



**HAL**  
open science

# Analyse de propagation d'ondes en géométrie complexe, problèmes discrets et continus

Antoine Benoit

► **To cite this version:**

Antoine Benoit. Analyse de propagation d'ondes en géométrie complexe, problèmes discrets et continus. Equations aux dérivées partielles [math.AP]. Université du littoral Côte d'Opale - ULCO, 2024. tel-04689489

**HAL Id: tel-04689489**

**<https://hal.science/tel-04689489v1>**

Submitted on 5 Sep 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITE DU LITTORAL COTE D'OPALE



---

Mémoire de synthèse en vue de l'  
**Habilitation à Diriger des Recherches**

Discipline  
**Mathématiques,**  
Ecole doctorale Sciences, Technologie, Santé (ED585).

---

**Analyse de propagation d'ondes en géométrie complexe, problèmes  
discrets et continus,**

par **Antoine BENOIT**

---

Soutenu publiquement le **24 juin 2024**, devant un jury constitué de :

- Mme **Sylvie BENZONI-GAVAGE**, Professeure des universités, Institut Camille Jordan, *examinatrice*
- M. **Jean-François COULOMBEL**, Directeur de recherches, Institut de Maths. de Toulouse, *examineur*
- M. **Jeffrey RAUCH**, Professeur des universités, University of Michigan, *examineur*
- M. **Miguel RODRIGUES**, Professeur des universités, Université de Rennes, *rapporteur*
- Mme **Carole ROSIER**, Professeure des universités, Université du Littoral Côte d'Opale, *examinatrice*.
- M. **Frédéric ROUSSET**, Professeur des universités, Université Paris-Saclay, *rapporteur*
- M. **Paolo SECCHI**, Professeur des universités, Università di Brescia, *rapporteur*



*Je ne sais pas ce qui est beau,  
mais je sais ce que j'aime  
et je trouve ça amplement suffisant,  
Boris Vian<sup>1</sup>,*

OINOS - *Mais, pendant cette dernière existence,  
j'avais rêvé que j'arriverais d'un seul coup  
à la connaissance de toutes choses,  
et du même coup au bonheur absolu.*  
AGATHOS - *Ah ce n'est pas dans la science qu'est le savoir,  
mais dans l'acquisition de la science !  
Savoir pour toujours, c'est l'éternelle béatitude ;  
mais, tout savoir, ce serait une damnation de démon.*  
Edgar Allan Poe<sup>2</sup>

---

<sup>1</sup>Référence inconnue

<sup>2</sup>*Nouvelles histoires extraordinaires*, 2018, Garnier Flammarion, p. 265, traduction : Charles Baudelaire

# Contents

<b>I</b>	<b>Partie introductive</b>	<b>6</b>
I.1	Introduction . . . . .	6
I.2	Organisation du mémoire . . . . .	11
I.3	Table des notations . . . . .	13
I.3.1	Notations générales . . . . .	13
I.3.2	Notations pour les problèmes continus . . . . .	14
I.3.3	Notations pour les problèmes discrets . . . . .	15
<b>II</b>	<b>Problèmes aux limites hyperboliques continus</b>	<b>16</b>
II.1	Motivations . . . . .	16
II.2	Le problème dans le demi-espace . . . . .	18
II.2.1	Le cadre idéal, fortement bien posé, la condition de Kreiss-Lopatinskii uniforme . . . . .	19
II.2.2	<i>Troubles begin</i> , le cadre faiblement bien posé . . . . .	24
II.2.3	Problèmes faiblement bien posés et développements d’optique géométrique . . . . .	29
	Une motivation des développements d’optique géométrique . . . . .	29
	Une application des développements d’optique géométrique, l’optimalité des pertes de dérivées . . . . .	34
II.3	Le problème dans la bande . . . . .	36
II.3.1	Théorie bien posée . . . . .	38
	La condition de Kreiss-Lopatinskii uniforme n’est plus assez fine . . . . .	38
	Une nouvelle condition liée à l’auto-interaction . . . . .	40
	Caractère suffisant de cette condition . . . . .	42
II.3.2	Illustration de cette condition <i>via</i> le d’optique géométrique . . . . .	45
II.3.3	Vers une nouvelle zoologie de problèmes faiblement bien posés . . . . .	48
II.3.4	De possibles extensions de la théorie fortement bien posée optimale dans la bande ? . . . . .	51
II.4	Le problème dans le quart d’espace . . . . .	53
II.4.1	Une théorie bien posée encore très en friche . . . . .	54
	(Bref) état de l’art . . . . .	54
	Le <i>symétriseur</i> d’Osher . . . . .	57
II.4.2	Un résultat de persistance de la régularité . . . . .	62
	Principale difficulté et analogie caractéristique . . . . .	62
	Description de la méthode dans le demi-espace . . . . .	63
	Le résultat de persistance . . . . .	65
	Ce qui reste à entreprendre . . . . .	69
II.4.3	Développements d’optique géométrique pour le quart d’espace . . . . .	70
	Heuristique . . . . .	70
	Ce que nous apprend le développement d’optique géométrique . . . . .	76
	Un dernier exemple étonnant de problème faiblement bien posé . . . . .	78
II.4.4	Vers des développements d’optique géométriques en géométrie bornée . . . . .	80

Développements d’optique géométrique dans un borné : le cas du triangle . . . . .	80
Retour sur les développements dans le cas du quart d’espace . . . . .	84
Les principales difficultés à surmonter et les potentielles extensions . . . . .	88
II.4.5 Quelques questions ouvertes . . . . .	89
Faire de l’analyse d’Osher une théorie aboutie . . . . .	89
Extension aux problèmes non linéaires . . . . .	91
L’approximation visqueuse et l’interaction des couches limites . . . . .	93
<b>III Schémas numériques de type différences finies</b>	<b>95</b>
III.1 Introduction et description des schémas . . . . .	95
III.2 Motivations . . . . .	98
III.3 Le cas de la demi-droite . . . . .	100
III.3.1 Description des conditions de bord numériques . . . . .	101
III.3.2 Définitions et commentaires sur la stabilité . . . . .	102
III.3.3 La théorie GKS en deux mots . . . . .	103
Les différentes réductions et la condition de Kreiss-Lopatinskii discrète . . . . .	103
III.3.4 Le ”graal” de la stabilité . . . . .	106
III.4 Un premier pas vers un domaine de résolution borné, le problème dans le segment . . . . .	107
III.4.1 Motivations . . . . .	107
III.4.2 Une discrétisation de l’analyse d’Osher . . . . .	107
Les modifications à apporter dans le cas de l’intervalle . . . . .	108
Le problème reformulé et la condition de stabilité . . . . .	109
Un <i>symétriseur</i> de type Osher discret . . . . .	112
Quelques mots sur la construction d’un <i>symétriseur</i> discret . . . . .	113
III.5 Vers un domaine de discrétisation borné, première étape le quart d’espace . . . . .	118
III.5.1 Stabilité du semi-groupe pour les conditions de Dirichlet . . . . .	118
III.5.2 Le schéma de Lax-Wendroff dans le quart d’espace . . . . .	120
La démonstration dans $\mathbb{Z}^2$ . . . . .	123
Les modifications dans le cas du quart d’espace avec Neumann d’ordre un . . . . .	124
Les modifications avec Neumann d’ordre deux . . . . .	125
III.5.3 Une discrétisation de l’analyse d’Osher ? . . . . .	126
III.6 Quelques autres questions ouvertes . . . . .	128
III.6.1 Établir un résultat probant de stabilité du problème dans le demi-espace . . . . .	128
III.6.2 Améliorer les conditions d’applicabilité du résultat de [Benoit, 2022] . . . . .	128
III.6.3 Une méthode numérique de vérification de la condition de stabilité dans l’intervalle à le Barbenchon ? . . . . .	130
<b>IV Autres résultats notables</b>	<b>132</b>
IV.1 Homogénéisation en temps long de l’équation des ondes . . . . .	132
IV.2 Contrôlabilité . . . . .	137
IV.2.1 Un résultat sur des équations paraboliques (dégénérées) . . . . .	137
IV.2.2 Contrôle pour l’équation des ondes . . . . .	140
IV.3 Un dernier exemple de développement d’optique géométrique pour conclure . . . . .	143
IV.3.1 Première approximation ”naturelle” . . . . .	145
IV.3.2 Quand continuer sur une bonne lancée conduit à l’accident . . . . .	146
IV.3.3 Apparition des correcteurs non caractéristiques . . . . .	148
<b>V Conclusion</b>	<b>150</b>

# Remerciements

En premier lieu, je tiens à remercier sincèrement l'ensemble des membres du jury d'avoir accepté de lire et d'évaluer ce travail d'habilitation. Mes remerciements les plus chaleureux ainsi que l'expression de ma profonde gratitude vont à Miguel Rodriguez, Frédéric Rousset et Paolo Secchi qui ont eu la lourde tâche de juger ce mémoire.

Merci à Miguel, d'avoir accepté de rapporter sur cette habilitation. J'espère de tout cœur que tu l'auras trouvé intéressant et que ce travail donnera lieu à de futures collaborations fécondes (de préférence au sein de la future ANR HEAD).

C'est un véritable plaisir d'avoir Frédéric, un de mes anciens professeurs, au sein de mon jury d'HDR. Ses cours de Master m'ont vraiment inspirés et ont très certainement contribué à faire ce que je suis actuellement (en provoquant au passage quelques crampes au poignet (et au cerveau)). Si je ne devais retenir qu'une chose de ces cours ça serait la phrase "faites les calculs comme vous l'entendez, puis essayez de leur rendre rigoureux dans un second temps".

Enfin je remercie sincèrement Paolo d'avoir accepté d'être rapporteur de cette habilitation. J'espère que le manuscrit en mauvais anglais était quand même lisible et que tout le travail engagé n'a pas été trop désagréable. J'espère que nous aurons à l'avenir l'occasion d'échanger sur les problèmes aux limites hyperboliques. *Grazie Mille !*

Je remercie ensuite Sylvie Benzoni-Gavage, Jean-François Coulombel, Jeffrey Rauch et Carole Rosier d'avoir accepté d'être examinateurs de ce travail postdoctoral.

J'aimerais commencer par remercier la directrice du LMPA, Carole. Plus qu'une simple participation à un jury, elle m'a permis depuis mon recrutement à l'ULCO de bénéficier des meilleures conditions de travail possibles. Si ce travail est soutenu aujourd'hui c'est en grande partie grâce à ces conditions de travail optimales pour lesquelles je lui suis profondément reconnaissant.

Les travaux de Jeffrey ont été (et je pense seront) toujours une grande source d'inspiration pour moi. C'est un très grand honneur de pouvoir le compter aujourd'hui parmi les membres de ce jury. Je crois que je serai plus que satisfait de ma carrière si jamais j'arrivais à obtenir le tiers du quart des résultats qu'il a obtenu durant sa carrière.

Je remercie Sylvie d'avoir joué le rôle de garante mais surtout de son intérêt concernant mes recherches qu'elle a manifesté depuis ma thèse. C'est un grand plaisir d'avoir ma "grand-mère" de thèse dans le jury, d'autant plus qu'elle est l'une des auteurs d'un livre qui peut être qualifié pour moi de livre de chevet.

Enfin, le dernier mais pas des moindres, mes plus chaleureux remerciements vont à mon ancien directeur de thèse, Jean-François. C'est un grand plaisir de pouvoir le compter dans le jury aujourd'hui. Lui aussi a grandement contribué à faire de moi le mathématicien que je suis à ce jour, notamment grâce à son encadrement doctoral sans faille, mais aussi, et cela est peut-être plus capital encore, grâce à la vision des mathématiques qu'il m'a transmis au fil de ces années. Je suis enchanté que nous collaborions ensemble et je suis sûr qu'à l'avenir nous aurons l'occasion de démontrer plein de beaux résultats d'analyse (numérique).

Je ne serais pas grand chose sans un bon cadre de travail. Je remercie donc chaleureusement mes collègues du LMPA qui tous les jours viennent partager leur bonne humeur et font que c'est un réel plaisir de venir travailler tous les jours.

Merci donc à mes collègues qui ont le bon goût de faire des maths compréhensibles Augustin, Christophe et Xavier. Merci à mes collègues qui font des maths que j'aimerais bien pouvoir faire mais que malheureusement je ne comprendrai jamais vraiment : Bruno, Christian, Isaar, Lucile et Romuald. Merci à Nicolas de faire des probabilités, je préfère que ça soit lui que moi ! Et enfin comme je ne suis pas sectaire, je remercie même mes algébristes préférés Jean, Loïc, Mélodie et Pierre-Louis. Là encore, je préfère que ça soit eux que moi ! Ou peut-être vaut il mieux que ça soit eux que moi ? Et comme je ne suis vraiment pas sectaire merci aux nombreux thésards qui viennent partager leurs enthousiasme et leur joie de vivre juvénile (Julie, Pierre, Vincent,...). Merci aussi à Gaëlle, Sandrine et Thierry avec qui c'est toujours un plaisir de partager un cours. Bref un grand merci à tous les membres du LMPA.

Enfin, un laboratoire de Maths n'est pas grand chose sans une bonne équipe administrative et technique et je tiens ici à remercier Isabelle, Gaëlle et Philippe pour tout ce qu'ils font pour nous tous les jours.

Je suis reconnaissant envers ma famille d'avoir contribué à faire de moi l'homme que je suis actuellement. On ne peut pas dire qu'ils soient d'une grande aide en maths, mais ils m'apportent beaucoup. Merci donc à ma mère, mon frère et ma sœur pour leur soutien. Merci infiniment à mon père qui a mis un point d'honneur à relire et à corriger les coquilles de ce manuscrit bien qu'il soit aux antipodes de son ancienne section CNU. Enfin, merci à mes neveux, Emile, Héloïse et Jules. Qui sait peut-être qu'un jour voudront ils suivre les traces de leur oncle et se lancer dans une carrière mathématique.

Enfin du fond du cœur je remercie Yoann. Heureusement que voici mon dernier mémoire universitaire car je commence à manquer d'inspiration pour te remercier de façon originale. J'aimerais donc juste te remercier d'être toi, d'être à mes côtés et pour tout ce que tu m'apportes tous les jours. Sans toi, je serais un peu comme un système hyperbolique sans symétriseur...



# Chapter I

## Partie introductive

### I.1 Introduction

On attribue à Kronecker la phrase : "Dieu créa les entiers naturels, le reste est l'œuvre de l'Homme". Plus je fais des mathématiques et plus je pense que cette phrase est fondée. Que malgré l'existence possible d'une quelconque vérité platonicienne sur laquelle reposent les Mathématiques, l'essentiel de ces dernières repose sur des constructions artistoteliciennes propres à l'Homme. Qu'à cela ne tienne ; ayant de toute façon bien plus de sympathie pour l'Homme que pour le Divin<sup>1</sup>, on se bornera dans la suite à ces constructions purement humaines, ces cathédrales de verre, sur lesquelles on espère pouvoir donner quelques éléments de réflexion. Sans transition, donc...

De très nombreuses équations de la physique s'écrivent sous la forme : pour  $d \geq 1$ ,

$$\partial_t u + \sum_{j=1}^d \partial_j f_j(u) = 0, \quad (\text{I.1})$$

où l'inconnue  $u \in \mathcal{U} \subset \mathbb{R}^N$  (un ensemble d'états physiquement admissibles) pour  $N \geq 1$ , et où les non-linéarités  $f_j : \mathcal{U} \rightarrow \mathbb{R}^N$ ,  $j \in \llbracket 1, d \rrbracket$ ,  $d$  étant la dimension de l'espace ambiant.

De tels systèmes, appelés systèmes de lois de conservation, ont la propriété que les différentes composantes du vecteur  $u$  sont conservées au fil du temps (comme on peut s'en convaincre en intégrant (I.1)). Ainsi, il est normal que les modèles physiques qui se basent sur la conservation de certaines quantités caractéristiques (la masse, la quantité de mouvement, l'énergie...) puissent s'écrire dans le formalisme de l'équation (I.1).

En particulier, les phénomènes physiques de nature ondulatoire, qu'il s'agisse d'ondes acoustiques (équation des ondes), lumineuses (équation de Maxwell), issues de la mécanique des fluides (équation d'Euler et ses nombreuses dérivées) peuvent s'écrire sous la forme (I.1) avec un choix idoine des coefficients  $f_j$ .

Il est à noter qu'une théorie assez systématique établissant l'existence de solutions pour des systèmes de lois de conservations posés en "grande" dimension n'est pas, à l'heure actuelle, disponible. Bien souvent, les résultats d'existence se font au cas par cas. On a essentiellement l'une des deux alternatives suivantes : ou bien on sait traiter le cas de systèmes mais unidimensionnels (on pense par exemple au schéma

---

<sup>1</sup>On se permet ici de reprendre les mots de Maxime Gorki : "Tout est dans l'homme, tout est pour l'homme ! L'homme seul existe, tout le reste est l'oeuvre de ses mains et de son cerveau. L'homme ! Quel mot magnifique ! Comme cela sonne fier ! Il faut respecter l'homme ! Pas le plaindre, pas l'humilier par la pitié, mais le respecter. Buons à l'homme, baron !" Maxime Gorki, Les Bas-fonds, ed l'Arche 2009, p 115

de Glimm [Glimm, 1965]) ou bien on sait traiter le cas scalaire multidimensionnel (voir par exemple [Kruzhkov, 1970]).

A la question de l'existence vient de plus s'ajouter la question de l'unicité de la solution. Cette dernière est loin d'être immédiate. En effet, par exemple, pour l'équation de Burgers :

$$\begin{cases} \partial_t u + u \partial_x u = 0 & \text{dans } \mathbb{R}_+ \times \mathbb{R}, \\ u|_{t=0}(x) = -\tanh x, \end{cases} \quad (\text{I.2})$$

il est classique que deux types de solutions coexistent, les ondes de chocs et les ondes de détente. Ainsi, on aimerait disposer d'un critère (de préférence physique) venant sélectionner une solution plutôt qu'une autre.

Dans le cas scalaire unidimensionnel, le fameux théorème de Kruzhkov [Kruzhkov, 1969] répond à ce problème. On a alors existence et unicité de la solution dite entropique, basée sur la vérification de relations d'entropie/flux d'entropie supplémentaires. On peut alors montrer que la solution "privilegiée" de (I.2) est celle qui présente un choc, c'est-à-dire, une discontinuité de la solution elle-même.

En particulier, on voit apparaître sur cet exemple simple une autre caractéristique propre des équations hyperboliques, qui les différencie drastiquement et fondamentalement de leurs cousines elliptiques ou paraboliques : l'équation peut venir générer en temps fini des singularités qui rendent sa solution moins régulière que les données du problème. Cela a pour conséquence que pour étudier une équation hyperbolique, il faut partir du principe qu'aucune régularisation par l'équation n'est à espérer. La solution sera (au mieux) aussi régulière que les données du problème. On ne pourra pas, ici, pouvoir bénéficier des gains de régularité qui ont lieu pour les équations elliptiques ou paraboliques.

Notons, ici, une autre propriété phare des équations hyperboliques : la vitesse de propagation à vitesse finie de l'information. Cette dernière assure que si la donnée initiale de (I.1) est à support compact, la solution de l'équation hérite alors de cette propriété. Comme on le verra dans la suite de ce mémoire, cette propriété est fort utile, car elle permet, bien souvent, de traquer la solution et d'avoir une bonne idée de l'endroit où se situe l'essentiel de l'information au fil du temps.

Toutefois, le revers de la médaille est alors que l'on ne peut pas espérer de "fuite" de masse à l'infini, ce qui peut, parfois, se révéler bien utile, notamment dans l'étude du caractère bien posé du problème, *via* l'établissement d'estimations d'énergie.

Pour une étude assez complète des lois de conservation scalaires ou bien des systèmes, on pourra, par exemple, citer les ouvrages de référence [Serre, 1996a]-[Serre, 1996b]. On donnera plus de détails dans l'introduction du Chapitre II concernant la construction de solutions d'un système de la forme (I.1) présentant une ou plusieurs interfaces de singularité.

On note toutefois que comme on l'a mentionné une étude caractérisation complète, comprenant un critère de sélection, des solutions d'un système de loi de conservation de la forme (I.1) est une question toujours active et largement ouverte.

Les résultats présentés ici ne sont pas aussi ambitieux et ils se borneront à un cas bien particulier de systèmes de lois de conservations, le cas le plus simple, à savoir, le cas linéaire. Cette simplification, très conséquente, n'est ni anodine, ni découplée de l'équation (I.1). En effet, avant de s'attaquer à un problème non linéaire il est capital de bien comprendre la dynamique linéaire qui sous-tend.

Une méthode classique est alors de traiter le cas non linéaire en se ramenant au cas linéaire par une méthode de point fixe. Cette méthode est rendue possible, si seulement on dispose d'une assez bonne théorie qualitative du problème linéarisé, notamment de sa théorie de persistance de la régularité.

L'équation d'évolution que l'on va considérer dans la suite<sup>2</sup> est alors donnée par

$$\begin{cases} L(\partial)u := \partial_t u + \sum_{j=1}^d A_j \partial_j u = 0 & \text{pour } (t, x) \in ]0, \infty[ \times \mathbb{R}^d, \\ u|_{t=0} = u_0 & \text{pour } x \in \mathbb{R}^d, \end{cases} \quad (\text{I.3})$$

où les matrices  $A_j \in \mathbf{M}_{N \times N}(\mathbb{R})$  et où la donnée initiale  $u_0$  est donnée dans un certain espace fonctionnel que l'on va préciser.

Le système (I.3) se résout facilement en utilisant l'analyse de Fourier lorsque  $u_0 \in L^2(\mathbb{R}^d)$ . De plus les résultats de Brenner [Brenner, 1966]-[Brenner, 1973] affirment qu'en dehors du cas bien particulier où les matrices  $A_j$  commutent deux à deux alors le problème (I.3) n'est pas bien posé dans  $L^p$  pour  $p \neq 2$ . Par conséquent, l'espace naturel pour étudier le problème linéaire est l'espace de Hilbert<sup>3</sup>  $L^2$ . Fort de cette remarque, on résout alors facilement du côté de Fourier en posant :

$$u(t, x) = \mathcal{F}_{\xi \rightarrow x}^{-1} e^{-itA(\xi)} \widehat{u}_0(\xi),$$

où  $A$  désigne le symbole (spatial) de  $L(\partial)$  ; pour  $\xi \in \mathbb{R}^d$ ,  $A(\xi) = \sum_{j=1}^d \xi_j A_j$ . Pour que la solution précédente ait un sens dans  $L^2$ , on voit alors apparaître les deux conditions suivantes :

- Pour tout  $\xi \in \mathbb{R}^d$ , le symbole  $A(\xi)$  est diagonalisable à valeurs propres réelles.
- La matrice de passage  $T \in \mathbf{GL}_N(\mathbb{R})$  (et son inverse) est bornée uniformément par rapport au paramètre de fréquence  $\xi$  à savoir :

$$\sup_{\xi \in \mathbb{R}^d} \|T(\xi)\| \times \|T^{-1}(\xi)\| < \infty.$$

D'après le théorème matriciel de Kreiss [Kreiss, 1959], les deux conditions précédentes caractérisent la notion d'hyperbolicité pour l'opérateur  $L(\partial)$  et donc le caractère bien posé du système (I.3).

Ainsi, on peut dire que le problème de Cauchy (I.3) est un problème bien compris et dans (presque) toute la suite de ce mémoire, on s'intéressera aux problèmes aux limites associés à l'opérateur  $L(\partial)$ . L'étude de tels problèmes sera motivée plus en détails au début du Chapitre II.

Les problèmes que l'on considérera s'écriront donc sous la forme :

$$\begin{cases} L(\partial)u = f & \text{dans } \mathbb{R}_+ \times \Gamma, \\ Bu|_{x \in \partial\Gamma} = g, \\ u|_{t=0} = u_0, \end{cases} \quad (\text{I.4})$$

où  $\Gamma \subset \mathbb{R}^d$  a vocation à être le demi-espace  $\mathbb{R}^{d-1} \times \mathbb{R}_+$ , ou encore le quart d'espace  $\mathbb{R}_+^2 \times \mathbb{R}^{d-1}$ . Dans le système (I.4), la condition de bord sur  $\partial\Gamma$  est générique au sens où le coefficient  $B$  est une matrice sur laquelle on impose juste que le nombre de lignes (et donc le nombre de composantes de la trace prescrites au bord) soit compatible avec l'opérateur  $L(\partial)$ . On peut facilement s'en convaincre en considérant le cas de  $N$  équations de transports découplées : dans un cadre hyperbolique seulement un certain nombre de composantes de la trace doivent être prescrites. Ceci est une propriété phare, sur laquelle on reviendra,

<sup>2</sup>Les travaux de l'auteur [Benoit, 2015] se cantonnaient au cadre linéaire. Les travaux présentés ici sont, pour leur majorité, dans la lignée directe de ces travaux de thèse. Il est donc assez naturel que l'on travaille, ici principalement dans le cas linéaire. Toutefois, après ces années à travailler dans un cadre linéaire, l'auteur se sent maintenant prêt à passer au monde non linéaire.

<sup>3</sup>Ce cadre fonctionnel, on ne peut plus rigide, constitue l'une des principales difficultés à l'obtention de résultats d'existence et d'unicité dans le cadre non linéaire. En effet, on a alors plus naturellement, à considérer des solutions qui sont  $L^1$  ou bien  $BV$ . Espaces qui se comparent mal à l'espace  $L^2$  du problème linéaire.

qui distingue les problèmes aux limites hyperboliques de leurs homologues elliptiques ou paraboliques où toutes les composantes doivent être prescrites.

Les questions que l'on s'est posées durant ces dernières années sur le système (I.4) peuvent alors être grossièrement résumées de la façon suivante :

- La caractérisation du caractère fortement bien posé. Plus précisément, peut-on caractériser (toutes) les conditions de bord  $B$  qui donnent lieu à un problème bien posé (au sens de Hadamard) ? Si oui, peut on expliciter cette caractérisation ? Si non, existe-t-il des conditions de bord qui donnent lieu à des notions bien posées plus faibles et si oui, pouvons nous alors décrire le comportement des solutions ?
- Une fois l'existence d'une solution (en un certain sens de caractère bien posé) établie, on aimerait construire, le plus explicitement possible, une solution approchée de la solution du problème (I.4). On aimerait, de plus, que la solution approchée le soit à un ordre de précision arbitrairement grand. Ceci afin de pouvoir utiliser cette solution approchée pour mettre en évidence des phénomènes qualitatifs sur la solution exacte.

C'est ces deux motivations, très générales, mais précisées au fil du texte, qui seront étudiées dans le Chapitre II.

Une fois une théorie du caractère bien posé pour le système (I.4) établie, se pose alors assez naturellement la question d'obtention de méthodes numériques implémentables qui permettent alors de venir approcher la solution exacte de (I.4). Nous devrions plutôt ici dire non pas la question de l'obtention mais plutôt de l'étude théorique de la convergence de la solution approchée vers la solution exacte.

En effet, les outils d'analyse numérique que l'on a mis en œuvre dans le Chapitre III sont les plus simples que l'on puisse concevoir à savoir des approximations de type différences finies. On renvoie à la Section III.2 pour quelques motivations qui nous ont poussé à étudier ces méthodes. On espère que le lecteur trouvera ces motivations suffisantes pour justifier pleinement l'emploi de méthodes aussi simples.

Là encore, comme pour ce qui nous a poussé à préférer travailler sur le cas simplifié de la linéarisation, (I.3), de (I.1), plutôt que sur le problème non linéaire, qui relève d'avantage de la Physique ; on a ici été motivé le souhait d'établir la théorie la plus concluante, précise et complète possible dans un cadre simple plutôt que d'essayer de développer des méthodes plus élaborées mais plus incomplètes dans un cadre plus complexe.

De plus, comme on le verra au fil de ce mémoire, même dans les cas simples que sont les cadres linéaires, ou bien les approximations par différences finies, beaucoup de choses belles et intéressantes restent à être comprises.

Par conséquent, il ne sera pas ici question de méthodes numériques élaborées, techniques ou encore astucieuses qui permettent d'approcher la solution exacte de systèmes complexes. Toutefois on mettra l'accent sur l'aspect théorique en tâchant d'obtenir une théorie satisfaisante de convergence, ou du moins ses éléments clés.

La question principale de cette synthèse est la suivante : On se place dans un domaine de résolution borné. Ce dernier vient naturellement lors de l'implémentation d'un schéma sur machine. Peut-on trouver, voir caractériser, des conditions de bord numériques qui assurent la convergence du schéma d'approximation par différences finies vers la solution exacte du problème (I.4) ?

Comme on le verra assez vite, nous sommes à l'heure actuelle bien loin d'un tel objectif et dans le Chapitre III, nous nous contenterons de fournir des éléments de réponse assez parcellaires, qui viennent compléter une littérature qui elle-même ne répond pas totalement à cette question.

On trouve toutefois assez marquant le "manque" de résultats traitants de l'analyse "théorique" de convergence pour des schémas aussi simples que des différences finies dès que ces derniers ont à prendre en compte des bords numériques qui pourtant existent au niveau de l'implémentatio). C'est assurément l'une des questions sur lesquelles nous avons l'intention de continuer à contribuer à l'avenir.

On conclut cette section d'introduction générale par deux paragraphes. Le premier décrit, plus en détails, l'organisation du mémoire, le second est une table (que l'on espère la plus exhaustive possible) des différentes notations rencontrées au fil du texte.

## I.2 Organisation du mémoire

Cette monographie est organisée de la façon suivante:

- Le Chapitre II traite de problèmes aux limites hyperboliques dans le cas continu. Il se compose de quatre sections distinctes. Ces parties sont auto-satisfaisantes et leur lecture peut se faire en grande partie indépendamment l'une de l'autre. Toutefois, puisque ces dernières correspondent à un niveau de difficulté croissant, on conseille l'ordre de lecture séquentiel d'apparition. Plus précisément, les quatre sections sont :
  1. La Section II.1. Elle vient donner des motivations poussant à l'étude des problèmes aux limites hyperboliques. Elle précise ainsi les termes de la partie introductive I.1.
  2. La Section II.2 traite de problèmes aux limites dans la géométrie canonique du demi-espace. Cette section commence par reprendre les éléments marquants de la théorie développée par Kreiss [Kreiss, 1970], puis introduit la notion de problèmes faiblement bien posés grâce à laquelle est illustrée l'importance et l'intérêt des développements d'optique géométrique. Le lecteur familier avec ces notions pourra très probablement passer cette section introductive.
  3. La Section II.3 regroupe les résultats que l'on a obtenus dans le cas du problème posé dans une bande. L'intérêt principal d'une telle géométrie est qu'il s'agit d'une géométrie intermédiaire entre le demi-espace et le quart d'espace en ceci que l'on garde la difficulté provoquée par le fait que la frontière admette plusieurs composantes. Toutefois, comme ces composantes sont paramétrisées par la même variable, l'analyse se trouve être simplifiée.
  4. Enfin, la Section II.4 décrit la littérature existante ainsi que nos contributions concernant le problème dans le quart d'espace. Comme on le verra la théorie bien posée pour les problèmes dans le quart d'espace est encore assez ouverte. C'est d'ailleurs assez surprenant, car ces problèmes apparaissent dans les années 1970 dans la littérature. On espère, ici, que nos diverses contributions, de nature assez variées (persistance de la régularité, développement d'optique géométrique...) permettent d'apporter des éléments favorisant la compréhension de tels problèmes.

Dans les trois dernières sections précédemment citées, on reprendra les deux questions principales exposées dans l'introduction I.1, à savoir que l'on discutera d'une part de la caractérisation des conditions de bord donnant lieu au caractère bien posé et qu'on illustrera d'autre part cette caractérisation sur des solutions approchées données par des développements d'optique géométrique.

- Le Chapitre III traite de l'analogue des problèmes décrits dans le Chapitre II, mais dans un cadre discret. On y met l'accent sur l'approximation de type différences finies qui est l'une des façons les plus simples de calculer numériquement une solution approchée de la solution exacte. L'ingrédient clé de la théorie de convergence étudiée dans le Chapitre III est celui de la stabilité de la méthode.

Ce chapitre se compose de six sections distinctes et l'on reprend ici l'ordre d'exposition du Chapitre II. Plus précisément

1. Les deux premières sections à savoir les Sections III.1 et III.2 sont introductives. Précisément la Section III.1 introduit et décrit les schémas de type différences finies dans la droite réelle et donne des éléments que l'on a jugé bon de rappeler en vue de l'extension à des bords numériques. On décrira des motivations qui nous poussent à étudier de tels schémas dans la Section III.2.
2. La Section III.3 reprend les éléments phares de l'analyse pionnière dans la demi-droite de [B. Gustafsson and Sundstrom., 1972]. Cette section reprend donc des éléments assez classiques et pourra parfaitement être survolée par le lecteur à l'aise avec la théorie GKS.

