.25) et (2.20) à la famille d'opérateurs holomorphes de type A  $\mathcal{R}(\mathbf{p}) + \tilde{\mathbf{A}}$ , il apparaît que  $\mathcal{R}(\mathbf{p}) + \tilde{\mathbf{A}}$  constitue une famille holomorphe d'opérateurs fermés à résolvantes compactes pour  $\mathbf{p} \in \mathbf{D}_1$ . Par conséquent, cet opérateur est soit singulier sur  $\mathbf{D}_1$  soit inversible pour tout  $\mathbf{p}$  appartenant à  $\mathbf{D}_1$  hormis un sous-ensemble discret et localement fini (eventuellement vide) de  $\mathbf{D}_1$ . Pour lever cette alternative, on remarque que  $P_\perp \mathbf{K}(\mathbf{p}_0, \cdot) P_\perp$  définit une forme sesquilinéaire coercive dans  $(\mathbf{grad}(H_0^1(\Omega)))^2$  car  $\mathbf{K}(\mathbf{p}_0, \cdot)$  est coercif dans  $L^2(\Omega)^6$ . Cela implique que  $P_\perp \mathbf{K}(\mathbf{p}_0, \cdot) P_\perp$  est inversible. Les opérateurs  $\mathbf{A}$  et  $\mathbf{K}(\mathbf{p}_0, \cdot)$  sont respectivement maximal monotone et monotone borné coercif dans  $L^2(\Omega)^6$ , il suit que l'opérateur  $\mathbf{A} + \mathbf{K}(\mathbf{p}_0, \cdot)$  est inversible. Enfin l'opérateur  $\mathcal{R}(\mathbf{p}_0) + \tilde{\mathbf{A}}$  est inversible dans  $(H(\mathbf{div}_0, \Omega))^2$ . Donc  $\mathcal{R}(\mathbf{p}) + \tilde{\mathbf{A}}$  est inversible pour tout  $p$  appartenant à  $\mathbf{D}_1$  hormis un sous-ensemble discret localement fini et éventuellement vide de  $\mathbf{D}_1$ . Cela montre que l'opérateur  $\mathbf{K}(\mathbf{p}, \cdot) + \mathbf{A}$  est inversible dans  $L^2(\Omega)^6$  pour tout  $\mathbf{p}$  appartenant à  $\mathbf{D}_1$  hormis un sous-ensemble discret localement fini (eventuellement vide) de  $\mathbf{D}_1$ .  $\square$

### Etude de la coercivité de la matrice PML

D'après ce qui précède, pour avoir l'existence et l'unicité de la solution, il faut donc vérifier qu'il existe un  $\mathbf{p}_0$  appartenant à  $\mathbb{C}$  tel que  $\mathbf{K}(\mathbf{p}_0, \mathbf{x})$  soit coercive. Si on exhibe un tel point, alors la résolvante est non vide et l'alternative est levée.

#### Propriété 2.28

La matrice  $\mathbf{K}(1, \mathbf{x})$  est symétrique définie positive pour tout  $x$  dans  $\mathbb{R}^3$ .

**Démonstration** : comme  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  est définie par la relation (2.20), on a

$$\begin{aligned} \mathbf{K}(\mathbf{p}, \mathbf{x})^T &= \left( \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) \mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} (\mathbf{J}(\mathbf{p}, \mathbf{x})^T)^{-1} \right)^T \\ &= \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) \mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} (\mathbf{J}(\mathbf{p}, \mathbf{x})^T)^{-1} \\ &= \mathbf{K}(\mathbf{p}, \mathbf{x}). \end{aligned}$$

Donc la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  est symétrique pour tout  $\mathbf{p} \in \mathbb{C}$ , donc a fortiori  $\mathbf{K}(1, \mathbf{x})$  est symétrique. Soit  $\mathbf{X} \in \mathbb{R}^3$  tel que  $\mathbf{X} \neq \mathbf{0}$ . Alors,

$$\begin{aligned} \mathbf{X}^T \mathbf{K}(1, \mathbf{x}) \mathbf{X} &= \mathbf{X}^T \left( \det(\mathbf{J}(1, \mathbf{x})) \mathbf{J}(1, \mathbf{x})^{-1} (\mathbf{J}(1, \mathbf{x})^T)^{-1} \right) \mathbf{X} \\ &= \det(\mathbf{J}(1, \mathbf{x})) \mathbf{X}^T \mathbf{J}(1, \mathbf{x})^{-1} (\mathbf{J}(1, \mathbf{x})^T)^{-1} \mathbf{X} \\ &= \det(\mathbf{J}(1, \mathbf{x})) \mathbf{X}^T \mathbf{J}(1, \mathbf{x})^{-1} (\mathbf{X}^T \mathbf{J}(1, \mathbf{x})^{-1})^T. \end{aligned}$$

Posons  $\mathbf{A} = \mathbf{X}^T \mathbf{J}(1, \mathbf{x})^{-1}$ . Alors  $\mathbf{X}^T \mathbf{K}(1, \mathbf{x}) \mathbf{X} = \det(\mathbf{J}(1, \mathbf{x})) \mathbf{A} \mathbf{A}^T$ .

Comme  $\mathbf{A} \mathbf{A}^T$  est définie positive pour toute matrice  $\mathbf{A}$ , on en déduit que  $\mathbf{K}(1, \mathbf{x})$  est définie positive si  $\det(\mathbf{J}(1, \mathbf{x})) > 0$ .

Or  $\mathbf{J}(1, \mathbf{x}) = \frac{D(\tilde{\mathbf{x}})}{D(\varphi(\mathbf{x}))} \frac{D(\varphi(\mathbf{x}))}{D(\mathbf{x})}$ . Soit  $\mathbf{F}(\mathbf{x}) = \frac{D(\varphi(\mathbf{x}))}{D(\mathbf{x})}$  la matrice du jacobien de  $\varphi$  et  $\mathbf{F}_0(\mathbf{x}) = \frac{D(\tilde{\mathbf{x}})}{D(\varphi(\mathbf{x}))}$ . Alors  $\det(\mathbf{J}(1, \mathbf{x})) = \det(\mathbf{F}(\mathbf{x})) \cdot \det(\mathbf{I}_3 + \mathbf{F}_0(\mathbf{x}))$ . La stricte positivité du déterminant sera garantie si les deux termes du produit sont strictement positifs. Comme  $\varphi$  est une fonction croissante qui représente le plongement dans une variété  $\mathbf{M}$  de  $\mathbb{C}^3$ , on a  $\det(\mathbf{F}(\mathbf{x})) > 0$ . Reste donc à montrer que  $\det(\mathbf{I}_3 + \mathbf{F}_0(\mathbf{x})) > 0$ . Soit  $(i, j) \in \llbracket 1, n \rrbracket$ ,  $\mathbf{F}_0(\mathbf{x})_{ij} = \delta_{ij} + f'(\varphi(\mathbf{x}))$ . La fonction  $f$  est strictement croissante, ces termes sont donc strictement positifs. La matrice  $\mathbf{K}(\mathbf{p}_0, \mathbf{x})$  de la formulation PML pour la valeur  $\mathbf{p}_0 = 1$  est donc bien symétrique définie positive.  $\square$

En exhibant une valeur de  $\mathbf{p}$  qui appartient au domaine d'holomorphie telle que la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  est coercive, on lève l'alternative de Fredholm et via le théorème (2.27), on garantit le caractère bien posé des équations de Maxwell écrites dans le milieu PML.

## 2.5 Formulations PML harmoniques

### 2.5.1 Une première formulation des domaines étoilés

La difficulté consiste donc à choisir correctement les fonctions  $\varphi$  et  $f$  en jeu. En effet, certains choix inspirés des formulations classiques peuvent générer des singularités sur la fonction de Hankel (cf. section (2.8)).

L'idée de la formulation suivante est de garder une couche de vide entre le domaine  $\Omega$  et les PML, dans laquelle il n'y aura pas d'atténuation des ondes. L'épaisseur de la couche peut être aussi petite que l'on souhaite, mais la condition CFL risque de se dégrader d'autant plus fortement que l'épaisseur tendra vers zéro. Une étude de stabilité est donc à envisager pour les résolutions en instationnaires.

Soit  $R = \max_{\theta} \rho_0(\theta)$ ,  $r = \min_{\theta} \rho_0(\theta)$  et  $k = \frac{R-r}{a}$ , avec  $a > 0$ . Le coefficient  $k$  déterminera l'épaisseur de la couche sans décroissance. Soit  $\rho_0$  la frontière du domaine  $\Omega$ . Pour s'adapter au domaine étoilé, l'immersion et le changement de variables seront donnés en coordonnées polaires, et laisseront  $\theta$  invariant.

$$\begin{cases} \widehat{\rho}(\mathbf{x}) = k \rho + (1 - k) \rho_0(\theta), \\ \widetilde{\rho}(\mathbf{x}) = \widehat{\rho}(\mathbf{x}) + \frac{1}{\mathbf{p}} f(\widehat{\rho}(\mathbf{x}) - R). \end{cases}$$

**Proposition 2.29**

La matrice harmonique associée à cette formulation est donnée par :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{Q} \begin{pmatrix} \frac{b \mathbf{p}(\mathbf{p} + \beta)}{k} + \frac{a^2 \mathbf{p}(\mathbf{p} + \alpha)}{k b} & -\frac{a \mathbf{p} + \alpha}{b} & 0 \\ \frac{a \mathbf{p} + \alpha}{b} & \frac{k \mathbf{p}(\mathbf{p} + \alpha)}{b} & 0 \\ 0 & 0 & k b (\mathbf{p} + \alpha)(\mathbf{p} + \beta) \end{pmatrix} \mathbf{Q}^T,$$

où  $\alpha = f'(\widehat{\rho}(\mathbf{x}) - R)$ ,  $\beta = \frac{f(\widehat{\rho}(\mathbf{x}) - R)}{\widehat{\rho}(\mathbf{x})}$ ,  $a = 2(1 - k) \frac{\rho'_0(\theta)}{\rho}$ ,  $b = \frac{\widehat{\rho}(\mathbf{x})}{\rho}$  et  $\mathbf{Q}$  la matrice orthogonale de rotation.

**Démonstration** : soit  $\mathbf{J}(\mathbf{p}, \mathbf{x})$  la matrice du jacobien de la transformation (2.24), alors  $\mathbf{J}(\mathbf{p}, \mathbf{x}) = \frac{\mathbf{D}(\tilde{x}, \tilde{y}, \tilde{z})}{\mathbf{D}(x, y, z)} = \frac{\mathbf{D}(\tilde{x}, \tilde{y}, \tilde{z})}{\mathbf{D}(\tilde{\rho}, \theta, z)} \frac{\mathbf{D}(\tilde{\rho}, \theta, z)}{\mathbf{D}(\widehat{\rho}, \theta, z)} \frac{\mathbf{D}(\widehat{\rho}, \theta, z)}{\mathbf{D}(\rho, \theta, z)} \frac{\mathbf{D}(\rho, \theta, z)}{\mathbf{D}(x, y, z)}$ .

$$\begin{aligned} \frac{\mathbf{D}(\tilde{\rho}, \theta, z)}{\mathbf{D}(\widehat{\rho}, \theta, z)} &= \begin{pmatrix} 1 + \frac{1}{\mathbf{p}} f'(\widehat{\rho}(\mathbf{x}) - R) & (1 - k) \rho'_0(\theta) & \left(1 + \frac{1}{\mathbf{p}} f'(\widehat{\rho}(\mathbf{x}) - R)\right) & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ \frac{\mathbf{D}(\widehat{\rho}, \theta, z)}{\mathbf{D}(\rho, \theta, z)} &= \begin{pmatrix} k & (1 - k) \rho'_0(\theta) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \frac{\mathbf{D}(\tilde{x}, \tilde{y}, \tilde{z})}{\mathbf{D}(\tilde{\rho}, \theta, z)} &= \begin{pmatrix} \cos\theta & -\tilde{\rho} \sin\theta & 0 \\ \sin\theta & \tilde{\rho} \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \tilde{\rho} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \frac{\mathbf{D}(\rho, \theta, z)}{\mathbf{D}(x, y, z)} &= \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ \rho & \rho & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\rho} & 0 \\ 0 & \rho & 1 \end{pmatrix}. \end{aligned}$$

Soit  $\mathbf{Q} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$  la matrice orthogonale de rotation. Afin d'alléger

les notations, notons  $\alpha = f'(\widehat{\rho}(\mathbf{x}) - R)$ ,  $\beta = \frac{f(\widehat{\rho}(\mathbf{x}) - R)}{\widehat{\rho}(\mathbf{x})}$ ,  $a = 2(1 - k) \frac{\rho'_0(\theta)}{\rho}$  et  $b = \frac{\widehat{\rho}(\mathbf{x})}{\rho}$ .

$$\begin{aligned}
\mathbf{J}(\mathbf{p}, \mathbf{x}) &= \mathbf{Q} \begin{pmatrix} 1 + \frac{1}{\mathbf{p}}\alpha & 0 & 0 \\ 0 & 1 + \frac{1}{\mathbf{p}}\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} k & a & 0 \\ 0 & b & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{Q}^T \\
\det(\mathbf{J}(\mathbf{p}, \mathbf{x})) &= kb \left(1 + \frac{1}{\mathbf{p}}\alpha\right) \left(1 + \frac{1}{\mathbf{p}}\beta\right) \\
\mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} &= \mathbf{Q} \begin{pmatrix} \frac{1}{k} & 0 & 0 \\ -a & \frac{1}{b} & 0 \\ \frac{k}{b} & \frac{1}{b} & 0 \end{pmatrix} \begin{pmatrix} \frac{\mathbf{p}}{\mathbf{p} + \beta} & 0 & 0 \\ 0 & \frac{\mathbf{p}}{\mathbf{p} + \alpha} & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{Q}^T.
\end{aligned}$$

L'équation (2.20) exprime la matrice de la formulation PML pour les équations de Maxwell sur  $(\mathbf{M}, g)$ , à savoir  $\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{p} \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) \mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} (\mathbf{J}(\mathbf{p}, \mathbf{x})^T)^{-1}$ . Le résultat proposé est alors obtenu directement.  $\square$

**Proposition 2.30**

*Le changement de variable (2.24) est injectif et la décroissance exponentielle de la solution est assurée dans la totalité des couches PML.*

**Démonstration** : nous avons déjà vérifié cette condition dans le cas général, moyennant des hypothèses sur  $\varphi$  et  $f$ . Néanmoins, nous proposons une démonstration explicite dans ce cas ci. Le noyau de Green s'écrivant  $\frac{\exp(-i\omega \{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\})}{4\pi \{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\}}$ , posons  $\tilde{d} = \{\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\}$  la racine carrée qui nous intéresse. Elle ne doit jamais s'annuler et rester à partie imaginaire strictement positive.

Soit  $\mathbf{x}_1 \in \Omega$  et  $\mathbf{x}_2$  un point du domaine PML. Alors  $\Delta\tilde{x} = \tilde{x}(\mathbf{x}_1) - \tilde{x}(\mathbf{x}_2)$ . Le changement de variable complexe est donné par :

$$\begin{cases} \hat{\rho}(\mathbf{x}) = k\rho + (1-k)\rho_0(\theta) \\ \tilde{\rho}(\mathbf{x}) = \hat{\rho}(\mathbf{x}) + \frac{1}{\mathbf{p}}f(\hat{\rho}(\mathbf{x}) - R), \end{cases}$$

$$\begin{aligned}
\Delta\tilde{x}^2 &= \tilde{\rho}_1^2 \cos^2(\theta_1) + \tilde{\rho}_2^2 \cos^2(\theta_2) - 2\tilde{\rho}_1\tilde{\rho}_2 \cos(\theta_1) \cos(\theta_2) \\
\Delta\tilde{y}^2 &= \tilde{\rho}_1^2 \sin^2(\theta_1) + \tilde{\rho}_2^2 \sin^2(\theta_2) - 2\tilde{\rho}_1\tilde{\rho}_2 \sin(\theta_1) \sin(\theta_2) \\
\tilde{d}^2 &= \tilde{\rho}_1^2 + \tilde{\rho}_2^2 - 2\tilde{\rho}_1\tilde{\rho}_2 \cos(\theta_1 - \theta_2).
\end{aligned}$$

On cherche à calculer les parties réelle et imaginaire de ce terme et à vérifier qu'elles ne s'annulent pas simultanément. Notons  $\Delta\theta = \theta_1 - \theta_2$ . Comme  $\mathbf{x}_1 \in \Omega$ , on a  $\tilde{\rho}_1 = \rho_1$ . A l'inverse,  $\tilde{\rho}_2 = \hat{\rho}_2 + \frac{1}{\mathbf{p}}f(\hat{\rho}_2 - R)$ . En considérant le cas limite  $\mathbf{p} = i\omega$ , on a :

$$\text{Im}(\tilde{d}^2) = \frac{2}{\omega} (\rho_1 \cos(\Delta\theta) - \hat{\rho}_2) f(\hat{\rho}_2 - R).$$

Comme  $f(\widehat{\rho}_2 - R) \geq 0$  et  $\omega > 0$ , la partie imaginaire s'annule quand  $\rho_1 \cos(\Delta\theta) = \widehat{\rho}_2$  et de  $f(\widehat{\rho}_2 - R) = 0$ .

**Premier cas :**  $f(\widehat{\rho}_2 - R) = 0$  donc  $\widehat{\rho}_2 = R$ .

$$\begin{aligned} \mathbf{Re}(\widetilde{d}^2) &= \rho_1^2 + \widehat{\rho}_2^2 - 2\rho_1\widehat{\rho}_2 \cos(\Delta\theta) - \frac{1}{\omega^2} f^2(\widehat{\rho}_2 - R) \\ &= \rho_1^2 + R^2 - 2R\rho_1 \cos(\Delta\theta) \\ &\geq \rho_1^2 + R^2 - 2R\rho_1 \\ &\geq (\rho_1 - R)^2 \geq 0. \end{aligned}$$

Comme  $\mathbf{x}_1 \in \Omega$ ,  $\rho_1 < \rho_0 \leq R$  donc la partie réelle ne s'annule jamais dans cette configuration.

**Deuxième cas :**  $\rho_1 \cos(\Delta\theta) = \widehat{\rho}_2$ .

$$\begin{aligned} \mathbf{Re}(\widetilde{d}^2) &= \rho_1^2 + \widehat{\rho}_2^2 - 2\rho_1\widehat{\rho}_2 \cos(\Delta\theta) - \frac{1}{\omega^2} f^2(\widehat{\rho}_2 - R) \\ &= \rho_1^2 + (\rho_1 \cos(\Delta\theta))^2 - 2(\rho_1 \cos(\Delta\theta))^2 - \frac{1}{\omega^2} f^2(\rho_1 \cos(\Delta\theta) - R) \\ &= \rho_1^2 - (\rho_1 \cos(\Delta\theta))^2 \\ &= \rho_1^2 (1 - \cos^2(\Delta\theta)) \geq 0. \end{aligned}$$

La partie réelle ne peut donc s'annuler que si  $\cos^2(\Delta\theta) = 1$  donc si  $\theta_1 = \theta_2[2\pi]$ . Mais alors, comme  $\rho_1 \cos(\Delta\theta) = \widehat{\rho}_2$ , on aurait  $\rho_1 = k\rho_2 + (1-k)\rho_0$  ce qui est impossible. La partie réelle ne s'annule jamais dans cette configuration, donc dans tous les cas, l'injectivité et la décroissance exponentielle du milieu sont garanties.  $\boxtimes$

## 2.5.2 Formulation hybride pour un domaine de forme $L$

On propose une décomposition d'un domaine de forme  $L$  en une partie convexe et une partie non convexe décrite par la figure (2.3). Ceci semble particulièrement intéressant afin de pouvoir utiliser des PML cartésiennes dans la partie convexe. L'immersion et le difféomorphisme introduits précédemment ne tiennent pas compte de ce fait. Cependant, ce procédé semble a priori plus complexe vu qu'il faut être capable de passer continûment d'une formulation à l'autre. Comme pour les PML précédentes, l'idée est de garder une couche de vide entre le domaine  $\Omega$  et les PML.

Supposons que le domaine  $\Omega$  est symétrique. Soit  $\varepsilon$  l'épaisseur des couches PML. Notons  $\mathbf{D}_c$  la partie convexe du domaine  $\Omega$  et  $\mathbf{D}_{nc} = \Omega \setminus \mathbf{D}_c$ . Il vient que :

$$\begin{aligned} \mathbf{D}_c &= ([-r - \varepsilon, R + \varepsilon] \times [-r - \varepsilon, r]) \cup ([-r - \varepsilon, r] \times [-r - \varepsilon, R + \varepsilon]) \\ &\quad \cup ([-r - \varepsilon, r] \times [-r - \varepsilon, R]) \cup ([-r - \varepsilon, R] \times [-r - \varepsilon, r]). \end{aligned}$$

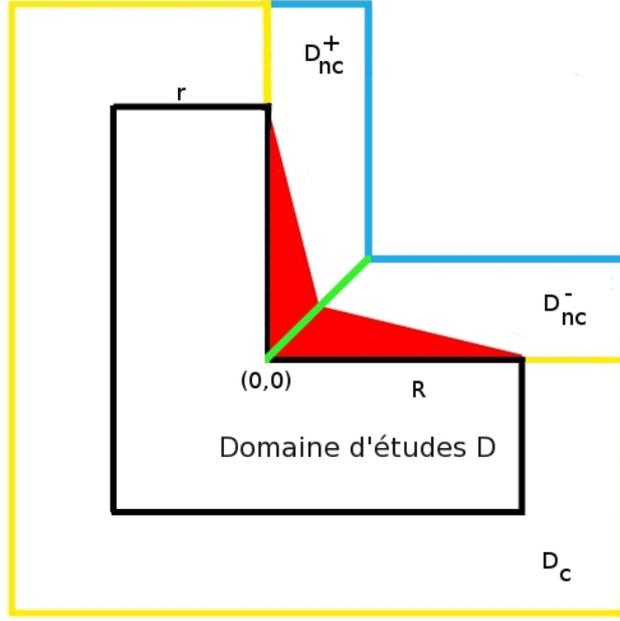


FIG. 2.3: Décomposition particulière d'un domaine de forme L

La zone non convexe peut être décomposée en deux parties, dépendant du signe de  $y - x$  :  $\mathbf{D}_{nc}^+ = \mathbf{D}_{nc} \cap \{y \geq x\}$  et  $\mathbf{D}_{nc}^- = \mathbf{D}_{nc} \cap \{x \geq y\}$ .

Le changement de variables et le difféomorphisme doivent garantir la continuité de la transition entre des PML convexes et non convexes. Soit  $\kappa > 0$  le coefficient déterminant la zone sans décroissance sur les PML non convexes. Posons

$$\begin{cases} \hat{x} = x + (\kappa - 1) \inf(x, y) \\ \hat{y} = y + (\kappa - 1) \inf(x, y) \end{cases} \quad \begin{cases} \tilde{x} = \hat{x} + \frac{1}{\mathbf{p}} \int \sigma_x(\tau) d\tau \\ \tilde{y} = \hat{y} + \frac{1}{\mathbf{p}} \int \sigma_y(\tau) d\tau \end{cases} \quad (2.24)$$

avec  $\sigma_x, \sigma_y : \mathbb{R} \rightarrow \mathbb{R}^+$  des fonctions au moins  $C^1$  telles que :

$$\begin{cases} \sigma_x(x, y) = 0 & \text{si } x \in [-r, R], \\ \sigma_y(x, y) = 0 & \text{si } y \in [-r, R]. \end{cases}$$

**Proposition 2.31**

La matrice harmonique associée à cette formulation est donnée par :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{T}(\mathbf{x}) \begin{pmatrix} \frac{\mathbf{p}(\mathbf{p} + \sigma_y(\mathbf{x}))}{\mathbf{p} + \sigma_x(\mathbf{x})} & 0 & 0 \\ 0 & \frac{\mathbf{p}(\mathbf{p} + \sigma_x(\mathbf{x}))}{\mathbf{p} + \sigma_y(\mathbf{x})} & 0 \\ 0 & 0 & \frac{(\mathbf{p} + \sigma_x(\mathbf{x}))(\mathbf{p} + \sigma_y(\mathbf{x}))}{\mathbf{p}} \end{pmatrix} \mathbf{T}(\mathbf{x})^T,$$

où la matrice triangulaire  $\mathbf{T}(\mathbf{x})$  dépend de la partie du domaine sur laquelle on la calcule.

**Démonstration** : soit  $\mathbf{J}(\mathbf{p}, \mathbf{x})$  la matrice du jacobien de la transformation (2.24).

$$\mathbf{J}(\mathbf{p}, \mathbf{x}) = \frac{\mathbf{D}(\tilde{x}, \tilde{y}, \tilde{z})}{\mathbf{D}(\hat{x}, \hat{y}, \hat{z})} \frac{\mathbf{D}(\hat{x}, \hat{y}, \hat{z})}{\mathbf{D}(x, y, z)}.$$

Le difféomorphisme (2.24) est différent selon la position par rapport à la première bissectrice du coin rentrant sur  $\mathbf{D}_{nc}$ . Soit  $\mathbf{T}_0$  la matrice triangulaire associée au changement de variable chapeau. Elle dépend de la partie du domaine sur laquelle on

la calcule : sur  $\mathbf{D}_{nc}^+$ ,  $\mathbf{T}_0^+(\mathbf{x}) = \begin{pmatrix} 1 & \kappa - 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  et sur  $\mathbf{D}_{nc}^-$ ,  $\mathbf{T}_0^-(\mathbf{x}) = \begin{pmatrix} \kappa & 0 & 0 \\ \kappa - 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ .

On a alors :

$$\begin{aligned} \mathbf{J}(\mathbf{p}, \mathbf{x}) &= \begin{pmatrix} 1 + \frac{1}{\mathbf{p}} \sigma_x(\mathbf{x}) & 0 & 0 \\ 0 & 1 + \frac{1}{\mathbf{p}} \sigma_y(\mathbf{x}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{T}_0(\mathbf{x}) \\ \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) &= \kappa \left( 1 + \frac{1}{\mathbf{p}} \sigma_x(\mathbf{x}) \right) \left( 1 + \frac{1}{\mathbf{p}} \sigma_y(\mathbf{x}) \right) \\ \mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} &= \mathbf{T}_0^{-1}(\mathbf{x}) \begin{pmatrix} \frac{\mathbf{p}}{\mathbf{p} + \sigma_x(\mathbf{x})} & 0 & 0 \\ 0 & \frac{\mathbf{p}}{\mathbf{p} + \sigma_y(\mathbf{x})} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ (\mathbf{J}(\mathbf{p}, \mathbf{x})^{-1})^T &= \begin{pmatrix} \frac{\mathbf{p}}{\mathbf{p} + \sigma_x(\mathbf{x})} & 0 & 0 \\ 0 & \frac{\mathbf{p}}{\mathbf{p} + \sigma_y(\mathbf{x})} & 0 \\ 0 & 0 & 1 \end{pmatrix} (\mathbf{T}_0^{-1}(\mathbf{x}))^T \end{aligned}$$

Posons  $\mathbf{T}(\mathbf{x}) = (\mathbf{T}_0(\mathbf{x}))^{-1}$ , c'est-à-dire sur  $\mathbf{D}_{nc}^+$ ,  $\mathbf{T}^+(\mathbf{x}) = \frac{1}{\kappa} \begin{pmatrix} 1 & 1 - \kappa & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa \end{pmatrix}$  et sur

$\mathbf{D}_{nc}^-$ ,  $\mathbf{T}^-(\mathbf{x}) = \frac{1}{\kappa} \begin{pmatrix} \kappa & 0 & 0 \\ 1 - \kappa & 1 & 0 \\ 0 & 0 & \kappa \end{pmatrix}$ . Le résultat proposé est alors obtenu directement,

étant donné que la matrice PML pour les équations de Maxwell sur la variété  $(\mathbf{M}, g)$  est donnée  $\mathbf{K}(\mathbf{p}, \mathbf{x}) = \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) \mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} (\mathbf{J}(\mathbf{p}, \mathbf{x})^T)^{-1}$ .  $\square$

Pour le domaine PML non convexe, il est intéressant de connaître exactement la frontière entre les zones avec et sans décroissance. Les lignes de transmission sans décroissance correspondent aux endroits où la partie imaginaire du changement de variables (2.24) s'annule. Soit  $f$  une fonction réelle au moins  $C^1$  telle que  $f$  soit positive croissante. Alors  $\sigma_x(x, y) = f'(\hat{x} - R)$  et  $\sigma_y(x, y) = f'(\hat{y} - R)$ . Comme la partie imaginaire est donnée par les fonctions  $\sigma$ , la limite entre la zone avec décroissance et la zone sans décroissance est, sur  $\mathbf{D}_{nc}^-$ ,  $\hat{x} = R$  et sur  $\mathbf{D}_{nc}^+$ ,  $\hat{y} = R$ .

D'où :

$$\text{sur } \mathbf{D}_{nc}^-, \quad y = \frac{1}{1-\kappa}(x-R) \quad \text{sur } \mathbf{D}_{nc}^+, \quad y = (1-\kappa)x + R.$$

Le point  $\left(\frac{R}{\kappa}, \frac{R}{\kappa}\right)$  donne donc l'épaisseur de la zone PML non convexe sans décroissance.

**Proposition 2.32**

*Le changement de variable (2.24) est injectif et la décroissance exponentielle de la solution est assurée dans la totalité des couches PML.*

**Démonstration** : soit  $\mathbf{x}_1 \in \Omega$  et  $\mathbf{x}_2$  un point du domaine PML. Le changement de variable complexe est donné par (2.24).

$$\begin{aligned} \Delta \tilde{x}^2 &= (\tilde{x}(\mathbf{x}_1) - \tilde{x}(\mathbf{x}_2))^2 = \tilde{x}_1^2 + \tilde{x}_2^2 - 2\tilde{x}_1\tilde{x}_2 \\ \Delta \tilde{y}^2 &= (\tilde{y}(\mathbf{x}_1) - \tilde{y}(\mathbf{x}_2))^2 = \tilde{y}_1^2 + \tilde{y}_2^2 - 2\tilde{y}_1\tilde{y}_2 \end{aligned}$$

D'où  $\tilde{d}^2 = \tilde{x}_1^2 + \tilde{y}_1^2 + \tilde{x}_2^2 + \tilde{y}_2^2 - 2(\tilde{x}_1\tilde{x}_2 + \tilde{y}_1\tilde{y}_2)$ .

Comme  $\mathbf{x}_1 \in \Omega$ , on a  $\tilde{x}_1 = x_1$  et  $\tilde{y}_1 = y_1$ . A l'inverse,  $\tilde{x}_2 = \hat{x}_2 + \frac{1}{\mathbf{p}}f(\hat{x}_2 - R)$  et  $\tilde{y}_2 = \hat{y}_2 + \frac{1}{\mathbf{p}}f(\hat{y}_2 - R)$ . En considérant le cas limite  $\mathbf{p} = i\omega$ , on a :

$$\mathbf{Im}(\tilde{d}^2) = \frac{2}{\omega} ((x_1 - \hat{x}_2)f(\hat{x}_2 - R) + (y_1 - \hat{y}_2)f(\hat{y}_2 - R)).$$

$\mathbf{x}_1 \in \Omega$  donc  $f(x_1 - R) = 0$  et  $f(y_1 - R) = 0$ . De plus, la fonction  $f$  a été choisie telle que  $\forall(\mathbf{t}, \mathbf{z}), \sum_i(\mathbf{t} - \mathbf{z})_i (f(\mathbf{t}_i - R) - f(\mathbf{z}_i - R)) \geq 0$ . Donc  $\mathbf{Im}(\tilde{d}^2)$  égale :

$$-\frac{2}{\omega} [(x_1 - \hat{x}_2)(f(x_1 - R) - f(\hat{x}_2 - R)) + (y_1 - \hat{y}_2)(f(y_1 - R) - f(\hat{y}_2 - R))] \leq 0.$$

La partie imaginaire s'annule donc quand  $f(\hat{x}_2 - R) = f(\hat{y}_2 - R) = 0$ , ce qui est équivalent à la condition  $\hat{x}_2 = \hat{y}_2 = R$ .

$$\begin{aligned} \mathbf{Re}(\tilde{d}^2) &= x_1^2 + y_1^2 + \hat{x}_2^2 + \hat{y}_2^2 - 2(x_1\hat{x}_2 + y_1\hat{y}_2) - \frac{1}{\omega^2}(f(\hat{x}_2 - R) + f(\hat{y}_2 - R)) \\ &= x_1^2 + y_1^2 + R^2 + R^2 - 2Rx_1 - 2Ry_1 = (x_1 - R)^2 + (y_1 - R)^2 \geq 0. \end{aligned}$$

La partie réelle ne peut s'annuler qu'en  $\mathbf{x}_1 = (R, R)$ . Or ce point n'appartient pas au domaine  $\Omega$ , donc la partie réelle et la partie imaginaire ne peuvent s'annuler simultanément, donc l'injectivité et la décroissance des PML sont assurées.

⊠

### 2.5.3 Amélioration de cette formulation

On souhaite réduire au maximum la zone sans décroissance. La deuxième partie du changement de variables (2.24) est dissymétrique en  $\hat{x}$  et  $\hat{y}$ . En introduisant une fonction symétrique, d'argument  $\hat{x} + \hat{y} - R$ , nous aurons égalité des dérivées partielles  $\frac{\partial f(\hat{x} + \hat{y} - R)}{\partial \hat{x}} = \frac{\partial f(\hat{x} + \hat{y} - R)}{\partial \hat{y}} = f'(\hat{x} + \hat{y} - R)$ . Le changement de variables considéré est

$$\begin{cases} \tilde{x} = \hat{x} + \frac{1}{\mathbf{p}} f(\hat{x} + \hat{y} - R) \\ \tilde{y} = \hat{y} + \frac{1}{\mathbf{p}} f(\hat{x} + \hat{y} - R). \end{cases} \quad (2.25)$$

**Proposition 2.33**

La matrice harmonique associée à cette formulation est

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{T}(\mathbf{x}) \begin{pmatrix} \frac{(\mathbf{p} + \sigma(\mathbf{x}))^2 + \sigma(\mathbf{x})^2}{\mathbf{p} + 2\sigma(\mathbf{x})} & -2\sigma(\mathbf{x}) \frac{\mathbf{p} + \sigma(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})} & 0 \\ -2\sigma(\mathbf{x}) \frac{\mathbf{p} + \sigma(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})} & \frac{(\mathbf{p} + \sigma(\mathbf{x}))^2 + \sigma(\mathbf{x})^2}{\mathbf{p} + 2\sigma(\mathbf{x})} & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{T}^T(\mathbf{x}),$$

où  $\sigma$  désigne la dérivée de la fonction  $f$ , et la matrice triangulaire  $\mathbf{T}(\mathbf{x})$  est la même que pour la formulation précédente.

**Démonstration** : elle se conduit comme précédemment. Une seule matrice est modifiée : le jacobien de la transformation tilde. Posons  $\sigma(\mathbf{x}) = f'(\hat{x} + \hat{y} - R)$ . Alors

$$\frac{\mathbf{D}(\tilde{x}, \tilde{y}, \tilde{z})}{\mathbf{D}(\hat{x}, \hat{y}, \hat{z})} = \mathbf{I}_3 + \frac{1}{\mathbf{p}} \begin{pmatrix} \sigma(\mathbf{x}) & \sigma(\mathbf{x}) & 0 \\ \sigma(\mathbf{x}) & \sigma(\mathbf{x}) & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\mathbf{J}(\mathbf{p}, \mathbf{x}) = \begin{pmatrix} 1 + \frac{1}{\mathbf{p}} \sigma(\mathbf{x}) & \frac{1}{\mathbf{p}} \sigma(\mathbf{x}) & 0 \\ \frac{1}{\mathbf{p}} \sigma(\mathbf{x}) & 1 + \frac{1}{\mathbf{p}} \sigma(\mathbf{x}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{T}_0(\mathbf{x})$$

$$\det(\mathbf{J}(\mathbf{p}, \mathbf{x})) = \kappa \left( 1 + \frac{2}{\mathbf{p}} \sigma(\mathbf{x}) \right)$$

$$\mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} = \mathbf{T}_0^{-1}(\mathbf{x}) \begin{pmatrix} \frac{\mathbf{p} + \sigma(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})} & \frac{-\sigma(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})} & 0 \\ \frac{-\sigma(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})} & \frac{\mathbf{p} + \sigma(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Comme  $\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{p} \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) \mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} (\mathbf{J}(\mathbf{p}, \mathbf{x})^T)^{-1}$ , on obtient le résultat annoncé.  $\square$

Dans ce cas, la partie imaginaire s'annule quand  $\sigma(x, y) = f'(\hat{x} + \hat{y} - R) = 0$ , donc quand  $\hat{x} + \hat{y} = R$ . Pour  $x = y$ , on obtient alors que le point  $\left(\frac{R}{2\kappa}, \frac{R}{2\kappa}\right)$  donne l'épaisseur de la zone PML non convexe sans décroissance. On a réduit de moitié cette zone par rapport à la formulation précédente.

**Proposition 2.34**

Le changement de variable (2.25) est injectif et la décroissance exponentielle de la solution est assurée dans la totalité des couches PML.

**Démonstration** : le changement de variable complexe est désormais donné par :

$$\begin{cases} \tilde{x} = x + (\kappa - 1) \mathbf{inf}(x, y) \\ \tilde{y} = y + (\kappa - 1) \mathbf{inf}(x, y) \end{cases} \quad \begin{cases} \tilde{x} = \hat{x} + \frac{1}{\mathbf{p}} f(\hat{x} + \hat{y} - R) \\ \tilde{y} = \hat{y} + \frac{1}{\mathbf{p}} f(\hat{x} + \hat{y} - R). \end{cases}$$

$$\begin{aligned} \tilde{d}^2 &= \tilde{x}_1^2 + \tilde{y}_1^2 + \tilde{x}_2^2 + \tilde{y}_2^2 - 2(\tilde{x}_1\tilde{x}_2 + \tilde{y}_1\tilde{y}_2) \\ \mathbf{Im}(\tilde{d}^2) &= \frac{2}{\omega} ((x_1 + \hat{x}_2) + (y_1 + \hat{y}_2)) f(\hat{x}_2 + \hat{y}_2 - R) \\ &= -\frac{2}{\omega} ((x_1 + y_1) - (\hat{x}_2 + \hat{y}_2)) (f(x_1 + y_1 - R) - f(\hat{x}_2 + \hat{y}_2 - R)) \leq 0. \end{aligned}$$

La partie imaginaire s'annule donc quand  $f(\hat{x}_2 + \hat{y}_2 - R) = 0$ . Cette condition correspond à la relation  $\hat{x}_2 + \hat{y}_2 = R$ .

$$\begin{aligned} \mathbf{Re}(\tilde{d}^2) &= x_1^2 + y_1^2 + \hat{x}_2^2 + \hat{y}_2^2 - 2(x_1\hat{x}_2 + y_1\hat{y}_2) - \frac{1}{\omega^2} (f(\hat{x}_2 - R) + f(\hat{y}_2 - R)) \\ &= x_1^2 + \hat{x}_2^2 - 2x_1\hat{x}_2 + y_1^2 + \hat{y}_2^2 - 2y_1\hat{y}_2 \\ &= (x_1 - \hat{x}_2)^2 + (y_1 - \hat{y}_2)^2 \geq 0. \end{aligned}$$

La partie réelle ne peut s'annuler qu'en  $\mathbf{x}_1 = \hat{\mathbf{x}}_2$ . Or le point  $\hat{\mathbf{x}}_2$  n'appartient pas au domaine  $\Omega$ , donc la partie réelle et la partie imaginaire ne peuvent s'annuler simultanément, donc l'injectivité et la décroissance des PML sont assurées.  $\square$

**Remarque** : le domaine présenté sur la figure (2.3) est en réalité beaucoup trop important. Dans des PML cartésiennes par exemple, la décroissance se fait perpendiculairement aux lignes d'exhaustions du domaine, à savoir horizontalement dans les zones latérales où  $\sigma_y = 0$  et verticalement dans les zones où  $\sigma_x = 0$ . Ici, les lignes d'exhaustions sont parallèles à la ligne rouge définissant la zone PML sans décroissance de la figure (2.3). Cette zone est déterminée par la valeur du paramètre  $\kappa$ . Pour la première formulation, son épaisseur maximale est atteinte pour le point  $(R/\kappa, R/\kappa)$ , et pour la seconde  $(R/2\kappa, R/2\kappa)$ . En reliant ce point aux coins du domaine  $\Omega$ , à savoir  $(0, R)$  et  $(R, 0)$ , on obtient cette ligne de séparation entre la zone

PML avec et sans damping. Comme pour les PML classiques, la décroissance se fait perpendiculairement aux lignes d'exhaustion, qui seront parallèles à la délimitation de la zone PML sans décroissance (cf. ce qui précède pour les isovaleurs nulles de  $\sigma_x, \sigma_y$  pour la première formulation et  $\sigma$  pour la seconde). Il est donc inutile d'utiliser un domaine au delà l'intersection de la ligne parallèle à cette dernière avec les coins délimitant la zone PML cartésiennes (cf. figure (2.4)). Cependant, le paramètre  $\kappa$  étant variable, plutôt que de générer un maillage par valeur de  $\kappa$ , nous avons choisi de compléter cette zone et avoir ainsi la liberté de faire varier  $\kappa$ . Dans les simulations numériques où  $\kappa$  sera choisi grand (autrement dit une fine zone sans décroissance), il y aura une zone inutile d'autant plus importante que l'allongement du L sera important.

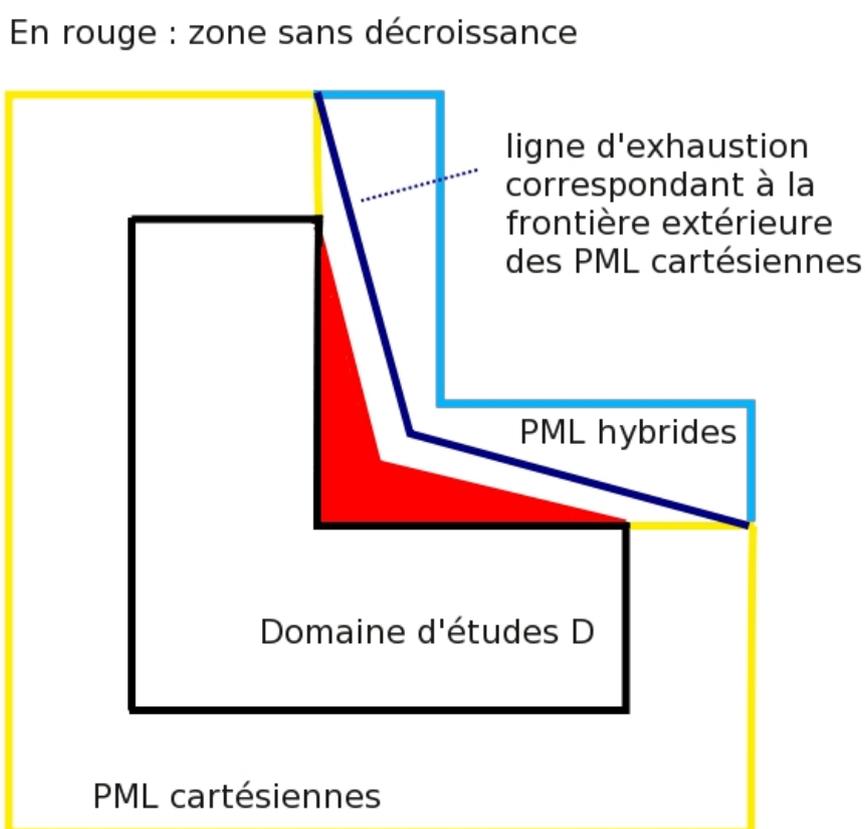


FIG. 2.4: Domaine de forme L et zones de décroissances

**Explications sur la figure (2.4) :** en rouge : la zone sans décroissance des PML non convexes, à savoir la zone dans laquelle les rayons lumineux seront particulièrement tordus et la vitesse de la lumière nettement accélérée. La ligne bleu foncée délimite la zone réelle d'action des PML non convexes. Enfin, est représentée en cyan la zone dite inutile décrite précédemment. Il y a même décroissance dans cette mince couche PML non convexes que dans l'épaisseur  $\varepsilon$  des PML cartésiennes.

## 2.6 Relèvement instationnaire

Nous avons donc établi dans la section précédente que les équations de Maxwell ou des Ondes dans le milieu PML pouvaient s'écrire in fine sous la forme

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) \begin{pmatrix} e \\ h \end{pmatrix} + \begin{pmatrix} 0 & -\nabla \times \\ \nabla \times & 0 \end{pmatrix} \begin{pmatrix} e \\ h \end{pmatrix} = \begin{pmatrix} -j \\ -m \end{pmatrix}$$

où  $(j, m)$  sont des seconds membres réguliers. Sous les hypothèses suivantes

- $\mathbf{K}(\mathbf{p}, \mathbf{x})$   $L^\infty$  en  $\mathbf{x}$
- $\mathbf{K}(\mathbf{p}, \mathbf{x})$  fraction rationnelle en  $\mathbf{p}$  :  $\mathbf{K}_{ij} = \frac{P_{ij}(\mathbf{p}, \mathbf{x})}{Q_{ij}(\mathbf{p}, \mathbf{x})}$  avec  $\mathbf{deg}P \leq \mathbf{deg}Q + 1$
- $\lim_{\mathbf{p} \rightarrow 0} \frac{1}{\mathbf{p}} \mathbf{K}(\mathbf{p}, \mathbf{x}) = \widetilde{\mathbf{A}}_0(\mathbf{x})$  uniformément définie positive,

le problème instationnaire est bien posé (cf chapitre 1) et peut s'écrire sous la forme d'un système de Friedrichs. L'existence et l'unicité de la solution découlent donc naturellement des propriétés de ce type de système.

### 2.6.1 Formulation étoilée

Rappelons l'immersion complexe (2.24)  $\begin{cases} \widehat{\rho}(\mathbf{x}) = k\rho + (1-k)\rho_0(\theta), \\ \widetilde{\rho}(\mathbf{x}) = \widehat{\rho}(\mathbf{x}) + \frac{1}{\mathbf{p}}f(\widehat{\rho}(\mathbf{x}) - R). \end{cases}$

**Proposition 2.35**

La matrice associée à cette formulation peut être décomposée sous la forme :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{Q}\mathbf{P}(\mathbf{x}) \left( \mathbf{p} \mathbf{I}(\mathbf{x}) + \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x}) \left( \mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}) \right)^{-1} \right) \mathbf{P}^T(\mathbf{x}) \mathbf{Q}^T, \quad (2.26)$$

où

$$\mathbf{P}(\mathbf{x}) = \begin{pmatrix} \frac{\widehat{\rho}(\mathbf{x})}{\rho} & -2(1-k)\frac{\rho'_0(\theta)}{\rho} & 0 \\ 0 & k & 0 \\ 0 & 0 & \kappa(\mathbf{x}) \end{pmatrix} \quad \mathbf{I}(\mathbf{x}) = \frac{1}{\kappa(\mathbf{x})} \mathbf{I}_3,$$

$$\mathbf{A}(\mathbf{x}) = \frac{1}{\kappa(\mathbf{x})} \mathbf{diag} \left( \beta(\mathbf{x}) - \alpha(\mathbf{x}), \alpha(\mathbf{x}) - \beta(\mathbf{x}), \alpha(\mathbf{x}) + \beta(\mathbf{x}) \right)$$

$$\mathbf{B}(\mathbf{x}) = \frac{1}{\kappa(\mathbf{x})} \mathbf{diag} \left( \alpha(\mathbf{x})(\alpha(\mathbf{x}) - \beta(\mathbf{x})), \beta(\mathbf{x})(\beta(\mathbf{x}) - \alpha(\mathbf{x})), \alpha(\mathbf{x})\beta(\mathbf{x}) \right)$$

$$\mathbf{F}(\mathbf{x}) = \mathbf{diag} \left( \alpha(\mathbf{x}), \beta(\mathbf{x}), 0 \right).$$

avec  $\kappa(\mathbf{x}) = \left( k \frac{\widehat{\rho}(\mathbf{x})}{\rho(\mathbf{x})} \right)^{-1}$ ,  $\alpha(\mathbf{x}) = f'(\widehat{\rho}(\mathbf{x}) - R)$ ,  $\beta(\mathbf{x}) = \frac{f(\widehat{\rho}(\mathbf{x}) - R)}{\widehat{\rho}(\mathbf{x})}$  et  $\mathbf{Q}$  matrice de rotation .

**Démonstration** : la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  est donnée par (2.29). Pour relever le système différentiel en temps, il faut décomposer cette matrice en éléments simples. Soit  $\alpha(\mathbf{x}) = f'(\widehat{\rho}(\mathbf{x}) - R)$  et  $\beta(\mathbf{x}) = \frac{f(\widehat{\rho}(\mathbf{x}) - R)}{\widehat{\rho}(\mathbf{x})}$ .

$$\begin{aligned} \frac{\mathbf{p}(\mathbf{p} + \alpha(\mathbf{x}))}{\mathbf{p} + \beta(\mathbf{x})} &= \mathbf{p} + \alpha(\mathbf{x}) - \beta(\mathbf{x}) + \frac{\beta(\mathbf{x})(\beta(\mathbf{x}) - \alpha(\mathbf{x}))}{\mathbf{p} + \beta(\mathbf{x})} \\ \frac{\mathbf{p}(\mathbf{p} + \beta(\mathbf{x}))}{\mathbf{p} + \alpha(\mathbf{x})} &= \mathbf{p} + \beta(\mathbf{x}) - \alpha(\mathbf{x}) + \frac{\alpha(\mathbf{x})(\alpha(\mathbf{x}) - \beta(\mathbf{x}))}{\mathbf{p} + \alpha(\mathbf{x})} \end{aligned}$$

Posons  $\kappa(\mathbf{x}) = \left(k \frac{\widehat{\rho}(\mathbf{x})}{\rho(\mathbf{x})}\right)^{-1}$ . La matrice de la formulation PML peut se décomposer en éléments simples sous la forme énoncée.  $\square$

Une telle formulation modifie la matrice devant l'opérateur temporel, ce qui affecte directement les valeurs des conditions CFL que nous nous proposons de déterminer explicitement. On rappelle que le cône d'évolution  $\Gamma_C$  associé à l'Equation aux Dérivées Partielles

$$\widetilde{\mathbf{A}}_0 \partial_t \phi + \sum_i \mathbf{A}_i \partial_i \phi = f,$$

est l'ensemble des  $\lambda \mathbf{n}$  avec  $\lambda \geq 0$  et  $\mathbf{n} \in \mathcal{S}^n$  tels que :

$$\forall \xi \in \mathbb{R}^n, \xi^T \left( \widetilde{\mathbf{A}}_0 n_t + \sum_i \mathbf{A}_i n_i \right) \xi \geq 0.$$

Si la matrice  $\widetilde{\mathbf{A}}_0$  est symétrique définie positive, le vecteur  $(1, 0, \dots, 0)$  appartient au cône d'évolution. On rappelle que le polaire d'un cône  $\Gamma$  une partie d'un espace de Hilbert  $\mathbf{H}$  est l'ensemble  $\overset{\circ}{\Gamma} = \{y \in \mathbf{H} | \forall x \in \Gamma, |(x|y)| \geq 0\}$ . Le cône de propagation  $\Gamma_p$  est le polaire du cône d'évolution  $\Gamma_c$  dans  $\mathbb{R}^{n+1}$ . D'où  $\Gamma_p = \overset{\circ}{\Gamma}_c$ .

Dans le cas des équations de Maxwell, la formulation PML modifie la matrice  $\mathbf{K}(\mathbf{x})$ . La partie constante en  $\mathbf{p}$  n'est plus l'identité. Soit  $[\varepsilon]$  la matrice correspondant à la partie constante de  $\mathbf{det}(\mathbf{J}(\mathbf{x}))(\mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x}))^{-1}$ .

### Propriété 2.36

La condition CFL associée à la formulation PML est :

$$\frac{dt}{d\mathbf{x}} \leq \alpha, \tag{2.27}$$

avec  $\alpha$  la plus petite valeur propre de  $[\varepsilon]$  où  $\widetilde{\mathbf{A}}_0 = \begin{pmatrix} [\varepsilon] & 0 \\ 0 & [\varepsilon] \end{pmatrix}$ .

**Démonstration** : dans le cas des équations de Maxwell modifiées par une formulation PML, le cône d'évolution est défini par :

$$\forall \xi \in \mathbb{R}^n, \xi^T \left( \frac{1}{2} [\varepsilon] (\mathbf{E} \cdot \mathbf{E} + \mathbf{H} \cdot \mathbf{H}) n_t + (\mathbf{E} \times \mathbf{H}) \cdot \mathbf{n} \right) \xi \geq 0.$$

Soit  $\alpha$  la plus petite valeur propre de  $[\varepsilon]$ . On souhaite minorer le plus précisément possible le rapport  $\frac{n_t}{|\mathbf{n}|}$  afin d'obtenir une majoration de  $\frac{dt}{d\mathbf{x}}$  et obtenir ainsi la meilleure condition CFL possible.

$$\frac{1}{2}[\varepsilon](\mathbf{E}.\mathbf{E} + \mathbf{H}.\mathbf{H}) \geq \frac{1}{2}\alpha(E^2 + H^2)$$

$$\begin{aligned} (\mathbf{E} \times \mathbf{H}).\mathbf{n} &\leq |\mathbf{E}\mathbf{H}| |\mathbf{n}| && \text{(inégalité de Schwartz)} \\ &\leq |\mathbf{E}| |\mathbf{H}| |\mathbf{n}| \\ &\leq \frac{1}{2}(E^2 + H^2) |\mathbf{n}|. \end{aligned}$$

Donc  $\frac{1}{2}(E^2 + H^2)n_t - \frac{1}{2\alpha}(E^2 + H^2) |\mathbf{n}| \leq 0$  ce qui implique que  $\frac{n_t}{|\mathbf{n}|} \leq \frac{1}{\alpha}$ . Comme  $dt n_t + d\mathbf{x}\mathbf{n} = 0$ ,  $\frac{d\mathbf{x}}{dt} = -\frac{n_t}{\mathbf{n}}$ . D'où  $\frac{d\mathbf{x}}{dt} \geq \frac{1}{\alpha}$ .  $\square$

**Proposition 2.37**

La condition CFL associée à cette formulation est la valeur maximale sur le domaine PML de :

$$\frac{\kappa^2 + 4 \left( (1 - \kappa) \frac{\rho'_0(\theta)}{\rho} \right)^2 + \left( \kappa + (1 - \kappa) \frac{\rho_0(\theta)}{\rho} \right)^2}{2 \kappa \left( \kappa + (1 - \kappa) \frac{\rho_0(\theta)}{\rho} \right)} - \sqrt{\frac{\left( \kappa^2 + 4 \left( (1 - \kappa) \frac{\rho'_0(\theta)}{\rho} \right)^2 + \left( \kappa + (1 - \kappa) \frac{\rho_0(\theta)}{\rho} \right)^2 \right)^2 - 1}{4 \kappa^2 \left( \kappa + (1 - \kappa) \frac{\rho_0(\theta)}{\rho} \right)^2}}.$$

**Démonstration :** d'après (2.26)

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{Q}\mathbf{P}(\mathbf{p} \mathbf{I} + \dots) \mathbf{P}^T \mathbf{Q}^T$$

avec  $\mathbf{I} = \frac{1}{\kappa b} \mathbf{I}_3$ ,  $\mathbf{Q}$  matrice orthogonale de rotation et  $\mathbf{P} = \begin{pmatrix} b & -a & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa b \end{pmatrix}$ , où

$a = 2(1 - \kappa) \frac{\rho'_0(\theta)}{\rho}$ ,  $b = \frac{\hat{\rho}}{\rho}$  et  $\hat{\rho} = \kappa \rho + (1 - \kappa) \rho_0(\theta)$ . Posons  $[\varepsilon]$  la matrice  $\frac{1}{\kappa b} \mathbf{P}\mathbf{P}^T$ . Ces valeurs propres sont les mêmes que la matrice  $\mathbf{Q}\mathbf{P}\mathbf{I}\mathbf{P}^T \mathbf{Q}^T$ . Soit  $\mathbf{M} = [\varepsilon] - \lambda \mathbf{I}_3$ .

$$M = \begin{pmatrix} \frac{\kappa}{b} + \frac{a^2}{\kappa b} - \lambda & \frac{a}{\kappa} & 0 \\ \frac{a}{\kappa} & \frac{b}{\kappa} - \lambda & 0 \\ 0 & 0 & \kappa b - \lambda \end{pmatrix}$$

$$\begin{aligned} \det(M) &= (\kappa b - \lambda) \left[ \left( \frac{\kappa}{b} + \frac{a^2}{\kappa b} - \lambda \right) \left( \frac{b}{\kappa} - \lambda \right) - \left( \frac{a}{\kappa} \right)^2 \right] \\ &= (\kappa b - \lambda) \left[ \lambda^2 - \frac{\kappa^2 + a^2 + b^2}{\kappa b} \lambda + 1 \right]. \end{aligned}$$

La première valeur propre est donc  $\kappa b$  et les deux autres valeurs propres sont solutions de l'équation  $x^2 - \frac{\kappa^2 + a^2 + b^2}{\kappa b} x + 1 = 0$ .

$$\begin{aligned} \Delta &= \left( \frac{\kappa^2 + a^2 + b^2}{\kappa b} \right)^2 - 4 \\ &= \left( \frac{1}{\kappa b} \right)^2 (\kappa^4 + a^4 + b^4 + 2a^2b^2 + 2a^2\kappa^2 - 2b^2\kappa^2) \\ &= \left( \frac{1}{\kappa b} \right)^2 ((\kappa^2 - b^2)^2 + a^2(a^2 + 2b^2 + 2\kappa^2)) \geq 0 \end{aligned}$$

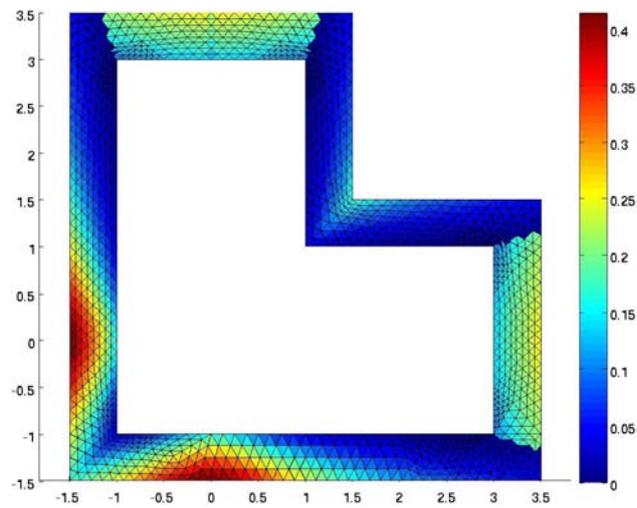
Les deux valeurs propres sont donc réelles, positives et inverses l'une de l'autre :

$$\begin{aligned} \lambda_1 &= \frac{\kappa^2 + a^2 + b^2}{2\kappa b} + \frac{\sqrt{(\kappa^2 + a^2 + b^2)^2 - 4\kappa^2 b^2}}{2\kappa b} \\ \lambda_2 &= \frac{\kappa^2 + a^2 + b^2}{2\kappa b} - \frac{\sqrt{(\kappa^2 + a^2 + b^2)^2 - 4\kappa^2 b^2}}{2\kappa b}. \end{aligned}$$

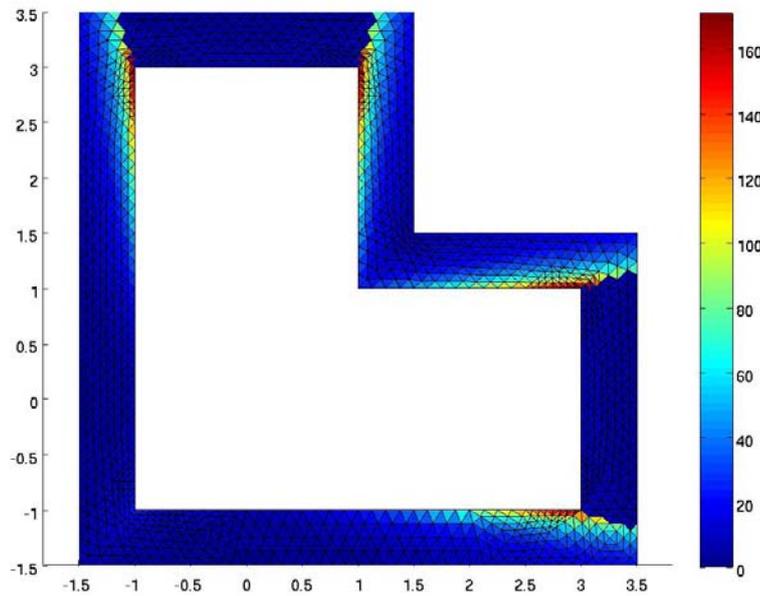
Comme  $\kappa \geq 1$ ,  $b \leq 1$ , on a la relation  $\lambda_2 \leq \kappa b \leq \lambda_1$ . La valeur propre qui nous intéresse est donc  $\lambda_2$ .  $\square$

Si  $a = 0$  soit donc  $\rho'_0(\theta) = 0$  alors  $\lambda_2 = \frac{b}{\kappa}$  et la condition CFL est alors  $\left( 1 + \left( \frac{1}{\kappa} - 1 \right) \frac{\rho_0(\theta)}{\rho} \right)$ . Pour les autres éléments du maillage, cas le plus général, la condition n'est pas aussi simple, et peut devenir très grande dans les zones à fort gradient pour la valeur de  $\rho'_0$ .