3. La Section III.4 décrit notre principale contribution à la théorie de la stabilité, à savoir une discrétisation de l'analyse d'Osher pour établir la stabilité d'un problème posé dans un segment. On pense qu'il est intéressant de lire cette section en parallèle avec la Section II.3, qui étudie le problème analogue dans le cas continu.
  4. La Section III.5 explore le cas d'un domaine de discrétisation borné avec comme première étude le cas d'un quart d'espace qui bien que non borné comprend la difficulté induite par la présence de deux conditions de bord numériques qui interagissent. Nous n'y décrivons pas de théorie générale, qui permettrait de conclure, à proprement parler. Cependant, nous donnons des exemples de conditions de bord particulières couplées parfois à des schémas d'approximation particuliers pour lesquels on peut conclure à la stabilité. Une fois de plus, il peut être instructif de lire les résultats du Paragraphe III.5.2 parallèlement à ceux du cas continu (c'est-à-dire ceux de la Section II.4).
- Le Chapitre IV est en grande partie indépendant des deux précédents. Il échappe, ou du moins tente d'échapper, à la théorie des systèmes hyperboliques. Il contient la description de résultats que nous avons obtenus au cours de ces dernières années. Ces questions peuvent toutefois être considérées comme proches de celles adressées dans le Chapitre II. En effet, l'équation étudiée est parfois purement hyperbolique (équation des ondes) ou obtenue comme perturbation visqueuse d'un problème hyperbolique. Ce chapitre se compose de trois sections distinctes et indépendantes :
    1. La Section IV.1 décrit des résultats d'homogénéisation (en temps long) de la solution de l'équation des ondes. Le but étant, ici, de décrire la propagation "moyennée" d'une onde dans une géométrie qui varie fortement selon une échelle  $\varepsilon^{-1}$ ,  $0 < \varepsilon \ll 1$ . On décrira, alors un résultat en temps long, plus précisément sur des échelles de temps de l'ordre de  $\varepsilon^{-k}T$ ,  $k \geq 1$ ,  $T > 0$ .
    2. La Section IV.2 expose des résultats de théorie du contrôle. Elle se compose de deux paragraphes indépendants, le premier traite du contrôle pour une équation de type chaleur (avec dégénérescence du coefficient de diffusion). Le second revient sur l'équation des ondes et donne une méthode de contrôlabilité basée sur le principe de Huygens.  
On notera que les deux méthodes de contrôle proposées sont des méthodes qui permettent de construire directement le contrôle. Elles ont alors des avantages (comme, par exemple, une simplification de la construction numérique des contrôles) comparées aux méthodes indirectes basées sur la dualité et l'observabilité.
    3. Enfin, la Section IV.3 expose des résultats qui permettent de venir décrire précisément la solution d'un système de loi de conservations de la forme (I.1) dans le cadre des ondes soniques de Métivier *via* des développements de type d'optique géométrique. Les résultats de cette section sont encore à l'étude, ils ne sont pas définitifs. Mais, on trouvait intéressants de les décrire quelque peu pour revenir sur le système (I.1) et ainsi boucler la boucle de ce mémoire.
  - Le Chapitre V dessine une conclusion qui reprend, en particulier, les principaux axes de recherche qui se dégagent et sur lesquels nous pensons qu'il serait intéressant de travailler à l'avenir, notamment ceux sur lesquels nous avons envie de travailler dans le futur.

## I.3 Table des notations

### I.3.1 Notations générales

- La lettre  $N \geq 1$ , désigne toujours le nombre d'équations du système considéré.
- La lettre  $d \geq 1$ , désigne toujours la dimension de l'espace ambiant dans lequel le système est posé.
- La lettre  $r \in \mathbb{N}$  désignera toujours le nombre de pas (en espace) d'un schéma vers la droite ( $r$  comme "right").
- La lettre  $\ell \in \mathbb{N}$  désignera toujours le nombre de pas (en espace) d'un schéma vers la gauche ( $\ell$  comme "left").<sup>4</sup>
- La lettre  $b \in \mathbb{N}$  (éventuellement complétée de l'indice 0 ou 1) désignera toujours le nombre de pas (en espace) de la condition de bord numérique d'un schéma.
- De façon générique  $\Omega$  désigne un espace/temps, son éventuel bord spatial sera noté  $\partial\Omega$ , tandis que  $\Gamma$  désigne seulement un espace.
- Pour  $x \in \mathbb{R}^d$  on écrira  $x'$  la partie tangentielle de  $x$ . Dans le cas du demi-espace on décomposera donc  $x = (x', x_d)$  où  $x' \in \mathbb{R}^{d-1}$ , dans le cas du quart d'espace on écrira  $x = (x_1, x_2, x')$  avec  $x' \in \mathbb{R}^{d-2}$ .
- La lettre  $C > 0$ , désigne une constante réelle autorisée à changer d'une ligne à l'autre sans changement de notation.
- Dans un espace métrique  $(X, d)$ , on note  $B(a, R)$ , la boule ouverte centrée en  $a \in X$  de rayon  $R > 0$ ;  $\overline{B(a, R)} := \{x \in X \mid d(a, x) < R\}$ . La boule fermée est notée  $\overline{B(a, R)}$ .
- Soit  $X$  un espace de Banach. Lorsque  $\dim X < \infty$ , on notera  $|u|$  pour la norme de l'élément  $u$ . Lorsque  $\dim X = \infty$ , on préférera la notation  $\|u\|$ .  
Soit  $X$  un espace de Hilbert, on notera de façon équivalente son produit scalaire  $\cdot$  ou  $\langle \cdot, \cdot \rangle$ . La notation  $\cdot$  est réservée à la dimension finie, la notation  $\langle \cdot, \cdot \rangle$  est systématiquement utilisée en dimension infinie.
- Soit  $X$  un espace vectoriel on note  $I$  l'opérateur identité sur  $X$ .
- Pour  $a, b \in \mathbb{R}$ ,  $\llbracket a, b \rrbracket$  désigne "l'intervalle d'entiers" entre  $a$  et  $b$ . On a  $\llbracket a, b \rrbracket := [a, b] \cap \mathbb{Z}$ .
- La notation  $\mathbf{M}_{p,q}(\mathbb{K})$  désigne le sous-espace vectoriel des matrices à  $p$  lignes et  $q$  colonnes à coefficients dans le corps  $\mathbb{K}$ . La notation  $\mathbf{GL}_N(\mathbb{K})$ , pour  $N \geq 1$ , désigne l'ensemble des matrices inversibles de  $\mathbf{M}_{N,N}(\mathbb{K})$ .
- Pour  $A \in \mathbf{M}_{p,q}(\mathbb{R})$ , on note  $A^T \in \mathbf{M}_{q,p}(\mathbb{R})$  la transposée de  $A$ . Pour  $A \in \mathbf{M}_{p,q}(\mathbb{C})$  on note  $A^* \in \mathbf{M}_{q,p}(\mathbb{R})$  la transcommutée de  $A$ . Par extension, pour un opérateur  $A$  linéaire, la notation  $A^*$  désigne l'opérateur adjoint de  $A$ .
- Pour  $A \in \mathbf{M}_{p,q}(\mathbb{C})$  on note  $\operatorname{Re}(A)$  la partie réelle de  $A$  définie par  $\operatorname{Re}(A) := \frac{1}{2}(A + A^*)$ . On généralise aussi cette définition pour un opérateur linéaire  $A$  en dimension infinie.
- Pour  $p \in ]1, \infty]$ ,  $L^p$  désigne l'espace usuel des fonctions de puissance  $p$ -ième intégrables. Les espaces de Sobolev classiques basés sur les espaces  $L^p$  seront notés classiquement  $W^{n,p}$ , où  $n \in \mathbb{N}$ . Classiquement on utilisera la notation  $H^n := W^{n,2}$ . On pose alors  $H^\infty := \bigcap_{n \geq 0} H^n$ . Bien souvent on se placera dans des espaces fonctionnels basés sur des versions pondérées par un poids exponentiel des espaces  $L^p$ . Ces espaces paramétrisés par  $\gamma > 0$ , hériteront du paramètre  $\gamma$  en indice.

<sup>4</sup>On notera ici que certains auteurs préfèrent étonnamment ici l'emploi de la lettre  $r$ ... Peut-être pour "ref" ?



- L'espace des fonctions à décroissance rapide de Schwartz sera noté  $\mathcal{S}$ . L'espace des fonctions  $\mathcal{C}^\infty$  à support compact sera noté  $\mathcal{D}$ . Par conséquent  $\mathcal{D}'$  désigne l'espace des distributions usuel.

### I.3.2 Notations pour les problèmes continus

- Pour  $j \in \llbracket 1, d \rrbracket$ , la lettre capitale  $A$  *via* l'écriture,  $A_j \in \mathbf{M}_{N,N}(\mathbb{R})$ , désigne toujours les coefficients de l'équation intérieure.
- Pour  $j \in \{0, 1, 2\}$ , la lettre capitale  $B$  *via* l'écriture,  $B_j \in \mathbf{M}_{p_j, N}(\mathbb{R})$ , désigne toujours les coefficients de l'équation sur le (les) bord(s).
- L'opérateur différentiel à l'intérieur sera noté  $L(\partial) := \partial_t + \sum_{j=1}^d A_j \partial_j$ . Son symbole (total) sera noté  $\mathcal{L}(\xi) := \xi_0 I + \sum_{j=1}^d \xi_j A_j$ , où  $\xi \in \mathbb{R}^{1+d}$ .
- La lettre  $\mathcal{I}$  désigne, de façon générique, l'ensemble des phases présentes dans un développement d'optique géométrique. Cet espace sera, très souvent, partitionné en différents sous-ensemble  $\mathcal{I}$ , les définitions étant rendus précises dans le texte.
- Pour  $k \in \mathcal{I}$ , associé à une fréquence hyperbolique, on définit la fonction de phase  $\varphi_k := \underline{\tau}t + \underline{\xi}_k \cdot x$ , où  $(\underline{\tau}, \underline{\xi}_k) \in \mathbb{R} \times \mathbb{R}^d$  désigne la fréquence associée à l'indice  $k$ .  
La phase d'oscillation sur le bord sera toujours notée  $\psi$ .
- Pour  $k \in \mathcal{I}$ , associé à une fréquence hyperbolique,  $\mathbf{v}_k$  désigne la vitesse de groupe associée à l'indice  $k$ .
- Pour  $k \in \mathcal{I}$ , associé à une fréquence hyperbolique,  $\Pi_k$  désigne le projecteur (usuel en optique géométrique) sur  $\ker \mathcal{L}(d\varphi_k)$ .
- La lettre  $f$  désigne toujours un terme source intérieur tandis que la lettre  $g$  (possiblement complétée d'un indice) désigne systématiquement un terme source de bord. Finalement, la donnée initiale sera toujours notée  $u_0$ .
- La lettre  $\Xi$  désigne toujours l'ensemble des fréquences ; il est défini par

$$\Xi := \left\{ \zeta := (\sigma := \gamma + i\tau, \eta) \in \mathbb{C} \times \mathbb{R}^{d-1} \setminus \gamma \geq 0 \right\}.$$

- Le symbole  $\Xi_0$  désigne le bord de l'espace des fréquences,  $\Xi_0 := \Xi \cap \{\gamma = 0\}$ .
- Pour  $\zeta \in \Xi$ , la notation  $\mathcal{A}(\zeta)$ , possiblement complétée par un indice  $j \in \{0, 1, 2\}$  désigne la matrice résolvante associée à  $L(\partial)$ .
- Pour  $\zeta \in \Xi$ , la notation  $\mathbf{E}^s(\zeta)$  (resp.  $\mathbf{E}^u(\zeta)$ ), possiblement complétée d'un indice,  $j \in \{0, 1, 2\}$ , désigne le sous-espace stable (resp. instable) de la matrice résolvante  $\mathcal{A}(\zeta)$ , au sens des systèmes dynamiques.
- Pour  $\zeta \in \Xi$ , on note  $\phi(\zeta) = \phi$ , notation possiblement complétée d'un indice, l'inverse de la matrice de bord  $B$  donnée par la condition de Kreiss-Lopatinskii uniforme.
- Pour  $\zeta \in \Xi$ , et  $p \neq q \in \{0, 1, 2\}$ , l'opérateur de traces d'Osher dans le cas continu sera noté par  $\overline{\mathbf{T}}_{p \rightarrow q}(\zeta)$ . Pour  $p = q$ ,  $\overline{\mathbf{T}}_{p \rightarrow p}(\zeta)$  désignera la composée  $\overline{\mathbf{T}}_{q \rightarrow p}(\zeta) \overline{\mathbf{T}}_{p \rightarrow q}(\zeta)$  obtenue pour  $q \neq p$ .

### I.3.3 Notations pour les problèmes discrets

- L'ensemble  $\mathcal{U} \subset \mathbb{C}$  désigne l'extérieur de la boule unité fermée,  $\mathcal{U} := \{z \in \mathbb{C} \mid |z| > 1\}$ .
- Pour  $z \in \mathbb{C} \setminus \{0\}$ , la matrice qui joue le rôle de la matrice résolvante dans le cas discret sera notée  $\mathbb{M}(z)$ . La matrice qui joue le rôle de la matrice au bord sera notée  $\mathbb{B}(z)$ , notation éventuellement complétée d'un indice.
- Pour  $z \in \mathcal{U}$ , l'espace vectoriel  $\mathbb{E}^s(z)$ , désignera le sous-espace stable de la matrice résolvante discrète  $\mathbb{M}(z)$ . L'espace vectoriel  $\mathbb{E}^u(z)$  désignera le sous-espace instable.
- Pour  $z \in \mathcal{U}$ , l'inverse de la matrice de bord  $\mathbb{B}(z)$ , restreinte au sous-espace stable  $\mathbb{E}^s(z)$ , donnée par la condition de Kreiss-Lopatinskii uniforme discrète sera notée  $\Phi(z) := \mathbb{B}_{|\mathbb{E}^s(z)}^{-1}$ . Cette notation pourra être complétée d'un indice.
- Pour  $z \in \mathcal{U}$  et  $p \neq q \in \{0, 1\}$ , l'opérateur de traces dans le cas discret sera noté par  $\overline{\mathbb{T}}_{p \rightarrow q}(z)$ . Pour  $p = q$ ,  $\overline{\mathbb{T}}_{p \rightarrow p}(z)$  désignera la composée  $\overline{\mathbb{T}}_{q \rightarrow p}(z) \overline{\mathbb{T}}_{p \rightarrow q}(z)$  obtenue pour  $q \neq p$ .

## Chapter II

# Problèmes aux limites hyperboliques continus

### II.1 Motivations

Dans ce paragraphe introductif, on se propose de donner quelques motivations qui expliquent l'intérêt de l'étude d'opérateurs hyperboliques posés dans un demi ou un quart d'espace. On décrira donc des exemples, formant une liste probablement non exhaustive, dans lesquels on est naturellement amené à rencontrer de tels problèmes.

Un premier exemple d'apparition des problèmes aux limites hyperboliques dans le demi-espace est issu de l'étude de la propagation d'un phénomène ondulatoire dans une géométrie bornée régulière. En effet, l'un des moyens simples d'attaquer ce problème est d'utiliser un argument de partition de l'unité. On se ramène alors à considérer une collection de problèmes d'un des deux types suivants :

- les premiers sont posés dans des ouverts n'intersectant pas la frontière  $\partial\Omega$ . Ils se comportent comme des problèmes de Cauchy.
- Les seconds sont posés dans des ouverts qui viennent couper  $\partial\Omega$ . Ils se comportent alors exactement comme des problèmes posés dans le demi-espace.

Bien sûr l'argument précédent nécessite de la régularité sur la frontière  $\partial\Omega$  afin de ramener les ouverts qui intersectent cette frontière à des problèmes hyperboliques dans le demi-espace. En effet, ce procédé n'est pas possible au voisinage d'un point  $\underline{x} \in \partial\Omega$  où la frontière est seulement lipschitzienne. Dans un tel cadre, le manque de régularité de la frontière doit être conservé et on se ramène alors à considérer un problème posé dans un domaine avec un coin.

Dans ce qui suit, on a débuté l'étude par un angle de taille  $\frac{\pi}{2}$  et donc on travaillera dans le cas du quart d'espace. Il serait toutefois intéressant de reprendre l'analyse pour un angle quelconque et voir en quoi cette dernière est modifiée. Il est à noter que si l'on dispose de résultats permettant des coefficients  $A_j$  variables alors on peut ramener une géométrie présentant un angle quelconque à un angle de taille  $\frac{\pi}{2}$  par un changement de variables.

Par conséquent, les problèmes aux limites dans le demi-espace et ou le quart d'espace sont de bons *toy-models* pour l'étude de propagation d'ondes dans des géométries bornées, plus complexes, ayant un bord éventuellement seulement lipschitzien.

Un autre exemple d'apparition dans la littérature des problèmes aux limites dans le demi-espace provient de la nature même des procédés qui permettent de construire des solutions de systèmes de

lois de conservation de la forme (I.1). On vient parfois chercher une solution  $u$  du système de lois de conservations (I.1) sous la forme d'une succession de deux états réguliers séparés par une interface  $\Gamma := \{(t, x) \mid \varphi(t, x) = 0\}$ , à travers laquelle, la solution peut développer une instabilité. Plus précisément, on écrit

$$u = \begin{cases} u_\ell & \text{si } 0 > \varphi(t, x), \\ u_r & \text{si } 0 < \varphi(t, x), \end{cases}$$

ainsi en supposant que la fonction  $\varphi$  est assez régulière, on peut redresser ce front de singularité et se ramener à étudier un problème dans le demi-espace de gauche pour l'état  $u_\ell$  et un problème dans le demi-espace de droite pour l'état  $u_r$ .

Tout le jeu est alors de comprendre comment les états de gauche et de droite peuvent communiquer l'un avec l'autre. En considérant un système de taille supérieure (en posant essentiellement  $u = [u_\ell, u_r]^T$ ) cela revient alors à imposer des conditions de transmission à l'interface maintenant redressée  $\{x_d = 0\}$ . Cette méthode a été menée à bien pour construire des solutions de systèmes de lois de conservation pouvant développer différents types de singularités le long de l'interface. On citera ici, à titre d'illustration, le cas des chocs, étudié par exemple dans [Guès et al., 2005a] ou encore dans le récent papier [Faye and Rodrigues, 2023], pour lesquels on a une discontinuité à travers l'interface ou encore le cas des ondes soniques de Métivier ([Métivier, 1991], voir aussi Paragraphe IV.3), pour lesquelles la solution est continue à travers l'interface, mais dont c'est la dérivée normale qui admet une discontinuité.

De la même façon, de nombreux problèmes à frontière libre ou encore de transition entre deux états (voir par exemple, de façon non exhaustive, [Coulombel, 2002a], [Benzoni-Gavage and Freistühler, 2004] ou encore [Benzoni-Gavage et al., 2012]) peuvent, après redressement du front, se ramener à l'étude d'un système hyperbolique posé dans le demi-espace. On voit donc se dessiner le rôle central de l'étude des problèmes dans la géométrie du demi-espace au sein de la théorie générale des lois de conservation.

Profitons ici de ce point pour motiver le fait que dans notre étude on permet des conditions de bord matricielles génériques qui sont bien loin de celles données par la physique (on pense ici aux conditions de Dirichlet et de Neumann). En effet, pour venir prendre en compte des conditions de transmission intéressantes on a besoin d'un tel formalisme. Par exemple, dans le cas des chocs, il est naturel d'imposer les conditions de Rankine-Hugoniot, conditions qui peuvent être complétées par celle de stabilité des chocs de Lax [Lax, 1973] (ces dernières étant équivalentes à la condition de Kreiss-Loaptinskii uniforme pour le problème aux limites en dimension supérieure), pour aboutir à la stabilité des ondes de chocs.

La précédente méthode permet donc de construire des solutions de systèmes de lois de conservation. Elle est particulièrement pertinente dans le cas unidimensionnel. Toutefois, dans un cadre multidimensionnel, l'étude du problème de Riemann tend à montrer que l'on s'attend à ce que la solution d'un système de lois de conservations génère autant de singularités que l'on a d'équations. Par exemple, dans le cas  $N = 2$ , on s'attend donc à devoir considérer des solutions présentant trois états  $u_\ell$ ,  $u_m$  et  $u_r$  séparés par deux interfaces  $\Gamma_\ell := \{\varphi = 0\}$  et  $\Gamma_r := \{\psi = 0\}$ . La solution recherchée s'écrit alors sous la forme

$$u = \begin{cases} u_\ell & \text{si } 0 > \varphi(t, x), \\ u_m & \text{si } \varphi(t, x) > 0 \text{ et } 0 > \psi(t, x), \\ u_r & \text{si } \psi(t, x) > 0. \end{cases}$$

On redresse encore une fois le front et on se ramène alors à étudier deux problèmes dans les demi-espaces de gauche et de droite pour les états  $u_\ell$  et  $u_r$ . Mais, maintenant on a de plus, un problème dans un domaine à coin pour l'état intermédiaire  $u_m$ . Cette approche a, par exemple, été utilisée par Métivier dans [Métivier, 1986] pour étudier l'interaction de deux ondes de choc.

Par ces quelques exemples, on espère avoir justifié pour le lecteur l'importance que peuvent revêtir les problèmes aux limites posés dans le demi-espace, ou encore dans le quart d'espace, avec conditions de bord génériques.

## II.2 Le problème dans le demi-espace

Dans ce paragraphe, on se concentrera sur le problème aux limites hyperbolique posé dans le demi-espace. En effet c'est dans cette géométrie que l'étude de ce problème débute historiquement par le travail fondateur de Kreiss [Kreiss, 1970].

Dans ce qui suit on considère donc le problème

$$\begin{cases} L(\partial)u = \partial_t u + \sum_{j=1}^d A_j \partial_j u = f & \text{dans } \mathbb{R} \times \mathbb{R}_+^d := \Omega, \\ Bu|_{x_d=0} = g & \text{sur } \mathbb{R} \times \mathbb{R}^{d-1} := \partial\Omega, \\ u|_{t=0} = u_0 & \text{sur } \mathbb{R}_+^d := \Gamma. \end{cases} \quad (\text{II.1})$$

Dans le système (II.1), l'inconnue  $u$  est à valeurs dans  $\mathbb{R}^N$ , pour  $N \geq 1$ . Les coefficients de l'équation intérieure de (II.1) sont ainsi des matrices  $A_j \in \mathbf{M}_{N \times N}(\mathbb{R})$ . De même on choisit de travailler avec une condition de bord générique c'est-à-dire que le coefficient  $B \in \mathbf{M}_{p \times N}(\mathbb{R})$  est une matrice qui encode juste le bon nombre de conditions de bord, à savoir le nombre de valeurs propres positives<sup>1</sup> du coefficient normal  $A_d$ .

Comme précédemment indiqué, le fait que l'on ait seulement un certain nombre de composantes de la trace sur le bord de  $\mathbb{R}_+^d$  qui soient prescrites est un phénomène propre aux équations hyperboliques. Ceci rend l'étude des problèmes aux limites hyperboliques plus riche mais aussi plus délicate que celle de leurs cousins elliptiques et paraboliques.

Avant toute chose, il faut donner un cadre dans lequel le système (II.1) est bien posé. On entendra ici bien posé au sens de Hadamard, c'est-à-dire que, dans un certain espace fonctionnel, la solution  $u$  existe, est unique, et peut être contrôlée de façon continue en fonction des données du problème.

Afin d'assurer le caractère bien posé du problème de Cauchy associé au problème (II.1), une hypothèse d'hyperbolicité sur les coefficients  $A_j$  doit être faite. Cette hypothèse a été brièvement mentionnée dans l'introduction. Elle peut venir prendre de nombreuses formulations selon le cas d'étude considéré (stricte hyperbolicité, à multiplicité constante, symétrie,...). Toutefois, ces dernières préoccupations n'ont pas été au cœur de notre recherche. C'est pourquoi, on ne décrira pas ici les différents notions d'hyperbolicité utilisées dans les énoncés qui suivent. On se permettra de rester assez vague sur l'hypothèse d'hyperbolicité sous laquelle on se place. En effet, l'intérêt de l'étude se place ailleurs.

Sans entrer dans des détails techniques inutiles, on rappelle que cette hypothèse d'hyperbolicité provient de l'analyse du problème de Cauchy par transformée de Fourier. L'hyperbolicité se décompose alors en deux points. D'abord en une propriété spectrale sur le symbole assurant l'intégrabilité des solutions propres associées, ensuite, en une propriété de diagonalisation (pour l'étude spectrale précédente) uniforme par rapport à la fréquence. Cette propriété uniforme permet alors de ramener les solutions propres précédentes dans la variable de position  $x$ . Comme on va le voir, ce caractère uniforme sera rencontré plusieurs fois dans ce manuscrit et jouera le rôle de fil rouge.

On suppose donc dans la suite que le problème de Cauchy associé à (II.1) est bien posé au sens de Hadamard. Les résultats de [Brenner, 1966] et [Brenner, 1973] impliquent que dans le cas général, on

---

<sup>1</sup>On peut facilement s'en convaincre en considérant le cas unidimensionnel pour un coefficient  $A_d$  diagonale. Le système (II.1) se découple alors en  $N$  équations de transports. Seules celles associées à une vitesse positive ont alors besoin d'une condition de bord. Les traces des transports sortants résultent de la résolution de l'équation.

doive se placer nécessairement dans un espace fonctionnel basé sur  $L^2(\Omega)$ . En particulier, les questions hyperboliques viennent avec de fortes contraintes structurelles et dans tout ce qui suit, on n'aura que peu de liberté sur le cadre fonctionnel de résolution.

En un sens, cette absence de choix est "rassurante" car nous saurons directement dans quel espace travailler, là où, la résolution d'une équation aux dérivées partielles revient, bien souvent, à trouver le bon espace fonctionnel de travail. Toutefois, toutes les méthodes qui nécessitent de s'extraire du cadre purement hilbertien  $L^2$  ne pourront pas être utilisées.

Le but premier de l'analyse du problème aux limites (II.1) est de caractériser les conditions de bord  $B$  qui donnent lieu à un problème fortement bien posé. Une telle caractérisation remonte aux travaux de Kreiss [Kreiss, 1970] et est décrite au Paragraphe II.2.1. Ensuite, l'existence de la solution étant établie, il est alors possible de chercher à décrire des propriétés qualitatives. On exposera de telles propriétés, obtenues par des développements de type optique géométrique, dans le Paragraphe II.2.3.

On conclut ce paragraphe introductif par une définition rigoureuse de ce que l'on entend par problème fortement bien posé dans la suite du texte. Ce concept sera toutefois amené à différer en fonction de la géométrie de notre espace de résolution. On demandera par exemple le contrôle de toutes les composantes de la trace lorsque le bord du domaine de résolution  $\Omega$  admet plusieurs faces.

Dans tous les cas, il s'agit de la notion de problème fortement bien posé utilisée (ou directement adaptée de) dans [Kreiss, 1970].

**Définition II.2.1 (Problème fortement bien posé)** *Le problème posé dans le demi-espace (II.1) est dit fortement bien posé si pour tout  $\gamma > 0$ , et pour tout terme source<sup>2</sup>  $f \in L^2_\gamma(\Omega)$ ,  $g \in L^2_\gamma(\partial\Omega)$ , il existe une unique solution  $u \in L^2_\gamma(\Omega)$ , à trace  $u|_{x_d=0} \in L^2_\gamma(\partial\Omega)$ , qui vérifie de plus l'estimation d'énergie : il existe  $C > 0$ , telle que pour tout  $\gamma > 0$ , on ait*

$$\gamma \|u\|_{L^2_\gamma(\Omega)}^2 + \|u|_{x_d=0}\|_{L^2_\gamma(\partial\Omega)}^2 \leq C \left( \frac{1}{\gamma} \|f\|_{L^2_\gamma(\Omega)}^2 + \|g\|_{L^2_\gamma(\partial\Omega)}^2 \right), \quad (\text{II.2})$$

où pour un sous-ensemble  $X \subset \Omega$ , on a posé<sup>3</sup>

$$L^2_\gamma(X) := \{u \in L^2(X) \mid e^{-\gamma t} u \in L^2(X)\}.$$

## II.2.1 Le cadre idéal, fortement bien posé, la condition de Kreiss-Lopatinskii uniforme

Dans ce paragraphe on rappelle la condition caractérisant les matrices de bord  $B$  qui conduisent au caractère fortement bien posé du problème (II.1) au sens de la Définition II.2.1. On ne donnera toutefois pas ici une étude approfondie de cette condition dite de Kreiss-Lopatinskii uniforme.

Pour établir une telle caractérisation, une piste très naturelle est de se raccrocher autant que faire se peut à l'analyse du problème de Cauchy associé à (II.1). Par conséquent, il est tentant d'essayer de se ramener à une équation différentielle ordinaire (matricielle) via l'analyse de Fourier, méthode qui permet de traiter aisément le cas du problème de Cauchy.

Toutefois, cette méthode rencontre rapidement l'obstruction structurelle que la variable  $x_d$  vit seulement dans une demi-droite. Par conséquent, on ne peut pas effectuer de transformée de Fourier pour cette variable. L'idée est alors "d'échanger" le rôle de la variable temporelle  $t$  et de la variable normale  $x_d$  et de se ramener à une équation différentielle ordinaire par rapport à la variable  $x_d$ . Pour ce faire, on

<sup>2</sup>On se place ici dans le cadre  $u_0 \equiv 0$ . Si ce n'est pas le cas, la norme  $L^2$  de cette donnée initiale apparaît dans le membre de droite. On demande aussi un contrôle  $L^\infty_t(e^{-2\gamma t} L^2_x)$  de la solution dans le membre de gauche.

<sup>3</sup>L'apparition de ce poids exponentiel en temps est un reliquat d'une transformée de Laplace, expliquée dans ce qui suit.

effectue une transformée de Laplace  $t \leftrightarrow \sigma := \gamma + i\tau$ , où  $\gamma > 0$  et  $\tau \in \mathbb{R}$  et une transformée de Fourier tangentielle  $x' \leftrightarrow \eta \in \mathbb{R}^{d-1}$  dans l'équation (II.1).

On considère ici plutôt le problème aux limites (pur) associé au problème aux limites (II.1). Ce dernier s'écrit

$$\begin{cases} L(\partial)u = f & \text{pour } (t, x) \in \mathbb{R} \times \mathbb{R}_+^d, \\ Bu|_{x_d=0} = g & \text{pour } (t, x') \in \mathbb{R} \times \mathbb{R}^{d-1}. \end{cases}$$

Du caractère fortement bien posé de ce problème sur  $\mathbb{R}$  peut alors se déduire le caractère fortement bien posé du problème (II.1) par causalité et en utilisant le résultat de continuation de [Rauch, 1972]. On renvoie, par exemple, à [[Benzoni-Gavage, 2007]-Section 4.5] pour plus de détails.

Le problème pour la transformée  $\hat{u}$ , s'écrit alors sous la forme

$$\begin{cases} \frac{d}{dx_d}\hat{u} = \mathcal{A}(\zeta)\hat{u} + A_d^{-1}\hat{f} & \text{pour } x_d > 0, \\ B\hat{u}|_{x_d=0} = \hat{g}, \end{cases} \quad (\text{II.3})$$

où l'on a posé pour simplifier  $\zeta := (\sigma, \eta)$  et où la matrice, dite résolvante,  $\mathcal{A}(\zeta)$  est définie par

$$\mathcal{A}(\zeta) := A_d^{-1} \left( \sigma I + i \sum_{j=1}^{d-1} \eta_j A_j \right).$$

L'espace des fréquences  $\Xi$  étant défini par

$$\Xi := \left\{ \zeta := (\gamma + i\tau, \eta) \in \mathbb{C} \times \mathbb{R}^{d-1} \mid \gamma \geq 0 \right\} \text{ et on introduit } \Xi_0 := \Xi \cap \{\gamma = 0\}.$$

Remarquons au passage que l'on se place ici implicitement souvent l'hypothèse d'inversibilité de la matrice  $A_d$ . Cette hypothèse qualifiée d'hypothèse de bord non caractéristique n'est toutefois pas nécessaire dans ce qui suit. On renvoie par exemple à [[Benzoni-Gavage, 2007]-Chapters 3 and 6] pour l'adaptation des résultats ci-dessous dans le cadre caractéristique.

Nous reviendrons aussi dans le Paragraphe II.4.2 pour plus de détails et une discussion sur cette hypothèse.

Le résultat suivant, qui est une conséquence directe de l'hypothèse d'hyperbolicité, assure que, tant que le paramètre de fréquence  $\gamma$  est strictement positif, la matrice résolvante  $\mathcal{A}(\zeta)$  a de bonnes propriétés spectrales. Plus précisément,

**Lemme II.2.1 (Hersh [Hersh, 1963])** *Pour tout  $\zeta$  vérifiant  $\gamma > 0$ , la matrice résolvante  $\mathcal{A}(\zeta)$  n'admet pas de valeurs propres imaginaires pures. On note  $\mathbf{E}^s(\zeta)$  (resp.  $\mathbf{E}^u(\zeta)$ ) le sous-espace propre, stable (resp. instable), engendré par les valeurs propres de partie réelle négative (resp. positive). On a alors  $\dim \mathbf{E}^s(\zeta) = p$ ,  $\dim \mathbf{E}^u(\zeta) = N - p$ , et donc*

$$\mathbb{C}^N = \mathbf{E}^s(\zeta) \oplus \mathbf{E}^u(\zeta). \quad (\text{II.4})$$

Grâce aux sous-espaces précédemment définis on peut alors énoncer la première condition nécessaire au caractère fortement bien posé au sens de la Définition II.2.1. En effet, considérons que le problème (II.1) avec termes sources homogènes,  $f \equiv g \equiv 0$ , soit fortement bien posé alors son unique solution est la solution nulle. Par transformée de Laplace-Fourier, il doit en être de même pour le problème résolvant (II.3) (toujours avec termes sources homogènes). Supposons maintenant qu'il existe un élément non nul,  $\underline{u}$  dans  $\ker B \cap \mathbf{E}^s(\zeta)$ . On définit alors

$$\hat{u}(\zeta, x_d) := e^{x_d \mathcal{A}(\zeta)} \underline{u}.$$

Une telle fonction est alors clairement solution du problème homogène (II.3). De plus, par définition de  $\mathbf{E}^s(\zeta)$  on a  $\hat{u} \in L^2_{x_d}([0, \infty[)$ . Par transformée de Laplace-Fourier inverse, on peut donc construire une solution non triviale au problème linéaire homogène (II.1). Ceci vient directement contredire le caractère fortement bien posé.

Par conséquent, on a justifié la proposition suivante

**Proposition II.2.1 (Condition de Kreiss-Lopatinskii (faible))** *Si le problème aux limites (II.1) est fortement bien posé au sens de la Définition II.2.1, alors la condition, dite de Kreiss-Lopatinskii (faible) est vérifiée. En d'autres termes on doit avoir*

$$\forall \zeta \in \Xi \setminus \Xi_0, \ker B \cap \mathbf{E}^s(\zeta) = \{0\}.$$

*De façon équivalente, la matrice  $B$  restreinte au sous-espace  $\mathbf{E}^s(\zeta)$  est inversible pour tout  $\zeta \in \Xi \setminus \Xi_0$ . Autrement dit, pour tout  $\zeta \in \Xi$ , il existe une constante  $C(\zeta) > 0$ , telle que*

$$\forall u \in \mathbf{E}^s(\zeta), |u| \leq C(\zeta)|Bu|. \quad (\text{II.5})$$

La grande contribution de Kreiss [Kreiss, 1970], à l'étude des problèmes aux limites hyperboliques, est d'avoir remarqué qu'une version de l'équation (II.5) dans laquelle la constante  $C$  est indépendante du paramètre de fréquence  $\zeta$ , suffit pour caractériser le caractère fortement bien posé du problème (II.1).

L'idée principale est ici qu'une équation de la forme (II.5) permet d'établir une estimation d'énergie *a priori* pour la solution du problème résolvant (II.3). Toutefois, si la constante  $C$  de l'estimation (II.5) dépend du paramètre de fréquence, on ne peut pas, *a priori*, revenir au problème (II.1) par transformée de Laplace-Fourier inverse. Si la constante  $C$  est indépendante de la fréquence un tel problème ne se pose plus. Nous avons déjà rencontré exactement le même type de comportement pour la notion d'hyperbolicité où l'on demandait une propriété de diagonalisation uniforme par rapport au paramètre de fréquences pour revenir dans la variable de position  $x$ .

Il est à indiquer, que démontrer une estimation de la forme

$$\forall \zeta \in \Xi \setminus \Xi_0, \forall u \in \mathbf{E}^s(\zeta), |u| \leq C|Bu|, \quad (\text{II.6})$$

où  $C > 0$  est indépendante de  $\zeta$ , pour un opérateur  $L(\partial)$  et une condition de bord  $B$  donnés est une tâche très ardue.

Ainsi, pour pallier ce problème, Kreiss propose l'approche suivante. Dejà, on peut remarquer que pour des questions d'homogénéité, l'estimation (II.6) n'a besoin d'être établie que sur la demi-sphère unité ouverte de  $\Xi$ . Puis, on vient compactifier cette demi-sphère ouverte en considérant le bord  $\{\gamma = 0\}$ . On cherche alors à établir des estimations (II.5) dépendant de façon continue du paramètre  $\zeta$ . Le caractère uniforme de la constante finale provient alors de la compacité de la demi-sphère fermée.

Toutefois, il est à noter que dans ce procédé, le lemme de Hersh qui est, en particulier, utilisé pour définir le sous-espace stable, vient dégénérer. La raison étant simplement que la limite d'une quantité signée n'a aucune raison de rester signée.

Par conséquent, l'analyse de Kreiss nécessite une étude approfondie du sous-espace stable (ou plus précisément de son prolongement par continuité) pour les fréquences du bord  $\{\gamma = 0\}$ . Cette étude se base sur le théorème suivant dit de structure par blocs initialement démontré dans [Kreiss, 1970] pour les opérateurs strictement hyperboliques puis étendu à d'autres classes d'hyperbolicité dans [Métivier, 2000] et [Métivier and Zumbrun, 2005].



**Théorème II.2.1** *Sous l'hypothèse d'hyperbolicité idoine<sup>4</sup>, et celle de non caractéristicité du bord, pour tout  $\underline{\zeta} \in \Xi$ , il existe  $\mathcal{V}$  un voisinage de  $\underline{\zeta}$  dans  $\Xi$ , un entier positif  $L$ , des entiers positifs  $\varrho_1, \dots, \varrho_L$  vérifiant  $\sum_{j=1}^L \varrho_j = N$  et enfin une application  $T$  régulière sur  $\mathcal{V}$  à valeurs dans  $\mathbf{GL}_N(\mathbb{R})$  tels que l'on puisse écrire :*

$$\forall \zeta \in \mathcal{V}, T(\zeta)^{-1} \mathcal{A}(\zeta) T(\zeta) := \text{diag}(\mathcal{A}_1(\zeta), \dots, \mathcal{A}_L(\zeta)),$$

où les différents blocs  $\mathcal{A}_j(\zeta)$  sont des matrices carrées de taille  $\varrho_j$  satisfaisant l'une des alternatives suivantes :

- i) Tous les éléments dans le spectre de  $\mathcal{A}_j(\underline{\zeta})$  ont une partie réelle strictement négative.
- ii) Tous les éléments dans le spectre de  $\mathcal{A}_j(\underline{\zeta})$  ont une partie réelle strictement positive.
- iii) On a  $\varrho_j = 1$ ,  $\mathcal{A}_j(\underline{\zeta}) \in i\mathbb{R}$  et enfin  $\partial_\gamma \mathcal{A}_j(\underline{\zeta}) \in \mathbb{R} \setminus \{0\}$ .
- iv) On a  $\varrho > 1$ , il existe alors  $k_j \in i\mathbb{R}$  tel que la matrice  $\mathcal{A}_j(\underline{\zeta})$  s'écrive sous la forme :

$$\mathcal{A}_j(\underline{\zeta}) = \begin{bmatrix} k_j & i & 0 \\ & \ddots & i \\ 0 & & k_j \end{bmatrix},$$

où le coefficient en bas à gauche de  $\partial_\gamma \mathcal{A}_j(\underline{\zeta}) \in \mathbb{R} \setminus \{0\}$ .

Ce théorème permet alors de définir les quatre types de fréquences que l'on va rencontrer dans l'ensemble  $\Xi_0$ .

**Définition II.2.2** *Une fréquence  $\underline{\zeta} \in \Xi_0$  est dite :*

- *Elliptique* si le Théorème II.2.1 pour la fréquence  $\underline{\zeta}$  est vérifié avec un bloc de type i) et un bloc de type ii) uniquement. On note  $\mathcal{E}$  l'ensemble des fréquences elliptiques.
- *Mixte* ou (elliptico-hyperbolique) si le Théorème II.2.1 pour la fréquence  $\underline{\zeta}$  est vérifié avec un bloc de type i), un de type ii) et au moins un bloc de type iii) (mais aucun de type iv)). On note  $\mathcal{M}$  l'ensemble des fréquences mixtes.
- *Hyperbolique* si le Théorème II.2.1 pour la fréquence  $\underline{\zeta}$  est vérifié avec des blocs de type iii) uniquement. On note  $\mathcal{H}$  l'ensemble des fréquences hyperboliques.
- *Glancing* (ou rasant) si le Théorème II.2.1 pour la fréquence  $\underline{\zeta}$  est vérifié avec au moins un bloc de type iv). On note  $\mathcal{G}$  l'ensemble des fréquences de glancing.

On a donc

$$\Xi_0 := \mathcal{E} \cup \mathcal{M} \cup \mathcal{H} \cup \mathcal{G}.$$

Ainsi, on a quatre types de fréquences, les elliptiques pour lesquelles la propriété de Hersh passe à la limite, les mixtes pour lesquelles au moins l'une des valeurs propres vient dégénérer sur l'axe des imaginaires purs, les hyperboliques pour lesquelles toutes les valeurs propres viennent dégénérer (à l'ordre un) et enfin les *glancing* pour lesquelles au moins l'une des valeurs propres vient dégénérer à un ordre supérieur à un.

---

<sup>4</sup>On entend ici strictement hyperbolique, hyperbolique à multiplicité constante ou bien encore géométriquement régulier au sens de [Métivier and Zumbrun, 2005]. Ce dernier cadre étant même équivalent au fait que la structure par blocs soit vérifiée.

La fin de l'analyse de Kreiss repose sur la construction d'un *symétriseur* qui comme pour le cas du problème de Cauchy associé à (II.1) permet d'établir une estimation d'énergie *a priori*.

L'intérêt d'un *symétriseur* est de se ramener à un cadre dans lequel on peut venir effectuer les différentes intégrations par parties dans le bilan d'énergie. En particulier, le *symétriseur* assure de bonnes propriétés de signes sur les termes de bord. On renvoie à la fin de ce paragraphe pour plus de détails sur un tel cadre d'étude. La construction du *symétriseur* se fait alors après localisation et dépend alors fortement du type de la fréquence  $\underline{\zeta}$  autour de laquelle on localise.

Les constructions au voisinage d'un élément elliptique, mixte ou hyperbolique sont assez directes, la principale difficulté se trouvant au voisinage des fréquences de *glancing*. On ne décrira pas plus avant cette construction et on renvoie par exemple à [[Chazarain and Piriou, 1981]-Section 7.5] ou encore à [[Benzoni-Gavage, 2007]-Chapter 5].

En utilisant alors cette estimation d'énergie *a priori* on peut alors montrer l'existence d'une solution faible, par exemple, via l'introduction d'un problème dual. Puis, on justifie que cette solution faible est unique car vérifiant l'estimation d'énergie *a priori* grâce à un lemme de type "fort=faible". On retrouve alors le résultat suivant qui constitue le travail fondateur de l'étude des problèmes aux limites hyperboliques.

**Théorème II.2.2 (Kreiss [Kreiss, 1970])** *Sous l'hypothèse d'hyperbolicité idoïne et celle de non caractéristicité du bord, le problème aux limites (II.1) est fortement bien posé au sens de la Définition II.2.1, si et seulement si, la condition de Kreiss-Lopatinskii uniforme est vérifiée. A savoir que*

$$\forall \zeta \in \Xi, \forall u \in \mathbf{E}^s(\zeta) \text{ on a } \ker B \cap \mathbf{E}^s(\zeta) = \{0\}.$$

Ce résultat a ensuite connu de nombreuses extensions. On peut par exemple citer le travail de thèse de Rauch [Rauch, 1972] qui permet de prendre en compte des données initiales non nulles et établit le contrôle de la norme  $L_t^\infty(L_x^2)$  de la solution. Une telle estimation d'énergie étant qualifiée d'estimation d'énergie du semi-groupe. Ou encore divers résultats traitant du cas où les coefficients de  $L(\partial)$  sont variables comme, par exemple, le résultat établissant une théorie de persistance de la régularité dans  $H^s$  de Massey et Rauch [Rauch and Massey, 1974].

On citera aussi les extensions au cas caractéristique (linéaire) considérées d'abord dans [Rauch, 1985] par Rauch puis par Secchi dans [Secchi, 1995], [Casella et al., 2006]. L'idée principale de tels résultats sera rappelée succinctement au Paragraphe II.4.2, lorsque l'on considérera la persistance de la régularité dans le quart d'espace.

La théorie de persistance de la régularité précédemment mentionnée peut être utilisée pour traiter des problèmes non linéaires comme c'est par exemple le cas dans un cadre non caractéristique dans [[Rauch and Massey, 1974]-Section 6]. Ou bien encore, pour des problèmes caractéristiques, comme initialement considérés par Guès dans [Guès, 1990], résultat ensuite raffiné par Secchi dans [Secchi, 1996], [Secchi, 1997] et les références ultérieures.

Notons toutefois que la majeure partie<sup>5</sup> des résultats précédents ne traite pas directement de problèmes satisfaisant la condition de Kreiss-Lopatinskii uniforme mais qu'ils s'appliquent uniquement à une classe de conditions de bord particulières, à savoir les conditions de bord strictement dissipatives. On rappelle ici la définition de telles conditions de bord, cette dernière étant appelée à être évoquée dans la suite.

---

<sup>5</sup>C'est du moins le cas pour les problèmes caractéristiques.

**Définition II.2.3** [Condition de bord strictement dissipative] On dit que la matrice de condition de bord  $B$  du problème (II.1) est strictement dissipative si l'inégalité suivante est vérifiée :

$$\forall u \in \ker B, u \neq 0, \langle A_d u, u \rangle < 0.$$

Ceci est équivalent à requérir l'existence de deux constantes  $\varepsilon, C > 0$  telles que

$$\forall u \in \mathbb{R}^N, \varepsilon |u|^2 + \langle A_d u, u \rangle \leq C |Bu|^2. \quad (\text{II.7})$$

Les avantages de telles conditions comparées à celles qui vérifient uniquement la condition de Kreiss-Lopatinskii uniforme sont multiples. En effet, en premier lieu, de telles conditions sont facilement à vérifier algébriquement tandis que la vérification de la condition de Kreiss-Lopatinskii uniforme constitue un travail bien plus ardu. Ensuite en ne conservant par projection que la partie de  $A_d$  associée à des valeurs propres négatives, il est clair que de telles conditions de bord existent toujours. Enfin, pour de telles conditions de bord, la démonstration de l'estimation d'énergie *a priori* qui constitue le point technique de la démonstration de Kreiss [Kreiss, 1970] est remplacée par une simple intégration par parties (dans laquelle le terme de bord sur  $\{x_d = 0\}$  se trouve exactement contrôlé par (II.7)) et est donc pour ainsi dire gratuite.

## II.2.2 *Troubles begin*, le cadre faiblement bien posé

Les résultats du paragraphe précédent, en particulier le théorème de Kreiss [Kreiss, 1970], sont très satisfaisants puisqu'ils permettent de totalement caractériser le caractère fortement bien posé en termes de la condition de Kreiss-Lopatinskii uniforme. Toutefois, des questions demeurent. Ces dernières sont de savoir si oui ou non, la condition de Kreiss-Lopatinskii uniforme est effectivement vérifiée en pratique ? Si oui, est-elle vérifiée "souvent" et si non, quelles sont les conséquences sur le caractère fortement bien posé ?

Comme on l'a mentionné précédemment, pour un opérateur hyperbolique donné, vérifier si telle condition de bord satisfait effectivement la condition de Kreiss-Lopatinskii uniforme est une tâche ardue. Ceci nécessite une étude propre à l'opérateur et la condition de bord considérés.

Cette difficulté est essentiellement liée au simple fait que l'on doit considérer des prolongements par continuité d'un sous-espace vectoriel pour lequel on ne dispose pas de formules explicites. En effet, ce dernier dépend des valeurs propres de la matrice résolvante  $\mathcal{A}(\zeta)$  qui ne sont pas explicitement connues lorsque les coefficients de  $L(\partial)$  n'ont pas de structure particulière. Ceci est d'autant plus vrai si  $N$  est grand.

A ce propos, on se permet de mentionner une piste de recherche que l'on aimerait étudier à l'avenir. Dans sa thèse sur la stabilité des schémas d'approximation de type différences finies (on renvoie à [Le Barbenchon et al., a] et [Le Barbenchon et al., b]) Le Barbenchon donne un moyen robuste de vérifier numériquement si une condition de bord (numérique) vérifie la condition de Kreiss-Lopatinskii uniforme discrète. Cette condition caractérise ici la stabilité du schéma. On peut se référer à la Section III.3 pour plus de détails pour la stabilité des schémas numériques. Ou bien le lecteur peut aussi remplacer le mot "stabilité" par "caractère fortement bien posé" et utiliser la théorie continue précédemment décrite.

Les articles [Le Barbenchon et al., a] et [Le Barbenchon et al., b] reposent sur l'introduction d'un déterminant de Lopatinskii dit intrinsèque  $\Delta_{int}$  qui a des propriétés d'holomorphic supérieures à celles du déterminant de Lopatinskii "usuel" défini par<sup>6</sup>  $\Delta(z) := \det B(e_1^s, \dots, e_p^s)(z)$ , où  $(e_1^s, \dots, e_p^s)(z)$  est une base du sous-espace stable du schéma écrit sous forme résolvante. Les zéros de  $\Delta$  sont alors des fréquences en lesquelles la condition de Kreiss-Lopatinskii uniforme discrète dégénère et donc l'existence de tels zéros

<sup>6</sup>Dans le cas des schémas on pose  $d = 1$  et ainsi  $\zeta = z \in \mathbb{C}$ .

conditionne la stabilité du schéma. Or, on construit  $\Delta_{int}$  de manière à ce que ce dernier ait les mêmes zéros que  $\Delta$ . Toutefois, ce dernier ayant de bonnes propriétés d'holomorphicité, le calcul de ses zéros peut, en utilisant des résultats classiques sur les fonctions holomorphes, se ramener au calcul d'un nombre d'entrelacements. Le Barbenchon utilise alors le fait qu'un tel nombre d'entrelacements est robuste à calculer numériquement.

Cette approche fournit, en particulier, un code de vérification numérique de la condition de Kreiss-Lopatinskii uniforme discrète.

Nous aimerions déterminer si la méthode de Le Barbenchon ne pourrait pas être adaptée aux cas de problèmes aux limites continus (II.1). Il s'agirait essentiellement de définir proprement ce qui joue le rôle pour les problèmes continus du déterminant de Lopatinskii intrinsèque. On a ensuite bon espoir que le calcul numérique du nombre d'entrelacements puisse se faire en grande partie comme dans [Le Barbenchon et al., a] et [Le Barbenchon et al., b].

L'essentiel de la difficulté est que dans le cadre continu nous aimerions pouvoir traiter, afin de quitter le cadre unidimensionnel, des paramètres de fréquences  $\zeta = (z, \eta)$  avec au mieux de l'holomorphicité par rapport à  $z$ . On doit donc comprendre comment traiter ces nouveaux paramètres.

Cette approche fournirait alors une méthode qui permet de vérifier numériquement si oui ou non une condition de bord vérifie la condition de Kreiss-Lopatinskii uniforme. De plus, cela donnerait un exemple pour lequel les méthodes dans le cas discret auraient été développées avant celles dans le cas continu, ce qui, à notre connaissance, constituerait une première dans la théorie.

Cet aparté sur la difficulté de vérifier la condition de Kreiss-Lopatinskii uniforme terminé, penchons nous maintenant sur la seconde question à savoir : quelles sont les conséquences d'une dégénérescence de la condition de Kreiss-Lopatinskii uniforme sur le caractère fortement bien posé ? Pour les voir on s'intéresse au système suivant qui provient d'une linéarisation de l'équation d'Euler isentropique autour d'un état subsonique sortant :

$$\begin{cases} \partial_t u + A_1 \partial_1 u + A_2 \partial_2 u = f & \text{dans } \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+, \\ Bu|_{x_2=0} = g & \text{sur } \mathbb{R} \times \mathbb{R}, \\ u|_{t \leq 0} = 0 & \text{sur } \mathbb{R} \times \mathbb{R}_+, \end{cases} \quad (\text{II.8})$$

où les coefficients sont donnés par

$$A_1 := \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2 := \begin{bmatrix} M & 0 & -1 \\ 0 & M & 0 \\ -1 & 0 & M \end{bmatrix} \quad \text{et } B := [1 \quad b_2 \quad b_3],$$

où  $M \in ]-1, 0[$  est le nombre de Mach, et où les coefficients au bord  $b_2, b_3 \in \mathbb{R}$ , sont deux paramètres libres. La Figure II.1, issue de [Benoit, 2014], illustre l'influence de la valeur des paramètres  $b_2$  et  $b_3$  sur le fait que la condition de Kreiss-Lopatinskii uniforme est, oui ou non, vérifiée.

Sur la Figure II.1, nous avons utilisé les abréviations *FI* pour Fortement Instable et *FS* pour Fortement Stable. Les problèmes dans la classe *FI*, représentant l'intérieur d'une demi-boule, sont ceux qui ne vérifient pas la condition de Kreiss-Lopatinskii (faible). On sait que ces problèmes génèrent une instabilité au sens de Hadamard et qu'ils n'ont aucune chance d'être bien posés dans aucun espace fonctionnel. A l'inverse, les problèmes dans la classe *FS*, représentant l'intérieur du cône, sont ceux qui sont fortement stables et donc qui en vertu du Théorème II.2.2 vérifient la condition de Kreiss-Lopatinskii uniforme.

Ainsi, si l'on fait abstraction des paramètres sur le bord du cône, nommément les droites  $\Omega_g^\pm$ , que l'on viendra considérer plus en détails dans la suite de ce mémoire, on se rend alors compte que dans leur "majorité", les conditions de bord n'entrent ni dans un cadre instable ni dans un cadre (fortement) stable. Ces

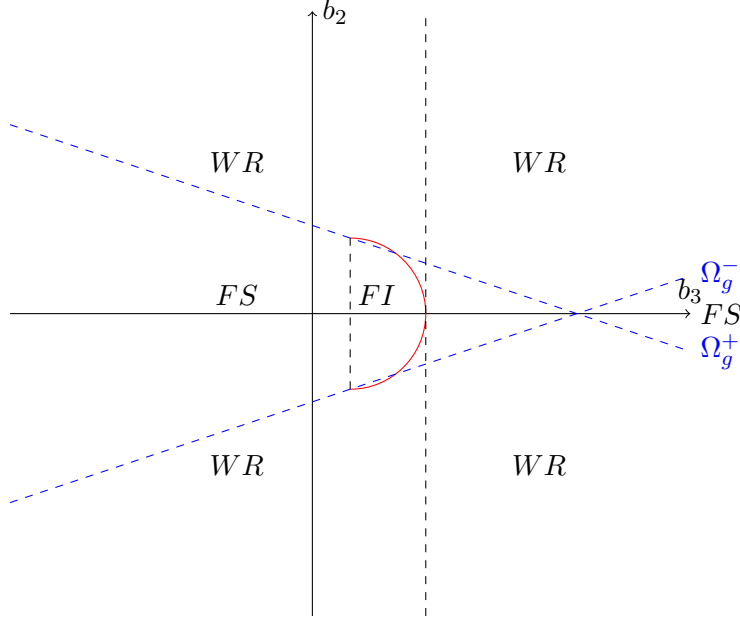


Figure II.1: Les différents régimes de condition de bord pour (II.8).

problèmes admettent des conditions de bord  $WR$  (Weakly Real) au sens de [Benzoni-Gavage et al., 2002] et en termes de caractère bien-posé ils développent des instabilités faibles au sens où l'on a toujours existence et unicité de la solution mais cette dernière vit dans un espace fonctionnel encodant moins de régularité que celui dans lequel les termes sources vivent. On dit aussi que la solution  $u$  perd des dérivées par rapport aux termes sources  $f$  et  $g$ .

Bien sûr cette discussion est menée sur un exemple mais reprenant les exemples traités dans l'article [Benzoni-Gavage et al., 2002], on peut assez facilement se convaincre que les conditions de bord dans la classe  $WR$  sont souvent prépondérantes (ou du moins qu'elles ne peuvent être écartées) dans l'analyse du caractère bien-posé d'un problème aux limites.

On peut toutefois s'interroger sur la réalité physique de telles conditions. En effet, rien, de prime abord, n'assure que l'intérêt de leur étude ne soit pas purement mathématique. Ceci n'est pas le cas. En effet, les travaux assez formels de Artola, Majda et Rosales (voir [Artola and Majda, 1987]-[Majda, 1988] et [Majda, 1983]) se basent sur la motivation que de tels problèmes apparaissent physiquement lorsque l'on regarde, par exemple, la formation des ondes de Mach. Une question naturelle qui se pose alors est de savoir s'il est possible de déterminer quelles sont les pertes de régularité possibles pour la solution.

Avant de répondre à cette question, il est bon de rappeler la définition des conditions de bords dans la classe  $WR$ . On donne ici une formulation équivalente (voir [Coulombel and Guès, 2010]) à la définition de [Benzoni-Gavage et al., 2002].

**Définition II.2.4 (Classe  $WR$ , [Benzoni-Gavage et al., 2002])** *On dit que la condition de bord  $B$  du problème aux limites (II.1) est dans la classe  $WR$  si les trois assertions suivantes sont vérifiées :*

- i) Pour tout  $\zeta \in \Xi \setminus \Xi_0$ , on a  $\ker B \cap \mathbf{E}^s(\zeta) = \{0\}$ .*
- ii) Soit  $\Upsilon := \{\zeta \in \Xi_0 \setminus \ker B \cap \mathbf{E}^s(\zeta) \neq \{0\}\}$  on a alors  $\Upsilon \neq \emptyset$  et  $\Upsilon \subset \mathcal{H}$ .*
- iii) Il existe  $\mathcal{V}$  un voisinage de  $\Upsilon$  dans  $\Sigma := \Xi \cap B(0,1)$ , une fonction  $\theta$ ,  $\mathcal{C}^\infty$  sur  $\mathcal{V}$ , une base  $(e_1(\zeta), \dots, e_p(\zeta))$  de  $\mathbf{E}^s(\zeta)$ , dépendant de façon  $\mathcal{C}^\infty$  de  $\zeta \in \mathcal{V}$ , et enfin une application  $T : \mathcal{V} \rightarrow$*

$\mathbf{GL}_p(\mathbb{C})$  tels que l'on puisse écrire

$$\forall \zeta \in \mathcal{V}, B[e_1(\zeta), \dots, e_p(\zeta)] = T(\zeta) \text{diag}(\gamma + i\theta(\zeta), 1, \dots, 1). \quad (\text{II.9})$$

Donnons quelques commentaires sur la Définition II.2.4. L'exigence *i*) est que si la condition de Kreiss-Lopatinskii uniforme dégénère cette dégénérescence ne peut avoir lieu que sur le bord de l'espace des fréquences, faute de quoi on est automatiquement en présence d'une instabilité forte. Le point *ii*) autorise alors des dégénérescences uniquement aux fréquences du bord qui sont hyperboliques (voir Définition II.2.2). Enfin le point *iii*), qui est de nature plus technique<sup>7</sup>, assure que la dégénérescence de la condition de Kreiss-Lopatinskii se fait à l'ordre un.

En effet, dans le membre de gauche de (II.9), on peut reconnaître un déterminant de Lopatinskii  $\Delta$ , introduit au début de ce paragraphe. Ainsi l'équation (II.9) assure essentiellement que pour  $\underline{\zeta} \in \Upsilon$ , on a  $\partial_\gamma \Delta(\underline{\zeta}) \neq 0$ , et on retrouve alors la définition de la classe  $WR$  de [Benzoni-Gavage et al., 2002].

Il est à noter que l'analyse rigoureuse des problèmes hyperboliques faiblement bien posés ne débute pas par la classe  $WR$  mais par les travaux de Sablé-Tougeron [Sablé-Tougeron, 1988]. Motivée par des considérations d'élastodynamique, l'auteure se place dans le cadre d'une dégénérescence de la Kreiss-Lopatinskii uniforme dans la zone elliptique. On entend par là que l'on a exactement les mêmes conditions sur la dégénérescence du déterminant de Lopatinskii que dans la Définition II.2.4, à ceci près que la dégénérescence a maintenant lieu dans la zone elliptique. Par conséquent on a alors  $\Upsilon \subset \mathcal{E}$  et non plus  $\Upsilon \subset \mathcal{H}$ .

Étant donné qu'il y a quatre zones distinctes dans le bord de l'espace des fréquences, il y a quatre types de dégénérescence possibles pour la condition de Kreiss-Lopatinskii uniforme et il y a fort à parier que les pertes de dérivées diffèrent d'une zone à une autre. C'est en effet le cas comme le montre la revue de la littérature suivante.

Pour une dégénérescence dans la zone elliptique, Sablé-Tougeron établit l'existence et l'unicité d'une solution qui présente une perte d'une dérivée sur le bord. Plus précisément, on entend par là que l'on a le contrôle de la trace dans  $H^{-1/2}$  au prix d'une régularité  $H^{1/2}$  sur la donnée au bord. On a alors l'estimation d'énergie

$$\|u\|_{L_\gamma^2(\Omega)}^2 + \|u|_{x_d=0}\|_{H_\gamma^{-1/2}(\partial\Omega)}^2 \leq \frac{C}{\gamma^2} \left( \|f\|_{L_\gamma^2(\Omega)}^2 + \|g\|_{H_\gamma^{1/2}(\partial\Omega)}^2 \right). \quad (\text{II.10})$$

Pour un autre exemple physique de dégénérescence de la condition de Kreiss-Lopatinskii uniforme en zone elliptique, on renvoie par exemple au problème de MagnétoHydroDynamique étudié par Olivier Pierre dans sa thèse de doctorat [Olivier, 2017].

Ensuite pour la zone mixte, motivé par un problème de transition de phases liquide/vapeur, Coulombel montre dans sa thèse de doctorat [Coulombel, 2002a], une estimation d'énergie comprenant une perte de dérivée sur le bord combinée à une perte d'une demi-dérivée à l'intérieur. Plus précisément, on a l'estimation d'énergie suivante :

$$\|u\|_{L_\gamma^2(\Omega)}^2 + \|u|_{x_d=0}\|_{H_\gamma^{-1/2}(\partial\Omega)}^2 \leq \frac{C}{\gamma^2} \left( \|f\|_{L_{x_d}^2(H_\gamma^{1/2}(\partial\Omega))}^2 + \|g\|_{H_\gamma^{1/2}(\partial\Omega)}^2 \right). \quad (\text{II.11})$$

Enfin, toujours dans son doctorat, Coulombel (publié dans [Coulombel, 2002b]) établit l'estimation d'énergie suivante pour les problèmes dans la classe  $WR$  : pour tout  $\gamma \geq 1$

$$\gamma \|u\|_{L_\gamma^2(\Omega)}^2 + \|u|_{x_d=0}\|_{L_\gamma^2(\partial\Omega)}^2 \leq C \left( \frac{1}{\gamma} \|f\|_{L_{x_d}^2(H_\gamma^1(\partial\Omega))}^2 + \|g\|_{H_\gamma^1(\partial\Omega)}^2 \right), \quad (\text{II.12})$$

<sup>7</sup>La formulation donnée par la Définition II.2.4 a un avantage certain quand on s'intéresse aux développements d'optique géométrique [Coulombel and Guès, 2010]. Il semble aussi que cette estimation soit assez pratique lorsque l'on s'intéresse aux estimations d'énergies comme le montre le récent travail de [Correa, ].

*i.e.* on voit apparaître une perte d’une dérivée à l’intérieur combinée à une perte d’une dérivée sur le bord.

Dans les trois cas, la méthode de démonstration est la même. En effet, il s’agit d’une adaptation locale autour de la fréquence de dégénérescence de la construction du *symétriseur* de Kreiss pour venir compenser des annulations de la condition de Kreiss-Lopatinskii uniforme de la forme

$$\forall v \in \mathbb{R}^N, |v| \leq C\gamma^\delta |Bv|,$$

où  $\delta > 0$ . Pour compenser ce nouveau facteur  $\gamma^\delta$ , on est alors amené à prendre un *symétriseur* de Kreiss qui dépend d’une nouvelle façon de  $\gamma$ . Ceci induit de nouveaux poids dans les autres variables de fréquences et donc de nouvelles normes Sobolev en revenant dans les variables d’état. On renvoie, par exemple à [Coulombel, 2002a] pour des constructions précises et explicites de ces nouveaux symétriseurs, expliquant les nouvelles pertes.

Ce qui est utilisé dans ces trois démonstrations est alors le fait que la construction du *symétriseur* de Kreiss est ”simple” au voisinage des fréquences elliptiques, mixtes et hyperboliques (on renvoie pour plus de détails par exemple à [[Benzoni-Gavage, 2007]-Chapter 5] ou encore à [[Chazarain and Piriou, 1981]-Section 7.5]). Ainsi adapter la construction du *symétriseur* en utilisant les poids idoines est relativement aisé.

Comme on l’a mentionné, la partie technique de la construction du *symétriseur* de Kreiss se fait au voisinage des points de *glancing*. Par conséquent, il est pour ainsi dire normal, qu’à l’heure actuelle, une telle adaptation de la méthode du *symétriseur* de Kreiss n’ait pas été développée pour une dégénérescence de la condition de Kreiss-Lopatinskii uniforme en zone de *glancing*. Pourtant, de telles dégénérescences comprennent des exemples physiques intéressants comme par exemple la condition de bord de Neumann pour l’équation des ondes (on renvoie à [Benoit and Loyer, 2023]).

On espère, un jour, contribuer à l’élaboration d’un tel résultat. Ce dernier semble toutefois assez technique et ambitieux.

Il se peut que les travaux de Eller [Eller, 2018] sur les estimations d’énergie sous la condition de Kreiss-Sakamoto [Sakamoto, 1972] soient une première source d’inspiration. De plus, avant cette étude le lien entre l’existence de conditions de bord conservatives au sens de [Eller, 2012] et la possible dégénérescence de la condition de Kreiss-Lopatinskii uniforme en zone de *glancing* devra être comprise.

Il est toutefois à noter que grâce à notre article [Benoit and Loyer, 2023], nous avons de bons candidats pour quelles pourraient être les pertes de dérivées lors d’une dégénérescence de la condition de Kreiss-Lopatinskii uniforme en zone de *glancing*. Ces résultats seront décrits plus avant dans le Paragraphe II.2.3.

Il est à noter que les diverses pertes de dérivées qui apparaissent dans les estimations (II.10), (II.11) et (II.12) sont de ”vraies” pertes de dérivées dans la mesure où l’on sait que la dégénérescence de la condition de Kreiss-Lopatinskii uniforme induit forcément une perte de régularité. Cela est à mettre en opposition avec les pertes de dérivées que l’on rencontre bien souvent dans la littérature qui, après des fois de longues années, se trouvent être améliorées pour au final converger vers une estimation sans perte.

Une question assez naturelle lorsque l’on est en présence d’estimations d’énergie avec pertes de dérivées est, toutefois, la question de l’optimalité des pertes qui sont établies dans les estimations (II.10), (II.11) et (II.12).

En effet, de prime abord, rien n’exclut le fait qu’en travaillant plus, nous puissions aboutir à des estimations analogues, mais avec des pertes de dérivées moindres. Cette question sera résolue dans la seconde partie du Paragraphe II.2.3, en utilisant des développements d’optique géométrique. Ces derniers font l’objet d’une courte introduction préalable décrivant leur philosophie et leurs motivations dans la première partie du Paragraphe II.2.3.