Considérons un maillage déstructuré non dégénéré mais quelconque. Nous allons donc calculer la valeur de  $\alpha$  (condition CFL) pour chaque élément et la visualiser. On constate que  $\alpha$  peut être fortement pénalisée selon les zones.

FIG. 2.5: Valeur de  $\alpha$  condition CFL

Pour mieux estimer la perte sur cette condition, nous allons représenter l'inverse la valeur de  $\alpha$  de la condition CFL (2.6).

FIG. 2.6: Inverse de la valeur de  $\alpha$  condition CFL

La condition CFL habituelle des Volumes Finis sera divisée par 180. En raffinant un facteur 200 sera atteint. Cependant, cette condition n'est forte que pour certains éléments, et comme prévu, dans les zones à fort gradient pour la valeur de  $\rho'_0$ . Une solution serait de travailler en implicite pour les éléments concernés.

**Propriété 2.38**

Les Equations de Maxwell peuvent être relevées sur  $\mathbb{R}^3 \times \mathbb{R}^+$  sous la forme :

$$\begin{cases} \bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}, t) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}, t) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}, t) = f_\varphi(\mathbf{x}, t) \\ \mathbf{T}(\mathbf{x})^{T-1} \frac{\partial \psi(\mathbf{x}, t)}{\partial t} + \mathbf{T}(\mathbf{x})^{T-1} \mathbf{F}(\mathbf{x}) \psi(\mathbf{x}, t) = \varphi(\mathbf{x}, t). \end{cases}$$

avec les matrices :

$$\begin{aligned} \mathbf{T}(\mathbf{x}) &= \mathbf{QP}(\mathbf{x}) & \bar{\mathbf{I}}(\mathbf{x}) &= \mathbf{T}(\mathbf{x}) \mathbf{I}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T \\ \bar{\mathbf{A}}(\mathbf{x}) &= \mathbf{T}(\mathbf{x}) \mathbf{A}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T & \bar{\mathbf{B}}(\mathbf{x}) &= \mathbf{T}(\mathbf{x}) \mathbf{B}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T \end{aligned}$$

**Démonstration** : soit  $\varphi$  la solution de ces Equations. D'après (2.26), les équations sont données sur le domaine PML par :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) \varphi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}) = f_\varphi(\mathbf{x}),$$

où  $f_\varphi$  est le second membre associé. Posons  $\mathbf{T}(\mathbf{x}) = \mathbf{QP}(\mathbf{x})$  et introduisons la notation barre qui signifiera  $\mathbf{T}(\mathbf{x}) \bullet \mathbf{T}(\mathbf{x})^T$ . Posons

$$\psi(\mathbf{x}) = \left( \mathbf{p} \mathbf{T}(\mathbf{x})^{T-1} + \mathbf{T}(\mathbf{x})^{T-1} \mathbf{F}(\mathbf{x}) \right)^{-1} \varphi(\mathbf{x}).$$

Viennent ainsi les Equations de Maxwell sur  $\mathbb{R}^3$  :

$$\begin{cases} \mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) \varphi(\mathbf{x}) + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}) = f_\varphi(\mathbf{x}) \\ \mathbf{p} \mathbf{T}(\mathbf{x})^{T-1} \psi(\mathbf{x}) + \mathbf{T}(\mathbf{x})^{T-1} \mathbf{F}(\mathbf{x}) \psi(\mathbf{x}) = \varphi(\mathbf{x}). \end{cases}$$

Par transformée de Laplace inverse, on obtient le résultat proposé.  $\square$

**2.6.2 Formulation hybride**

Le difféomorphisme et le changement de variable sont donnés par (2.24) :

$$\begin{cases} \hat{x} = x + (\kappa - 1) \inf(x, y) \\ \hat{y} = y + (\kappa - 1) \inf(x, y) \end{cases} \quad \begin{cases} \tilde{x} = \hat{x} + \frac{1}{\mathbf{p}} \int \sigma_x(\tau) d\tau \\ \tilde{y} = \hat{y} + \frac{1}{\mathbf{p}} \int \sigma_y(\tau) d\tau \end{cases}$$

**Proposition 2.39**

La matrice PML associée à cette formulation peut être décomposée sous la forme :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{T}(\mathbf{x}) \left( \mathbf{p} \mathbf{I} + \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x}) (\mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}))^{-1} \right) \mathbf{T}(\mathbf{x})^T, \quad (2.28)$$

$$\begin{aligned} \text{où } \mathbf{I} &= \frac{1}{\kappa} \mathbf{I}_3 & \mathbf{F}(\mathbf{x}) &= \text{diag}(\sigma_x(\mathbf{x}), \sigma_y(\mathbf{x}), 0) \\ \mathbf{A}(\mathbf{x}) &= \frac{1}{\kappa} \text{diag}(\sigma_y(\mathbf{x}) - \sigma_x(\mathbf{x}), \sigma_x(\mathbf{x}) - \sigma_y(\mathbf{x}), \sigma_x(\mathbf{x}) + \sigma_y(\mathbf{x})) \\ \mathbf{B}(\mathbf{x}) &= \frac{1}{\kappa} \text{diag}(\sigma_x(\mathbf{x})(\sigma_x(\mathbf{x}) - \sigma_y(\mathbf{x})), \sigma_y(\mathbf{x})(\sigma_y(\mathbf{x}) - \sigma_x(\mathbf{x})), \sigma_x(\mathbf{x}) \sigma_y(\mathbf{x})) \end{aligned}$$

et la matrice triangulaire  $\mathbf{T}(\mathbf{x})$  dépend de la partie du domaine considérée : sur  $\mathbf{D}_{nc}^+$ ,

$$\mathbf{T}^+ = \begin{pmatrix} 1 & 1 - \kappa & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa \end{pmatrix} \text{ et sur } \mathbf{D}_{nc}^-, \quad \mathbf{T}^- = \begin{pmatrix} \kappa & 0 & 0 \\ 1 - \kappa & 1 & 0 \\ 0 & 0 & \kappa \end{pmatrix}.$$

**Démonstration** : la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  est donnée par (2.31). Pour relever le système différentiel en temps, il faut décomposer cette matrice en éléments simples.

$$\begin{aligned} \frac{\mathbf{p}(\mathbf{p} + \sigma_x(\mathbf{x}))}{\mathbf{p} + \sigma_y(\mathbf{x})} &= \mathbf{p} + \sigma_x(\mathbf{x}) - \sigma_y(\mathbf{x}) + \frac{\sigma_y(\mathbf{x})(\sigma_y(\mathbf{x}) - \sigma_x(\mathbf{x}))}{\mathbf{p} + \sigma_y(\mathbf{x})} \\ \frac{\mathbf{p}(\mathbf{p} + \sigma_y(\mathbf{x}))}{\mathbf{p} + \sigma_x(\mathbf{x})} &= \mathbf{p} + \sigma_y(\mathbf{x}) - \sigma_x(\mathbf{x}) + \frac{\sigma_x(\mathbf{x})(\sigma_x(\mathbf{x}) - \sigma_y(\mathbf{x}))}{\mathbf{p} + \sigma_x(\mathbf{x})} \end{aligned}$$

La matrice de la formulation PML peut se décomposer en éléments simples sous la forme énoncée.  $\square$

**Proposition 2.40**

La condition CFL associée à (2.28) est  $\frac{1}{\kappa} + \kappa - 1 - \left(1 - \frac{1}{\kappa}\right) \sqrt{1 + \kappa^2}$ .

**Démonstration** : Posons  $[\varepsilon]$  la matrice  $\mathbf{p} \mathbf{T}^T(\mathbf{x}) \mathbf{T}(\mathbf{x})$ . Elle représente l'évolution temporelle des équations. Une fois le système PML relevé en temps, nous aurons  $[\varepsilon] \frac{\partial}{\partial t} \varphi(\mathbf{x}, t) + \dots$ . La condition CFL dépend donc des valeurs propres de cette matrice.

La constante  $\kappa$  est supérieure ou égale à 1. Le cas limite  $\kappa = 1$  implique  $T = I_3$  soit  $[\varepsilon] = \mathbf{I}_3$ . La condition CFL est alors égale à 1. Ce cas n'est cependant pertinent : le domaine non convexe est souhaité le plus petit possible, ce qui correspond à de grandes valeurs de  $\kappa$ . Supposons donc  $\kappa \gg 1$ .

**Premier cas** :  $x \leq y$ . Posons  $\tau = \frac{1}{\kappa}$ .  $[\varepsilon] = \begin{pmatrix} \tau & \tau - 1 & 0 \\ \tau - 1 & \tau - 2 + 2\kappa & 0 \\ 0 & 0 & \kappa \end{pmatrix}$ .

$$\begin{aligned} \det([\varepsilon] - \lambda \mathbf{I}_3) &= \begin{vmatrix} \tau - \lambda & \tau - 1 & 0 \\ \tau - 1 & \tau - 2 + 2\kappa - \lambda & 0 \\ 0 & 0 & \kappa - \lambda \end{vmatrix} \\ &= (\kappa - \lambda) [(\tau - \lambda)(\tau - 2 + 2\kappa - \lambda) - (\tau - 1)^2] \\ &= (\kappa - \lambda) [\tau^2 - 2\tau + 2 + \lambda^2 - \lambda(2\tau + 2\kappa - 2) - (\tau - 1)^2] \\ &= (\kappa - \lambda) [\lambda^2 - 2(\tau + \kappa - 1)\lambda + 1]. \end{aligned}$$

Une première valeur propre est  $\kappa$ , les deux suivantes sont solutions de l'équation  $x^2 - 2(\tau + \kappa - 1)x + 1 = 0$ .

$$\begin{aligned}\Delta &= 4(\tau + \kappa - 1)^2 - 4 = 4(\tau + \kappa - 1 - 1)(\tau + \kappa - 1 + 1) \\ &= 4\kappa(\tau + \kappa)(\tau^2 - 2\tau + 1) = 4\kappa(\tau + \kappa)(\tau - 1)^2 > 0\end{aligned}$$

Cette équation admet donc deux racines réelles positives :

$$\begin{aligned}\lambda_1 &= \frac{1}{\kappa} + \kappa - 1 + \left(1 - \frac{1}{\kappa}\right)\sqrt{1 + \kappa^2} \\ \lambda_2 &= \frac{1}{\kappa} + \kappa - 1 - \left(1 - \frac{1}{\kappa}\right)\sqrt{1 + \kappa^2}.\end{aligned}$$

De plus,  $\lambda_1 \lambda_2 = 1$  et d'après ce qui précède,  $\lambda_2 < \lambda_1 = \frac{1}{\lambda_2}$  donc  $\lambda_2 < 1$  et par conséquent  $\lambda_2 < \kappa$ . La plus petite valeur propre de la matrice  $[\varepsilon]$  est donc  $\frac{1}{\kappa} + \kappa - 1 - \left(1 - \frac{1}{\kappa}\right)\sqrt{1 + \kappa^2}$  d'où la condition CFL proposée.

**Deuxième cas :**  $y \leq x$ .

$$[\varepsilon] = \begin{pmatrix} \tau - 2 + 2\kappa & \tau - 1 & 0 \\ \tau - 1 & \tau & 0 \\ 0 & 0 & \kappa \end{pmatrix}$$

$$\begin{aligned}\det([\varepsilon] - \lambda \mathbf{I}_3) &= \begin{vmatrix} \tau - \lambda & \tau - 1 & 0 \\ \tau - 1 & \tau - 2 + 2\kappa - \lambda & 0 \\ 0 & 0 & \kappa - \lambda \end{vmatrix} \\ &= (\kappa - \lambda) [\lambda^2 - 2(\tau + \kappa - 1)\lambda + 1].\end{aligned}$$

On obtient les mêmes valeurs propres que dans le cas précédent. \(\boxtimes\)

**Propriété 2.41** (Comportement asymptotique de la condition CFL)

La valeur de la condition CFL tend vers  $\frac{1}{2\kappa}$  quand  $\kappa$  tend vers l'infini.

**Démonstration :** Supposons que  $\kappa \rightarrow \infty$ . Alors  $\tau \rightarrow 0^+$ . Effectuons un développement limité.

$$\begin{aligned}\sqrt{1 + \frac{1}{\tau}} &= \frac{1}{\tau} \sqrt{\tau^2 + 1} = \frac{1}{\tau} \left(1 + \frac{1}{2}\tau^2 - \frac{1}{8}\tau^4 + o(\tau^4)\right) \\ \lambda &= \frac{1}{\kappa} + \kappa - 1 - \left(1 - \frac{1}{\kappa}\right)\sqrt{1 + \kappa^2} \\ &= \tau + \frac{1}{\tau} - 1 - (1 - \tau)\frac{1}{\tau} \left(1 + \frac{1}{2}\tau^2 - \frac{1}{8}\tau^4 + o(\tau^4)\right) \\ &= \tau + \frac{1}{\tau} - 1 + 1 + \frac{1}{2}\tau^2 - \frac{1}{\tau} - \frac{1}{2}\tau + \frac{1}{8}\tau^3 + o(\tau^3) \\ &= \frac{1}{2}\tau + \frac{1}{2}\tau^2 + \frac{1}{8}\tau^3 + o(\tau^3).\end{aligned}$$

D'où  $\lambda \underset{\tau \rightarrow 0^+}{\sim} \frac{1}{2}\tau$ . \(\boxtimes\)

**Remarque :** la condition CFL ne dépend pas de l'allongement du L. Dans la formulation étoilée précédente, cette dernière est pénalisée à cause des forts gradients de la dérivée de  $\rho_0$ . Ici elle est majorée par  $\frac{1}{2\kappa}$ . Ce sera d'autant plus intéressant que le domaine sera non convexe.

**Propriété 2.42**

Les Equations de Maxwell peuvent être relevées sur  $\mathbb{R}^3 \times \mathbb{R}^+$  sous la forme :

$$\begin{cases} \bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}, t) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}, t) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}, t) = f_\varphi(\mathbf{x}, t) \\ \mathbf{T}(\mathbf{x})^{T^{-1}} \frac{\partial \psi(\mathbf{x}, t)}{\partial t} + \mathbf{T}(\mathbf{x})^{T^{-1}} \mathbf{F}(\mathbf{x}) \psi(\mathbf{x}, t) = \varphi(\mathbf{x}, t). \end{cases}$$

avec les matrices :

$$\bar{\mathbf{I}}(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \mathbf{I}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T \quad \bar{\mathbf{A}}(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \mathbf{A}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T \quad \bar{\mathbf{B}}(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \mathbf{B}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T$$

**Démonstration :** soit  $\varphi$  solution des ces équations.

D'après (2.28)  $\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{T}(\mathbf{x}) (\mathbf{p} \mathbf{I} + \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x}) (\mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}))^{-1}) \mathbf{T}(\mathbf{x})^T$  et les équations sont données sur le domaine PML par :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) \varphi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}) = f_\varphi(\mathbf{x}),$$

où  $f_\varphi$  est le second membre associé. La notation barre signifiera  $\mathbf{T}(\mathbf{x}) \bullet \mathbf{T}(\mathbf{x})^T$ .

Posons  $\psi(\mathbf{x}) = (\mathbf{p} \mathbf{T}(\mathbf{x})^{T^{-1}} + \mathbf{T}(\mathbf{x})^{T^{-1}} \mathbf{F}(\mathbf{x}))^{-1} \varphi(\mathbf{x})$ . Viennent ainsi les Equations de Maxwell sur  $\mathbb{R}^3$  :

$$\begin{cases} \mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) \varphi(\mathbf{x}) + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}) = f_\varphi(\mathbf{x}) \\ \mathbf{p} \mathbf{T}(\mathbf{x})^{T^{-1}} \psi(\mathbf{x}) + \mathbf{T}(\mathbf{x})^{T^{-1}} \mathbf{F}(\mathbf{x}) \psi(\mathbf{x}) = \varphi(\mathbf{x}). \end{cases}$$

Par transformée de Laplace inverse, on obtient le résultat proposé. \(\boxtimes\)

### 2.6.3 Deuxième formulation hybride

Dans ce cas, le changement de variables (2.24) était devenu (2.25), soit une symétrisation en  $x$  et  $y$ . Posons  $\sigma_x := \sigma_y := \sigma$ .

**Proposition 2.43**

La matrice associée à cette formulation peut être décomposée sous la forme :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{T}^T(\mathbf{x}) \left( \mathbf{p} \mathbf{I} + \mathbf{A}(\mathbf{x}) + \frac{1}{\mathbf{p} + 2\sigma(\mathbf{x})} \mathbf{B}(\mathbf{x}) \right) \mathbf{T}(\mathbf{x}), \quad (2.29)$$

où  $\mathbf{I} = \frac{1}{\kappa} \mathbf{I}_3$ ,  $\mathbf{A}(\mathbf{x}) = \frac{2\sigma(\mathbf{x})}{\kappa} \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ,  $\mathbf{B}(\mathbf{x}) = \frac{2\sigma^2(\mathbf{x})}{\kappa} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  et la matrice triangulaire  $\mathbf{T}(\mathbf{x})$  reste inchangée par rapport à la formulation précédente.

**Démonstration** : la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  est donnée par (2.33). Pour relever le système différentiel en temps, il faut décomposer cette matrice en éléments simples.

$$\frac{\mathbf{p} + \sigma(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})} = 1 - \frac{\sigma(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})}$$

$$\frac{(\mathbf{p} + \sigma(\mathbf{x}))^2 + \sigma^2(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})} = \mathbf{p} + \frac{2\sigma^2(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})}$$

La matrice de la formulation PML peut se décomposer en éléments simples sous la forme énoncée.  $\square$

Dans ce cas, la partie imaginaire s'annule quand  $\sigma(x, y) = f'(\hat{x} + \hat{y} - R) = 0$ , donc quand  $\hat{x} + \hat{y} = R$ . Pour  $x = y$ , on obtient alors que le point  $\left(\frac{R}{2\kappa}, \frac{R}{2\kappa}\right)$  donne l'épaisseur de la zone PML non convexe sans décroissance. On a réduit de moitié cette zone par rapport à la formulation précédente.

#### Proposition 2.44

La condition CFL associée à (2.29) est  $\frac{1}{\kappa} + \kappa - 1 - \left(1 - \frac{1}{\kappa}\right) \sqrt{1 + \kappa^2}$ , qui tend vers  $\frac{1}{2\kappa}$  quand  $\kappa$  tend vers l'infini.

**Démonstration** : la condition CFL associée à cette formulation est la même que pour la formulation précédente. En effet, la matrice associée à cette formulation peut être décomposée sous la forme (2.29) avec  $\mathbf{I} = \frac{1}{\kappa} \mathbf{I}_3$  et la matrice triangulaire  $\mathbf{T}$  reste inchangée par rapport à précédemment.  $\square$

#### Propriété 2.45

Les Equations de Maxwell peuvent être relevées sur  $\mathbb{R}^3 \times \mathbb{R}^+$  sous la forme :

$$\begin{cases} \bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}, t) + \psi(\mathbf{x}, t) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}, t) = f_\varphi(\mathbf{x}, t) \\ \frac{\partial \psi(\mathbf{x}, t)}{\partial t} + 2\sigma(\mathbf{x}) \psi(\mathbf{x}, t) = \bar{\mathbf{B}}(\mathbf{x}) \varphi(\mathbf{x}, t). \end{cases}$$

avec les matrices :

$$\bar{\mathbf{I}}(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \mathbf{I}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T \quad \bar{\mathbf{A}}(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \mathbf{A}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T \quad \bar{\mathbf{B}}(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \mathbf{B}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T$$

**Démonstration** : la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  est donnée par (2.33) autrement dit

$$\mathbf{T}(\mathbf{x}) \left( \mathbf{p} \mathbf{I} + \mathbf{A}(\mathbf{x}) + \frac{1}{\mathbf{p} + 2\sigma(\mathbf{x})} \mathbf{B}(\mathbf{x}) \right) \mathbf{T}(\mathbf{x})^T.$$

En procédant comme la formulation précédente, on obtient les Equations de Maxwell sur  $\mathbb{R}^3$  :

$$\begin{cases} \mathbf{p} \bar{\mathbf{I}}(\mathbf{x})\varphi(\mathbf{x}) + \bar{\mathbf{A}}(\mathbf{x})\varphi(\mathbf{x}) + \psi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}) = f_\varphi(\mathbf{x}) \\ \mathbf{p} \psi(\mathbf{x}) + 2\sigma(\mathbf{x}) \psi(\mathbf{x}) = \bar{\mathbf{B}}(\mathbf{x})\varphi(\mathbf{x}). \end{cases}$$

Par transformée de Laplace inverse, on obtient le résultat proposé.  $\square$

## 2.6.4 Conclusions et généralisation

La formulation harmonique des Equations de Maxwell peut être mise sous la forme

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) \begin{pmatrix} e \\ h \end{pmatrix} + \begin{pmatrix} 0 & -\nabla \times \\ \nabla \times & 0 \end{pmatrix} \begin{pmatrix} e \\ h \end{pmatrix} = \begin{pmatrix} -j \\ -m \end{pmatrix} \quad (2.30)$$

où  $(j, m)$  sont des seconds membres réguliers. Sous certaines hypothèses ( $\mathbf{K}(\mathbf{p}, \mathbf{x})$   $L^\infty$  dans  $\mathbf{x}$ ,  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  holomorphe en  $\mathbf{p} \in \mathbf{H} \supset i\mathbb{R}_*^+$ ,  $\exists \mathbf{p}_0 \in \mathbf{H}$  tel que  $\mathbf{Re}(\mathbf{K}(\mathbf{p}_0, \mathbf{x}))$  est coercive et  $\exists \mathbf{p}_1 \in \mathbf{H}$  tel que  $\mathbf{K}(\mathbf{p}_1, \mathbf{x}) + \mathbf{A}_i^c \partial_i$  soit à résolvante compacte), il existe une unique solution à ce problème pour  $\mathbf{p} \in \mathbf{H}/\mathbf{S}$  où  $\mathbf{S}$  est un ensemble localement fini, pouvant être nul, avec des conditions aux limites maximales monotones.

### Propriété 2.46

La matrice des équations de Maxwell dans le milieu PML pour des problèmes harmoniques est

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \det \mathbf{T} \left( \mathbf{T}^{-1} \mathbf{P}^* \mathbf{M} \mathbf{P} \mathbf{T}^{*-1} \right)$$

$$\text{avec } \mathbf{T} = \frac{D\varphi}{D\mathbf{x}}, \mathbf{M} = \text{diag} \left( \prod_k^i (\mathbf{p} + \lambda_k) \right) \text{ et } \mathbf{P}^* = \mathbf{P}^{-1}.$$

**Démonstration** : le calcul de la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  a été établie à chaque formulation. Néanmoins, étant donné le changement de variable générique

$$\tilde{\mathbf{x}} = \varphi(\mathbf{x}) + \frac{1}{\mathbf{p}} f(\varphi(\mathbf{x})),$$

il est possible d'obtenir une formulation générale. Soit la matrice jacobienne  $\mathbf{J}(\mathbf{p}, \mathbf{x})$ . Elle s'écrit  $\frac{D\tilde{\mathbf{x}}}{D\mathbf{x}} = \left( \mathbf{I} + \frac{1}{\mathbf{p}} \frac{Df}{D\varphi} \right) \frac{D\varphi}{D\mathbf{x}}$ . Comme  $f$  est le gradient d'une fonction convexe par rapport à  $\varphi$ , il existe  $\mathbf{P}$  une matrice unitaire  $\mathbf{P}^* = \mathbf{P}^{-1}$  telle que

$$\frac{Df}{D\varphi} = \mathbf{P}^* \left( \mathbf{I} + \frac{1}{\mathbf{p}} \mathbf{D} \right) \mathbf{P}$$

où  $\mathbf{D}$  est une matrice diagonale. Soit  $\lambda_i$ ,  $i = 1, 2, 3$  les valeurs propres de cette matrice. Posons  $\mathbf{T} = \frac{D\varphi}{D\mathbf{x}}$ , alors  $\mathbf{J}(\mathbf{p}, \mathbf{x}) = \mathbf{P}^* \left( \mathbf{I} + \frac{1}{\mathbf{p}} \mathbf{D} \right) \mathbf{P} \mathbf{T}$ . La transformée de Hodge

associée est  $\star = \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) \mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} (\mathbf{J}(\mathbf{p}, \mathbf{x})^T)^{-1}$ . Le résultat proposé est alors obtenu en exprimant  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  fonction de la transformée de Hodge pour les Equations de Maxwell, à savoir  $\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{p}\star$ .  $\square$

**Théorème 2.47**

Les Equations de Maxwell dans le milieu PML s'écrivent en instationnaire

$$\bar{\mathbf{I}}(\mathbf{x}) \frac{\partial}{\partial t} \begin{pmatrix} e \\ h \end{pmatrix} + \bar{\mathbf{A}}(\mathbf{x}) \begin{pmatrix} e \\ h \end{pmatrix} + \bar{\mathbf{B}}(\mathbf{x}) \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 & -\nabla \times \\ \nabla \times & 0 \end{pmatrix} \begin{pmatrix} e \\ h \end{pmatrix} = \begin{pmatrix} -j \\ -m \end{pmatrix},$$

et  $(u, v)$  sont solutions des EDO

$$\begin{pmatrix} \mathbf{T}^* \mathbf{P}^{-1} & 0 \\ 0 & \mathbf{T}^* \mathbf{P}^{-1} \end{pmatrix} \frac{\partial}{\partial t} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \mathbf{T}^* \mathbf{P}^{-1} \mathbf{F}(\mathbf{x}) & 0 \\ 0 & \mathbf{T}^* \mathbf{P}^{-1} \mathbf{F}(\mathbf{x}) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} e \\ h \end{pmatrix},$$

avec les matrices déduites (2.46)  $\mathbf{F}(\mathbf{x}) = \text{diag}(\lambda_i)$ ,  $\mathbf{A}(\mathbf{x}) = \text{diag} \left( \sum_k^i \lambda_k - \lambda_i \right)$ ,

$$\mathbf{B}(\mathbf{x}) = \text{diag} \left( \prod_k^i (\lambda_i - \lambda_k) \right), \text{ et}$$

$$\begin{aligned} \bar{\mathbf{I}}(\mathbf{x}) &= \begin{pmatrix} \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{P} \mathbf{T}^{*-1} & 0 \\ 0 & \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{P} \mathbf{T}^{*-1} \end{pmatrix} \\ \bar{\mathbf{A}}(\mathbf{x}) &= \begin{pmatrix} \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{A}(\mathbf{x}) \mathbf{P} \mathbf{T}^{*-1} & 0 \\ 0 & \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{A}(\mathbf{x}) \mathbf{P} \mathbf{T}^{*-1} \end{pmatrix} \\ \bar{\mathbf{B}}(\mathbf{x}) &= \begin{pmatrix} \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{B}(\mathbf{x}) \mathbf{P} \mathbf{T}^{*-1} & 0 \\ 0 & \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{B}(\mathbf{x}) \mathbf{P} \mathbf{T}^{*-1} \end{pmatrix} \end{aligned}$$

**Démonstration** : soit  $(i, j, k)$  deux à deux différents. Comme

$$\frac{(\mathbf{p} + \lambda_i)(\mathbf{p} + \lambda_j)}{\mathbf{p} + \lambda_k} = \mathbf{p} + (\lambda_i + \lambda_j - \lambda_k) - \frac{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}{\mathbf{p} + \lambda_k},$$

la matrice  $\begin{pmatrix} \frac{(\mathbf{p} + \lambda_2)(\mathbf{p} + \lambda_3)}{\mathbf{p} + \lambda_1} & 0 & 0 \\ 0 & \frac{(\mathbf{p} + \lambda_1)(\mathbf{p} + \lambda_3)}{\mathbf{p} + \lambda_2} & 0 \\ 0 & 0 & \frac{(\mathbf{p} + \lambda_1)(\mathbf{p} + \lambda_2)}{\mathbf{p} + \lambda_3} \end{pmatrix}$  se décompose

en éléments simples sous la forme  $\mathbf{p} \mathbf{I}_3 + \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x}) (\mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}))^{-1}$  où

$$\mathbf{F}(\mathbf{x}) = \text{diag}(\lambda_i) \quad \mathbf{A}(\mathbf{x}) = \text{diag} \left( \sum_k^i \lambda_k - \lambda_i \right) \quad \mathbf{B}(\mathbf{x}) = \text{diag} \left( \prod_k^i (\lambda_i - \lambda_k) \right).$$

Posons  $\mathbf{T} = \begin{pmatrix} \mathbf{D}\varphi \\ \mathbf{D}\mathbf{x} \end{pmatrix}$ . On en déduit que la matrice PML peut être écrite comme :

$$\det(\mathbf{T}) \left( \mathbf{T}^{-1} \mathbf{P}^* (\mathbf{p} \mathbf{I}_3 + \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x}) (\mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}))^{-1}) \mathbf{P} \mathbf{T}^{*-1} \right). \quad (2.31)$$

Remarquons que  $(\mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}))^{-1} \mathbf{P} \mathbf{T}^{*-1} = (\mathbf{T}^* \mathbf{P}^{-1} (\mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x})))^{-1}$ . Posons

$$\begin{aligned} \bar{\mathbf{I}}(\mathbf{x}) &= \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{I}_6 \mathbf{P} \mathbf{T}^{*-1} \\ \bar{\mathbf{A}}(\mathbf{x}) &= \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \begin{pmatrix} \mathbf{A}(\mathbf{x}) & 0 \\ 0 & \mathbf{A}(\mathbf{x}) \end{pmatrix} \mathbf{P} \mathbf{T}^{*-1} \\ \bar{\mathbf{B}}(\mathbf{x}) &= \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \begin{pmatrix} \mathbf{B}(\mathbf{x}) & 0 \\ 0 & \mathbf{B}(\mathbf{x}) \end{pmatrix} \mathbf{P} \mathbf{T}^{*-1} \\ \tilde{\mathbf{F}}(\mathbf{x}) &= \mathbf{T}^* \mathbf{P}^{-1} \begin{pmatrix} \mathbf{F}(\mathbf{x}) & 0 \\ 0 & \mathbf{F}(\mathbf{x}) \end{pmatrix} \\ \tilde{\mathbf{I}}(\mathbf{x}) &= \mathbf{T}^* \mathbf{P}^{-1} \mathbf{I}_6 \end{aligned}$$

Soit  $\begin{pmatrix} e \\ h \end{pmatrix}$  la solution des Equations de Maxwell et  $\begin{pmatrix} -j \\ -m \end{pmatrix}$  le second membre. Les équations sont données sur le domaine PML par :

$$\left( \mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) + \bar{\mathbf{A}}(\mathbf{x}) + \bar{\mathbf{B}}(\mathbf{x}) (\mathbf{p} \tilde{\mathbf{I}}(\mathbf{x}) + \tilde{\mathbf{F}}(\mathbf{x}))^{-1} \right) \begin{pmatrix} e \\ h \end{pmatrix} + \begin{pmatrix} 0 & -\nabla \times \\ \nabla \times & 0 \end{pmatrix} \begin{pmatrix} e \\ h \end{pmatrix} = \begin{pmatrix} -j \\ -m \end{pmatrix}.$$

Posons  $\begin{pmatrix} u \\ v \end{pmatrix} = (\mathbf{p} \tilde{\mathbf{I}}(\mathbf{x}) + \tilde{\mathbf{F}}(\mathbf{x}))^{-1} \begin{pmatrix} e \\ h \end{pmatrix}$ . Il s'agit donc de la solution d'une Equation Différentielle Ordinaire. Introduite dans l'équation précédente, on obtient le système régissant les Equations de Maxwell sur  $\mathbb{R}^6$  :

$$\begin{cases} \mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) \begin{pmatrix} e \\ h \end{pmatrix} + \bar{\mathbf{A}}(\mathbf{x}) \begin{pmatrix} e \\ h \end{pmatrix} + \bar{\mathbf{B}}(\mathbf{x}) \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 & -\nabla \times \\ \nabla \times & 0 \end{pmatrix} \begin{pmatrix} e \\ h \end{pmatrix} = \begin{pmatrix} -j \\ -m \end{pmatrix} \\ \mathbf{p} \tilde{\mathbf{I}}(\mathbf{x}) \begin{pmatrix} u \\ v \end{pmatrix} + \tilde{\mathbf{F}}(\mathbf{x}) \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} e \\ h \end{pmatrix}. \end{cases}$$

En utilisant la transformée de Laplace inverse, on obtient in fine sur  $\mathbb{R}^6 \times \mathbb{R}^+$

$$\begin{cases} \bar{\mathbf{I}}(\mathbf{x}) \partial_t \begin{pmatrix} e \\ h \end{pmatrix} + \bar{\mathbf{A}}(\mathbf{x}) \begin{pmatrix} e \\ h \end{pmatrix} + \bar{\mathbf{B}}(\mathbf{x}) \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 & -\nabla \times \\ \nabla \times & 0 \end{pmatrix} \begin{pmatrix} e \\ h \end{pmatrix} = \begin{pmatrix} -j \\ -m \end{pmatrix} \\ \tilde{\mathbf{I}}(\mathbf{x}) \partial_t \begin{pmatrix} u \\ v \end{pmatrix} + \tilde{\mathbf{F}}(\mathbf{x}) \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} e \\ h \end{pmatrix}. \end{cases}$$

Il s'agit bien d'un système de Friedrichs, facilement implémentable, et dont l'existence et l'unicité de la solution sont prouvés via les outils du chapitre 1 de ce mémoire.  $\square$

**Remarque importante** : ce théorème met en lumière une propriété particulièrement importante des PML : la partie spatiale de l'opérateur  $\begin{pmatrix} 0 & -\nabla \times \\ \nabla \times & 0 \end{pmatrix}$  n'est pas affectée dans le milieu PML. Cela garantit donc la possibilité de choisir librement la méthode d'approximation numérique (Elements Finis, Volumes Finis, Galerkin Discontinu, ...). On pourrait également écrire une formulation de Gedney plus appropriée à une approximation FDTD pour des schémas leap frog.

La formulation précédente a été établie pour des Equations de Maxwell tridimensionnelles. Pour obtenir les Equations bidimensionnelles, il suffit de réaliser une descente de Hadamard sur les Equations précédentes. Dans les résultats numériques présentés dans (2.10), nous travaillerons en bidimensionnel. Nous allons donc expliciter ces dernières.

**Propriété 2.48**

Soit  $\varphi = (E_x, E_y, H_z)^T$ . Les équations de Maxwell bidimensionnelles sont données sur le domaine PML par le système de Friedrichs sur  $\mathbb{R}^3 \times \mathbb{R}^+$  :

$$\begin{cases} \bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}, t) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}, t) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}, t) = f_\varphi(\mathbf{x}, t) \\ \tilde{\mathbf{I}}(\mathbf{x}) \frac{\partial \psi(\mathbf{x}, t)}{\partial t} + \tilde{\mathbf{F}}(\mathbf{x}) \psi(\mathbf{x}, t) = \varphi(\mathbf{x}, t) \end{cases}$$

avec les matrices suivantes :

$$\begin{aligned} \bar{\mathbf{I}}(\mathbf{x}) &= \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{P} \mathbf{T}^{*-1} \\ \bar{\mathbf{A}}(\mathbf{x}) &= \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{A}(\mathbf{x}) \mathbf{P} \mathbf{T}^{*-1} \\ \bar{\mathbf{B}}(\mathbf{x}) &= \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{B}(\mathbf{x}) \mathbf{P} \mathbf{T}^{*-1} \\ \tilde{\mathbf{F}}(\mathbf{x}) &= \mathbf{T}^* \mathbf{P}^{-1} \mathbf{F}(\mathbf{x}) \\ \tilde{\mathbf{I}}(\mathbf{x}) &= \mathbf{T}^* \mathbf{P}^{-1} \end{aligned}$$

*Remarque* : dans le cas où  $\varphi = (H_x, H_y, E_z)^T$ , il faudrait remplacer la matrice  $\begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix}$  par  $\begin{pmatrix} 0 & 0 & \partial_y \\ 0 & 0 & -\partial_x \\ \partial_y & -\partial_x & 0 \end{pmatrix}$ .

**Démonstration** : l'opérateur  $\nabla \times$  découple naturellement les 6 équations de Maxwell une fois la « descente de Hadamard » effectuée, c'est à dire en remplaçant l'opérateur  $\partial_z$  par zéro, et en considérant comme nulle la valeur propre de la matrice diagonale du jacobien de  $F$  par rapport à  $\varphi$  (ceci est vrai en l'absence de PML sur l'axe  $\mathbf{z}$ ). De plus, à la vue des matrices mises en jeu dans la formulation tridimensionnelle (diagonales par bloc), le système sera facilement découplable.

$$\nabla \times \mathbf{E} = \begin{pmatrix} \partial_y E_z \\ -\partial_x E_z \\ \partial_x E_y - \partial_y E_x \end{pmatrix} \quad -\nabla \times \mathbf{H} = \begin{pmatrix} -\partial_y H_z \\ \partial_x H_z \\ \partial_y H_x - \partial_x H_y \end{pmatrix}.$$

De fait, les équations se décomposent en deux systèmes dit mode transverse électrique  $\varphi = (E_x, E_y, H_z)^T$  ou magnétique  $\varphi = (H_x, H_y, E_z)^T$ . En procédant de manière identique à la démonstration tridimensionnelle, la décomposition en éléments simples de la matrice PML étant (2.31), posons :

$$\begin{aligned}\bar{\mathbf{I}}(\mathbf{x}) &= \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{P} \mathbf{T}^{*-1} \\ \bar{\mathbf{A}}(\mathbf{x}) &= \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{A}(\mathbf{x}) \mathbf{P} \mathbf{T}^{*-1} \\ \bar{\mathbf{B}}(\mathbf{x}) &= \det(\mathbf{T}) (\mathbf{T}^{-1} \mathbf{P}^*)^{-1} \mathbf{B}(\mathbf{x}) \mathbf{P} \mathbf{T}^{*-1} \\ \tilde{\mathbf{F}}(\mathbf{x}) &= \mathbf{T}^* \mathbf{P}^{-1} \mathbf{F}(\mathbf{x}) \\ \tilde{\mathbf{I}}(\mathbf{x}) &= \mathbf{T}^* \mathbf{P}^{-1}\end{aligned}$$

Soit  $f_\varphi$  le second membre. Les équations sont données sur le domaine PML par :

$$\left( \mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) + \bar{\mathbf{A}}(\mathbf{x}) + \bar{\mathbf{B}}(\mathbf{x}) \left( \mathbf{p} \tilde{\mathbf{I}}(\mathbf{x}) + \tilde{\mathbf{F}}(\mathbf{x}) \right)^{-1} \right) \varphi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}) = f_\varphi(\mathbf{x}).$$

Posons  $\psi(\mathbf{x}) = \left( \mathbf{p} \tilde{\mathbf{I}}(\mathbf{x}) + \tilde{\mathbf{F}}(\mathbf{x}) \right)^{-1} \varphi(\mathbf{x})$ . Viennent ainsi les Equations de Maxwell sur  $\mathbb{R}^3$  :

$$\begin{cases} \mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) \varphi(\mathbf{x}) + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}) = f_\varphi(\mathbf{x}) \\ \mathbf{p} \tilde{\mathbf{I}}(\mathbf{x}) \psi(\mathbf{x}) + \tilde{\mathbf{F}}(\mathbf{x}) \psi(\mathbf{x}) = \varphi(\mathbf{x}). \end{cases}$$

En utilisant la transformée de Laplace inverse, on obtient in fine sur  $\mathbb{R}^3 \times \mathbb{R}^+$

$$\begin{cases} \bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}, t) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}, t) + \begin{pmatrix} 0 & 0 & -\partial_y \\ 0 & 0 & \partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} \varphi(\mathbf{x}, t) = f_\varphi(\mathbf{x}, t) \\ \tilde{\mathbf{I}}(\mathbf{x}) \frac{\partial \psi(\mathbf{x}, t)}{\partial t} + \tilde{\mathbf{F}}(\mathbf{x}) \psi(\mathbf{x}, t) = \varphi(\mathbf{x}, t). \end{cases}$$

On retrouve, là encore, un système de Friedrichs, et la partie spatiale de l'opérateur n'est pas modifiée. Ces deux points garantissent donc une implémentation numérique simple, et laisse le choix de la méthode d'approximation spatiale.  $\square$

En 2001, dans [20], cette approche de la théorie PML avec des variétés pseudo riemanniennes à fibrés tangent et cotangent complexes avait déjà été abordée pour des problèmes harmoniques. Les différences majeures de ce travail avec celui présenté dans cette thèse résident en quatre points.

1. Pour la restriction à un domaine borné, nous avons, dans le cas harmonique, la potentielle présence de fréquences de résonances (rappel :  $\mathbf{p}$  appartient à un domaine d'holomorphie  $\mathbf{H}$  privé de  $\mathbf{S}$  ensemble localement fini pouvant être vide. Lassas *et al.* n'ont pas cette restriction.

2. Cette différence se paye : leur théorie implique de placer la frontière artificielle « assez loin » du domaine  $\mathbf{D}$ . Dans notre cas, elle peut être placée aussi près que l'on veut.
3. Lassas *et al.* requierent une hypothèse supplémentaire pour définir leurs variétés pseudo riemanniennes absorbantes : la  $\eta$ -euclidianité asymptotique de la métrique. Autrement dit, ils imposent une stricte convexité et que la métrique et la connexion soient majorées asymptotiquement par une métrique euclidienne. Nous n'avons pas besoin de cette contrainte.
4. Dernier point : nous proposons des formulations temporelles. Leur travail se limitait en effet aux formulations harmoniques.

## 2.7 Equation des Ondes

Soit  $\varphi(\mathbf{x}, t) = (u(\mathbf{x}, t), v(\mathbf{x}, t), \rho(\mathbf{x}, t))^T$  solution du problème à coefficients constants bidimensionnel pour l'aéroacoustique à l'ordre 1 suivant :

$$\frac{\partial}{\partial t} \varphi(\mathbf{x}, t) + \begin{pmatrix} \mathbf{O} & \nabla \\ \nabla^T & 0 \end{pmatrix} \varphi(\mathbf{x}, t) = 0. \quad (2.32)$$

Montrons que ce type d'Equation admet bien une formulation intrinsèque et établissons les Equations PML.

**Propriété 2.49** (Formulation intrinsèque)

Soit la 2-forme  $W$  et les 1-formes  $(U, V)$  suivantes :

$$\begin{cases} U = -vdx + udy \\ V = udx + vdy \\ W = \rho dx \wedge dy \end{cases}$$

L'équation harmonique de (2.32) écrite extrinsèquement

$$\mathbf{p} \varphi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix} \varphi(\mathbf{x}) = 0$$

peut être écrite sous la forme intrinsèque suivante :

$$\mathbf{p} \begin{pmatrix} W \\ V \end{pmatrix} + \mathbf{d} \begin{pmatrix} U \\ \rho \end{pmatrix} = 0. \quad (2.33)$$

**Démonstration** : comme  $U = -vdx + udy$ , on a :

$$\begin{aligned} \mathbf{d}U &= \mathbf{d}(-vdx + udy) \\ &= -dv \wedge dx + du \wedge dy \\ &= (\partial_y v + \partial_x u) dx \wedge dy. \end{aligned}$$

De plus,  $\mathbf{p}W + \mathbf{d}U = 0 \Leftrightarrow (\mathbf{p}\rho) dx \wedge dy + (\partial_y v + \partial_x u) dx \wedge dy = 0$ . D'où :

$$(\mathbf{p}\rho + \partial_y v + \partial_x u) dx \wedge dy = 0. \quad (2.34)$$

Par ailleurs,  $d\rho = \partial_x \rho dx + \partial_y \rho dy$  et  $\mathbf{p}V + d\rho = 0$ . Par identification de coefficients sur la base  $(x, y)$  :

$$(\mathbf{p}u + \partial_x \rho) dx + (\mathbf{p}v + \partial_y \rho) dy = 0dx + 0dy.$$

On a donc bien équivalence entre la formulation extrinsèque et la formulation intrinsèque.

### **Théorème 2.50**

La matrice de la formulation PML pour l'Equation des Ondes est :

$$\begin{aligned} \mathbf{K}(\mathbf{p}, \mathbf{x}) &= \begin{pmatrix} \frac{1}{\sqrt{g}} g_{ij} & \mathbf{0}_2 \\ 0 & \sqrt{g} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\det(\mathbf{J}(\mathbf{p}, \mathbf{x}))} \mathbf{J}(\mathbf{p}, \mathbf{x})^T \mathbf{J}(\mathbf{p}, \mathbf{x}) & \mathbf{0}_2 \\ 0 & \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) \end{pmatrix}, \end{aligned}$$

où  $\mathbf{J}(\mathbf{p}, \mathbf{x})$  est la matrice jacobienne du changement de variable complexe.

**Démonstration** soit la formulation extrinsèque de (2.32) dans la base  $(x, y)$  de  $\mathbb{R}^2$  exprimée plus haut. Soit une variété  $\mathbf{M}$  réelle à fibrés tangent et cotangent complexes que l'on immerge dans  $\mathbb{C}^2$ . Les nouvelles inconnues seront notées  $(\tilde{x}, \tilde{y})$ . Soit  $\mathbf{J}(\mathbf{p}, \mathbf{x})$  est la matrice jacobienne du changement de variable complexe.

Soit  $\tilde{u}, \tilde{v}, \tilde{\rho}$  les inconnues exprimées dans la base  $(x, y)$  et  $\tilde{\tilde{u}}, \tilde{\tilde{v}}, \tilde{\tilde{\rho}}$  dans la base  $(\tilde{x}, \tilde{y})$ .

$$\begin{aligned} \tilde{U} &= -\tilde{v} d\tilde{x} + \tilde{u} d\tilde{y} = -\tilde{v} dx + \tilde{u} dx \\ \tilde{W} &= \tilde{\rho} d\tilde{x} \wedge d\tilde{y} = \tilde{\rho} dx \wedge dy. \end{aligned}$$

En notant  $z_{ij}$  les coefficients de la matrice  $\mathbf{J}(\mathbf{p}, \mathbf{x})$ , il vient que :

$$\begin{pmatrix} d\tilde{x} \\ d\tilde{y} \end{pmatrix} = \mathbf{J}(\mathbf{p}, \mathbf{x}) \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} z_{11} dx + z_{12} dy \\ z_{22} dx + z_{21} dy \end{pmatrix}$$

On en déduit que  $\tilde{\tilde{\rho}} = (z_{11}z_{22} - z_{12}z_{21}) \tilde{\rho}$  et  $\begin{pmatrix} \tilde{\tilde{u}} \\ \tilde{\tilde{v}} \end{pmatrix} = \begin{pmatrix} z_{22} & -z_{12} \\ -z_{21} & z_{11} \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}$ . Par ailleurs,  $\begin{pmatrix} z_{22} & -z_{12} \\ -z_{21} & z_{11} \end{pmatrix}^{-1} = \frac{1}{z_{11}z_{22} - z_{12}z_{21}} \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix}$  d'où :

$$\begin{cases} \begin{pmatrix} \tilde{\tilde{u}} \\ \tilde{\tilde{v}} \end{pmatrix} = \frac{1}{\det(\mathbf{J}(\mathbf{p}, \mathbf{x}))} \mathbf{J}(\mathbf{p}, \mathbf{x}) \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} \\ \tilde{\tilde{\rho}} = \frac{1}{\det(\mathbf{J}(\mathbf{p}, \mathbf{x}))} \tilde{\rho}. \end{cases}$$

En utilisant l'écriture intrinsèque de l'équation des ondes (2.33) dans la base  $(\tilde{x}, \tilde{y})$ , on obtient :

$$\begin{aligned} \mathbf{p}\tilde{W} + \mathbf{d}\tilde{U} &= \mathbf{p}\tilde{\rho} dx \wedge dy + (\partial_y \tilde{v} + \partial_x \tilde{u}) dx \wedge dy = \mathbf{p}\tilde{\rho} + [\partial_x \quad \partial_y] \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = 0 \\ \mathbf{p}\tilde{V} + \mathbf{d}\tilde{\rho} &= \mathbf{p}(\tilde{u}d\tilde{x} + \tilde{v}d\tilde{y}) + \partial_{\tilde{x}}\tilde{\rho}d\tilde{x} + \partial_{\tilde{y}}\tilde{\rho}d\tilde{y} \\ &= \mathbf{p} \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} + [\partial_x \quad \partial_y]^T \tilde{\rho} \\ &= \mathbf{p} \frac{1}{\det(\mathbf{J}(\mathbf{p}, \mathbf{x}))} \mathbf{J}(\mathbf{p}, \mathbf{x}) \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} + [\partial_x \quad \partial_y]^T \frac{1}{\det(\mathbf{J}(\mathbf{p}, \mathbf{x}))} \tilde{\rho} = 0. \end{aligned}$$

Comme  $\begin{bmatrix} \partial_{\tilde{x}} \\ \partial_{\tilde{y}} \end{bmatrix} = \mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} \begin{bmatrix} \partial_x \\ \partial_y \end{bmatrix}$  il vient que :

$$\mathbf{p} \begin{pmatrix} \frac{1}{\det(\mathbf{J}(\mathbf{p}, \mathbf{x}))} \mathbf{J}(\mathbf{p}, \mathbf{x})^T \mathbf{J}(\mathbf{p}, \mathbf{x}) & \mathbf{0}_2 \\ 0 & \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\rho} \end{pmatrix} + \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\rho} \end{pmatrix} = 0 \quad (2.35)$$

**Remarque** : une autre solution aurait consisté à écrire l'opérateur du premier ordre des ondes sous la forme  $\mathcal{W} = \mathbf{p}\mathbf{I}_{n+1} + \begin{pmatrix} \mathbf{0}_{n+1} & \nabla \\ \nabla^T & 0 \end{pmatrix}$ . A partir de la formulation intrinsèque ci dessus, on obtient que

$$\mathcal{W} = \begin{pmatrix} \star^{-1} \mathbf{p} & \mathbf{d} \\ \mathbf{d} & \star \mathbf{p} \end{pmatrix} \quad (2.36)$$

où les inconnues appartiennent à  $\Omega^{n-1} \times \Omega^0$ . En procédant alors de la même façon que pour les Equations de Maxwell, on retrouve le résultat ci dessus.

**Remarque** : en notant  $(f, g)$  le second membre pour l'équation des ondes, et  $(u, \rho)$  les inconnues, si  $\delta f = 0$  sur  $\Omega$ , on retrouve bien l'équation des ondes scalaires. En effet, comme  $\rho$  est une 0-forme,  $\delta \rho = 0$  donc  $\Delta^0 \rho = (\delta \mathbf{d} + \mathbf{d} \delta) \rho = \mathbf{d} \rho$  d'où :

$$\mathbf{p}^2 \rho + \delta \mathbf{d} \rho = \star g = \mathbf{p}^2 \rho - \Delta^0 \rho.$$

## 2.7.1 Formulation étoilée

Rappelons le changement de variable (2.24)

$$\begin{cases} \hat{\rho}(\mathbf{x}) = k \rho + (1 - k) \rho_0(\theta), \\ \tilde{\rho}(\mathbf{x}) = \hat{\rho}(\mathbf{x}) + \frac{1}{\mathbf{p}} f(\hat{\rho}(\mathbf{x}) - R). \end{cases}$$

**Proposition 2.51**

La matrice PML harmonique associée à cette formulation pour l'Equation des Ondes est donnée par :

$$\mathbf{p} \mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{Q}^T \begin{pmatrix} \frac{k \mathbf{p}(\mathbf{p} + \alpha)}{b \mathbf{p} + \beta} & \frac{a \mathbf{p} + \alpha}{b \mathbf{p} + \beta} & 0 \\ \frac{a \mathbf{p} + \alpha}{b \mathbf{p} + \beta} & \frac{b \mathbf{p}(\mathbf{p} + \beta)}{k \mathbf{p} + \alpha} + \frac{a^2 \mathbf{p}(\mathbf{p} + \alpha)}{k b \mathbf{p} + \beta} & 0 \\ 0 & 0 & k b (\mathbf{p} + \alpha)(\mathbf{p} + \beta) \end{pmatrix} \mathbf{Q},$$

où  $\alpha = f'(\widehat{\rho}(\mathbf{x}) - R)$ ,  $\beta = \frac{f(\widehat{\rho}(\mathbf{x}) - R)}{\widehat{\rho}(\mathbf{x})}$ ,  $a = 2(1 - k) \frac{\rho'_0(\theta)}{\rho}$ ,  $b = \frac{\widehat{\rho}(\mathbf{x})}{\rho}$  et  $\mathbf{Q}$  la matrice orthogonale de rotation.