### II.2.3 Problèmes faiblement bien posés et développements d'optique géométrique

Ce paragraphe décrit un exemple d'application des développements d'optique géométrique appliqués aux problèmes aux limites (II.1). On s'intéressera, en particulier, à l'optimalité des estimations d'énergie avec pertes de dérivées dans la seconde partie. Ce paragraphe reprend les résultats de [Benoit, 2014] et [Benoit and Loyer, 2023]. Toutefois, avant cela, on donne une brève introduction aux développements d'optique géométrique, ou BKW (pour Brillouin-Kramers-Wentzel).

#### Une motivation des développements d'optique géométrique

Les développements d'optique géométrique de la solution d'un problème (pas forcément aux limites) hyperbolique font l'objet d'une littérature très abondante. Les principaux contributeurs à cette étude sont Joly, Métivier et Rauch dans leur série d'articles des années 1990. On cite ici quelques-uns de leurs résultats remarquables [Joly et al., 1993], [Joly et al., 1995a], [Joly et al., 1995b] et [Joly et al., 2000]. La liste étant très loin d'être exhaustive.

Dans ces résultats, les auteurs s'intéressent, dans une majorité des cas, à des développements d'optique géométrique pour des problèmes de Cauchy quasi ou semi-linéaires qui ont des difficultés qui leurs sont propres. Par exemple, s'il ne faut en citer qu'un, la génération d'harmoniques *via* les résonances induites par les non-linéarités.

Puisque nous sommes ici dans un cadre purement linéaire, de tels phénomènes sont absents et les développements d'optique géométrique sont plus simples. En effet, ces derniers contiennent moins de phases (encore que...) que dans le cadre non linéaire. La philosophie que l'on va décrire ci-après sur ces problèmes simplifiés contient toutefois une grande partie des idées générales qui sous-tendent à la construction de n'importe quel développement d'optique géométrique. On renvoie le lecteur intéressé par de tels développements pour les problèmes non linéaires dans le demi-espace aux travaux de Coulombel, Guès et Williams (voir de façon non exhaustive [Coulombel et al., 2011], [Coulombel et al., 2014] et [Coulombel and Williams, 2020]) ou encore au récent travail de Kilque [Kilque, 2022].

Les développements d'optique géométrique s'appliquent à des problèmes (possiblement aux limites) de la forme (II.1) pour lesquels les termes sources sont fortement oscillants. Construire un développement d'optique géométrique pour un tel problème, c'est alors construire un développement approchant la solution exacte dans le régime des hautes fréquences. Pour préciser cela, on s'intéresse au problème aux limites suivant<sup>8</sup>

$$\begin{cases} L(\partial)u^\varepsilon = 0 & \text{dans } \Omega, \\ Bu^\varepsilon|_{x_d=0} = g^\varepsilon & \text{sur } \partial\Omega, \\ u^\varepsilon|_{t \leq 0} = 0 & \text{sur } \Gamma, \end{cases} \quad (\text{II.13})$$

où le paramètre  $0 < \varepsilon \ll 1$ , décrit la longueur d'onde typique du problème. Le terme source  $g^\varepsilon$  est choisi fortement oscillant par rapport à  $\varepsilon$  (voir ci-dessous). On se restreindra dans ce qui suit à des oscillations planes. Le cas d'oscillations plus génériques peut aussi être considéré mais ce dernier peut être soumis à un temps de résolution maximal induit par la résolution de l'équation eikonale sous-jacente (voir (II.16)). On écrit donc  $g^\varepsilon$  sous la forme :

$$g^\varepsilon(t, x') = e^{\frac{i}{\varepsilon}\psi(t, x')} g(t, x'), \text{ où } \psi(t, x') := \underline{\tau}t + \underline{\eta} \cdot x',$$

pour une amplitude  $g$  donnée dans un espace fonctionnel assez régulier (typiquement  $H^\infty(\partial\Omega)$ ) et pour une fréquence de bord  $(\underline{\tau}, \underline{\eta}) \in \mathbb{R}^{1+d-1}$  fixée. On demande aussi que l'amplitude  $g$  vérifie les conditions

<sup>8</sup>Dans un souci de simplification, on a choisi ici de prendre un terme source nul à l'intérieur. En effet, puisque nous sommes en présence d'un problèmes aux limites, nous sommes plus intéressé par l'influence d'un terme source dans la condition de bord de (II.1) que par un terme source interne.



de compatibilités nécessaires.

L'heuristique d'un développement d'optique géométrique est alors de faire le pari que la solution  $u^\varepsilon$  du système (II.13) hérite du caractère fortement oscillant du terme source  $g^\varepsilon$ . On cherche donc à exprimer  $u^\varepsilon$  comme un développement asymptotique (en termes de puissance de  $\varepsilon$ ) "d'un certain nombre" de paquets d'ondes, chacun oscillant selon une fonction de phase  $\varphi_k$ . Ainsi on postule pour  $u^\varepsilon$ , l'ansatz :

$$u^\varepsilon(t, x) \sim \sum_{n \geq 0} \varepsilon^n \sum_{k \in \mathcal{K}} e^{\frac{i}{\varepsilon} \varphi_k(t, x)} u_{n,k}(t, x). \quad (\text{II.14})$$

Construire un développement d'optique géométrique revient alors à :

- Déterminer quelles sont les fonctions de phase  $\varphi_k := \psi + \xi_k x_d$  pertinentes, afin de prendre en compte les différentes oscillations présentes dans la solution  $u^\varepsilon$ . On voudra, en particulier, préciser leur ensemble de définition  $\mathcal{K}$ .
- Déterminer les amplitudes des différentes oscillations, à savoir, les termes  $u_{n,k}$ .

Puis, dans un second temps, il s'agit de justifier que la solution approchée fournie par l'ansatz tronqué (II.14) est effectivement une solution approchée de la solution exacte  $u^\varepsilon$ .

Les intérêts d'un tel développement sont alors multiples. D'une part, du moins dans le cas linéaire, il est à noter (ceci est même assez remarquable dans notre opinion) que les équations qui régissent les amplitudes  $u_{n,k}$  sont bien plus simples que le système régissant la solution  $u^\varepsilon$ . En effet, comme on va le voir, il s'agit essentiellement d'équations de transport, donc pouvant être résolues de façon explicite. Ces dernières pouvant, toutefois, faire intervenir des dérivées de la donnée  $g$  ainsi que certains objets purement algébriques qui sont toujours explicitement calculables. Par conséquent, un développement d'optique géométrique fournit une solution approchée on ne peut plus calculable, et donc implémentable en pratique.

D'autre part, les développements d'optique géométrique permettent aussi de mettre la lumière sur certaines propriétés qualitatives de la solution exacte  $u^\varepsilon$ . En effet, si l'on veut montrer que la solution exacte,  $u^\varepsilon$  vérifie une certaine propriété, on va chercher à montrer que cette propriété est vraie pour son développement d'optique géométrique. Cela est souvent bien plus aisé car comme on l'a indiqué on connaît le développement d'optique géométrique explicitement. Puis, si l'on dispose d'une convergence assez fine de la solution d'optique géométrique vers la solution exacte, dans un cadre fonctionnel assez fin, on peut en déduire que la solution exacte  $u^\varepsilon$  vérifie alors nécessairement la propriété ciblée par un raisonnement par l'absurde.

A titre d'exemples de telles propriétés qui peuvent être montrées sur le développement d'optique géométrique, on peut, par exemple, citer la propriété phare des équations hyperboliques, de propagation à vitesse finie de l'information. Cette dernière est vérifiée par l'approximation d'optique géométrique car, elle est construite comme une somme de solutions d'équations de transport.

Mais, encore pour établir l'optimalité des pertes de dérivées dans les estimations (II.10), (II.11) et (II.12), comme on le verra plus en détails à la fin de ce paragraphe.

On conclut ce discours introductif en donnant les grandes lignes qui se dégagent dans la construction du développement d'optique géométrique pour le problème (II.13). Ceci est motivé par le fait que c'est un excellent travail préparatoire. En effet il contient tous les ingrédients principaux pour les constructions dans les géométries plus complexes, qui seront étudiées dans la suite de ce mémoire (voir Paragraphes II.3.3, II.4.4).

Dans ce qui suit, on expose la construction du développement d'optique géométrique pour le cas d'une fréquence de bord  $(i\mathcal{T}, \underline{\eta}) \in \mathcal{H}$ . Ceci induira, en particulier, que les amplitudes  $u_{n,k}$  vérifient alors toutes le même type d'équation. Les constructions pour les autres types de fréquence de bord peuvent être trouvées comme un cas particulier du travail pour une phase de *glancing*, dans le cas semi-linéaire de Williams [Williams, 1996] ensuite généralisé, à tous ordres, dans [Williams, 2000] (voir aussi [Hernandez, 2015]). Pour le traitement des modes elliptiques ou mixtes, on se doit aussi de mentionner les travaux de Marcou [Marcou, 2010] et de Lescarret [Lescarret, 2007].

On injecte l'ansatz (II.14) dans l'équation d'évolution de (II.13). On regroupe les différentes amplitudes en termes de puissance de  $\varepsilon$  et en supposant une propriété d'indépendance linéaire des phases  $\varphi_k$ , on est amené à considérer la cascade d'équations :

$$\begin{cases} \mathcal{L}(d\varphi_k)u_{0,k} = 0 & \forall k \in \mathcal{K}, \text{ (pour le terme en } \varepsilon^{-1}), \\ i\mathcal{L}(d\varphi_k)u_{n+1,k} + L(\partial)u_{n,k} = 0 & \forall k \in \mathcal{K}, \forall n \in \mathbb{N}, \text{ (pour le terme en } \varepsilon^n), \end{cases} \quad (\text{II.15})$$

où l'on a introduit le symbole  $\mathcal{L}$  de  $L(\partial)$  défini par : pour tout  $\xi \in \mathbb{R}^{1+d}$ ,  $\mathcal{L}(\xi) := \xi_0 I + \sum_{j=1}^d \xi_j A_j$ .

On s'intéresse, dans un premier temps, à la première équation de (II.15). Elle est satisfaite, si et seulement si, l'amplitude  $u_{0,k}$  est dans le noyau de  $\mathcal{L}(d\varphi_k)$ . Encore faut il que ce noyau ne soit pas trivial. C'est pourquoi, dans la suite, on choisira les phases  $\varphi_k$ , de telle sorte que ce noyau soit non trivial. Cela revient à choisir les paramètres  $\xi_k$ , comme les racines dans la variable  $\xi$ , de la relation de dispersion

$$\det \mathcal{L}(\mathcal{T}, \underline{\eta}, \xi) = 0. \quad (\text{II.16})$$

A ce stade, on a déterminé les fonctions de phase  $\varphi_k$ , ainsi que leur ensemble de paramétrisation  $\mathcal{K}$ . Si l'on définit  $\Pi_k$ , le projecteur sur  $\ker \mathcal{L}(d\varphi_k)$ , le long de son image, on retrouve alors la condition, bien connue en optique géométrique, de polarisation de l'amplitude principale  $\Pi_k u_{0,k} = u_{0,k}$ .

La seconde équation de (II.15) (pour  $n = 0$ ) devient donc, après composition par  $\Pi_k$ , en utilisant la condition de polarisation précédente :

$$\Pi_k L(\partial) \Pi_k u_{0,k} = 0,$$

et l'on se retrouve en position d'appliquer l'un des résultats cruciaux en optique géométrique, à savoir le lemme de Lax [Lax, 1957]. Il affirme que l'opérateur précédent est un opérateur de transport à la vitesse dite de groupe.

**Lemme II.2.2 (Lax [Lax, 1957])** *Sous une hypothèse de régularité des solutions de l'équation eikonale (II.16), pour tout  $k \in \mathcal{K}$ , on a*

$$\Pi_k L(\partial) \Pi_k = (\partial_t + \mathbf{v}_k \cdot \nabla_x) \Pi_k,$$

où  $\mathbf{v}_k$  désigne la vitesse de groupe associée à la phase  $\varphi_k$ .

Par conséquent, on s'est ramené à résoudre les équations de transport (découplées)

$$(\partial_t + \mathbf{v}_k \cdot \nabla_x) u_{0,k} = 0, \quad (\text{II.17})$$

et c'est à ce stade de l'analyse que l'on voit apparaître l'influence de la géométrie du demi-espace et l'importance de la condition de Kreiss-Lopatinskii uniforme. En effet, selon le signe de la dernière composante de la vitesse de groupe,  $\mathbf{v}_{k,d}$ , le comportement de l'équation de transport (II.17) change drastiquement. On a la dichotomie suivante<sup>9</sup> :

<sup>9</sup>Le cas  $\mathbf{v}_{k,d} = 0$  correspondant à une fréquence de bord de *glancing*. Ce dernier est exclu par choix de la phase au bord  $\psi$ . Il est néanmoins considéré dans les travaux de Williams [Williams, 1996] et [Williams, 2000].

- Si  $\mathbf{v}_{k,d} < 0$ , le transport se fait "de la droite vers la gauche". Il est qualifié de sortant. On note  $\mathcal{O}$ , l'ensemble des indices sortants. Pour de tels indices, la valeur de la trace  $u_{0,k}|_{x_d=0}$  est obtenue par résolution de l'équation de transport. En particulier, cette dernière ne nécessite pas de condition de bord. On a donc à résoudre une équation de transport homogène dont la solution est automatiquement nulle.
- Si  $\mathbf{v}_{k,d} > 0$ , le transport se fait "de la gauche vers la droite". Il est qualifié d'entrant. On note  $\mathcal{I}$ , l'ensemble des indices entrants. Pour de tels indices, le terme source sur le bord  $\{x_d = 0\}$  induit un phénomène de transport à l'intérieur. Pour le déterminer, une condition de bord est requise et on a donc besoin d'une condition de bord en  $\{x_d = 0\}$ .

Pour expliciter ces conditions de bord, on doit d'abord injecter l'ansatz (II.14) dans la condition de bord de (II.13) et découpler les différents termes. On obtient

$$\forall n \in \mathbb{N}, B \left[ \sum_{k \in \mathcal{I}} u_{n,k} \right]_{|x_d=0} = \delta_{n,0} g - B \left[ \sum_{\substack{k \in \mathcal{O} \\ =0}} u_{n,k} \right]_{|x_d=0}, \quad (\text{II.18})$$

où  $\delta_{\cdot,\cdot}$  désigne le symbole de Kronecker.

Et donc, en particulier, pour  $n = 0$ , on obtient, grâce à la condition de polarisation, la condition de bord couplée

$$B \left[ \sum_{k \in \mathcal{I}} \Pi_k u_{0,k} \right]_{|x_d=0} = g. \quad (\text{II.19})$$

La seule façon de déterminer, à partir de (II.19), les valeurs des traces  $u_{0,k}|_{x_d=0}$  est, pour ainsi dire, d'être en mesure d'inverser la matrice  $B$ . Cette matrice n'est clairement pas inversible. Toutefois on rappelle que la condition de Kreiss-Lopatinskii uniforme n'est rien autre qu'imposer que la restriction de  $B$  au sous-espace stable  $\mathbf{E}^s(\zeta)$  soit inversible. Or, il s'avère ici que l'on a (on renvoie par exemple à [Coulombel and Guès, 2010] pour une démonstration) l'égalité ensembliste  $\mathbf{E}^s(i\tau, \eta) = \bigoplus_{k \in \mathcal{I}} \ker \mathcal{L}(d\varphi_k)$ . En d'autres termes, on peut inverser  $B$  pour obtenir les valeurs des traces recherchées à savoir

$$\forall k \in \mathcal{I}, \Pi_k u_{0,k}|_{x_d=0} = \Pi_k \phi(\underline{\zeta}) g,$$

où l'on a posé  $\phi(\underline{\zeta}) := B|_{\mathbf{E}^s(i\tau, \eta)}^{-1}$ , l'inverse donnée par la condition de Kreiss-Lopatinskii uniforme.

Par conséquent, les amplitudes principales dans le développement d'optique géométrique sont alors déterminées en intégrant le long des caractéristiques les équations de transport

$$\forall k \in \mathcal{O}, \begin{cases} (\partial_t + \mathbf{v}_k \cdot \nabla_x) u_{0,k} = 0 & \text{dans } \Omega, \\ u_{0,k}|_{t \leq 0} = 0 & \text{sur } \Gamma, \end{cases} \quad (\text{II.20})$$

et

$$\forall k \in \mathcal{I}, \begin{cases} (\partial_t + \mathbf{v}_k \cdot \nabla_x) u_{0,k} = 0 & \text{dans } \Omega, \\ u_{0,k}|_{x_d=0} = \Pi_k \phi(\underline{\zeta}) \left( g - B \sum_{k \in \mathcal{O}} u_{0,k}|_{x_d=0} \right) & \text{sur } \partial\Omega, \\ u_{0,k}|_{t \leq 0} = 0 & \text{sur } \Gamma. \end{cases} \quad (\text{II.21})$$

Comme précédemment indiqué, le fait que l'on soit en présence d'une équation homogène dans (II.20) implique automatiquement que les amplitudes sortantes sont nulles. Ce n'est toutefois pas le cas dans le cas général, si l'équation de transport (II.20) a un terme source non nul. Dans un tel cadre, le point capital est alors que les équations sur les modes sortants (II.20) peuvent être résolues indépendamment (et donc en particulier avant) celles déterminant les modes entrants (II.21).

Une fois cette résolution effectuée, on connaît alors le terme source dans la condition de bord de (II.21). On peut donc déterminer les modes entrants, ce qui conclut la détermination de l'amplitude principale du développement d'optique géométrique.

Comme on le verra dans la suite de ce mémoire, notamment aux Paragraphes II.3.2 et II.4.3, ce point assurant la déterminations des phases sortantes avant et indépendamment des phases entrantes n'est plus vrai si l'on se place dans des géométries plus élaborées. C'est à ce stade que naîtra le phénomène d'auto-interaction des phases.

Cette remarque anodine est, il nous semble, au coeur de la construction des développements d'optique géométrique pour des problèmes aux limites, qui, à notre avis, nécessite toujours les deux ingrédients clés suivants<sup>10</sup> :

- L'identification d'une équation de la cascade que l'on est en mesure de résoudre avant toutes les autres (dans le cas précédent il s'agit de l'équation (II.20)). Résolution qui sert d'initialisation à la résolution totale.
- Un ordre de résolution qui permet ensuite de résoudre les autres équations de la cascade (II.15), en fonction de leur type, en se basant sur la connaissance de la solution de l'équation d'initialisation.

Dans l'exemple précédent, l'ordre est alors le suivant : sortant puis entrant. Il s'agit d'un ordre on ne peut plus simple. Comme on va le voir dans la suite (voir Paragraphes II.4.3 et II.4.4) lorsque l'on considère des géométries plus évoluées que le demi-espace, l'ordre précédemment mentionné peut s'avérer devenir plus complexe et peut-être moins naturel de prime abord.

Toutefois, il est important de garder en mémoire que, bien que les choses puissent sembler devenir parfois (très) techniques, les deux piliers de nos constructions, à savoir l'initialisation et la détermination d'un ordre de résolution, resteront toujours exactement les mêmes.

Une fois que l'on a déterminé les amplitudes principales dans le développement d'optique géométrique, on peut alors, en utilisant un algorithme bien éprouvé et rodé (voir par exemple [Rauch, 1999]) déterminer autant de correcteurs  $u_{n,k}$ ,  $n \geq 1$ , que l'on veut. On ne décrira pas ici spécifiquement cette méthode. On insiste sur le fait que la construction de tels correcteurs nécessite essentiellement uniquement de calculer des dérivées des amplitudes précédentes, ainsi que des pseudo-inverses de l'opérateur  $\mathcal{L}(d\varphi_k)$ . Par conséquent, la construction des correcteurs est on ne peut plus explicitable.

Le problème étant supposé fortement bien posé, on montre alors facilement par une analyse d'erreur classique que la solution approchée que l'on a construite (II.14) est effectivement une solution approchée de la solution exacte de (II.13). On résume les résultats de ce paragraphe dans le théorème suivant :

**Théorème II.2.3** *On se place sous les hypothèses d'hyperbolicité idoïne, de non caractéristicité du bord et sous la condition de Kreiss-Lopatinskiï uniforme. Pour simplifier, supposons que la fréquence dans le terme source de bord de (II.13) soit hyperbolique. Alors, pour tout  $n \in \mathbb{N}$ , les cascades d'équations (II.15) et (II.18) admettent une unique solution  $(u_{n,k})_{n \in \mathbb{N}, k \in \mathcal{I} \cup \mathcal{O}} \subset H^\infty(\Omega)$ . Cette solution vérifie de plus que, pour tout  $n \in \mathbb{N}$ , et  $k \in \mathcal{I} \cup \mathcal{O}$ ,  $u_{n,k}|_{t \leq 0} = 0$ .*

Soit  $N_0 \geq 0$ , si l'on définit le développement d'optique géométrique tronqué,

$$u_{app,N_0}^\varepsilon := \sum_{n=0}^{N_0} \varepsilon^n \sum_{k \in \mathcal{I} \cup \mathcal{O}} e^{i\varepsilon\varphi_k} u_{n,k},$$

alors on a l'estimation d'erreur<sup>11</sup>

$$\|u^\varepsilon - u_{app,N_0}^\varepsilon\|_{L^2(\Omega_T)} \leq C_T \varepsilon^{N_0+1},$$

<sup>10</sup>On paraphrase ici les propos de Coulombel qui parle de résolution d'un système triangulaire. Savoir si ce système est inférieur ou supérieur fait toujours l'objet d'un débat avec Olivier Pierre.

<sup>11</sup>Utilisant un résultat de persistance de la régularité pour le problème aux limites (II.1), *i.e.* un résultat analogue au Théorème II.2.2 dans les espaces de Sobolev, l'estimation d'erreur précédente passe en norme Sobolev et en particulier en norme  $L^\infty$ .

où  $C_T$  ne dépend pas de  $\varepsilon$ .

Ce théorème établit donc l'existence d'un développement de type optique géométrique et justifie que ce développement est une approximation de la solution exacte,  $u^\varepsilon$ , à tous ordres.

### Une application des développements d'optique géométrique, l'optimalité des pertes de dérivées

Dans ce paragraphe, on donne une application des développements d'optique géométrique pour l'étude de propriétés qualitatives de la solution de problèmes aux limites. Plus précisément, il sera question de l'optimalité des pertes de dérivées de la littérature pour les problèmes faiblement bien posés. Ce paragraphe reprend donc les résultats des articles [Coulombel and Guès, 2010], [Benoit, 2014] et [Benoit and Loyer, 2023].

On décrit donc dans ce qui suit comment la méthode de résolution du paragraphe précédent doit être adaptée pour traiter des problèmes pour lesquels la condition de Kreiss-Lopatinskii uniforme dégénère. On utilise alors ensuite le développement d'optique géométrique pour établir que les pertes de dérivées démontrées dans la littérature s'avèrent être optimales pour trois des quatre types de dégénérescence possibles.

Dans le dernier cas, celui de la dégénérescence en zone de *glancing*, comme précédemment indiqué, on ne dispose pas d'une telle estimation d'énergie. Par conséquent, on ne pourra pas conclure à l'optimalité. Toutefois, l'application de la méthode précédente nous donne de bons candidats sur ce que pourraient être les pertes optimales.

L'idée générale dans tout ce qui suit, qui est initialement dûe à [Coulombel and Guès, 2010] (voir aussi [Majda, 1988] et [Majda, 1983]), est qu'une perte de régularité de la solution exacte  $u^\varepsilon$  du problème aux limites (II.13) doit se lire sur son développement d'optique comme un décalage d'amplitude. Dans le cas fortement bien posé, pour un terme source en  $O(1)$ , on avait alors un développement d'optique géométrique de même amplitude.

Dans le cas faiblement bien posé, intéressons nous, par exemple, aux pertes de dérivées le long du bord. On postule qu'un terme source régulier et de faible amplitude donc se comportant en  $O(\varepsilon^\alpha)$ , avec  $\alpha > 0$ , va être amplifié et va donner naissance à un développement d'optique géométrique en  $O(1)$ . Puisque dans les estimations d'énergie de la littérature, on rencontre une perte d'une dérivée le long du bord, le bon candidat sera ici  $\alpha = 1$ .

On illustre alors avec plus de détails quelle est l'influence de la dégénérescence de la condition de Kreiss-Lopatinskii uniforme sur la détermination des amplitudes du développement. L'ansatz (II.14) est inchangé. En effet, on ne s'attend pas à ce que la dégénérescence de la condition de Kreiss-Lopatinskii uniforme induise de nouvelles phases. En particulier, on a toujours la cascade d'équations à l'intérieur (II.15). Celle sur le bord est alors modifiée en

$$\forall n \in \mathbb{N}, B \left[ \sum_{k \in \mathcal{I}} u_{n,k} \right]_{|x_d=0} = \delta_{n,1} g - B \left[ \sum_{k \in \mathcal{O}} u_{n,k} \right]_{|x_d=0}. \quad (\text{II.22})$$

En particulier, l'équation, pour  $n = 0$ , s'écrit (on peut facilement justifier que les amplitudes sortantes sont nulles)

$$B \left[ \sum_{k \in \mathcal{I}} u_{0,k} \right]_{|x_d=0} = 0. \quad (\text{II.23})$$

Equation pour laquelle nous ne sommes plus en mesure, comme cela a été fait au paragraphe précédent, d'inverser la matrice  $B$  pour en déduire la nullité des traces. On peut seulement en déduire la valeur des

traces au bord modulo un élément de  $\ker B \cap \mathbf{E}^s(\zeta) \neq \{0\}$ , mais qui est toutefois un sous-espace vectoriel de dimension un si l'on se place dans le cadre d'une dégénérescence dictée par la Définition II.2.4.

Pour simplifier l'exposition, et sans grande perte de généralité, supposons que  $\#\mathcal{S} = 1$ . Ainsi, (II.22) revient à imposer que pour  $k \in \mathcal{S} := \{\underline{k}\}$

$$u_{0,\underline{k}|_{x_d=0}} = \alpha_0(t, x')e, \text{ où } e \text{ est un générateur de } \ker B \cap \mathbf{E}^s(\zeta).$$

L'amplitude principale est donc totalement déterminée par la connaissance de son amplitude au bord  $\alpha_0$ , quantité pour laquelle on a besoin d'une nouvelle équation pour lever l'indétermination.

Afin de déterminer une telle équation, ceci indépendamment du type de dégénérescence que l'on est en train de considérer, la méthode est la même dans [Coulombel and Guès, 2010], [Benoit, 2014], [Benoit and Loyer, 2023] ou encore [Marcou, 2010]. On considère l'unique condition de bord qui fait intervenir  $u_{0,k}$ . Il s'agit de la condition de bord d'ordre un. Ceci n'est pas forcément clair de prime abord. Toutefois comme on l'a mentionné précédemment le premier correcteur  $u_{1,k}$  dépend de l'amplitude précédente. On ne décrira pas plus dans le détail cette dépendance. On se contentera ici de mentionner que la partie non polarisée du premier correcteur dépend des dérivées de l'amplitude  $u_{0,k}$  (la partie polarisée étant invisible dans la condition de bord) et d'un pseudo-inverse.

Par conséquent, dans les quatre cas de dégénérescence possibles, l'équation (II.22) peut en fait se réécrire sous la forme (on renvoie le lecteur intéressé à [Coulombel and Guès, 2010], [Benoit, 2014] ou [Benoit and Loyer, 2023] pour les réécritures explicites):

$$\mathcal{L}\alpha_0 = \mathcal{F}(g, (u_{0,k})_{k \in \mathcal{O}}, (u_{1,k})_{k \in \mathcal{O}}), \quad (\text{II.24})$$

où  $\mathcal{F}(\cdot, \cdot, \cdot)$  est une transformation explicite et où  $\mathcal{L}$  est un opérateur agissant sur l'inconnue  $\alpha_0$ . Par conséquent,  $\alpha_0$  est déterminée, si l'opérateur au bord  $\mathcal{L}$  est inversible dans un espace fonctionnel idoine.

Le point commun des dégénérescences de la condition de Kreiss-Lopatinskii uniforme dans les zones elliptique, mixte et hyperbolique est que l'on peut montrer (on renvoie à [Coulombel and Guès, 2010] et [Marcou, 2010]) que l'opérateur  $\mathcal{L}$  est un opérateur de transport. Par conséquent, on peut aisément résoudre (II.24) en intégrant le long des caractéristiques et la construction de l'amplitude principale est complète. Dans le cas d'une dégénérescence de la condition de Kreiss-Lopatinskii uniforme en zone de *glancing*, on a besoin de considérer un développement d'optique géométrique à double échelle, c'est-à-dire typiquement une échelle  $\frac{1}{\varepsilon}$  pour les modes elliptiques et hyperboliques et typiquement une échelle  $\frac{1}{\sqrt{\varepsilon}}$  pour les modes de *glancing*.

La situation se trouve légèrement modifiée car les modes de *glancing* vérifient des équations de type Schrödinger (voir [Williams, 2000]) et ainsi les phénomènes de transport ne sont plus prépondérants. On peut toutefois montrer que l'opérateur  $\mathcal{L}$  est un multiplicateur de Fourier à symbole non nul. Par conséquent, on peut, aussi dans un tel cadre, en inversant le symbole, facilement inverser cet opérateur et lever l'indétermination sur  $\alpha_0$ .

Une fois le développement d'optique géométrique construit on peut alors facilement l'utiliser pour venir saturer les estimations d'énergie dont on dispose dans la littérature pour montrer leur optimalité. Ceci est fait dans [Coulombel and Guès, 2010] et [Benoit, 2014] pour les trois premières dégénérescences.

L'idée est la suivante : on procède par l'absurde en supposant que l'estimation d'énergie n'est pas optimale et donc qu'elle pourrait être établie pour un  $s > \underline{s}$  ( $\underline{s}$  étant la perte donnée par l'estimation de la littérature [Coulombel, 2002a], [Sablé-Tougeron, 1988] ou [Coulombel, 2002b]).

L'analyse d'erreur, dans laquelle on a un léger gain sur l'exposant de Sobolev,  $s$ , indique alors que le développement d'optique géométrique est une bonne approximation de la solution exacte  $u^\varepsilon$ . Toutefois en raison du *scaling* en  $\varepsilon$  des termes de bord, l'estimation d'énergie implique aussi que la solution exacte  $u^\varepsilon$  est en  $O(\varepsilon^\alpha)$ ,  $\alpha > 0$ , dans l'espace fonctionnel idoine.

L'inégalité triangulaire induit alors que le développement d'optique géométrique devrait être en  $O(\varepsilon^\alpha)$ , et en particulier que, son amplitude principale devrait s'annuler. Ceci entre en contradiction avec le fait qu'un terme source de taille  $\varepsilon$  vient allumer une amplitude principale par le biais de l'équation (II.24).

On conclut la discussion précédente par le théorème suivant qui donne la classification la plus complète (à l'heure actuelle) des pertes de dérivées pour les problèmes aux limites faiblement bien posés dans le demi-espace.

**Théorème II.2.4** *[[Coulombel and Guès, 2010],[Benoit, 2014],[Benoit and Loyer, 2023]] En se plaçant sous les hypothèses d'hyperbolicité idoïne<sup>12</sup>, de non caractéristicité du bord, on suppose aussi que la condition de Kreiss-Lopatinskii uniforme dégénère de façon convenable. On a alors la classification suivante des problèmes aux limites faiblement bien posés :*

- *Si la condition de Kreiss-Lopatinskii uniforme vient à dégénérer en zone elliptique, alors l'estimation d'énergie de Sablé-Tougeron [Sablé-Tougeron, 1988] est optimale.*
- *Si la condition de Kreiss-Lopatinskii uniforme dégénère en zone mixte<sup>13</sup>, sous une hypothèse de structure assurant la propagation de l'instabilité (voir [Benoit, 2014] pour plus de détails), alors l'estimation d'énergie de Coulombel [Coulombel, 2002a] est optimale.*
- *Si la condition de Kreiss-Lopatinskii uniforme dégénère en zone hyperbolique, alors l'estimation d'énergie de [Coulombel, 2002b] est optimale.*
- *Si la condition de Kreiss-Lopatinskii uniforme dégénère en zone de glancing, alors on a le résultat suivant. Supposons qu'il existe  $s_1, s_2 > 0$ , tels que le problème aux limites (II.1) vérifie une estimation d'énergie de la forme*

$$\|u\|_{L^2(\Omega_T)}^2 \leq C_T \left( \|f\|_{L^2_{x_d}(H^s_{t,x})(]-\infty, T] \times \mathbb{R}^{d-1})}^2 + \|g\|_{H^{s_2}(\partial\Omega_T)}^2 \right).$$

*Alors on a l'alternative :*

- *Sous la même hypothèse de structure, qui assure la propagation des singularités internes, que dans [Benoit, 2014], on a  $s_1 \geq \frac{1}{2}$ .*
- *Si le problème n'admet pas de perte à l'intérieur, i.e.  $s_1 = 0$ , alors  $s_2 \geq \frac{1}{2}$ .*
- *Dans le cas général, on a  $s_1 \geq \frac{1}{2}$  ou  $s_2 \geq \frac{1}{2}$ .*

## II.3 Le problème dans la bande

Les résultats de cette section ont été publiés dans les quatre articles [Benoit, 2020a], [Benoit, 2020c], [Benoit, 2020b] et [Benoit, 2021]. Ils concernent l'étude de systèmes hyperboliques (linéaires) dans le même formalisme que celui rencontré dans la Section II.2. Toutefois, ces systèmes ne sont maintenant plus posés dans l'espace entier mais seulement dans une bande. L'une des variables de direction (disons  $x_d$ ) vit maintenant dans un intervalle. Sans perte de généralité, on fixera l'intervalle comment étant égal à  $]0, 1[$ .

<sup>12</sup>On parle ici typiquement d'hypothèse d'hyperbolicité à multiplicité constante.

<sup>13</sup>On entend ici que la dégénérescence se fait sur la composante hyperbolique de la zone mixte. Une dégénérescence sur la composante elliptique de la zone mixte n'a jamais été étudiée en détails mais devrait se dérouler exactement comme une dégénérescence en zone purement elliptique.

De tels problèmes s'écrivent donc sous la forme

$$\begin{cases} L(\partial)u = f & \text{pour } (t, x', x_d) \in \mathbb{R} \times \mathbb{R}^{d-1} \times ]0, 1[ := \Omega, \\ B_0 u|_{x_d=0} = g_0 & \text{pour } (t, x') \in \mathbb{R} \times \mathbb{R}^{d-1} := \partial\Omega_0, \\ B_1 u|_{x_d=1} = g_1 & \text{pour } (t, x') \in \mathbb{R} \times \mathbb{R}^{d-1} := \partial\Omega_1, \\ u|_{t \leq 0} = 0, & \text{pour } x \in \mathbb{R}^{d-1} \times ]0, 1[ := \Gamma, \end{cases} \quad (\text{II.25})$$

où l'opérateur hyperbolique  $L(\partial)$  est le même qu'à la Section II.2 et où les matrices de bords sont arbitraires. Elles imposent, néanmoins, le bon nombre de conditions de bord, à savoir  $B_0 \in \mathbf{M}_{p,N}(\mathbb{R})$  et  $B_1 \in \mathbf{M}_{N-p,N}(\mathbb{R})$ , où  $p$  désigne le nombre de valeurs propres (strictement) positives du coefficient normal  $A_d$  (on se place dans tout ce qui suit dans un cadre non caractéristique).

Avant de décrire plus avant les résultats que nous avons obtenus sur les problèmes de la forme (II.25), on va décrire les motivations qui nous ont poussé à étudier de telles équations. On s'intéressera plus particulièrement à la nouvelle géométrie.

Il est à noter, qu'à notre connaissance, la littérature ne comprend pas de résultats concernant l'étude de systèmes sous la forme (II.25). Ce fait peut probablement être expliqué par la raison suivante : pour traiter de tels problèmes un argument de partition de l'unité est suffisant.

En effet, pour étudier, par exemple, le caractère fortement bien posé du problème aux limites (II.25), on peut se ramener, en définissant une partition de l'unité idoine de l'intervalle  $[0, 1]$ , à étudier trois problèmes aux limites : un problème intérieur, un problème ne faisant intervenir que le bord gauche et un problème ne faisant intervenir que le bord droit.

En se plaçant alors sous des hypothèses garantissant le caractère fortement bien posé de chacun des trois problèmes (essentiellement la condition de Kreiss-Lopatinskii uniforme sur chacun des bords et l'hyperbolicité de l'opérateur), on conclut alors au caractère fortement bien posé du problème dans la bande.

La démonstration précédente à l'avantage certain de la simplicité. Toutefois, en raison de sa nature elle est sourde à la conversation que peuvent avoir les deux faces du domaine l'une avec l'autre. En effet, puisque l'opérateur  $L(\partial)$  est hyperbolique, ce dernier, sauf si on lui adjoint un terme d'amortissement drastique, a tendance à transporter l'information et ainsi les deux conditions de bord se parlent. On entend par là que l'on peut parfaitement concevoir une information initialement présente sur le bord  $\partial\Omega_0$ , transportée vers  $\partial\Omega_1$ , réfléchi contre cette face qui retourne ensuite vers la face  $\partial\Omega_0$ . Cette information y subit une nouvelle réflexion et ainsi de suite. Phénomène de "ping-pong", que l'argument de partition de l'unité est incapable de prendre en compte.

Ce qui nous a poussé à étudier le problème aux limites (II.25) est précisément d'essayer de comprendre ce phénomène. En effet, nous avons commencé, dans notre thèse [Benoit, 2015], par l'étude du problème dans le quart d'espace, problème qui sera décrit plus avant dans la Section II.4. Pour un tel problème, comprendre précisément ce dialogue entre les faces est ardu et c'est encore, à l'heure actuelle, assez mystérieux. Mais, cela constitue néanmoins, l'une des clés principales pour résoudre le problème dans un quart d'espace.

Ainsi, en passant du problème dans le quart d'espace à celui dans la bande nous l'avons simplifié en ceci que les deux faces interagissent *via* la même variable normale  $x_d$  (et non pas *via* deux variables  $x_1$  et  $x_2$  comme pour le quart d'espace). Cependant nous avons conservé la difficulté principale qui est l'existence de deux composantes dans la frontière du domaine de résolution. En un certain sens le problème dans la bande (II.25) peut donc être vu comme un *toy-model* de l'interaction des faces ou encore un travail préparatoire pour l'étude du problème dans un domaine à coin.



En effet, comme on va le voir dans la suite de cette section, en comparaison avec la Section II.4, le problème dans la bande (II.25) partage avec le problème dans le quart d'espace (II.39), les mêmes grandes étapes dans les démonstrations.

Cependant les objets mathématiques que l'on va être amené à manipuler dans le cas de la bande, bien que ces derniers aient leur total analogue pour le problème dans le quart d'espace, sont de nature beaucoup plus simple. Ils deviennent ainsi bien plus manipulables. On passe essentiellement de matrices (opérateurs de multiplication) dans le cas de la bande à des opérateurs intégraux de Fourier dans le cas du quart d'espace (voir les équations (II.30) et (II.43)).

Par conséquent, on sera à même de conclure dans le cas de la bande, là où, le cas du quart d'espace nécessite encore, à l'heure actuelle, du travail.

A notre avis, l'intérêt de la géométrie de la bande ne se limite pas à être seulement un *toy-model* pour le problème dans le quart d'espace<sup>14</sup>. En effet, de nombreux problèmes issus de la physique sont posés dans une telle géométrie. On pense par exemple, à la propagation d'une vague (régie par l'équation KdV par exemple) dans un canal voir par exemple [Laurent et al., 2010]), ou encore, la propagation de la lumière dans une fibre optique, ou plus généralement la propagation d'une onde dans un guide d'onde. De tels problèmes imposent, bien souvent, les mêmes conditions de bord sur les deux faces de la bande. Ainsi, notre étude donne un cadre où l'on peut venir prescrire des conditions différentes sur les deux faces. Un jour, nous aimerions nous pencher, plus spécifiquement, sur l'intérêt physique que peut apporter un tel formalisme mathématique.

Cette section est organisée de la manière suivante, le Paragraphe II.3.1 traite du caractère fortement bien posé optimal du problème (II.25) et analyse donc les résultats de l'article [Benoit, 2020a]. On y montre essentiellement que la caractéristique fortement bien posé optimal est équivalent à une nouvelle condition d'inversibilité.

Le Paragraphe II.3.2 reprend les résultats de l'article [Benoit, 2020c] et établit le développement d'optique géométrique du problème (II.25). On donne ainsi une illustration microlocale de la nouvelle condition rencontrée dans le paragraphe précédent.

Le Paragraphe II.3.3 revient sur la notion de problème faiblement bien posé dans un tel cadre. En reprenant les articles [Benoit, 2020b] et [Benoit, 2021], on montre, en particulier que de nouveaux types de problèmes faiblement bien posés peuvent apparaître.

Enfin, le Paragraphe II.3.4 dessine une conclusion et donne quelques perspectives de recherche à propos du problème dans la bande.

### II.3.1 Théorie bien posée

#### La condition de Kreiss-Lopatinskii uniforme n'est plus assez fine

Comme pointé dans l'introduction de cette section, imposer la condition de Kreiss-Lopatinskii uniforme donne une théorie bien posée simple, mais assez peu fine. Pour illustrer ceci tâchons d'établir l'estimation d'énergie *a priori*  $L^2$  de (II.25).

On veut la même estimation que pour le demi-espace, à ceci près que l'on demande maintenant un contrôle des deux composantes de la trace, à savoir, des termes  $\|u_{|x_d=0}\|_{L^2(\partial\Omega_0)}$  et  $\|u_{|x_d=1}\|_{L^2(\partial\Omega_1)}$ .

Si l'on introduit une fonction régulière  $\chi_0$  (resp.  $\chi_1$ ) valant 1 proche de zéro (resp. un) et 0 proche de un (resp. zéro). En notant  $\chi_{1/2} = 1 - \chi_0 - \chi_1$  et en introduisant  $u_j := \chi_j u$  où  $u$  désigne une solution de (II.25), on montre alors facilement que l'on a l'estimation d'énergie intérieure

$$\gamma \|u_{1/2}\|_{L^2_\gamma(\Omega)}^2 \leq \frac{C}{\gamma} \left[ \|f\|_{L^2_\gamma(\Omega)}^2 + \|u\|_{L^2_\gamma(\Omega)}^2 \right],$$

<sup>14</sup>Le rapporteur de [Benoit, 2020a], maintenant en retraite, a qualifié cette problématique de question qui aurait dû être posée depuis un certain temps déjà.

qui est conséquence de l'hyperbolicité.

On montre, ensuite, que l'on a les deux estimations suivantes pour les problèmes à un seul bord

$$\text{pour } j \in \{0, 1\}, \quad \gamma \|u_j\|_{L^2_\gamma(\Omega)}^2 + \|u_{|x_j=0}\|_{L^2_\gamma(\partial\Omega_j)}^2 \leq C \left( \frac{1}{\gamma} \left[ \|f\|_{L^2_\gamma(\Omega)}^2 + \|u\|_{L^2_\gamma(\Omega)}^2 \right] + \|g_j\|_{L^2_\gamma(\partial\Omega_j)}^2 \right),$$

conséquences de la condition de Kreiss-Lopatinskii uniforme sur chaque face.

Ainsi, par l'inégalité triangulaire on peut venir estimer  $\|u\|_{L^2(\Omega)}^2$  grâce aux  $\|u_j\|_{L^2_\gamma(\Omega)}^2$  et, si on prend  $\gamma$  assez grand, absorber le membre de droite des équations précédentes dans le membre de gauche ce qui conduit à une estimation d'énergie *a priori*  $L^2$ .

On obtient donc l'estimation d'énergie, il existe  $C, \gamma_0 > 0$  tels que pour tout  $\gamma \geq \gamma_0$ ,

$$\gamma \|u\|_{L^2_\gamma(\Omega)}^2 + \|u_{|x_1=0}\|_{L^2_\gamma(\partial\Omega_1)}^2 + \|u_{|x_2=0}\|_{L^2_\gamma(\partial\Omega_2)}^2 \leq C \left( \frac{1}{\gamma} \|f\|_{L^2_\gamma(\Omega)}^2 + \|g_1\|_{L^2_\gamma(\partial\Omega_1)}^2 + \|g_2\|_{L^2_\gamma(\partial\Omega_2)}^2 \right). \quad (\text{II.26})$$

Ce qui est toutefois assez peu satisfaisant dans cette estimation est que, puisqu'elle ne tient pas pour tout  $\gamma > 0$ , elle implique qu'*a priori* la solution peut avoir une croissance exponentielle non triviale par rapport à la variable temporelle<sup>15</sup>

Considérons un exemple simple constitué de deux équations de transports découplées l'une allant vers la gauche et l'autre vers la droite, le couplage de ces deux équations ayant lieu au niveau des conditions de bord. On peut calculer explicitement la solution d'un tel exemple. Elle dépend principalement des coefficients de réflexion  $c_0, c_1 \in \mathbb{R}$  lors d'un rebond contre la face  $\partial\Omega_0$  et  $\partial\Omega_1$ . On peut voir que la solution se comporte alors essentiellement (par rapport au temps) comme  $(c_0 c_1)^t$  et on a alors deux comportements possibles :

- Ou bien, globalement lors des réflexions, les conditions de bord font augmenter la norme, à savoir  $|c_0 c_1| > 1$ , alors, dans ce cas, on a bien une solution qui va avoir une croissance exponentielle non triviale en temps et l'estimation d'énergie (II.26) est admissible.
- Ou bien, globalement les deux conditions de bords diminuent (ou conservent) la norme à savoir  $|c_0 c_1| \leq 1$ . Auquel cas, la norme  $L^2$  de la solution viendra diminuer avec le temps (ou sera conservée). Dans un tel cas de figure, l'estimation d'énergie (II.26) est bien loin d'être optimale.

Nous avons donc voulu, pour comprendre si une telle croissance exponentielle de la solution par rapport au temps, issue de l'interaction des deux conditions de bord, avait lieu, caractériser les problèmes dans la bande (II.25) qui admettent une estimation d'énergie de la forme (II.26) mais avec  $\gamma_0 = 0$ . Pour ce faire, en plus de la condition de Kreiss-Lopatinskii uniforme sur chaque face, on a été amené dans [Benoit, 2020a] à imposer une nouvelle condition qui détermine si, oui ou non, on peut venir choisir  $\gamma_0 = 0$  dans (II.26).

Cette nouvelle condition est obtenue en adaptant, à la nouvelle géométrie de la bande, des techniques développées par [Osher, 1973] pour travailler dans le quart d'espace. Ainsi, non contente d'être intéressante en elle-même car elle constitue la clé de voûte de l'étude des problèmes dans la bande, l'exposition de cette condition sera aussi un excellent travail préparatoire à la Section II.4.

<sup>15</sup>Un tel comportement était absent pour le problème dans le demi-espace pour lequel, l'analyse de [Kreiss, 1970], en l'absence de terme d'ordre zéro, montre que la solution admet, au pire, une croissance sous-exponentielle par rapport à la variable temporelle. Ceci n'est plus le cas si l'on considère à la place de  $L(\partial)$ , l'opérateur "amorti"  $L(\partial) + D$ ,  $D \in \mathbf{M}_{N \times N}(\mathbb{R})$  où  $D$  n'a pas de propriétés de signe, ce qui implique que l'on doit *a priori* pouvoir prendre  $\gamma$  assez grand pour manger le terme  $\langle Du, u \rangle$  dans le bilan d'énergie.

## Une nouvelle condition liée à l'auto-interaction

Ce paragraphe expose principalement les résultats de l'article [Benoit, 2020a] qui établissent le caractère fortement bien posé "sous exponentiel" du problème dans la bande (II.25). Le coeur de l'analyse se base sur une approche de type *symétriseur de Kreiss*. Nous avons ici adapté du quart d'espace vers la bande, les symétriseurs utilisés dans [Osher, 1973] pour établir l'estimation d'énergie *a priori*. La construction de tels symétriseurs nécessitant la nouvelle condition d'inversibilité que l'on a mentionnée dans le paragraphe précédent.

L'obtention de cette condition et son caractère (du moins "partiellement") nécessaire<sup>16</sup> est assez simple à décrire.

Pour commencer décrivons le caractère fortement bien posé qui va nous intéresser dans la suite

**Définition II.3.1 (Problème bien posé dans la bande)** *On dit que*

- le problème (II.25) est fortement bien posé si il existe  $\gamma_0 \geq 0$ , tel que pour tout  $\gamma > \gamma_0$  et pour tous termes sources  $f \in L^2_\gamma(\Omega)$ ,  $(g_0, g_1) \in L^2_\gamma(\partial\Omega_0) \times L^2_\gamma(\partial\Omega_1)$ , le problème (II.25) admet une unique solution  $u \in L^2_\gamma(\Omega)$ , à traces  $u|_{x_d=j} \in L^2_\gamma(\partial\Omega_j)$ ,  $j \in \{0, 1\}$  qui vérifie de surcroît l'estimation d'énergie suivante : il existe  $C > 0$  telle que pour tout  $\gamma > \gamma_0$ ,

$$\gamma \|u\|_{L^2(\Omega)}^2 + \|u|_{x_d=0}\|_{L^2_\gamma(\partial\Omega_0)}^2 + \|u|_{x_d=1}\|_{L^2_\gamma(\partial\Omega_1)}^2 \leq C \left( \frac{1}{\gamma} \|f\|_{L^2_\gamma(\Omega)}^2 + \|g_0\|_{L^2_\gamma(\partial\Omega_0)}^2 + \|g_1\|_{L^2_\gamma(\partial\Omega_1)}^2 \right).$$

- Le problème est optimalement (ou sous-exponentiellement) fortement bien posé si l'on peut prendre  $\gamma_0 = 0$  dans la définition précédente.

On suppose que le problème (II.25), pour lequel on conserve uniquement comme source non triviale le terme  $g_0$ , admet une solution (régulière)  $u$ . On considère l'extension de  $u$  par zéro pour  $x_d > 1$ . Cette extension, toujours notée  $u$ , vérifie le problème dans le demi-espace :

$$\begin{cases} L(\partial)u = -\delta_{x_d=1}A_d u|_{x_d=1} & \text{pour } (t, x) \in \mathbb{R} \times \mathbb{R}_+^d, \\ B_0 u|_{x_d=0} = g_0 & \text{pour } (t, x') \in \mathbb{R} \times \mathbb{R}^{d-1}, \\ u|_{t \leq 0} = 0 & \text{pour } x \in \mathbb{R}_+^d, \end{cases}$$

où  $\delta$  désigne la masse de Dirac.

Sous des hypothèses de structures classiques<sup>17</sup>, on peut réitérer l'analyse du demi-espace. Pour ce faire, on effectue une transformée de Laplace-Fourier  $(t, x') \rightsquigarrow \zeta := (\sigma, \eta) \in \mathbb{C}_+ \times \mathbb{R}^{d-1}$  de  $u$ . On obtient alors l'équation différentielle ordinaire

$$\begin{cases} \frac{d\hat{u}}{dx_d}(\zeta, x_d) = \mathcal{A}(\zeta)\hat{u}(\zeta, x_d) - \delta_{x_d=1}A_d\hat{u}|_{x_d=1}(\zeta) & \text{pour } x_d > 0, \\ B_0\hat{u}|_{x_d=0}(\zeta) = \hat{g}_0(\zeta) & \text{en } x_d = 0, \end{cases} \quad (\text{II.27})$$

et où la matrice résolvante  $\mathcal{A}(\zeta) := -A_d^{-1} \left( \sigma I + i \sum_{j=1}^{d-1} \eta_j A_j \right)$  admet de bonnes propriétés spectrales d'après [Hersh, 1963].

Tant que le paramètre de Laplace a une partie réelle non triviale, l'espace  $\mathbb{C}^N$  s'écrit alors comme la somme directe d'un sous-espace stable,  $\mathbf{E}^s(\zeta)$  et d'un sous-espace instable,  $\mathbf{E}^u(\zeta)$ . On notera  $\Pi^s(\zeta)$  et

<sup>16</sup>On précisera les caractères nécessaire et suffisant de la condition au Théorème II.3.1.

<sup>17</sup>On entend plus précisément par là une hypothèse d'hyperbolicité à multiplicité constante et une de non-caractérisité du bord.

$\Pi^u(\zeta)$  les projecteurs sur le sous-espace stable et le sous-espace instable respectivement, ce relativement à la décomposition précédente.

En utilisant la formule de Duhamel, on peut alors venir écrire  $\widehat{u}$ , la solution de (II.27), comme  $\widehat{u} = (\Pi^s \widehat{u})(\zeta) + (\Pi^u \widehat{u})(\zeta)$  avec :

$$\Pi^s(\zeta) \widehat{u}(\zeta, x_d) = e^{x_d \mathcal{A}(\zeta)} \Pi^s(\zeta) \widehat{u}|_{x_d=1}(\zeta) - \int_0^{x_d} \delta_{s=1} e^{(x_d-s) \mathcal{A}(\zeta)} ds \Pi^s(\zeta) \widehat{u}|_{x_d=1}(\zeta), \quad (\text{II.28})$$

$$\Pi^u(\zeta) \widehat{u}(\zeta, x_d) = \int_{x_d}^{\infty} \delta_{s=1} e^{(x_d-s) \mathcal{A}(\zeta)} ds \Pi^u(\zeta) \widehat{u}|_{x_d=1}(\zeta). \quad (\text{II.29})$$

Ces equations donnent, en particulier, les valeurs des parties stable et instable de la trace en  $\{x_d = 0\}$ . Toutefois, rappelons maintenant qu'au niveau de la trace, la condition de Kreiss-Lopatinskii uniforme permet d'exprimer la partie stable de la trace en fonction de la partie instable. Plus précisément, on a

$$\Pi^s(\zeta) \widehat{u}|_{x_d=0}(\zeta) = \phi_0(\zeta) g_0 - \phi_0(\zeta) B_0 \Pi^u(\zeta) \widehat{u}|_{x_d=0}(\zeta),$$

où  $\phi_0(\zeta) := B_{0|\mathbf{E}^s(\zeta)}^{-1}$  est l'inverse donné par la condition de Kreiss-Lopatinskii uniforme sur la face  $\partial\Omega_0$ . On vient utiliser cette équation dans (II.28) et (II.29), pour obtenir une relation donnant la trace sur  $\{x_d = 0\}$  en fonction de celle sur  $\{x_d = 1\}$ . Plus précisément, il vient

$$\widehat{u}|_{x_d=0}(\zeta) = \phi_0(\zeta) g_0 + \underbrace{(\phi_0(\zeta) B_0 - I) e^{-\mathcal{A}(\zeta)}}_{:= \overline{\mathbf{T}}_{1 \rightarrow 0}(\zeta)} \Pi^u(\zeta) \widehat{u}|_{x_d=1}(\zeta). \quad (\text{II.30})$$

Procéder de manière similaire, mais en étendant cette fois la solution par zéro pour  $x_d < 0$ , donne alors une relation reliant la trace  $\widehat{u}|_{x_d=1}$  en termes de celle de  $\widehat{u}|_{x_d=0}$ , de la forme

$$\widehat{u}|_{x_d=1}(\zeta) = \overline{\mathbf{T}}_{0 \rightarrow 1}(\zeta) \Pi^s(\zeta) \widehat{u}|_{x_d=0}(\zeta), \quad (\text{II.31})$$

où  $\overline{\mathbf{T}}_{0 \rightarrow 1}(\zeta)$  est défini de façon similaire à  $\overline{\mathbf{T}}_{1 \rightarrow 0}(\zeta)$ , à changement près des indices.

Combiner les équations (II.30) et (II.31) nous amène (après quelques manipulations algébriques pour faire disparaître les projecteurs) alors à considérer la condition suivante : si la trace  $\widehat{u}|_{x_d=0}(\zeta)$  est bien définie, alors elle doit vérifier l'équation de compatibilité

$$(I - \overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta)) \widehat{u}|_{x_d=0}(\zeta) = \phi_0(\zeta) g_0(\zeta), \quad (\text{II.32})$$

où l'on a posé  $\overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta) := \overline{\mathbf{T}}_{0 \rightarrow 1}(\zeta) \overline{\mathbf{T}}_{1 \rightarrow 0}(\zeta)$ . Le membre de droite de (II.32) étant arbitraire, il est alors naturel d'imposer l'inversibilité de l'opérateur de multiplication  $I - \overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta)$  sur  $L^2(\partial\Omega_0)$ .

En fait, lors des manipulations algébriques nous permettant d'ôter les projecteurs qui figurent dans les équations (II.30) et (II.31) (voir [[Benoit, 2020a]-Paragraphe 3.2]), on est amené à considérer une autre condition d'inversibilité et à préciser les espaces sur lesquels on impose les conditions d'inversibilité.

La nouvelle condition s'écrit sous la même forme que (II.32), mais pour un nouvel opérateur de multiplication  $\mathbf{T}_{0 \rightarrow 0}(\zeta)$ , défini de manière analogue à  $\overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta)$ , si ce n'est que dans les définitions des opérateurs  $\mathbf{T}_{1 \rightarrow 0}(\zeta)$  et  $\mathbf{T}_{0 \rightarrow 1}(\zeta)$ , qui composent  $\overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta)$ , nous n'avons plus les identités en facteur du terme exponentiel<sup>18</sup>.

Ainsi, en raison de la définition de l'inverse donné par la condition de Kreiss-Lopatinskii uniforme, l'opérateur  $\mathbf{T}_{0 \rightarrow 0}(\zeta)$  agit sur la partie de l'espace associée au sous-espace stable  $\mathbf{E}^s(\zeta)$ , là où, l'opérateur  $\overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta)$  agit sur le noyau de la condition de bord  $\ker B_0$ .

<sup>18</sup>Dans la présente discussion, la nouvelle condition sur  $I - \mathbf{T}_{0 \rightarrow 0}$  ne sera pas capitale. C'est pourquoi on se permet de ne pas définir proprement cet opérateur. On renvoie le lecteur intéressé à [[Benoit, 2020a]-Paragraphe 3.2]

En effet, on rappelle que l'on a, d'après la condition de Kreiss-Lopatinskii uniforme,  $\mathbb{C}^N := \ker B_0 \oplus \mathbf{E}^s(\zeta)$ , donc  $I - \phi_0(\zeta)B_0$  à ses valeurs dans  $\ker B_0$ , et à l'inverse  $\phi_0(\zeta)B_0$  à ses valeurs dans  $\mathbf{E}^s(\zeta)$ . Donc les deux opérateurs précédents sont, en un certain sens, complémentaires.

On obtient donc la condition (presque)<sup>19</sup> nécessaire pour assurer le caractère fortement bien posé optimal.

**Hypothèse II.3.1** *On se place sous l'hypothèse d'hyperbolicité à multiplicité constante, de non caractéristicité du bord et la condition de Kreiss-Lopatinskii uniforme sur chaque face. On suppose la propriété suivante :*

1. *Pour tout paramètre  $\zeta \in \Xi \setminus \Xi_0$ , l'opérateur  $I - \mathbf{T}_{0 \rightarrow 0}(\zeta)$  est un isomorphisme de  $L^2(\mathbf{E}^s(\zeta))$  vers lui-même. De plus, l'inverse  $(I - \mathbf{T}_{0 \rightarrow 0}(\zeta))^{-1}$  est uniformément borné par rapport à  $\zeta$ .*
2. *Pour tout paramètre  $\zeta \in \Xi \setminus \Xi_0$ , l'opérateur  $I - \overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta)$  est un isomorphisme de  $L^2(\ker B_0)$  vers lui-même. De plus, l'inverse  $(I - \overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta))^{-1}$  est uniformément borné par rapport à  $\zeta$ .*

Nous pensons qu'ainsi formulée, l'Hypothèse II.3.1 doit être comparée à la formulation de la condition de Kreiss-Loaptinskii uniforme :

$$\forall \zeta \in \Xi, v \in \mathbf{E}^s(\zeta), |v| \leq C|Bv|,$$

dans laquelle  $C$  est indépendante de  $\zeta \in \Xi \setminus \Xi_0$ . C'est-à-dire, la condition de Kreiss-Lopatinskii uniforme obtenue sans prolongation du sous-espace stable. Par conséquent, l'Hypothèse II.3.1 semble très difficile à établir en pratique. Il y a fort à parier que cette difficulté puisse se trouver surmontée en venant considérer une version "compactifiée" de l'Hypothèse II.3.1. On fera plus de commentaires sur ce point à la fin de cette section.

### Caractère suffisant de cette condition

On peut alors montrer le caractère suffisant de l'Hypothèse II.3.1 pour assurer le caractère optimalement fortement bien posé du problème aux limites (II.25). On donne ici les grandes lignes de la démonstration apparaissant dans [[Benoit, 2020a]-Paragraph 3.3]. Le schéma de démonstration est classique et suit celui du demi-espace. Il s'agit d'une démonstration en trois points

1. L'établissement d'une estimation d'énergie *a priori* pour le problème (II.25).
2. La construction d'une solution faible du problème (II.25).
3. L'établissement d'un lemme de type "*fort=faible*" qui assure que la solution faible précédente est forte. En particulier, elle vérifie alors l'estimation d'énergie. Elle est donc unique.

Nous ne nous étendrons pas sur le point 3. précédemment cité, car, dans la géométrie de la bande, ce dernier est obtenu *via* les mêmes techniques de régularisation de la solution que dans le cas du demi-espace. On verra toutefois dans le Paragraphe II.4.2, que ce point devient central dans le cas du quart d'espace.

De la même manière, le point 2. du schéma de preuve précédent est fait de manière assez classique par dualité [Lax and Phillips, 1960]. On introduit un problème dual pour (II.25) dont la solution joue le rôle d'une fonction test dans la formulation faible. Si l'on arrive alors à démontrer que ce problème dual (qui par construction, est de même nature que le problème primal (II.25)) vérifie l'estimation d'énergie *a*

---

<sup>19</sup>Dans [Benoit, 2020a], on démontre seulement que l'inversibilité des opérateurs précédents sur les sous-espaces idoines est nécessaire. Le caractère uniforme de l'inverse n'est pas montré comme étant nécessaire, même si l'on a de fortes raisons de penser qu'il s'avère l'être.

*priori*, alors on obtient un résultat d'unicité de la solution pour le problème dual. De ce résultat se déduit alors classiquement un résultat d'existence pour le problème primal.

On peut aussi noter que dans le cas particulier où l'opérateur  $\mathbf{T}_{0 \rightarrow 0}(\zeta)$  est une contraction, une solution faible de (II.25) peut alors être construite à la main comme une série de Neumann (voir [[Benoit, 2020a]-Proposition 3.27]).

Le point le plus délicat dans la démonstration précédente est, comme dans le cas du demi-espace, l'établissement de l'estimation d'énergie *a priori*. On ne décrira pas ici précisément la démonstration. Indiquons toutefois que cette dernière ressemble à celle du demi-espace.

En effet, elle fait intervenir un *symétriseur* qui permet de se ramener au cas de systèmes symétriques pour lesquels les termes de bord qui apparaissent dans les intégrations par parties sont signés. Dans un tel cas, de figure la démonstration de l'estimation d'énergie est alors triviale (voir, par exemple, [[Benzoni-Gavage, 2007]-Chapter 3]).

Le *symétriseur* que l'on est amené à construire dans le cas de la bande est très fortement inspiré de celui introduit par Osher dans [Osher, 1973] pour établir l'estimation d'énergie *a priori* dans le cas du quart d'espace. On regroupe les propriétés que l'on va imposer à cet objet dans la définition suivante (voir Définition 3.17 et 3.19 de [Benoit, 2020a])

**Définition II.3.2 (*Symétriseur dans la bande*)** *On dit que  $\mathcal{S} : \Xi \setminus \Xi_0 \times [0, 1] \rightarrow \mathbf{M}_{N \times N}(\mathbb{C})$  est un symétriseur pour le problème (II.25) si les conditions suivantes sont satisfaites :*

- i) *Pour tout  $\zeta \in \Xi \setminus \Xi_0$  et pour tout  $x_d \in [0, 1]$ , la matrice  $\Sigma(\zeta, x_d) := \mathcal{S}(\zeta, x_d)A_d$  est hermitienne.*
- ii) *Pour tout  $\zeta \in \Xi \setminus \Xi_0$ , l'opérateur  $\Sigma(\zeta, 0)$  (resp.  $\Sigma(\zeta, 1)$ ) est uniformément borné (par rapport à  $\zeta$ ) sur  $\mathbf{E}^s(\zeta)$  (resp.  $\mathbf{E}^u(\zeta)$ ).*
- iii) *Pour tout  $\zeta \in \Xi \setminus \Xi_0$ ,  $\mathcal{S}(\cdot, \zeta)$  est solution de l'équation différentielle*

$$\frac{1}{2} \frac{d}{dx_d} \mathcal{S}(\zeta, x_d) - \text{Re}(\mathcal{S}(\zeta, x_d)A_d \mathcal{A}(\zeta)) = 0, \text{ pour } x_d \in [0, 1].$$

- iv) *Pour tout  $v \in \mathbb{C}^N$  et pour tout  $\zeta \in \Xi \setminus \Xi_0$ , on a l'inégalité*

$$\langle v, \mathcal{S}(\zeta, 0)v \rangle \leq C_0 |B_0 v|^2 - \varepsilon_0 |v|^2,$$

*où les constantes  $C_0, \varepsilon_0 > 0$  sont indépendantes de  $\zeta$ .*

- v) *Pour tout  $w \in \mathbb{C}^N$  et pour tout  $\zeta \in \Xi \setminus \Xi_0$ , on a l'inégalité*

$$-\langle w, \mathcal{S}(\zeta, 1)w \rangle \leq C_1 |B_1 w|^2 - \varepsilon_1 |w|^2,$$

*où les constantes  $C_1, \varepsilon_1 > 0$  sont indépendantes de  $\zeta$ .*

On peut alors facilement montrer (voir [[Benoit, 2020a]-Proposition 3.8]) que si le problème (II.25) admet un *symétriseur*, au sens de la Définition II.3.2, alors il est optimalement fortement bien posé.

Expliquons les analogies et les différences entre le *symétriseur* de la Définition II.3.2 et le *symétriseur* de Kreiss [Kreiss, 1970].

Il est avant tout à noter que le terme de *symétriseur* est, ici, quelque peu abusif. En effet, il nous faut préciser que dans [Benoit, 2020a], on utilise l'hypothèse de symétrie des coefficients  $A_j$ . Ainsi, le *symétriseur* ne symétrise pas. La terminologie de "positivité" soufflée par Williams serait peut-être plus pertinente mais l'auteur a du mal à construire un néologisme lui-même basé sur un néologisme.

Une autre raison, qui nous a poussé à continuer d'utiliser le terme de *symétriseur* est que ces objets jouent un rôle central dans la théorie des problèmes (aux limites ou de Cauchy) hyperboliques. Comme on va le voir, l'objet de la Définition II.3.2 partage des propriétés communes avec le *symétriseur* de Kreiss, donc la terminologie n'est pas si abusive. On donnera plus de détails sur ce que serait un "vrai" *symétriseur* dans le Paragraphe II.3.4.