**Démonstration :** soit  $\mathbf{J}(\mathbf{p}, \mathbf{x})$  la matrice du jacobien de la transformation (2.24). A la différence des Equations de Maxwell écrites en tridimensionnel, ici,  $\mathbf{x} \in \mathbb{R}^2$ , d'où :

$$\mathbf{J}(\mathbf{p}, \mathbf{x}) = \frac{\mathbf{D}(\tilde{x}, \tilde{y})}{\mathbf{D}(x, y)} = \frac{\mathbf{D}(\tilde{x}, \tilde{y})}{\mathbf{D}(\tilde{\rho}, \theta)} \frac{\mathbf{D}(\tilde{\rho}, \theta)}{\mathbf{D}(\widehat{\rho}, \theta)} \frac{\mathbf{D}(\widehat{\rho}, \theta)}{\mathbf{D}(\rho, \theta)} \frac{\mathbf{D}(\rho, \theta)}{\mathbf{D}(x, y)}.$$

$$\frac{\mathbf{D}(\tilde{\rho}, \theta, z)}{\mathbf{D}(\widehat{\rho}, \theta, z)} = \begin{pmatrix} 1 + \frac{1}{\mathbf{p}} f'(\widehat{\rho}(\mathbf{x}) - R) & (1 - k) \rho'_0(\theta) & \left(1 + \frac{1}{\mathbf{p}} f'(\widehat{\rho}(\mathbf{x}) - R)\right) & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\frac{\mathbf{D}(\widehat{\rho}, \theta, z)}{\mathbf{D}(\rho, \theta, z)} = \begin{pmatrix} k & (1 - k) \rho'_0(\theta) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Soit  $\mathbf{Q} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$  matrice unitaire, donc  $\mathbf{Q}^{-1} = \mathbf{Q}^T$ . Comme précédemment, notons  $\alpha = f'(\widehat{\rho}(\mathbf{x}) - R)$ ,  $\beta = \frac{f(\widehat{\rho}(\mathbf{x}) - R)}{\widehat{\rho}(\mathbf{x})}$ ,  $a = 2(1 - k) \frac{\rho'_0(\theta)}{\rho}$  et  $b = \frac{\widehat{\rho}(\mathbf{x})}{\rho}$ . Alors :

$$\mathbf{J}(\mathbf{p}, \mathbf{x}) = \mathbf{Q} \begin{pmatrix} 1 + \frac{1}{\mathbf{p}} \alpha & 0 \\ 0 & 1 + \frac{1}{\mathbf{p}} \beta \end{pmatrix} \begin{pmatrix} k & a \\ 0 & b \end{pmatrix} \mathbf{Q}^T$$

$$\det(\mathbf{J}(\mathbf{p}, \mathbf{x})) = k b \left(1 + \frac{1}{\mathbf{p}} \alpha\right) \left(1 + \frac{1}{\mathbf{p}} \beta\right).$$

Etant donné que la matrice de la formulation PML est donnée par (2.35), il faut

former la partie  $2 \times 2$  de la matrice avec le terme

$$\begin{aligned} \frac{1}{\det(\mathbf{J}(\mathbf{p}, \mathbf{x}))} \mathbf{J}(\mathbf{p}, \mathbf{x})^T (\mathbf{J}(\mathbf{p}, \mathbf{x})) &= \mathbf{Q}^T \begin{pmatrix} k & 0 \\ a & b \end{pmatrix} \begin{pmatrix} \frac{\mathbf{p} + \alpha}{\mathbf{p} + \beta} & 0 \\ 0 & \frac{\mathbf{p} + \beta}{\mathbf{p} + \alpha} \end{pmatrix} \begin{pmatrix} k & a \\ 0 & b \end{pmatrix} \mathbf{Q} \\ &= \mathbf{Q}^T \begin{pmatrix} \frac{k \mathbf{p}(\mathbf{p} + \alpha)}{b \mathbf{p} + \beta} & \frac{a \mathbf{p} + \alpha}{b \mathbf{p} + \beta} \\ \frac{a \mathbf{p} + \alpha}{b \mathbf{p} + \beta} & \frac{b \mathbf{p}(\mathbf{p} + \beta)}{k \mathbf{p} + \alpha} + \frac{a^2 \mathbf{p}(\mathbf{p} + \alpha)}{k b \mathbf{p} + \beta} \end{pmatrix} \mathbf{Q} \end{aligned}$$

En notant désormais  $\mathbf{Q}$  la matrice de rotation dans  $\mathbb{R}^3$  et en rajoutant pour le terme (3, 3) le déterminant du jacobien, on obtient le résultat proposé.  $\square$

**Proposition 2.52**

La matrice PML associée à cette formulation pour l'Equation des Ondes peut être décomposée sous la forme :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{Q} \mathbf{P}(\mathbf{x}) \left( \mathbf{p} \mathbf{I}(\mathbf{x}) + \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x}) \left( \mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}) \right)^{-1} \right) \mathbf{P}^T(\mathbf{x}) \mathbf{Q}^T,$$

où

$$\mathbf{P}(\mathbf{x}) = \begin{pmatrix} k & -2(1-k) \frac{\rho'_0(\theta)}{\rho} & 0 \\ 0 & \frac{\widehat{\rho}(\mathbf{x})}{\rho} & 0 \\ 0 & 0 & \kappa(\mathbf{x}) \end{pmatrix} \quad \mathbf{I}(\mathbf{x}) = \frac{1}{\kappa(\mathbf{x})} \mathbf{I}_3,$$

$$\mathbf{A}(\mathbf{x}) = \frac{1}{\kappa(\mathbf{x})} \text{diag} \left( \alpha(\mathbf{x}) - \beta(\mathbf{x}), \beta(\mathbf{x}) - \alpha(\mathbf{x}), \alpha(\mathbf{x}) + \beta(\mathbf{x}) \right)$$

$$\mathbf{B}(\mathbf{x}) = \frac{1}{\kappa(\mathbf{x})} \text{diag} \left( \alpha(\mathbf{x})(\alpha(\mathbf{x}) - \beta(\mathbf{x})), \beta(\mathbf{x})(\beta(\mathbf{x}) - \alpha(\mathbf{x})), \alpha(\mathbf{x}) \beta(\mathbf{x}) \right)$$

$$\mathbf{F}(\mathbf{x}) = \text{diag} \left( \alpha(\mathbf{x}), \beta(\mathbf{x}), 0 \right).$$

avec  $\kappa(\mathbf{x}) = \left( k \frac{\widehat{\rho}(\mathbf{x})}{\rho(\mathbf{x})} \right)^{-1}$ ,  $\alpha(\mathbf{x}) = f'(\widehat{\rho}(\mathbf{x}) - R)$ ,  $\beta(\mathbf{x}) = \frac{f(\widehat{\rho}(\mathbf{x}) - R)}{\widehat{\rho}(\mathbf{x})}$  et  $\mathbf{Q}$  matrice de rotation.

**Démonstration :** Soit  $\alpha(\mathbf{x}) = f'(\widehat{\rho}(\mathbf{x}) - R)$  et  $\beta(\mathbf{x}) = \frac{f(\widehat{\rho}(\mathbf{x}) - R)}{\widehat{\rho}(\mathbf{x})}$ .

$$\begin{aligned} \frac{\mathbf{p}(\mathbf{p} + \alpha(\mathbf{x}))}{\mathbf{p} + \beta(\mathbf{x})} &= \mathbf{p} + \alpha(\mathbf{x}) - \beta(\mathbf{x}) + \frac{\beta(\mathbf{x})(\beta(\mathbf{x}) - \alpha(\mathbf{x}))}{\mathbf{p} + \beta(\mathbf{x})} \\ \frac{\mathbf{p}(\mathbf{p} + \beta(\mathbf{x}))}{\mathbf{p} + \alpha(\mathbf{x})} &= \mathbf{p} + \beta(\mathbf{x}) - \alpha(\mathbf{x}) + \frac{\alpha(\mathbf{x})(\alpha(\mathbf{x}) - \beta(\mathbf{x}))}{\mathbf{p} + \alpha(\mathbf{x})} \end{aligned}$$

Posons  $\kappa(\mathbf{x}) = \left( k \frac{\widehat{\rho}(\mathbf{x})}{\rho(\mathbf{x})} \right)^{-1}$ . La matrice de la formulation PML peut se décomposer en éléments simples sous la forme énoncée.  $\square$

**Propriété 2.53**

Les Equations des Ondes peuvent être relevées sur  $\mathbb{R}^3 \times \mathbb{R}^+$  sous la forme :

$$\begin{cases} \bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}, t) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}, t) + \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix} \varphi(\mathbf{x}, t) = 0 \\ \mathbf{T}(\mathbf{x})^{T^{-1}} \frac{\partial \psi(\mathbf{x}, t)}{\partial t} + \mathbf{T}(\mathbf{x})^{T^{-1}} \mathbf{F}(\mathbf{x}) \psi(\mathbf{x}, t) = \varphi(\mathbf{x}, t). \end{cases}$$

avec les matrices :

$$\begin{aligned} \mathbf{T}(\mathbf{x}) &= \mathbf{Q}\mathbf{P}(\mathbf{x}) & \bar{\mathbf{I}}(\mathbf{x}) &= \mathbf{T}(\mathbf{x})\mathbf{I}(\mathbf{x})\mathbf{T}(\mathbf{x})^T \\ \bar{\mathbf{A}}(\mathbf{x}) &= \mathbf{T}(\mathbf{x})\mathbf{A}(\mathbf{x})\mathbf{T}(\mathbf{x})^T & \bar{\mathbf{B}}(\mathbf{x}) &= \mathbf{T}(\mathbf{x})\mathbf{B}(\mathbf{x})\mathbf{T}(\mathbf{x})^T \end{aligned}$$

**Démonstration** : la démonstration se conduit de manière identique à celle pour les Equations de Maxwell. Soit  $\varphi$  la solution de l'Equation des Ondes sur le domaine PML :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) \varphi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix} \varphi(\mathbf{x}) = 0.$$

Posons  $\mathbf{T}(\mathbf{x}) = \mathbf{Q}\mathbf{P}(\mathbf{x})$  et introduisons la notation barre qui signifiera  $\mathbf{T}(\mathbf{x}) \bullet \mathbf{T}(\mathbf{x})^T$ . Posons  $\psi(\mathbf{x}) = \left( \mathbf{p} \mathbf{T}(\mathbf{x})^{T^{-1}} + \mathbf{T}(\mathbf{x})^{T^{-1}} \mathbf{F}(\mathbf{x}) \right)^{-1} \varphi(\mathbf{x})$ . D'où sur  $\mathbb{R}^3$  :

$$\begin{cases} \mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) \varphi(\mathbf{x}) + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix} \varphi(\mathbf{x}) = 0 \\ \mathbf{p} \mathbf{T}(\mathbf{x})^{T^{-1}} \psi(\mathbf{x}) + \mathbf{T}(\mathbf{x})^{T^{-1}} \mathbf{F}(\mathbf{x}) \psi(\mathbf{x}) = \varphi(\mathbf{x}). \end{cases}$$

Par transformée de Laplace inverse, on obtient le résultat proposé.  $\square$

**2.7.2 Formulation hybride**

Le changement est cette fois ci donné par (2.24)

$$\begin{cases} \hat{x} = x + (\kappa - 1) \mathbf{inf}(x, y) \\ \hat{y} = y + (\kappa - 1) \mathbf{inf}(x, y) \end{cases} \quad \begin{cases} \tilde{x} = \hat{x} + \frac{1}{\mathbf{p}} \int \sigma_x(\tau) d\tau \\ \tilde{y} = \hat{y} + \frac{1}{\mathbf{p}} \int \sigma_y(\tau) d\tau \end{cases}$$

**Proposition 2.54**

La matrice PML harmonique associée à cette formulation pour l'Equation des Ondes est donnée par :

$$\mathbf{p}\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{T}(\mathbf{x}) \begin{pmatrix} \frac{\mathbf{p}(\mathbf{p} + \sigma_x(\mathbf{x}))}{\mathbf{p} + \sigma_y(\mathbf{x})} & 0 & 0 \\ 0 & \frac{\mathbf{p}(\mathbf{p} + \sigma_y(\mathbf{x}))}{\mathbf{p} + \sigma_x(\mathbf{x})} & 0 \\ 0 & 0 & \frac{(\mathbf{p} + \sigma_x(\mathbf{x}))(\mathbf{p} + \sigma_y(\mathbf{x}))}{\mathbf{p}} \end{pmatrix} \mathbf{T}(\mathbf{x})^T,$$

où la matrice triangulaire  $\mathbf{T}(\mathbf{x})$  dépend de la partie du domaine sur laquelle on la calcule.

**Démonstration** : soit  $\mathbf{J}(\mathbf{p}, \mathbf{x})$  la matrice du jacobien de la transformation (2.24).

$$\mathbf{J}(\mathbf{p}, \mathbf{x}) = \frac{\mathbf{D}(\tilde{x}, \tilde{y})}{\mathbf{D}(\hat{x}, \hat{y})} \frac{\mathbf{D}(\hat{x}, \hat{y})}{\mathbf{D}(x, y)}.$$

Comme pour la formulation des Equations de Maxwell, introduisons  $\mathbf{T}_0$  la matrice triangulaire associée au changement de variable chapeau qui dépend de la partie du domaine sur laquelle on la calcule : sur  $\mathbf{D}_{nc}^+$ ,  $\mathbf{T}_0^+(\mathbf{x}) = \begin{pmatrix} 1 & \kappa - 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  et sur  $\mathbf{D}_{nc}^-$ ,  $\mathbf{T}_0^-(\mathbf{x}) = \begin{pmatrix} \kappa & 0 & 0 \\ \kappa - 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ . On désignera également par  $\mathbf{T}_0 := \mathbf{T}_0(1 : 2, 1 : 2)$

dans les calculs suivants. Comme on ne travaille qu'avec des matrices carrées, cette surcharge de notation ne sera pas pénalisante pour la lisibilité des calculs. On a alors :

$$\begin{aligned} \mathbf{J}(\mathbf{p}, \mathbf{x}) &= \begin{pmatrix} 1 + \frac{1}{\mathbf{p}} \sigma_x(\mathbf{x}) & 0 \\ 0 & 1 + \frac{1}{\mathbf{p}} \sigma_y(\mathbf{x}) \end{pmatrix} \mathbf{T}_0(\mathbf{x}) \\ \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) &= \kappa \left( 1 + \frac{1}{\mathbf{p}} \sigma_x(\mathbf{x}) \right) \left( 1 + \frac{1}{\mathbf{p}} \sigma_y(\mathbf{x}) \right) \end{aligned}$$

Posons  $\mathbf{T}(\mathbf{x}) = (\mathbf{T}_0(\mathbf{x}))^{-1}$ , c'est-à-dire sur  $\mathbf{D}_{nc}^+$ ,  $\mathbf{T}^+(\mathbf{x}) = \frac{1}{\kappa} \begin{pmatrix} 1 & 1 - \kappa & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa \end{pmatrix}$  et sur

$\mathbf{D}_{nc}^-$ ,  $\mathbf{T}^-(\mathbf{x}) = \frac{1}{\kappa} \begin{pmatrix} \kappa & 0 & 0 \\ 1 - \kappa & 1 & 0 \\ 0 & 0 & \kappa \end{pmatrix}$ . La matrice PML s'obtient alors comme dans la sous section précédente.  $\square$

**Proposition 2.55**

La matrice PML associée à cette formulation pour l'Equation des Ondes peut être décomposée sous la forme :

$$\mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{T}(\mathbf{x}) (\mathbf{p} \mathbf{I} + \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x}) (\mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}))^{-1}) \mathbf{T}(\mathbf{x})^T,$$

$$\text{où } \mathbf{I} = \frac{1}{\kappa} \mathbf{I}_3$$

$$\mathbf{F}(\mathbf{x}) = \text{diag}(\sigma_y(\mathbf{x}), \sigma_x(\mathbf{x}), 0)$$

$$\mathbf{A}(\mathbf{x}) = \frac{1}{\kappa} \text{diag}(\sigma_x(\mathbf{x}) - \sigma_y(\mathbf{x}), \sigma_y(\mathbf{x}) - \sigma_x(\mathbf{x}), \sigma_x(\mathbf{x}) + \sigma_y(\mathbf{x}))$$

$$\mathbf{B}(\mathbf{x}) = \frac{1}{\kappa} \text{diag}(\sigma_y(\mathbf{x})(\sigma_y(\mathbf{x}) - \sigma_x(\mathbf{x})), \sigma_x(\mathbf{x})(\sigma_x(\mathbf{x}) - \sigma_y(\mathbf{x})), \sigma_x(\mathbf{x}) \sigma_y(\mathbf{x}))$$

et la matrice triangulaire  $\mathbf{T}(\mathbf{x})$  dépend de la partie du domaine considérée : sur  $\mathbf{D}_{nc}^+$ ,

$$\mathbf{T}^+ = \begin{pmatrix} 1 & 1 - \kappa & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa \end{pmatrix} \text{ et sur } \mathbf{D}_{nc}^-, \quad \mathbf{T}^- = \begin{pmatrix} \kappa & 0 & 0 \\ 1 - \kappa & 1 & 0 \\ 0 & 0 & \kappa \end{pmatrix}.$$

**Démonstration :**

$$\begin{aligned} \frac{\mathbf{p}(\mathbf{p} + \sigma_x(\mathbf{x}))}{\mathbf{p} + \sigma_y(\mathbf{x})} &= \mathbf{p} + \sigma_x(\mathbf{x}) - \sigma_y(\mathbf{x}) + \frac{\sigma_y(\mathbf{x})(\sigma_y(\mathbf{x}) - \sigma_x(\mathbf{x}))}{\mathbf{p} + \sigma_y(\mathbf{x})} \\ \frac{\mathbf{p}(\mathbf{p} + \sigma_y(\mathbf{x}))}{\mathbf{p} + \sigma_x(\mathbf{x})} &= \mathbf{p} + \sigma_y(\mathbf{x}) - \sigma_x(\mathbf{x}) + \frac{\sigma_x(\mathbf{x})(\sigma_x(\mathbf{x}) - \sigma_y(\mathbf{x}))}{\mathbf{p} + \sigma_x(\mathbf{x})} \end{aligned}$$

La matrice de la formulation PML peut se décomposer en éléments simples sous la forme énoncée.  $\square$

### Propriété 2.56

Les Equations des Ondes peuvent être relevées sur  $\mathbb{R}^3 \times \mathbb{R}^+$  sous la forme :

$$\begin{cases} \bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}, t) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}, t) + \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix} \varphi(\mathbf{x}, t) = 0 \\ \mathbf{T}(\mathbf{x})^{T-1} \frac{\partial \psi(\mathbf{x}, t)}{\partial t} + \mathbf{T}(\mathbf{x})^{T-1} \mathbf{F}(\mathbf{x}) \psi(\mathbf{x}, t) = \varphi(\mathbf{x}, t). \end{cases}$$

avec les matrices :

$$\bar{\mathbf{I}}(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \mathbf{I}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T \quad \bar{\mathbf{A}}(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \mathbf{A}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T \quad \bar{\mathbf{B}}(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \mathbf{B}(\mathbf{x}) \mathbf{T}(\mathbf{x})^T$$

**Démonstration :** se conduit de la même manière que celle pour les Equations de Maxwell.  $\square$

### 2.7.3 Deuxième formulation hybride

Le changement de variables considéré est (2.25), à savoir :

$$\begin{cases} \hat{x} = x + (\kappa - 1) \inf(x, y) \\ \hat{y} = y + (\kappa - 1) \inf(x, y) \end{cases} \quad \begin{cases} \tilde{x} = \hat{x} + \frac{1}{\mathbf{p}} f(\hat{x} + \hat{y} - R) \\ \tilde{y} = \hat{y} + \frac{1}{\mathbf{p}} f(\hat{x} + \hat{y} - R). \end{cases}$$

### Proposition 2.57

La matrice PML harmonique associée à cette formulation pour l'Equation des Ondes est donnée par :

$$\mathbf{p} \mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{T}(\mathbf{x}) \begin{pmatrix} \left( \frac{\mathbf{p} + \sigma(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})} \right)^2 & \frac{\sigma(\mathbf{x})^2}{\mathbf{p} + 2\sigma(\mathbf{x})} & 0 \\ \frac{\sigma(\mathbf{x})^2}{\mathbf{p} + 2\sigma(\mathbf{x})} & \left( \frac{\mathbf{p} + \sigma(\mathbf{x})}{\mathbf{p} + 2\sigma(\mathbf{x})} \right)^2 & 0 \\ 0 & 0 & \kappa(\mathbf{p} + 2\sigma(\mathbf{x})) \end{pmatrix} \mathbf{T}^T(\mathbf{x}),$$

où  $\sigma$  désigne la dérivée de la fonction  $f$ , et la matrice triangulaire  $\mathbf{T}(\mathbf{x})$  est la même que pour la formulation précédente.

**Démonstration** : elle se conduit comme pour la formulation précédente. Le jacobien de la transformation tilde contient la seule modification.

Posons  $\sigma(\mathbf{x}) = f'(\hat{x} + \hat{y} - R)$ . Alors  $\frac{D(\tilde{x}, \tilde{y})}{D(\hat{x}, \hat{y})} = \mathbf{I}_2 + \frac{1}{\mathbf{p}} \begin{pmatrix} \sigma(\mathbf{x}) & \sigma(\mathbf{x}) \\ \sigma(\mathbf{x}) & \sigma(\mathbf{x}) \end{pmatrix}$ .

$$\mathbf{J}(\mathbf{p}, \mathbf{x}) = \begin{pmatrix} 1 + \frac{1}{\mathbf{p}}\sigma(\mathbf{x}) & \frac{1}{\mathbf{p}}\sigma(\mathbf{x}) & 0 \\ \frac{1}{\mathbf{p}}\sigma(\mathbf{x}) & 1 + \frac{1}{\mathbf{p}}\sigma(\mathbf{x}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{T}_0(\mathbf{x})$$

$$\det(\mathbf{J}(\mathbf{p}, \mathbf{x})) = \kappa \left( 1 + \frac{2}{\mathbf{p}}\sigma(\mathbf{x}) \right) \quad \boxtimes$$

## 2.7.4 Généralisation

De la même façon que pour les Equations de Maxwell, nous proposons une écriture générale indépendamment de la formulation PML choisie.

### Propriété 2.58

La matrice  $\mathbf{p} \mathbf{K}(\mathbf{p}, \mathbf{x})$  pour l'équation des ondes peut s'écrire sous la forme

$$\begin{pmatrix} \frac{1}{\det \mathbf{T}} \left( \mathbf{T}^T \mathbf{P}^T \begin{pmatrix} \mathbf{p} \frac{\mathbf{p} + \lambda_1}{\mathbf{p} + \lambda_2} & 0 \\ 0 & \mathbf{p} \frac{\mathbf{p} + \lambda_2}{\mathbf{p} + \lambda_1} \end{pmatrix} \mathbf{P}^{*-1} \mathbf{T} \right) & 0 \\ 0 & \det \mathbf{T} \frac{(\mathbf{p} + \lambda_1)(\mathbf{p} + \lambda_2)}{\mathbf{p}} \end{pmatrix}$$

$$\text{où } \mathbf{T} = \frac{D\varphi}{D\mathbf{x}}, \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \text{ et } \mathbf{P}^* = \mathbf{P}^{-1}.$$

**Démonstration** : le calcul de la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$  a été établie à chaque formulation. Néanmoins, étant donné le changement de variable générique

$$\tilde{\mathbf{x}} = \varphi(\mathbf{x}) + \frac{1}{\mathbf{p}}f(\varphi(\mathbf{x})),$$

il est possible d'obtenir une formulation générale. Soit la matrice jacobienne  $\mathbf{J}(\mathbf{p}, \mathbf{x})$ . Elle s'écrit  $\frac{D\tilde{\mathbf{x}}}{D\mathbf{x}} = \left( \mathbf{I} + \frac{1}{\mathbf{p}} \frac{Df}{D\varphi} \right) \frac{D\varphi}{D\mathbf{x}}$ . Comme  $f$  est le gradient d'une fonction convexe par rapport à  $\varphi$ , il existe  $\mathbf{P}$  une matrice unitaire  $\mathbf{P}^* = \mathbf{P}^{-1}$  telle que

$$\frac{Df}{D\varphi} = \mathbf{P}^* \left( \mathbf{I} + \frac{1}{\mathbf{p}} \mathbf{D} \right) \mathbf{P}$$

où  $\mathbf{D}$  est une matrice diagonale. Soit  $\lambda_i, i = 1, 2$  les valeurs propres de cette matrice. Posons  $\mathbf{T} = \frac{D\varphi}{D\mathbf{x}}$ , alors  $\mathbf{J}(\mathbf{p}, \mathbf{x}) = \mathbf{P}^* \left( \mathbf{I} + \frac{1}{\mathbf{p}} \mathbf{D} \right) \mathbf{P} \mathbf{T}$ . La transformée de Hodge associée est  $\star = \det(\mathbf{J}(\mathbf{p}, \mathbf{x})) \mathbf{J}(\mathbf{p}, \mathbf{x})^{-1} (\mathbf{J}(\mathbf{p}, \mathbf{x})^T)^{-1}$ . L'opérateur des ondes se traduit

de façon intrinsèque dans la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x}) = \begin{pmatrix} \mathbf{p} \star^{-1} & 0 \\ 0 & \mathbf{p} \star \end{pmatrix}$ .

$$\mathbf{K}(\mathbf{p}, \mathbf{x})(1 : 2, 1 : 2) = \det(\mathbf{J}(\mathbf{p}, \mathbf{x}))^{-1} \mathbf{J}(\mathbf{p}, \mathbf{x})^T \mathbf{J}(\mathbf{p}, \mathbf{x}).$$

$$\mathbf{K}(\mathbf{p}, \mathbf{x})(3, 3) = \det(\mathbf{J}(\mathbf{p}, \mathbf{x})).$$

D'où le résultat proposé. \(\square\)

### **Théorème 2.59**

*Les Equations des Ondes bidimensionnelles dans le milieu PML s'écrivent en instationnaire sous la forme*

$$\bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}, t) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}, t) + \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix} \varphi(\mathbf{x}, t) = 0,$$

et  $\psi$  est solution d'une Equation aux Dérivées Ordinaires :

$$\bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \psi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{F}}(\mathbf{x}) \psi(\mathbf{x}, t) = \varphi(\mathbf{x}, t),$$

avec les matrices déduites de la décomposition précédente

$$\begin{aligned} \bar{\mathbf{I}}(\mathbf{x}) &= \begin{pmatrix} \frac{1}{\det \mathbf{T}} \mathbf{T}^T \mathbf{P}^T \mathbf{P}^{*-1} \mathbf{T} & 0 \\ 0 & \det \mathbf{T} \end{pmatrix} \\ \bar{\mathbf{A}}(\mathbf{x}) &= \begin{pmatrix} \frac{1}{\det \mathbf{T}} \mathbf{T}^T \mathbf{P}^T \mathbf{A}(\mathbf{x}) \mathbf{P}^{*-1} \mathbf{T} & 0 \\ 0 & \det \mathbf{T}(\lambda_1 + \lambda_2) \end{pmatrix} \\ \bar{\mathbf{B}}(\mathbf{x}) &= \begin{pmatrix} \frac{1}{\det \mathbf{T}} \mathbf{T}^T \mathbf{P}^T \mathbf{B}(\mathbf{x}) \mathbf{P}^{*-1} \mathbf{T} & 0 \\ 0 & \det \mathbf{T}(\lambda_1 \lambda_2) \end{pmatrix} \\ \bar{\mathbf{F}}(\mathbf{x}) &= \begin{pmatrix} \frac{1}{\det \mathbf{T}} \mathbf{T}^T \mathbf{P}^T \mathbf{F}(\mathbf{x}) \mathbf{P}^{*-1} \mathbf{T} & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

**Démonstration** : soit  $i \neq j$ . Comme  $\mathbf{p} \frac{\mathbf{p} + \lambda_i}{\mathbf{p} + \lambda_j} = \mathbf{p} + (\lambda_i - \lambda_j) - \frac{(\lambda_i - \lambda_j)\lambda_j}{\mathbf{p} + \lambda_j}$ ,

la matrice  $\begin{pmatrix} \mathbf{p} \frac{\mathbf{p} + \lambda_1}{\mathbf{p} + \lambda_2} & 0 \\ 0 & \mathbf{p} \frac{\mathbf{p} + \lambda_2}{\mathbf{p} + \lambda_1} \end{pmatrix}$  se décompose en éléments simples sous la forme  $\mathbf{p} \mathbf{I}_3 + \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x}) (\mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}))^{-1}$  où

$$\begin{aligned} \mathbf{F}(\mathbf{x}) &= \mathbf{diag}(\lambda_2, \lambda_1) & \mathbf{A}(\mathbf{x}) &= \mathbf{diag}(\lambda_1 - \lambda_2, \lambda_2 - \lambda_1) \\ \mathbf{B}(\mathbf{x}) &= \mathbf{diag}(-(\lambda_1 - \lambda_2)\lambda_2, -(\lambda_2 - \lambda_1)\lambda_1). \end{aligned}$$

Posons :

$$\begin{aligned}\bar{\mathbf{I}}(\mathbf{x}) &= \begin{pmatrix} \frac{1}{\det \mathbf{T}} \mathbf{T}^T \mathbf{P}^T \mathbf{P}^{*-1} \mathbf{T} & 0 \\ 0 & \det \mathbf{T} \end{pmatrix} \\ \bar{\mathbf{A}}(\mathbf{x}) &= \begin{pmatrix} \frac{1}{\det \mathbf{T}} \mathbf{T}^T \mathbf{P}^T \mathbf{A}(\mathbf{x}) \mathbf{P}^{*-1} \mathbf{T} & 0 \\ 0 & \det \mathbf{T}(\lambda_1 + \lambda_2) \end{pmatrix} \\ \bar{\mathbf{B}}(\mathbf{x}) &= \begin{pmatrix} \frac{1}{\det \mathbf{T}} \mathbf{T}^T \mathbf{P}^T \mathbf{B}(\mathbf{x}) \mathbf{P}^{*-1} \mathbf{T} & 0 \\ 0 & \det \mathbf{T}(\lambda_1 \lambda_2) \end{pmatrix} \\ \bar{\mathbf{F}}(\mathbf{x}) &= \begin{pmatrix} \frac{1}{\det \mathbf{T}} \mathbf{T}^T \mathbf{P}^T \mathbf{F}(\mathbf{x}) \mathbf{P}^{*-1} \mathbf{T} & 0 \\ 0 & 0 \end{pmatrix}\end{aligned}$$

Soit  $\varphi$  la solution des Equations des Ondes données sur le domaine PML par :

$$\left( \mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) + \bar{\mathbf{A}}(\mathbf{x}) + \bar{\mathbf{B}}(\mathbf{x}) (\mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) + \bar{\mathbf{F}}(\mathbf{x}))^{-1} \right) \varphi + \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix} \varphi = 0.$$

Posons  $\psi = (\mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) + \bar{\mathbf{F}}(\mathbf{x}))^{-1} \varphi$ . Il s'agit donc de la solution d'une Equation Différentielle Ordinaire. Introduite dans l'équation précédente, on obtient le système régissant les Equations des Ondes sur  $\mathbb{R}^3$  :

$$\begin{cases} \mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) \varphi(\mathbf{x}) + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}) + \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix} \varphi(\mathbf{x}) = 0 \\ \mathbf{p} \bar{\mathbf{I}}(\mathbf{x}) \psi(\mathbf{x}) + \bar{\mathbf{F}}(\mathbf{x}) \psi(\mathbf{x}) = \varphi(\mathbf{x}). \end{cases}$$

Par transformée de Laplace inverse, on obtient sur  $\mathbb{R}^3 \times \mathbb{R}^+$

$$\begin{cases} \bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}, t) + \bar{\mathbf{B}}(\mathbf{x}) \psi(\mathbf{x}, t) + \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ \partial_x & \partial_y & 0 \end{pmatrix} \varphi(\mathbf{x}, t) = 0 \\ \bar{\mathbf{I}}(\mathbf{x}) \frac{\partial \psi(\mathbf{x}, t)}{\partial t} + \bar{\mathbf{F}}(\mathbf{x}) \psi(\mathbf{x}, t) = \varphi(\mathbf{x}, t). \end{cases}$$

Il s'agit bien d'un système de Friedrichs, facilement implémentable, et dont l'existence et l'unicité de la solution sont prouvés via les outils du chapitre 1 de ce mémoire.  $\square$

**Remarque** : comme pour les Equations de Maxwell, la partie spatiale de l'opérateur n'est pas affectée dans le milieu PML.

**Remarque** : le travail explicité ci dessus peut être généralisé en  $n$  dimensions pour l'Equation des Ondes. Il suffit d'écrire l'opérateur de façon intrinsèque (comme dans (2.36)) à savoir  $\mathcal{W} = \mathbf{p} \mathbf{I}_{n+1} + \begin{pmatrix} \mathbf{0}_{n+1} & \nabla \\ \nabla^T & 0 \end{pmatrix} \begin{pmatrix} \star^{-1} \mathbf{p} & \mathbf{d} \\ \mathbf{d} & \star \mathbf{p} \end{pmatrix}$  où les inconnues appartiennent à  $\mathbb{R}^{n-1} \times \mathbb{R}$ .

## 2.8 Etude des formulations PML « naïves »

On pourrait penser généraliser la théorie PML des domaines non convexes chaque fois que l'on peut trouver sur cette frontière un champ de vecteurs « s'écartant » et en écrivant un damping le long de chacun de ces vecteurs. Soit  $\mathbf{x} = (\rho, \theta, z) \in \mathbb{R}^3$ ,  $\varphi$  un  $C^1$  difféomorphisme de  $\mathbb{R}^3$  et  $f$  une fonction  $C^1$  qui vérifie les hypothèses (2.5). Soit  $\rho_0$  une fonction  $C^1$  qui représente la frontière du domaine d'étude  $\Omega$ . Dans le cas d'une géométrie étoilée, le seul paramètre de cette fonction est l'angle. On choisit l'immersion laissant  $\theta$  et  $z$  invariants et telle que :

$$\tilde{\rho} = \rho + \frac{1}{\mathbf{p}} f(\rho - \rho_0(\theta)). \quad (2.37)$$

Nous allons montrer que cette formulation est non correcte.

### Proposition 2.60

La matrice harmonique associée à cette formulation est donnée par  $\mathbf{p} \mathbf{K}(\mathbf{p}, \mathbf{x})$ , soit :

$$\mathbf{Q} \begin{pmatrix} \frac{\mathbf{p}(\mathbf{p} + \alpha(\mathbf{x}))}{\mathbf{p} + \beta(\mathbf{x})} + \frac{\frac{\rho'_0(\theta)^2}{\rho^2} \beta(\mathbf{x})^2 \mathbf{p}}{(\mathbf{p} + \beta(\mathbf{x}))(\mathbf{p} + \alpha(\mathbf{x}))} & \frac{\frac{\rho'_0(\theta)}{\rho} \beta(\mathbf{x}) \mathbf{p}}{\mathbf{p} + \alpha(\mathbf{x})} & 0 \\ \frac{\frac{\rho'_0(\theta)}{\rho} \beta(\mathbf{x}) \mathbf{p}}{\mathbf{p} + \alpha(\mathbf{x})} & \frac{\mathbf{p}(\mathbf{p} + \beta(\mathbf{x}))}{\mathbf{p} + \alpha(\mathbf{x})} & 0 \\ 0 & 0 & \frac{(\mathbf{p} + \beta(\mathbf{x}))(\mathbf{p} + \alpha(\mathbf{x}))}{\mathbf{p}} \end{pmatrix} \mathbf{Q}^T,$$

où  $\alpha(\mathbf{x}) = \frac{f(\rho - \rho_0(\theta))}{\rho}$ ,  $\beta(\mathbf{x}) = f'(\rho - \rho_0(\theta))$  et  $\mathbf{Q} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ .

**Démonstration** : soit  $\mathbf{J}$  la matrice jacobienne. Les nouvelles variables cartésiennes sont données par :  $\tilde{x} = \tilde{\rho} \cos(\theta)$   $\tilde{y} = \tilde{\rho} \sin(\theta)$ .

$$\begin{aligned} \mathbf{J} &= \frac{\mathbf{D}(\tilde{x}, \tilde{y}, \tilde{z})}{\mathbf{D}(x, y, z)} = \frac{\mathbf{D}(\tilde{x}, \tilde{y}, \tilde{z})}{\mathbf{D}(\tilde{\rho}, \theta, z)} \frac{\mathbf{D}(\tilde{\rho}, \theta, z)}{\mathbf{D}(\rho, \theta, z)} \frac{\mathbf{D}(\rho, \theta, z)}{\mathbf{D}(x, y, z)}. \\ \frac{\mathbf{D}(\tilde{\rho}, \theta, z)}{\mathbf{D}(\hat{\rho}, \theta, z)} &= \begin{pmatrix} 1 + \frac{1}{\mathbf{p}} f'(\rho - \rho_0(\theta)) & -\rho'_0(\theta) \frac{1}{\mathbf{p}} f'(\rho - \rho_0(\theta)) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \frac{\mathbf{D}(\tilde{x}, \tilde{y}, \tilde{z})}{\mathbf{D}(\tilde{\rho}, \theta, z)} &= \begin{pmatrix} \cos\theta & -\tilde{\rho} \sin\theta & 0 \\ \sin\theta & \tilde{\rho} \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \tilde{\rho} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \frac{\mathbf{D}(\rho, \theta, z)}{\mathbf{D}(x, y, z)} &= \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ \frac{\rho}{0} & \frac{\rho}{0} & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\rho} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

En posant  $\mathbf{Q} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$  matrice de passage orthogonale entre les coordonnées polaires et cartésiennes, on obtient :

$$\begin{aligned} \mathbf{J} &= \mathbf{Q} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \tilde{\rho} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 + \frac{1}{\mathbf{p}} f'(\rho - \rho_0(\theta)) & -\rho'_0(\theta) \frac{1}{\mathbf{p}} f'(\rho - \rho_0(\theta)) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\rho} & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{Q}^T \\ &= \mathbf{Q} \begin{pmatrix} 1 + \frac{1}{\mathbf{p}} f'(\rho - \rho_0(\theta)) & -\frac{\rho'_0(\theta)}{\rho} \frac{1}{\mathbf{p}} f'(\rho - \rho_0(\theta)) & 0 \\ 0 & \tilde{\rho} & 0 \\ 0 & \frac{\rho}{\mathbf{p}} & 1 \end{pmatrix} \mathbf{Q}^T = \mathbf{Q} \mathbf{F}_0 \mathbf{Q}^T. \end{aligned}$$

$$\mathbf{Com}(\mathbf{J}) = \mathbf{Com}(\mathbf{Q} \mathbf{F}_0 \mathbf{Q}^T) = \mathbf{Com}(\mathbf{Q}) \mathbf{Com}(\mathbf{F}_0) \mathbf{Com}(\mathbf{Q}^T) = \mathbf{Q} \mathbf{Com}(\mathbf{F}_0) \mathbf{Q}^T.$$

$$\frac{\tilde{\rho}}{\rho} = 1 + \frac{1}{\mathbf{p}} \frac{f(\rho - \rho_0(\theta))}{\rho}. \text{ Posons } \alpha(\mathbf{x}) = \frac{f(\rho - \rho_0(\theta))}{\rho} \text{ et } \beta(\mathbf{x}) = f'(\rho - \rho_0(\theta)).$$

$$\begin{aligned} \mathbf{Com}(\mathbf{F}_0) &= \begin{pmatrix} 1 + \frac{1}{\mathbf{p}} \alpha(\mathbf{x}) & 0 & 0 \\ \frac{\rho'_0(\theta)}{\rho} \frac{1}{\mathbf{p}} \beta(\mathbf{x}) & 1 + \frac{1}{\mathbf{p}} \beta(\mathbf{x}) & 0 \\ 0 & 0 & \left(1 + \frac{1}{\mathbf{p}} \alpha(\mathbf{x})\right) \left(1 + \frac{1}{\mathbf{p}} \beta(\mathbf{x})\right) \end{pmatrix} \\ \det(\mathbf{F}_0) &= \left(1 + \frac{1}{\mathbf{p}} \alpha(\mathbf{x})\right) \left(1 + \frac{1}{\mathbf{p}} \beta(\mathbf{x})\right) \\ \mathbf{F}_0^{-1} &= \begin{pmatrix} \frac{\mathbf{p}}{\mathbf{p} + \beta(\mathbf{x})} & \frac{\rho'_0(\theta)}{\rho} \frac{\mathbf{p} \beta(\mathbf{x})}{(\mathbf{p} + \alpha(\mathbf{x})) (\mathbf{p} + \beta(\mathbf{x}))} & 0 \\ 0 & \frac{\mathbf{p}}{\mathbf{p} + \alpha(\mathbf{x})} & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

La matrice de la formulation PML  $\mathbf{p} \mathbf{K}(\mathbf{p}, \mathbf{x})$  est alors donnée par la relation :

$$\mathbf{p} \mathbf{K}(\mathbf{p}, \mathbf{x}) = \mathbf{p} \mathbf{Com}(\mathbf{J}) \mathbf{J}^{-1} = \mathbf{p} \mathbf{Q} \mathbf{Com}(\mathbf{F}_0) \mathbf{F}_0^{-1} \mathbf{Q}^T.$$

D'où le résultat proposé.  $\square$

### Difficulté afférente à cette formulation

On avait effectué l'immersion de  $\mathbf{M}$  dans  $\mathbb{C}^3$  de la façon suivante :

$$\tilde{\mathbf{x}} = \mathbf{x} + \frac{1}{\mathbf{p}} \frac{\mathbf{x}}{\|\mathbf{x}\|} f(\|\mathbf{x}\| - \rho_0(\mathbf{x})), \quad (2.38)$$

avec  $\rho_0(\mathbf{x})$  la distance du point  $\mathbf{x}$  à la frontière du scattered, qui dépend de l'angle  $\theta$  en deux dimensions, et des angles d'Euler en trois dimensions. Soit  $\mathcal{S}^{n-1}$  la sphère

unité de dimension  $(n - 1)$ .  $\mathbb{R}^n$  est toujours difféomorphe à  $\mathcal{S}^{n-1} \times \mathbb{R}^+$ .

Cette immersion ne convient pas puisqu'elle ne donne pas un bon prolongement de la solution élémentaire étant donné qu'elle ne vérifie pas la condition suivante :

$$(\tilde{\mathbf{x}} - \tilde{\mathbf{y}})^2 = 0 \quad \Rightarrow \quad \mathbf{x} = \mathbf{y}. \quad (2.39)$$

L'immersion garantit cependant la décroissance asymptotique. On pourrait s'interroger sur la nécessité de la condition (2.39). Pour réaliser une formulation PML correcte, il faudra étendre la solution à un système de coordonnées curvilignes dans  $\mathbb{R}^n$ , ce qui implique un changement de variables. Pour passer de  $\mathbb{R}^n$  à notre variété  $\mathbf{M}$ , il faut avoir une bijection (difféomorphisme). Il est possible de ne faire qu'un simple changement de variables à condition qu'il soit compatible avec la convolution. Soit  $\mathbf{y} \in \mathbb{C}^n$ . Considérons une carte locale. Il faut garantir sur tout  $\mathbb{C}^n$  et non seulement sur  $\mathbf{M}$ , une structure de translation. La partie réelle ne doit pas être dégénérée, autrement dit :

$$\forall \mathbf{v} \in T_{\mathbf{y}}^{\mathbb{R}} \mathbf{H}, \quad g(\mathbf{v}, \mathbf{v}) = \mathbf{H}^T \mathbf{H} \cdot \mathbf{v} \neq 0, \quad (2.40)$$

avec  $\mathbf{H} \in \mathcal{M}_n(\mathbb{C})$ . Si cette condition est respectée, on sait qu'il existe une fonction  $\alpha$  telle que

$$g_{\mathbf{y}}(\mathbf{v}, \mathbf{v}) \alpha(\mathbf{y}) > 0.$$

On peut alors définir une fonction  $\phi_{\varepsilon} \in C^{\infty}(\mathbb{C}, \mathbb{C})$  analytique pour  $\mathbf{x}$  et  $\mathbf{y}$  telle que :

$$\phi_{\varepsilon} = \frac{1}{4\pi \sqrt{(\tilde{\mathbf{x}} - \tilde{\mathbf{y}})^2 + \xi \varepsilon^2}}.$$

La condition (2.40) est nécessaire pour immerger le changement de variables dans  $\mathbb{C}^n$ . Il est toujours possible de faire un changement de variables mais l'immersion de la variété dans  $\mathbb{C}^n$  est plus délicate. Il faut avoir stabilité du signe de la partie imaginaire de la racine  $\sqrt{(\tilde{\mathbf{x}} - \tilde{\mathbf{y}})^2 + \xi \varepsilon^2}$  et la non nullité en dehors du cas  $\mathbf{x} = \mathbf{y}$ .

Soit  $\mathbf{x}$  un point du domaine PML et  $\mathbf{y} \in \Omega$ . Soit la normale  $\mathbf{n}(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|}$ . Alors  $f(\|\mathbf{y}\| - \rho_0(\mathbf{y})) = 0$ . On souhaite calculer la partie imaginaire de la racine. La partie réelle provient des termes de type  $\mathbf{x}_k \mathbf{y}_k$  uniquement. Soit  $\mathbf{p} = \varepsilon + i\omega$ . Après avoir fixé une direction d'étoilage, on obtient :

$$\begin{aligned} \sum_k \tilde{\mathbf{x}}_k \mathbf{y}_k &= \sum_k \left( \mathbf{x}_k + \frac{1}{\varepsilon + i\omega} \mathbf{n}_k f(\|\mathbf{x}\| - \rho_0(\mathbf{x})) \right) \mathbf{y}_k \\ &= \sum_k \left( \mathbf{x}_k \mathbf{y}_k + \frac{\mathbf{n}_k \mathbf{y}_k \varepsilon}{\varepsilon^2 + \omega^2} f(\|\mathbf{x}\| - \rho_0(\mathbf{x})) + i \mathbf{n}_k \mathbf{y}_k \frac{-\omega}{\varepsilon^2 + \omega^2} f(\|\mathbf{x}\| - \rho_0(\mathbf{x})) \right). \end{aligned}$$

Le problème survient quand on se rapproche du domaine  $\Omega$ . Quand  $\mathbf{n}$  est orthogonale à  $\mathbf{y}$ , la partie imaginaire s'annule. On change donc de signe pour la partie

imaginaire dans le carré. Cependant, il serait possible que celui de la racine reste invariant.  $f$  est une fonction croissante en  $\mathbf{x}$ , plus rapidement que  $\mathbf{x}$ . En l'infini, la partie imaginaire est positive, puis s'annule et enfin devient négative quand  $\mathbf{n} \cdot \mathbf{y} < 0$ . La partie imaginaire change donc de signe et évolue comme le produit scalaire  $\mathbf{n} \cdot \mathbf{y}$ . Elle ne change pas de signe quand sa projection est dans le domaine  $\Omega$  (où la fonction  $f$  est nulle).

Etant donné la nécessité de conserver la partie imaginaire de signe constant, il faut trouver une courbure qui ne passe pas par la coupure. De plus, cette coupure doit intervenir entre  $]0, \pi]$  pour conserver le bon signe. Or avec (2.38) c'est impossible. Avec cette immersion, à l'extérieur du domaine, il existe toujours une fréquence  $\omega$  telle que la partie réelle et la partie imaginaire s'annulent au même endroit  $\mathbf{y}$  appartenant à  $\mathbf{D}^c \cap \mathbf{Conv}_H(\mathbf{D})$ . Le noyau de Green est singulier en  $\mathbf{y}$  : la décroissance est donc impossible à obtenir autour de  $\mathbf{y}$ .

### Les singularités de la fonction de Hankel

Nous allons expliquer pourquoi les résultats numériques obtenus avec cette formulation posent problème. Dans certaines configurations géométriques et selon la position de la source (condition initiale), il était possible de voir apparaître des instabilités, alors que pour d'autres paramètres, la solution semblait tout à fait correcte. Le côté apparemment « aléatoire » de l'apparition de ces problèmes étant surprenant, une étude locale a été entreprise.

Pour l'équation des ondes, nous connaissons le noyau de Schwartz. A l'aide des fonctions de Hankel de type 1 ou 2 selon la convention de signe choisie pour définir la variable de Laplace  $\mathbf{p} = \pm i\omega + 0$ , il est possible d'exprimer la solution en tout point du domaine. Elle est obtenue à l'aide de  $H_0^1(\omega \sqrt{(\tilde{\mathbf{x}} - \mathbf{x}_0)^2 + (\tilde{\mathbf{y}} - \mathbf{y}_0)^2})$ . On appelle  $d = \sqrt{(\tilde{\mathbf{x}} - \mathbf{x}_0)^2 + (\tilde{\mathbf{y}} - \mathbf{y}_0)^2}$  cette racine complexe qui pose clairement un problème de détermination.

On rappelle que l'on définit notre variété pseudo Riemmanienne à l'aide d'une immersion complexe définie par le changement de variable  $\tilde{\rho} = \rho + \frac{1}{\mathbf{p}} f(\rho - \rho_0(\theta))$  où  $f$  est une fonction positive croissante respectant certaines propriétés énoncées plus haut. Alors :

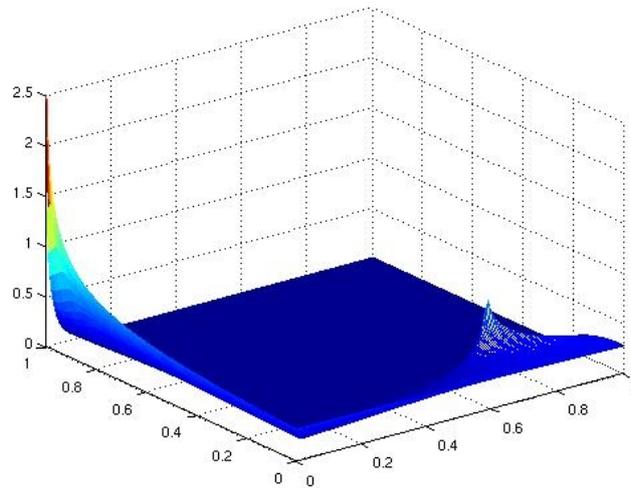
$$\begin{cases} \tilde{x} = \tilde{\rho} \cos(\theta) \\ \tilde{y} = \tilde{\rho} \sin(\theta) \end{cases} \quad \begin{cases} x_0 = \rho_0 \cos(\theta_0) \\ y_0 = \rho_0 \sin(\theta_0) \end{cases}$$

On obtient alors  $d^2 = \tilde{\rho}^2 + \rho_0^2 - 2\rho_0\tilde{\rho} \cos(\delta\theta)$  et la partie imaginaire de la racine est donc égale à :

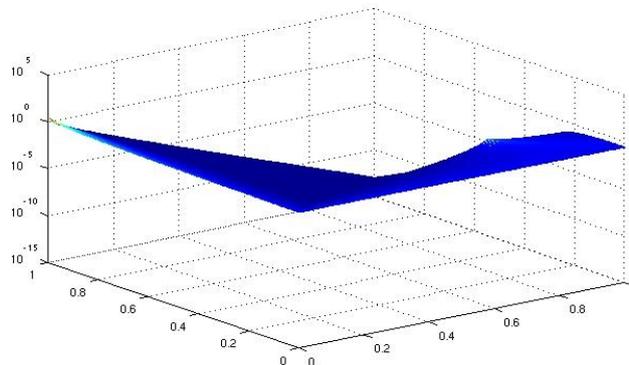
$$\mathbf{Im}(d^2) = \frac{2f}{\omega}(\rho - \rho_0 \cos(\delta\theta))$$

Le problème est alors que les parties réelle et imaginaire peuvent s'annuler au même endroit : le noyau de Schwartz peut alors tendre vers l'infini à l'intérieur

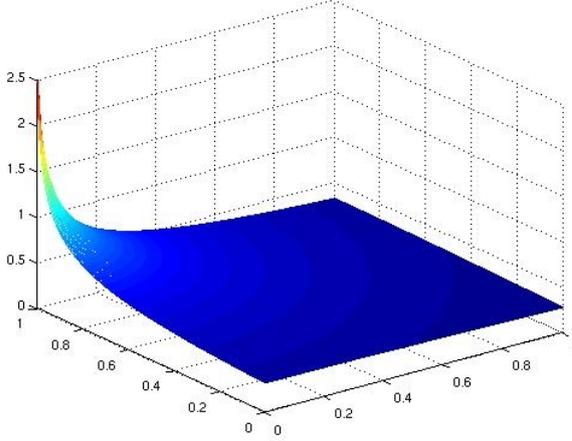
même du domaine PML! Si  $\rho$  est assez grand, il n'y aura pas de problème, ce qui correspond bien à la notion de décroissance exponentielle en asymptotique que nous avons déjà montré. Mais, localement, pour des  $\rho$  assez faibles, on peut rencontrer des difficultés variant avec les décroissances voulues, le choix de la fonction  $f$ , les coefficients d'amortissement à l'intérieur des PML, mais aussi en fonction des formes du domaine au travers du choix du  $\rho_0$  ou des positions des termes sources. Se pose alors un problème de décroissance non uniforme qui peut générer des instabilités. Afin de visualiser clairement ce problème, nous avons réalisé un cas test : nous représenterons la zone convexe (l'intérieur du L). Les axes gradués donnent l'orientation du L, et la zone plottée est le carré qui compléterait le L pour le transformer lui-même en carré. Nous allons visualiser le noyau de Schwartz qui est une fonction de Hankel à argument complexe quand on place une source dans le coin haut gauche du L (de coordonnées  $(0, 1)$ ). L'instabilité apparaît très clairement pour la formulation naïve sur la figure (a). Les figures (b) et (d) présentent ces mêmes résultats en échelle log pour mieux évaluer la décroissance exponentielle de la solution.



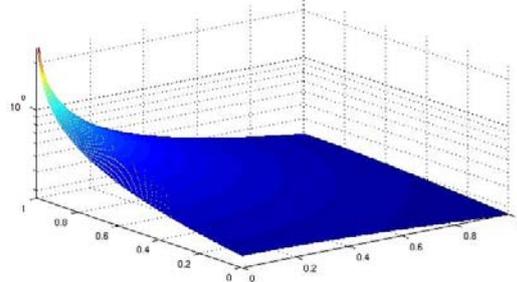
(a) Fonction de Hankel pour les PML naïves



(b) Echelle log



(c) Fonction de Hankel pour les PML corrigées



(d) Echelle log

## 2.9 Cas des Equations d'Euler Linéarisées Barotropes

Rappelons les Equations d'Euler Linéarisées Barotropes établies au chapitre I (1.9) à coefficients constants dans  $\mathbb{R}^+ \times \mathbb{R}^3$  :

$$\frac{\partial}{\partial t} \varphi + c \frac{\partial}{\partial x} \varphi + \begin{bmatrix} \mathbf{0} & \nabla \\ \nabla^T & 0 \end{bmatrix} \varphi = f. \quad (2.41)$$

Le choix de l'axe de la direction de la convection a pour but de simplifier les calculs. Le profil de vitesse  $c$  est choisi constant tel que :  $0 \leq c < 1$ . Le second membre  $f$  a son support spatial contenu dans un parallélépipède  $\mathbf{D}_f = ]-x_0, x_0[ \times ]-y_0, y_0[ \times ]-z_0, z_0[$ , et  $f \in L^2(\mathbb{R}^+, \mathbf{D}_f)$ .

**Remarque :** quand  $c = 0$ , la différence essentielle avec les équations traitées dans le chapitre précédent, outre la difficulté d'écrire les équations de façon intrinsèque, est que le déterminant du symbole n'est pas elliptique. En effet, soit  $\mathcal{C} = \mathbf{p} + c \frac{\partial}{\partial x}$  et  $\mathcal{A} = \mathcal{C} + \begin{bmatrix} \mathbf{0} & \nabla \\ \nabla^T & 0 \end{bmatrix}$ . Le problème harmonique associé à (2.41) est :

$$\mathbf{p} \varphi + \mathcal{A} \varphi = F(\mathbf{p}), \quad (2.42)$$

avec  $\mathbf{p} = \varepsilon + i\omega$ , et  $F$  la transformée de Laplace de  $f$ . Le problème sera étudié en absorption limite ( $\mathbf{p} = \lim_{\varepsilon \rightarrow 0} (\varepsilon + i\omega)$ ). On souhaite inverser l'opérateur  $\mathcal{A}$ .

### Propriété 2.61

Le déterminant de l'opérateur est donné par  $\det(\mathcal{A}) = \mathcal{C}^2(\mathcal{C}^2 - \Delta)$ .

**Démonstration** : la matrice associée à  $\mathcal{A}$  est  $\begin{pmatrix} \mathcal{C} & 0 & 0 & \partial_x \\ 0 & \mathcal{C} & 0 & \partial_y \\ 0 & 0 & \mathcal{C} & \partial_z \\ \partial_x & \partial_y & \partial_z & \mathcal{C} \end{pmatrix}$ . Donc :

$$\begin{aligned} \begin{vmatrix} \mathcal{C} & 0 & 0 & \partial_x \\ 0 & \mathcal{C} & 0 & \partial_y \\ 0 & 0 & \mathcal{C} & \partial_z \\ \partial_x & \partial_y & \partial_z & \mathcal{C} \end{vmatrix} &= \mathcal{C}(\mathcal{C}I_3) - \partial_x \begin{vmatrix} 0 & 0 & \partial_x \\ \mathcal{C} & 0 & \partial_y \\ 0 & \mathcal{C} & \partial_z \end{vmatrix} + \partial_x \begin{vmatrix} \mathcal{C} & 0 & \partial_x \\ 0 & 0 & \partial_y \\ 0 & \mathcal{C} & \partial_z \end{vmatrix} - \partial_x \begin{vmatrix} \mathcal{C} & 0 & \partial_x \\ 0 & \mathcal{C} & \partial_y \\ 0 & 0 & \partial_z \end{vmatrix} \\ &= \mathcal{C}^4 + \partial_x \mathcal{C}(-\partial_x \mathcal{C}) + \partial_y \mathcal{C}(-\partial_y \mathcal{C}) - \partial_z \mathcal{C}(\partial_z \mathcal{C}) \\ &= \mathcal{C}^2(\mathcal{C}^2 - (\partial_x^2 + \partial_y^2 + \partial_z^2)). \quad \square \end{aligned}$$

**Remarque** : le déterminant n'est donc pas elliptique. En effet,  $\mathcal{C}^2 - (\partial_x^2 + \partial_y^2 + \partial_z^2)$  est elliptique mais  $\mathcal{C}$  ne l'est pas. Au sens de Hormander, il est cependant partiellement elliptique par rapport à  $x$ . Il va donc falloir chercher une formulation PML à la main.

**Propriété 2.62**

$$\mathcal{A}^{-1} = \frac{1}{\mathcal{C}(\mathcal{C}^2 - \Delta)} \begin{pmatrix} \mathcal{C}^2 - \partial_y^2 - \partial_z^2 & \partial_x \partial_y & \partial_x \partial_z & -\mathcal{C} \partial_x \\ \partial_x \partial_y & \mathcal{C}^2 - \partial_x^2 - \partial_z^2 & \partial_y \partial_z & -\mathcal{C} \partial_y \\ \partial_x \partial_z & \partial_y \partial_z & \mathcal{C}^2 - \partial_x^2 - \partial_y^2 & -\mathcal{C} \partial_z \\ -\mathcal{C} \partial_x & -\mathcal{C} \partial_y & -\mathcal{C} \partial_z & \mathcal{C}^2 \end{pmatrix}.$$

**Démonstration** : à l'aide de la proposition (2.61), on peut alors inverser formellement l'équation (2.42) pour obtenir la solution suivante :

$$\varphi = \det(\mathcal{A}) \mathbf{Com}(\mathcal{A})^T F = \frac{1}{\mathcal{C}^2(\mathcal{C}^2 - \Delta)} \mathbf{Com}(\mathcal{A})^T F,$$

où  $\mathbf{Com}(\mathcal{A})$  représente la matrice des cofacteurs de  $\mathcal{A}$ . Soit  $a_{ij}$ ,  $(i, j) \in \llbracket 1, 4 \rrbracket^2$  les coefficients de cette matrice. Alors  $a_{ij} = (-1)^{i+j-1} m_{ij}$  où  $m_{ij}$  est le mineur associé à la  $i^e$  ligne et la  $j^e$  colonne.

$$\begin{aligned} a_{11} &= \begin{vmatrix} \mathcal{C} & 0 & \partial_y \\ 0 & \mathcal{C} & \partial_z \\ \partial_y & \partial_z & \mathcal{C} \end{vmatrix} = \mathcal{C}(\mathcal{C}^2 - \partial_y^2 - \partial_z^2) & a_{12} &= - \begin{vmatrix} 0 & 0 & \partial_y \\ 0 & \mathcal{C} & \partial_z \\ \partial_x & \partial_z & \mathcal{C} \end{vmatrix} = \mathcal{C} \partial_x \partial_y \\ a_{13} &= \begin{vmatrix} 0 & 0 & \partial_x \\ \mathcal{C} & 0 & \partial_y \\ \partial_y & \partial_z & \mathcal{C} \end{vmatrix} = \mathcal{C} \partial_x \partial_z & a_{14} &= - \begin{vmatrix} 0 & 0 & \mathcal{C} \\ \partial_x & \partial_y & \partial_z \\ \mathcal{C} & 0 & \partial_x \end{vmatrix} = -\mathcal{C}^2 \partial_x \\ a_{21} &= - \begin{vmatrix} 0 & \mathcal{C} & \partial_z \\ \partial_y & \partial_z & \mathcal{C} \\ \mathcal{C} & 0 & \partial_x \end{vmatrix} = \mathcal{C} \partial_x \partial_y & a_{22} &= \begin{vmatrix} 0 & \mathcal{C} & \partial_z \\ \partial_x & \partial_z & \mathcal{C} \\ \mathcal{C} & 0 & 0 \end{vmatrix} = \mathcal{C}(\mathcal{C}^2 - \partial_x^2 - \partial_z^2) \\ a_{23} &= - \begin{vmatrix} 0 & 0 & \partial_z \\ \partial_x & \partial_y & \mathcal{C} \\ 0 & 0 & \partial_x \end{vmatrix} = \mathcal{C} \partial_y \partial_z & a_{24} &= \begin{vmatrix} 0 & 0 & \mathcal{C} \\ \partial_x & \partial_y & \partial_z \\ \mathcal{C} & \partial_x & \end{vmatrix} = -\mathcal{C}^2 \partial_y \\ a_{31} &= \begin{vmatrix} \mathcal{C} & 0 & \partial_y \\ \partial_y & \partial_z & \mathcal{C} \\ 0 & 0 & \partial_x \end{vmatrix} = \mathcal{C} \partial_x \partial_z & a_{32} &= - \begin{vmatrix} 0 & 0 & \partial_y \\ \partial_x & \partial_z & \mathcal{C} \\ \mathcal{C} & 0 & 0 \end{vmatrix} = \mathcal{C} \partial_y \partial_z \end{aligned}$$

$$\begin{aligned}
a_{33} &= \begin{vmatrix} \mathcal{C} & 0 & \partial_x \\ 0 & \mathcal{C} & \partial_y \\ \partial_x & \partial_y & \mathcal{C} \end{vmatrix} = \mathcal{C}(\mathcal{C}^2 - \partial_x^2 - \partial_y^2) & a_{34} &= - \begin{vmatrix} \mathcal{C} & 0 & 0 \\ 0 & \mathcal{C} & 0 \\ \partial_x & \partial_y & \partial_z \end{vmatrix} = -\mathcal{C}^2 \partial_z \\
a_{41} &= - \begin{vmatrix} 0 & 0 & \partial_x \\ \mathcal{C} & 0 & \partial_y \\ 0 & \mathcal{C} & \partial_z \end{vmatrix} = -\mathcal{C}^3 \partial_x & a_{42} &= \begin{vmatrix} \mathcal{C} & 0 & \partial_x \\ 0 & 0 & \partial_y \\ 0 & \mathcal{C} & \partial_z \end{vmatrix} = -\mathcal{C}^3 \partial_y \\
a_{43} &= - \begin{vmatrix} \mathcal{C} & 0 & \partial_x \\ 0 & \mathcal{C} & \partial_y \\ 0 & 0 & \partial_z \end{vmatrix} = -\mathcal{C}^2 \partial_z & a_{44} &= \begin{vmatrix} \mathcal{C} & 0 & 0 \\ 0 & \mathcal{C} & 0 \\ 0 & 0 & \mathcal{C} \end{vmatrix} = \mathcal{C}^3
\end{aligned}$$

On remarque que  $\forall (i, j) \in \llbracket 1, 4 \rrbracket^2$ ,  $a_{ij} = a_{ji}$  donc  $\mathbf{Com}(\mathcal{A})$  est symétrique. De plus, tous les  $a_{ij}$  ont un facteur  $\mathcal{C}$  commun qui se simplifie avec une partie de  $\mathbf{det}(\mathcal{A})$  pour obtenir la formule proposée.  $\square$

On peut donc trouver un cadre général d'inversion du déterminant compatible avec le calcul matriciel. Soit  $\mathcal{C} = \mathbf{p} + c \frac{\partial}{\partial x}$  avec  $\mathbf{p} = \varepsilon + i\omega$ , et  $*$  la convolution par rapport à la variable  $x$ . On suppose que le second membre  $F$  appartient au module de l'algèbre des convolveurs  $\mathcal{S}'$ . On a :

**Propriété 2.63**

Soit  $|\cdot|_c$  une norme adaptée à notre problème, à savoir :

$$\forall \mathbf{x} \in \mathbb{R}^3, \quad |\mathbf{x}|_c = \sqrt{\frac{x^2}{1-c^2} + y^2 + z^2}.$$

Les inverses des opérateurs sont donnés par les relations :

$$\begin{aligned}
(\mathcal{C}^2 - \Delta)^{-1} &= \frac{1}{4\pi\sqrt{1-c^2}} \frac{1}{|\mathbf{x}|_c} \mathbf{exp} \left( \frac{-\mathbf{p}}{\sqrt{1-c^2}} \left( \frac{-cx}{\sqrt{1-c^2}} + |\mathbf{x}|_c \right) \right) * \\
\mathcal{C}^{-1} &= \frac{1}{c} \mathbf{exp} \left( -\mathbf{p} \frac{x}{c} \right) \mathbf{1}(x) \otimes \delta(y) \otimes \delta(z) *.
\end{aligned}$$

On a alors la solution :

$$\varphi = \mathcal{C}^{-2} * (\mathcal{C}^2 - \Delta)^{-1} * \mathbf{M} * F,$$

où  $\mathbf{M}$  est la matrice de dérivation de dirac associée à la matrice  $\mathcal{C}^{-1} \mathbf{Com}(\mathcal{A})^T$ .

**Démonstration** :  $\mathcal{C} = \mathbf{p} + c \frac{\partial}{\partial x}$ . Son inverse s'obtient directement : c'est le produit tensoriel d'une fonction à décroissance rapide par une distribution à support ponctuel. Le calcul de l'inverse de  $\mathcal{C}^2 - \Delta$  est plus compliqué. Soit  $\xi = \mathcal{F}(x)$ ,  $\nu = \mathcal{F}(y)$  et  $\varepsilon = \mathcal{F}(z)$  les variables transformées par Fourier. En appliquant la transformée de Fourier sur l'opérateur  $\mathcal{C}^2 - \Delta$ , on obtient que :

$$\begin{aligned}
\mathcal{F}(\mathcal{C}^2 - \Delta) &= (\mathbf{p} + ic\xi)^2 + \xi^2 + \nu^2 + \varepsilon^2 \\
&= \frac{\mathbf{p}^2}{1-c^2} + \left( \sqrt{1-c^2} \xi + i\mathbf{p} \frac{c}{\sqrt{1-c^2}} \right)^2 + \nu^2 + \varepsilon^2.
\end{aligned}$$

On sait que  $\mathcal{F}(SoJ) = \frac{1}{|\det J|} \mathcal{F}(S) o(J^*)^{-1}$  (formule de l'isomorphie). On utilise ensuite la formule du retard pour se ramener à la solution de  $\left(\frac{\mathbf{p}}{\sqrt{1-c^2}}\right)^2 - \Delta = \delta$ . D'où la solution proposée.  $\square$

Comme  $\mathbf{Re}(\mathbf{p}) > 0$  et que  $c < 1$ , on a :

$$\lim_{|\mathbf{x}| \rightarrow \infty} \frac{-c x}{\sqrt{1-c^2}} + |\mathbf{x}|_c = +\infty,$$

donc l'inverse de  $\mathcal{C}^2 - \Delta$  appartient à  $L^1_{loc}$  et est à décroissance rapide. Les inverses sont uniques pour ces deux éléments dans l'algèbre des convoleurs  $\mathcal{O}'_c$ . Comme  $\mathcal{S}'$  est un module sur  $\mathcal{O}'_c$ , on peut inverser le système (2.42) puisque l'on sait inverser le déterminant dans l'algèbre.

Cette proposition est établie pour  $\mathbf{Re}(\mathbf{p}) > 0$  et le passage à la limite quand  $\mathbf{Re}(\mathbf{p}) \rightarrow 0$  exige plus quand au comportement du second membre en l'infini.

Introduire des PML sur cette équation revient à modifier l'équation (2.42) en :

$$\mathbf{p} \tilde{\varphi} + \tilde{\mathcal{A}} \tilde{\varphi} = \tilde{F}. \quad (2.43)$$

On cherche alors une immersion telle que la solution  $\tilde{\varphi}$  soit à décroissance rapide pour tout  $\tilde{F}$  à support compact, même quand  $\mathbf{Re}(\mathbf{p}) \rightarrow 0$ . De plus, il faut que le support de  $\tilde{F}$  soit inclus dans  $\mathbf{D}_f$  et que  $\tilde{F} = F$ . Ainsi on aura  $\tilde{\varphi}|_{\mathbf{D}_f} = \varphi|_{\mathbf{D}_f}$ . Enfin, on souhaite que le comportement asymptotique de  $\tilde{\varphi}$  soit indépendant de  $\mathbf{p}$ . Autrement dit, il faut que la solution du problème PML  $\tilde{\varphi}$  soit majorée par une fonction à décroissance rapide à l'infini indépendante de  $\mathbf{p}$ .

Si l'on remplace  $\mathbf{p}$  par  $i\omega$ , on obtient que  $\hat{\mathcal{C}} = i\omega + c \frac{\partial}{\partial x}$ . Si  $c \neq 0$ , le déterminant n'est pas elliptique mais  $\hat{\mathcal{C}}$  est partiellement elliptique en  $x$  et  $\hat{\mathcal{C}}^2 - \Delta$  elliptique. Si  $c = 0$ , le déterminant est elliptique (exemple : équation de Helmholtz). Malgré cela, il est quand même possible d'interpréter l'équation (2.43) comme étant l'équation (2.42) écrite sur une variété immergée dans  $\mathbb{C}^3$  à l'aide du changement de variable complexe :

$$\tilde{\mathbf{x}}_i = \tilde{\mathbf{x}}_i(\mathbf{x}) = \mathbf{x}_i + \frac{1}{\mathbf{p}} \int_0^{\mathbf{x}_i} \sigma_i(\tau) d\tau, \quad \forall i \in \llbracket 1, 3 \rrbracket,$$

où les fonctions  $\sigma_i \in C^\infty$ , croissantes sur  $\mathbb{R}^+$ , décroissantes sur  $\mathbb{R}^-$ , toujours positives ou nulles. On aura de plus les relations :

$$\sigma_1(\mathbf{x}) = 0 \text{ si } x \in [-x_0, x_0] \quad \sigma_2(\mathbf{x}) = 0 \text{ si } y \in [-y_0, y_0] \quad \sigma_z(\mathbf{x}) = 0 \text{ si } z \in [-z_0, z_0].$$

Pour démontrer la décroissance rapide en l'infini, nous ferons en outre l'hypothèse que les fonctions  $\sigma_i$  tendent lentement vers l'infini quand  $|\mathbf{x}|$  tend vers l'infini. Soit  $\mathbf{M}$  la variété définie par cette immersion. Elle est localement confondue avec  $\mathbf{D}_f$ ,

est à fibré tangent complexe, engendré par les  $\partial_{x_i}$  et cotangent complexe, engendré par les  $dx_i$ . Sur cette variété, il existe une pseudo-métrique (isomorphisme entre le fibré tangent et cotangent) donnée par :

$$g_{ij} = \sum_{\omega} \frac{\partial \tilde{\omega}}{\partial x_i} \frac{\partial \tilde{\omega}}{\partial x_j} = \delta_{ij} \left(1 + \frac{1}{\mathbf{p}} \sigma_i(\mathbf{x})\right)^{-1} \left(1 + \frac{1}{\mathbf{p}} \sigma_j(\mathbf{x})\right)^{-1}.$$

La matrice des  $g_{ij}$  est diagonale donc inversible, et son déterminant  $g$  a une racine globalement définie  $\sqrt{g} = \left(1 + \frac{1}{\mathbf{p}} \sigma_1(\mathbf{x})\right)^{-1} \left(1 + \frac{1}{\mathbf{p}} \sigma_2(\mathbf{x})\right)^{-1} \left(1 + \frac{1}{\mathbf{p}} \sigma_3(\mathbf{x})\right)^{-1}$ .

On peut faire les calculs à la main pour des PML cartésiennes mais il est impossible de les faire de façon intrinsèque. Il nous faudrait pour cela un calcul différentiel et tensoriel défini sur les variétés Riemanniennes (produits extérieurs et intérieurs de tenseurs ou de formes, transformée de Hodge, expression des dérivées co et contra variantes). Or comme ces équations sont les linéarisées d'équations intrinsèques, elles ne peuvent pas s'écrire de façon intrinsèque. De plus, on sait que les Equations d'Euler linéarisées n'admettent pas de formulation dispersive [31]. Il faut donc exhiber le changement de variables, l'appliquer et se placer dans les bonnes algèbres. L'équation (2.43) peut alors être redéfinie de la façon suivante :

**Définition 2.10**

On appelle système PML cartésien associé à (2.42), le système défini sur  $\mathbf{M}$  par :

$$p\tilde{\Phi} + \tilde{\mathcal{A}}\tilde{\Phi} = \tilde{F}, \quad (2.44)$$

où  $\tilde{\mathcal{A}}$  est obtenu en remplaçant  $\partial_{x_i}$  par  $\partial_{\tilde{x}_i}$  dans  $\mathcal{A}$ , et  $\tilde{F}$  est à support compact (de forme  $F(\tilde{x}(x), \tilde{y}(y), \tilde{z}(z))$  où l'application  $(x, y, z) \rightarrow \tilde{F}$  est  $L^2$  et à support compact dans  $\mathbb{R}^3$ ).

Nous allons exhiber une solution fondamentale à droite régulière de ce système. Pour ce faire, intéressons nous tout d'abord aux noyaux des opérateurs  $\tilde{\mathcal{C}}$  et  $\tilde{\mathcal{C}}^2 - \Delta$  obtenus à partir de  $\mathcal{C}$  et  $\mathcal{C}^2 - \Delta$  en remplaçant  $\partial_{x_i}$  par  $\partial_{\tilde{x}_i}$ .

**Proposition 2.64**

Les opérateurs  $\tilde{\mathcal{C}}$  et  $\tilde{\mathcal{C}}^2 - \Delta$  admettent des noyaux fondamentaux symétriques à droite réguliers, avec respectivement :

$$G_{\tilde{\mathcal{C}}*\tilde{\varphi}} = \frac{1}{c} \int \mathbf{1}(x - x') \exp\left(-\frac{\mathbf{p}}{c}(\tilde{x} - \tilde{x}')\right) \varphi(\tilde{x}', \tilde{y}, \tilde{z}) d\tilde{x}'.$$

$$G_{\tilde{\mathcal{C}}^2 - \Delta*\tilde{\varphi}} = \frac{1}{4\pi\sqrt{1-c^2}} \int \frac{1}{\tilde{r}} \exp\left(-\frac{\mathbf{p}}{\sqrt{1-c^2}}\left(\frac{-c(\tilde{x} - \tilde{x}')}{\sqrt{1-c^2}} + \tilde{r}\right)\right) \tilde{\varphi} d\tilde{x} \wedge d\tilde{y} \wedge d\tilde{z},$$

où  $\tilde{r} = \sqrt{\frac{(\tilde{x} - \tilde{x}')^2}{1-c^2} + (\tilde{y} - \tilde{y}')^2 + (\tilde{z} - \tilde{z}')^2}$  (la racine étant définie par sa détermination en partie réelle positive).

**Démonstration** :  $G_{\widetilde{\mathcal{C}^2 - \Delta}}$  est obtenu par prolongement analytique à  $\mathbb{C}^3/\{0\}$  ( $\mathcal{C}^2 - \Delta$  est elliptique pour  $c < 1$ ). On montre que  $G_{\widetilde{\mathcal{C}^2 - \Delta}}^*$  est un noyau fondamental régulier (et même très régulier puisque  $C^\infty$  sauf sur la diagonale  $X' = X$ ) de  $\widetilde{\mathcal{C}^2 - \Delta}$  de façon tout à fait analogue à ce qui est fait pour  $\mathbf{p}^2 - \Delta$  dans [10], [5].  $\mathcal{C}^2 - \Delta$  est au moins hypoelliptique. Soit :

$$\phi = G_{\widetilde{\mathcal{C}^2 - \Delta}}^* \widetilde{\varphi} = \frac{1}{c} \exp\left(-\frac{\mathbf{p}}{c} \widetilde{x}\right) \int_{-\infty}^x \exp\left(\frac{\mathbf{p}}{c} \widetilde{\tau}\right) S_x(\tau) \varphi(\widetilde{\tau}, \widetilde{y}, \widetilde{z}) d\widetilde{\tau}.$$

On a  $c\partial_{\widetilde{x}}\phi = -\mathbf{p}\phi + \varphi$  donc  $\widetilde{\mathcal{C}}_0 G_{\widetilde{\mathcal{C}^2 - \Delta}}^* \widetilde{\varphi} = \widetilde{\varphi}$  et  $\phi$  est évidemment  $C^\infty$ .  $\square$

**Conclusion** : dans le cas des Equations d'Euler Linéarisées Barotropes il est impossible d'obtenir une formulation PML générale. La non ellipticité du déterminant du symbole couplé à l'impossibilité d'écrire les équations de façon intrinsèque oblige à revenir à des méthodes au cas par cas. Néanmoins, les PML cartésiennes sont réalisables sur ce type de problème (par exemple pour un écoulement dans un guide d'onde avec des profils de vitesses en tangente hyperbolique).

## 2.10 Simulations Numériques

Dans cette section seront détaillés les cas tests numériques réalisés. Après avoir explicité les conditions aux limites admissibles et la formulation variationnelle, nous détaillerons les diverses formulations. Les cas tests seront réalisés pour des problèmes instationnaires et harmoniques, sur des maillages déstructurés.

### 2.10.1 Relèvement des conditions aux limites

Dans l'exemple qui suit, nous allons étudier les conditions aux limites admissibles, avant de rechercher un relèvement pour une condition aux limites donnée. Soit  $d = 2$  ou 3 la dimension d'espace considérée. L'opérateur associé aux problèmes de type diffraction étudiés ici est  $\mathbf{I}_{d+1}\partial_t + \begin{pmatrix} \mathbf{O}_d & \nabla \\ \nabla^T & 0 \end{pmatrix}$ .

On note alors  $\mathbf{K} = \mathbf{I}_{d+1}$  et  $\forall i \in \llbracket 1, n \rrbracket$ ,  $\mathbf{A}^i = \left( \begin{array}{ccc|c} & & & \delta_{1i} \\ & & & \vdots \\ & \mathbf{O}_{d,d} & & \delta_{di} \\ \hline \delta_{1i} & \cdots & \delta_{di} & 0 \end{array} \right)$ . On est

dans le cas d'un système symétrique à coefficients constants.

Au bord, si  $\mathbf{n} = (n_1, \dots, n_d)$  est la normale unitaire sortante au domaine en un point de sa frontière, la matrice  $\mathbf{A} \mathbf{n} = \mathbf{A}^i \mathbf{n}_i$  s'écrit :  $\mathbf{A} \mathbf{n} = \begin{pmatrix} \mathbf{O}_d & \mathbf{n} \\ \mathbf{n}^T & 0 \end{pmatrix}$ . On cherche à expliciter le noyau de cette matrice. Comme  $\mathbf{n} \neq \mathbf{0}$ , un vecteur  $V = {}^t(v_1, \dots, v_d, u) \in \mathbb{R}^{d+1}$  est dans  $\ker(\mathbf{A} \mathbf{n})$  si et seulement si  $v \cdot \mathbf{n} = 0$  et  $u = 0$ . On constate alors que  $\ker(\mathbf{A} \mathbf{n})$  est un espace de dimension  $d - 1$  et que, d'après la remarque précédente, on peut expliciter le noyau de la manière suivante :

$$\ker(\mathbf{A} \mathbf{n}) = \begin{bmatrix} \mathbf{n} \\ 0 \end{bmatrix}^\perp \cap \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}^\perp = \left( \mathbb{C} \begin{bmatrix} \mathbf{n} \\ 0 \end{bmatrix} + \mathbb{C} \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \right)^\perp$$

On remarque que  $\mathbf{A} \mathbf{n} \begin{bmatrix} \mathbf{n} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}$  et que  $\mathbf{A} \mathbf{n} \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{n} \\ 0 \end{bmatrix}$ . La base orthonormée de vecteurs propres de  $\mathbf{A} \mathbf{n}$  est alors :  $(K_1, \dots, K_{d-1}, U_1, U_{-1})$  où les vecteurs  $K_1, \dots, K_{d-1}$  forment une base de  $\ker(\mathbf{A} \mathbf{n})$ .  $U_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{n} \\ 1 \end{bmatrix}$  est le vecteur propre associé à la valeur propre 1 et  $U_{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{n} \\ -1 \end{bmatrix}$  est le vecteur propre associé à la valeur propre  $-1$ . Les deux vecteurs précédents ne sont jamais nuls si la frontière est régulière et ils sont normés orthogonaux.

Lorsque  $\mathbf{A}$  est une matrice symétrique réelle dans un  $\mathbb{R}$  ev de dimension  $n$ , si  $(v_1, \dots, v_n)$  désigne une base orthonormale de vecteurs propres de  $\mathbf{A}$  associés aux valeurs propres  $(\lambda_1, \dots, \lambda_n)$ , il est facile de vérifier que  $\mathbf{A}$  s'écrit :  $\mathbf{A} = \sum_{i=1}^n \lambda_i v_i \otimes v_i$ .

Dans le cas complexe, cette relation devient :  $\mathbf{A} = \sum_{i=1}^n \lambda_i v_i \otimes \bar{v}_i$ . Cette somme peut être restreinte aux valeurs propres non nulles. On a alors :

$$\mathbf{A}^\pm = \sum_{\lambda \in sp(\mathbf{A}) \cap \mathbb{R}_\pm^*} \lambda v_\lambda \otimes v_\lambda \quad |\mathbf{A}| = \sum_{\lambda \in sp(\mathbf{A})} |\lambda| v_\lambda \otimes v_\lambda = \mathbf{A}^+ - \mathbf{A}^-.$$

Dans notre cas,  $\mathbf{A} \mathbf{n} = U_1 \otimes U_1 - U_{-1} \otimes U_{-1}$ . On obtient alors :

$$\begin{aligned} \mathbf{A} \mathbf{n}^+ &= U_1 \otimes U_1 = \frac{1}{2} \begin{pmatrix} \mathbf{n} \otimes \mathbf{n} & \mathbf{n} \\ \mathbf{n}^T & 1 \end{pmatrix} \\ \mathbf{A} \mathbf{n}^- &= -U_{-1} \otimes U_{-1} = \frac{1}{2} \begin{pmatrix} -\mathbf{n} \otimes \mathbf{n} & \mathbf{n} \\ \mathbf{n}^T & -1 \end{pmatrix} \\ |\mathbf{A} \mathbf{n}| &= \mathbf{A} \mathbf{n}^+ - \mathbf{A} \mathbf{n}^- = \begin{pmatrix} \mathbf{n} \otimes \mathbf{n} & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Construire un espace vectoriel maximal positif à partir de la décomposition de  $\mathbb{C}^{d+1}$  en  $\ker(\mathbf{A} \mathbf{n}) \oplus \mathbb{C}U_1 \oplus \mathbb{C}U_{-1}$  revient à choisir une droite dans le cône de positivité de  $\mathbf{A} \mathbf{n}$  dans le plan  $\mathbb{C}U_1 \oplus \mathbb{C}U_{-1}$ . Soit  $(\lambda, \mu) \in \mathbb{C}^2$ . Comme  $(\mathbf{A} \mathbf{n} U_{(\lambda, \mu)} | U_{(\lambda, \mu)}) = |\lambda|^2 - |\mu|^2$ , le vecteur  $U_{(\lambda, \mu)} = \lambda U_1 + \mu U_{-1}$  est dans le cône de positivité de  $\mathbf{A} \mathbf{n}$  si et seulement si  $|\lambda| \geq |\mu|$ . Comme  $(\mathbf{A} \mathbf{n} U_{(\lambda, \mu)} | U_{(\lambda, \mu)}) \geq 0$ ,  $\lambda \neq 0$  et ainsi  $\left| \frac{\mu}{\lambda} \right| \leq 1$ . Donc, pour construire un espace vectoriel adéquat, il faut et il suffit d'adjoindre à  $\ker(\mathbf{A} \mathbf{n})$  la droite complexe engendrée par un vecteur  $U_{(1, \beta)} = U_1 + \beta U_{-1}$ ,  $\beta \in \mathbb{C}$ ,  $|\beta| \leq 1$ .

Soit  $\mathcal{N} = \ker(\mathbf{A} \mathbf{n}) \oplus \mathbb{C}U_{(1, \beta)}$  l'espace vectoriel maximal positif. Soit  $U \in \mathcal{N}$ . Alors  $U$  s'écrit sous la forme  $K + \lambda U_{(1, \beta)}$ , avec  $K \in \ker(\mathbf{A} \mathbf{n})$  et  $\lambda \in \mathbb{C}$ .

Donc  $(\mathbf{A} \mathbf{n} K | K) = |\lambda|^2(1 - |\beta|^2) \geq 0$  et ainsi  $(\mathbf{A} \mathbf{n} U | U) \geq 0$ .

Soit  $V \in [\mathbf{A} \mathbf{n}(\mathcal{N})]^\perp$ . Comme  $\mathbf{A} \mathbf{n}(\mathcal{N}) = \{0\} + \mathbb{C} \mathbf{A} \mathbf{n} U_{(1,\beta)} = \mathbb{C} U_{(1,-\beta)}$ , on a  $[\mathbf{A} \mathbf{n}(\mathcal{N})]^\perp = [\mathbb{C} U_{(1,-\beta)}]^\perp = \ker(\mathbf{A} \mathbf{n}) \oplus \mathbb{C} U_{(\beta,1)}$ . Donc  $V$  est de la forme  $K + \lambda U_{(\beta,1)}$ , avec  $K \in \ker(\mathbf{A} \mathbf{n})$  et  $\lambda \in \mathbb{C}$ . D'où  $(\mathbf{A} \mathbf{n} V | V) = |\lambda|^2(1 - |\beta|^2) \geq 0$  et ainsi  $(\mathbf{A} \mathbf{n} V | V) \leq 0$ .

Pour le problème primal, puisque  $\mathcal{N} = \ker(\mathbf{A} \mathbf{n}) \oplus \mathbb{C} U_{(1,\beta)} = [\mathbb{C} U_{(-\beta,1)}]^\perp$ , la condition  $U \in \mathcal{N}$  s'écrit  $U^T \cdot U_{(-\beta,1)} = 0$ , avec les vecteurs  $U^T = (v_1, \dots, v_d, u)$  et  $U_{(-\beta,1)} = \frac{1}{\sqrt{2}} \begin{bmatrix} (-\beta + 1)n \\ -\beta - 1 \end{bmatrix}$ . La condition aux limites correspondante au choix d'un  $\beta$  dans le disque unité s'écrit alors :

$$(\beta - 1)v \cdot \mathbf{n} + (\beta + 1)u = 0. \quad (2.45)$$

On retrouve les conditions de type Dirichlet lorsque  $\beta = 1$  et les conditions de type Neumann lorsque  $\beta = -1$ .

Pour le problème dual, on a  $[\mathbf{A} \mathbf{n}(\mathcal{N})]^\perp = [\mathbb{C} U_{(1,-\beta)}]^\perp$ . Auquel cas, la condition  $V \in [\mathbf{A} \mathbf{n}(\mathcal{N})]^\perp$  s'écrit  $V^T \cdot U_{(1,-\beta)} = 0$ , avec les notations  $V^T = (w_1, \dots, w_d, v)$  et  $U_{(1,-\beta)} = \frac{1}{\sqrt{2}} \begin{bmatrix} (1 - \beta)n \\ 1 + \beta \end{bmatrix}$ , soit :

$$(1 - \beta)w \cdot \mathbf{n} + (1 + \beta)v = 0. \quad (2.46)$$

Pour relever la condition aux limites, on cherche, en tout point  $x$  de la frontière, une matrice  $\mathbf{M}(x)$ , carrée de taille  $d + 1$ , vérifiant  $\ker(\mathbf{M}(x)) = \mathcal{N}(x)$ .

Or  $\text{rang}(\mathbf{M}) = 1$  donc  $\mathbf{M}$  est de la forme  $U \otimes V$  où  $U$  et  $V$  sont deux vecteurs de  $\mathbb{C}^{d+1}$ . On a alors :

$$\ker(\mathbf{M}) = V^\perp \quad \text{Im}(\mathbf{M}) = \mathbb{C}U.$$

D'où  $\mathbf{M}$  s'écrit sous la forme  $W \otimes U_{(-\beta,1)}$ . Il y a correspondance biunivoque entre le choix d'un vecteur dans le cône de positivité de  $\mathbf{A} \mathbf{n}$  (que l'on adjoint à  $\ker(\mathbf{A} \mathbf{n})$  pour former l'espace vectoriel  $\mathcal{N}$ ) et le choix d'un vecteur dans le cône de négativité de  $\mathbf{A} \mathbf{n}$  dont l'orthogonal sera le noyau de  $\mathbf{M}$ . Si on choisit arbitrairement un vecteur dans le cône de négativité et  $\beta'$  dans le disque unité de telle sorte que le vecteur  $U_{(\beta',1)}$  détermine le noyau de  $\mathbf{M}$ , on obtient que  $\mathcal{N} = \ker(\mathbf{A} \mathbf{n}) \oplus \mathbb{C} U_{(1,-\beta')}$  et les conditions aux limites pour le problème primal seront celles obtenues dans [38] :

$$(\beta' + 1)v \cdot \mathbf{n} + (\beta' - 1)u = 0.$$

La matrice  $\mathbf{M}$  dépendant de  $\beta$ , nous la désignerons par la suite comme :

$$\mathbf{M}_\beta = \frac{1}{2} \begin{pmatrix} (\beta + 1)\mathbf{n} \otimes \mathbf{n} & (\beta - 1)\mathbf{n} \\ -(1 + \beta)\mathbf{n}^T & 1 - \beta \end{pmatrix}$$

### 2.10.2 Formulation variationnelle du problème

Pour définir la formulation variationnelle des problèmes de diffraction, s'affranchir des problèmes de traces et justifier les intégrations par parties, on va se placer dans l'espace  $\tilde{E}$  défini par

$$\varphi \in \tilde{E} \Leftrightarrow \varphi \in (H_{mex}^{k+1}(\Omega \times [0, T]))^m, \quad (2.47)$$

$k \in \mathbb{N}$ , ce qui signifie que les composantes de  $\varphi$  sont  $H^{k+1}$  par morceaux dans  $\Omega \times [0, T]$  à valeurs complexes, ou plus précisément qu'il existe une famille finie d'ouverts  $(U_e)$  vérifiant les propriétés suivantes. Les  $U_e$  sont connexes, deux à deux disjoints, non dégénérés et de mesure de Lebesgue uniformément bornée.  $\bigcup_e U_e = \overline{\Omega \times [0, T]}$ .  $\partial U_e$  est de classe  $C_{mex}^{k+1}$ , et  $U_e$  est localement d'un seul côté de la frontière  $\partial U_e$ . Pour tout  $e$ , les restrictions de  $\varphi$  à  $U_e$  sont dans  $H^{k+1}(U_e)$ .

En pratique, nous considérons une sous-partition  $(V_e)$  en ouverts de la partition  $(U_e)$ , vérifiant les mêmes propriétés. Posons  $S\varphi = \bigcup_e \partial V_e \setminus \partial Q$   $R\varphi = \bigcup_e V_e$ .  $S\varphi$  représente l'ensemble des discontinuités de la distribution  $\varphi$  et  $R\varphi$  son complémentaire dans  $\Omega$ .

Considérons un problème de diffraction dont la forme générale est établie au chapitre 1 (1.14), à savoir dans  $\mathbb{R}^n \times \mathbb{R}^+$ , avec  $m$  le nombre d'équations :

$$\mathbf{A}_0(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \sum_i \mathbf{A}_i(\mathbf{x}) \frac{\partial \varphi(\mathbf{x}, t)}{\partial \mathbf{x}_i} + \mathbf{B}(\mathbf{x}) \varphi(\mathbf{x}, t) = f(\mathbf{x}, t).$$

#### **Théorème 2.65**

Soit  $\mathbf{n} = (n_i)$  un vecteur unitaire normal sortant à  $\partial V_e$ .

$\varphi_\beta$  est solution du problème de diffraction dans  $\tilde{E}$  si et seulement si  $\forall \psi \in \tilde{E}$ , on a :

$$\begin{aligned} \int_{R\varphi \cap R\psi} ((\mathbf{A}_0 \partial_t + \mathbf{A}^i \partial_i + \mathbf{B})\varphi, \psi) + \int_{\widetilde{S\varphi \cup S\psi}} (\mathbf{A}^i \mathbf{n}_i^- (\varphi^+ - \varphi^-), \psi^-) \\ + \int_{\partial(\Omega \times [0, T])} (\mathbf{M}_\beta \varphi, \psi) = \int_{\Omega \times [0, T]} (f, \psi). \end{aligned} \quad (2.48)$$

où  $\widetilde{S\varphi \cup S\psi}$  signifie que l'intégration correspondante doit s'effectuer suivant les deux orientations de  $\mathbf{n}$  (des deux côtés de la frontière). Si  $\mathbf{x} \in (S\varphi \cup S\psi)$ , on désigne par  $\varphi_\beta^+(\mathbf{x})$  la limite quand  $\varepsilon \rightarrow 0^+$  de  $\varphi(\mathbf{x} + \varepsilon \mathbf{n})$  et  $\varphi_\beta^-(\mathbf{x})$  celle de  $\varphi(\mathbf{x} - \varepsilon \mathbf{n})$ , et cette limite est au moins définie dans  $H^{\frac{1}{2}}(S\varphi \cup S\psi)$ .

En choisissant une seule orientation de la normale, on obtient que :

$$\int_{\widetilde{S\varphi \cup S\psi}} (\mathbf{A}^i \mathbf{n}_i^- (\varphi^+ - \varphi^-), \psi^-) = \int_{S\varphi \cup S\psi} (\mathbf{A}^i \mathbf{n}_i^- (\varphi^+ - \varphi^-), \psi^-) + (\mathbf{A}^i \mathbf{n}_i^+ (\varphi^+ - \varphi^-), \psi^+).$$

**Démonstration** : l'équation principale du système (1.14) s'écrit au sens des distributions :

$$\begin{cases} \mathbf{A}_0(\mathbf{x})\partial_t\varphi + \mathbf{A}_i(\mathbf{x})\partial_i\varphi + \mathbf{B}(\mathbf{x})\varphi = f \text{ presque partout sur } R\varphi \\ \mathbf{A}^i\mathbf{n}_i[\varphi] = 0 \text{ sur } S\varphi. \end{cases}$$

$$\text{En conséquence, } \forall \psi \in \tilde{E}, \int_{R\varphi \cap R\psi} ((\mathbf{A}_0\partial_t + \mathbf{A}^i\partial_i + \mathbf{B})\varphi, \psi) = \int_{R\varphi \cap R\psi} (f, \psi).$$

De plus, comme  $\ker(\mathbf{A}^i\mathbf{n}_i) \subset \ker(\mathbf{A}^i\mathbf{n}_i)^+$ ,  $\ker(\mathbf{A}^i\mathbf{n}_i) \subset \ker(\mathbf{A}^i\mathbf{n}_i)^-$  et que  $(\mathbf{A}^i(-\mathbf{n}_i))^+ = -(\mathbf{A}^i\mathbf{n}_i)^-$ , on a l'équivalence entre

$$\mathbf{A}^i\mathbf{n}_i[\varphi] = 0 \text{ sur } S\varphi \Leftrightarrow \mathbf{A}^i\mathbf{n}_i^-(\varphi^+ - \varphi^-) = 0 \text{ en } \mathbf{n}_i \text{ et en } -\mathbf{n}_i, \quad (2.49)$$

donc, pour tout  $\psi \in \tilde{E}$ , on a  $\int_{\widetilde{S\varphi \cup S\psi}} (\mathbf{A}^i\mathbf{n}_i^-(\varphi^+ - \varphi^-), \psi^-) = 0$ . La contribution de la condition aux limites permet de conclure, à savoir que  $\varphi$  solution de (1.14) implique que  $\varphi$  vérifie (2.48). Un choix judicieux de fonctions tests  $\psi$  qui permet d'annuler successivement les intégrales mises en jeu dans (2.48), ainsi que l'équivalence (2.49) montrent trivialement la réciproque.  $\square$

Une discrétisation naturelle symétrique espace/temps de cette formulation variationnelle avec des éléments produits cartésiens espace/temps aboutirait à un schéma totalement implicite. Si chacune des composantes des fonctions de base est un polynôme de degré inférieur à  $k$ , les auteurs de [8, 15] obtiennent l'estimation de l'erreur de consistance en la norme  $||| \cdot |||$  associée au produit scalaire

$$(\varphi, \psi)_{(L^2(Q))^m} + \int_{\widetilde{S\varphi \cup S\psi}} (|\mathbf{A}^i\mathbf{n}_i|[\varphi], [\psi])_{\mathbb{C}^m} d\sigma,$$

où  $[\varphi]$  représente le saut des traces de  $\varphi$  sur  $S\varphi$ , suivante :

$$|||\varphi - \varphi_h||| \leq cste \|\varphi\|_{H_{m,cx}^{k+1}} h^{k+\frac{1}{2}}, \quad (2.50)$$

où  $\varphi_h$  est la solution du problème associé au schéma totalement implicite et  $h$  la dimension caractéristique du maillage du domaine spatio-temporel  $Q$ . Pour l'approximation temporelle, nous choisirons un schéma implicite, inconditionnellement stable.

Dans cette section nous allons présenter les résultats numériques obtenus pour valider la théorie PML. Nous étudions deux types de problèmes : le problème en condition initiale, et la réflexion d'une onde plane sur un scatterer parfait. Pour les problèmes instationnaires, étant donné que toutes les fréquences jouent un rôle équivalent, elles ont toutes la même énergie. Ainsi, un problème instationnaire en condition initiale est le cas le plus difficile pour cette théorie.

**Remarque** : la formulation variationnelle ci dessus sera utilisée aussi bien dans ce chapitre que dans le chapitre 3. Le flux splitting réalisé est en effet de type Van

Leer-Roe, et en choisissant des fonctions  $\varphi$  et  $\psi$  comme des polynômes de degré  $k$  en espace et à coefficients dépendant du temps, on obtiendra des formulations de type Galerkin Discontinu implicite.

### 2.10.3 Problème harmonique

Nous allons dans un premier temps simuler numériquement les PML étoilées pour l'Equation de Maxwell bidimensionnelle. Nous utiliserons la décomposition de la matrice  $\mathbf{K}(\mathbf{p}, \mathbf{x})$ , plus facilement implémentable obtenue dans la section (2.6). Soit un maillage  $\bigcup \omega_e$  déstructuré où les éléments seront supposés non dégénérés.

Dans le cas des PML étoilées, la proposition (2.26) donne la matrice sous la forme

$$\mathbf{QP}(\mathbf{x}) \left( \mathbf{p} \mathbf{I}(\mathbf{x}) + \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x}) \left( \mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}) \right)^{-1} \right) \mathbf{P}^T(\mathbf{x}) \mathbf{Q}^T.$$

La formulation harmonique s'obtient en remplaçant  $\mathbf{p}$  par  $i\omega$ . On rappelle que  $\omega = 2\pi f = \frac{2\pi}{T}$  et  $f = \frac{c}{\lambda}$  donc  $\omega = \frac{2\pi c}{\lambda}$ . On jouera sur cette fréquence pour resserer ou éloigner les isovaleurs.

**Proposition 2.66**

La matrice PML complexe se décompose sur  $\mathbb{R}^2$  sous la forme :

$$\mathbf{Re}(\mathbf{K}) = \mathbf{QP}(\mathbf{A} + \mathbf{BFD}) \mathbf{P}^T(\mathbf{x}) \mathbf{Q}^T \quad \mathbf{Im}(\mathbf{K}) = \omega \mathbf{QP}(\mathbf{I} - \mathbf{BD}) \mathbf{P}^T(\mathbf{x}) \mathbf{Q}^T$$

$$\text{avec } \mathbf{D} = \begin{pmatrix} \frac{1}{\omega^2 + \alpha^2} & 0 & 0 \\ 0 & \frac{1}{\omega^2 + \beta^2} & 0 \\ 0 & 0 & \frac{1}{\omega^2} \end{pmatrix}.$$

**Démonstration** : découle naturellement de la décomposition en éléments simples de la matrice PML.  $\square$

Pour les simulations numériques, une méthode de type Volumes Finis a été choisie. La formulation variationnelle du problème est décrite dans la sous section précédente.

**Proposition 2.67**

La formulation PML étoilées harmonique des Equations de Maxwell bidimensionnelle revient à la résolution du système  $\mathbf{M}\phi = \mathbf{F}$  où  $\mathbf{Re}(\mathbf{M}) = \mathbf{Re}(\mathbf{K})$  plus la contribution des termes venant de l'intégrale de bord  $(\mathbf{A}_i \mathbf{n}_i)$ ,  $\mathbf{Im}(\mathbf{M}) = \mathbf{Im}(\mathbf{K})$ ,  $\mathbf{F} = (\text{aire}_{\omega_e} f(x_{\omega_e}^G, y_{\omega_e}^G))$  représente la condition initiale choisie, à savoir la régularisée du Dirac sous la forme  $f(x, y) = \exp \left( 1 + \frac{1}{\left(\frac{x-x_c}{R}\right)^2 + \left(\frac{y-y_c}{R}\right)^2 - 1} \right) \mathbb{1}_D(x, y)$ .

La fréquence sera prise à  $2\pi$ , la source au centre d'étoilage du L et le rayon sera pris en  $\frac{\lambda}{10}$ . L'apparente discontinuité des résultats de la figure (2.7) est due à de forts gradients de  $\rho'_0$  mal pris en compte tant par la géométrie du maillage que par la visualisation des résultats.

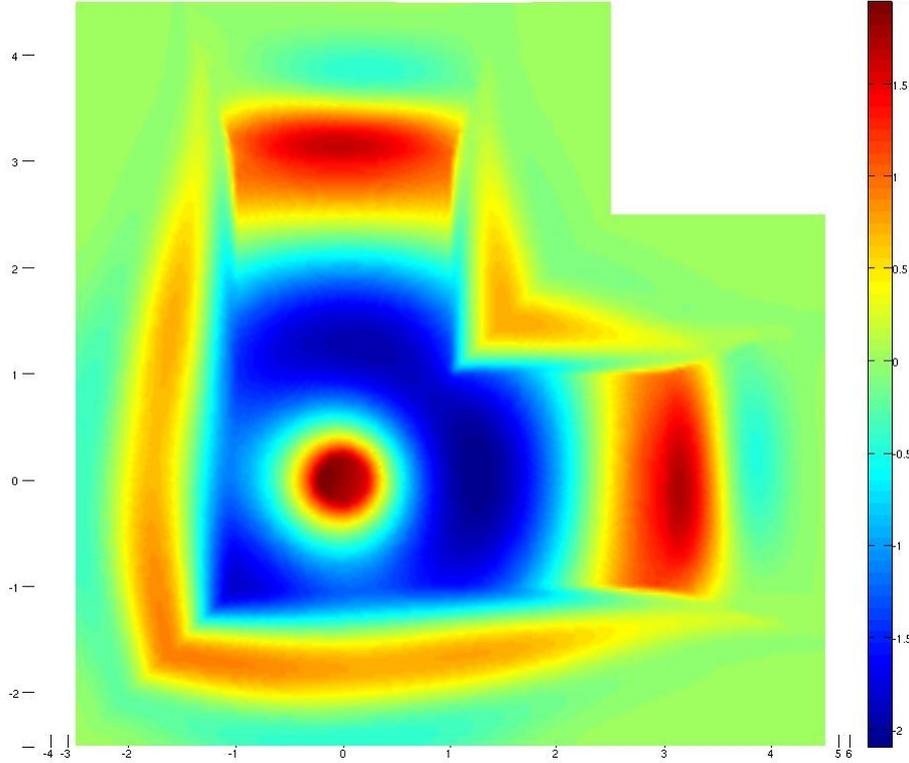


FIG. 2.7: Partie Imaginaire de  $H_z$

Dans le cas des PML hybrides, la proposition (2.28) donne la matrice sous la forme

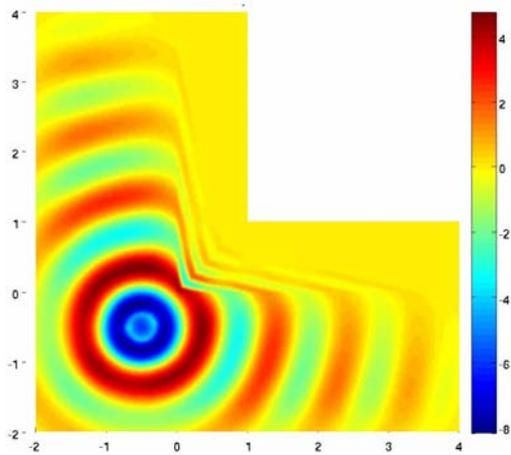
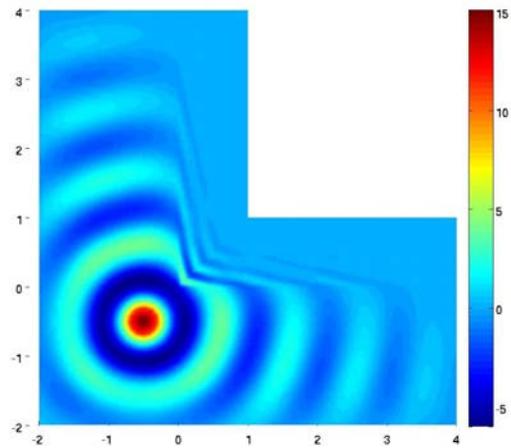
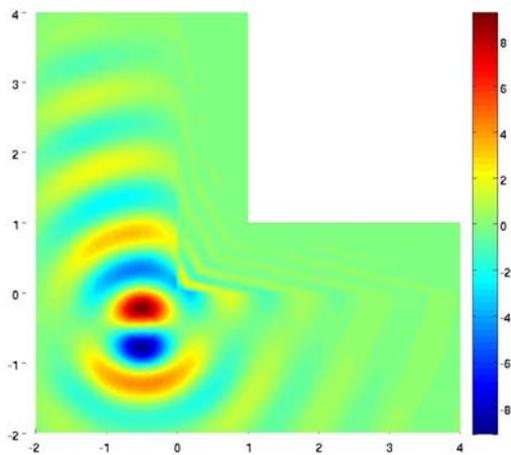
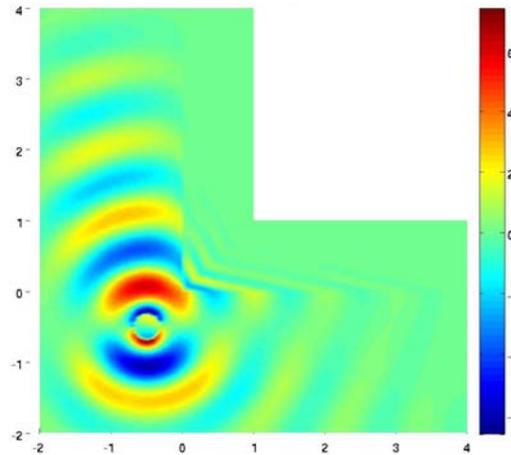
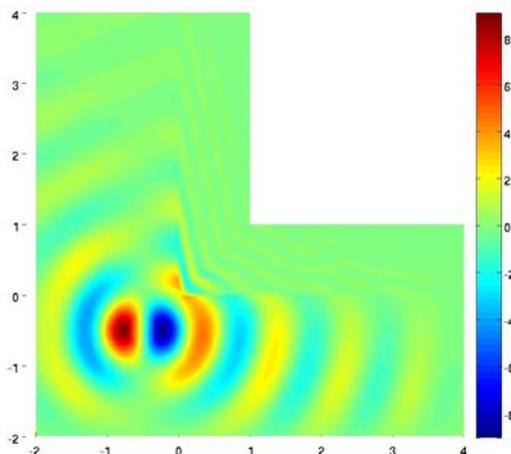
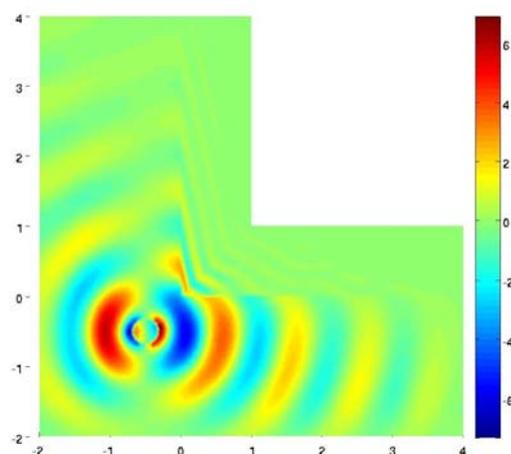
$$\mathbf{T}(\mathbf{x}) (\mathbf{p} \mathbf{I} + \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x}) (\mathbf{p} \mathbf{I}_3 + \mathbf{F}(\mathbf{x}))^{-1}) \mathbf{T}(\mathbf{x})^T.$$

**Proposition 2.68**

La matrice PML complexe se décompose sur  $\mathbb{R}^2$  sous la forme :

$$\begin{cases} \operatorname{Re}(\mathbf{K}) = \mathbf{T}(\mathbf{A} + \mathbf{BFD}) \mathbf{T}^T \\ \operatorname{Im}(\mathbf{K}) = \omega \mathbf{T}(\mathbf{I} - \mathbf{BD}) \mathbf{T}^T(\mathbf{x}) \end{cases} \quad \mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{\omega^2 + \sigma_x^2} & 0 & 0 \\ 0 & \frac{1}{\omega^2 + \sigma_y^2} & 0 \\ 0 & 0 & \frac{1}{\omega^2} \end{pmatrix}$$

En augmentant la fréquence par rapport à la figure (2.7), on obtient pour les trois composantes du mode transverse électrique ( $E_x, E_y, H_z$ ) les figures suivantes. Il est possible d'y observer très nettement les transports d'information et la vive torsion des rayons lumineux.

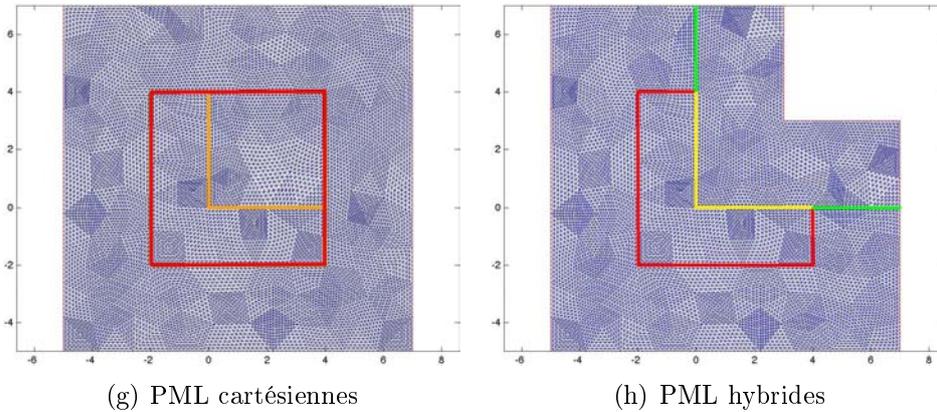
(a) Partie Imaginaire de  $H_z$ (b) Partie Réelle de  $H_z$ (c) Partie Imaginaire de  $E_x$ (d) Partie Réelle de  $E_x$ (e) Partie Imaginaire de  $E_y$ (f) Partie Réelle de  $E_y$

**Proposition 2.69**

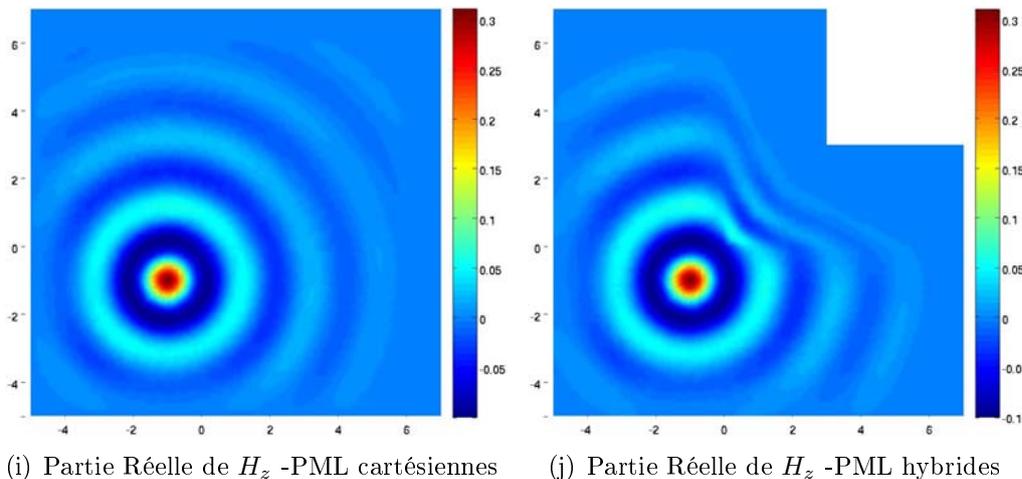
La matrice PML complexe pour la deuxième formulation hybride se décompose sur  $\mathbb{R}^2$  sous la forme :

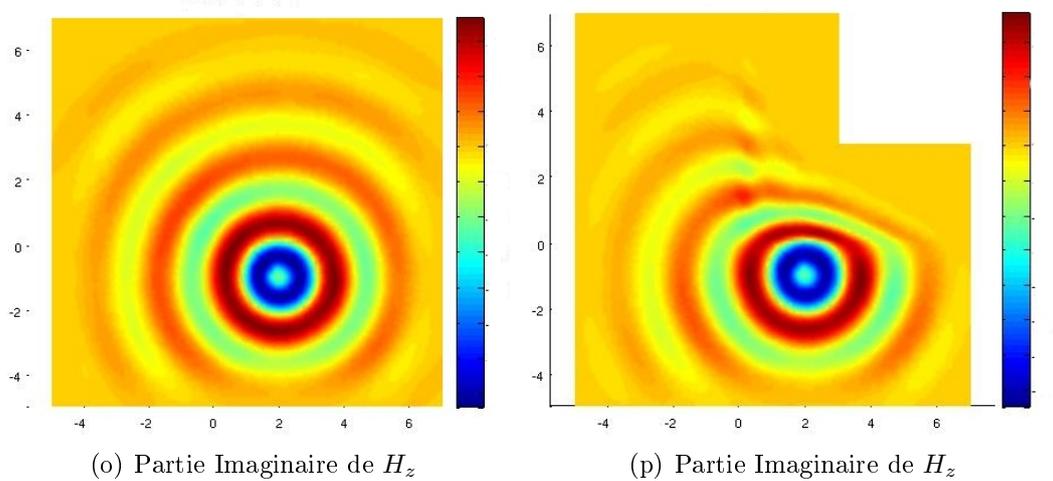
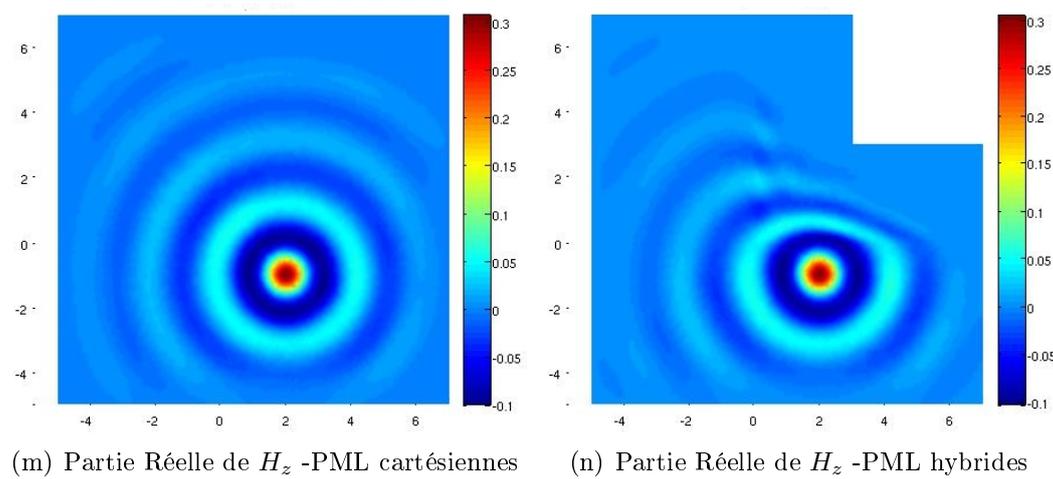
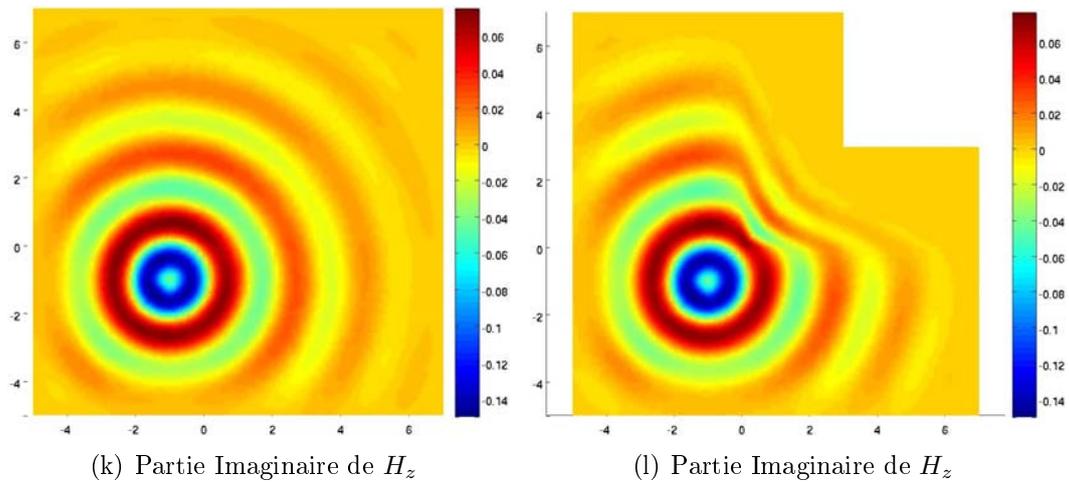
$$\operatorname{Re}(K) = T \left( A + \frac{\sigma}{4\sigma^2 + \omega^2} B \right) T^T \quad \operatorname{Im}(K) = \omega T \left( I - \frac{1}{4\sigma^2 + \omega^2} B \right) T^T(\mathbf{x}).$$

Nous allons désormais réaliser une comparaison entre les PML hybrides et les PML cartésiennes (solution de référence). Soit le maillage déstructuré proposé par la figure (g). En rouge : l'interface milieu réel - PML cartésiennes. En orange : frontière transparente milieu - milieu. Il s'agit de la zone ajoutée par rapport à des PML non convexes. Ce maillage fonctionne de paire avec celui de la figure (h) : ils présentent une zone commune, le L intérieur. En vert : l'interface PML non convexes - PML cartésiennes et en jaune l'interface milieu - PML non convexes.