Les points *i)* et *iii)* de la Définition II.3.2 sont des analogues de la condition de symétrisation de l'équation à l'intérieur du demi-espace du *symétriseur* de Kreiss. Il est toutefois à noter que, là où, le *symétriseur* de Kreiss est défini de façon à satisfaire une inégalité matricielle, le *symétriseur* dans la bande requiert de résoudre une équation différentielle ordinaire. Ceci est lié au fait que le *symétriseur* doit pouvoir dépendre de la variable  $x_d$  puisque ce dernier doit pouvoir imposer deux conditions de positivité sur le bord, l'une en  $\{x_d = 0\}$  et l'autre en  $\{x_d = 1\}$ .

Ces conditions de positivité sont les deux conditions décrites aux points *iv)* et *v)* de la Définition II.3.2. Elles sont des généralisations directes de la condition de positivité sur le bord imposée dans le *symétriseur* de Kreiss<sup>20</sup>.

Enfin, la condition *ii)* de la Définition II.3.2 est une nouveauté comparée au *symétriseur* de Kreiss. Cette condition de borne uniforme est faite pour assurer que l'on puisse, à la fin des fins, revenir dans les variables de position et donc revenir du système résolvant (II.27) au système (II.25), *via* la formule de Plancherel.

La démonstration de l'existence d'un *symétriseur* est la partie technique de [Benoit, 2020a] et elle se base, *mutatis mutandis*, sur l'adaptation de la construction du *symétriseur* d'Osher pour le cas du quart d'espace [Osher, 1973]. Nous ne décrirons pas, ici, cette construction en détails. On renvoie à [[Benoit, 2020a]-Paragraphe 3.3.2] ou encore au Paragraphe III.4.2 où les points marquants de cette construction seront abordés dans un cadre discret.

Un point remarquable de cette construction est que, pour sa définition, le *symétriseur*  $\mathcal{S}$  utilise l'inverse de  $I - \overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta)$  donné par l'Hypothèse II.3.1. Le caractère uniforme, donné par l'Hypothèse II.3.1, permet alors de démontrer le point *ii)* de la Définition II.3.2.

Un autre point remarquable, absent de l'analyse d'Osher [Osher, 1973], est le fait que l'on n'utilise pour construire le *symétriseur* que l'inverse de  $I - \overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta)$  sur le sous-ensemble associé à  $\ker B_0$ . Dans le quart d'espace, cette propriété est demandée sur l'espace  $\mathbb{C}^N$  en entier. Ce raffinement a une conséquence importante concernant la construction de solution faible c'est-à-dire le point 2. de notre schéma de démonstration. En effet, on peut montrer que la condition d'inversibilité 2. de l'Hypothèse II.3.1 écrite pour le problème dual est équivalente<sup>21</sup> à l'hypothèse d'inversibilité 1. de l'Hypothèse II.3.1 pour le problème primal. Ce résultat n'étant vrai que si l'on considère les restrictions idoines pour les opérateurs  $I - \mathbf{T}_{0 \rightarrow 0}(\zeta)$  et  $I - \overline{\mathbf{T}}_{0 \rightarrow 0}(\zeta)$ .

Ainsi en supposant 2. de l'Hypothèse II.3.1, on a l'estimation d'énergie *a priori* et en supposant 1. de l'Hypothèse II.3.1, on a l'estimation d'énergie duale dont on déduit l'existence d'une solution faible. Comme précédemment indiqué un lemme "fort=fable" conclut la démonstration.

Le résultat que l'on obtient est alors le suivant

**Théorème II.3.1 ([Benoit, 2020a])** *On suppose que le problème (II.25) est hyperbolique à multiplicité constante, que les bords sont non caractéristiques et qu'ils vérifient les conditions de Kreiss-Lopatinskii uniforme idoines.*

<sup>20</sup>On exprime le plus souvent la condition de positivité sur le bord pour le *symétriseur* de Kreiss après une dernière réduction du problème à la donnée de bord triviale  $g = 0$ . De la même façon, les conditions de positivité *iv)* et *v)* ont une formulation réduite.

<sup>21</sup>Ceci tend à montrer que les deux conditions dans l'Hypothèse II.3.1 s'avèrent être équivalentes.

1. Si le problème est optimalement fortement bien posé au sens de la Définition II.3.1, alors l'Hypothèse II.3.1, sans le caractère uniforme des inverses, est vérifiée.
2. Réciproquement, si le problème (II.25) vérifie l'Hypothèse II.3.1, alors il est optimalement fortement bien posé au sens de la Définition II.3.1.

### II.3.2 Illustration de cette condition *via* le d'optique géométrique

Ce paragraphe décrit les résultats de l'article [Benoit, 2020c]. Le fait le plus marquant de cette contribution est, à notre avis, que comme la condition de Kreiss-Lopatinskii uniforme, la condition d'inversibilité décrite dans le paragraphe précédent, admet une version microlocalisée nécessaire et suffisante pour construire le développement d'optique géométrique de la solution.

Ainsi, une fois de plus, le développement BKW permet de rendre plus explicites et tangibles les conditions utilisées pour la démonstration du caractère fortement bien posé.

On considère le problème fortement oscillant associé à (II.25) où l'on s'intéresse particulièrement à un terme source de bord fortement oscillant  $g_0^\varepsilon = e^{\frac{i}{\varepsilon}\psi(t,x')}g$ , où l'amplitude  $g$  est régulière, nulle pour les temps négatifs.

Le système prend alors la forme

$$\begin{cases} L(\partial)u^\varepsilon = 0 & \text{dans } \Omega, \\ B_0u^\varepsilon|_{x_d=0} = g_0^\varepsilon & \text{sur } \partial\Omega_0, \\ B_1u^\varepsilon|_{x_d=1} = 0 & \text{sur } \partial\Omega_1, \\ u^\varepsilon|_{t \leq 0} = 0 & \text{sur } \Gamma, \end{cases} \quad (\text{II.33})$$

système pour lequel on cherche une solution approchée sous la forme d'une somme de paquets d'ondes fortement oscillants (voir le Paragraphe II.2.3). Supposons, pour simplifier et dans un premier temps, que la fréquence  $(\tau, \eta)$  associée à la phase de bord  $\psi := \tau t + \eta \cdot x'$ , ne soit pas de *glancing*, i.e.  $(i\tau, \eta) \notin \mathcal{G}$ .

Ainsi l'équation eikonale  $\det \mathcal{L}(d\varphi_k) = 0$  admet seulement deux types de phases  $\varphi_k := \psi + \xi_k x_d$  :

- Les phases elliptiques pour lesquelles on a  $\text{Re } \xi_k \neq 0$ . Suivant l'analyse de [Lescarret, 2007] ces phases induisent des couches limites évanescentes (resp. explosives) qui sont exponentiellement décroissantes (resp. croissantes) selon le signe de cette partie réelle.
- Les phases hyperboliques pour lesquelles on a  $\xi_k \in i\mathbb{R}$ . D'après le lemme de Lax [Lax, 1957], ces phases induisent des phénomènes de transport. Ces derniers peuvent aller de :
  - la gauche vers la droite et la phase est alors qualifiée d'entrante<sup>22</sup>.
  - la droite vers la gauche et la phase est alors qualifiée de sortante<sup>23</sup>.

Pour simplifier encore l'exposition supposons, ceci est suffisant pour comprendre les phénomènes en présence, que l'équation eikonale n'admet qu'une phase entrante,  $\varphi_+$ , et une phase sortante,  $\varphi_-$ .

Afin de décrire l'ansatz de (II.33), commençons par la simple (mais cruciale) remarque que notre système étant hyperbolique, il satisfait le principe de propagation à vitesse finie de la solution. Par conséquent, le terme source non trivial localisé sur le bord  $\partial\Omega_0$  ne peut pas influencer directement la solution loin de cette face. Ainsi, il est naturel de considérer que, du moins pendant un temps court, le

<sup>22</sup>On fixe, ici, arbitrairement  $\partial\Omega_0$  comme frontière de référence.

<sup>23</sup>Le cas d'un transport purement tangentiel aux faces du bord est exclu car on a supposé que l'on se plaçait autour d'une fréquence non *glancing*



problème (II.33) ne voit pas la condition de bord sur  $\partial\Omega_1$ . Donc on commence par prendre pour ansatz celui pour le problème dans le demi-espace à savoir :

$$u^\varepsilon(t, x) \sim \sum_{n \geq 0} \varepsilon^n e^{\frac{i}{\varepsilon} \varphi_+} u_{n,+}(t, x) + \sum_{n \geq 0} \varepsilon^n e^{\frac{i}{\varepsilon} \varphi_-} u_{n,-}(t, x) + \sum_{n \geq 0} \varepsilon^n e^{\frac{i}{\varepsilon} \psi} U_{n,ev} \left( \frac{x_d}{\varepsilon} \right), \quad (\text{II.34})$$

développement dans lequel les  $u_{n,\pm}$  sont liées aux deux phases hyperboliques et les  $U_{n,ev}$  regroupent les contributions des éventuelles phases évanescents. Puisque l'on cherche une solution  $L^2(\Omega)$ , on a exclu les phases explosives.

Injecter cet ansatz dans l'équation d'évolution de (II.33) nous donne, de façon classique, essentiellement grâce au lemme de Lax [Lax, 1957], que l'amplitude principale  $u_{0,+}$  (resp.  $u_{0,-}$ ) vérifie une équation de transport entrante (resp. sortante), à savoir pour laquelle la vitesse de groupe vérifie  $\mathbf{v}_{\pm,d} > \pm 0$ . L'amplitude principale liée à la partie évanescence vérifie, quant à elle, une équation différentielle linéaire, ordinaire dans la variable rapide  $X_d := \frac{x_d}{\varepsilon}$ . Ainsi, sa solution aura un comportement exponentiellement décroissant par rapport à  $x_d$ .

Remarquons au passage qu'en tant que solution d'une équation de transport sortante totalement homogène (car on a  $g_1 \equiv 0$ ), l'amplitude  $u_{0,-}$  pourrait être (initialement) négligée dans l'ansatz (II.34). Toutefois, cela n'est pas vrai pour tout temps. En effet, l'information transportée associée à  $u_{0,+}$  va après un temps (explicite) de transport, frapper la face  $\partial\Omega_1$  et y "allumer" un terme source. On peut alors réitérer l'argument de propagation à vitesse finie pour se ramener à étudier le problème dans le demi-espace  $\{x_d < 1\}$ . Pour un tel problème, la phase  $\varphi_-$  est maintenant entrante, elle a donc maintenant toute sa place dans l'ansatz. On ajoute aussi à ce dernier une composante pour prendre en compte les modes explosifs initialement écartés. Ces derniers étant maintenant évanescents pour la face  $\partial\Omega_1$ . L'ansatz final s'écrit donc

$$u^\varepsilon \sim \sum_{n \geq 0} \varepsilon^n e^{\frac{i}{\varepsilon} \varphi_+} u_{n,+} + \sum_{n \geq 0} \varepsilon^n e^{\frac{i}{\varepsilon} \varphi_-} u_{n,-} + \sum_{n \geq 0} \varepsilon^n e^{\frac{i}{\varepsilon} \psi} U_{n,ev} \left( \frac{x_d}{\varepsilon} \right) + \sum_{n \geq 0} \varepsilon^n e^{\frac{i}{\varepsilon} \psi} U_{n,ex} \left( \frac{x_d - 1}{\varepsilon} \right). \quad (\text{II.35})$$

On peut décrire toutes les équations qui gouvernent chacune des amplitudes de (II.35) à l'intérieur du domaine. Pour cela, on reprend les calculs exposés au Paragraphe II.2.3 et on utilise, par exemple, l'analyse de Lescarret [Lescarret, 2007] pour traiter les phases elliptiques. Ces dernières étant de moindre importance on se permet de ne pas décrire précisément la détermination de leurs équations internes.

Pour conclure, il suffit de considérer les conditions de bord. La condition de bord en  $\{x_d = 0\}$  donne au premier ordre

$$B_0 [u_{0,+} + u_{0,-}]_{|x_d=0} + B_0 U_{0,ev}|_{x_d=0} + \underbrace{B_0 U_{0,ex} \left( -\frac{1}{\varepsilon} \right)}_{=0} = g_0,$$

car en utilisant la décroissance rapide de  $U_{0,ex}$ , ce terme contribue en  $O(\varepsilon^\infty)$  et peut donc être négligé dans une condition d'ordre  $O(1)$ . On procède de même sur la face  $\partial\Omega_1$ . Enfin, on utilise la condition de Kreiss-Lopatinskii uniforme sur chaque face pour obtenir le système suivant pour les amplitudes hyperboliques

$$\begin{cases} \partial_t u_{0,\pm} + \mathbf{v}_\pm \cdot \nabla_x u_{0,\pm} = 0 & \text{dans } \Omega, \\ u_{0,+}|_{x_d=0} = g_0 - \phi_0 B_0 u_{0,-}|_{x_d=0} & \text{sur } \partial\Omega_0, \\ u_{0,-}|_{x_d=1} = -e^{\frac{i}{\varepsilon} \tilde{\psi}} \phi_1 B_1 u_{0,+}|_{x_d=1} & \text{sur } \partial\Omega_1, \\ u_{0,\pm}|_{t \leq 0} = 0 & \text{sur } \Gamma, \end{cases} \quad (\text{II.36})$$

où<sup>24</sup> le couplage a lieu à travers les conditions de bord.

<sup>24</sup>Le terme  $e^{\frac{i}{\varepsilon} \tilde{\psi}}$  est un facteur de phase qui est un reliquat de l'évaluation des phases  $\varphi_\pm$  en  $\{x_d = 1\}$ . Son expression exacte n'a pas d'influence dans ce qui suit.

On peut alors facilement résoudre ce système par un argument d'analyse/synthèse. On suppose connue la trace (inconnue),  $u_{0,-|x_d=0}$ . La condition de bord sur  $\partial\Omega_0$  donne  $u_{0,+|x_d=0}$ . On résout l'équation de transport associée et on calcule, en particulier,  $u_{0,+|x_d=1}$ . Ceci donne  $u_{0,-|x_d=1}$ . Une seconde intégration le long des caractéristiques détermine alors  $u_{0,-}$ . Prenant alors la trace de la solution sur  $\partial\Omega_0$ , on obtient alors que l'inconnue  $u_{0,-|x_d=0}$  vérifie une équation de la forme

$$(I - \mathbf{T})u_{0,-|x_d=0} = F(g), \quad (\text{II.37})$$

où  $F(g)$  dépend explicitement de l'amplitude de la source  $g_0^\varepsilon$  et où l'opérateur  $\mathbf{T}$  s'écrit sous la forme

$$(\mathbf{T}u)(t, x') := e^{\frac{i}{\varepsilon}\tilde{\psi}} Ru(t - \alpha, x - \beta').$$

Les paramètres  $\alpha \in \mathbb{R}_+$  et  $\beta' \in \mathbb{R}^{d-1}$  étant explicites en termes des deux vitesses de groupe (voir [[Benoit, 2020c]-equation (42)]) et la matrice  $R$  dépend explicitement des conditions de bord.

Le paramètre  $\alpha$  correspond au temps nécessaire pour régénérer la phase auto-interagissante par rebonds successifs. La matrice  $R$  est, quant à elle, reliée au "coefficient" d'amplification lors des deux rebonds nécessaires pour régénérer la phase.

Ainsi, pour initialiser la résolution de la cascade d'équation du développement d'optique géométrique, on a besoin de résoudre l'équation (II.37). En d'autres termes, d'inverser (dans un espace fonctionnel adapté) l'opérateur  $I - \mathbf{T}$ . On retrouve alors une version microlocalisée de l'Hypothèse II.3.1 utilisée dans la caractérisation des problèmes optimalement fortement bien posés. Notons que cette hypothèse d'inversion peut toujours être satisfaite si l'on veut construire le développement BKW seulement sur un temps fini  $t \leq T < \infty$ . Dans un tel cadre, les termes dans le développement en série de Neumann de  $(I - \mathbf{T})^{-1}u$  s'écrivent alors sous la forme  $C_n u(t - n\alpha)$  et viennent s'annuler pour tout  $n$  assez grand car  $\alpha \in \mathbb{R}_+$ . Un autre cadre favorable pour faire converger la série est celui où l'opérateur  $\mathbf{T}$  est une contraction ce qui se traduit par un petit coefficient (global) de réflexion ( $\|R\| < 1$ ).

Dans le cas général où l'on permet à l'équation eikonale d'avoir de multiples racines, la méthode décrite précédemment s'applique et l'on obtient une condition d'inversibilité analogue à (II.37), bien que plus technique. On renvoie le lecteur intéressé à [[Benoit, 2020c]-Paragraph 5.1] pour plus de détails. De la même manière dans [Benoit, 2020c], on considère aussi le cas où la fréquence au bord est de *glancing* en utilisant les résultats de [Williams, 1996]. Ces résultats permettent, en construisant les amplitudes de *glancing* comme des transports tangentiels, de construire à moindre coût une solution approchée (avec faible précision)<sup>25</sup> de la solution de (II.33). Le point intéressant dans cette analyse est que les modes de *glancing* sont liés à des phénomènes de transport tangentiel ou au pire à des couches limites. Ils n'interviennent pas simultanément dans les deux conditions de bord. Ils ne sont donc pas réfléchis et par conséquent, comme les modes elliptiques, ils n'entrent pas en compte dans la condition d'inversibilité de  $I - \mathbf{T}$ . On reviendra sur cette remarque dans le Paragraphe d'ouverture II.3.4.

Le théorème de [Benoit, 2020c] que nous avons discuté dans ce paragraphe s'écrit donc :

**Théorème II.3.2 ([Benoit, 2020c]-Theorems 7.2 et 7.3)** *On suppose que le problème (II.25) est hyperbolique à multiplicité constante, qu'il admet des bords non caractéristiques et que chaque face de la bande vérifie la condition de Kreiss-Lopatinskiï uniforme. On suppose de surcroît que l'équation de compatibilité*

<sup>25</sup>Le travail de [Williams, 1996] ne fournit que la construction du premier correcteur. Il a ensuite été amélioré à tous ordres dans [Williams, 2000]. En particulier, les résultats de [Benoit, 2020c], lorsque l'on a une fréquence de *glancing*, ne sont pas optimaux.

(voir (II.37)) admet une unique solution dans un espace fonctionnel idoine<sup>26</sup> Dans ce qui suit, on note  $\gamma_0 \geq 0$  la croissance exponentielle minimale en temps de la solution de (II.25).

- Si la fréquence au bord  $(i_{\mathcal{T}}, \eta) \notin \mathcal{G}$ , on considère l'ansatz (II.35). On peut construire des amplitudes  $(u_{n,k})_{n \geq 0}, (U_{ev,n})_{n \geq 0}$  et  $(U_{ex,n})_{n \geq 0}$  telles que l'ansatz (II.35) vérifie la cascade d'équation d'optique géométrique à tous ordres. Pour  $N_0 \in \mathbb{N}$ , l'ansatz tronqué à l'ordre  $N_0$ , nommé  $u_{N_0,app}^\varepsilon$ , vérifie, qu'il existe  $C > 0$ , telle que pour tout  $\gamma \geq \gamma_0$ , on ait

$$\|u^\varepsilon - u_{N_0,app}^\varepsilon\|_{L^2_\gamma(\Omega)} \leq C\varepsilon^{N_0+1}.$$

- Si la fréquence au bord  $(i_{\mathcal{T}}, \eta) \in \mathcal{G}$ , on considère l'ansatz (II.35) auquel on adjoint un terme de transport (cru et imprécis)  $u_{n,g,k}^\varepsilon(t, x)$  pour gérer les modes de glancing. On peut alors construire un ansatz au premier ordre,  $u_{app}^\varepsilon$ , tel qu'il existe  $C > 0$ , telle que pour tout  $\gamma \geq \gamma_0$ , on ait

$$\|u^\varepsilon - u_{app}^\varepsilon\|_{L^2_\gamma(\Omega)} \leq C\varepsilon^{1/4}.$$

### II.3.3 Vers une nouvelle zoologie de problèmes faiblement bien posés

Ce paragraphe synthétise les résultats des articles [Benoit, 2020b] et [Benoit, 2021]. Ils montrent que l'ajout d'une condition de bord dans un problème aux limites, peut venir totalement chambouler, ou au contraire, ne pas modifier, le comportement que l'on est en droit d'attendre pour un problème faiblement bien posé.

On explique brièvement comment le procédé de génération d'instabilités faibles (*i.e.* la perte d'un nombre fini de dérivées) est influencé par la présence d'un second bord. On reprend donc le problème fortement oscillant (II.33), mais dans lequel on suppose maintenant que la condition de Kreiss-Lopatinskii uniforme n'est pas satisfaite pour la face  $\partial\Omega_0$ . Ainsi, comme expliqué dans le Paragraphe II.2.3, le terme source sur le bord  $g_0^\varepsilon$  s'écrit maintenant  $\varepsilon^\beta e^{\frac{i}{\varepsilon}\psi} g_0$ , où  $\beta > 0$ , est témoin de la perte de régularité de la solution. La face  $\partial\Omega_1$  vérifie (dans un souci de simplification) la condition de Kreiss-Lopatinskii uniforme. Le procédé de génération des instabilités est alors fortement dépendant de la nature de la phase de bord  $\psi$ .

Dans [Benoit, 2020b], nous considérons le cas (plus favorable) d'une phase au bord elliptique. Alors, reprenant notre argument de propagation à vitesse finie de l'information on peut initialement se rapporter à l'analyse dans le demi-espace. On s'attend à ce qu'un terme source au bord en  $\varepsilon$  génère un terme principal dans le développement d'optique géométrique d'ordre  $\varepsilon^0$ . La question naturelle est de se demander si cette instabilité peut venir être amplifiée lors de rebonds successifs contre les faces du domaine de résolution ?

On rappelle que les amplitudes associées aux phases elliptiques ont un comportement exponentiellement décroissant dans la variable normale  $x_d$ . Par conséquent, l'amplitude elliptique principale (d'ordre zéro) est générée proche de la face  $\partial\Omega_0$  et quand elle frappe la face  $\partial\Omega_1$ , elle le fait en  $O(\varepsilon^\infty)$ . Elle n'est pour ainsi dire pas réfléchiée et par conséquent, ne peut pas être renvoyée vers la face  $\partial\Omega_0$  pour y subir une nouvelle perte de dérivée.

Par conséquent, pour une dégénérescence de la condition de Kreiss-Lopatinskii uniforme dans la zone elliptique, le comportement de la solution du problème dans la bande est exactement le même que pour le problème dans le demi-espace.

<sup>26</sup>Un bon cadre fonctionnel est ici donné par l'espace  $H_{\natural,\gamma}^\infty(\partial\Omega_0) := \cap_{s \in \mathbb{N}} H_{\natural,\gamma}^s(\partial\Omega_0)$  pour  $\gamma > \bar{\gamma}_0$ , avec  $\bar{\gamma}_0 \geq 0$  donné et vérifiant  $\bar{\gamma}_0 \leq \gamma_0$  où  $\gamma_0$  désigne l'éventuel facteur de croissance exponentielle (en temps) du système (II.25). Pour  $s \in \mathbb{N}$ , on définit

$$H_{\natural,\gamma}^s(X) := \{u \in \mathcal{D}'(X) \setminus ue^{-\gamma t} \in H^s(X) \text{ et } u|_{t \leq 0} = 0\},$$

pour  $X \subset \Omega$ .

Ce phénomène devient toutefois totalement faux lorsque la dégénérescence de la condition de Kreiss-Lopatinskii uniforme se fait dans la zone hyperbolique. Ceci fait l’objet de l’article [Benoit, 2021]. La raison en est simple, ces amplitudes ne sont plus gouvernées par des équations à décroissance rapide mais par des équations de transport, donc les instabilités se propagent. L’étude de [Coulombel and Guès, 2010] dans le demi-espace, nous invite à choisir (initialement)  $\beta = 1$ . On génère donc proche de  $\partial\Omega_0$  une amplitude d’ordre  $\varepsilon^0$ . Cette dernière étant hyperbolique, elle est transportée vers  $\partial\Omega_1$ , y est réfléchié toujours en  $\varepsilon^0$ , car on est en présence de la condition de Kreiss-Lopatinskii uniforme contre la face  $\partial\Omega_1$ .

Ensuite, cette amplitude reprend sa route vers la face  $\partial\Omega_0$ . Lors de sa réflexion, elle subit une nouvelle amplification en raison de l’absence de la condition de Kreiss-Lopatinskii uniforme. Changeant l’échelle du terme de bord en  $\varepsilon^2$ , on est donc amené à considérer une amplitude principale se comportant comme

$$u_0 = \begin{cases} O(\varepsilon) & \text{si } t < \alpha, \\ O(1) & \text{si } t \geq \alpha, \end{cases} \quad \text{où } \alpha \text{ est le temps nécessaire pour faire un aller-retour vers la face } \partial\Omega_0.$$

Ainsi, en réitérant cet argument, si l’on veut construire le développement d’optique géométrique jusqu’à un temps  $T$  donné, on va prendre un terme de bord d’ordre  $\varepsilon^{1+[T/\alpha]}$ , où  $[\cdot]$  désigne la partie entière. La détermination des différentes amplitudes dans le développement d’optique géométrique est ici non triviale, technique et calculatoire. On renvoie à l’article [Benoit, 2021] pour les détails de cette construction, ou bien à la discussion qui suit pour ses grandes lignes.

On montre donc, en utilisant le développement d’optique géométrique, que si le problème dans la bande admet une estimation d’énergie, alors cette dernière doit comprendre un nombre de pertes (sur le bord) qui est une fonction croissante du temps final de résolution  $T$ . Ce comportement diffère alors beaucoup de celui dans le demi-espace où une unique perte est observée. Un tel comportement de la solution est, à notre avis, un exemple de la richesse et de la beauté des systèmes hyperboliques (faiblement bien posés). Notons que ce genre de perte de régularité croissante en fonction du temps, que certains pourraient qualifier de pathologique, n’est pas nouveau. En effet, il apparaît déjà pour le problème de Cauchy mais pour des problèmes à coefficients très peu réguliers (à savoir dans la classe log-Zygmund). On citera ici la série de travaux de Colombini, Del Santo, Franeli et Métivier : [Colombini et al., 2013a]-[Colombini et al., 2015] ou [Colombini et al., 2013b]. Cette perte peut aussi se dérouler pour le problème de Cauchy lorsque l’on se place sous une hypothèse d’hyperbolicité faible (on pourra consulter [Morisse, 2021]).

Nous pensons que l’exemple donné dans [Benoit, 2021] est assez notable, car d’une part, il s’agit d’un problème fortement hyperbolique et à coefficients constants et d’autre part, car en raison de la simplicité du phénomène sous-jacent il est plus facile à appréhender que les exemples de la littérature.

Il est ici assez juste de nuancer le terme de ”simplicité”. En effet, une étude rigoureuse de ce phénomène de propagation des singularités dans la bande vient nécessairement avec un minimum de technique. Minimum qui a peut-être été quelque peu augmenté dans [Benoit, 2021] où nous avons cherché à établir le résultat le plus général possible (hyperbolicité à multiplicité constante, nombre arbitraire de phases hyperboliques...). Ainsi, puisque les temps nécessaires pour faire un aller-retour ne sont pas les mêmes et que la perte de régularité ne concerne qu’une composante de la partie hyperbolique, nous avons eu à gérer beaucoup de phases qui coexistent à des échelles diverses.

Le point principal de la méthode de construction est toutefois le suivant. On commence la résolution de la cascade d’équations du développement d’optique géométrique par la détermination des amplitudes principales, nommément les  $u_{0,k}$ , en d’autres termes les amplitudes qui portent le plus de singularités. Ce choix peut sembler de prime abord un peu contre-intuitif, toutefois il ne l’est pas. En effet, il n’y a pour ainsi dire qu’un unique moyen de venir allumer un terme portant  $1 + [T/\alpha]$  pertes de dérivées. Il ne peut provenir que de l’information initiale qui a eu le temps de faire suffisamment de tours de boucle. Un point délicat de l’analyse étant ici que de prime abord dans le terme  $u_{0,k}$  il pourrait y avoir une contribution du terme  $u_{1,k}$  qui proviendrait du premier correcteur qui a fait un tour de boucle. Une étude précise des temps de parcours et des différents temps d’apparition des amplitudes montre toutefois qu’une telle

situation est impossible.

Ce choix peut sembler de prime abord un peu contre-intuitif, toutefois il ne l'est pas. En effet, il n'y a pour ainsi dire qu'un unique moyen de venir allumer un terme portant  $1 + [T/\alpha]$  pertes de dérivées. Il ne peut provenir que de l'information initiale qui a eu le temps de faire suffisamment de tours de boucle. Un point délicat de l'analyse étant ici que de prime abord dans le terme  $u_{0,k}$  il pourrait y avoir une contribution du terme  $u_{1,k}$  qui proviendrait du premier correcteur qui a fait un tour de boucle. Une étude précise des temps de parcours et des différents temps d'apparition des amplitudes montre toutefois qu'une telle situation est impossible.

Puis, on détermine classiquement les amplitudes  $u_{1,k}$  qui portent  $[T/\alpha]$  pertes de dérivées. Ces dernières dépendent des  $u_{0,k}$ , *via* leur partie non polarisée et leur partie polarisée dépend du terme source qui a effectué  $[T/t_*] - 1$  tours de boucle. On conclut alors la construction par itérations.

En somme, les résultats de [Benoit, 2020b] et [Benoit, 2021] montrent que l'ajout d'une nouvelle condition de bord dans un problème aux limites, vérifiant pourtant la condition de Kreiss-Lopatinskii uniforme, peut avoir un sérieux impact sur les phénomènes d'instabilité en présence. On le résume dans le théorème suivant :

**Théorème II.3.3** ([Benoit, 2020b], [Benoit, 2021]) *For  $T > 0$ , on suppose que le problème dans la bande*

$$\begin{cases} L(\partial)u = f & \text{dans } ]-\infty, T] \times \Gamma := \Omega_T, \\ B_0 u|_{x_d=0} = g_0 & \text{sur } ]-\infty, T] \times \mathbb{R}^{d-1} := \partial\Omega_{0,T}, \\ B_1 u|_{x_d=1} = g_1 & \text{sur } ]-\infty, T] \times \mathbb{R}^{d-1} := \partial\Omega_{1,T}, \\ u|_{t \leq 0} = 0 & \text{sur } \Gamma, \end{cases} \quad (\text{II.38})$$

*est hyperbolique, à bords non caractéristiques. On suppose que la condition de bord sur la face  $\partial\Omega_1$  vérifie la condition de Kreiss-Lopatinskii uniforme, tandis que celle sur la face  $\partial\Omega_0$  viole la condition de Kreiss-Lopatinskii uniforme. On a alors les alternatives suivantes :*

- *Si la condition de Kreiss-Lopatinskii uniforme sur  $\partial\Omega_0$  dégénère dans la zone elliptique, alors le problème (II.25) vérifie l'analogie de l'estimation d'énergie dans le demi-espace (voir (II.10)). Plus précisément, il existe  $C_T > 0$ , telle que la solution  $u$  de (II.38) vérifie :*

$$\begin{aligned} \|u\|_{L^2(\Omega_T)}^2 + \|u|_{x_d=0}\|_{H^{-1/2}(\partial\Omega_{0,T})}^2 + \|u|_{x_d=1}\|_{L^2(\partial\Omega_{1,T})}^2 \\ \leq C_T \left( \|f\|_{L^2(\Omega_T)}^2 + \|g_0\|_{H^{1/2}(\partial\Omega_{0,T})}^2 + \|g_1\|_{L^2(\partial\Omega_{1,T})}^2 \right) \end{aligned}$$

*De plus, le 1/2 dans le membre de droite de l'inégalité précédente est optimal.*

- *Soit  $K > 0$  et deux termes sources  $f \in L^2(\Omega_T)$ ,  $g_0 \in H^K(\partial\Omega_{0,T})$  (on pose  $g_1 \equiv 0$ ). On suppose que le problème (II.38) est faiblement bien posé au sens suivant, il admet une unique solution  $u \in L^2(\Omega_T)$  qui vérifie l'estimation d'énergie : il existe  $C_T > 0$  telle que*

$$\|u\|_{L^2(\Omega)}^2 \leq C_T \left[ \|f\|_{L^2(\Omega_T)}^2 + \|g_0\|_{H^K(\partial\Omega_{0,T})}^2 \right],$$

*alors on a nécessairement  $K \geq 1 + [\frac{T}{\alpha}]$ , où  $\alpha > 0$  désigne le temps minimal de régénération. Il dépend de façon explicite des vitesses de groupes.*

On conclut ce paragraphe en donnant deux possibles directions de recherche :

- Dans un souci de complétude de la théorie, on pourrait avoir envie d'étudier dans la bande une dégénérescence de la condition de Kreiss-Lopatinskii uniforme en zone de *glancing* en se basant sur l'article [Benoit and Loyer, 2023]. Toutefois, étant donné que les modes de *glancing* sont associés à des couches limites on peut aisément conjecturer que l'instabilité ne pourra pas être amplifiée. Par conséquent, le comportement de la solution doit être le même que dans le cas du demi-espace. Ainsi il n'y a pas de nouveau phénomène qualitatif à espérer.
- Le développement d'optique géométrique pour un problème dont l'une des faces est dans la classe  $WR$  montre que l'estimation d'énergie, si elle existe, exhibe forcément un nombre de pertes de régularité croissante en temps. Ceci est un résultat assez négatif dans la mesure où il exclut (sauf à se placer dans le cadre régulier) une existence de la solution en temps infini. On aimerait toutefois, en se plaçant sur un temps fixe  $T$ , trouver un moyen d'établir une estimation d'énergie avec une perte de dérivées dépendant de  $T$ , et plus généralement établir la théorie faiblement bien posée associée. Une telle étude n'est pas vraiment motivée par des considérations d'intérêt physique mais plutôt par son côté mathématique attrayant.

### II.3.4 De possibles extensions de la théorie fortement bien posée optimale dans la bande ?

Les extensions possibles de la caractérisation des problèmes optimalement bien posés dans la bande sont nombreuses mais ces dernières doivent être quelque peu nuancées.