Dans un premier temps, on place la condition initiale (régularisée du Dirac) au centre d'étoile du L. Ensuite, la source est décentrée. Ainsi le rectangle vertical du L ne communique pas directement avec la source. Cela permet de visualiser la torsion des rayons lumineux et de vérifier qu'il y a bien apport d'information sans ajout d'erreurs numériques ni retard ni réflexions parasites : les restriction de la solution au domaine commun sont comparables.





Nous allons maintenant nous intéresser au cas du problème de scattering. Soit un obstacle parfait  $\Gamma \subset \Omega$  où  $\Omega \subset \mathbb{R}^2$  est le domaine de vide étudié de forme L. On utilisera un maillage déstructuré composé de  $N$  triangles  $(\omega_k)_{k \in \llbracket 1, N \rrbracket}$  non dégénéré, représenté sur la figure (2.8).

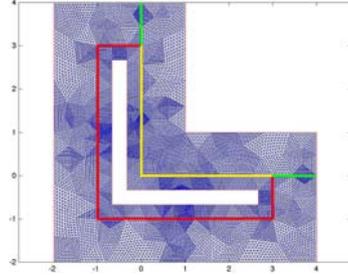
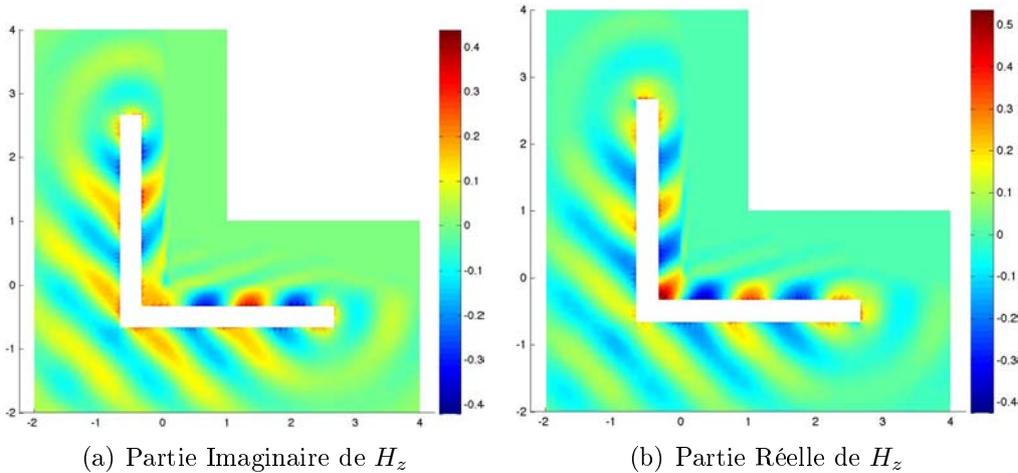


FIG. 2.8: Maillage pour un problème de scattering

Nous allons étudier la propagation d'une onde plane incidente  $\varphi_{inc}$  réfléchi par  $\Gamma$  de la forme  $\varphi_{inc}(\mathbf{x}) = \varphi_0 e^{\omega(-\mathbf{k} \cdot \mathbf{x})}$ . Comme  $\varphi_0$  est solution des Equations dans le vide, on en déduit que  $\varphi_0 = (-k_y, k_x, 1)^T$ , où  $\mathbf{k} = (k_x, k_y)^T$  est le vecteur directeur de l'onde incidente. Pour les conditions aux limites, on choisit sur  $\partial\Omega$  des conditions aux limites transparentes, et sur  $\partial\Gamma$  des conditions réfléchissantes, soit donc :

$$\begin{cases} \forall \mathbf{x} \in \partial\Gamma, & f(\mathbf{x}) = \int_{\partial\omega_k \cap \partial\Gamma} -M_1 \varphi_{inc}(\mathbf{x}) \mathbf{exp}(-\omega \mathbf{k} \cdot \mathbf{x}), \\ \forall \mathbf{x} \in \partial\Omega, & \mathbf{A}_i \mathbf{n}_i^-(\varphi^+(\mathbf{x})) = 0. \end{cases}$$

Nous avons déjà expliqué que le domaine PML non convexe était plus grand que nécessaire (2.4). Sur les résultats numériques obtenus pour le problème de scattering, la « zone inutile » apparaît très clairement (figures (a)). Le vecteur incident choisi pour ces expériences est  $\mathbf{k} = \left( \frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} \right)^T$ .



(a) Partie Imaginaire de  $H_z$

(b) Partie Réelle de  $H_z$

### 2.10.4 Problème instationnaire

**Proposition 2.70**

Les Equations PML cartésiennes pour les Equations de Maxwell 2D sont données par

$$\begin{cases} \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \mathbf{A} \varphi(\mathbf{x}, t) + \mathbf{B} \psi(\mathbf{x}, t) + (\mathbf{A}_1 \partial_x + \mathbf{A}_2 \partial_y) \varphi(\mathbf{x}, t) = f(\mathbf{x}, t) \\ \frac{\partial \psi(\mathbf{x}, t)}{\partial t} + \mathbf{F} \psi(\mathbf{x}, t) - \varphi(\mathbf{x}, t) = 0. \end{cases}$$

avec

$$\begin{aligned} \mathbf{A} &= \mathbf{diag}(\sigma_y - \sigma_x, \sigma_x - \sigma_y, \sigma_x + \sigma_y) \\ \mathbf{F} &= \mathbf{diag}(\sigma_x, \sigma_y, 0) \\ \mathbf{B}_1 &= \mathbf{diag}(\sigma_x(\sigma_x - \sigma_y), -\sigma_y(\sigma_x - \sigma_y), \sigma_x \sigma_y), \end{aligned}$$

et les matrices  $\mathbf{A}_i$  dépendant du choix du mode transverse étudié. Par la suite, nous considérerons le mode transverse électrique, et les matrices  $\mathbf{A}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$  et

$$\mathbf{A}_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Le système étant bien posé, on en cherche maintenant une approximation. Contrairement au cas de conditions aux limites classiques, il nous faut considérer un couple d'équations, chacune étant approchées par un schéma numérique convergent et stable. Cependant, la stabilité de chacun des deux schémas ne permet pas de conclure quant à celle de leur couplage. C'est pourquoi on préférera utiliser, lorsque cela n'est pas trop contraignant, la formulation implicite des schémas, beaucoup plus stable que l'explicite. Soit un maillage  $\bigcup \omega_e$  déstructuré où les éléments seront supposés non dégénérés.

$$\begin{cases} \frac{\partial \varphi_{\omega_e}(t)}{\partial t} + \mathbf{A} \varphi_{\omega_e}(t) + \mathbf{B} \psi_{\omega_e}(t) = F(\varphi_{\omega_e}(t)) \\ \frac{\partial \psi_{\omega_e}(t)}{\partial t} + \mathbf{F} \psi_{\omega_e}(t) - \varphi_{\omega_e}(t) = 0. \end{cases}$$

où  $F(\varphi_{\omega_e}(t))$  représente la contribution du terme  $-(\mathbf{A}_1 \partial_x + \mathbf{A}_2 \partial_y) \varphi(\mathbf{x}, t)$  (cf. sous section précédente sur la formulation variationnelle) et du second membre calculé pour l'élément  $\omega_e$ . Soit  $(\varphi^n, \psi^n)$  les solutions à l'instant  $n$  et  $\Delta t$  le pas de discrétisation temporelle.

**Proposition 2.71**

La solution pour un schéma de type Euler Explicite de la formulation PML cartésiennes des Equations de Maxwell est donnée par :

$$\begin{pmatrix} \varphi_{\omega_e}^{n+1} \\ \psi_{\omega_e}^{n+1} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_3 - \Delta t \mathbf{A} & -\Delta t \mathbf{B} \\ \Delta t \mathbf{I}_3 & \mathbf{I}_3 - \Delta t \mathbf{F} \end{pmatrix} \begin{pmatrix} \varphi_{\omega_e}^n \\ \psi_{\omega_e}^n \end{pmatrix} + \begin{pmatrix} \Delta t F(\varphi_{\omega_e}^n) \\ \mathbf{O}_3 \end{pmatrix}$$

**Démonstration :**

$$\begin{cases} \frac{\varphi_{\omega_e}^{n+1} - \varphi_{\omega_e}^n}{\Delta t} + \mathbf{A} \varphi_{\omega_e}^n + \mathbf{B} \psi_{\omega_e}^n = F(\varphi_{\omega_e}^n) \\ \frac{\psi_{\omega_e}^{n+1} - \psi_{\omega_e}^n}{\partial t} + \mathbf{F} \psi_{\omega_e}^n - \varphi_{\omega_e}^n = 0. \end{cases}$$

D'où le résultat proposé. \(\square\)

**Proposition 2.72**

La solution pour un schéma de type Euler Implicite de la formulation PML cartésiennes des Equations de Maxwell est donnée par :

$$\begin{pmatrix} \varphi_{\omega_e}^{n+1} \\ \psi_{\omega_e}^{n+1} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_3 + \Delta t \mathbf{A} & \Delta t \mathbf{B} \\ -\Delta t \mathbf{I}_3 & \mathbf{I}_3 + \Delta t \mathbf{F} \end{pmatrix}^{-1} \left( \begin{pmatrix} \varphi_{\omega_e}^n \\ \psi_{\omega_e}^n \end{pmatrix} + \begin{pmatrix} \Delta t F(\varphi_{\omega_e}^n) \\ \mathbf{O}_3 \end{pmatrix} \right)$$

**Démonstration :** à la différence de la formulation explicite, l'écriture de ce schéma nécessite une inversion matricielle mais garantit la stabilité.

$$\begin{cases} \frac{\varphi_{\omega_e}^{n+1} - \varphi_{\omega_e}^n}{\Delta t} + \mathbf{A} \varphi_{\omega_e}^{n+1} + \mathbf{B} \psi_{\omega_e}^{n+1} = F(\varphi_{\omega_e}^n) \\ \frac{\psi_{\omega_e}^{n+1} - \psi_{\omega_e}^n}{\partial t} + \mathbf{F} \psi_{\omega_e}^{n+1} - \varphi_{\omega_e}^{n+1} = 0. \end{cases}$$

D'où le système matriciel :

$$\begin{pmatrix} \frac{1}{\Delta t} \mathbf{I}_3 + \mathbf{A} & \mathbf{B} \\ -\mathbf{I}_3 & \frac{1}{\Delta t} \mathbf{I}_3 + \mathbf{F} \end{pmatrix} \begin{pmatrix} \varphi_{\omega_e}^{n+1} \\ \psi_{\omega_e}^{n+1} \end{pmatrix} = \begin{pmatrix} \frac{1}{\Delta t} \mathbf{I}_3 & \mathbf{O}_3 \\ \mathbf{O}_3 & \frac{1}{\Delta t} \mathbf{I}_3 \end{pmatrix} \begin{pmatrix} \varphi_{\omega_e}^n \\ \psi_{\omega_e}^n \end{pmatrix} + \begin{pmatrix} F(\varphi_{\omega_e}^n) \\ \mathbf{O}_3 \end{pmatrix}$$

\(\square\)

Les résultats pour les formulations instationnaires seront présentés pour des problèmes de scattering. Soit un obstacle parfait  $\Gamma \subset \mathbb{R}^2$ . Soit  $[0, T]$  l'intervalle de simulation. Soit  $\Omega \subset \mathbb{R}^2$  le domaine de vide étudié. Le maillage est identique à celui utilisé pour la formulation harmonique, il est déstructuré et composé de  $N$  triangles  $(\omega_k)_{k \in [1, N]}$ . Nous allons étudier la propagation d'une onde plane incidente  $\varphi_{inc}$  réfléchiée par un obstacle parfait  $\Gamma$ . Formulation du problème :

$$\begin{aligned} \iint_{\omega_k} (\partial_t \varphi + \mathbf{A}_i \partial_i \varphi) + \int_{\partial \omega_k} \mathbf{A}_i \mathbf{n}_i^- (\varphi^+ - \varphi) &= \iint_{\omega_k} f = \int_{\partial \omega_k \cap \partial \Gamma} \mathbf{M}_1(\varphi + \varphi_{inc}), \\ \varphi_{inc}(\mathbf{x}, t) &= \varphi_0 \mathbf{sin}(\omega(t - \mathbf{k} \cdot \mathbf{x} + \phi)) \mathbf{1}_{t \leq t_\phi}. \end{aligned}$$

Comme  $\varphi_0$  est solution de l'équation des ondes dans le vide, on en déduit que  $\varphi_0 = (-k_y, k_x, 1)^T$ , où  $\mathbf{k} = (k_x, k_y)^T$  est le vecteur directeur de l'onde incidente. Par ailleurs, on choisit les grandeurs caractéristiques du maillage de façon à obtenir

une longueur d'onde pour la largeur de ce dernier, donc  $\lambda = 1$ , ce qui implique une fréquence  $\omega = 2\pi$ . Pour la phase  $\phi$ , on souhaite  $\forall \mathbf{x} \in \Omega$ ,  $\varphi_{inc}(\mathbf{x}, 0) = 0$ . Soit  $\mathbf{x}_i = (x_i, y_i)^T$  le premier point de l'objet  $\Gamma$  touché. Alors on aura  $\phi = 2\pi \mathbf{k} \cdot \mathbf{x}_i$ , d'où l'onde incidente  $\varphi_{inc}(\mathbf{x}, t) = \varphi_0 \sin(2\pi(t - \mathbf{k} \cdot \mathbf{x} + \phi) \mathbb{1}_{t \leq \mathbf{k} \cdot (\mathbf{x} - \mathbf{x}_i)})$ . Pour les conditions aux limites, on choisit sur  $\partial\Omega$  des conditions aux limites transparentes, et sur  $\partial\Gamma$  des conditions réfléchissantes, soit donc :

$$\begin{cases} \forall (\mathbf{x}, t) \in \partial\Gamma \times [0, T], & f(\mathbf{x}, t) = \int_{\partial\omega_{\mathbf{k}} \cap \partial\Gamma} \mathbf{M}_1(\varphi(\mathbf{x}, t) + \varphi_{inc}(\mathbf{x}, t)), \\ \forall (\mathbf{x}, t) \in \partial\Omega \times [0, T], & \mathbf{A}_i \mathbf{n}_i^-(\varphi^+(\mathbf{x}, t)) = 0. \end{cases}$$

Le choix d'une formulation implicite a été fait. Le vecteur incident choisi est  $\mathbf{k} = (-1/\sqrt{2}, -1/\sqrt{2})^T$ .

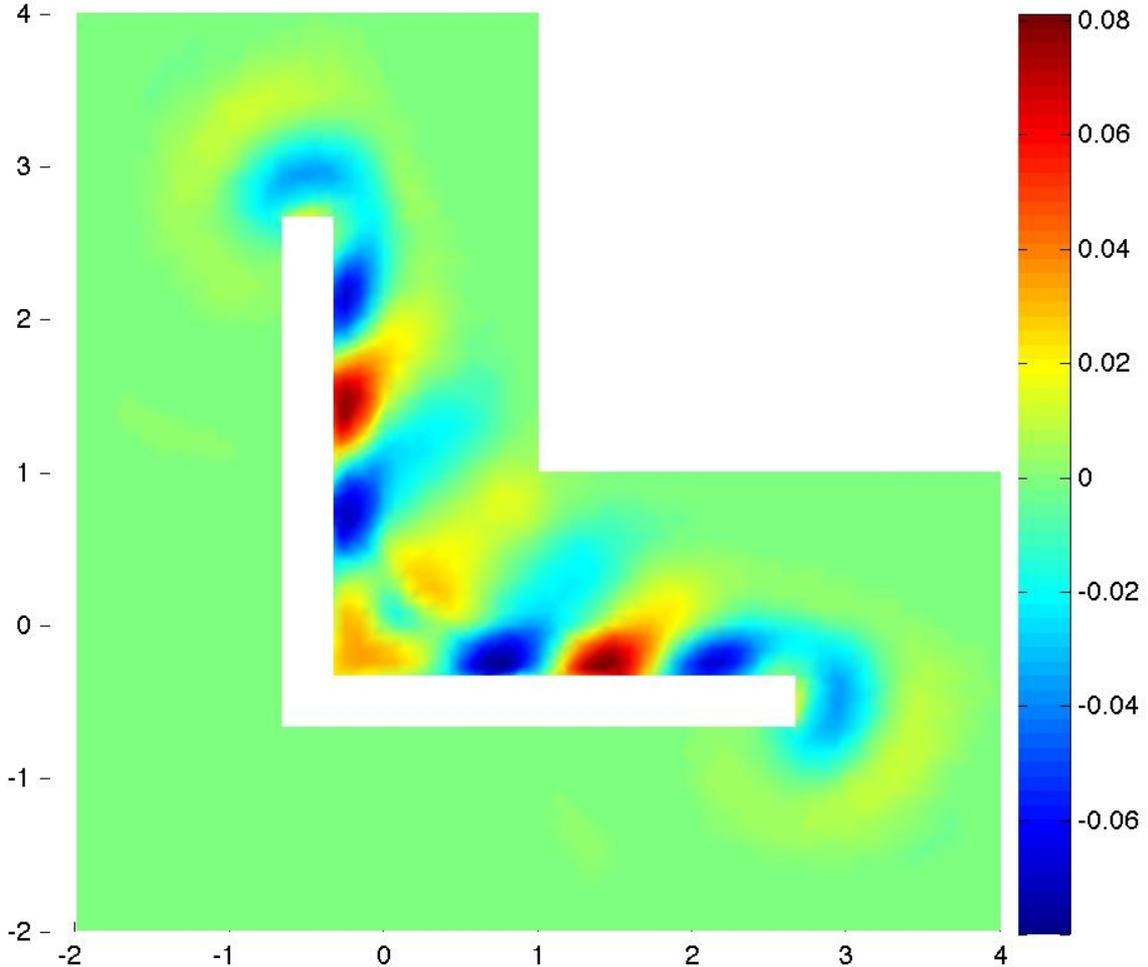


FIG. 2.9:  $H_z$  au temps  $t_3$



# Chapitre 3

## Une nouvelle Approximation de Haute Précision

Un schéma numérique de type Volumes Finis garantit une bonne stabilité, une bonne convergence, et n'impose que peu de contrainte sur le choix de la condition CFL. Mais le choix d'approximer la solution par une constante sur chaque cellule n'est pas toujours acceptable. Une bonne méthode généralement utilisée pour monter en ordre est le Galerkin Discontinu. Après calcul des conditions CFL pour un problème d'advection simple monodimensionnel à coefficients constants avec différentes bases polynômiales, nous ferons ensuite une étude de complexité et de coût d'un schéma Galerkin Discontinu en espace et Runge-Kutta en temps dans le cadre général. Prouvant ainsi que ce type de schéma s'avère coûteux, nous présenterons une nouvelle approximation de haute précision qui saura conserver des conditions CFL correctes, dérivée de la méthode des Volumes Finis.

### 3.1 Méthodes de Galerkin Discontinu : complexité et conditions CFL

Soit  $\Omega$  un segment de  $\mathbb{R}$  et  $T \in \mathbb{R}^+ \setminus \{0\}$ . Considérons comme cas d'étude l'équation d'advection monodimensionnelle :

$$\forall (x, t) \in \Omega \times [0, T], \quad \frac{\partial u(x, t)}{\partial t} + a \frac{\partial u(x, t)}{\partial x} = f(x, t). \quad (3.1)$$

Soit  $h \in \mathbb{R}^+ \setminus \{0\}$  le pas de discrétisation spatiale. Soit  $(\omega_k)_{k \in \llbracket 1, N \rrbracket}$  la partition de  $\Omega$  associée. Chaque  $\omega_k$  est alors un segment  $[\alpha_k, \beta_k]$  tel que  $\beta_k - \alpha_k = h$ , et on a  $\Omega = \bigcup_k \omega_k$ . Soit  $\varphi \in C^1(\mathbb{R})$  une fonction test. Si  $u \in C^1(H_{mcx}^1(\Omega) \times [0, T])$ , notons  $u_h(x, t)$  la solution approximée et  $k \in \llbracket 1, N \rrbracket$ . Rappelons la formulation variationnelle (2.48) de Galerkin Discontinu que nous allons utiliser, établie dans la section (2.10.2) pour un système de Friedrichs général en dimension  $n$  :

$$\int_{Ru_h \cap R\varphi} ((\mathbf{A}_0 \partial_t + \mathbf{A}^i \partial_i + \mathbf{B})u_h, \varphi) + \int_{\widetilde{Su_h \cup S\varphi}} (\mathbf{A}^i \mathbf{n}_i^- (u_h^+ - u_h^-), \varphi^-) = \int_{\Omega} (f, \varphi).$$

Ici  $\mathbf{A}_0 = 1$ ,  $\mathbf{A}_1 = a$  et  $\mathbf{B} = 0$ .

La normale sortante gauche à l'élément  $k$  (au niveau de  $\alpha_k$ ) est  $n = -1$ . A droite, pour  $\beta_k$ , la normale est  $n = 1$ . De fait,  $\mathbf{A}^i \mathbf{n}_i^- = \mathbf{inf}(-a, 0)$  en  $\alpha_k$  et  $\mathbf{A}^i \mathbf{n}_i^+ = \mathbf{inf}(a, 0)$  en  $\beta_k$ .

Donc, si  $u \in C^1(H_{m_{cx}}^1(\Omega) \times [0, T])$ , la formulation variationnelle de Galerkin Discontinu associée à (3.1) que nous considérons est :

$$\int_{\omega_k} \left( \frac{\partial}{\partial t} u_h + a \frac{\partial}{\partial x} u_h, \varphi \right) + \mathbf{inf}(a, 0)(u_h(\beta_k + 0) - u_h(\beta_k - 0))\varphi(\beta_k - 0) \\ + \mathbf{sup}(a, 0)(u_h(\alpha_k + 0) - u_h(\alpha_k - 0))\varphi(\alpha_k + 0) = \int_{\omega_k} (f, \varphi).$$

Dans le cas où  $a > 0$  qui est celui que nous considérerons dans la suite (sans perte de généralité), cette formulation s'écrit plus simplement avec  $u_h^+ = u_h(\alpha_k + 0)$  et  $u_h^- = u_h(\alpha_k - 0)$  pour tout élément  $k \in \llbracket 1, N \rrbracket$

$$\int_{\omega_k} \left( \frac{\partial}{\partial t} u_h + a \frac{\partial}{\partial x} u_h, \varphi \right) + a(u_h^+ - u_h^-)\varphi^- = \int_{\omega_k} (f, \varphi). \quad (3.2)$$

### 3.1.1 Calcul de la condition CFL

#### **Théorème 3.1**

La condition CFL pour l'équation (3.1) approximée avec une méthode de type Volumes Finis en espace et Euler Explicite en temps est :

$$\Delta t \leq \frac{h}{a}.$$

**Démonstration :** Pour une approximation spatiale de type Volumes Finis, notons  $\forall k \in \llbracket 1, N \rrbracket$ ,  $u_k(t)$  la solution approchée par une constante sur l'élément  $\omega_k$ . (3.2) devient alors, en prenant une fonction test constante par morceaux égale à 1 :

$$\forall k \in \llbracket 1, N \rrbracket, h \frac{d}{dt} u_k + a(u_k(t) - u_{k-1}(t)) = f_k.$$

Soit  $\Delta t \in \mathbb{R}^+ \setminus \{0\}$  le pas de discrétisation temporelle. Soit  $\cup_n [n\Delta t, (n+1)\Delta t]$  la partition de  $[0, T]$  associée. Soit  $N_t = \frac{T}{\Delta t}$ . Avec une approximation temporelle de type Euler Explicite, l'équation précédente devient :

$$\forall (k, n) \in \llbracket 1, N \rrbracket \times \llbracket 1, N_t \rrbracket, h \frac{u_k^{(n+1)} - u_k^{(n)}}{\Delta t} + a(u_k^{(n)} - u_{k-1}^{(n)}) = f_k^{(n)}.$$

□

**Définition 3.1**

Soit le semi discrétisé en espace d'un système d'Équations aux Dérivées Partielles de type :

$$\frac{du}{dt} = \mathbf{L}(u).$$

On dit que le schéma d'Euler Explicite est fortement stable, si pour une méthode d'approximation spatiale donnée (où l'opérateur linéaire de discrétisation spatiale est noté  $\mathbf{L}$ ), il existe un  $\Delta t_E$  strictement positif et une certaine norme tels que :

$$\forall \Delta t < \Delta t_E, \quad \| u^n + \Delta t \mathbf{L}(u^n) \| \leq (1 + \mathcal{O}(\Delta t)) \| u^n \|. \quad (3.3)$$

Habituellement, la norme de dépend pas du maillage. Néanmoins, un problème avec les méthodes de Galerkin Discontinu est justement que l'on soit amené à faire dépendre la norme du pas de discrétisation spatiale. On adoptera ici les mêmes techniques que [9].

Soit  $\mathbf{U}^{(n)} = (u_k^{(n)})_{k \in [1, N]}$  le vecteur de la solution discrétisée. Soit  $\mathbf{F}^{(n)}$  le second membre . Soit  $\mathbf{I}_N$  la matrice identité et  $\mathbf{J}_N$  la matrice de Jordan, à savoir la matrice des  $\delta_{i+1, j}$ . Posons  $\mathbf{A} = (1 - a \frac{\Delta t}{h}) \mathbf{I}_N + a \frac{\Delta t}{h} \mathbf{J}_N$ . L'équation peut alors s'écrire sous la forme matricielle suivante :

$$\mathbf{U}^{(n+1)} = \mathbf{A} \mathbf{U}^{(n)} + \frac{\Delta t}{h} \mathbf{F}^{(n)}.$$

On souhaite que quand  $n$  tende vers l'infini, la norme  $\| \mathbf{U}^{(n)} \|$  reste bornée dans  $\mathbb{R}^N$ .

**Définition 3.2**

On appelle  $\| \cdot \|_\gamma$  norme subordonnée à la norme  $\gamma$  de  $\mathbb{R}^N$  la norme définie par  $\inf_{X \in \mathbb{R}^n, X \neq 0} \frac{|\mathbf{A} X|_\gamma}{|X|_\gamma} = \| \mathbf{A} \|_\gamma$ . Pour les matrices, la notion de norme subordonnée est équivalente à la notion de norme sous multiplicative, soit :

$$\forall (\mathbf{A}, \mathbf{B}), \quad \| \mathbf{A} \mathbf{B} \| \leq \| \mathbf{A} \| \| \mathbf{B} \|.$$

**Théorème 3.2**

Soit  $\mathbf{A}$  une matrice carrée de dimension  $N$ . Il existe une norme subordonnée telle que :

$$\forall \varepsilon > 0, \quad \| \mathbf{A} \| \leq \rho(\mathbf{A}) + \varepsilon,$$

où  $\rho(\mathbf{A})$  représente le rayon spectral de la matrice  $\mathbf{A}$ .

L'inégalité (3.3) sera donc vérifiée si il existe un  $\Delta t_E$  tel que :

$$\forall \Delta t < \Delta t_E, \quad \| \mathbf{A} \| \leq (1 + C)\Delta t,$$

soit encore sous la condition  $\forall \Delta t < \Delta t_E, \quad \rho(\mathbf{A}) \leq (1 + C)\Delta t$ , puisque dans ce cas, on pourra assurer l'existence d'une norme subordonnée telle que  $\| \mathbf{A} \| \leq \rho(\mathbf{A})$ .

Pour que ce schéma soit stable il faut que la matrice  $\mathbf{A}$  soit positive. Il faut donc garantir que  $(1 - a \frac{\Delta t}{h}) \mathbf{I}_N$  soit positive, donc que  $(1 - a \frac{\Delta t}{h}) \geq 0$ , d'où le résultat du théorème.  $\square$

Pour une méthode de type Galerkin Discontinu de degré  $k$ , il faut choisir des fonctions de base polynômiales de degré  $k$ . Nous allons dans un premier temps choisir des polynômes de Lagrange ou les monômes. Une telle méthode entraîne l'apparition de matrices dites matrice de masse  $\mathbf{M}$  et matrice de rigidité  $\mathbf{K}$ . Le système différentiel (3.2) s'écrit alors sous la forme matricielle suivante :

$$\mathbf{M} \frac{dU}{dt} + \mathbf{K}U = F. \quad (3.4)$$

Posons  $\mathbf{A} = \mathbf{I}_N - \Delta t \mathbf{M}^{-1} \mathbf{K}$ . Le système (3.4) devient sous une approximation de type Euler Explicite :

$$\mathbf{U}^{(n+1)} = \mathbf{A} \mathbf{U}^{(n)} + \Delta t \mathbf{F}^{(n)}$$

Il faut donc pour assurer la stabilité de l'approximation, avoir  $\rho(\mathbf{A}) \leq 1$ , et cette inégalité déterminera la condition CFL. On utilisera la propriété que la matrice de rigidité globale est triangulaire par blocs pour le calcul spectral (cf démonstration du théorème suivant).

### **Théorème 3.3**

*La condition CFL pour l'équation (3.1) approximée avec une méthode de type Galerkin Discontinu d'ordre 1 avec une base de polynômes de type Lagrange et Euler Explicite en temps est :*

$$\Delta t \leq \frac{2}{3} \frac{h}{a}.$$

**Démonstration :** Pour une base de type Lagrange, à l'ordre 1, les fonctions de bases sont :

$$\begin{cases} b_1(x) = -\frac{x}{h} + \frac{1}{2}, \\ b_2(x) = \frac{x}{h} + \frac{1}{2}. \end{cases}$$

Soit  $(u_k^{(1)}(t), u_k^{(2)}(t))$  les valeurs de la solution approchée  $u_k(t)$  en  $-h/2$  et  $h/2$  et  $\varphi$  une fonction test. On a alors :

$$\begin{cases} u_k(t) &= u_k^{(1)}(t) b_1(x) + u_k^{(2)}(t) b_2(x) \\ \partial_t u \cdot \varphi &:= \begin{pmatrix} b_1(x)^2 & b_1(x)b_2(x) \\ b_1(x)b_2(x) & b_2(x)^2 \end{pmatrix} \begin{pmatrix} \frac{du_k^{(1)}(t)}{dt} & \frac{du_k^{(2)}(t)}{dt} \end{pmatrix}^T \\ \partial_x u \cdot \varphi &:= (u_k^{(1)}(t) \partial_x b_1(x) + u_k^{(2)}(t) \partial_x b_2(x)) = \begin{pmatrix} -b_1(x)/h & b_1(x)/h \\ -b_2(x)/h & -b_2(x)/h \end{pmatrix} \begin{pmatrix} u_k^{(1)}(t) \\ u_k^{(2)}(t) \end{pmatrix} \end{cases}$$

La formulation variationnelle appliquée à un élément  $\omega_k$  donne

$$\begin{aligned} & \int_{-h/2}^{h/2} \begin{pmatrix} b_1(x)^2 & b_1(x)b_2(x) \\ b_1(x)b_2(x) & b_2(x)^2 \end{pmatrix} dx \cdot \left( \frac{du_k^{(1)}(t)}{dt}, \frac{du_k^{(2)}(t)}{dt} \right)^T \\ & + a \int_{-h/2}^{h/2} \begin{pmatrix} -b_1(x)/h & b_1(x)/h \\ -b_2(x)/h & -b_2(x)/h \end{pmatrix} dx \cdot \left( u_k^{(1)}(t), u_k^{(2)}(t) \right)^T \\ & + a \begin{pmatrix} b_1(-h/2)^2 & b_1(-h/2)b_2(-h/2) \\ b_1(-h/2)b_2(-h/2) & b_2(-h/2)^2 \end{pmatrix} \cdot \left( u_k^{(1)}(t), u_k^{(2)}(t) \right)^T \\ & + a \begin{pmatrix} b_1(h/2)^2 & b_1(h/2)b_2(h/2) \\ b_1(h/2)b_2(h/2) & b_2(h/2)^2 \end{pmatrix} \cdot \left( u_{k-1}^{(1)}(t), u_{k-1}^{(2)}(t) \right)^T = \int_{-h/2}^{h/2} f \varphi. \end{aligned}$$

La contribution des termes de bord amène donc  $a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_k^{(1)}(t) \\ u_k^{(2)}(t) \end{pmatrix}$  et une partie dépendant de l'élément voisin précédent uniquement  $a \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} u_{k-1}^{(1)}(t) \\ u_{k-1}^{(2)}(t) \end{pmatrix}$ .

La matrice globale de la formulation sera donc triangulaire par blocs. Pour calculer son rayon spectral, on ne s'intéressera donc qu'à la partie diagonale par bloc, c'est-à-dire la contribution des inconnues intérieures à l'élément même pour le terme de bord. Les blocs composant les matrices de masse et rigidité sont donnés par :

$$\left\{ \begin{array}{l} \mathbf{M} = \int_{-h/2}^{h/2} \begin{pmatrix} b_1(x)^2 & b_1(x)b_2(x) \\ b_1(x)b_2(x) & b_2(x)^2 \end{pmatrix} dx \\ \mathbf{K} = a \int_{-h/2}^{h/2} \begin{pmatrix} -b_1(x)/h & b_1(x)/h \\ -b_2(x)/h & -b_2(x)/h \end{pmatrix} dx \\ \quad + a \begin{pmatrix} b_1(-h/2)^2 & b_1(-h/2)b_2(-h/2) \\ b_1(-h/2)b_2(-h/2) & b_2(-h/2)^2 \end{pmatrix}. \end{array} \right.$$

$$\text{D'où } \left\{ \begin{array}{l} \mathbf{M} = \frac{h}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \\ \mathbf{K} = \frac{a}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \end{array} \right. \text{ Posons } \lambda = \frac{a \Delta t}{h}, \text{ on obtient } \mathbf{A} = \begin{pmatrix} 3\lambda - 1 & \lambda \\ -3\lambda & \lambda - 1 \end{pmatrix}.$$

La condition  $\rho(\mathbf{A}) \leq 1$  devient alors  $6\lambda^2 - 4\lambda + 1 \leq 1$  soit donc  $\lambda(6\lambda - 4) \leq 0$ . Comme la condition CFL est non nulle, on obtient  $\lambda \leq 2/3$ .  $\square$

#### **Théorème 3.4**

*La condition CFL pour l'équation (3.1) approximée avec une méthode de type Galerkin Discontinu d'ordre 1 avec la base naturelle des polynômes et Euler Explicite en temps est :*

$$\Delta t \leq \frac{2}{3} \frac{h}{a}.$$

**Démonstration** : soit  $b_1(x) = 1$  et  $b_2(x) = x$ . On note  $(a_k(t), b_k(t))$  la valeur de la pente et la valeur au centre de l'élément  $\omega_k$ . Alors  $u_k(t) = a_k(t)\frac{x}{h} + b_k(t)$ . Les matrices de masse et rigidité deviennent :

$$\begin{cases} \mathbf{M} = \int_{-h/2}^{h/2} \begin{pmatrix} 1 & 0 \\ 0 & x^2 \end{pmatrix} dx & = \frac{h}{12} \begin{pmatrix} 12 & 0 \\ 0 & 1 \end{pmatrix} \\ \mathbf{K} = a \int_{-h/2}^{h/2} \begin{pmatrix} 0 & 1 \\ 0 & x \end{pmatrix} dx + a \begin{pmatrix} 1 & -h/2 \\ -h/2 & h^2/4 \end{pmatrix} & = \frac{a}{4} \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}. \end{cases}$$

Posons  $\lambda = \frac{a \Delta t}{h}$ , on obtient  $\mathbf{A} = \begin{pmatrix} 1 - \lambda & -\lambda/2 \\ 6\lambda & 1 - 3\lambda \end{pmatrix}$ . La condition de stabilité est donc décrite par l'inéquation  $6\lambda^2 - 4\lambda + 1 \leq 1$ . D'où  $\lambda \leq 2/3$ .  $\square$

Nous savons qu'augmenter l'ordre du schéma d'approximation temporel permet de se rapprocher de la condition CFL des Volumes Finis. Nous allons donc procéder à ces calculs en changeant l'approximation temporelle Euler Explicite par des méthodes de Runge-Kutta. Mais au préalable, étant donné que sur les exemples précédents que les valeurs propres de la matrice de formulation du problème  $\mathbf{M}^{-1}\mathbf{K}$  restent inchangées par rapport aux bases, nous allons démontrer que les matrices sont semblables et amènent une formulation équivalente (et la même condition CFL). Il sera alors possible de choisir la base induisant les calculs les plus simples (ou les plus sympathiques).

**Proposition 3.5 (Equivalence des bases)**

*Soit  $p$  un ordre d'approximation spatiale. Alors, pour toutes les bases polynômiales de degré  $p$ , on obtiendra une unique condition CFL et des matrices de formulation de type  $\mathbf{M}^{-1}\mathbf{K}$  semblables.*

**Démonstration** : soient  $(b_i)$  et  $(b'_j)$  deux bases polynômiales de degré  $p$  fixé. Ces bases (du même espace vectoriel) sont reliées grâce à une matrice de passage à coefficients constants  $\alpha_{ij}$ , donc  $b'_r = \sum \alpha_{rj} b_j$ . Les termes généraux de la matrice de masse sont de la forme  $\int b_i b_j$ ,  $\int b'_i b'_j$ , ceux de la matrice de rigidité de la forme  $\int b_i \frac{\partial b_j}{\partial x}$ ,  $\int b'_i \frac{\partial b'_j}{\partial x}$ . Les intégrales de bords, quant à elles, sont de la forme  $\int b_i$ ,  $\int b'_i$ . On constate donc que les coefficients  $\alpha_{ij}$  vont sortir linéairement de tous ces termes, et formeront les matrices  $\mathbf{P}$  et  $\mathbf{P}^*$  qui relieront les matrices de formulation des différentes bases. Les matrices  $\mathbf{A}$  finales sont semblables donc les rayons spectraux sont identiques. Ainsi, à un ordre d'approximation spatiale  $p$  fixé correspond une unique condition CFL.  $\square$

**Remarque** : on peut donc choisir la base qui nous convient le mieux, et dans notre cas, au vue de la simplicité des calculs engendrés, nous choisirons la base naturelle des polynômes.

Soit  $p$  l'ordre d'approximation spatiale. La base naturelle des polynômes s'écrit alors :  $(1, x, x^2, \dots, x^p)$ . La solution est alors approximée sous la forme :

$$u(x, t) = \sum_{i=0}^p \alpha_{p-i}(t) x^i.$$

Calculons les matrices de masse et de rigidité associées. Il faut faire attention à homogénéiser les formules, c'est pourquoi on effectuera le changement de variables  $\alpha_i = h^i \beta_i$  et on prendra soin de diviser la ligne concernée par le jacobien de la transformation, à savoir le pas d'espace  $h$ . On obtient ainsi les coefficients de la matrice de masse :

$$\begin{cases} m_{ij} = 0 & \text{si } i + j \text{ impair,} \\ m_{ij} = \frac{1}{2^{k+i-j}(p+1+i-j)} & \text{sinon .} \end{cases}$$

La matrice de rigidité se décompose comme somme de deux matrices, celle de la contribution des flux (indice 1) , et celle issue de  $\int \frac{\partial u}{\partial x} \varphi$  (indice 2).

$$\begin{cases} k_{ij}^{(1)} = \left(\frac{-1}{2}\right)^{p+i-j} \\ k_{ij}^{(2)} = 0 & \text{si } i + j \text{ pair,} \\ k_{ij}^{(2)} = \frac{p+1-i}{2^{p-1+i-j}(p+i-j)} & \text{sinon .} \end{cases}$$

On discrétise désormais l'opérateur temporel. Pour cela, nous allons considérer plusieurs schémas de type Runge-Kutta. Ce sont des méthodes à pas, stables et consistantes. Dans [14], les auteurs généralisent des conditions de stabilité obtenues pour des schémas à un ordre d'approximation quelconque en espace et Euler Explicite en temps, à des conditions CFL obtenues pour le même ordre d'approximation spatial mais pour des approximations temporelles d'ordre plus élevées (des méthodes de Runge-Kutta préservant la propriété de stabilité forte des approximations de type Euler Explicite du premier ordre). Le résultat principal énoncé est le suivant :

***Théorème 3.6***

*Si le schéma de résolution est fortement stable pour une méthode de type Euler Explicite, alors pour un  $r \geq 2$  donné, il existe un schéma de Runge-Kutta optimal d'ordre  $r$  qui préserve cette propriété de forte stabilité pour la même norme et pour le même ordre d'approximation spatiale.*

De plus, dans [14], les auteurs déterminent les coefficients à introduire dans l'algorithme de type Runge-Kutta multi-pas tels que l'algorithme choisi d'ordre  $m$  soit unique, précis d'ordre  $m$ , et conduise à la meilleure condition CFL possible.

Pour des schémas de la forme  $\frac{du}{dt} = \mathbf{L}(u)$ , ils définissent le schéma de Runge-Kutta d'ordre  $m$  de la façon récursive suivante :

$$\forall 1 \leq i \leq m-1, \begin{cases} u^{(i)} = u^{(i-1)} + \Delta t \mathbf{L}u^{(i-1)} \\ u^{(m)} = \sum_{k=0}^{m-2} \alpha_{m,k} u^{(k)} + \alpha_{m,m-1} (u^{(m-1)} + \Delta t \mathbf{L}u^{(m-1)}), \end{cases}$$

avec pour  $k = 1, \dots, m-1$  les  $\alpha_{m,k}$  donnés par le tableau 3.1.

TAB. 3.1: Coefficients du schéma Runge-Kutta optimal selon son ordre.

Ordre m	$\alpha_{m,0}$	$\alpha_{m,1}$	$\alpha_{m,2}$	$\alpha_{m,3}$	$\alpha_{m,4}$	$\alpha_{m,5}$	$\alpha_{m,6}$	$\alpha_{m,7}$
1	1							
2	$\frac{1}{2}$	$\frac{1}{2}$						
3	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$					
4	$\frac{3}{8}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{24}$				
5	$\frac{11}{30}$	$\frac{3}{8}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{120}$			
6	$\frac{53}{144}$	$\frac{11}{30}$	$\frac{3}{16}$	$\frac{1}{189}$	$\frac{1}{48}$	$\frac{1}{720}$		
7	$\frac{103}{280}$	$\frac{53}{144}$	$\frac{11}{60}$	$\frac{3}{48}$	$\frac{1}{72}$	$\frac{1}{240}$	$\frac{1}{5040}$	
8	$\frac{2119}{5760}$	$\frac{103}{280}$	$\frac{53}{288}$	$\frac{11}{180}$	$\frac{1}{64}$	$\frac{1}{360}$	$\frac{1}{1440}$	$\frac{1}{40320}$

On remarque que la formulation d'ordre  $k$  se confond avec le développement de l'exponentielle de la matrice  $\Delta t \mathbf{M}^{-1} \mathbf{K}$ , et que, de fait, l'ordre  $q$  avec  $q \leq k$  s'obtient par troncature à l'ordre  $q$  de la formule à l'ordre  $k$ . Les coefficients sont donnés par le tableau 3.2.

TAB. 3.2: Développement tronqué de l'exponentielle fonction de l'ordre du RK.

---

1	$I - \mathbf{A}$
2	$I - \mathbf{A} + \frac{\mathbf{A}^2}{2}$
3	$I - \mathbf{A} + \frac{\mathbf{A}^2}{2} - \frac{\mathbf{A}^3}{6}$
4	$I - \mathbf{A} + \frac{\mathbf{A}^2}{2} - \frac{\mathbf{A}^3}{6} + \frac{\mathbf{A}^4}{24}$
5	$I - \mathbf{A} + \frac{\mathbf{A}^2}{2} - \frac{\mathbf{A}^3}{6} + \frac{\mathbf{A}^4}{24} - \frac{\mathbf{A}^5}{120}$
6	$I - \mathbf{A} + \frac{\mathbf{A}^2}{2} - \frac{\mathbf{A}^3}{6} + \frac{\mathbf{A}^4}{24} - \frac{\mathbf{A}^5}{120} + \frac{\mathbf{A}^6}{720}$
7	$I - \mathbf{A} + \frac{\mathbf{A}^2}{2} - \frac{\mathbf{A}^3}{6} + \frac{\mathbf{A}^4}{24} - \frac{\mathbf{A}^5}{120} + \frac{\mathbf{A}^6}{720} - \frac{\mathbf{A}^7}{5040}$
8	$I - \mathbf{A} + \frac{\mathbf{A}^2}{2} - \frac{\mathbf{A}^3}{6} + \frac{\mathbf{A}^4}{24} - \frac{\mathbf{A}^5}{120} + \frac{\mathbf{A}^6}{720} - \frac{\mathbf{A}^7}{5040} + \frac{\mathbf{A}^8}{40320}$
9	$I - \mathbf{A} + \frac{\mathbf{A}^2}{2} - \frac{\mathbf{A}^3}{6} + \frac{\mathbf{A}^4}{24} - \frac{\mathbf{A}^5}{120} + \frac{\mathbf{A}^6}{720} - \frac{\mathbf{A}^7}{5040} + \frac{\mathbf{A}^8}{40320} - \frac{\mathbf{A}^9}{362880}$
10	$I - \mathbf{A} + \frac{\mathbf{A}^2}{2} - \frac{\mathbf{A}^3}{6} + \frac{\mathbf{A}^4}{24} - \frac{\mathbf{A}^5}{120} + \frac{\mathbf{A}^6}{720} - \frac{\mathbf{A}^7}{5040} + \frac{\mathbf{A}^8}{40320} - \frac{\mathbf{A}^9}{362880} + \frac{\mathbf{A}^{10}}{3628800}$

---

**Théorème 3.7 (Théorème spectral)**

Soit  $f : \mathcal{M}_n(\mathbb{C}) \rightarrow \mathcal{M}_n(\mathbb{C})$  une fonction continue. Soit  $A \in \mathcal{M}_n(\mathbb{C})$ . Notons  $\sigma(A)$  le spectre de la matrice  $A$ . Alors :

$$f(\sigma(A)) = \sigma(f(A)).$$

Etant donné que le développement tronqué de l'exponentielle est un polynôme, il est continu et assure ainsi que, connaissant les valeurs propres de la matrice  $\mathbf{M}^{-1}\mathbf{K}$ , on obtienne facilement celle de la matrice de formulation globale. Soit  $\lambda \in \text{sp}(\Delta t \mathbf{M}^{-1}\mathbf{K})$  telle que  $\rho(\Delta t \mathbf{M}^{-1}\mathbf{K}) = |\lambda|$ , où  $\rho(\Delta t \mathbf{M}^{-1}\mathbf{K})$  représente le rayon spectral de la matrice  $(\Delta t \mathbf{M}^{-1}\mathbf{K})$ . Si  $\lambda \in \mathbb{R}$ , alors le rayon spectral de la matrice de la formulation globale est la valeur absolue de  $g(\lambda)$ , où  $g$  est le développement tronqué de l'exponentielle. Dans le cas complexe, c'est son module.

Soit  $\mathbf{A} = \Delta t \mathbf{M}^{-1} \mathbf{K}$  et  $\mathbf{I}$  la matrice identité. On peut désormais calculer les différentes conditions CFL pour les algorithmes de Runge-Kutta connaissant les matrices de masse et de rigidité associées au schéma de Galerkin Discontinu considéré. Soit  $\rho(\mathbf{A})$  le rayon spectral de la matrice  $\mathbf{A}$  et  $\lambda \in \mathbf{sp}(\mathbf{A})$  la valeur propre associée. La condition CFL est donnée par l'intermédiaire de :

$$|g(\lambda)| \leq 1.$$

Nous allons tracer les courbes  $|g(\lambda)|$  pour des ordres de Runge-Kutta  $r$  variant jusqu'à l'ordre 13, et ce, pour des schémas de type Galerkin Discontinu d'ordre 1 à 7. En prenant l'intersection de ces courbes avec l'axe  $y = 1$ , nous aurons la condition CFL maximale pour chaque couple (Runge-Kutta, Galerkin Discontinu). Ces courbes sont présentées de (3.1) à (3.7) et les résultats regroupés dans le tableau (3.3). Les conditions CFL obtenues sont vraiment très optimistes. En effet, si on prend comme méthode de calcul des conditions CDL la norme  $\mathcal{L}^2$  des polynômes ce ne peut qu'être pire. De plus, pour certaines équations (cf. sous section 3.1.3) la méthode de Galerkin Discontinu d'ordre 1 n'a pas de condition de stabilité.