Ceci a déjà été mentionné et peut être considéré comme un peu anecdotique, mais dans un souci de "beauté générale" de la théorie nous aimerions disposer d'un *symétriseur* "qui symétrise" réellement l'équation. Nous pourrions alors nous passer de l'hypothèse de symétrie des coefficients  $A_j$ . Dans [Benoit, 2020a], l'hypothèse de symétrie des coefficients est utilisée pour pouvoir se ramener à un problème homogène à l'intérieur en utilisant le caractère fortement bien posé d'un problème auxiliaire (qui est symétrique et admet des conditions de bord strictement dissipatives).

La principale difficulté est donc que sans cette hypothèse, on doit imposer que l'équation différentielle dans le point *iii*) définissant le *symétriseur* (voir Définition II.3.2) soit vérifiée avec un terme source quelconque. Ceci n'est en soit pas un problème au regard de la résolution même de l'équation. Toutefois, cela rajoute des termes qui se répercutent à chaque stade de la construction du *symétriseur*. Il faudrait alors, reprendre posément la construction (voir Paragraphe III.4.2) et étudier l'influence de ces nouveaux termes. On pense, en particulier, à la propriété de borne *ii*) de la Définition II.3.2.

Ensuite, réussir à construire un *symétriseur* dans le cas où les coefficients  $A_j$  dépendent des variables d'états  $(t, x)$  est aussi une direction à étudier. En étant conscient que l'intérêt de l'étude n'est pas ici dans le résultat en tant que tel, mais dans sa possible transposabilité aux problèmes dans le quart d'espace. En effet, lorsque l'on considère un opérateur à coefficients variables défini par  $L(\partial)u = \partial_t u + \sum_{j=1}^d \partial_j(A_j(t, x)u)$  alors une part importante de l'analyse se fait sur un opérateur de la forme  $\partial_t u + \sum_{j=1}^d A_j(t, x)\partial_j u + Du$  où  $D$  est un opérateur d'ordre zéro prenant en compte les variations des  $A_j$ .

Or, pour de tels opérateurs la présence du terme d'ordre zéro  $D$  (qui n'a aucune raison d'être signé) induit naturellement la présence d'un seuil minimal de validité de l'estimation d'énergie (II.26) et donc une croissance exponentielle par rapport au temps de la solution.

Ainsi, le but d'une telle théorie bien posée serait d'identifier les problèmes pour lesquels la croissance exponentielle de la solution est liée non pas à des variations internes des coefficients mais à des répétitions successives sur le bord.

Cette question est intéressante, mais elle peut sembler demander beaucoup d'efforts dans un cadre où, dans tous les cas, le phénomène qualitatif de croissance exponentielle ne peut pas être exclu. Cependant, si l'on prend en compte la grande ressemblance entre le problème dans la bande avec celui dans le quart

d'espace, il est à parier que la construction d'un *symétriseur* à coefficients variables dans la bande serait un excellent travail préparatoire à l'établissement d'un tel *symétriseur* dans le quart d'espace qui, lui, constitue un résultat très désirable.

Le cas caractéristique, que cela soit pour la théorie bien posée du problème (II.25) ou bien pour son développement d'optique géométrique, serait aussi une piste à considérer si l'on veut par exemple pouvoir appliquer nos résultats à l'équation de Maxwell. Une fois de plus, ceci serait un travail préparatoire en vue de l'ambitieux projet de traiter des systèmes caractéristiques dans le quart d'espace.

Enfin une dernière piste d'étude que l'on aimerait mentionner concerne l'étude de problèmes non linéaires. Puisque pour de tels problèmes, l'existence de la solution n'est *a priori* garantie que sur un intervalle de temps fini  $[0, T_*]$ , où  $T_*$  est typiquement le temps d'apparition de la singularité, ou encore d'explosion de la norme  $H^s$ , on peut se demander en quoi s'efforcer d'obtenir des estimations d'énergies optimales (qui sont vraiment intéressantes en temps infini) est pertinent. La méthode canonique pour montrer l'existence de la solution de tels problèmes est de procéder par point fixe sur le problème linéarisé, d'utiliser les estimations d'énergie et de forcer le point fixe à être contractant en choisissant un temps final assez petit.

Remarquons qu'en temps fini, une estimation non optimale se traduit au pire par l'apparition d'un terme  $e^{\gamma T}$  dans la constante du membre de droite de l'estimation d'énergie. Ainsi, en utilisant une estimation non optimale issue de la méthode simple de localisation, on pourra typiquement rendre la méthode contractante en imposant  $CT_*e^{\gamma T_*} < 1$ , au lieu de  $CT_* < 1$ . Par conséquent, avec une telle approche on obtiendrait un temps maximal admissible, mais qui est, *a priori*, inférieur à celui dans un cadre où le phénomène de croissance exponentielle est exclu. Récupérer l'estimation d'énergie optimale (si elle a lieu) devrait donc permettre de récupérer un temps d'existence optimal.

Plus généralement, nous pensons qu'il serait intéressant de mener l'étude lorsque le système linéarisé admet une estimation d'énergie bien posée optimale, afin de déterminer rigoureusement si, dans un tel cadre, le temps maximal d'existence ne pourrait pas égalier voir dépasser celui d'apparition de la singularité pour un problème sans bord ou dans un demi-espace.

Nous pensons qu'il devrait être possible de caractériser des conditions de bord qui jouent le rôle d'un amortissement et qui permettraient ainsi d'augmenter la durée de vie du problème dans la bande comparé, au problème dans le demi-espace ou le problème sans bord.

En effet, si l'on considère l'équation de Burgers

$$\begin{cases} \partial_t u + u \partial_x u = 0 & \text{pour } (t, x) \in ]0, T[ \times \mathbb{R}, \\ u|_{t=0} = u_0 & \text{pour } x \in \mathbb{R}, \end{cases}$$

où, par exemple  $u_0 \in \mathcal{S}(\mathbb{R})$ . Sa solution peut s'écrire  $u(t, x) = u_0(\mathbf{x}(t))$ , où la caractéristique  $\mathbf{x}$  est définie comme solution de

$$\begin{cases} \frac{d}{dt} \mathbf{x} = u(t, \mathbf{x}(t)) & \text{pour } t > 0, \\ \mathbf{x}(0) = x. \end{cases}$$

Ainsi

$$\mathbf{x}(t) = x + tu_0(x).$$

On retrouve donc le temps de résolution maximal, correspondant à l'intersection des caractéristiques  $T_* := -\inf_{x \in \mathbb{R}} \frac{1}{u'_0(x)}$ .

Mais si maintenant on considère l'équation de Burgers amortie

$$\begin{cases} \partial_t u^a + u^a \partial_x u^a + a u^a = 0 & \text{pour } (t, x) \in ]0, T[ \times \mathbb{R}, \\ u^a|_{t=0} = u_0, \end{cases}$$

avec  $a > 0$ , on peut reprendre la même analyse et montrer que sa solution s'écrit sous la forme  $u(t, x) = e^{-at} u_0(\mathbf{x}^a(t))$  où la nouvelle caractéristique  $\mathbf{x}^a$  est maintenant donnée par

$$\mathbf{x}^a(t) = x + \frac{1}{a} (1 - e^{-at}) u_0(x).$$

Le temps d'intersection des caractéristiques est maintenant déterminé par  $T_*^a = -\inf_{x \in \mathbb{R}} \frac{1}{a} \ln(1 + \frac{a}{u_0'(x)})$ .

Par conséquent, on obtient que pour tout  $a > 0$ , on a  $T_* < T_*^a$ . La solution de l'équation de Burgers amortie vit donc plus longtemps que celle de l'équation de Burgers classique. Comme pour le problème dans la bande, le caractère fortement bien posé optimal ressemble assez (d'un point de vu qualitatif) à l'ajout d'un amortissement interne, cela semble indiquer que, parfois, un problème dans la bande non linéaire peut vivre plus longtemps que son analogue sans bande. On indiquera ici aussi le travail de [Lebeau, 1996] sur l'influence d'un terme d'amortissement dans l'équation des ondes.

Dans un souci d'exhaustivité, concluons en mentionnant une perspective qui, toujours pour des problèmes non linéaires dans la bande, constitue un pont entre le problème dans la bande et son analogue dans le quart d'espace dont on parlera dans la suite de cette partie : nous aimerions parvenir à construire rigoureusement des développements d'optique géométrique pour des problèmes non linéaires posés dans la bande. Cette perspective étant un pont entre le monde de la bande et celui du quart d'espace, elle sera décrite avec plus de détails dans le Paragraphe II.4.5.

## II.4 Le problème dans le quart d'espace

Dans la dernière section de ce chapitre, on étudiera l'analogue du problème (II.25), mais maintenant posé dans le quart d'espace  $\{x_1, x_2 > 0\}$ . Le système prend donc la forme : pour  $d \geq 2$

$$\begin{cases} L(\partial)u = f & \text{pour } (t, x_1, x_2, x') \in \mathbb{R} \times \mathbb{R}_+^2 \times \mathbb{R}^{d-2} := \Omega, \\ B_1 u|_{x_1=0} = g_1 & \text{si } (t, x_2, x') \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^{d-2} := \partial\Omega_1, \\ B_2 u|_{x_2=0} = g_2 & \text{si } (t, x_1, x') \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^{d-2} := \partial\Omega_1, \\ u|_{t \leq 0} = 0 & \text{si } x \in \mathbb{R}_+^2 \times \mathbb{R}^{d-2} := \Gamma, \end{cases} \quad (\text{II.39})$$

Cette section est organisée de la façon suivante : dans le Paragraphe II.4.1, on donne un état de l'art concernant la théorie fortement bien posée pour le problème dans le quart d'espace. Le Paragraphe II.4.2 décrit, quant à lui, le premier résultat, à notre connaissance, établissant un caractère fortement bien posé en grande régularité et non pas uniquement dans l'espace  $L^2(\Omega)$ .

Le Paragraphe II.4.3 expose comment construire le développement d'optique géométrique pour la solution de (II.39) (sous des hypothèses simplificatrices). Il revient, en particulier, sur l'heuristique développée dans [Benoit, 2015] pour déterminer les phases obtenues par rebonds successifs, puis indique en quoi le développement d'optique géométrique peut venir illustrer la théorie fortement bien posée comme c'était le cas dans la géométrie de la bande. Enfin, ce paragraphe donne un dernier exemple de traitement des problèmes faiblement bien posés *via* des techniques d'optique géométrique.

Le Paragraphe II.4.4 introduit une nouvelle piste de recherche que nous avons récemment débutée à savoir l'établissement de développements d'optique géométrique dans un domaine borné. Pour ce faire, nous avons été amené, en particulier, à perfectionner les techniques de [Benoit, 2015]. Enfin, le Paragraphe II.4.5 contient des questions ouvertes, ainsi que des pistes de recherche que nous aimerions étudier à l'avenir.



## II.4.1 Une théorie bien posée encore très en friche

### (Bref) état de l'art

La théorie fortement bien posée pour les problèmes aux limites hyperboliques dans le quart d'espace est une théorie relativement ancienne qui apparaît essentiellement dans les années 1970 *via* les travaux d'Osher [Osher, 1973]-[Osher, 1974b]. Cette théorie a connu un regain d'intérêt récemment grâce aux travaux indépendants et complémentaires de Huang-Temam [Huang and Temam, 2014], Huang [Huang, 2015], Halpern-Rauch [Halpern and Rauch, 2016], Métivier-Rauch [Métivier and Rauch, 2017] et enfin l'auteur [Benoit, 2015]. Le but de ce paragraphe est de donner un panorama, que l'on espère assez exhaustif, de ces contributions. Toutefois, comme on le verra, il est à souligner que leur ensemble ne fournit (encore) qu'une théorie assez parcellaire.

A notre connaissance, les premiers travaux sur les problèmes hyperboliques posés dans un domaine à coin sont dûs à Sarason dans [Sarason, 1962]. Dans cette première mais fondamentale contribution, l'auteur étudie la validité d'un lemme de type "fort=faible" pour la solution d'un problème aux limites posé dans domaine spatial dont la frontière est anguleuse. Le résultat s'applique à des coins en dimension deux et est établi après la mise sous forme normale

$$A_1 \partial_1 u + A_2 \partial_2 u + \sigma u \rightsquigarrow \partial_1 u + A_1^{-1} A_2 \partial_2 u + \sigma A_1^{-1} u,$$

où  $\sigma$  est la variable duale de  $t$  par transformée de Laplace. En particulier, la non-caractéristicité des bords est supposée. A la connaissance de l'auteur, les problèmes dans le quart d'espace admettant un bord caractéristique constituent un domaine d'étude inexploré.

Dans la forme normale précédente, Sarason se ramène alors, à une analyse au cas par cas sur le spectre de la matrice  $A_1^{-1} A_2$ . On étudie alors des équations de transport, quand on est sur une valeur propre réelle simple, des équations de type parabolique, lorsque la valeur propre est réelle, associée à un bloc de Jordan et enfin des équations elliptiques, lorsque la valeur propre étudiée est complexe.

Le cas dans lequel la démonstration ne conclut pas et qui doit donc être exclu *via* une hypothèse est celui où la matrice au coin  $A_1^{-1} A_2$  admet un bloc de Jordan associé à une valeur propre réelle positive. L'hypothèse de Sarason, qui assure donc un résultat de type "fort=faible" pour la solution du problème hyperbolique (II.39), est donc la suivante :

**Hypothèse II.4.1 ([Sarason, 1962] Sarason)** *Les valeurs propres de la matrice dite au coin  $A_1^{-1} A_2$  associées à des blocs de Jordan sont négatives.*

Comme on le verra dans la suite de ce mémoire, notamment lors de l'établissement de la théorie de la persistance de la régularité du Paragraphe II.4.2, nous avons été amené à rencontrer une variante de l'Hypothèse de Sarason II.4.1 dans nos travaux.

Le contre-exemple de [Métivier and Rauch, 2017] montre aussi le rôle central de cette hypothèse et plus particulièrement de la matrice au coin  $A_1^{-1} A_2$  dans la théorie globale. En effet, dans l'acte de conférence [Métivier and Rauch, 2017], Métivier et Rauch donnent un contre-exemple à l'unicité de la solution d'un problème à coin lorsque la matrice au coin admet un bloc de Jordan réel (indépendamment de son signe).

Un tel résultat de non unicité de la solution d'un problème, qui on le rappelle est **linéaire**, illustre un caractère on ne peut plus pathologique du problème. Toutefois, on se doit d'insister sur le fait que le contre-exemple précédent dépend de ce que l'on entend par problème fortement bien posé. En effet, dans [Métivier and Rauch, 2017], la non unicité est obtenue en construisant deux solutions. L'une avec des traces  $L^2$  et la seconde avec des traces qui ne sont pas dans  $L^2$ . Si l'on définit le caractère fortement bien posé comme la généralisation de celui de Kreiss [Kreiss, 1970] et que l'on impose que les traces soient  $L^2$  alors on évite un tel contre-exemple.

Il n'en demeure pas moins que le contre-exemple de [Métivier and Rauch, 2017] illustre la propension du problème (II.39) à se comporter de façon pathologique.

Une variante du lemme "fort=fiable" de Sarason peut aussi être trouvée dans l'article de Peyser [Peyser, 1965] dans le cas particulier pour les matrices  $A_1$  et  $A_2$  commutent. Bien que cette hypothèse soit rarement vérifiée par les systèmes issus de la physique, car elle induit que l'on travaille essentiellement sur des équations de transport, cet article est intéressant car même s'il ne s'avère pas très (physiquement) applicable, sa démonstration est beaucoup plus facilement lisible que celle de [Sarason, 1962].

Les travaux de Sarason sur le caractère fortement bien posé du problème dans le quart d'espace sont assez nombreux et l'on doit citer [Sarason, 1977] et [Sarason, 6970], articles dans lesquels l'auteur se concentre sur les problèmes à coin symétriques, sans toutefois parvenir à un résultat de caractérisation du caractère fortement bien posé totalement satisfaisant. On peut aussi citer le résultat [Sarason, 1967] dans lequel Sarason passe par une technique de régularisation elliptique pour le problème.

Un autre résultat de Sarason conjoint avec Smoller est l'article [Sarason and Smoller, 7475], article pionnier et exploratoire dans lequel les auteurs donnent de précieuses intuitions sur quel devrait être le développement d'optique géométrique de la solution d'un problème à coin. Dans cet article, sans toutefois donner des développements rigoureux, la notion d'auto-interaction des phases, que l'on a rencontrée pour le problème dans la bande et qui devient une notion centrale pour le problème à coin, est introduite. Une construction rigoureuse des développements d'optique géométrique se basant sur les travaux de [Sarason and Smoller, 7475] peut être trouvée dans [Benoit, 2016] ou bien dans la thèse de l'auteur [Benoit, 2015].

A notre avis, la principale contribution dans l'établissement d'une caractérisation des problèmes fortement bien posés dans le quart d'espace, caractérisation à la Kreiss [Kreiss, 1970] est due à Osher dans sa série d'articles du début des années 1970, à savoir les trois articles [Osher, 1973],[Osher, 1974b] et [Osher, 1974a]. Décrivons succinctement ces résultats avant de revenir plus en détails dessus dans le Paragraphe II.4.1.

Osher considère des opérateurs hyperboliques quelconques et des conditions de bord matricielles génériques. Le but est alors de caractériser les conditions de bord qui donnent lieu à un problème fortement bien posé au sens de la généralisation naturelle de celle dans le demi-espace (voir Définition II.2.1). Sans grande surprise, ceci est démontré sur l'exemple explicite d'un système de transport interne découplé avec couplage sur le bord dans [Osher, 1974a] : imposer la condition de Kreiss-Lopatinskii uniforme sur chaque face du quart d'espace est une condition nécessaire mais non suffisante pour assurer le caractère fortement bien posé. Dans [Osher, 1973], Osher propose donc une nouvelle condition (on renvoie au Paragraphe II.4.1 pour plus de détails), grâce à laquelle on est en mesure de démontrer des estimations d'énergie *a priori*. La caractérisation des problèmes aux limites hyperboliques dans le quart d'espace aurait pu se conclure à ce stade, si ce n'est les deux écueils rencontrés par les travaux d'Osher. Plus précisément :

- Rien n'assure l'existence d'une solution (fiable).
- L'estimation d'énergie *a priori* proposée par Osher comprend un nombre de pertes de dérivées **non explicite**. Ainsi on est "assez" loin d'une estimation d'énergie établissant le caractère fortement bien posé dans  $L^2$ .

Malgré cela, le travail d'Osher reste le résultat le plus abouti dans la direction d'une caractérisation complète des problèmes dans le quart d'espace fortement bien posés.

Nous aimons particulièrement le fait que cette analyse tente, autant que faire se peut, de mimer au maximum l'analyse du problème dans le demi-espace notamment en se basant sur la construction

d'un *symétriseur*. Même si cette théorie n'a pas abouti depuis les années 1970, nous pensons que cette approche par *symétriseur* est une approche qui ne doit pas être négligée, ne serait-ce que parce qu'elle est extrêmement esthétique.

Nous ne donnerons pas, ici, plus de détails sur ces travaux, renvoyant au Paragraphe II.4.1 pour une description de la condition au coin d'Osher et au Paragraphe II.4.5 pour des pistes de recherche liées à cette analyse.

Nous pourrions nous dire, au vu du manque de résultats définitifs dans la littérature, que l'obtention d'une caractérisation complète des conditions de bord donnant lieu à des problèmes fortement bien posés est peut-être un problème trop ambitieux et qu'il serait de bon ton de réduire (dans un premier temps) nos ambitions.

Cet axe de recherche "simplifié" a été développé par plusieurs auteurs en restreignant l'étude à l'équation des ondes. On peut citer [Blokhnin and Tkachev, 1996a]-[Blokhnin and Tkachev, 1996b] dans lesquels les auteurs obtiennent des estimations d'énergie *a priori* pour des conditions de bord de type Robin. Toutefois, l'existence de la solution n'est pas établie. On se doit ici aussi de mentionner l'article [Blokhnin and Tkachev, 1996c] dans lequel il est montré, encore une fois sur l'équation des ondes, que la condition de Kreiss-Lopatinskii uniforme sur chaque face n'est pas suffisante pour avoir une estimation d'énergie *a priori*. On retrouve ainsi dans un cadre particulier la remarque d'Osher [Osher, 1974a]. Pour être exhaustif, on se doit aussi de citer les travaux antérieurs de [Kupka and Osher, 1971] complétés par les résultats de [Kojima and Taniguchi, 1980] et [Taniguchi, 1978].

En conclusion, à l'heure actuelle, on ne dispose pas d'une théorie fortement bien posée totalement convaincante pour l'équation des ondes posée dans un quart d'espace avec conditions de bord génériques. On pense que l'établissement d'une telle théorie, en se basant sur les résultats précédemment cités, serait toutefois intéressante et qu'elle constitue une piste que l'on aimerait développer à l'avenir.

Dans les travaux plus récents mentionnés au début de cette revue de la littérature, le problème a aussi été simplifié non pas en se restreignant à une équation particulière mais en se restreignant à une classe de conditions de bord (qui ne serait pas forcément optimale) pour laquelle on pourrait montrer le caractère fortement bien posé. Comme on l'a mentionné, l'hypothèse assurant que les conditions de bord sont strictement dissipatives simplifie grandement l'analyse dans le demi-espace. Ainsi, c'est souvent cette hypothèse simplificatrice qui a été retenue.

Dans l'article [Huang and Temam, 2014], Huang et Temam se placent dans un rectangle. Ainsi, ils n'ont pas deux conditions de bord à considérer mais quatre. Ils exposent alors des conditions de bord bien particulières pour lesquelles on est en mesure d'appliquer le théorème de Hille-Yosida. Ils obtiennent alors le caractère fortement bien posé, avec estimation du semi-groupe.

Ce travail a ensuite été légèrement généralisé dans [Huang, 2015] à des conditions de bord plus génériques où l'auteur montre le caractère fortement bien posé en appliquant, une nouvelle fois, le théorème de Hille-Yosida. Ceci se fait sous une condition purement algébrique sur les conditions de bord. Toutefois, le sens de la condition algébrique n'est pas discuté et il n'est pas clair que cette dernière recouvre les conditions de bord strictement dissipatives. Le résultat de Huang [Huang, 2015] se rapproche donc d'un résultat déjà existant de Peyser [Peyser, 1975] où le caractère bien posé est étudié via une condition dite de torsion sur les matrices  $A_1$  et  $A_2$ .

Pour des conditions de bord strictement dissipatives, le résultat le plus abouti est probablement dû à Halpern et Rauch dans [Halpern and Rauch, 2016]. Dans cet article, en raison de motivations de nature numérique, plus particulièrement l'étude des *Perfectly Matched Layers*, les auteurs se placent directement dans le cadre très intéressant d'un coin en trois dimensions. Ils supposent que les conditions de bord autour du coin sont toutes strictement dissipatives. Grâce à cela, par un argument de régularisation du quart d'espace en un demi-espace, ils parviennent à un caractère fortement bien posé dans l'espace  $L^2$ .

Ce résultat est très satisfaisant. Toutefois, en raison de la nature même des techniques utilisées pour sa démonstration, on pense en particulier à la régularisation du bord, il ne peut s'appliquer qu'à des problèmes aux limites ayant une propriété "d'ellipticité cachée" et qui ont en particulier le même nombre de conditions de bord sur chaque face du bord. Ceci exclut donc de nombreux cas de la mécanique des fluides où l'on peut être amené à linéariser autour d'un état qui est entrant pour l'une des faces et sortant pour la seconde.

Toujours pour les conditions strictement dissipatives, on conclura cette revue de la littérature en mentionnant le résultat de [Benoit, 2015] qui établit le caractère fortement bien posé des problèmes aux limites (symétriques) avec conditions de bord strictement dissipatives dans le cas d'un coin bidimensionnel. Mais comme on dit, *Redde Caesari quae sunt Caesaris et quae sunt Dei Deo*<sup>27</sup>, le mérite de ce résultat revient à Sarason. En effet, dans l'analyse de [Benoit, 2015], on se base de façon non triviale sur le résultat "fort=faible" de Sarason [Sarason, 1962]. Les autres points de la démonstration à savoir l'établissement d'une estimation d'énergie *a priori* et l'existence d'une solution faible étant assez directs, classiques et déjà connus, dans le cas des conditions de bord strictement dissipatives. On renvoie par exemple à [[Benzoni-Gavage, 2007]-Chapter 3].

### Le *symétriseur* d'Osher

On décrit ici avec plus de détails, les résultats sur le caractère fortement bien posé du problème aux limites (II.39) obtenus dans [Osher, 1973]. Comme mentionné précédemment ils ne sont pas suffisants pour conclure, mais l'approche par *symétriseur* à l'avantage d'être cohérente avec la théorie de Kreiss [Kreiss, 1970] dans le demi-espace.

Comme on va le voir, le coeur de l'analyse d'Osher dans [Osher, 1973] se base sur une nouvelle condition d'inversibilité proche de celle que l'on a décrite dans le Paragraphe II.3.1. Toutefois, là où la condition dans la bande permet de conclure, car elle se ramène à la dimension finie, celle dans le quart d'espace, qui fait intervenir des espaces de dimension infinie, ne permet pas (à l'heure actuelle) de conclure.

Des pistes de recherche concernant l'analyse d'Osher peuvent cependant être trouvées à la fin de ce paragraphe ainsi qu'au Paragraphe II.4.5.

Commençons par décrire la condition utilisée par Osher pour établir son estimation d'énergie *a priori*. L'obtention de cette dernière est à mettre en parallèle de celle dans le cas de la bande, voir Paragraphe II.3.1.

On choisit  $f, g_1 \equiv 0$  dans (II.39) et on suppose que ce problème à coin (II.39) est fortement bien posé<sup>28</sup>.

Soit  $u$  la solution de (II.39), on note  $v$  le prolongement de  $u$  par zéro pour les  $x_1$  négatifs. Alors  $v$  vérifie le problème aux limites (dans le demi-espace supérieur):

$$\begin{cases} A_1 \partial_1 v + A_2 \partial_2 v + [\sigma I + iA'(\eta)] v = \delta_0(x_1) A_1 u|_{x_1=0} & \text{pour } (x_1, x_2) \in \mathbb{R} \times \mathbb{R}_+, \\ B_2 v|_{x_2=0} = g_2 & \text{pour } x_1 \in \mathbb{R}. \end{cases} \quad (\text{II.40})$$

Pour obtenir le problème (II.40) à partir de (II.39) on a effectué la transformée de Laplace  $t \leftrightarrow \sigma$  et la transformée de Fourier tangentielle  $x' \leftrightarrow \eta \in \mathbb{R}^{d-2}$ . Le symbole tronqué dans (II.40) étant défini par  $A'(\eta) := \sum_{j=3}^d \eta_j A_j$ .

Dans le problème (II.40), la variable  $x_1$  vivant dans  $\mathbb{R}$ , on peut effectuer une transformée de Fourier

<sup>27</sup>Il faut rendre à César ce qui est à César et à Dieu ce qui est à Dieu. Marc, XII, 13-17

<sup>28</sup>On entend par là une adaptation directe du concept définissant le caractère fortement bien posé dans la bande de la Définition II.3.1 avec les traces idoines.

$x_1 \rightsquigarrow \eta_1$  pour se ramener à étudier le problème sous forme résolvente

$$\begin{cases} \frac{d}{dx_2} \widehat{v}(\eta_1) = \mathcal{A}_2(\eta_1) \widehat{v} + A_2^{-1} A_1 u|_{x_1=0} & \text{pour } x_2 > 0, \\ B_2 \widehat{v}|_{x_2=0}(\eta_1) = \widehat{g}_2(\eta_1), \end{cases} \quad (\text{II.41})$$

où l'on a posé  $\mathcal{A}_2(\eta_1) := \mathcal{A}_2(\sigma, \eta_1, \eta) := -A_2^{-1} [\sigma I + i\eta_1 A_1 + iA'(\eta)]$ .

Fort de l'analyse des problèmes aux limites dans le demi-espace, utilisant la formule de Duhamel, on peut alors décomposer  $\widehat{v}$  de la façon suivante :

$$\widehat{v}(\eta_1) = \Pi_2^s(\eta_1) \widehat{v}(\eta_1) + \Pi_2^u(\eta_1) \widehat{v}(\eta_1),$$

où chaque terme est donné par<sup>29</sup>

$$\begin{aligned} (\Pi_2^s \widehat{v})(\eta_1) &:= e^{x_2 \mathcal{A}_2(\eta_1)} \Pi_2^s(\eta_1) \widehat{v}|_{x_2=0} + \int_0^{x_2} e^{(x_2-y) \mathcal{A}_2(\eta_1)} \Pi_2^s(\eta_1) A_2^{-1} A_1 u|_{x_1=0}(y) dy \\ (\Pi_2^u \widehat{v})(\eta_1) &:= - \int_{x_2}^{\infty} e^{(x_2-y) \mathcal{A}_2(\eta_1)} \Pi_2^u(\eta_1) A_2^{-1} A_1 u|_{x_1=0}(y) dy, \end{aligned}$$

où  $\Pi_2^s$  (resp.  $\Pi_2^u$ ) désigne le projecteur sur le sous-espace stable (resp. instable) de  $\mathcal{A}_2$ , relativement à la décomposition de Hersh,  $\mathbb{C}^N = \mathbf{E}_2^s \oplus \mathbf{E}_2^u$ .

On utilise ensuite la condition de Kreiss-Lopatinskii uniforme pour la face  $\partial\Omega_2$ . La partie stable de la trace sur  $\partial\Omega_2$  s'exprime alors comme fonction de la partie instable. Plus précisément, on en déduit que

$$\widehat{v}|_{x_2=0}(\eta_1) = (\phi_2(\eta_1) B_2 - I) \int_0^{\infty} e^{-y \mathcal{A}_2(\eta_1)} \Pi_2^u(\eta_1) A_2^{-1} A_1 u|_{x_1=0}(y) dy.$$

Par transformée de Fourier inverse, on obtient alors une condition reliant la valeur de la trace  $u|_{x_2=0}$  en fonction de la trace  $u|_{x_1=0}$ . Cette dernière s'écrit

$$u|_{x_2=0}(x_1) = (\mathbf{T}_{1 \rightarrow 2} u|_{x_1=0})(x_1) + (\mathbf{S} g_2)(x_1), \quad (\text{II.42})$$

où l'on a défini  $\mathbf{T}_{1 \rightarrow 2} := \mathbf{T}_{1 \rightarrow 2}(\sigma, \eta)$  et  $\mathbf{S} := \mathbf{S}(\sigma, \eta)$ ,

$$\begin{aligned} (\mathbf{T}_{1 \rightarrow 2} u)(x_1) &:= \int_{\mathbb{R}} e^{ix_1 \eta_1} (\phi_2(\eta_1) B_2 - I) \int_0^{\infty} e^{-y \mathcal{A}_2(\eta_1)} \Pi_2^u(\eta_1) A_2^{-1} A_1 u(y) dy d\eta_1, \\ (\mathbf{S} g)(x_1) &:= \int_{\mathbb{R}} e^{ix_1 \eta_1} \phi_2(\eta_1) \widehat{g}(\eta_1) d\eta_1. \end{aligned} \quad (\text{II.43})$$

Réitérer des calculs similaires, mais après une extension par zéro pour les  $x_2$  négatifs, conduit à la relation :

$$u|_{x_1=0}(x_2) = (\mathbf{T}_{2 \rightarrow 1} u|_{x_2=0})(x_2), \quad (\text{II.44})$$

où  $\mathbf{T}_{2 \rightarrow 1}$  a une définition analogue à celle de  $\mathbf{T}_{1 \rightarrow 2}$ , *modulo* un changement d'indices.

Composer (II.42) et (II.44) donne alors la condition de compatibilité sur la valeur de la trace sur  $\partial\Omega_1$  suivante :

$$(I - \mathbf{T}_{1 \rightarrow 1}) u|_{x_1=0} = \mathbf{T}_{2 \rightarrow 1} \mathbf{S} g_2, \quad (\text{II.45})$$

<sup>29</sup>Les formules suivantes sont à lire en parallèle de celles dans la bande (voir (II.28) et (II.29)). On retrouve exactement la même structure. Toutefois, là où la partie du terme source portant la trace à l'intérieur de (II.27) était indépendant de la variable pour laquelle on applique la formule de Duhamel, ce n'est plus le cas pour celui de (II.41). Par conséquent, l'intégrale ne se découple pas comme dans (II.28) et (II.29). Comme on va le voir, ceci impliquera en particulier que les opérateurs de traces dans le cas du quart d'espace sont maintenant donnés par des opérateurs intégraux de Fourier.

où l'on a posé  $\mathbf{T}_{1 \rightarrow 1} := \mathbf{T}_{2 \rightarrow 1} \mathbf{T}_{1 \rightarrow 2}$ .

Ainsi, on remarque que l'on a une condition définissant la trace sur  $\partial\Omega_1$  très analogue à celle de la bande (voir Paragraphe II.3.1). La nouvelle condition fait toutefois intervenir des opérateurs intégraux de Fourier. A notre avis, le fait que la condition de détermination de la trace (II.45) fasse intervenir des opérateurs aussi techniques est l'une des principales raisons expliquant pourquoi, à l'heure actuelle, la caractérisation des conditions de bord donnant des problèmes fortement bien posés n'a pas encore été achevée. Comme on va le voir, cette caractérisation pourrait passer par l'analyse d'Osher combinée à une analyse poussée des opérateurs de trace  $\mathbf{T}_{1 \rightarrow 2}$  et  $\mathbf{T}_{2 \rightarrow 1}$ .