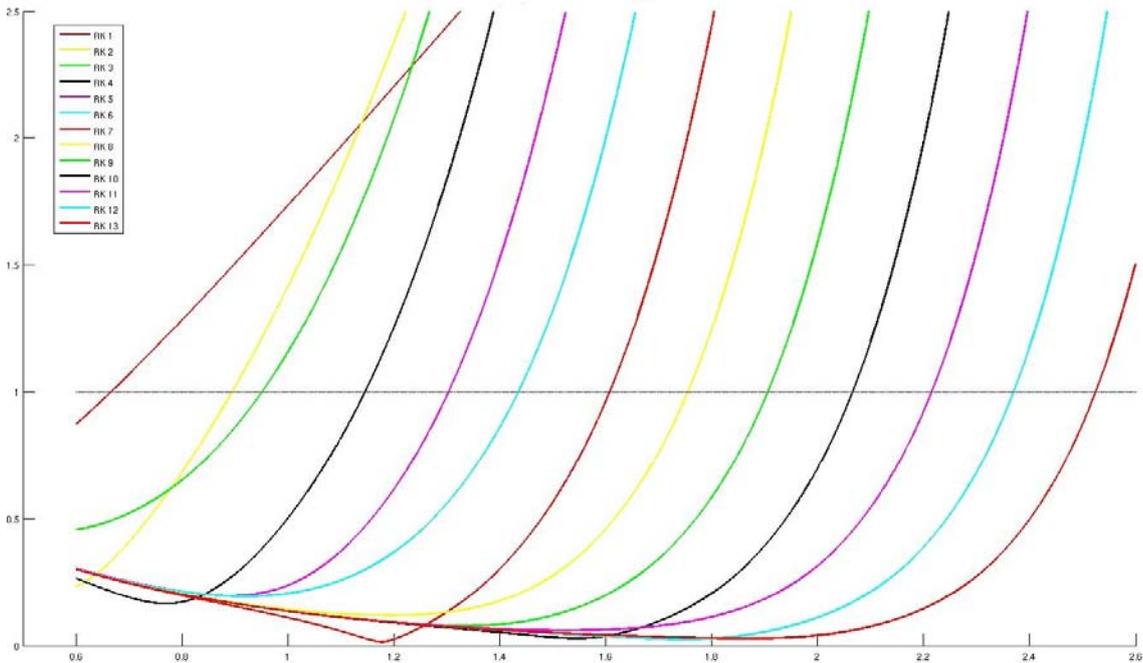


FIG. 3.1: Galerkin Discontinu d'ordre 1

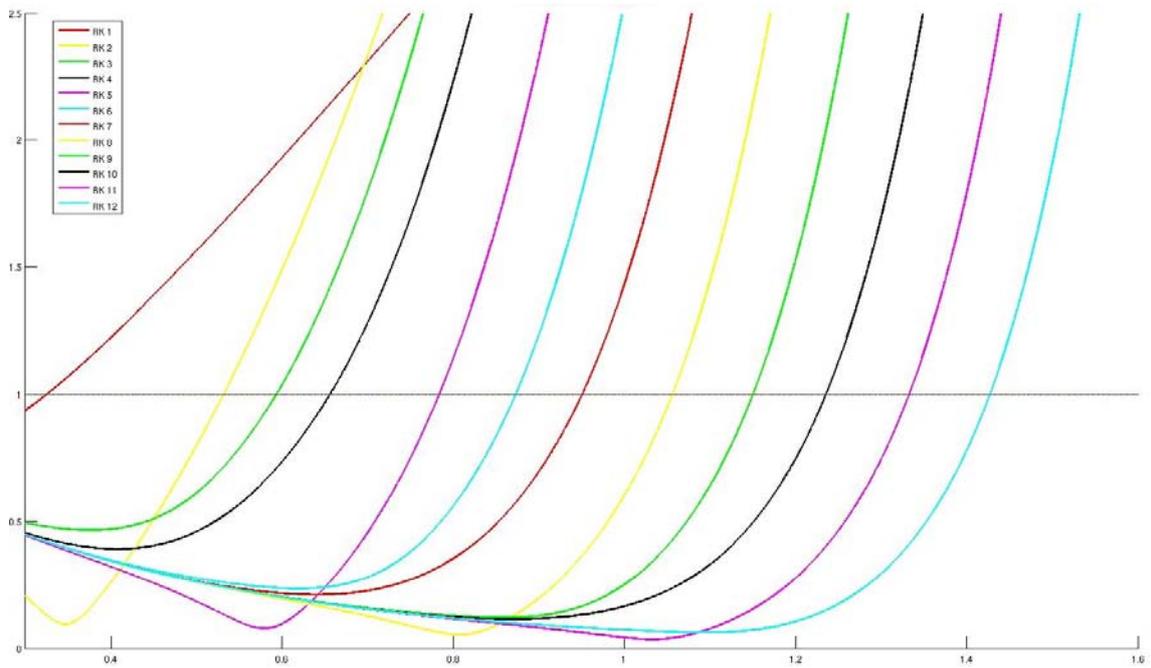


FIG. 3.2: Galerkin Discontinu d'ordre 2

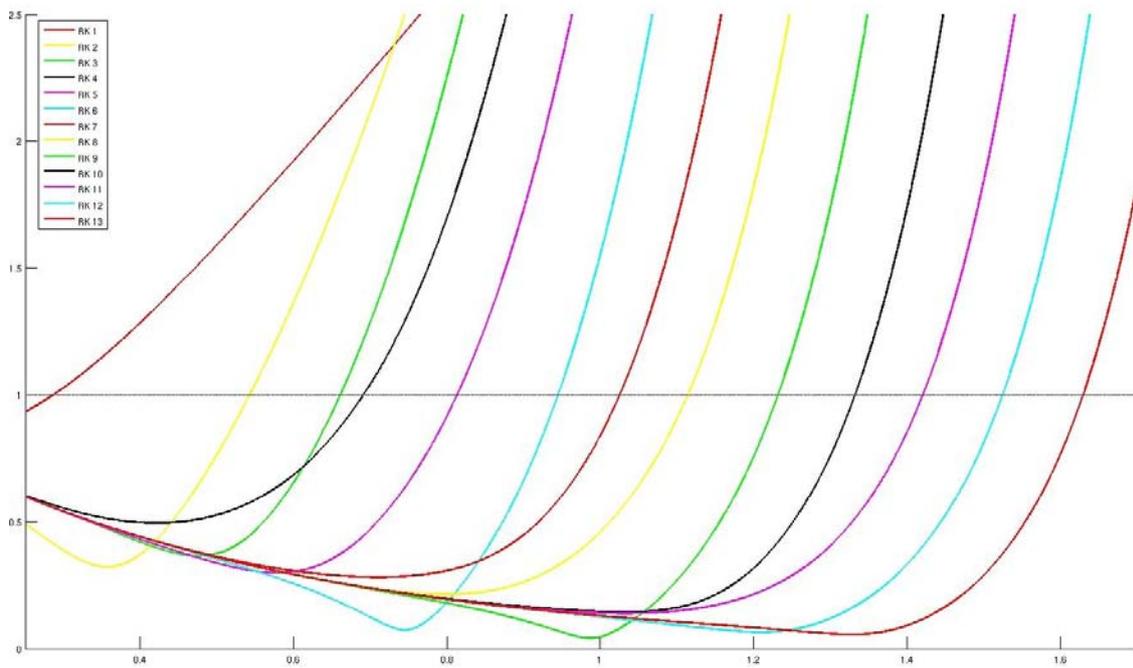


FIG. 3.3: Galerkin Discontinu d'ordre 3

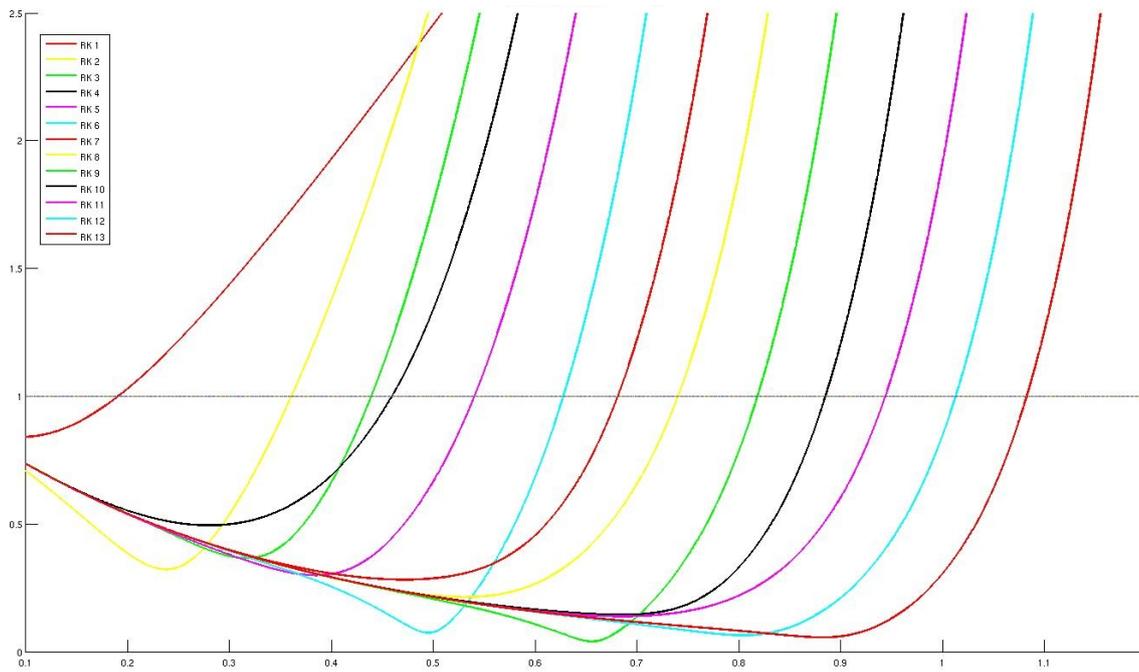


FIG. 3.4: Galerkin Discontinuu d'ordre 4

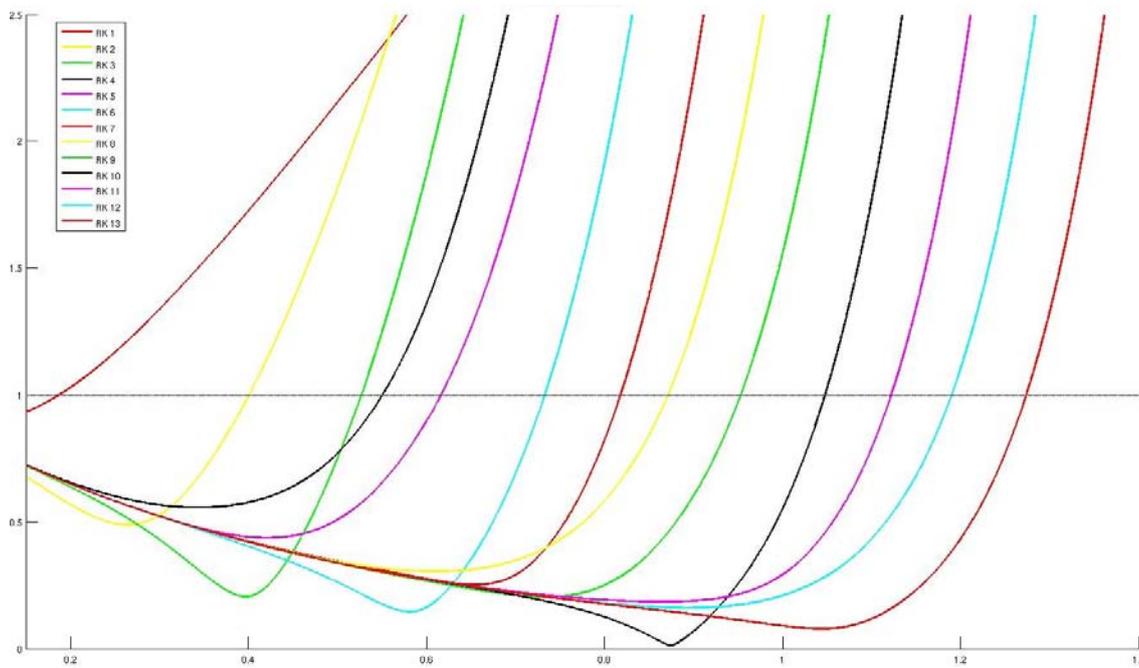


FIG. 3.5: Galerkin Discontinuu d'ordre 5

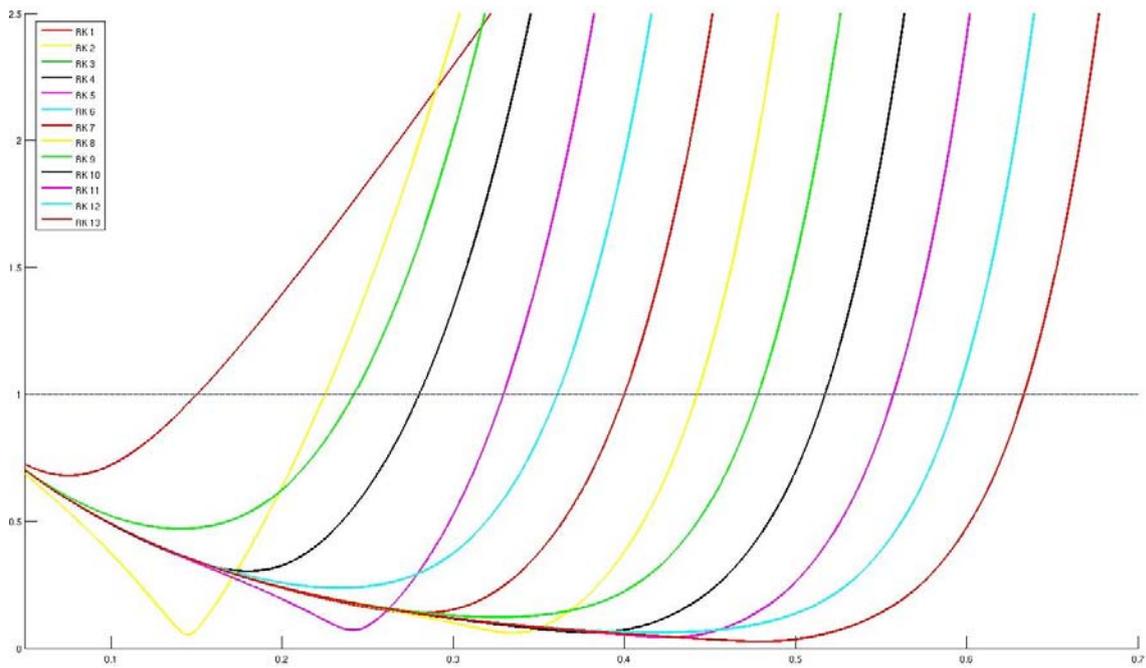


FIG. 3.6: Galerkin Discontinu d'ordre 6

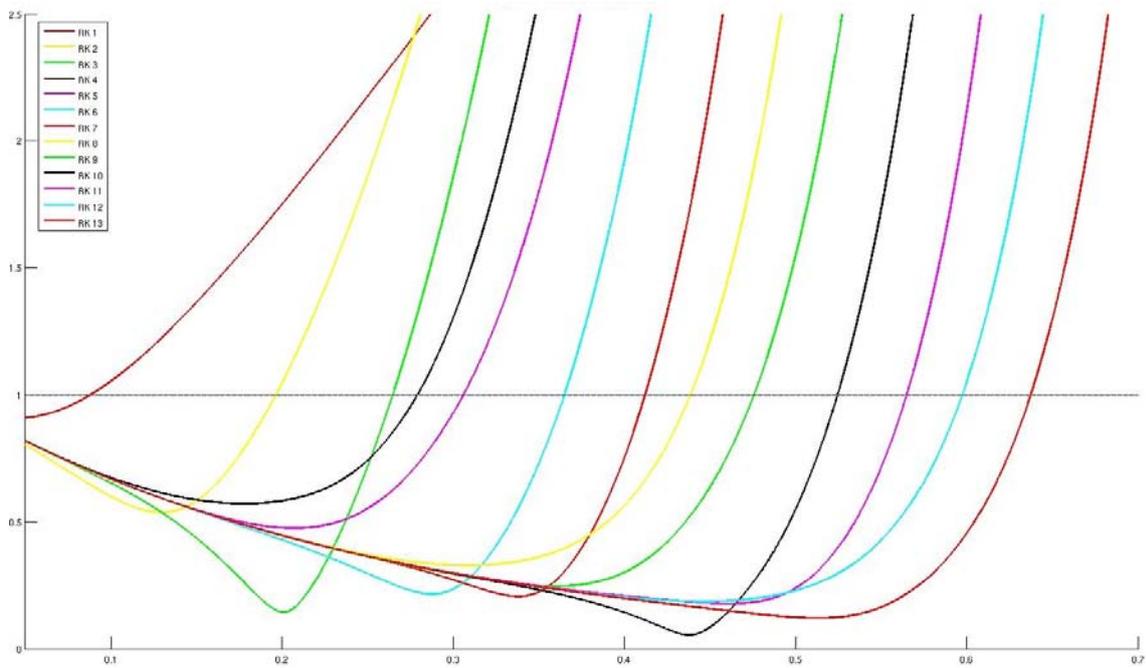


FIG. 3.7: Galerkin Discontinu d'ordre 7

TAB. 3.3: Conditions CFL selon l'ordre du RK (1-13) et du GD(1-7).

$RK \backslash GD$	1	2	3	4	5	6	7
1	0.667	0.325	0.289	0.193	0.19	0.154	0.091
2	0.893	0.532	0.546	0.363	0.404	0.227	0.198
3	0.967	0.596	0.662	0.442	0.529	0.244	0.266
4	1.147	0.656	0.696	0.463	0.554	0.282	0.282
5	1.307	0.784	0.816	0.544	0.62	0.331	0.309
6	1.44	0.876	0.948	0.63	0.736	0.363	0.367
7	1.613	0.954	1.028	0.684	0.82	0.402	0.414
8	1.76	1.061	1.117	0.742	0.874	0.444	0.441
9	1.907	1.153	1.235	0.821	0.954	0.479	0.477
10	2.067	1.241	1.336	0.888	1.05	0.512	0.527
11	2.219	1.335	1.425	0.946	1.125	0.559	0.567
12	2.376	1.43	1.529	1.015	1.192	0.596	0.598
13	2.532	1.25	1.634	1.087	1.278	0.635	0.64

**Théorème 3.8**

Soit  $k$  un ordre d'approximation spatiale. Alors il existe un ordre  $r$  pour une méthode de type Runge-Kutta tel que la condition CFL soit aussi grande que l'on veut, c'est-à-dire, tel qu'il existe  $C \in \mathbb{R}^+ \setminus \{0\}$  telle que  $CFL \geq C$ .

**Démonstration :** Soit  $a$  la valeur propre complexe telle que sa partie réelle est strictement négative. Soit  $x \in \mathbb{R}^+$ . Soit  $g_{n-1}$  le développement de l'exponentielle tronqué à l'ordre  $n - 1$ . Alors :

$$e^{ax} = g_{n-1}(ax) + R(x),$$

avec  $R(x) \leq \frac{|a|^n x^n}{n!}$  d'après la formule de Taylor Mac Laurin. Posons  $x = Kn$ , avec  $K$  une constante strictement positive. Alors :

$$|g_{n-1}(ax)| \leq |e^{ax}| + |R(x)| \Rightarrow |g_{n-1}(aKn)| \leq |e^{aKn}| + \frac{|a|^n K^n n^n}{n!}.$$

En utilisant la formule de Stirling,  $n! \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n}\right)$ , et à partir d'un certain rang, on a :

$$|g_{n-1}(aKn)| \leq |e^{aKn}| + \frac{|a|^n K^n e^n}{\sqrt{2\pi n}}.$$

Or  $|e^{aKn}| \xrightarrow{n \rightarrow \infty} 0$  et  $\frac{|a|^n K^n e^n}{\sqrt{2\pi n}} \xrightarrow{n \rightarrow \infty} 0$  si et seulement si  $|a|^n K^n e^n \leq 1$ , donc

$$K \leq \frac{1}{|a|e} \Rightarrow CFL \geq \frac{n+1}{|a|e}.$$

⊠

Il convient désormais de se demander si on obtient la convergence d'une approximation explicite stable comme celle-ci pour le système différentiel. Pour une méthode de type Volumes Finis, on démontre que l'ordre d'approximation du système différentiel est conservé dans l'explicitation. Pour un schéma implicite ou le système différentiel, on obtient une convergence en  $h^\alpha$ . L'explicitation du système différentiel stable dans la même norme que la norme de convergence peut faire perdre sur l'exposant  $\alpha$  de la convergence. La norme dans laquelle on contrôle la stabilité dépend de  $h$ . Il existe toujours une norme dans laquelle la stabilité s'exprime en termes de CFL hyperbolique. Cependant, la stabilité d'un schéma explicite au sens du rayon spectral de la matrice ne suffit pas à assurer la convergence du schéma explicite vers la solution du problème en norme  $\mathcal{L}^2$  quand  $h$  tend vers 0, et encore moins à exprimer l'ordre de convergence en  $h^\alpha$ .

### 3.1.2 Etude de la complexité d'un schéma de type Runge-Kutta couplé à une méthode de type Galerkin Discontinu

Soit  $r$  l'ordre du Runge-Kutta et  $k$  celui du Galerkin Discontinu. Nous travaillons sur des équations de types équations des Ondes écrites sous la forme suivante :

$$\mathbf{A}_0 \frac{\partial \varphi(\mathbf{x}, t)}{\partial t} + \sum_{i=1}^{2 \text{ ou } 3} \mathbf{A}_i \frac{\partial \varphi(\mathbf{x}, t)}{\partial \mathbf{x}_i} = f(\mathbf{x}, t). \quad (3.5)$$

Etudier la complexité d'un tel schéma sur une équation aussi générale implique de faire quelques hypothèses. Nous prendrons comme condition CFL celle donnée par le calcul monodimensionnel précédent(3.3) ramenée en bidimensionnel et tridimensionnel par rapport aux conditions CFL des Volumes Finis. Cette hypothèse est d'ailleurs particulièrement optimiste (cf. limites de la méthode de Galerkin Discontinu plus loin). Elle est correcte pour des maillages structurés mais trop favorable pour des maillages déstructurés, où augmenter la dimension de l'espace équivaut à diminuer la condition CFL. On suppose que les inversions des matrices de masses sont stockées. Les calculs seront menés sous deux cas différents : le premier, où il faut inverser les matrices de masses, le second, où elle seront supposées condensées (cette hypothèse favorisera les grands ordres). Enfin, la matrice  $\mathbf{A}_0$  sera supposée symétrique et ne dépendant pas de paramètres spatiaux (cette dernière hypothèse

permet notamment d'assurer qu'il existe une base, celle des polynômes de Legendre, qui convienne pour obtenir des matrices de masses condensées).

Calculons le coût d'un schéma Runge-Kutta optimal d'ordre  $r$  par élément. Il équivaut du point de vue de la complexité à  $r$  résolutions d'un schéma d'Euler Explicite où la première résolution peut être d'une complexité différente.

Soit  $n$  le nombre d'inconnues scalaires par élément et  $p$  le nombre d'équation du système différentiel. Soit l'inversion d'une matrice diagonale la référence d'une complexité nulle. Si la matrice de masse est condensée, le coût d'un schéma d'Euler est  $p$  fois celui d'une multiplication matrice-vecteur :  $C_1 = p \times (2r - 1) n^2$ . En revanche, si il faut inverser la matrice de masse, nous aurons un produit matrice de masse-matrice de rigidité, qui induit une complexité en  $n^3$ . D'où un coût en  $C_2 = p \times (n^3 + 2(r - 1) n^2)$ . Nous allons maintenant mener les études de complexité à partir de ces deux formules.

Rappelons la formule reliant le nombre d'inconnues scalaires par élément à l'ordre du Galerkin Discontinu et la dimension de l'espace, notée  $j$ .

$$n = \frac{(k + 1)(k + 2) \dots (k + j)}{j!}. \quad (3.6)$$

Pour calculer l'erreur de projection, on étudie le cas où la solution oscille de façon isotrope dans toutes les directions en se plaçant sur un segment/carré/cube de une longueur d'onde. Soit  $N$  le nombre d'éléments par longueur d'onde. Soit une solution de la forme  $\sin(2\pi x) \times \sin(2\pi y) \times \sin(2\pi z)$ . On fait en cela une hypothèse défavorable quant aux « isotropiques » oscillations de la solution : en effet, cette solution est appropriée pour des calculs dans une cavité, mais cette hypothèse devient très pénalisante dans le cas d'un problème extérieur ou même guidé.

[1] utilise l'estimation a priori qui relie l'erreur de consistance à l'erreur de projection, et estime l'erreur de projection en majorant le plus finement possible le reste de la fonction de Taylor avec reste intégral. En suivant cette méthode, on obtient la formulation :

$$\varepsilon_{proj} = \alpha \times \frac{2\pi^{k+1}}{(k + 1)!} \times \frac{1}{N^{k+\frac{1}{2}}}, \quad (3.7)$$

où la constante  $\alpha$  ne dépend ni de l'ordre du Galerkin Discontinu, ni du temps. Cette inconnue peut être levée en travaillant à erreur donnée. En effet, on désire simplement une comparaison entre les méthodes de Galerkin Discontinu et celles des Volumes Finis. Il suffit alors de travailler à précision égale sur les deux modèles. Soit  $\delta$  la précision souhaitée. En Volumes Finis, le nombre d'éléments par longueur d'onde est l'inverse de la précision. L'erreur de projection des Volumes Finis est donnée par la relation précédente en prenant  $k = 0$  d'où :

$$\varepsilon_{proj VF} = \alpha \times 2\pi\sqrt{\delta}. \quad (3.8)$$

On normalise la complexité des Volumes Finis à 1 et on étudie ainsi le gain de complexité. Si la valeur de la condition CFL augmente, la complexité diminue. Dans [29], il est montré que les ordres de Runge-Kutta optimaux sont les ordres 2 et 3, aussi allons faire les calculs pour ces ordres. Cela ne signifie cependant pas qu'il ne faille pas utiliser des méthodes de type Runge-Kutta d'ordre supérieur. Par exemple, si la solution est un polynôme de degré 4, un seul élément en espace est nécessaire si on utilise une approximation de type Runge-Kutta d'ordre 4 en temps.

Calculons le gain d'inconnues par élément et par longueur d'onde. En cherchant à obtenir la même erreur pour les méthodes de type Volumes Finis et Galerkin Discontinu, donc grâce à la formule (3.7), nous pouvons relier le nombre d'inconnues pour les deux méthodes :

$$N = \left( \frac{2\pi^k}{(k+1)! \sqrt{\delta}} \right)^{\frac{1}{k + \frac{1}{2}}} \quad (3.9)$$

En tridimensionnel, le nombre d'éléments par longueur d'onde pour une méthode de Galerkin Discontinu est en fait le nombre d'éléments par cube de côté de une longueur d'onde. Pour une méthode de type Volumes Finis, le nombre d'élément est l'inverse de la précision  $\delta$ . Soit  $G_i$  le gain d'inconnues entre les deux méthodes.

$$G_i = N^3 \times \delta^3 = \left( \left( \frac{2\pi^k}{(k+1)!} \right)^{\frac{1}{k + \frac{1}{2}}} \times \delta^{-2k} \right)^3. \quad (3.10)$$

TAB. 3.4: Gain d'inconnues par élément et par longueur d'ondes fonction de l'ordre  $k$  du GD et de la précision.

$\delta$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
$\frac{1}{10}$	$9.87e^{-2}$	$3.82e^{-2}$	$1.99e^{-2}$	$1.19e^{-2}$	$7.8e^{-3}$	$5.4e^{-3}$	$3.9e^{-3}$
$\frac{1}{50}$	$3.95e^{-3}$	$8.0e^{-4}$	$3.16e^{-4}$	$1.62e^{-4}$	$9.6e^{-5}$	$6.2e^{-5}$	$4.3e^{-5}$
$\frac{1}{100}$	$9.87e^{-4}$	$1.5e^{-4}$	$5.33e^{-5}$	$2.56e^{-5}$	$1.46e^{-5}$	$9.2e^{-6}$	$6.2e^{-6}$
$\frac{1}{500}$	$3.95e^{-5}$	$3.19e^{-6}$	$8.5e^{-7}$	$3.51e^{-7}$	$1.81e^{-7}$	$1.07e^{-7}$	$6.8e^{-8}$
$\frac{1}{1000}$	$9.87e^{-6}$	$6.05e^{-7}$	$1.43e^{-7}$	$5.53e^{-8}$	$2.73e^{-8}$	$1.56e^{-8}$	$9.8e^{-9}$

Pour étudier le gain complexité, nous considérerons deux possibilités : soit la matrice est condensée, soit l'inversion est réalisée en preprocessing. Soit  $G_{C_1}$  le gain

de complexité associé à la première hypothèse, et  $G_{C_2}$  associé à la seconde. Comme le gain de complexité est le produit du gain d'inconnues, de la valeur de la condition CFL et du coût du schéma Runge-Kutta (donné par  $C_1$  et  $C_2$ ), on obtient :

$$\begin{cases} G_{C_1} &= G_i (2r - 1) n^2 CFL(k, r) \\ G_{C_2} &= G_i (n^3 + 2(r - 1) n^2) CFL(k, r). \end{cases} \quad (3.11)$$

Nous allons faire varier l'ordre  $k$  de la méthode type Galerkin Discontinu entre 1 et 7. La précision  $\delta$  sera prise variant régulière de 0.1 à 0.001. Les gains d'inconnues et de complexité seront regroupés dans les tableaux (3.4) à (3.6). Après étude de ces tableaux, on constate qu'une approximation d'ordre 2 donne de bons résultats même à faible précision, bien que ce cas favorise très nettement l'ordre 1. L'ordre 3 peut devenir intéressant à haute précision. Les autres ordres des méthodes de type Galerkin Discontinu donnent des gains de complexité et d'inconnues du même ordre mais sachant que ces schémas impliquent des matrices de taille nettement supérieure et une complexité de mise en oeuvre plus grande, il faudrait, pour justifier leur emploi, gagner au moins un ordre de grandeur. Il n'apparaît donc pas utile, au moins dans les problèmes confinés, de monter l'ordre des méthodes de Galerkin Discontinu (ni celui des méthodes de Runge-Kutta). On se contentera donc de travailler avec des ordres 2 ou 3. On prêtera une attention particulière au choix des conditions CFL afin de s'assurer de la convergence et de la stabilité de ces méthodes. Ces résultats montrent également que la recherche d'une autre façon de monter en ordre moins pénalisante en terme de complexité et de condition CFL semble devoir être étudiée.

TAB. 3.5: Gain de complexité  $G_{c_1}$  fonction de l'ordre  $k$  du GD et de la précision.

$\delta$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
$\frac{1}{10}$	4.23	6.09	13.01	15.88	29.57	25.94	33.48
$\frac{1}{50}$	0.169	0.128	0.207	0.217	0.367	0.301	0.37
$\frac{1}{100}$	$4.23e^{-2}$	$2.42e^{-2}$	$3.49e^{-2}$	$3.42e^{-2}$	$5.54e^{-2}$	$4.41e^{-2}$	$5.31e^{-2}$
$\frac{1}{500}$	$1.69e^{-3}$	$5.1e^{-4}$	$5.5e^{-4}$	$4.7e^{-4}$	$6.9e^{-4}$	$5.1e^{-4}$	$5.9e^{-4}$
$\frac{1}{1000}$	$4.23e^{-4}$	$9.7e^{-5}$	$9.4e^{-5}$	$7.4e^{-5}$	$1.04e^{-4}$	$7.5e^{-5}$	$8.4e^{-5}$

TAB. 3.6: Gain de complexité  $G_{c_2}$  fonction de l'ordre  $k$  du GD et de la précision.

$\delta$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
$\frac{1}{10}$	8.46	24.37	95	196	572	744	1362
$\frac{1}{50}$	$3.38e^{-1}$	$5.12e^{-1}$	1.52	2.68	7.1	8.6	15.03
$\frac{1}{100}$	$8.5e^{-2}$	$9.7e^{-2}$	$2.56e^{-1}$	$4.22e^{-1}$	1.07	1.27	2.16
$\frac{1}{500}$	$4.0e^{-3}$	$2.0e^{-2}$	$4.0e^{-3}$	$6e^{-3}$	$1.3e^{-2}$	$1.5e^{-2}$	$2.4e^{-2}$
$\frac{1}{1000}$	$8.46e^{-4}$	$3.86e^{-4}$	$96.87e^{-4}$	$9.1e^{-4}$	$2.0e^{-3}$	$2.15e^{-3}$	$3.42e^{-3}$

### 3.1.3 Autres limites des méthodes de Galerkin Discontinu

Soit le cas favorable de la convergence du problème semi-discrétisé en  $h^{k+\frac{1}{2}}$  en  $C^1_{([0,T],L^2)}$  avec un second membre assez régulier, où  $h$  représente le pas d'espace. Soit  $u_h$  la solution du système différentiel semi-discrétisé. Alors :

$$\| u - u_h \|_{L^2} \leq C \times h^{k+\frac{1}{2}},$$

où la constante  $C$  dépend de la forme du domaine, du second membre ainsi que des conditions aux limites, prises en compte au travers de la norme  $H^{k+1}$ . Autrement dit, cette constante dépend de la solution  $\| u \|_{H^{k+1}}$ .

On appelle  $u_{h,r}$  la solution du système discrétisé approximée par une méthode de Runge-Kutta d'ordre  $r$ , polynômiale par morceau. Si cette solution est stable, on a une convergence de  $\| u_{h,r} - u_h \|_h$  vers 0. Dans [14], pour un système linéaire, la méthode de type Runge Kutta optimale (au sens de la formule de Taylor en temps) est aussi celle qui conserve la stabilité forte  $L^2$  du schéma d'Euler. Il sera important de s'assurer que l'approximation ne perde pas en qualité par rapport à l'approximation explicite du degré polynômial, et qu'il n'y ait pas de décroissance en  $h$ . Les seuls travaux existants sont ceux de [7] pour des systèmes de Friedrichs avec des conditions aux limites visqueuses sous une approximation par une méthode de type Volumes Finis et un schéma temporel Euler Explicite.

Les deux normes  $\| \cdot \|_h$  et  $\| \cdot \|_{L^2}$  ne restant pas équivalentes quand le pas d'espace tend vers zéro, on ne peut pas, par ce moyen, montrer la convergence de  $u_{h,r}$  vers  $u$ . Dans [7], les auteurs montrent que pour un schéma de type Euler - Galerkin Discontinu d'ordre 1, il est impossible de converger dans une norme équivalente à la norme  $\| \cdot \|_{L^2}$  (exemple de l'équation d'advection monodimensionnelle à laquelle on ajoute des Conditions aux Limites de type Von Neumann). Pendant longtemps, la

condition CFL de  $\frac{1}{9}$  a été avancée mais des simulations numériques ont été réalisées prouvant que pour environ 500000 inconnues il y avait perte de la stabilité (il faut compter environ 100 inconnues si l'on choisit la condition CFL de 0.5)

### 3.2 Les « Volumes Finis aux dérivées »

Considérons des Equations aux Dérivées Partielles de la forme :

$$\forall(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}^+, \quad \mathbf{A}_0 \frac{\partial}{\partial t} u(\mathbf{x}, t) + \sum_{i=1}^n \mathbf{A}_i \frac{\partial}{\partial \mathbf{x}_i} u(\mathbf{x}, t) + \mathbf{B}u(\mathbf{x}, t) = f(\mathbf{x}, t), \quad (3.12)$$

avec  $u(0, \mathbf{x}) = u_0(\mathbf{x})$  et où  $\mathbf{A}_i \in C^{k+2+\varepsilon}$  et  $\mathbf{B}, f, u_0 \in C^{k+1+\varepsilon}$  sont des matrices symétriques. Ces hypothèses garantissent la régularité  $H_{mca}^{k+1}$  pour ce problème instationnaire [32]. Pour approximer ce type d'Equation avec un degré élevé, il est nécessaire d'avoir certaines régularités. Pour les méthodes de type Galerkin Discontinu avec un flux splitting dissipatif on a  $\|u - u_h\| \leq c(\|u\|_{H^{k+1}}) h^{k+\frac{1}{2}}$ . Comme vu dans la section précédente, on a de plus une chute des conditions CFL en fonction de l'ordre de l'approximation et même, dans certain cas, absence de stabilité pour des schémas d'Euler.

L'idée de la nouvelle approximation d'ordre élevé que nous proposons repose sur le principe que, si  $u$  est solution de (3.12), chacune de ses dérivées partielles l'est aussi. Comme  $u$  est une fonction assez régulière, les dérivées partielles le sont aussi et permettent cette écriture,

$$\begin{aligned} \forall j \in \llbracket 1, n \rrbracket, \forall(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}^+, \quad \mathbf{A}_0 \frac{\partial}{\partial t} \left( \frac{\partial}{\partial \mathbf{x}_j} u(\mathbf{x}, t) \right) + \mathbf{A}_i \frac{\partial}{\partial \mathbf{x}_i} \left( \frac{\partial}{\partial \mathbf{x}_j} u(\mathbf{x}, t) \right) \\ + \left( \frac{\partial}{\partial \mathbf{x}_j} \mathbf{A}_i \right) \frac{\partial}{\partial \mathbf{x}_i} u(\mathbf{x}, t) + \mathbf{B} \frac{\partial}{\partial \mathbf{x}_j} u(\mathbf{x}, t) + \frac{\partial}{\partial \mathbf{x}_j} \mathbf{B}u(\mathbf{x}, t) = \frac{\partial}{\partial \mathbf{x}_j} f(\mathbf{x}, t). \end{aligned}$$

Posons  $u_\alpha = D_\alpha u$  avec  $\alpha = (\alpha_1, \dots, \alpha_n)$  et  $|\alpha| \leq k$ , où  $k$  représente l'ordre de la méthode souhaité. En procédant par récurrence, on obtient alors les équations suivantes, notées  $(E_\alpha)$ , avec  $\beta = (\beta_1, \dots, \beta_n)$ ,

$$\forall(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}^+, \quad \mathbf{A}_0 \frac{\partial}{\partial t} u_\alpha(\mathbf{x}, t) + \sum_{i=1}^n \mathbf{A}_i \frac{\partial}{\partial \mathbf{x}_i} u_\alpha(\mathbf{x}, t) + \sum_{|\beta| \leq k} \mathbf{C}_\beta u_\beta(\mathbf{x}, t) = D_\alpha f(\mathbf{x}, t), \quad (3.13)$$

où les matrices  $\mathbf{C}_\beta$  sont au moins  $C^{k+1+\varepsilon}$  et dépendent des dérivées partielles de  $\mathbf{A}_i$  et  $\mathbf{B}$ . Ces équations seront approximées par une méthode type Volumes Finis (où les fonctions tests sont les fonctions caractéristiques du maillage). Ceci garantit une condition CFL constante et une certaine facilité d'implémentation du schéma.

**Remarque :** on aurait les mêmes résultats avec des fonctions régulières par morceaux sous les hypothèses  $\mathbf{B}$  régulière par morceaux et  $\mathbf{A}_0$  constante par morceaux

(la partie principale de l'opérateur est la même pour tout  $\alpha$ ). Aussi, nous considérons que la solution  $u$  est  $H_{mex}^{k+1}$ , à savoir que les discontinuités sont cylindriques en temps. Notons  $\Sigma_{u_\alpha}$  les surfaces de discontinuité de  $u_\alpha$ . Ces surfaces sont en inclusion avec les surfaces de discontinuité de  $u$ .

Outre les gains de condition CFL et de complexité que ce type d'approximation promet, d'autres points viennent conforter cet aspect prometteur. Pour avoir une régularité correcte avec des méthodes de type Galerkin Discontinu, il faut choisir les matrices  $(\mathbf{A}_i, \mathbf{B})$  dans  $C^\alpha \times C^{\alpha-1}$ , avec  $\alpha > k + 2$  afin d'assurer la consistance du schéma dans  $H_{tan}^{s k+1}$ . Les hypothèses permettent donc de dériver partiellement sans rajouter des conditions de régularité supplémentaires. Les dérivées partielles de type  $D_\alpha u$  sont solutions de systèmes de même partie principale que la solution et donc les fronts d'ondes sur les dérivées partielles se propagent de la même façon que ceux de la solution originelle  $\forall \alpha, |\alpha| < k$ . La forme d'écriture polynômiale pourrait a priori gêner mais on aura in fine autant d'équations que d'inconnues donc la résolution est assurée.

### **Théorème 3.9**

Soit un maillage régulier  $\Omega = \bigcup_e \omega_e$  tel que pour tout  $\alpha$  tel que  $|\alpha| \leq k$  on ait  $\Sigma_{u_\alpha} \subset \bigcup_e \delta\omega_e$ . Alors toute solution  $u_\alpha$  solution (au sens des distributions) de  $(E_\alpha)$  vérifie pour tout  $\alpha$  tel que  $|\alpha| \leq k$  et pour tout élément du maillage  $\omega_e$ ,

$$\int_{\omega_e} \mathbf{A}_0 \partial_t u_\alpha + \int_{\omega_e} \left( \mathbf{A}^i \partial_i u_\alpha + \sum_{|\beta| \leq k} \mathbf{C}_\beta u_\beta \right) + \int_{\partial\omega_e} \mathbf{A}^i \mathbf{n}_i^- (u_\alpha^+ - u_\alpha^-) = \int_{\omega_e} D_\alpha f$$

avec les conditions de saut sur  $u_\alpha$  décrites par  $\mathbf{A}_i \mathbf{n}_i^- [u_\alpha] = 0$ .

Soit  $u_h = \sum_e u_h^e$  et  $u_h^e = \sum_{|\alpha| \leq k} U_\alpha^e (x - x_g)^\alpha$ . Pour tout  $\omega_e$  et  $\alpha$  tel que  $|\alpha| \leq k$ ,

$$\int_{\omega_e} \mathbf{A}_0 \partial_t D_\alpha u_h^e + \int_{\omega_e} \left( \mathbf{A}^i \partial_i D_\alpha u_h^e + \sum_{|\beta| \leq k} \mathbf{C}_\beta^\alpha D_\beta u_h^e \right) + \int_{\partial\omega_e} \mathbf{A}^i \mathbf{n}_i^- [D_\alpha u_h^e] = \int_{\omega_e} D_\alpha f.$$

Ceci conduit donc naturellement à un système d'EDO linéaire en  $U = \{u_h^e\}$  que l'on peut écrire sous la forme

$$\mathbf{M} \frac{dU}{dt} + \mathbf{K}U = F,$$

avec  $N_e \times N_i \times \frac{(n+1) \dots (n+k)}{k!}$  inconnues et équations, où  $N_e$  représente le nombre d'éléments et  $N_i$  le nombre d'Equations. Ce système étant sous forme de système de Friedrichs, l'existence et l'unicité de la solution sont assurées et son implémentation est très simple (passant par le calcul de  $\mathbf{M}^{-1} \mathbf{K}$  explicite pour tout élément  $\omega_e$ ).

**Démonstration** : soit une Equation aux Dérivées Partielles de la forme (3.12). Si les coefficients sont suffisamment réguliers, on dérive les équations. Si la solution

est  $H^{s,k+1}$  par morceaux, on dérivera seulement par morceaux (ce sera par exemple le cas lorsque la matrice  $\mathbf{A}_0$  présentera des discontinuités). Soit un maillage  $\Omega = \cup \omega_e$  déstructuré composé d'éléments non dégénérés. Soit  $u$  solution de régularité  $H_{mex}^{k+1}$  telle que les discontinuités éventuelles sont portées par la frontière du domaine  $\Omega$ , à savoir pour tout  $n$  uplet  $\alpha$  tel que  $|\alpha| \leq k$  et pour tout élément du maillage, on a  $\Sigma_{u_\alpha} \subset \cup_e \delta\omega_e$ . Comme le problème  $\partial_i f^i = 0$ ,  $f^i \in H_{mex}^{k+1}$  est équivalent au sens des distribution à  $\partial_i f^i = 0$  p.p., où  $f^i$  défini en tant que fonction  $H_{mex}^{k+1}$  plus la condition de nullité du saut de  $f^i$  sur les lignes de discontinuité de  $f$ , on a équivalence (théorème de localisation) au sens des distributions entre la formulation variationnelle

$$\int_{\omega_e} \mathbf{A}_0 \partial_t u_\alpha + \int_{\omega_e} \left( \mathbf{A}^i \partial_i u_\alpha + \sum_{|\beta| \leq k} \mathbf{C}_\beta u_\beta \right) + \int_{\partial\omega_e} \mathbf{A}^i \mathbf{n}_i^- (u_\alpha^+ - u_\alpha^-) = \int_{\omega_e} D_\alpha f$$

et

$$\begin{cases} \mathbf{A}_0 \partial_t u_\alpha + \left( \mathbf{A}^i \partial_i u_\alpha + \sum_{|\beta| \leq k} \mathbf{C}_\beta u_\beta \right) + \mathbf{A}^i \mathbf{n}_i^- (u_\alpha^+ - u_\alpha^-) = D_\alpha f & p.p. \\ \text{sauts nuls sur la frontière de la partie régulière, à savoir } \mathbf{A}_i \mathbf{n}_i^- [u_\alpha] = 0 \end{cases}$$

Cela est encore vrai pour les conditions inhomogènes d'obstacle parfait si le champ incident est assez régulier. Dans le cas d'un champ diffracté, tous les  $\mathbf{A}_i \mathbf{n}_i^- u$  sont des champs diffractés donc la condition  $\mathbf{A}_i \mathbf{n}_i^- [u_\alpha] = 0$  est bien une Absorbing Boundary Condition pour  $u$  et toutes les dérivées partielles  $u_\alpha$ . Donc en présence de conditions aux limites de type obstacle parfait sur les champs incidents, les conditions aux limites seront les mêmes sur toutes les équations dérivées en dérivant le champ incident si ce dernier est suffisamment régulier.

Donc sur les éléments où la solution est régulière les intégrales sont nulles. Sinon, les discontinuités sont portées par le cylindre temporel, donc pour un tel élément  $\omega_e$ , la condition du saut appartient au noyau de  $\mathbf{A}_i \mathbf{n}_i^-$  et de l'autre côté de sa frontière au noyau de  $\mathbf{A}_i (-\mathbf{n}_i)^- = \mathbf{A}_i \mathbf{n}_i^+$ . Mais  $\ker(\mathbf{A}_i \mathbf{n}_i) = \ker(\mathbf{A}_i \mathbf{n}_i^-) \cap \ker(\mathbf{A}_i \mathbf{n}_i^+)$  donc les sauts doivent appartenir au noyau de  $\mathbf{A}_i \mathbf{n}_i$  (a fortiori celui de  $\mathbf{A}_i \mathbf{n}_i^-$ ).  $\square$

Pour fixer les idées, et réaliser des démonstrations de consistance et de stabilité, nous allons nous intéresser à certains types d'Equations aux Dérivées Partielles en particulier : l'Equation d'Advection monodimensionnelle et les systèmes de type Friedrichs à coefficients constants en bidimensionnel.

### 3.2.1 Equation d'Advection monodimensionnelle

Replaçons nous dans le cas de la section (3.1). Soit  $\Omega \subset \mathbb{R}$  et  $T \in \mathbb{R}_+^*$ . Le problème considéré est l'équation d'advection monodimensionnelle (3.1) avec  $a = 1$  et  $f \equiv 0$

pour simplifier les calculs :  $\frac{\partial u(x, t)}{\partial t} + \frac{\partial u(x, t)}{\partial x} = 0$ . Soit le maillage  $\Omega = \sum_{k=1}^{N_h} \omega_k$ , avec  $\omega_k = [\alpha_k, \beta_k]$  segments de longueur  $h \in \mathbb{R}^+ \setminus \{0\}$ .

**Lemme 3.10**

*L'approximation Volumes Finis aux Dérivées à l'ordre 1 est solution de*

$$\begin{cases} h \frac{du_k(t)}{dt} + (u_k(t) - u_{k-1}(t)) - \frac{h}{2} (u_{xk}(t) - u_{xk-1}(t)) = 0, \\ u_k(t) = u_h(x, t), & x \in \omega_k \\ u_{xk}(t) = (u_x)_h(x, t), & x \in \omega_k. \end{cases}$$

**Démonstration** : soit  $k \in \llbracket 1, N \rrbracket$ . Considérons l'élément d'indice  $k$  de ce maillage. La solution approchée  $u_h(x, t)$  à l'ordre 1 est recherchée sur cet élément sous la forme  $u_k(t) = u_k^{(0)}(t) + u_k^{(1)}(t)x$ . Les fonctions tests sont prises constantes (égales à 1) par morceaux. La formulation variationnelle (3.2) appliquée à cette équation est

$$\begin{aligned} & \int_{-h/2}^{h/2} \left( \frac{d(u_k^{(0)}(t) + u_k^{(1)}(t)x)}{dt} + u_k^{(1)}(t) \right) dx + \left( u_k^{(0)}(t) - \frac{h}{2} u_k^{(1)}(t) - (u_{k-1}^{(0)}(t) + \frac{h}{2} u_{k-1}^{(1)}(t)) \right) \\ &= h \frac{du_k^{(0)}(t)}{dt} + \left( u_k^{(0)}(t) - u_{k-1}^{(0)}(t) \right) - \frac{h}{2} \left( u_k^{(1)}(t) - u_{k-1}^{(1)}(t) \right) = 0. \end{aligned}$$

L'idée de cette méthode est de dériver l'équation initiale autant de fois que l'ordre d'approximation souhaitée. Donc ici il convient de rajouter l'équation sur  $u_x(x, t) = \frac{\partial u(x, t)}{\partial x}$  à savoir  $\frac{\partial u_x(x, t)}{\partial t} + \frac{\partial u_x(x, t)}{\partial x} = 0$ . Sur l'élément  $\omega_k$ ,  $u_x$  est alors approximée par  $u_{xk}(t) = u_k^{(1)}(t)$ , autrement dit par une méthode équivalente à une méthode de type Volumes Finis. La formulation variationnelle de l'équation pour l'élément  $\omega_k$  est donc

$$\int_{-h/2}^{h/2} \frac{du_k^{(1)}(t)}{dt} dx + \left( u_k^{(1)}(t) - u_{k-1}^{(1)}(t) \right) = h \frac{du_k^{(1)}(t)}{dt} + \left( u_k^{(1)}(t) - u_{k-1}^{(1)}(t) \right) = 0.$$

En multipliant cette dernière équation par  $x$  et en la sommant avec la relation obtenue sur  $\frac{du_k^{(0)}(t)}{dt}$ , on obtient le résultat proposé.  $\square$

**Lemme 3.11**

*La méthode Volumes Finis aux Dérivées à l'ordre 2 s'écrit, pour tout  $\omega_k$ ,*

$$h \frac{du_k(t)}{dt} + (u_k(t) - u_{k-1}(t)) - \frac{h}{2} (u_{xk}(t) - u_{xk-1}(t)) - \frac{h^2}{12} (u_{xxk}(t) - u_{xxk-1}(t)) = 0.$$

**Démonstration** : pour l'ordre 2, la solution sur l'élément  $k$  est de la forme  $u_k^{(0)}(t) + u_k^{(1)}(t)x + u_k^{(2)}(t)\frac{x^2}{2}$ . Il y aura donc 3 équations, portant sur  $u_k$ ,  $u_{xk}$  et  $u_{xxk}$ .

D'où le système à résoudre

$$\left\{ \begin{array}{l} \int_{-h/2}^{h/2} \left( \frac{d}{dt} \left( u_k^{(0)}(t) + u_k^{(1)}(t)x + u_k^{(2)}(t)\frac{x^2}{2} \right) + u_k^{(1)}(t) + u_k^{(2)}(t)x \right) dx \\ + \left( u_k^{(0)}(t) - u_k^{(1)}(t)\frac{h}{2} + u_k^{(2)}(t)\frac{h^2}{4} - \left( u_{k-1}^{(0)}(t) + u_{k-1}^{(1)}(t)\frac{h}{2} + u_{k-1}^{(2)}(t)\frac{h^2}{4} \right) \right) = 0 \\ \int_{-h/2}^{h/2} \left( \frac{d}{dt} \left( u_k^{(1)}(t) + u_k^{(2)}(t)x \right) + u_k^{(2)}(t) \right) dx \\ + \left( u_k^{(1)}(t) - \frac{h}{2}u_k^{(2)}(t) - \left( u_{k-1}^{(1)}(t) + \frac{h}{2}u_{k-1}^{(2)}(t) \right) \right) = 0 \\ \int_{-h/2}^{h/2} \frac{du_k^{(2)}(t)}{dt} dx + \left( u_k^{(2)}(t) - u_{k-1}^{(2)}(t) \right) = 0. \end{array} \right.$$

Simplifions le système précédent, nous obtenons,

$$\left\{ \begin{array}{l} h \frac{du_k^{(0)}(t)}{dt} + \frac{h^3}{24} \frac{du_k^{(2)}(t)}{dt} \\ + \left( u_k^{(0)}(t) - u_{k-1}^{(0)}(t) \right) - \frac{h}{2} \left( u_k^{(1)}(t) - u_{k-1}^{(1)}(t) \right) + \frac{h^2}{4} \left( u_k^{(2)}(t) - u_{k-1}^{(2)}(t) \right) = 0 \\ h \frac{du_k^{(1)}(t)}{dt} + \left( u_k^{(1)}(t) - u_{k-1}^{(1)}(t) \right) - \frac{h}{2} \left( u_k^{(2)}(t) - u_{k-1}^{(2)}(t) \right) = 0 \\ h \frac{du_k^{(2)}(t)}{dt} + \left( u_k^{(2)}(t) - u_{k-1}^{(2)}(t) \right) = 0. \end{array} \right. \quad (3.14)$$

L'intervalle d'intégration symétrique apporte des simplifications qui dépendent de la parité de la puissance de  $x$ . En rajoutant une puissance de  $x$  paire, on obtient une contribution de la dérivée temporelle du coefficient  $u_k^{(2)}(t)$  qui n'apparaissait pas pour l'approximation d'ordre 1. Aussi, il faut, quand on souhaite écrire le schéma d'un ordre pair à un ordre impair, substituer le terme issu de la dérivée temporelle par son approximation Volumes Finis. Ici, on remplace donc  $\frac{du_k^{(2)}(t)}{dt}$  par  $-\frac{u_k^{(2)}(t) - u_{k-1}^{(2)}(t)}{h}$  dans la première équation de (3.14), d'où

$$h \frac{du_k^{(0)}(t)}{dt} + \left( u_k^{(0)}(t) - u_{k-1}^{(0)}(t) \right) - \frac{h}{2} \left( u_k^{(1)}(t) - u_{k-1}^{(1)}(t) \right) - \frac{h^2}{12} \left( u_k^{(2)}(t) - u_{k-1}^{(2)}(t) \right) = 0.$$

La preuve se termine en ajoutant à cette équation  $x$  fois la deuxième équation de (3.14) et  $\frac{x^2}{2}$  fois la troisième équation de (3.14).  $\square$

**Remarque :** la dérivée d'ordre le plus élevée est approximée par une méthode de type Volumes Finis et celle d'ordre inférieur par une méthode de Volumes Finis aux Dérivées d'ordre 1. Cette méthode est donc récursive, ce qui permettra une facilité de montée en ordre dans les codes numériques. Une fois que la méthode d'ordre  $n$  est implémentée, il n'y a qu'une équation à rajouter pour implémenter l'ordre  $n+1$ .

Généralisons maintenant ces résultats pour une approximation d'ordre  $n$ . Sur l'élément  $\omega_k$ , la solution approchée  $u_h(x, t)$  est recherchée sous la forme polynomiale

$$u_k(t) = \sum_{i=0}^n u_k^{(i)}(t) \frac{x^i}{i!}.$$

L'approximation est constituée de  $n + 1$  équations obtenues par dérivations successives de l'équation de départ

$$\forall (x, t) \in \mathbb{R} \times [0, T], \forall i \in \llbracket 0, n \rrbracket, \quad \frac{\partial}{\partial t} \frac{\partial^i}{\partial x^i} u(x, t) + \frac{\partial^{i+1}}{\partial x^{i+1}} u(x, t) = 0.$$

**Lemme 3.12**

L'approximation Volumes Finis aux Dérivées d'ordre  $n$  est solution, pour tout élément  $\omega_k$  de

$$\frac{du_k(t)}{dt} + \sum_{i=0}^n \alpha_i h^i (u_k^{(i)}(t) - u_{k-1}^{(i)}(t)) = 0,$$

où les coefficients  $\alpha_i$  sont construits par récurrence :  $\alpha_0 = 1$ ,  $\alpha_1 = 1/2$  et  $\forall i \geq 2$ ,

$$\begin{cases} \alpha_i = 0 & \text{si } i \text{ est impair,} \\ \alpha_i = \frac{1}{2^i i!} + \sum_{j \text{ pair}}^{i-1} \frac{-\alpha_j}{2^{i-j} (i-j+1)!} & \text{sinon.} \end{cases}$$

**Démonstration** : les dérivées partielles de la solution approchée sur le segment  $\omega_k$  s'écrivent

$$\forall j \in \llbracket 0, n \rrbracket, \quad u_k^{(j)}(t) = \sum_{i=0}^{n-j} u_k^{(i+j)}(t) \frac{x^i}{i!}.$$

Le système à résoudre est donc composé de la formulation variationnelle de l'équation et des équations dérivées, d'où

$$\int_{-h/2}^{h/2} \left( \frac{du_k^{(j)}(t)}{dt} + u_k^{(j+1)}(t) \right) dx + (u_k^{(j)}(t) - u_{k-1}^{(j)}(t)) = 0, \quad j \in \llbracket 0, n \rrbracket.$$

Ce qui donne

$$\int_{-h/2}^{h/2} \left( \sum_{i=0}^{n-j} \frac{du_k^{(j+i)}(t)}{dt} \frac{x^i}{i!} + \sum_{i=1}^{n-j} u_k^{(j+i)}(t) \frac{x^{i-1}}{(i-1)!} \right) dx + \sum_{i=0}^{n-j} \left[ u_k^{(j+i)}(t) \frac{(-h)^i}{2^i i!} - u_{k-1}^{(j+i)}(t) \frac{h^i}{2^i i!} \right] = 0.$$

Comme  $[-h/2, h/2]$  est un intervalle symétrique, les intégrales de fonctions impaires sont nulles, donc

$$\begin{aligned} & \sum_{i \text{ pair}}^{n-j} \frac{du_k^{(j+i)}(t)}{dt} \frac{h^{i+1}}{2^i (i+1)!} + \sum_{i \text{ pair}}^{n-j} (u_k^{(j+i)}(t) - u_{k-1}^{(j+i)}(t)) \frac{h^i}{2^i i!} \\ & + \sum_{i \text{ impair}}^{n-j} u_k^{(j+i)}(t) \frac{h^i}{2^{i-1} i!} - \sum_{i \text{ impair}}^{n-j} (u_k^{(j+i)}(t) + u_{k-1}^{(j+i)}(t)) \frac{h^i}{2^i i!} = 0. \end{aligned}$$

Le système d'équations se simplifie ensuite sous la forme, pour tout  $j \in \llbracket 0, n \rrbracket$ ,

$$\sum_{i=0}^{E[(n-j)/2]} \frac{du_k^{(j+2i)}(t)}{dt} \frac{h^{2i+1}}{2^{2i}(2i+1)!} + \sum_{i=0}^{n-j} (u_k^{(j+i)}(t) - u_{k-1}^{(j+i)}(t)) \frac{h^i}{2^{2i}i!} = 0. \quad (3.15)$$

Le système obtenu est triangulaire supérieur. Multiplier la  $j^e$  équation par  $\frac{x^j}{j!}$  et sommer les  $n+1$  équations termine la preuve.  $\square$

**Théorème 3.13 (Consistance)**

*L'approximation d'ordre  $n$  pour l'équation d'advection monodimensionnelle est consistante d'ordre  $n+2$  si  $n$  pair,  $n+1$  sinon.*

**Démonstration :** utilisons le résultat du lemme (3.12). Appliquons la formule de Taylor sur les termes de la forme  $(u_k^{(i)}(t) - u_{k-1}^{(i)}(t))$ . Soit  $p$  l'ordre du reste choisi,

$$(u_k^{(i)}(t) - u_{k-1}^{(i)}(t)) = \sum_{j=0}^p u_k^{(j+i)}(t) \frac{(-1)^{j+1}}{j!} h^j + o(h^{p+i}).$$

Comme  $u$  est solution de l'équation d'advection (ce qui permet d'éliminer le terme de dérivation temporelle), l'erreur de consistance est

$$\varepsilon = \sum_{i=0}^p \beta_i h^i + o(h^{p+i}), \text{ où } \begin{cases} \beta_i = \sum_{j=0}^{\min(i-1, n)} \frac{(-1)^{i-j+1}}{(i-j)!} \alpha_j & \text{si } i < n+2, \\ \beta_i = 0 & \text{sinon.} \end{cases}$$

Le schéma à l'ordre  $n$  pair est donc consistant d'ordre  $n+2$ . La nullité des coefficients  $\alpha_i$  impairs assure que le schéma pour un  $n$  impair est le même que celui de l'entier pair précédent, la consistance reste donc la même pour l'ordre impair qui suit.  $\square$

**Théorème 3.14 (Stabilité)**

*La stabilité de l'approximation d'ordre  $n$  pour l'Equation d'Advection monodimensionnelle est la même que celle d'une méthode Volumes Finis.*

**Démonstration par récurrence :** soit la proposition  $\mathcal{P}(j)$  « la composante d'ordre  $n-j$  est stable ».

*Initialisation :* la composante  $u_k^{(n)}$  qui correspond à la dérivée la plus élevée d'ordre  $n$  est déterminée par des approximations de type Volumes Finis dont elle hérite la stabilité  $l^2$ . Soit la convention de notation suivante :  $u_k^{(n)N}$  désigne l'approximation (type Euler) au  $N^e$  pas de temps de la composante  $u_k^{(n)}$  pour l'élément  $\omega_k$ . Soit  $N_t$  le nombre final d'itérations temporelles et  $\Delta t$  le pas de temps. Alors il existe une constante  $c_n$  appartenant à  $]0, 1]^2$  telle que  $\forall N \in \llbracket 1, N_t \rrbracket$  on ait

$$\| u_k^{(n)N} \| \leq (c_n + \mathcal{O}(\Delta t)) \| u_k^{(n)0} \|.$$

Donc la propriété  $\mathcal{P}(0)$  est vraie.

*Réurrence* : supposons  $\mathcal{P}(j-1)$  vraie. Dans la démonstration du lemme (3.12), une fois que l'équation (3.15) est établie, il est possible d'écrire le résultat sous une autre forme que celle proposée :  $\frac{du_k^{(j)}(t)}{dt} + \sum_{i=0}^{n-j} \beta_i h^i (u_k^{(j+i)}(t) - u_{k-1}^{(j+i)}(t)) = 0$ , pour tout  $j \in \llbracket 0, n \rrbracket$ , et les coefficients  $\beta_i$  sont des constantes calculables par récurrence. Donc pour la  $j^e$  composante, l'équation s'écrit

$$\frac{du_k^{(j)}(t)}{dt} + (u_k^{(j)}(t) - u_{k-1}^{(j)}(t)) + \sum_{i=1}^{n-j} \beta_i h^i (u_k^{(j+i)}(t) - u_{k-1}^{(j+i)}(t)) = 0.$$

La première partie correspond à une approximation de type Volumes Finis. Notons  $v_k^{(j)}$  cette partie. Il existe donc  $c_j \in ]0, 1]$  telle que

$$\forall i \in \llbracket 1, N_t \rrbracket, \|v_k^{(j)N}\| \leq (c_j + \mathcal{O}(\Delta t)) \|v_k^{(j)0}\|.$$

L'autre partie de l'approximation, donnée par  $\sum_{i=1}^{n-j} \beta_i h^i (u_k^{(j+i)}(t) - u_{k-1}^{(j+i)}(t))$  est un terme de degré 0 en approximation spatiale qui ne dépendra que des dérivées d'ordre supérieur. Après l'approximation temporelle de type Euler, cette somme, notée  $w_k^{(j)}$  sera multiplié par  $\Delta t$  et

$$\|w_k^{(j)N}\| \leq \Delta t g(c_{j+1}, c_{j+2}, \dots, c_n, \Delta t, \|u_k^{(j+1)0}\|, \|u_k^{(j+2)0}\|, \dots, \|u_k^{(n)0}\|),$$

où  $g$  est une fonction linéaire de chacun de ces arguments. Comme  $\mathcal{P}(j-1)$  est vraie, on a donc stabilité de  $u_k^{(j+1)}$  à  $u_k^{(n)}$ , donc  $\|w_k^{(j)N}\| \leq \mathcal{O}(\Delta t) \|u_k^{(j)0}\|$ . D'où

$$\|u_k^{(j)N}\| \leq \|v_k^{(j)}\| + \|w_k^{(j)}\| \leq (c_j + \mathcal{O}(\Delta t)) \|u_k^{(j)0}\|.$$

On a donc bien la stabilité  $l^2$  de cette approximation. Soit la norme définie par la somme sur tous les éléments des normes  $H^k$  (condition  $(L^2)^m$  sur toutes les dérivées). Nous avons démontré la stabilité de l'approximation au sens de la norme  $l^2$  des valeurs des fonctions au centre des éléments. La solution est approchée sous la forme

$$u_k(t) = \sum_{j=0}^n \frac{x^j}{j!} u_k^{(j)}(t). \text{ Les fonctions } u_j \text{ qui correspondent aux dérivées partielles sont}$$

prises constantes par morceaux et dépendent de  $h$ . Si les éléments ont des volumes comparables (pour tout élément du maillage son volume est borné), ce qui est le cas pour un maillage structuré, la norme  $l^2$  est équivalente à la norme  $L^2$  des fonctions constantes par éléments. Au final, on a donc équivalence entre la norme  $l^2$  et  $L^2$  puisqu'on travaille avec un maillage structuré à pas constant. Cette approximation est donc  $L^2$  stable. De plus, [14] démontre que tout schéma de type Runge-Kutta est stable si l'approximation de type Euler l'est.  $\square$

**Remarque** : dans le cas 2D ou 3D, si les éléments sont assez ronds au sens où le rapport  $\text{vol}(\omega_e)/\text{vol}(\delta\omega_e)$  est borné, on a équivalence avec la norme  $L^2$  des dérivées de polynômes et la norme  $L^2$  des fonctions constantes par éléments. Dans toute la suite de ce chapitre, nous démontrerons donc uniquement la stabilité au sens de la norme  $l^2$  des valeurs des fonctions au centre des éléments.

Grâce aux résultats précédents de consistance (3.13) et de stabilité (3.14), on obtient le théorème suivant comme conséquence immédiate du théorème de Lax-Richtmyer-Morton [33].

**Théorème 3.15 (Convergence)**

*L'approximation d'ordre  $n$  étant consistante et stable, alors elle est convergente.*

### 3.2.2 Systèmes de Friedrichs à coefficients constants

$$\forall(\mathbf{x}, t) \in \mathbb{R}^2 \times \mathbb{R}^+, \quad \frac{\partial u(\mathbf{x}, t)}{\partial t} + \mathbf{A}_1 \frac{\partial u(\mathbf{x}, t)}{\partial x} + \mathbf{A}_2 \frac{\partial u(\mathbf{x}, t)}{\partial y} = 0, \quad (3.16)$$

avec  $(\mathbf{A}_1, \mathbf{A}_2)$  à coefficients constants. Dans le cas contraire, des termes de multiplication apparaissent, mais ne modifient pas les propriétés ni l'écriture de ce schéma. Dans un premier temps, nous considérons des maillages structurés. Cette hypothèse permet d'utiliser un développement de type Taylor pour démontrer la consistance des schémas obtenus.

Soit  $\omega_{k,l}$  l'élément de référence du maillage, indexé par un couple d'entiers  $(k, l)$ , rectangulaire et inclus dans  $[-\Delta x/2, \Delta x/2] \times [-\Delta y/2, \Delta y/2]$ , où  $(\Delta x, \Delta y)$  sont les pas de discrétisation en espace. Soient  $(\lambda_j^i)$  les valeurs propres des matrices  $\mathbf{A}_i$  alors  $\mathbf{A}_i = \mathbf{P} \text{diag}(\lambda_j^i) \mathbf{P}^{-1}$ . En posant  $\lambda_j^{i+} = \max(\lambda_j^i, 0)$  et  $\lambda_j^{i-} = \min(\lambda_j^i, 0)$ , une décomposition de  $\mathbf{A}_i$  est  $\mathbf{A}_i^- + \mathbf{A}_i^+$  avec  $\mathbf{A}_i^+ = \mathbf{P} \text{diag}(\lambda_j^{i+}) \mathbf{P}^{-1}$  et  $\mathbf{A}_i^- = \mathbf{P} \text{diag}(\lambda_j^{i-}) \mathbf{P}^{-1}$ .

**Propriété 3.16**

*L'approximation Volumes Finis aux Dérivées d'ordre 1 pour le problème (3.16) est un polynôme vectoriel de degré 1 solution sur chaque  $\omega_{k,l}$  de*

$$\begin{aligned} & \frac{du_{k,l}(t)}{dt} + \mathbf{A}_2^- \frac{u_{k,l+1}(t) - u_{k,l}(t)}{\Delta y} - \mathbf{A}_2^+ \frac{u_{k,l-1}(t) - u_{k,l}(t)}{\Delta y} \\ & + \mathbf{A}_1^- \frac{u_{k+1,l}(t) - u_{k,l}(t)}{\Delta x} - \mathbf{A}_1^+ \frac{u_{k-1,l}(t) - u_{k,l}(t)}{\Delta x} \\ & - \mathbf{A}_2^- \frac{u_{y_{k,l+1}}(t) - u_{y_{k,l}}(t)}{2} - \mathbf{A}_2^+ \frac{u_{y_{k,l-1}}(t) - u_{y_{k,l}}(t)}{2} \\ & - \mathbf{A}_1^- \frac{u_{x_{k+1,l}}(t) - u_{x_{k,l}}(t)}{2} - \mathbf{A}_1^+ \frac{u_{x_{k-1,l}}(t) - u_{x_{k,l}}(t)}{2} \equiv 0. \end{aligned}$$

**Démonstration** : à l'ordre 1, la solution approchée  $u_a(x, y, t)$  est sous la forme  $u^{(0,0)}(t) + u^{(1,0)}(t)x + u^{(0,1)}(t)y$ . Les fonctions  $u^{(i,j)}$  sont les nouvelles inconnues. Il

y a 3 inconnues pour 3 équations. Les dérivées partielles  $\frac{\partial u_a(x, y, t)}{\partial x} = u^{(1,0)}(t)$  et  $\frac{\partial u_a(x, y, t)}{\partial y} = u^{(0,1)}(t)$  correspondent au centre de gravité de l'élément aux coefficients  $u_{k,l}^{(1,0)}$  et  $u_{k,l}^{(0,1)}$ . Les équations pour ces dérivées partielles (celles d'ordre le plus élevé) sont les mêmes que celles obtenues avec une méthode de Volumes Finis

$$\iint_{\omega_{k,l}} \frac{du_{k,l}^{(1,0)}(t)}{dt} + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- (u_{ext}^{(1,0)}(t) - u_{k,l}^{(1,0)}(t)) = 0 \quad (3.17)$$

$$\iint_{\omega_{k,l}} \frac{du_{k,l}^{(0,1)}(t)}{dt} + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- (u_{ext}^{(0,1)}(t) - u_{k,l}^{(0,1)}(t)) = 0. \quad (3.18)$$

La formulation variationnelle (3.2) appliquée à cette équation est

$$\begin{aligned} & \iint_{\omega_{k,l}} \frac{d}{dt} \left( u_{k,l}^{(0,0)}(t) + u_{k,l}^{(1,0)}(t)x + u_{k,l}^{(0,1)}(t)y \right) \\ & + \iint_{\omega_{k,l}} \left( \mathbf{A}_1 \frac{\partial}{\partial x} + \mathbf{A}_2 \frac{\partial}{\partial y} \right) \left( u_{k,l}^{(0,0)}(t) + u_{k,l}^{(1,0)}(t)x + u_{k,l}^{(0,1)}(t)y \right) \\ & + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- \left( u_{ext}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t) \right) + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- \left( u_{ext}^{(1,0)}(t)x_{ext}(s) - u_{k,l}^{(1,0)}(t)x(s) \right) ds \\ & + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- \left( u_{ext}^{(0,1)}(t)y_{ext}(s) - u_{k,l}^{(0,1)}(t)y(s) \right) ds = 0. \end{aligned}$$

Comme les intervalles d'intégration sont symétriques, des simplifications apparaissent.

$$\begin{aligned} & \Delta x \Delta y \left( \frac{du^{(0,0)}(t)}{dt} + \mathbf{A}_1 \frac{du^{(1,0)}(t)}{dt} + \mathbf{A}_2 \frac{du^{(0,1)}(t)}{dt} \right) + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- \left( u_{ext}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t) \right) \\ & + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- \left( u_{ext}^{(1,0)}(t)x_{ext}(s) - u_{k,l}^{(1,0)}(t)x(s) \right) ds \\ & + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- \left( u_{ext}^{(0,1)}(t)y_{ext}(s) - u_{k,l}^{(0,1)}(t)y(s) \right) ds = 0. \end{aligned} \quad (3.19)$$

Le terme  $FV_{k,l}(t) = \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- \left( u_{ext}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t) \right)$  apparaît également dans les formulations Volumes Finis et

$$\begin{aligned} FV_{k,l}(t) &= \mathbf{A}_2^- \frac{u_{k,l+1}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t)}{\Delta y} - \mathbf{A}_2^+ \frac{u_{k,l-1}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t)}{\Delta y} \\ &+ \mathbf{A}_1^- \frac{u_{k+1,l}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t)}{\Delta x} - \mathbf{A}_1^+ \frac{u_{k-1,l}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t)}{\Delta x}. \end{aligned} \quad (3.20)$$

A ce terme se rajoute la contribution des intégrales de bord, homogène au carré du pas d'espace.

$$\begin{aligned}
I_{k,l}(t) &= \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- (u_{ext}^{(1,0)}(t)x_{ext}(s) - u_{k,l}^{(1,0)}(t)x(s)) ds \\
&\quad + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- (u_{ext}^{(0,1)}(t)y_{ext}(s) - u_{k,l}^{(0,1)}(t)y(s)) ds \\
&= \mathbf{A}_2^- \left( u_{k,l+1}^{(0,1)}(t) - u_{k,l}^{(0,1)}(t) \right) + \mathbf{A}_2^+ \left( u_{k,l-1}^{(0,1)}(t) - u_{k,l}^{(0,1)}(t) \right) \\
&\quad + \mathbf{A}_1^- \left( u_{k+1,l}^{(1,0)}(t) - u_{k,l}^{(1,0)}(t) \right) + \mathbf{A}_1^+ \left( u_{k-1,l}^{(1,0)}(t) - u_{k,l}^{(1,0)}(t) \right). \quad (3.21)
\end{aligned}$$

En multipliant (3.17) par  $x$ , (3.18) par  $y$ , et en sommant les deux équations obtenues avec (3.19), on obtient la formule proposée.  $\square$

Les démonstrations de consistance et de stabilité requièrent les mêmes techniques que celles mises en jeu pour l'équation d'advection monodimensionnelle.

**Théorème 3.17 (Consistance)**

*L'approximation Volumes Finis aux Dérivées d'ordre 1 (propriété (3.16)) est consistante à l'ordre 2.*

**Démonstration** : on applique la formule de Taylor pour obtenir l'erreur de consistance. Par exemple,  $u_{k,l+1}^{(0,1)} - u_{k,l}^{(0,1)} = \Delta y u_{k,l}^{(0,2)} + \frac{\Delta y^2}{2} u_{k,l}^{(0,3)} + o(\Delta y^3)$ . Après simplifications,

$$\Delta t u_{k,l}^{(0,0)} + \mathbf{A}_1 u_{k,l}^{(1,0)} + \mathbf{A}_2 u_{k,l}^{(0,1)} - \frac{\Delta y^2}{12} \mathbf{A}_1 u_{k,l}^{(0,3)} - \frac{\Delta x^2}{12} \mathbf{A}_2 u_{k,l}^{(0,3)} + o(\Delta x^3) + o(\Delta y^3) = 0.$$

Comme  $u$  est solution de l'équation (3.16),  $\Delta t u_{k,l}^{(0,0)} + \mathbf{A}_1 u_{k,l}^{(1,0)} + \mathbf{A}_2 u_{k,l}^{(0,1)} = 0$ . Soit  $\varepsilon$  l'erreur de consistance, alors

$$\varepsilon = \frac{\Delta y^2}{12} \mathbf{A}_1 u_{k,l}^{(0,3)} + \frac{\Delta x^2}{12} \mathbf{A}_2 u_{k,l}^{(0,3)} + o(\Delta x^2) + o(\Delta y^2). \quad (3.22)$$

L'approximation proposée est donc consistante à l'ordre 2.  $\square$

**Théorème 3.18 (Stabilité)**

*La stabilité de l'approximation Volumes Finis aux Dérivées d'ordre 1 (propriété (3.16)) est la même que celle d'une méthode de type Volumes Finis.*

**Démonstration** : remarquons tout d'abord que les composantes  $u^{(0,1)}$  et  $u^{(1,0)}$  sont déterminées par des approximations Volumes Finis dont elles héritent la stabilité  $l^2$ . Soit  $u_{k,l}^{(i,j)N}$  l'approximation (type Euler) au  $N^e$  pas de temps de la composante  $u_{k,l}^{(i,j)}$  où  $(i,j) \in \{0,1\}$ ,  $N_t$  le nombre final d'itérations temporelles et  $\Delta t$  le pas de temps. Alors il existe deux constantes  $(c_{0,1}, c_{1,0})$  appartenant à  $]0,1]^2$  telles que  $\forall N \in \llbracket 1, N_t \rrbracket$  on ait

$$\begin{aligned}
\| u_{k,l}^{(0,1)N} \| &\leq (c_{0,1} + \mathcal{O}(\Delta t)) \| u_{k,l}^{(0,1)0} \|, \\
\| u_{k,l}^{(1,0)N} \| &\leq (c_{1,0} + \mathcal{O}(\Delta t)) \| u_{k,l}^{(1,0)0} \|.
\end{aligned}$$

Pour  $u_{k,l}^{(0,0)}$ , d'après les équations précédentes, une partie du schéma, notée  $FV_{k,l}$ , équation (3.20), correspond à un schéma de type Volumes Finis. Il existe  $c_{0,0} \in ]0, 1]$  telle que

$$\forall N \in \llbracket 1, N_t \rrbracket, \| FV_{k,l}^N \| \leq (c_{0,0} + \mathcal{O}(\Delta t)) \| u_{k,l}^{(0,0)0} \|.$$

L'autre partie de l'approximation  $I_{k,l}$ , donnée par (3.21), est un terme de degré 0 en approximation spatiale qui ne dépendra que des dérivées d'ordre supérieur, à savoir ici  $u^{(0,1)}$  et  $u^{(1,0)}$ . Après l'approximation temporelle de type Euler, ce terme est multiplié par  $\Delta t$  et

$$\| I_{k,l}^N \| \leq \Delta t g(c_{0,1}, c_{1,0}, \Delta t, \| u_{k,l}^{(1,0)0} \|, \| u_{k,l}^{(0,1)0} \|) \leq \mathcal{O}(\Delta t) \| u_{k,l}^{(0,0)0} \|,$$

où  $g$  est une fonction linéaire de chacun de ces arguments. Donc  $\forall n \in \llbracket 1, N_t \rrbracket$ ,

$$\| u_{k,l}^{(0,0)N} \| \leq \| FV_{k,l}^N \| + \Delta t \| I_{k,l}^N \| \leq (c_{0,0} + \mathcal{O}(\Delta t)) \| u_{k,l}^{(0,0)0} \|.$$

D'où la stabilité  $l^2$  de cette approximation. Comme il s'agit d'un maillage structuré,  $\Delta x$  et  $\Delta y$  sont constants sur le maillage, on a donc équivalence entre la norme  $l^2$  et  $L^2$ . Cette approximation est donc  $L^2$  stable.  $\square$

Comme pour l'équation d'advection monodimensionnelle, grâce aux résultats précédents de consistance et de stabilité, la convergence de la méthode est une conséquence immédiate du théorème de Lax-Richtmyer-Morton [33].

### **Théorème 3.19 (Convergence)**

*L'approximation Volumes Finis aux Dérivées d'ordre 1, donnée par la propriété (3.16), est convergente.*

Nous allons maintenant nous intéresser à l'ordre 2. Pour l'ordre 1, à la différence d'une méthode de type Galerkin Discontinu, nous n'utilisons pas de fonctions de base en produit croisé, à savoir  $xy$ . Ce terme n'apparaît qu'à partir de l'ordre 2, mais bien que cette fonction soit une fonction de base, la dérivée croisée  $\partial_x \partial_y$  n'intervient pas dans l'approximation du polynôme vectoriel.

### **Propriété 3.20**

*L'approximation Volumes Finis aux Dérivées à l'ordre 2 pour le problème (3.16) est le polynôme vectoriel de degré 2 solution pour chaque élément  $\omega_{k,l}$  de*

$$\begin{aligned} & \frac{du_{k,l}(t)}{dt} + \mathbf{A}_2^- \frac{1}{\Delta y} (u_{k,l+1}(t) - u_{k,l}(t)) - \mathbf{A}_2^+ \frac{1}{\Delta y} (u_{k,l-1}(t) - u_{k,l}(t)) \\ & + \mathbf{A}_1^- \frac{1}{\Delta x} (u_{k+1,l}(t) - u_{k,l}(t)) - \mathbf{A}_1^+ \frac{1}{\Delta x} (u_{k-1,l}(t) - u_{k,l}(t)) \\ & - \mathbf{A}_2^- \frac{1}{2} (u_{y_{k,l+1}}(t) - u_{y_{k,l}}(t)) - \mathbf{A}_2^+ \frac{1}{2} (u_{y_{k,l-1}}(t) - u_{y_{k,l}}(t)) \\ & - \mathbf{A}_1^- \frac{1}{2} (u_{x_{k+1,l}}(t) - u_{x_{k,l}}(t)) - \mathbf{A}_1^+ \frac{1}{2} (u_{x_{k-1,l}}(t) - u_{x_{k,l}}(t)) \\ & - \mathbf{A}_2^- \frac{\Delta x}{12} (u_{yy_{k,l+1}}(t) - u_{yy_{k,l}}(t)) - \mathbf{A}_2^+ \frac{\Delta x}{12} (u_{yy_{k,l-1}}(t) - u_{yy_{k,l}}(t)) \\ & - \mathbf{A}_1^- \frac{\Delta y}{12} (u_{xx_{k+1,l}}(t) - u_{xx_{k,l}}(t)) - \mathbf{A}_1^+ \frac{\Delta y}{12} (u_{xx_{k-1,l}}(t) - u_{xx_{k,l}}(t)) \equiv 0. \end{aligned}$$

**Démonstration** : soit la solution approchée à l'ordre 2 sur  $\omega_{k,l}$  l'élément

$$u_a(\mathbf{x}, t) = u_{k,l}^{(0,0)}(t) + u_{k,l}^{(1,0)}(t)x + u_{k,l}^{(0,1)}(t)y + u_{k,l}^{(1,1)}(t)xy + u_{k,l}^{(2,0)}(t)\frac{x^2}{2} + u_{k,l}^{(0,2)}(t)\frac{y^2}{2},$$

où les fonctions  $u_{k,l}^{(i,j)}$  correspondent, au centre de gravité, aux dérivées partielles d'ordre  $i$  en  $x$  et  $j$  en  $y$ . Cette approximation sera composée de 6 inconnues pour 6 équations. Les dérivées partielles d'ordre 2 sont

$$\frac{\partial^2 u_a(\mathbf{x}, t)}{\partial x^2} = u_{k,l}^{(2,0)}(t) \quad \frac{\partial^2 u_a(\mathbf{x}, t)}{\partial y^2} = u_{k,l}^{(0,2)}(t) \quad \frac{\partial^2 u_a(\mathbf{x}, t)}{\partial xy} = u_{k,l}^{(1,1)}(t)$$

Les équations sur ces fonctions sont les mêmes que celles obtenues avec une formulation Volumes Finis classiques, à savoir les 3 équations

$$\begin{aligned} \iint_{\omega_{k,l}} \frac{du_{k,l}^{(2,0)}(t)}{dt} + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- \left( u_{ext}^{(2,0)}(t) - u_{k,l}^{(2,0)}(t) \right) &= 0 \\ \iint_{\omega_{k,l}} \frac{du_{k,l}^{(0,2)}(t)}{dt} + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- \left( u_{ext}^{(0,2)}(t) - u_{k,l}^{(0,2)}(t) \right) &= 0 \\ \iint_{\omega_{k,l}} \frac{du_{k,l}^{(1,1)}(t)}{dt} + \int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- \left( u_{ext}^{(1,1)}(t) - u_{k,l}^{(1,1)}(t) \right) &= 0 \end{aligned} \quad (3.23)$$

Les dérivées partielles d'ordre 1 correspondent à une approximation de degré 1

$$\begin{cases} \frac{\partial u_a(\mathbf{x}, t)}{\partial x} = u_{k,l}^{(1,0)}(t) + u_{k,l}^{(2,0)}(t)x + u_{k,l}^{(1,1)}(t)y \\ \frac{\partial u_a(\mathbf{x}, t)}{\partial y} = u_{k,l}^{(0,1)}(t) + u_{k,l}^{(1,1)}(t)x + u_{k,l}^{(0,2)}(t)y. \end{cases}$$

Les deux équations qui leur sont associées sont celles du schéma d'ordre 1, données par la propriété précédente (3.16). Ces fonctions ne possèdent aucune fonction de la forme  $x^i y^j$  avec  $i$  et  $j$  tous les deux pairs, autrement dit,

$$\iint_{\omega_{k,l}} \mathbf{A}_i \frac{\partial u_a(\mathbf{x}, t)}{\partial \mathbf{x}_i} = 0.$$

Cette intégrale n'intervient donc pas dans l'équation d'ordre 0 sur  $u_a(\mathbf{x}, t)$  qui s'obtient comme précédemment. L'intégrale de surface sur la dérivée temporelle fait apparaître des simplifications par zéro si la fonction  $x^i y^j$  est impaire.

$$\iint_{\omega_{k,l}} \frac{du_a(\mathbf{x}, t)}{dt} := \Delta x \Delta y \frac{du_{k,l}^{(0,0)}(t)}{dt} + \frac{\Delta x^3 \Delta y}{24} \frac{du_{k,l}^{(2,0)}(t)}{dt} + \frac{\Delta x \Delta y^3}{24} \frac{du_{k,l}^{(0,2)}(t)}{dt}.$$

La présence des termes d'ordre 2 dans la forme polynômiale de la solution approchée fait apparaître des termes supplémentaires pour la dérivée temporelle. Il faut

expliciter  $\frac{du_{k,l}^{(2,0)}}{dt}$  et  $\frac{du_{k,l}^{(0,2)}}{dt}$  grâce à leurs approximations par la méthode Volumes Finis (équation (3.23)),

$$\begin{aligned}\Delta x \Delta y \frac{du_{k,l}^{(2,0)}(t)}{dt} &= -\Delta y \mathbf{A}_1^- \left( u_{k+1,l}^{(2,0)}(t) - u_{k,l}^{(2,0)}(t) \right) - \Delta y \mathbf{A}_1^+ \left( u_{k-1,l}^{(2,0)}(t) - u_{k,l}^{(2,0)}(t) \right) \\ &\quad - \Delta x \mathbf{A}_2^- \left( u_{k,l+1}^{(2,0)}(t) - u_{k,l}^{(2,0)}(t) \right) - \Delta x \mathbf{A}_2^+ \left( u_{k,l-1}^{(2,0)}(t) - u_{k,l}^{(2,0)}(t) \right) \\ \Delta x \Delta y \frac{du_{k,l}^{(0,2)}(t)}{dt} &= -\Delta y \mathbf{A}_1^- \left( u_{k+1,l}^{(0,2)}(t) - u_{k,l}^{(0,2)}(t) \right) - \Delta y \mathbf{A}_1^+ \left( u_{k-1,l}^{(0,2)}(t) - u_{k,l}^{(0,2)}(t) \right) \\ &\quad - \Delta x \mathbf{A}_2^- \left( u_{k,l+1}^{(0,2)}(t) - u_{k,l}^{(0,2)}(t) \right) - \Delta x \mathbf{A}_2^+ \left( u_{k,l-1}^{(0,2)}(t) - u_{k,l}^{(0,2)}(t) \right).\end{aligned}$$

Les intégrales de bords sont de la forme  $\int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- (u_{ext}(\mathbf{x}, t) - u_a(\mathbf{x}, t))$ , ie

$$\begin{aligned}\Delta y \mathbf{A}_1^- (u_{ak+1,l}(\mathbf{x}, t) - u_{ak,l}(\mathbf{x}, t)) + \Delta y \mathbf{A}_1^+ (u_{ak-1,l}(\mathbf{x}, t) - u_{ak,l}(\mathbf{x}, t)) \\ + \Delta x \mathbf{A}_2^- (u_{ak,l+1}(\mathbf{x}, t) - u_{ak,l}(\mathbf{x}, t)) + \Delta x \mathbf{A}_2^+ (u_{ak,l-1}(\mathbf{x}, t) - u_{ak,l}(\mathbf{x}, t)).\end{aligned}$$

Par exemple,

$$\begin{aligned}u_{ak+1,l}(\mathbf{x}, t) - u_{ak,l}(\mathbf{x}, t) &= u_{k+1,l}^{(0,0)}(t) - \frac{\Delta x}{2} u_{k+1,l}^{(1,0)}(t) + \frac{\Delta x^2}{8} u_{k+1,l}^{(2,0)}(t) + \frac{\Delta y^2}{24} u_{k+1,l}^{(0,2)}(t) \\ &\quad - u_{k,l}^{(0,0)}(t) - \frac{\Delta x}{2} u_{k,l}^{(1,0)}(t) - \frac{\Delta x^2}{8} u_{k,l}^{(2,0)}(t) - \frac{\Delta y^2}{24} u_{k,l}^{(0,2)}(t).\end{aligned}$$

Les termes  $\frac{\Delta y^2}{24} u_{k+1,l}^{(0,2)}(t)$  et  $-\frac{\Delta y^2}{24} u_{k,l}^{(0,2)}(t)$  se simplifient avec ceux issus de la formulation Volumes Finis de  $\frac{du_{k,l}^{(0,2)}(t)}{dt}$  et  $\frac{\Delta x^2}{8} u_{k+1,l}^{(2,0)}(t)$  et  $-\frac{\Delta x^2}{8} u_{k,l}^{(2,0)}(t)$  prennent alors le coefficient de  $1/8 - 1/24 = 1/12$ . Ce qui donne après simplification par l'aire  $\Delta x \Delta y$ ,

$$\begin{aligned}\frac{du_{k,l}^{(0,0)}}{dt} + \mathbf{A}_2^- \frac{1}{\Delta y} (u_{k,l+1}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t)) &\quad - \mathbf{A}_2^+ \frac{1}{\Delta y} (u_{k,l-1}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t)) \\ + \mathbf{A}_1^- \frac{1}{\Delta x} (u_{k+1,l}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t)) &\quad - \mathbf{A}_1^+ \frac{1}{\Delta x} (u_{k-1,l}^{(0,0)}(t) - u_{k,l}^{(0,0)}(t)) \\ - \mathbf{A}_2^- \frac{1}{2} (u_{k,l+1}^{(0,1)}(t) - u_{k,l}^{(0,1)}(t)) &\quad - \mathbf{A}_2^+ \frac{1}{2} (u_{k,l-1}^{(0,1)}(t) - u_{k,l}^{(0,1)}(t)) \\ - \mathbf{A}_1^- \frac{1}{2} (u_{k+1,l}^{(1,0)}(t) - u_{k,l}^{(1,0)}(t)) &\quad - \mathbf{A}_1^+ \frac{1}{2} (u_{k-1,l}^{(1,0)}(t) - u_{k,l}^{(1,0)}(t)) \\ - \mathbf{A}_2^- \frac{\Delta x}{12} (u_{k,l+1}^{(0,2)}(t) - u_{k,l}^{(0,2)}(t)) &\quad - \mathbf{A}_2^+ \frac{\Delta x}{12} (u_{k,l-1}^{(0,2)}(t) - u_{k,l}^{(0,2)}(t)) \\ - \mathbf{A}_1^- \frac{\Delta y}{12} (u_{k+1,l}^{(2,0)}(t) - u_{k,l}^{(2,0)}(t)) &\quad - \mathbf{A}_1^+ \frac{\Delta y}{12} (u_{k-1,l}^{(2,0)}(t) - u_{k,l}^{(2,0)}(t)) = 0.\end{aligned}$$

En multipliant, pour tout  $(i, j) \in \llbracket 0, 2 \rrbracket^2$  tels que  $i + j \leq 2$ , l'équation sur  $\frac{\partial^{i+j} u_a(\mathbf{x}, t)}{\partial x^i \partial y^j}$  par  $\frac{x^i y^j}{i! j!}$  et en les sommant, on recompose ainsi l'équation pour la

solution  $u_a$ . En effet, la fonction  $u_{k,l}^{(i,j)}$  s'identifie à  $\frac{\partial^{i+j} u_a(\mathbf{x}, t)}{\partial x^i \partial y^j}$  au centre de gravité de l'élément, d'où la formulation (3.20).  $\square$

**Théorème 3.21 (Consistance)**

L'approximation Volumes Finis à l'ordre  $n$ , établie dans la proposition (3.20), est consistante à l'ordre 4.

**Démonstration** : comme précédemment, cette démonstration s'appuie sur les formules de Taylor. Après applications successives, on obtient

$$\Delta t u_{k,l}^{(0,0)} + \mathbf{A}_1 u_{k,l}^{(1,0)} + \mathbf{A}_2 u_{k,l}^{(0,1)} + \frac{\Delta x^4}{720} \mathbf{A}_1 u_{k,l}^{(0,5)} + \frac{\Delta y^4}{720} \mathbf{A}_2 u_{k,l}^{(5,0)} + o(\Delta x^4) + o(\Delta y^4) = 0.$$

$u^{(0,0)}$  est solution de l'équation (3.16), donc l'erreur de consistance  $\varepsilon$  est

$$\varepsilon = -\frac{\Delta x^4}{720} \mathbf{A}_1 u_{k,l}^{(0,5)} - \frac{\Delta y^4}{720} \mathbf{A}_2 u_{k,l}^{(5,0)} + o(\Delta x^4) + o(\Delta y^4).$$

L'approximation est donc consistante à l'ordre 4.  $\square$

**Théorème 3.22 (Stabilité)**

L'approximation Volumes Finis à l'ordre  $n$  a la même stabilité qu'une méthode de type Volumes Finis.

**Démonstration** : les composantes  $u^{(1,1)}$ ,  $u^{(2,0)}$  et  $u^{(0,2)}$  sont approximées de la même façon que pour un schéma Volumes Finis. Donc, il existe des constantes  $(c_{1,1}, c_{2,0}, c_{0,2}) \in ]0, 1]^3$ , telles que  $\forall (i, j) \in \{(1, 1), (0, 2), (2, 0)\}$ ,  $\forall N \in \llbracket 1, N_t \rrbracket$ ,

$$\| u_{k,l}^{(i,j)N} \| \leq (c_{i,j} + \mathcal{O}(\Delta t)) \| u_{k,l}^{(i,j)0} \|.$$

Les composantes d'ordre 1 sont déterminées par les Volumes Finis aux Dérivées d'ordre 1. D'après le théorème (3.18),  $\forall (i, j) \in \{(0, 1), (1, 0)\}$ ,  $\forall N \in \llbracket 1, N_t \rrbracket$ ,  $\exists c_{i,j} \leq 1$  telle que

$$\| FV_{k,l}^{(i,j)N} \| \leq (c_{i,j} + \mathcal{O}(\Delta t)) \| u_{k,l}^{(i,j)0} \|.$$

De plus, on sait que  $\| I_{k,l}^{(i,j)N} \| \leq \mathcal{O}(\Delta t) \| u_{k,l}^{(i,j)0} \|$ . D'où la stabilité recherchée sur les dérivées partielles

$$\| u_{k,l}^{(i,j)N} \| \leq (c_{i,j} + \mathcal{O}(\Delta t)) \| u_{k,l}^{(i,j)0} \|.$$

Reste à traiter  $_{k,l}u^{(0,0)}$ . Comme pour l'approximation précédente d'ordre 1, on distingue une contribution identique à celle d'une approximation Volumes Finis,  $FV_{k,l}^{(0,0)N}$ , et un terme de degré 0 en pas d'espace  $I_{k,l}^{(0,0)N}$  (équation (3.21)), et pour ce schéma-ci, une partie de degré 1 en pas d'espace définie par

$$\begin{aligned} II_{k,l}^{(0,0)}(t) &= \mathbf{A}_2^- \frac{\Delta y (u_{k,l+1}^{(0,2)}(t) - u_{k,l}^{(0,2)}(t))}{12} - \mathbf{A}_2^+ \frac{\Delta y (u_{k,l-1}^{(0,2)}(t) - u_{k,l}^{(0,2)}(t))}{12} \\ &+ \mathbf{A}_1^- \frac{\Delta x (u_{k+1,l}^{(2,0)}(t) - u_{k,l}^{(2,0)}(t))}{12} - \mathbf{A}_1^+ \frac{\Delta x (u_{k-1,l}^{(2,0)}(t) - u_{k,l}^{(2,0)}(t))}{12}. \end{aligned}$$

Une fois discrétisé, ce terme devient majorable selon la même méthode que  $I_{k,l}^{0,0}$  à savoir  $\forall N \in \llbracket 1, N_t \rrbracket$ ,

$$\| II_{k,l}^{(0,0) N} \| \leq \Delta t f(\Delta x, \Delta y, \Delta t, c_{0,2}, c_{2,0}, \| u_{k,l}^{(0,2) 0} \|, \| u_{k,l}^{(2,0) 0} \|) \leq \mathcal{O}(\Delta t) \| u_{k,l}^{(0,0) 0} \|$$

$$\| u_{k,l}^{(0,0) N} \| \leq \| FV_{k,l}^{(0,0) N} \| + \| I_{k,l}^{(0,0) N} \| + \| II_{k,l}^{(0,0) N} \| \leq (c_{0,0} + \mathcal{O}(\Delta t)) \| u_{k,l}^{(0,0) 0} \|.$$

On a donc la stabilité  $L^2$  de ce schéma via la stabilité  $l^2$  établie ci dessus.  $\square$

Par le théorème de Lax-Richtmyer, la stabilité et la consistance assurent la convergence.

### **Théorème 3.23 (Convergence)**

*L'approximation Volumes Finis aux Dérivées d'ordre  $n$  (proposition (3.20)) est convergente.*

### 3.2.3 Généralisation de la méthode en bidimensionnel

Soit  $n$  l'ordre d'approximation souhaitée et  $\omega_e$  un maillage quelconque de  $\Omega$ . La solution approchée sur chaque élément  $\omega_e$  est sous la forme polynômiale

$$(u_a(\mathbf{x}, t))_e = \sum_{(i,j)}^{i+j \leq n} u_e^{(i,j)}(t) \frac{x^i y^j}{i! j!}. \quad (3.24)$$

- Si  $i + j = n$ , les fonctions  $u_e^{(i,j)}$  sont approximées par le schéma Volumes Finis.
- Si  $i + j = n - m$ ,  $1 < m < n$ , les fonctions  $u_e^{(i,j)}$  sont approximées par une méthode de type Volumes Finis aux Dérivées d'ordre  $m$ .
- L'écriture de l'approximation  $u_e^{(0,0)}$  se conduit de la même manière que précédemment, en remplaçant  $u_a$  par son approximation dans

$$\iint_{\omega_e} \frac{d(u_a(\mathbf{x}, t)_e)}{dt} + \iint_{\omega_e} \mathbf{A}_i \frac{\partial(u_a(\mathbf{x}, t)_e)}{\partial \mathbf{x}_i} + \int_{\partial \omega_e} \mathbf{A}_i \mathbf{n}_i^- ((u_a(\mathbf{x}, t))_{ext} - (u_a(\mathbf{x}, t))_e) = 0.$$

Il y a donc  $\sum_1^{n+1} k = \frac{(n+1)(n+2)}{2}$  inconnues et autant d'équations.

**Remarque :** la généralisation proposée ne tient pas compte du type de maillage. Cependant, dans un premier temps, le maillage est supposé structuré, où chaque élément  $\omega_{k,l}$  avec  $(k, l) \in \llbracket 1, N_x \rrbracket \times \llbracket 1, N_y \rrbracket$  est un rectangle de coté  $\Delta x$  et  $\Delta y$ . Cette hypothèse permet de simplifier l'écriture du schéma, mais surtout d'obtenir la démonstration de consistance à l'aide de la formule de Taylor. Quant à la stabilité, comme elle est héritée de la méthode Volumes Finis, la démonstration peut être conduite aussi bien pour des maillages structurés que déstructurés, tant que les éléments de ce dernier sont assez réguliers (volumes et rapports volume/périmètre).

**Lemme 3.24**

L'approximation Volumes Finis aux Dérivées d'ordre  $n$  pour les équations (3.16) est solution pour tout élément  $\omega_{k,l}$  de

$$\begin{aligned} \frac{du_{k,l}(t)}{dt} + \sum_{i=0}^n \alpha_i \left( \Delta x^{i-1} \left[ \mathbf{A}_1^- \left( u_{k+1,l}^{(i,0)}(t) - u_{k,l}^{(i,0)}(t) \right) + (-1)^i \mathbf{A}_1^+ \left( u_{k-1,l}^{(i,0)}(t) - u_{k,l}^{(i,0)}(t) \right) \right] \right. \\ \left. + \Delta y^{i-1} \left[ \mathbf{A}_2^- \left( u_{k,l+1}^{(0,i)}(t) - u_{k,l}^{(0,i)}(t) \right) + (-1)^i \mathbf{A}_2^+ \left( u_{k,l-1}^{(0,i)}(t) - u_{k,l}^{(0,i)}(t) \right) \right] \right) \equiv 0. \end{aligned} \quad (3.25)$$

où les coefficients  $\alpha_i$  sont construits par récurrence :  $\alpha_0 = 1$ ,  $\alpha_1 = 1/2$  et  $\forall i \geq 2$ ,

$$\begin{cases} \alpha_i = 0 & \text{si } i \text{ est impair,} \\ \alpha_i = \frac{1}{2^i i!} + \sum_{j=0, j \text{ pair}}^{i-1} \frac{-\alpha_j}{2^{i-j}(i-j+1)!} & \text{sinon.} \end{cases}$$

**Démonstration** : la solution approchée à l'ordre  $n$  s'écrit sur l'élément  $\omega_{k,l}$  sous la forme polynômiale

$$\sum_{i,j}^{i+j \leq n} u_{k,l}^{(i,j)}(t) \frac{x^i y^j}{i! j!}.$$

On remplace cette expression dans la formulation variationnelle associée à (3.16)

$$\iint_{\omega_{k,l}} \frac{\partial u_a(\mathbf{x}, t)}{\partial t} + \iint_{\omega_{k,l}} \mathbf{A}_i \frac{\partial u_a(\mathbf{x}, t)}{\partial \mathbf{x}_i} + \int_{\partial \omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- (u_a(\mathbf{x}, t)_{ext} - u_a(\mathbf{x}, t)) = 0.$$

Les dérivées partielles d'ordre 1 de  $u_a$  sur  $\omega_{k,l}$  sont

$$\frac{\partial u_a(\mathbf{x}, t)}{\partial x} = \sum_{i,j}^{i+j \leq n} u_{k,l}^{(i,j)}(t) \frac{x^{i-1} y^j}{(i-1)! j!} \quad \frac{\partial u_a(\mathbf{x}, t)}{\partial y} = \sum_{i,j}^{i+j \leq n} u_{k,l}^{(i,j)}(t) \frac{x^i y^{j-1}}{i! (j-1)!}.$$

Comme les intervalles d'intégration sont symétriques, les intégrales doubles de type  $\iint_{\omega_{k,l}} x^i y^j dx dy = \int_{-\Delta x/2}^{-\Delta x/2} x^i dx \cdot \int_{-\Delta y/2}^{-\Delta y/2} y^j dy$  sont non nulles sous certaines conditions de parité pour  $i$  et  $j$  (dans ce cas, que si  $i$  impair et  $j$  pair). D'où les simplifications

$$\begin{cases} \iint_{\omega_{k,l}} \mathbf{A}_1 \frac{\partial u_a(\mathbf{x}, t)}{\partial x} = \mathbf{A}_1 \sum_{i \text{ impair}, j \text{ pair}}^{i+j \leq n} u_{k,l}^{(i,j)}(t) \frac{\Delta x^i \Delta y^{j+1}}{2^{i+j-1} i! (j+1)!} \\ \iint_{\omega_{k,l}} \mathbf{A}_2 \frac{\partial u_a(\mathbf{x}, t)}{\partial y} = \mathbf{A}_2 \sum_{i \text{ pair}, j \text{ impair}}^{i+j \leq n} u_{k,l}^{(i,j)}(t) \frac{\Delta x^{i+1} \Delta y^j}{2^{i+j-1} (i+1)! j!} \\ \iint_{\omega_{k,l}} \frac{\partial u_a(\mathbf{x}, t)}{\partial t} = \sum_{i,j \text{ pairs}}^{i+j \leq n} \frac{du_{k,l}^{(i,j)}(t)}{dt} \frac{\Delta x^{i+1} \Delta y^{j+1}}{2^{i+j} (i+1)! (j+1)!} \end{cases}$$

Les intégrales de bord  $\int_{\partial\omega_{k,l}} \mathbf{A}_i \mathbf{n}_i^- (u_a(\mathbf{x}, t)_{ext} - u_a(\mathbf{x}, t))$  se décomposent en quatre intégrales vu que les éléments sont des rectangles. Par exemple, pour la partie droite,  $I_d(t) = \int_{-\Delta y/2}^{-\Delta y/2} \mathbf{A}_1^- \sum_{i,j}^{i+j \leq n} \left[ u_{k+1,l}^{(i,j)}(t) \frac{(-\Delta x)^i y^j}{2^i i! j!} - u_{k,l}^{(i,j)}(t) \frac{\Delta x^i y^j}{2^i i! j!} \right] dy$ , les puissances de  $j$  impaires ne contribuent pas à cette intégrale, de même que les puissances paires de  $i$ . D'où

$$\left\{ \begin{array}{l} I_d(t) = -\mathbf{A}_1^- \sum_{i,j \text{ pair}}^{i+j \leq n} \left( (-1)^i u_{k+1,l}^{(i,j)}(t) - u_{k,l}^{(i,j)}(t) \right) \frac{\Delta x^i \Delta y^{j+1}}{2^{i+j} i! (j+1)!} \\ I_g(t) = -\mathbf{A}_1^+ \sum_{i,j \text{ pair}}^{i+j \leq n} \left( u_{k-1,l}^{(i,j)}(t) - (-1)^i u_{k,l}^{(i,j)}(t) \right) \frac{\Delta x^i \Delta y^{j+1}}{2^{i+j} i! (j+1)!} \\ I_h(t) = -\mathbf{A}_2^- \sum_{i \text{ pair}, j}^{i+j \leq n} \left( (-1)^j u_{k,l+1}^{(i,j)}(t) - u_{k,l}^{(i,j)}(t) \right) \frac{\Delta x^{i+1} \Delta y^j}{2^{i+j} (i+1)! j!} \\ I_b(t) = -\mathbf{A}_2^+ \sum_{i \text{ pair}, j}^{i+j \leq n} \left( u_{k,l-1}^{(i,j)}(t) - (-1)^j u_{k,l}^{(i,j)}(t) \right) \frac{\Delta x^{i+1} \Delta y^j}{2^{i+j} (i+1)! j!} \end{array} \right.$$

Comme  $\mathbf{A}_1 = \mathbf{A}_1^- + \mathbf{A}_1^+$ , l'équation (3.26), portant sur les indices impairs de  $i$  permet d'éliminer les  $(-1)^i$ , d'où  $I_d(t) + I_g(t) + \iint_{\omega_{k,l}} \mathbf{A}_1 \frac{\partial u_a(\mathbf{x}, t)}{\partial x}$  égale in fine

$$-\frac{\Delta x^i \Delta y^{j+1}}{2^{i+j} i! (j+1)!} \sum_{i,j \text{ pair}}^{i+j \leq n} \left[ (-1)^i \mathbf{A}_1^- \left( u_{k+1,l}^{(i,j)}(t) - u_{k,l}^{(i,j)}(t) \right) + \mathbf{A}_1^+ \left( u_{k-1,l}^{(i,j)}(t) - u_{k,l}^{(i,j)}(t) \right) \right].$$

Donc

$$\begin{aligned} & \frac{du_{k,l}^{(0,0)}(t)}{dt} + \sum_{(i,j)=(2,2) \text{ pairs}}^{i+j \leq n} \frac{\Delta x^i \Delta y^j}{2^{i+j} (i+1)! (j+1)!} \frac{du_{k,l}^{(i,j)}(t)}{dt} \\ &= \sum_{j \text{ impair}}^{i+j \leq n} \frac{\Delta x^{i-1} \Delta y^j}{2^{i+j} i! (j+1)!} \left( \mathbf{A}_1^- (-1)^i \left( u_{k+1,l}^{(i,j)}(t) - u_{k,l}^{(i,j)}(t) \right) + \mathbf{A}_1^+ \left( u_{k-1,l}^{(i,j)}(t) - u_{k,l}^{(i,j)}(t) \right) \right) \\ &+ \sum_{i \text{ impair}}^{i+j \leq n} \frac{\Delta x^i \Delta y^{j-1}}{2^{i+j} (i+1)! j!} \left( \mathbf{A}_2^- (-1)^j \left( u_{k,l+1}^{(i,j)}(t) - u_{k,l}^{(i,j)}(t) \right) + \mathbf{A}_2^+ \left( u_{k,l-1}^{(i,j)}(t) - u_{k,l}^{(i,j)}(t) \right) \right). \end{aligned} \quad (3.26)$$

Il s'agit d'un système triangulaire, qui pourra être mis sous la forme annoncée dans le lemme. Le calcul des  $\alpha_i$  se mène de façon identique au calcul général proposé pour l'Equation d'Advection monodimensionnelle.  $\square$

### **Théorème 3.25 (Consistance)**

L'approximation proposée (3.24) est consistante à l'ordre  $n+2$  si  $n$  est pair,  $n+1$  sinon.

**Démonstration** : l'idée de cette preuve est encore une fois l'utilisation successives de formules de Taylor. Utilisons pour la preuve le lemme (3.24) en appliquant la formule de Taylor sur les termes différences. Soit  $p$  l'ordre du reste choisi. Sachant que  $u$  est solution (3.16) (ce qui permet d'éliminer le terme en  $\frac{d}{dt}$ ), l'erreur de consistance  $\varepsilon$  est donnée par

$$\varepsilon = \sum_{i=0}^p \beta_i ((\Delta x)^i + (\Delta y)^i) + o(\Delta x^{p+i}) + o(\Delta y^{p+i}),$$

où

$$\begin{cases} \beta_i = \sum_{j=0}^{\min(i-1, n)} \frac{(-1)^{i-j+1}}{(i-j)!} \alpha_j^- & \text{si } i < n+2, \\ \beta_i = 0 & \text{sinon.} \end{cases}$$

Le schéma à l'ordre  $n$  pair est donc consistant d'ordre  $n+2$ . La nullité des coefficients  $\alpha_i$  impairs assure que le schéma pour un  $n$  impair est le même que celui de l'entier pair précédent, la consistance reste donc la même pour l'ordre impair qui suit.  $\square$

### **Théorème 3.26 (Stabilité)**

*L'approximation (3.24) a la même stabilité qu'une méthode de type Volumes Finis.*

**Démonstration** : Notons  $\mathcal{P}(p)$  la proposition : les composantes d'ordre  $n-p$  sont  $l^2$  stables.

*Initialisation* : si  $i+j=n$ , la fonction  $\frac{\partial^n \varphi}{\partial x^i \partial y^j}$  correspond au centre de gravité de l'élément au coefficient  $u^{(i,j)}$  et son approximation se fait par une méthode de type Volumes Finis. La stabilité  $l^2$  et la convergence sont donc assurées par cette méthode. Soit  $u_{k,l}^{(i,j)(N)}$  désigne l'approximation (type Euler) au  $N^e$  pas de temps de la composante  $u_{k,l}^{(i,j)}$ ,  $N_t$  le nombre final d'itérations temporelles et  $\Delta t$  le pas de temps. Alors il existe des constantes  $c_{i,j} \in ]0, 1]$ , où  $i+j=n$  telles que  $\forall N \in \llbracket 1, N_t \rrbracket$  on ait

$$\| u_{k,l}^{(i,j)(N)} \| \leq (c_{i,j} + \mathcal{O}(\Delta t)) \| u_{k,l}^{(i,j)0} \|.$$

Donc la propriété  $\mathcal{P}(0)$  est vraie.

*Réurrence* : supposons  $\mathcal{P}(p-1)$  vraie. La procédure est la même que pour l'Equation d'Advection monodimensionnelle. On se sert de la stabilité établie des dérivées d'ordre supérieur et de la forme particulière des équations de cette approximation : si  $i+j < n$ , le schéma peut être décomposé grâce à (3.26) sous la forme d'une partie similaire à une formulation de type Volumes Finis auquel s'ajoutent des termes de la forme  $\mathcal{O}(\Delta x)(\dots) + \mathcal{O}(\Delta y)(\dots)$ . L'approximation sous un schéma d'Euler explicite conduit à la stabilité  $\| u_{k,l}^{(i,j)(N)} \| \leq (1 + \mathcal{O}(\Delta t)) \| u_{k,l}^{(i,j)0} \|$ .  $\square$

**Remarque :** cette démonstration peut être conduite de manière analogue en déstructurée. En effet, si pour tout élément  $\omega_e$  on a  $\mathbf{vol}(\omega_e)$  borné et  $\frac{\mathbf{vol}(\omega_e)}{\mathbf{vol}(\delta\omega_e)}$ , il y a équivalence entre la norme  $l^2$  des coefficients et la norme  $L^2$  des polynômes.

Par le théorème de Lax-Richtmyer, la stabilité et la consistance assurent la convergence.

**Théorème 3.27 (Convergence)**

*L'approximation proposée (3.24) est convergente.*

### 3.3 Simulations numériques

Les diverses simulations numériques seront réalisées pour l'Equation des Ondes bidimensionnelle, donc à coefficients constants, avec

$$\mathbf{A}_0 = \mathbf{I}_3 \quad \mathbf{A}_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \mathbf{A}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

#### Problème en condition initiale - Maillage structuré

Nous choisissons la résolution d'un problème où la condition initiale est une unité approchée à support dans un disque  $D$  de centre  $(x_c, y_c)$  et de rayon  $R$ .

$$\rho_0(x, y) = \rho_0(x, y) \mathbf{1}_D(x, y) = \exp\left(\frac{(x - x_c)^2 + (y - y_c)^2}{(x - x_c)^2 + (y - y_c)^2 - R^2}\right) \mathbf{1}_D(x, y). \quad (3.27)$$

Nous disposons de résultats monodimensionnels. Il est possible de ramener le cas bidimensionnel au cas monodimensionnel en étudiant, dans un premier temps, le mode transverse. Ce mode consiste en une invariance en  $y$ . Il y aura donc sur  $y$  une condition d'obstacle parfait. Soit  $\Gamma_1$  la frontière verticale avec des conditions de Silver Müller. Soit  $\Gamma_2$  la frontière horizontale avec des conditions d'obstacle parfait. Posons

$$\mathbf{M}_\beta = \frac{1}{2} \begin{pmatrix} (\beta + 1)\mathbf{n} \times \mathbf{n} & (\beta - 1)\mathbf{n} \\ -(\beta + 1)\mathbf{n}^T & (1 - \beta) \end{pmatrix} \text{ avec } \beta \in [-1, 1]. \text{ Alors } \mathbf{M}_1 = \begin{pmatrix} \mathbf{n} \times \mathbf{n} & 0 \\ -\mathbf{n}^T & 0 \end{pmatrix} \text{ et}$$

$$\mathbf{A}_i \mathbf{n}_i^- = \mathbf{M}_0 = -\frac{1}{2} \begin{pmatrix} \mathbf{n} \times \mathbf{n} & -\mathbf{n} \\ -\mathbf{n}^T & 1 \end{pmatrix}. \text{ La formulation variationnelle associée est donc,}$$

en notant  $\Gamma = \Gamma_1 \cup \Gamma_2$ ,

$$\begin{aligned} \iint_{\omega_k} \left( \frac{\partial u(\mathbf{x}, t)}{\partial t} + \mathbf{A}_1 \frac{\partial u(\mathbf{x}, t)}{\partial x} \right) + \int_{\partial\omega_k/\Gamma} \mathbf{A}_i \mathbf{n}_i^- (u_{ext}(\mathbf{x}, t) - u(\mathbf{x}, t)) \\ + \int_{\partial\omega_k \cap \Gamma_1} -\mathbf{A}_i \mathbf{n}_i^- (u(\mathbf{x}, t)) + \int_{\partial\omega_k \cap \Gamma_2} \mathbf{M}_1 u(\mathbf{x}, t) = 0. \end{aligned}$$

La condition initiale est donc modifiée pour ne dépendre que de  $x$  à support dans un canal.

$$\rho_0(x) = \mathbf{exp} \left( \frac{(x - x_c)^2}{(x - x_c)^2 - R^2} \right) \mathbb{1}_D(x, y).$$

Cette condition initiale est bien dérivable : ses dérivées partielles seront les conditions initiales des équations sur les dérivées. Pour la méthode d'ordre 1, la solution est approximée sous la forme

$$u(\mathbf{x}, t) = u^{(0,0)}(t) + u^{(1,0)}(t)x + u^{(0,1)}(t)y.$$

Les équations sur  $u^{(1,0)}$  et  $u^{(0,1)}$  sont les mêmes que pour une méthode de Volumes Finis classiques mais les conditions initiales seront les dérivées partielles de  $\rho_0(\mathbf{x})$ . Comme  $\frac{\partial \rho_0(\mathbf{x})}{\partial y} = 0$ , il n'y aura pas de condition initiale pour l'équation sur  $u^{(0,1)}$ . A l'inverse,  $\frac{\partial \rho_0(\mathbf{x})}{\partial x} = \frac{-2R^4(x - x_c)}{(R^4 - (x - x_c)^2)} \mathbf{exp} \left( \frac{(x - x_c)^2}{(x - x_c)^2 - R^2} \right) \mathbb{1}_D(x, y)$ , sera la condition initiale pour l'équation sur  $u^{(1,0)}$ . Pour l'approximation temporelle, on choisit une méthode de Runge-Kutta d'ordre 2, qui sera programmée sous la forme de deux schémas Euler Explicites successifs. Soit  $\Delta t$  le pas de temps,  $U_0$  la solution au temps précédent, et  $U_1$  la solution intermédiaire. Si le problème peut s'écrire sous la forme  $\frac{dU}{dt} = \mathbf{M}U_0$ , le schéma Runge-Kutta choisi s'écrit :

$$\begin{cases} \frac{U_1 - U_0}{\Delta t} = \mathbf{M}U_0 \\ \frac{U_2 - U_1}{\Delta t} = \mathbf{M}U_1 \\ U = \frac{U_2 + U_0}{2} \end{cases}$$

En recomposant la solution, on voit bien que  $U = U_0 + \Delta t \mathbf{M}U_0 + \frac{\Delta t^2}{2} \mathbf{M}^2 U_0$  est d'ordre 2. La condition CFL associée à cette méthode pour une méthode de type Volumes Finis est de  $1/2$ .

Par ailleurs, pour ce problème, il est possible de calculer la solution analytique. Elle nous donnera ainsi un critère de comparaison des plus efficaces, et on implé-

mentera aussi une méthode de type Volumes Finis classique.

$$\begin{cases} \frac{\partial u(x, y, t)}{\partial t} + \frac{\partial \rho(x, y, t)}{\partial x} & = 0 \\ \frac{\partial v(x, y, t)}{\partial t} + \frac{\partial \rho(x, y, t)}{\partial y} & = 0 \\ \frac{\partial \rho(x, y, t)}{\partial t} + \frac{\partial u(x, y, t)}{\partial x} + \frac{\partial v(x, y, t)}{\partial y} & = 0. \end{cases}$$

Comme on est invariant en  $y$ , on se ramène au système

$$\begin{cases} \frac{\partial u(x, t)}{\partial t} + \frac{\partial \rho(x, t)}{\partial x} & = 0 \\ \frac{\partial v(x, t)}{\partial t} & = 0 \\ \frac{\partial \rho(x, t)}{\partial t} + \frac{\partial u(x, t)}{\partial x} & = 0. \end{cases}$$

Les inconnues  $w = \rho - u$  et  $\mu = \rho + u$  sont alors solutions des équations de transport

$$\begin{cases} \frac{\partial w(x, t)}{\partial t} - \frac{\partial w(x, t)}{\partial x} & = 0 \\ \frac{\partial \mu(x, t)}{\partial t} + \frac{\partial \mu(x, t)}{\partial x} & = 0. \end{cases}$$

On rappelle que la résolution des Equations de transport  $\frac{\partial \varphi(x, t)}{\partial t} + c \frac{\partial \varphi(x, t)}{\partial x} = 0$  passe par la méthode des caractéristiques et donne comme solution, en notant  $\varphi_0$  la condition initiale,  $\varphi(x, t) = \varphi_0(x - ct)$ . Ici, les vitesses  $c$  valent respectivement 1 et  $-1$ , d'où les solutions suivantes :

$$\begin{cases} w(x, t) = \rho_0(x - t) \\ \mu(x, t) = \rho_0(x + t) \end{cases}$$

$$\begin{cases} u(x, t) = \frac{1}{2}(\rho_0(x - t) - \rho_0(x + t)) \\ v(x, t) = 0 \\ \rho(x, t) = \frac{1}{2}(\rho_0(x - t) + \rho_0(x + t)). \end{cases}$$

Soit un carré de longueur  $5 \times 5$  avec 100 éléments par longueur. Les résultats numériques présentés portent sur l'Equation des Ondes bidimensionnelle en mode Transverse avec différents schémas en temps : pour la figure (3.8) il s'agit d'un schéma d'Euler Explicite alors que pour les figures suivantes, une méthode de Runge-Kutta d'ordre 2 aura été choisie. Afin de réaliser un test le plus difficile possible, nous ne prendrons que peu d'éléments sur le support de la condition initiale : le nombre d'éléments variera en effet entre 10 éléments (figures (3.8) et (a)) et 20 pour

la figure (b). Le centre de la condition initiale sera en  $(2.5, 2.5)$  pour un rayon allant donc de 0.1 ou 0.2. Les courbes présentées sont la solution pour une condition initiale au centre du domaine en un point d'abscisse 2.9 fonction du nombre d'itérations en temps.

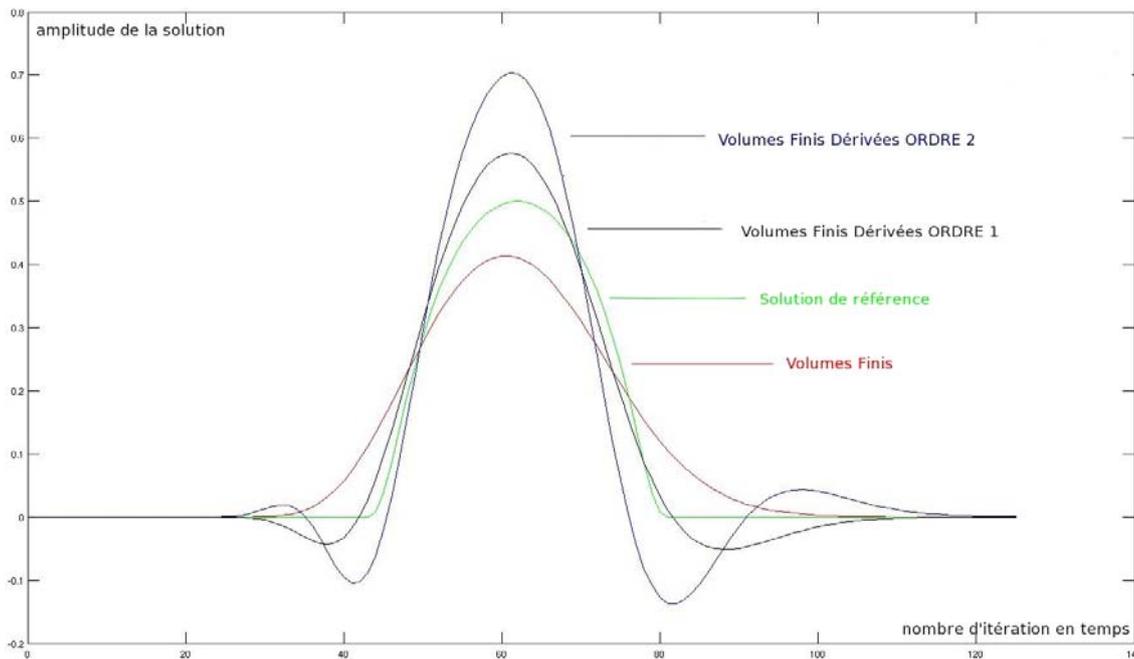
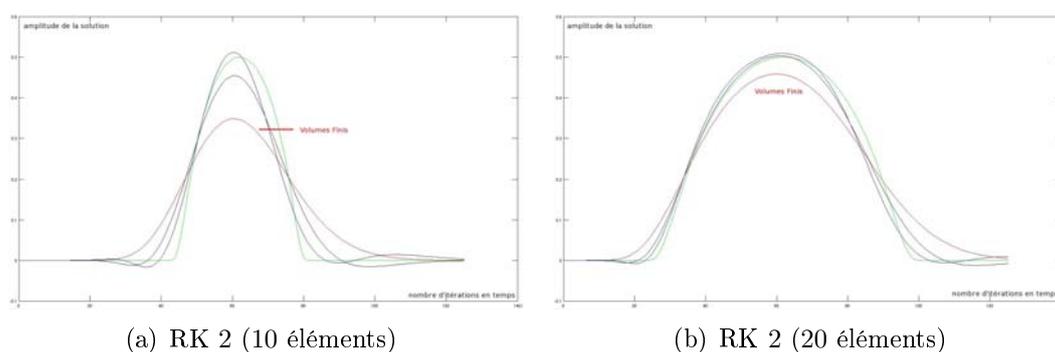


FIG. 3.8: EE(10 éléments)

L'approximation est polynômiale. Or, sur la frontière de la condition initiale, à savoir  $(x-x_c)^2 = R^2$ , la série de Taylor n'est pas convergente (coefficients nuls), donc la fonction n'est pas analytique dans ce voisinage et l'approximation polynômiale est difficile : elle manifeste alors des overshoots et des undershoots. La méthode de type Volumes Finis respecte certes la positivité mais pénalise fortement l'amplitude de la solution. Si on augmente l'approximation temporelle (figure (a)), on constate que les over/under - shoots sont compensés. Et en augmentant le nombre d'élément sur le support de la condition initiale (figure (b)), on constate la convergence entre la solution analytique et les solutions de cette méthode, alors que la méthode de type Volumes Finis reste d'amplitude inférieure.



Nous allons maintenant établir une comparaison avec les méthodes de type Galerkin Discontinu. Le cas test est toujours un carré de dimension  $5 \times 5$  avec cette fois-ci 200 éléments par longueur. On choisit une approximation temporelle de type Euler Explicite. La condition initiale est la régularisée d'un dirac (3.27), placée au centre du carré et de rayon 0.2. Le point de visualisation est le point (2.9, 2.5). Pour une méthode de Galerkin d'ordre 1, la solution est approximée par un polynôme de type  $a_0 + a_1x + a_2y + a_3xy$ . La méthode des Volumes Finis aux Dérivées à l'ordre 1 ne comporte pas de terme croisé  $xy$  : l'approximation est moins précise que la méthode  $Q1$ . A l'ordre 2 cette différence est comblée, et la méthode présente l'avantage d'induire un undershoot moins important que la méthode  $Q1$ .

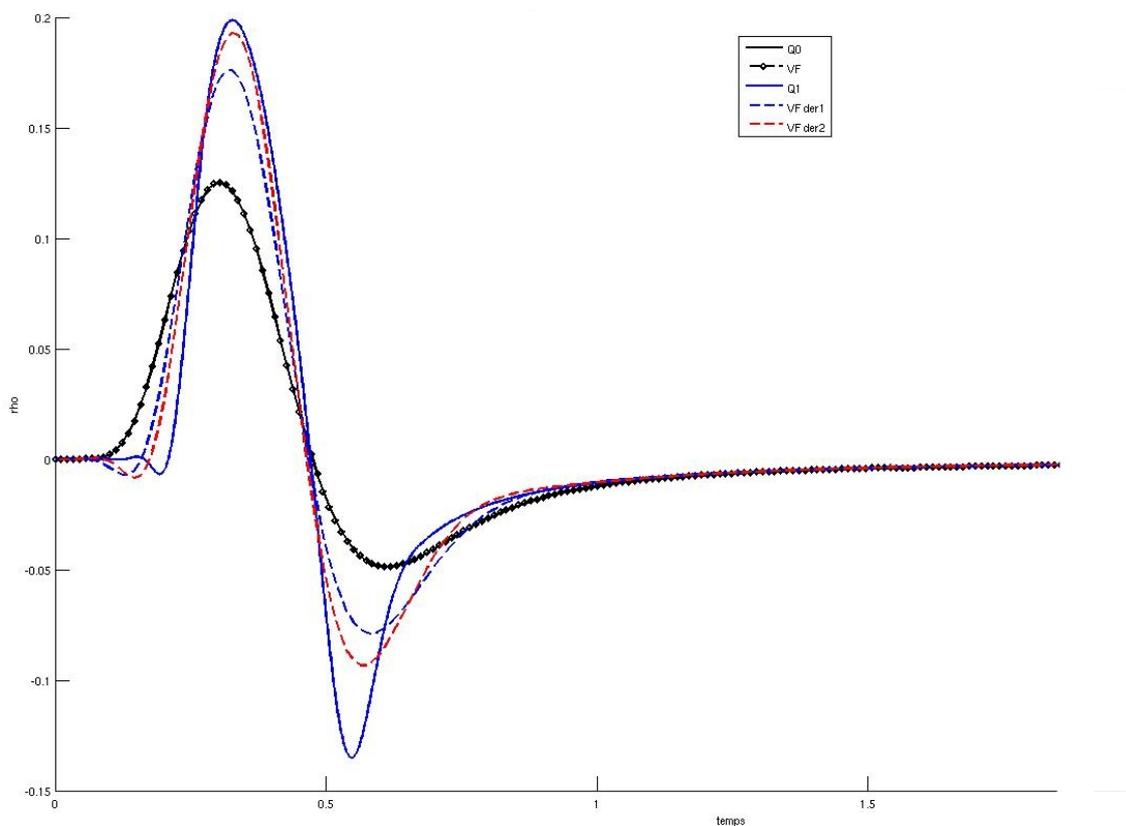


FIG. 3.9: Comparaison Volumes Finis aux Dérivées - Galerkin Discontinu

### Problème en condition initiale - Maillage déstructuré

Nous allons maintenant nous intéresser au cas de maillages déstructurés. La cellule de référence est un triangle  $T$  supposé non dégénéré. Soit  $\mathbf{n}$  la normale sortante à  $T$  et  $\varphi$  la solution de la formulation variationnelle :

$$\iint_T \left( \frac{\partial u(\mathbf{x}, t)}{\partial t} + \mathbf{A}_1 \frac{\partial u(\mathbf{x}, t)}{\partial x} \right) + \int_{\partial T} \mathbf{A}_i \mathbf{n}_i^- (u_{ext} - u) = 0. \quad (3.28)$$

Nous avons deux possibilités pour exprimer le schéma. La première est de se placer dans un repère lié à l'élément  $T$  considéré, en plaçant l'origine au centre de gravité du triangle. Ainsi, nous aurons  $\iint_T x = \iint_T y = 0$ . La deuxième est de garder le repère classique. Les deux formulations sont équivalentes, mais nous préférons la seconde.

#### Propriété 3.28 (Implémentation pour la méthode à l'ordre 1)

Soit un maillage déstructuré  $\kappa = \bigcup_{i \in \mathbf{I}} T_i$ . Soit un système de Friedrichs du type (3.28).

Soit l'approximation de la solution sous la forme d'un polynôme de degré 1, notée  $u(\mathbf{x}, t)_{T_i} = u_{T_i}^{(0,0)}(t) + u_{T_i}^{(1,0)}(t)x + u_{T_i}^{(0,1)}(t)y$ . Soit  $G_i$  le centre de gravité de l'élément,  $V_i$  le triangle voisin opposé au sommet  $S_i$ ,  $l_i$  la longueur de la frontière entre les deux éléments, et  $a_i$  l'aire du triangle  $T_i$ . Alors  $\forall i \in \mathbf{I}$ ,

$$\left\{ \begin{array}{l} \frac{du_{T_i}^{(0,0)}(t)}{dt} - \sum_{i=1}^3 x_{G_i} l_i \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(1,0)}(t) - u_{T_i}^{(1,0)}(t)) \\ \quad - \sum_{i=1}^3 y_{G_i} l_i \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(0,1)}(t) - u_{T_i}^{(0,1)}(t)) = 0, \\ a_i \frac{du_{T_i}^{(1,0)}(t)}{dt} + \sum_{i=1}^3 l_i \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(1,0)}(t) - u_{T_i}^{(1,0)}(t)) = 0, \\ a_i \frac{du_{T_i}^{(0,1)}(t)}{dt} + \sum_{i=1}^3 l_i \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(0,1)}(t) - u_{T_i}^{(0,1)}(t)) = 0. \end{array} \right.$$

**Démonstration :** Soit un triangle  $T$ . La solution est approchée à l'ordre 1 par  $u_T = u_T^{(0,0)}(t) + u_T^{(1,0)}(t)x + u_T^{(0,1)}(t)y$ . La formulation variationnelle (3.28) donne alors trois équations, celles sur les dérivées d'ordre 1 correspondant à des équations de type Volumes Finis. Nous noterons  $V_i$  le triangle voisin opposé au sommet  $S_i$ ,  $l_i$  la longueur de la frontière entre les deux éléments, et  $a$  l'aire du triangle  $T$ .

$$\left\{ \begin{array}{l} \iint_T \frac{du_T^{(1,0)}(t)}{dt} + \int_{\partial T} \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(1,0)}(t) - u_T^{(1,0)}(t)) = 0, \\ \iint_T \frac{du_T^{(0,1)}(t)}{dt} + \int_{\partial T} \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(0,1)}(t) - u_T^{(0,1)}(t)) = 0. \end{array} \right.$$

D'où

$$\begin{cases} \frac{du_T^{(1,0)}(t)}{dt} + \sum_{i=1}^3 \frac{l_i}{a} \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(1,0)}(t) - u_T^{(1,0)}(t)) = 0, \\ \frac{du_T^{(0,1)}(t)}{dt} + \sum_{i=1}^3 \frac{l_i}{a} \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(0,1)}(t) - u_T^{(0,1)}(t)) = 0. \end{cases}$$

Détaillons maintenant celle pour le terme d'ordre 0.

$$\begin{aligned} \frac{\partial u(\mathbf{x}, t)}{\partial t} &= \frac{du_T^{(0,0)}(t)}{dt} + \frac{du_T^{(1,0)}(t)}{dt} x + \frac{du_T^{(0,1)}(t)}{dt} y \\ \Rightarrow \iint_T \frac{\partial u(\mathbf{x}, t)}{\partial t} dx dy &= a \frac{du_T^{(0,0)}(t)}{dt} + \frac{du_T^{(1,0)}(t)}{dt} \iint_T x dx dy + \frac{du_T^{(0,1)}(t)}{dt} \iint_T y dx dy \end{aligned}$$

Il faut donc calculer ces intégrales de surfaces. Pour se faire, utilisons une transformation affine du plan  $F$  qui transforme le triangle  $T = (S_1, S_2, S_3)$  en un triangle  $\hat{T} = (\hat{S}_1, \hat{S}_2, \hat{S}_3)$  isocèle rectangle de coté 1 donc le sommet  $\hat{S}_1$  est l'origine.

$$\begin{pmatrix} x \\ y \end{pmatrix} = F \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix}.$$

Soit  $\mathbf{J}$  la matrice jacobienne de la transformation  $F$ , alors

$$\begin{aligned} \iint_T x dx dy &= \iint_{\hat{T}} |\det(\mathbf{J})| (x_1 + (x_2 - x_1)\hat{x} + (x_3 - x_1)\hat{y}) d\hat{x} d\hat{y}. \\ |\det(\mathbf{J})| &= |(x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)| = 2a. \end{aligned}$$

De plus, sur l'élément  $\hat{T}$ , on peut calculer aisément les intégrales de surfaces :

$$\iint_{\hat{T}} d\hat{x} d\hat{y} = \frac{1}{2}, \quad \iint_{\hat{T}} \hat{x} d\hat{x} d\hat{y} = \frac{1}{6}, \quad \iint_{\hat{T}} \hat{y} d\hat{x} d\hat{y} = \frac{1}{6}.$$

On obtient alors

$$\begin{aligned} \iint_T x dx dy &= a \frac{x_1 + x_2 + x_3}{3} = ax_G \\ \iint_T y dx dy &= a \frac{y_1 + y_2 + y_3}{3} = ay_G \end{aligned}$$

$$\text{D'où } \iint_T \frac{\partial u(\mathbf{x}, t)}{\partial t} dx dy = a \left[ \frac{du_T^{(0,0)}(t)}{dt} + \frac{du_T^{(1,0)}(t)}{dt} x_G + \frac{du_T^{(0,1)}(t)}{dt} y_G \right].$$

Or comme les dérivées d'ordre 1 sont solutions de (3.29), on en déduit

$$\begin{aligned} \iint_T \frac{\partial u(\mathbf{x}, t)}{\partial t} dx dy &= a \frac{du_T^{(0,0)}(t)}{dt} - \sum_{i=1}^3 x_G l_i \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(1,0)}(t) - u_T^{(1,0)}(t)) \\ &\quad - \sum_{i=1}^3 y_G l_i \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(0,1)}(t) - u_T^{(0,1)}(t)). \end{aligned} \quad (3.29)$$

Comme  $\iint_T \mathbf{A}_i \frac{\partial u(\mathbf{x}, t)}{\partial \mathbf{x}_i} dx dy = a(\mathbf{A}_1 u_T^{(1,0)}(t) + \mathbf{A}_2 u_T^{(0,1)}(t))$ , il ne reste à calculer que la contribution des intégrales de bords.

$$\begin{aligned} \int_{\partial T} \mathbf{A}_i \mathbf{n}_i^- (u_{ext}(\mathbf{x}, t) - u(\mathbf{x}, t)) &= \sum_{i=1}^3 l_i \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(0,0)}(t) - u_T^{(0,0)}(t)) \\ &+ \int_{\partial T} \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(1,0)}(t) x_{V_i} - u_T^{(1,0)}(t) x_T) + \int_{\partial T} \mathbf{A}_i \mathbf{n}_i^- (u_{V_i}^{(0,1)}(t) y_{V_i} - u_T^{(0,1)}(t) y_T). \end{aligned}$$

Le calcul des intégrales de bords à l'aide de la méthode de Simpson est exacte pour des polynômes d'ordre 1. Elle nécessite les coordonnées du point milieu de l'arête, que nous noterons  $I_i$  pour la frontière entre le triangle  $T$  et son voisin  $V_i$ .

$$\begin{cases} \int_{\partial T} x_{V_i} = \sum_{i=1}^3 (l_i x_{I_i} - l_i (x_{GV_i} - x_{GT})) \\ \int_{\partial T} y_{V_i} = \sum_{i=1}^3 (l_i y_{I_i} - l_i (y_{GV_i} - y_{GT})). \end{cases}$$

$$\begin{cases} \int_{\partial T} x_T = \sum_{i=1}^3 l_i x_{I_i} \\ \int_{\partial T} y_T = \sum_{i=1}^3 l_i y_{I_i}. \end{cases}$$

$$\begin{aligned} \int_{\partial T} \mathbf{A}_i \mathbf{n}_i^- (u_{ext}(\mathbf{x}, t) - u(\mathbf{x}, t)) &= \sum_{i=1}^3 l_i \mathbf{A}_i \mathbf{n}_i^- \left[ \left( u_{V_i}^{(0,0)}(t) - u_T^{(0,0)}(t) \right) \right. \\ &+ x_{I_i} \left( u_{V_i}^{(1,0)}(t) - u_T^{(1,0)}(t) \right) + y_{I_i} \left( u_{V_i}^{(0,1)}(t) - u_T^{(0,1)}(t) \right) \\ &\left. - \left( (x_{GV_i} - x_{GT}) u_{V_i}^{(1,0)}(t) + (y_{GV_i} - y_{GT}) u_{V_i}^{(0,1)}(t) \right) \right]. \end{aligned} \quad (3.30)$$

En sommant les équations (3.29) et (3.30), on obtient alors l'équation sur le terme constant (d'ordre 0). D'où le résultat énoncé.  $\square$

Pour les maillages déstructurés, nous n'avons pas prouvé la consistance de notre approximation. Soit le maillage de longueur  $5 \times 5$  avec 8000 éléments. Il est composé de carrés régulièrement coupés en deux, et permet ainsi de faire tourner le code pour les maillages structurés (figure 3.11) comme le code pour les maillages déstructurés (figure 3.12). La source est placée au centre, en  $(2.5, 2.5)$  de rayon 0.4. La condition CFL choisie est 0.5. Les deux solutions de la méthode à l'ordre 1 sont identiques. Comparées à une solution obtenue avec une méthode de type Volumes Finis (figure 3.10), on constate un gain d'amplitude significatif et un aspect plus net des isocontours.

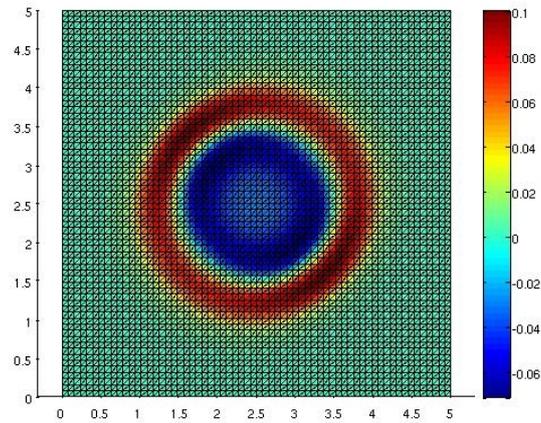


FIG. 3.10: Schéma Explicite. Volumes Finis. Code déstructuré

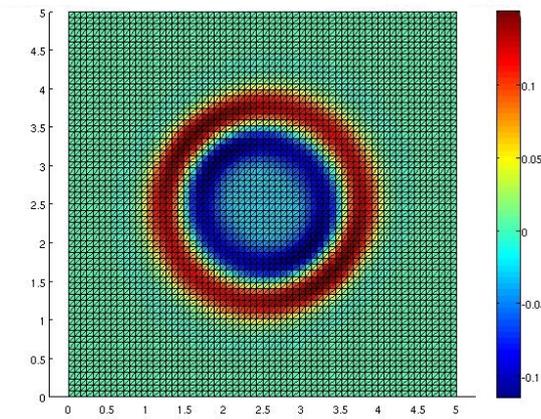


FIG. 3.11: Schéma Explicite. Volumes Finis Dérivées ORDRE 1. Code structuré

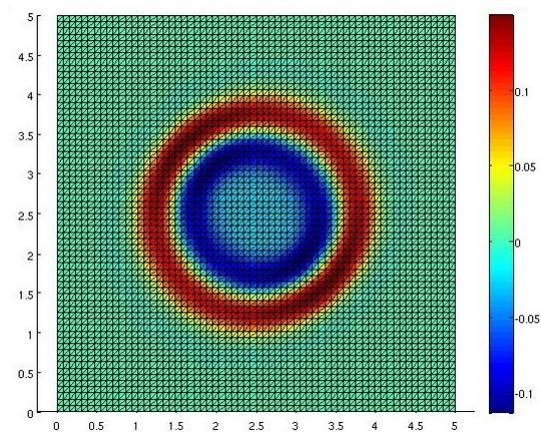


FIG. 3.12: Schéma Explicite. Volumes Finis Dérivées ORDRE 1. Code déstructuré

### Problème de scatterer - Maillage déstructuré

Soit un obstacle parfait  $\Gamma \subset \mathbb{R}^2$ . Soit  $[0, T]$  l'intervalle de simulation. Soit  $\Omega \subset \mathbb{R}^2$  le domaine de vide étudié. Soit un maillage non structuré de  $\Omega$  composé de  $N$  triangles  $(\omega_k)_{k \in [1, N]}$ . Soit l'Equation aux Dérivées Partielles :

$$\forall (\mathbf{x}, t) \in \Omega \times [0, T], \frac{\partial}{\partial t} u(\mathbf{x}, t) + \begin{pmatrix} 0 & 0 & \frac{\partial}{\partial x} \\ 0 & 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & 0 \end{pmatrix} u(\mathbf{x}, t) = f(\mathbf{x}, t).$$

Nous allons étudier la propagation d'une onde plane incidente  $u_{inc}$  réfléchiée par un obstacle parfait  $\Gamma$ . Formulation du problème :

$$\iint_{\omega_k} (\partial_t u + \mathbf{A}_i \partial_i u) + \int_{\partial \omega_k} \mathbf{A}_i \mathbf{n}_i^- (u_{ext}^+ - u) = \iint_{\omega_k} f = \int_{\partial \omega_k \cap \partial \Gamma} \mathbf{M}_1(u + u_{inc}), \quad (3.31)$$

$$u_{inc}(\mathbf{x}, t) = u_0 \mathbf{sin}(\omega(t - \mathbf{k} \cdot \mathbf{x} + \phi) \mathbb{1}_{t \leq t_\phi}) = (u_0, v_0, \rho_0)^T \mathbf{sin}(\omega(t - \mathbf{k} \cdot \mathbf{x} + \phi) \mathbb{1}_{t \leq t_\phi}).$$

Comme  $u_{(0)}$  est solution de l'équation des ondes dans le vide, on en déduit que  $u_0 = (k_x, k_y, 1)^T$ , où  $\mathbf{k} = (k_x, k_y)^T$  est le vecteur directeur de l'onde incidente. Par ailleurs, on choisit les grandeurs caractéristiques du maillage de façon à obtenir une longueur d'onde pour la largeur de ce dernier, donc  $\lambda = 1$ , ce qui implique une fréquence  $\omega = 2\pi$ . Pour la phase  $\phi$ , on souhaite  $\forall \mathbf{x} \in \Omega$ ,  $u_{inc}(\mathbf{x}, 0) = 0$ . Soit  $\mathbf{x}_i = (x_i, y_i)^T$  le premier point de l'objet  $\Gamma$  touché. Alors on aura  $\phi = 2\pi \mathbf{k} \cdot \mathbf{x}_i$ , d'où l'onde incidente

$$u_{inc}(\mathbf{x}, t) = u_0 \mathbf{sin} (2\pi(t - \mathbf{k} \cdot \mathbf{x} + \phi) \mathbb{1}_{t \leq \mathbf{k} \cdot (\mathbf{x} - \mathbf{x}_i)}).$$

Pour les conditions aux limites, on choisit sur  $\partial \Omega$  des conditions aux limites transparentes, et sur  $\partial \Gamma$  des conditions réfléchissantes, soit donc

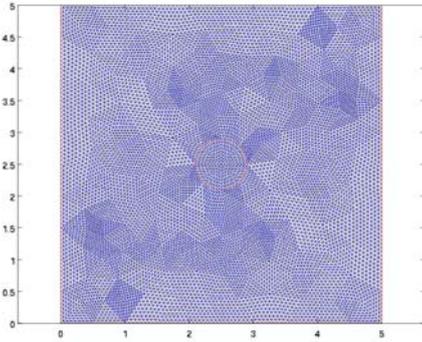
$$\begin{cases} \forall (\mathbf{x}, t) \in \partial \Gamma \times [0, T], & f(\mathbf{x}, t) = \int_{\partial \omega_k \cap \partial \Gamma} \mathbf{M}_1(u(\mathbf{x}, t) + u_{inc}(\mathbf{x}, t)), \\ \forall (\mathbf{x}, t) \in \partial \Omega \times [0, T], & \mathbf{A}_i \mathbf{n}_i^- (u^+(\mathbf{x}, t)) = 0. \end{cases}$$

Pour cette approximation, il faut être prudent sur les conditions aux limites. Elles doivent pouvoir être dérivées pour s'appliquer sur les équations des dérivées partielles. Ce type de condition « obstacle parfait » convient, et donne les conditions sur les dérivées en dérivant l'onde incidente. Par exemple, pour la méthode Volumes Finis aux Dérivées d'ordre 1, la solution approchée est de la forme  $u_T^{(0,0)}(t) + u_T^{(1,0)}(t)x + u_T^{(0,1)}(t)y$ . Il faut rajouter aux équations de la propriété (3.31) les seconds membres

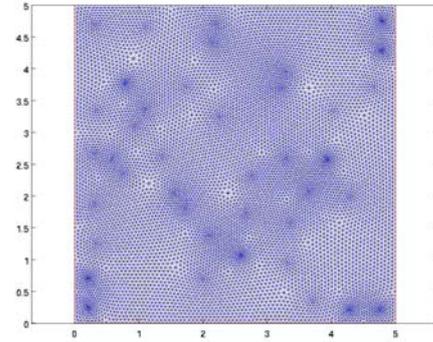
$$\left\{ \begin{array}{l} \int_{\partial\omega_k \cap \partial\Gamma} \mathbf{M}_1 \left( u_T^{(0,0)}(t) + u_{inc}^{(0,0)}(t)_T \right) = 0 \\ \int_{\partial\omega_k \cap \partial\Gamma} \mathbf{M}_1 \left( u_T^{(1,0)}(t) + u_{inc}^{(1,0)}(t)_T \right) = 0 \\ \int_{\partial\omega_k \cap \partial\Gamma} \mathbf{M}_1 \left( u_T^{(0,1)}(t) + u_{inc}^{(0,1)}(t)_T \right) = 0, \end{array} \right.$$

où  $u_{inc}^{(0,0)}(t)_T = u_{inc}(\mathbf{x}, t)$  calculé au centre de gravité du triangle  $T$ , de même que  $u_{inc}^{(1,0)}(t)_T = \frac{\partial u_{inc}(\mathbf{x}, t)}{\partial x} \Big|_T$  et  $u_{inc}^{(0,1)}(t)_T = \frac{\partial u_{inc}(\mathbf{x}, t)}{\partial y} \Big|_T$ .

Le problème de scattering considéré ici sera la réflexion par un disque placé au centre du maillage (figure (a)) d'une onde incidente de direction  $\mathbf{k} = (1, 0)^T$ . Les maillages sont destructurés, et une approximation de type Euler Explicite est choisie. On constate que la méthode à l'ordre 1 compense la perte d'amplitude d'une méthode de type Volumes Finis classiques. Pour le problème de scattering, l'amplitude maximale de l'onde dépasse les 0.6 alors qu'avec la méthode de type Volumes Finis elle est de l'ordre de 0.4, soit un gain de presque 50%.



(a) Maillage Problème de Scattering



(b) Maillage Problème en Condition Initiale

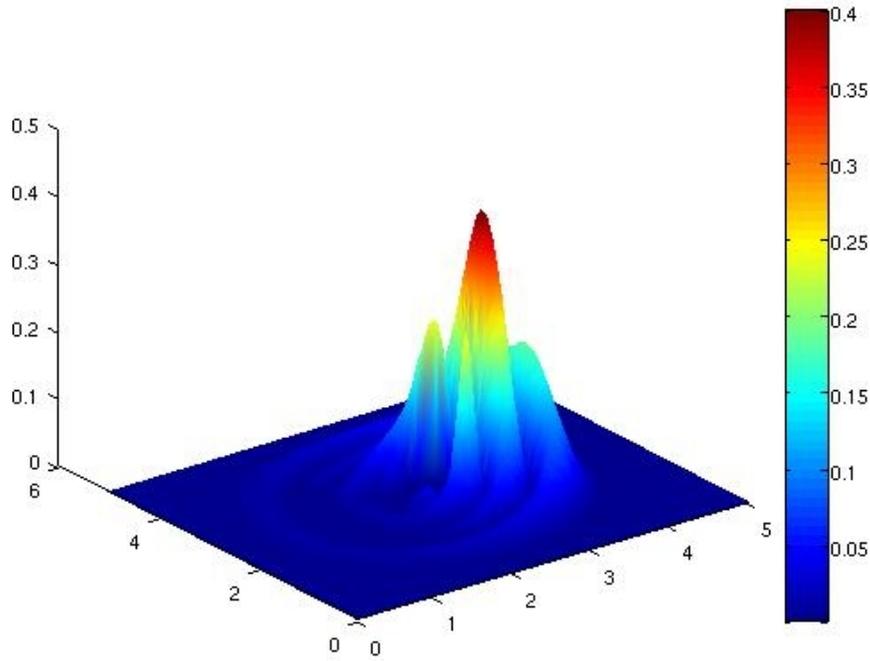


FIG. 3.13: Schéma Explicite. Volumes Finis. Problème de Scattering

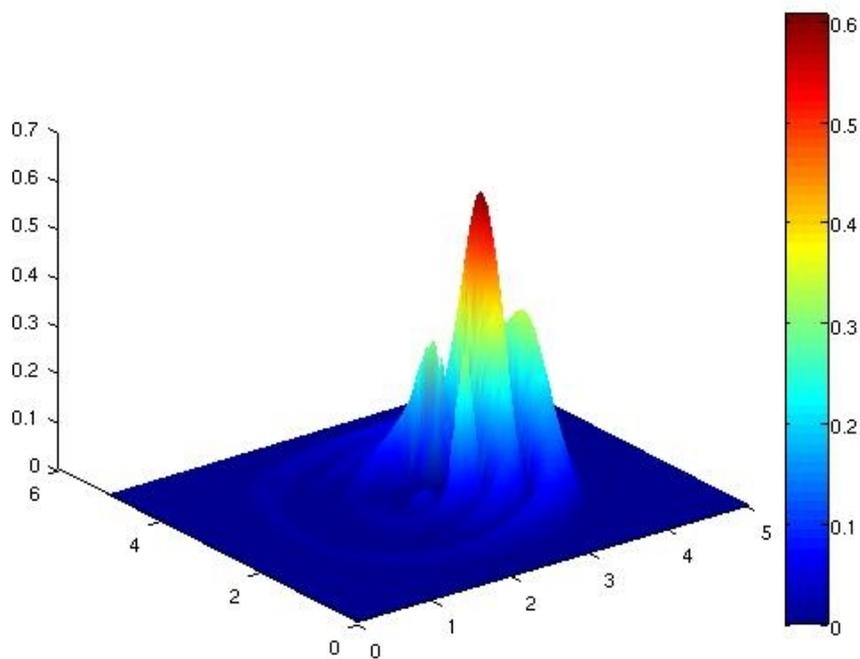


FIG. 3.14: Schéma Explicite. Volumes Finis Dérivées ORDRE 1. Problème de Scattering

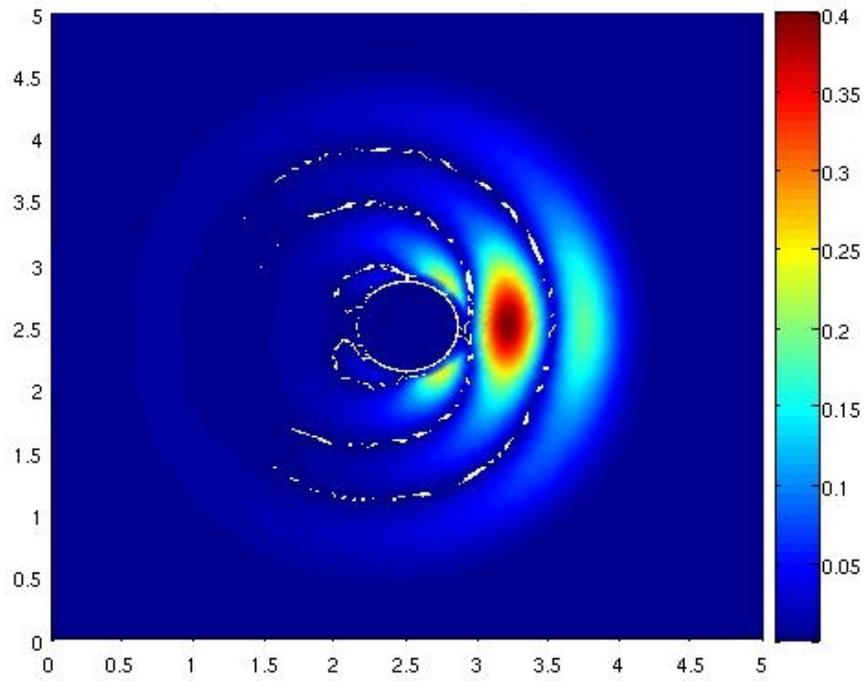


FIG. 3.15: Couple horizontale de (3.13)

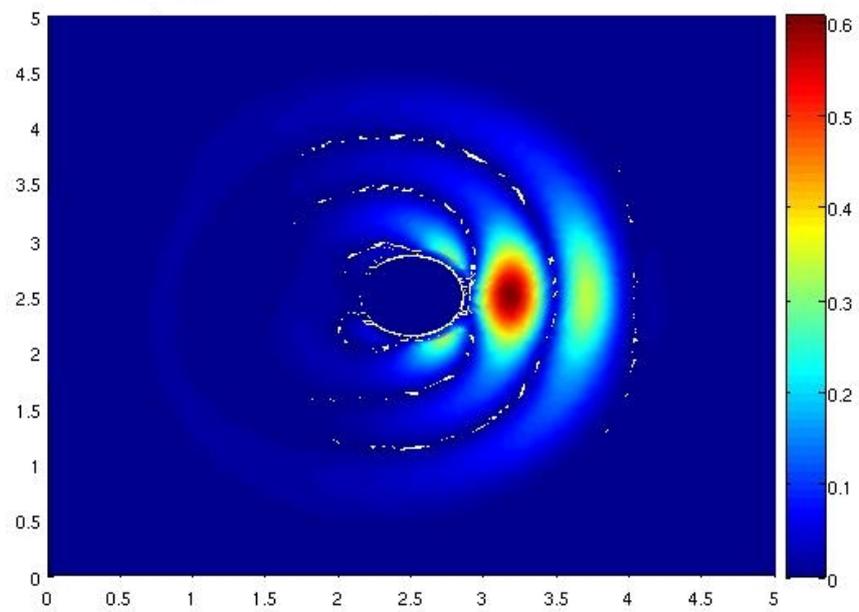


FIG. 3.16: Couple horizontale de (3.14)

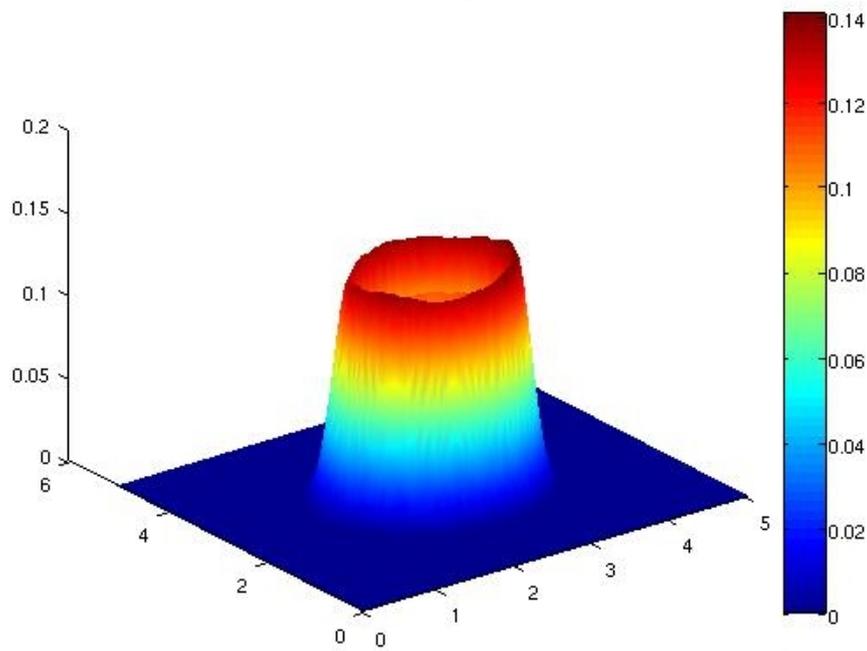


FIG. 3.17: Schéma Explicite. Volumes Finis. Problème en espace libre

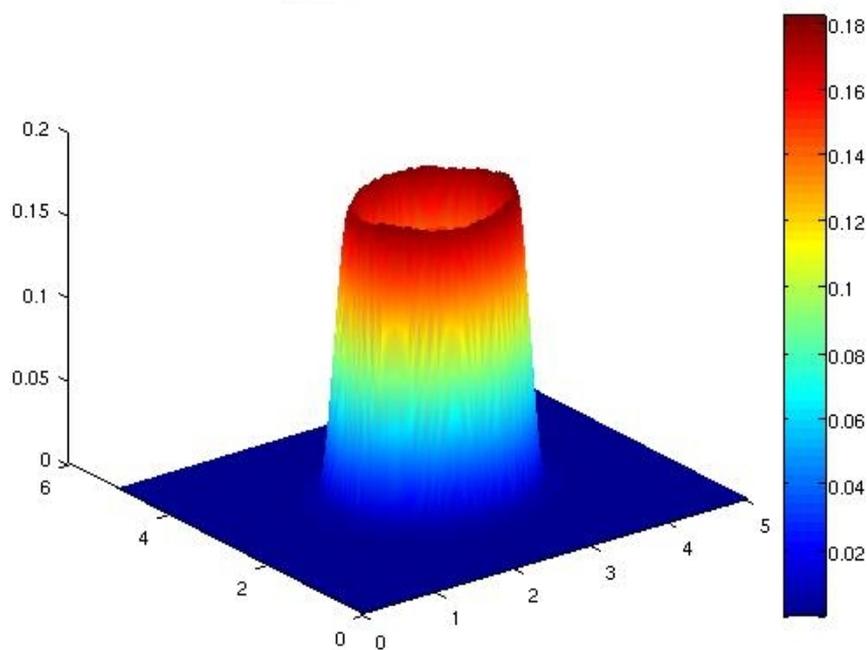


FIG. 3.18: Schéma Explicite. Volumes Finis Dérivées ORDRE 1. Problème en espace libre

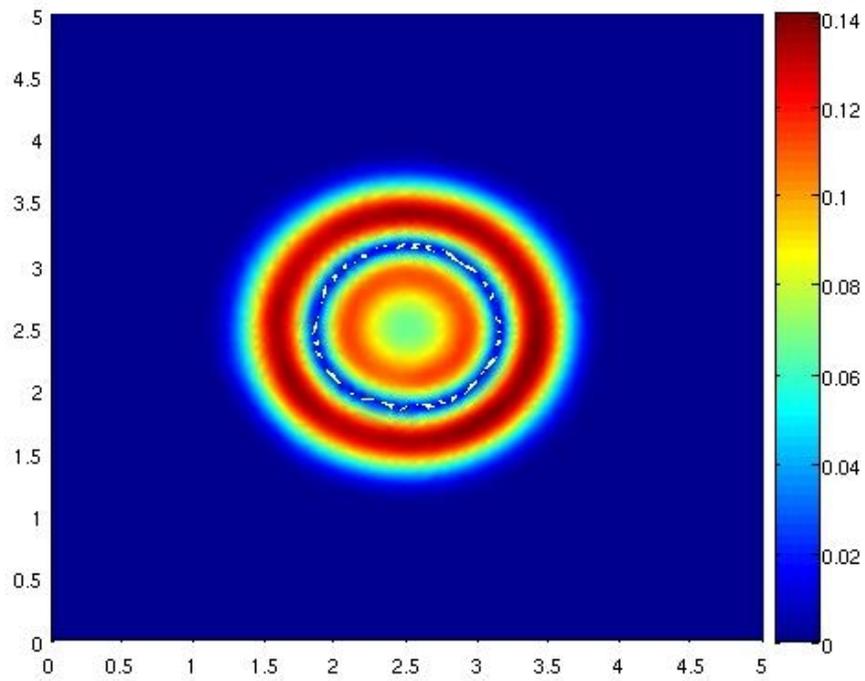


FIG. 3.19: Couple horizontale de (3.17)

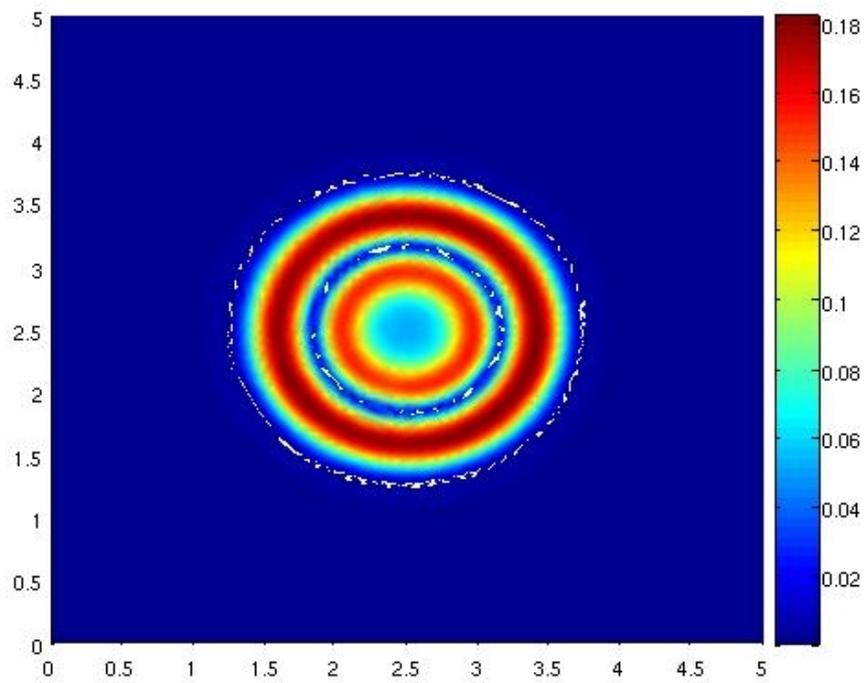


FIG. 3.20: Couple horizontale de (3.18)



# Conclusions et Perspectives

## Principaux résultats obtenus sur les milieux fictifs absorbants

Si l'on examine théoriquement les formulations fréquentielles citées ci-dessus elles consistent simplement à relever les équations écrites dans  $\mathbb{R}^n$  dans une variété totalement réelle à  $N$  dimensions plongée isométriquement dans la variété  $\mathbb{C}^n$  munie de la structure presque riemannienne (ou pseudo euclidienne) canonique de façon à ce que cette variété coïncide avec le domaine effectif de calcul en tant qu'ouvert de  $\mathbb{R}^n$  muni de la structure euclidienne habituelle. On peut donc réécrire les systèmes de Friedrichs de propagation d'ondes si ceux-ci admettent une formulation intrinsèque en tant que dérivation d'algèbre dans les algèbres graduées des fibrés cotangents de produits extérieurs de forme et tenter par cette voie d'obtenir un nouveau système dont on saurait exprimer le noyau de Schwartz par des opérations analogues aux algèbres de convolution conduisant à l'existence de l'unicité de la solution et à une décroissance au moins exponentielle des perturbations (y compris celle consistant à borner artificiellement le domaine sur cette variété par des conditions aux limites homogènes adéquates). Cette démarche initiée par Lassas [20] l'avait conduit à considérer des prolongements amenant à des problèmes bien posés, dans le domaine fréquentiel et autour d'un obstacle parfait, tant en domaine borné que non borné. En reprenant et complétant ce travail, nous avons établi les mêmes résultats en omettant une propriété de la variété exigée par les travaux de Lassas (asymptotiquement euclidienne) au prix du fait, qu'en domaine borné, l'existence et l'unicité de la solution en fréquentiel n'est obtenue qu'à l'exception d'un ensemble localement fini de fréquences près. Notre attention s'est bien entendu portée sur la possibilité de réduire le domaine effectif de calcul même si l'obstacle n'est pas convexe. Il est à noter qu'une telle éventualité peut sembler surprenante a priori (considération du type propagation d'ondes planes, conditions nécessaires à la décomposition de Nirenberg, ...). Nous avons procédé par conditions suffisantes et presque « nécessaires » de façon à obtenir la généralisation la plus large possible. On retrouve pour les équations de Maxwell et les ondes pluridimensionnelles l'incidence du caractère convexe de l'obstacle mais à un difféomorphisme de  $\mathbb{R}^3$  près (et un difféomorphe de convexe peut être loin d'être convexe!). Bien que, par ce biais, nous puissions placer la complexification des dérivées, et donc la décroissance du noyau de Schwartz, aussi près que l'on veut de l'obstacle, que nous puissions réaliser un relèvement temporel dispersif conduisant à un système de Friedrichs bien posé avec convergence des approximations usuelles, des inconvénients demeurent. En effet l'introduction de difféomorphismes conduit, dans le milieu dispersif équivalent, à des vitesses de

propagation qui dans la partie non convexe du domaine tendent vers l'infini quand la complexification des dérivées se rapproche du bord de l'obstacle...Il va sans dire que lors des méthodes explicites d'intégration temporelle ce que l'on gagne en complexité par réduction du domaine de calcul effectif on le perd par décroissance du CFL. On peut toutefois envisager dans « les parties non convexes » une résolution temporelle implicite (les matrices associées étant de relative petite taille et à bande étroite). Quant au problème harmonique, on peut penser que la modification du système bicaractéristique du système temporel associé exige, pour conserver la précision, des éléments plus petits. (Toutefois les tests numériques réalisés autour de domaines en forme de « L » conduisent à de bon résultats). Si cela est le cas, il faudrait associer ce type de PML avec un raffinement de maillage de type « goal oriented ». De toute façon, on pourrait considérer un calcul avec de tels milieux fictifs comme un préconditionneur efficace à la résolution du problème couplé EDP équations intégrales pour un objet diffractant hétérogène de forme non canonique et non convexe. En résumé pour les équations admettant une « formulation intrinsèque » (en dehors de la partie principale de l'opérateur qui elle l'admet toujours) nous avons exploré les conditions les plus générales pour lesquelles l'écriture de milieux fictifs absorbants est possible. On sait que les équations d'Euler linéarisées, à coefficient constant, n'admettent pas, en présence de convection de l'écoulement porteur, de formulation dispersive. Une intéressante question ouverte (qui pourrait concerner l'élastodynamique) est : quels sont les systèmes de Friedrichs pour lesquels une telle formulation existe ? Ce que suggère nos travaux est la recherche du prolongement du domaine effectif de calcul dans une variété l'incluant, le calcul différentiel sur un fibré au dessus de celle-ci structuré par un groupe de Lie autre que le groupe linéaire de  $\mathbb{C}^n$  ou une réduction de celui-ci...

## Principaux résultats obtenus sur la nouvelle approximation proposée

Remarquons tout d'abord que l'ordre élevé des approximations n'est théoriquement opérant que si la solution est suffisamment régulière au moins par morceaux cylindriques en temps. Un théorème dû à J. Rauch [32] décrit l'obtention d'une telle régularité en rapport avec la régularité des coefficients, du domaine et des seconds membres. La régularité nécessaire aux estimations à priori des approximations de type Galerkin Discontinu par des polynômes de degré  $K$  conduit au caractère bien posé des systèmes de Friedrichs ayant pour inconnues les dérivées partielles jusqu'à l'ordre  $K$  de la solution du système originel. Une idée naturelle est alors de considérer les systèmes d'équation obtenus sur l'inconnu et toutes ses dérivées jusqu'à l'ordre  $K$  en les intégrant, sur une famille totale « d'éléments », avec un fluxsplitting adéquat sur leurs frontières. En considérant sur ces éléments une interpolation de la solution par des polynômes de degré  $K$  exprimés dans leur base naturelle (autour du centre de gravité) on obtient formellement un système d'EDO, avec autant d'équations que d'inconnues, dont la matrice de masse, évidemment triangulaire et très creuse, est aisément inversible. Quand  $K = 0$  on obtient évidemment l'approximation Volumes

Finis. La stabilité  $\mathcal{L}^2$  d'une approximation explicite du système d'EDO ainsi obtenu est indépendante de  $K$  et la même que quand  $K$  est égal à zéro ! Cela ne permet évidemment pas de conclure quant à la convergence et à fortiori l'ordre de convergence du schéma. Toutefois, dans le cas où le maillage est structuré cartésien, nous avons montré par récurrence que le schéma Différence Finie compact ainsi obtenu était consistant et donc, par le théorème d'équivalence de Lax, convergeant. Ceci n'indique pas l'ordre de convergence mais la consistance est de degré  $K + 1$ . Des expériences numériques en approximations déstructurées pour l'équation des ondes bidimensionnelles conduisent cependant à des résultats satisfaisants. Pour de telles approximations, la consistance ne peut évidemment pas être étudiée par le biais de la formule de Taylor. Diverses récentes généralisations du théorème d'équivalence de Lax pourraient amener à une démonstration de la convergence mais, ne disposant pas des habituelles propriétés de coercivité et d'orthogonalité de l'erreur le problème d'obtention d'estimations à priori reste largement ouvert.



# Bibliographie

- [1] M. Ainsworth. Discrete Dispersion Relation for hp-Version Finite Element Approximation at High Wave Number. *SIAM Journal on Numerical Analysis*, 42(2) :553–575, 2004.
- [2] J. P. Aubin. Approximation des espaces de distributions et des opérateurs différentiels. *Mémoires de la S.M.F*, 12 :3–139, 1967.
- [3] J.P. Berenger. A Perfectly Matched Layer for the Absorption of Electromagnetic Waves. *Journal of Computational Physics*, 114 :185–200, 1994.
- [4] W.C. Chew and W.H. Weedon. A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates. *Microwave and Optical Technology Letters*, 7(13), 1994.
- [5] F. Collino and P. Monk. The Perfectly Matched Layer in Curvilinear Coordinates. *SIAM Journal on Scientific Computing*, 19(6) :2061–2090, 1998.
- [6] D.L. Colton and R. Kress. *Integral Equation Methods in Scattering Theory*. John Wiley & Sons Inc, 1983.
- [7] Y. Coudiere, J.P. Vila, and P. Villedieu. Convergence d'un schéma volumes finis explicite en temps pour les systèmes hyperboliques linéaires symétriques en domaines bornés. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, 331(1) :95–100, 2000.
- [8] J.P. Croisille. *Contribution à l'étude théorique et à l'approximation par éléments finis du système hyperbolique de la dynamique des gaz multidimensionnelle et multiespèces*. PhD thesis, Paris 6, 1990.
- [9] J.P. Croisille. Résolution des équations d'Euler par une méthode de type Galerkin-discontinu. Application à l'écoulement autour d'une aube bidimensionnelle. *Technical Report ONERA*, 1990.
- [10] B. Dah. *Sur la modélisation de milieux fictifs absorbants de type couches de Bérenger*. PhD thesis, Université Paul Sabatier, 2001.
- [11] R. Dautray and J.L. Lions. *Analyse mathématique et calcul numérique pour les sciences et les techniques*, volume 5. Masson, Paris, 1985.

- [12] G. de Rham. Variétés différentiables. Formes, courants, formes harmoniques. *Actualites Sci. Ind.*, 1222, 1955.
- [13] S.K. Godunov. *Lois de conservation et intégrales d'énergie des équations hyperboliques*, volume 1270. Springer Lecture Notes in Mathematics, 1987.
- [14] S. Gottlieb, C.W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM review*, 43 :89–112, 2001.
- [15] P. Helluy. *Résolution numérique des équations de Maxwell harmoniques par une méthode d'éléments finis discontinus*. PhD thesis, Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, 1994.
- [16] L. Hormander. *Linear partial differential operators*, volume 2. Berlin, 1969.
- [17] K. Kasahara and MY Yamaguti. Strongly hyperbolic systems of partial diff. equations with constant coefficients. *Mem. Coll. Sci. Univ. Kyoto*, 33 :1–23, 1960.
- [18] T. Kato. *Perturbation theory for linear operators*. Springer, 1966.
- [19] V.K. Khoan. *Distributions, analyse de Fourier, opérateurs aux dérivées partielles*, volume 2. Vuibert, 1972.
- [20] M. Lassas, J. Liukkonen, and E. Somersalo. Complex Riemannian metric and absorbing boundary conditions. *Journal de mathématiques pures et appliquées*, 80(7) :739–768, 2001.
- [21] M. Lassas and E. Somersalo. On the existence and convergence of the solution of PML equations. *Computing*, 60(3) :229–241, 1998.
- [22] S. Laurens. Présentation générale de la méthode PML, géométrie différentielle associée. *Technical Report ONERA*, pages 37–70, Décembre 2006.
- [23] S. Laurens. Milieux Fictifs parfaitement absorbants : une méthode de résolution des équations de Maxwell en instationnaires pour des domaines non convexes. *Technical Report ONERA*, pages 78–99, Décembre 2007.
- [24] S. Laurens. Un nouveau schéma de haute précision, les Volumes Finis aux Dérivées ; Conditions PML non convexes. *Technical Report ONERA*, pages 81–99, Décembre 2008.
- [25] J. Liukkonen. Generalized electromagnetic scattering in a complex geometry. *Journal of Mathematical Analysis and Applications*, 254(2) :498–514, 2001.
- [26] A. J. Majda. Coercive Inequalities for Nonelliptic Symmetric Systems. *Communications on pure and applied Mathematics*, XXVIII :49–89, 1975.

- 
- [27] P.A. Mazet, L. Segui, and B. Dah. Sur l'existence et l'unicité des solutions pour le système de Maxwell harmonique en présence de couches de Bérenger. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, 333(6) :599–604, 2001.
- [28] V. Mouysset. *Résolution des équations de Maxwell instationnaires*. PhD thesis, Université Paul Sabatier, 2006.
- [29] J.T. Oden, S. Prudhomme, and L. Demkowicz. A posteriori error estimation for acoustic wave propagation problems. *Archives of Computational Methods in Engineering*, 12(4) :343–389, 2005.
- [30] V. Petkov. *Scattering theory for hyperbolic operators*, volume 21. North Holland, 1989.
- [31] A. Rahmouni. *Des modèles PML bien posés pour divers problèmes hyperboliques*. PhD thesis, Université Paris XIII, 2000.
- [32] J. Rauch. Symmetric positive systems with boundary characteristic of constant multiplicity. *Transactions of the American Mathematical Society*, 291(1) :167–187, September 1985.
- [33] R.D. Richtmyer. *Difference methods for initial value problems*. Interscience, 1957.
- [34] L. Schwartz. *Théorie des Distributions*. Hermann, 1966.
- [35] D. Serre. *Systèmes de lois de conservation*, volume 1. Diderot Editeur, 1996.
- [36] F.L. Teixeira and W.C. Chew. General closed-form PML constitutive tensors to match arbitrary bianisotropic and dispersive linear media. *IEEE Microwave and Guided Wave Letters*, 8(6) :223–225, 1998.
- [37] F.L. Teixeira, K.P. Hwang, W.C. Chew, and J.M. Jin. Conformal PML-FDTD schemes for electromagnetic field simulations : A dynamic stability study. *IEEE Transactions on Antennas and Propagation*, 49(6) :902–907, 2001.
- [38] Y. Ventribout. *Contrôle des perturbations aéroacoustiques par impédances de parois : application à un modèle de matériaux poreux*. PhD thesis, Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, 2006.
- [39] C.H. Wilcox. *Scattering theory for the d'Alembert equation in exterior domains*, volume 442. Springer Verlag, 1975.