L'hypothèse d'Osher est alors la suivante<sup>30</sup> :

**Hypothèse II.4.2 (Osher [Osher, 1973])** *Pour tout paramètre de fréquence  $\zeta \in \Xi \setminus \Xi_0$ , l'opérateur  $I - \mathbf{T}_{1 \rightarrow 1}(\zeta)$  est uniformément inversible de  $L^2(\mathbb{R}_+)$  dans  $L^2(\mathbb{R}_+)$ . En particulier, il existe  $C > 0$ , telle que pour tout  $\zeta \in \Xi \setminus \Xi_0$ , et pour tout  $u \in L^2(\mathbb{R}_+)$ , on ait*

$$\|u\|_{L^2(\mathbb{R}_+)} \leq C \|(I - \mathbf{T}_{1 \rightarrow 1}(\zeta))u\|_{L^2(\mathbb{R}_+)}.$$

Donnons ici quelques commentaires sur cette hypothèse :

- Il est assez remarquable que l'Hypothèse II.4.2 soit une condition d'inversibilité uniforme exactement comme c'était le cas dans le demi-espace pour la condition de Kreiss-Lopatinskii uniforme. Ceci confère un caractère assez cohérent à la théorie globale des problèmes aux limites hyperboliques.
- Toutefois, là où les conditions de Kreiss-Lopatinskii uniforme et celle dans la bande sont des conditions d'inversibilité sur un certain sous-espace, dans sa formulation actuelle la condition d'Osher est une condition que l'on impose sur tout l'espace. Nous aimerions, à l'avenir, voir si nous ne pourrions pas utiliser les arguments développés dans la bande pour raffiner l'Hypothèse II.4.2. Cette nouvelle condition demanderait alors l'inversibilité seulement sur un sous-espace de  $L^2$ , mais elle permettrait toujours de construire un *symétriseur*. Cela pourrait possiblement permettre de construire une solution au problème dans le quart d'espace (II.39). En effet, l'obstruction actuelle est que l'on ne sait pas montrer que l'Hypothèse d'Osher II.4.2 passe au problème dual. Dans le cas de la bande, on a vu qu'une hypothèse d'inversibilité sur tout l'espace ne passait pas au problème dual, mais que la version raffinée, elle, passait au dual. Démontrer l'existence d'une solution faible pour le problème à coin serait à notre avis un achèvement important dans la théorie fortement bien posée.
- Indiquons toutefois que l'Hypothèse II.4.2 est vraiment très loin d'être anodine et confère à la théorie actuelle un air de *if Theorem*. En effet, la condition de Kreiss-Lopatinskii (qui revient juste à vérifier l'inversibilité d'une matrice) est déjà, comme on l'a dit, très difficile à vérifier en pratique, notamment en raison du caractère uniforme de l'inversibilité. Ici, nous sommes en dimension infinie, ainsi le problème est forcément plus complexe. Une piste de recherche pour éviter ce caractère uniforme de la borne pourrait être de procéder comme pour la condition de Kreiss-Lopatinskii uniforme en compactifiant la boule de l'espace des fréquences et de considérer les fréquences du bord  $\Xi_0$ . On a déjà mentionné une motivation similaire dans le cas de la bande. En particulier, il faudrait comprendre comment réagissent les opérateurs de trace au voisinage de fréquences de *glancing* pour lesquelles leur définition est loin d'être claire (les projecteurs  $\Pi_{1,2}^{s,u}$  n'étant plus définis). C'est en partie pour cela que l'on a cherché, grâce aux développements d'optique géométrique, à comprendre comment pourraient réagir ces opérateurs pour des fréquences du bord. On donnera plus de détails sur ces développements dans les Paragraphes II.4.3 et II.4.4.

---

<sup>30</sup>Il est ici à noter que le caractère nécessaire dans [Osher, 1973] n'est pas totalement bien établi. On a toutefois bon espoir que l'adaptation des arguments que l'on a été amené à développer dans la bande dans [Benoit, 2020a] puisse rendre ce point de la démonstration totalement rigoureux.

Grâce à cette hypothèse, Osher construit un *symétriseur* qui permet d'établir une estimation d'énergie *a priori*. On ne décrira pas dans le détail ici la construction du *symétriseur* qui s'avère assez technique. Là encore, on pourra se faire une bonne idée de cette construction en considérant le cas discret décrit au Paragraphe III.4.2.

Nous trouvons remarquable que cette construction ait pu être juste légèrement modifiée afin de traiter le problème dans la bande. Ceci est à notre sens un fort argument en faveur du fait que cette construction devrait aussi opérer dans le cas du quart d'espace.

La construction du *symétriseur* d'Osher se fait dans les variables polaires qui semblent être un choix très naturel, on verra en particulier tout l'intérêt de ce choix de coordonnées dans le Paragraphe II.4.2, pour étudier le problème dans le quart d'espace (II.39). Après passage dans les variables polaires, le problème (II.39) (toujours après transformation de Laplace-Fourier) s'écrit

$$\begin{cases} \mathcal{A}(\theta)\partial_\theta u + \mathbf{M}u = 0 & \text{pour } (\theta, \mathbf{r}) \in ]0, \frac{\pi}{2}[ \times \mathbb{R}_+, \\ B_1 u|_{\theta=\frac{\pi}{2}} = 0 & \text{pour } \mathbf{r} \in \mathbb{R}_+, \\ B_1 u|_{\theta=0} = g_2 & \text{pour } \mathbf{r} \in \mathbb{R}_+, \end{cases} \quad (\text{II.46})$$

où l'opérateur  $\mathbf{M}$  est défini par

$$\mathbf{M} := \mathbf{M}(\zeta, \theta, \mathbf{r}) = (\sin \theta A_2 + \cos \theta A_1) \mathbf{r} \partial_{\mathbf{r}} + \mathbf{r} (\sigma I + i A'(\eta)),$$

et où la matrice  $\mathcal{A}$  est définie par  $\mathcal{A}(\theta) := \cos \theta A_2 - \sin \theta A_1$ .

Un avantage certain de cette formulation dans les variables polaires est qu'elle permet de ramener les deux conditions de bord sur  $\partial\Omega_1$  et sur  $\partial\Omega_2$  comme des conditions de bord exprimées selon la même variable, à savoir, la variable angulaire  $\theta$ . On échange donc un problème dans le quart d'espace contre un problème dans une bande<sup>31</sup>.

L'apport remarquable de l'analyse d'Osher [Osher, 1973] est d'avoir identifié, puis construit, un bon *symétriseur*. Les exigences sur cet objet sont les suivantes :

**Définition II.4.1** [*Symétriseur d'Osher, [Osher, 1973]*] Un opérateur  $\mathcal{S}(\zeta, \theta)$  pour  $\theta \in [0, \frac{\pi}{2}]$  et  $\zeta \in \Xi \setminus \Xi_0$  est appelé *symétriseur d'Osher*, si il vérifie les propriétés suivantes :

- i) L'opérateur  $\mathcal{S}\mathcal{A}$  est symétrique sur  $L^2(\mathbb{R}_+)$ .
- ii) L'opérateur  $\mathcal{S}$  est borné sur  $L^2(\mathbb{R}_+)$ , uniformément par rapport aux variables  $(\zeta, \theta) \in (\Xi \setminus \Xi_0) \times [0, \frac{\pi}{2}]$ .
- iii) L'opérateur  $\mathcal{S}$  est solution de l'équation différentielle

$$\frac{1}{2} \partial_\theta (\mathcal{S}\mathcal{A}) + \text{Re}(\mathcal{S}\mathbf{M}) = 0, \quad (\text{II.47})$$

où  $\text{Re}(\cdot) := \frac{1}{2}(\cdot + \cdot^*)$  désigne la partie réelle d'un opérateur.

- iv) Pour  $\theta = \frac{\pi}{2}$ , il existe deux constantes  $\varepsilon_1, C_1 > 0$ , telles que indépendamment de  $\zeta$ , l'on ait :

$$\forall u \in L^2(\mathbb{R}_+), \langle u, \mathcal{S}(\zeta, \frac{\pi}{2})u \rangle \geq \varepsilon_1 \|u\|_{L^2(\mathbb{R}_+)}^2 - C_1 \|B_1 u\|_{L^2(\mathbb{R}_+)}^2. \quad (\text{II.48})$$

- v) Pour  $\theta = 0$ , il existe deux constantes  $\varepsilon_2, C_2 > 0$ , telles que indépendamment de  $\zeta$ , l'on ait :

$$\forall u \in L^2(\mathbb{R}_+), \langle u, \mathcal{S}(\zeta, 0)u \rangle \geq \varepsilon_2 \|u\|_{L^2(\mathbb{R}_+)}^2 - C_2 \|B_2 u\|_{L^2(\mathbb{R}_+)}^2. \quad (\text{II.49})$$

---

<sup>31</sup>Remarquons ici que comme l'opérateur  $\mathbf{M}$  n'est pas hyperbolique, les résultats du Paragraphe II.2 ne s'appliquent malheureusement pas.

On peut assez aisément se convaincre qu'avec un tel *symétriseur* en main, la démonstration de l'estimation d'énergie *a priori* se ramène à celle dans le cas symétrique à conditions de bord strictement dissipatives.

On donne dans ce qui suit quelques remarques sur le *symétriseur* d'Osher :

- Les points *iv)* et *v)* de la Définition II.3.2 servent à contrôler la solution sur le bord du domaine. Ils sont totalement similaires aux points *iv)* et *v)* de la Définition II.3.2. Un tel contrôle au bord est aussi présent dans le *symétriseur* de Kreiss [Kreiss, 1970] (voir aussi [[Chazarain and Piriou, 1981]-Section 7.5] ou [[Benzoni-Gavage, 2007]-Chapter 5]). Ainsi, ils sont totalement naturels dans le cas du quart d'espace.
- Les points *i)* et *iii)* de la Définition II.4.1 servent à assurer que le *symétriseur* se comporte bien pour l'équation intérieur et qu'il permet ainsi, de se ramener au cas symétrique à conditions de bord strictement dissipatives. Là encore, il faut lire en parallèle le point *iii)* de la Définition II.4.1 avec le point *iii)* de la Définition II.3.2 dans la bande.
- D'un certain point de vue, comme on vient de le voir, les *symétriseurs* dans la bande et dans le quart d'espace doivent vérifier essentiellement les mêmes conditions. Ainsi ce n'est pas surprenant que l'on ait pu adapter assez facilement la construction donnée par Osher dans le cas du quart d'espace [Osher, 1973] pour établir le caractère fortement bien posé optimal dans la bande dans [Benoit, 2020a].

Mais, d'un autre point de vue, ces deux types de *symétriseurs* diffèrent de façon essentielle. Il ne s'agit pas des mêmes objets mathématiques. En effet, là où le *symétriseur* dans la bande est un opérateur en dimension infinie, mais de multiplication son analogue dans le cas du quart d'espace agit pleinement sur des espaces de dimension infinie. Ceci a pour conséquence que le point *ii)* de la Définition II.4.1, point qui assure une borne uniforme sur le *symétriseur*, devient délicat à démontrer dans le cas de la dimension infinie tandis que ce dernier était relativement aisé pour les opérateurs de multiplication. Comme on va le voir, ce point est au cœur du fait que l'on n'arrive pas, à l'heure actuelle, à caractériser les problèmes à coin fortement bien posés. On rappelle que ce point est utilisé de façon non triviale pour assurer une borne uniforme du *symétriseur*. Grâce à cette dernière, on peut utiliser l'estimation d'énergie *a priori* pour le problème résolvant afin de revenir dans les variables de position  $(t, x')$ .

Sous les hypothèses de structure idoines (hyperbolicité, non-caractéristicité,...) et surtout en se placant sous l'Hypothèse II.4.2, Osher démontre que le problème à coin (II.39) admet un *symétriseur* au sens de la Définition II.4.1. Il obtient alors une estimation d'énergie *a priori* de la forme

$$\begin{aligned} \gamma \|u\|_{L^2_\gamma(\Omega)}^2 + \|u|_{x_1=0}\|_{L^2_\gamma(\partial\Omega_1)}^2 + \|u|_{x_2=0}\|_{L^2_\gamma(\partial\Omega_2)}^2 \\ \leq C \left( \frac{1}{\gamma} \|f\|_{H^s_\gamma(\Omega)}^2 + \|g_1\|_{H^{s_1}_\gamma(\partial\Omega_1)}^2 + \|g_2\|_{H^{s_2}_\gamma(\partial\Omega_2)}^2 \right), \end{aligned} \quad (\text{II.50})$$

où les réels  $s, s_1, s_2 > 0$ , ne sont pas explicités dans la démonstration de [Osher, 1973].

Expliquons d'où proviennent ces pertes de dérivées. Elles sont dues à la construction même du *symétriseur* d'Osher et plus particulièrement dans la vérification de la borne uniforme *ii)* de la Définition II.4.1.

Le *symétriseur* dépend de l'opérateur  $(I - \mathbf{T}_{1 \rightarrow 1})^{-1}$ , sur lequel, par hypothèse, on a un bon contrôle uniforme de la norme  $L^2$ , mais aussi, et c'est là où le bât blesse, des opérateurs de traces  $\mathbf{T}_{1 \rightarrow 2}$  et  $\mathbf{T}_{2 \rightarrow 1}$  qui ne sont pas, par hypothèse, uniformément bornés dans  $L^2(\mathbb{R}_+)$ . Pour les contrôler Osher paye des pertes de dérivées qui se répercutent ensuite dans (II.50). Ainsi, en reprenant proprement les travaux



d'Osher, on pourrait éventuellement établir une estimation d'énergie *a priori* avec un nombre de pertes de dérivées rendu explicite. Résultat non optimal, mais qui aurait le mérite de faire avancer les lignes.

L'un des apports de [[Benoit, 2015]-Chapitre 5] est d'avoir montré que si les opérateurs de traces  $\mathbf{T}_{1 \rightarrow 2}$  et  $\mathbf{T}_{2 \rightarrow 1}$  étaient sans pertes, c'est-à-dire bornés uniformément de  $L^2(\mathbb{R}_+)$  à valeurs dans  $L^2(\mathbb{R}_+)$ , alors l'estimation d'énergie *a priori* (II.50) était en fait sans pertes. Nous avons aussi montré dans [Benoit, 2015], sous des hypothèses de structure<sup>32</sup> sur les matrices de Sarason  $A_1^{-1}A_2$  et  $A_2^{-1}A_1$ , le caractère borné  $L^2$  à valeurs dans  $L^2$  de ces opérateurs. A l'heure actuelle, seul le caractère uniforme de la borne manque pour conclure. Plus de détails sur cette borne seront donnés au Paragraphe II.4.5.

## II.4.2 Un résultat de persistance de la régularité

Ce paragraphe décrit l'article [Benoit, 2024] qui donne un résultat de persistance de la régularité de la solution du problème posé dans le quart d'espace. A travers les démonstrations, afin de pallier de nouvelles difficultés induites par la géométrie du quart d'espace (voir la première partie du Paragraphe II.4.2), on a été amené à reconsidérer la théorie de persistance de la régularité pour la solution du problème dans le demi-espace dans l'article [Benoit, c]. Cette description occupe la seconde partie du Paragraphe II.4.2.

L'adaptation de cette méthode au quart d'espace est décrite dans la troisième partie du Paragraphe II.4.2. Enfin, le Paragraphe II.4.2 comprend des perspectives liées à notre résultat de persistance de la régularité.

### Principale difficulté et analogie caractéristique

Les deux grandes difficultés dans l'étude de la persistance de la régularité de la solution du problème dans le quart d'espace (II.39) sont les suivantes :

1. "Il y a trop de directions normales" pour régulariser.
2. "Il y a trop de dérivées normales" qui se comportent "mal" avec l'équation.

Pour bien illustrer ces difficultés, il est utile de se rappeler la méthode pour obtenir la persistance de la régularité de la solution d'un problème non caractéristique dans le demi-espace :

- tout d'abord dans le cas du demi-espace, tous les opérateurs  $\partial_t$  et  $\partial_j$ ,  $j \in \llbracket 1, d-1 \rrbracket$  commutent avec l'équation. Ainsi on n'a pas la difficulté 2., car seule la dérivée normale  $\partial_d$  se comporte mal avec l'équation en raison de la condition de bord en  $\{x_d = 0\}$ .
- Ensuite, en convolant par une approximation de l'unité dépendant des variables tangentielles  $(t, x')$ , on peut venir régulariser les dérivées tangentielles. Là encore, notons qu'en raison de la condition de bord en  $\{x_d = 0\}$ , une régularisation par rapport à  $x_d$  viendrait modifier la condition de bord. On peut donc facilement montrer que la solution  $u \in H_{t,x'}^s(\mathbb{R} \times \mathbb{R}^{d-1})$ .
- Pour récupérer la régularité manquante on utilise alors l'équation interne grâce à laquelle on vient exprimer

$$\partial_d u := A_d^{-1} \left[ f - \partial_t u - \sum_{j=1}^{d-1} A_j \partial_j u \right], \quad (\text{II.51})$$

qui montre donc que  $\partial_d u \in L^2$  et donc par itérations que  $u \in H_{t,x}^s(\mathbb{R} \times \mathbb{R}_+^d)$ .

---

<sup>32</sup>On se ramène dans cet argument à une analyse mode par mode pour la matrice résolvante  $\mathcal{A}_1(0,1) := A_1^{-1}A_2$ . Le cas qui échappe à l'analyse et qui doit être exclu par hypothèse, est celui où l'on a une fréquence de *glancing*, hypothèse physiquement admissible.

Notons, pour de futures discussions, que l'on utilise ici fortement le fait que le problème soit non caractéristique. En effet, la difficulté pour les problèmes caractéristiques est que l'équation interne ne fournit plus la régularité manquante de la solution.

Dans le cas du quart d'espace, nous avons donc deux points à garder à l'esprit :

1. d'une part, les opérateurs  $\partial_1$  et  $\partial_2$  se comportent mal avec l'équation. On devra donc leur suppléer d'autres opérateurs qui eux se comportent mieux.
2. D'autre part, on ne peut pas régulariser par rapport ni à  $x_1$  ni à  $x_2$ . Par conséquent, l'équation interne implique uniquement que  $A_1\partial_1u + A_2\partial_2u \in L^2$ , qui est bien loin d'impliquer que  $\partial_1u$  et  $\partial_2u$  sont  $L^2$ . Il est toutefois important de remarquer que ce problème est totalement similaire au cas caractéristique dans le demi-espace. On entend par là que dans les deux cas l'équation interne ne fournit pas la régularité manquante.

Pour pallier la difficulté 2. nous avons choisi une méthode, quelque peu radicale, qui vise à obtenir une solution régulière sans utiliser la méthode de régularisation par une approximation de l'unité. Cette méthode est décrite dans le paragraphe suivant. Le point 1. est, quant à lui, inspiré de l'analyse du problème caractéristique pour lequel on sait voir par exemple [Rauch, 1985] qu'un bon opérateur à considérer est l'opérateur tangentiel  $x_1\partial_1$ . On donne plus de détails dans le second paragraphe.

### Description de la méthode dans le demi-espace

Nous exposons ici les résultats de [Benoit, c]. Ils établissent l'existence d'une solution régulière du problème (caractéristique ou non) dans le demi-espace sous la condition que ce dernier soit  $L^2$  bien posé. En particulier, les résultats de [Benoit, c] ne sont pas nouveaux et peuvent par exemple être retrouvés dans [Rauch and Massey, 1974], [Rauch, 1985], comme un cas particulier de [Guès, 1990]. Nos résultats ne retrouvent pas, à l'heure actuelle, les résultats optimaux de Secchi [Secchi, 1995], ou encore les articles de Secchi et de ses coauteurs précédemment cités.

Toutefois, comme indiqué précédemment, le fait que l'on évite d'utiliser la méthode classique de régularisation par approximation de l'unité introduit un nouveau procédé, qui a l'intérêt certain de pouvoir s'étendre facilement au cas du quart d'espace, là où, la régularisation précédente bloque devant les trop nombreuses "directions normales".

De plus, nous pensons que cette méthode, parce qu'elle évite des points de régularisation techniques non essentiels, fournit une démonstration plus directe et peut-être plus esthétique des résultats déjà présents dans la littérature.

Deux points doivent toutefois être formulés avec clarté. Tout d'abord, l'idée de l'une de nos méthodes développées dans [Benoit, c], à savoir l'usage du théorème de Hille-Yosida, se base très fortement sur une idée déjà présente dans [[Benzoni-Gavage, 2007]-Chapter 3]. Nous l'avons simplement étendue de l'espace  $L^2$  vers l'espace de Sobolev  $H^s$  et généralisée à une plus grande classe de problèmes. D'autre part, à l'heure actuelle, les résultats de [Benoit, c] ne permettent pas de prendre en compte le cas de coefficients  $A_j$  variant avec le temps, ce qui réduit leur applicabilité, par exemple à des problèmes quasi-linéaires. Toutefois il est important de rappeler que nous ne cherchions pas une méthode qui permette de retrouver des résultats quasi-linéaires déjà connus. Nous cherchions une méthode, qui, elle, permettait de passer à la géométrie du quart d'espace.

On propose toutefois des pistes pour lever cette obstruction à la fin de ce paragraphe.

L'article [Benoit, c] fournit deux méthodes qui permettent d'obtenir une solution régulière du problème aux limites dans le demi-espace. Le problème que l'on étudie peut-être caractéristique ou éventuellement

à coefficients variables par rapport à la variable d'espace  $x$ . Ces deux méthodes se basent sur les méthodes déjà existantes suivantes :

- La méthode usuelle de [Lax and Phillips, 1960], qui donne l'existence de la solution dans  $L^2$  par dualité. Cette méthode s'applique en particulier à des conditions de bord inhomogènes. On a ici raffiné cette méthode en l'adaptant pour obtenir directement l'existence d'une solution dans un espace de Sobolev de grand ordre. L'hypothèse fondamentale est de supposer que le problème dual admet une estimation d'énergie *a priori* en grande norme. Ceci est facilement vérifiable dans le cas des systèmes symétriques à conditions de bord strictement dissipatives (ou encore pour la condition de Kreiss-Lopatinskii uniforme) pour lesquels ces hypothèses assurant l'estimation *a priori* primale passent au dual.
- Le théorème de Hille-Yosida dans son énoncé du à Lumer et Philips [Lumer and Phillips, 1961] dont un exemple d'applicabilité dans le cas  $L^2$ , non caractéristique, peut être trouvé dans l'ouvrage [[Benzoni-Gavage, 2007]-Chapter 3]. Cette méthode s'applique uniquement (pour l'instant) à des conditions de bord homogènes et dissipatives. Toutefois, elle permet de traiter directement des conditions initiales inhomogènes, sans avoir à passer par le principe de prolongation  $L^2$  de Rauch [Rauch, 1972].

Dans ce qui suit, on exposera uniquement les idées principales de la deuxième méthode. Toute l'idée consiste ici à voir le problème dans le demi-espace comme un problème d'évolution dans la variable temporelle. On se ramène donc à considérer l'équation

$$\begin{cases} \frac{d}{dt}u = Au & \text{pour } t > 0, \\ u|_{t=0} = u_0, \end{cases} \quad (\text{II.52})$$

où  $A$  est l'opérateur défini par  $A := \sum_{j=1}^d A_j \partial_j$  et où la condition de bord est prise en compte dans le domaine de définition  $D(A)$  de  $A$ . En particulier, pour que  $D(A)$  ait une structure d'espace vectoriel la condition de bord doit être homogène<sup>33</sup>.

On cherche alors à appliquer la théorie des semi-groupes au système (II.52). Pour cela, utilisant le théorème de Lumer-Philips [Lumer and Phillips, 1961], il est suffisant de trouver un espace de Banach  $X$  tel que l'opérateur  $A : D(A) \rightarrow X$  soit fermé, à domaine dense et dissipatif, ainsi qu'un réel  $\lambda_0 \geq 0$  tel que  $\text{Im}(\lambda_0 I - A) = X$ .

Si l'on se restreint pour  $X$  à un espace de Hilbert, la définition de l'opérateur de  $A$  ne pose pas de difficulté et la seconde hypothèse peut-être remplacée par la dissipativité de l'opérateur dual. En effet, rappelons que la dissipativité d'un opérateur  $A$  requiert que

$$\forall \lambda > 0, \forall u \in D(A), \|(\lambda I - A)u\|_X \geq \lambda \|u\|_X,$$

qui implique en particulier l'injectivité de l'opérateur  $\lambda I - A$ . Une telle injectivité pour l'opérateur dual donne alors, la surjectivité de l'opérateur primal recherchée.

Pour conclure à un résultat de persistance de la régularité, il est alors suffisant de se placer dans l'espace  $X = H_{x'}^s(\mathbb{R}^{d-1})$  et de montrer que les opérateurs  $A$  et  $A^*$  sont dissipatifs sur cet espace. Ce résultat est facilement obtenu par une intégration par parties dans laquelle les termes de bord ont le bon goût d'avoir le signe voulu, si l'on se place dans le cadre des conditions de bord strictement dissipatives. En particulier la dissipativité de  $A^*$  se base sur le fait que l'opérateur  $A$  est anti-autoadjoint. Dans le

<sup>33</sup>Notons ici, au passage, que si les coefficients  $A_j$  dépendent du temps alors il en est de même pour le domaine de définition  $D(A)$ . Chose que l'on a cherché à éviter.

cas non caractéristique, la régularité manquante par rapport à  $x_d$  est alors obtenue en utilisant l'équation interne comme expliqué dans le Paragraphe II.4.2.

Pour étendre notre résultat au cas des coefficients variables (en espace) il suffit de reprendre la même méthode mais en remplaçant la notion de dissipativité par celle de quasi-dissipativité<sup>34</sup>. De même, notre méthode se généralise sans changement notable dans la structure de la démonstration au cas caractéristique. Il suffit, en effet, de changer l'espace de Hilbert de base en  $X = H_{tan}^s$  défini par

$$H_{tan}^s(\Omega) := \left\{ u \in \mathcal{D}'(\Omega) \setminus \forall \alpha \in \mathbb{N}^d, |\alpha| \leq s, \partial_1^{\alpha_1} \cdots \partial_{d-1}^{\alpha_{d-1}}(x_d \partial_d)^{\alpha_d} u \in L^2(\Omega) \right\},$$

pour retrouver le résultat de régularité tangentielle dans le cas caractéristique de [Rauch, 1985].

On conclut ce paragraphe en donnant des éléments de réflexion avant de pouvoir généraliser nos méthodes au cas de coefficients variables par rapport au temps. Comme déjà indiqué cette généralisation n'est, à notre avis, pas capitale dans le cas du demi-espace, puisqu'elle est principalement utilisée pour traiter des problèmes quasi-linéaires. Pour de tels problèmes, la théorie est bien développée par des méthodes déjà existantes et bien rodées. Obtenir une méthode à coefficients variables en temps serait intéressant pour sa transposabilité au quart d'espace, cette piste nous semble être importante à étudier.

Pour ce faire commençons par noter qu'une approche "naïve" et quelque peu non optimale se basant sur un doublement des variables peut être employée. Cette méthode consiste à ne pas considérer l'équation  $\frac{d}{dt}u = A(t, x)u$  mais l'équation  $\frac{d}{dt}\tilde{u}(t, s, x) = \tilde{A}(s, x)\tilde{u}(t, s, x)$  suppléée de conditions de bords idoines, à lui appliquer la méthode précédemment décrite dans un espace de Sobolev incluant les variables  $x$  et  $s$  et à obtenir alors la solution recherchée en prenant la trace  $\tilde{u}|_{s=t}$ . Cette dernière pouvant être contrôlée par injection de Sobolev.

Le point noir de cette méthode est qu'elle va typiquement nécessiter pour fonctionner une régularité des coefficients en  $W^{s+p, \infty}$  (au lieu de  $W^{s, \infty}$ ) et plus gênant, elle va induire une perte de  $p$  dérivées par rapport aux données. Dans ce qui précède  $p \geq \frac{d}{2}$  est choisi assez grand pour utiliser la dernière injection de Sobolev. Par conséquent, cette méthode n'est pas optimale. Elle a cependant l'avantage de la simplicité et indique que l'usage de notre méthode dans le cas de coefficients variables en temps est possible.

Pour obtenir une méthode optimale nous avons, à l'heure actuelle, deux grandes pistes de recherche. La première consiste à s'intéresser plus en détails à la littérature abondante (on pense aux nombreux travaux de Kato dont [Kato, 1956] constitue un exemple) des années 1970 sur les équations de la forme  $\frac{d}{dt}u = A(t)u$  et essayer de trouver l'un des théorèmes d'existence (abstrait) qui s'appliquerait à notre cas particulier du problème dans le demi-espace à conditions de bord strictement dissipatives. La seconde, qui vient rejoindre la première, proposée par J. Rauch, consiste à revenir aux fondamentaux concernant l'étude des semi-groupes pour les problèmes aux limites hyperboliques en étudiant l'article de [Massey, 1972] dans lequel ce résultat est obtenu, pour des coefficients dépendant du temps, dans l'espace de Sobolev  $H^1(\mathbb{R}_+^d)$ . Il s'agirait alors de généraliser ce résultat aux ordres supérieurs.

## Le résultat de persistance

Dans ce paragraphe on décrit plus avant les résultats de [Benoit, 2024]. En particulier, la résolution des deux difficultés principales exposées dans la première partie du paragraphe est discutée. Commençons par revenir sur la grande ressemblance de notre problème avec le problème caractéristique dans le demi-espace. On rappelle que dans un tel cadre, le "bon" opérateur différentiel à considérer n'est pas  $\partial_d$ , mais plutôt  $x_d \partial_d$ .

<sup>34</sup>Un léger travail est aussi nécessaire pour montrer que l'opérateur dual est essentiellement anti-autoadjoint au sens où il s'écrit  $A^* = -A + \Phi$ , où  $\Phi$  est linéaire, continu.

En effet, on a alors  $L(\partial)(x_d \partial_d u) = L(\partial)f + A_d \partial_d u$ , équation où l'on peut venir exprimer le second terme du membre de droite de façon indépendante de  $\partial_d u$  en venant lui substituer la relation donnée par l'équation  $L(\partial)u = f$ . De plus, on a clairement que  $x_d \partial_d u$  vérifie la condition de bord homogène. On peut donc appliquer la théorie  $L^2$  bien posée à  $x_d \partial_d u$ .

Fort de cela, on remarque rapidement que dans le cas du quart d'espace un "bon" opérateur différentiel à considérer est l'opérateur de dérivation radiale  $\mathbf{r}\partial_{\mathbf{r}} := x_1 \partial_1 + x_2 \partial_2$ . Ce dernier vérifie alors l'équation

$$L(\partial)\mathbf{r}\partial_{\mathbf{r}}u = \underbrace{A_1 \partial_1 u + A_2 \partial_2 u}_{=L(\partial)u - \sum_{j=3}^d A_j \partial_j u} + \mathbf{r}\partial_{\mathbf{r}}L(\partial)u,$$

équation dont le membre de droite peut-être supposé dans  $L^2$  par régularisation tangentielle. On a aussi les conditions de bord

$$\begin{cases} B_1(\mathbf{r}\partial_{\mathbf{r}}u)|_{x_1=0} = (x_2 \partial_2)B_1 u|_{x_1=0}, \\ B_2(\mathbf{r}\partial_{\mathbf{r}}u)|_{x_2=0} = (x_1 \partial_1)B_2 u|_{x_2=0}. \end{cases}$$

Par conséquent, on peut venir appliquer la théorie  $L^2$  bien posée pour obtenir que  $x_1 \partial_1 u + x_2 \partial_2 u \in L^2(\Omega)$ . Bien sûr, ce résultat n'est pas suffisant de prime abord pour récupérer que chaque dérivée tangentielle  $x_1 \partial_1 u$  et  $x_2 \partial_2 u$  est dans  $L^2(\Omega)$ .

Tout le jeu dans ce qui suit, va être de donner une méthode simple qui permette d'obtenir une telle conclusion. L'idée<sup>35</sup> est alors de chercher un contrôle de la dérivée angulaire  $\partial_{\theta} := -x_2 \partial_1 + x_1 \partial_2$ . Pour cela, on effectue un changement de variables en coordonnées polaires  $(\mathbf{r}, \theta) \in \mathbb{R}_+ \times ]0, \frac{\pi}{2}[$  dans l'équation d'évolution de (II.39). On obtient que  $\tilde{u}(\mathbf{r}, \theta) := u(\mathbf{r} \cos \theta, \mathbf{r} \sin \theta)$  est solution de l'équation :

$$\mathcal{A}(\theta)\partial_{\theta}\tilde{u} = \tilde{f} - \mathcal{M}(\theta)\mathbf{r}\partial_{\mathbf{r}}\tilde{u} + \mathbf{B}\tilde{u}, \quad (\text{II.53})$$

où les matrices  $\mathcal{A}(\theta)$  et  $\mathcal{M}(\theta)$  sont explicitement données par

$$\mathcal{A}(\theta) := \cos \theta A_1 - \sin \theta A_2 \text{ et } \mathcal{M}(\theta) := \cos \theta A_1 + \sin \theta A_2,$$

et où  $\mathbf{B}$  est un opérateur explicite ne faisant intervenir que les dérivées tangentielles  $\partial_j u$ ,  $j \in \{t\} \cup [3, d]$ .

Ainsi, si l'on suppose que la matrice  $\mathcal{A}(\theta)$  est inversible, ou de façon équivalente si on fait le renforcement suivant sur l'hypothèse de Sarason sur la matrice  $A_1^{-1}A_2$  :

**Hypothèse II.4.3** *On suppose que le problème (II.39) est non caractéristique pour chaque bord, alors les valeurs propres réelles de la matrice  $A_1^{-1}A_2$  (si elles existent) sont négatives.*

On inverse alors  $\mathcal{A}(\theta)$  pour récupérer, grâce à (II.53), le contrôle de la dérivée angulaire  $\partial_{\theta}\tilde{u}$ . Cette équation vient alors en un certain sens, jouer le rôle de l'équation explicite (II.51) dans le cas non caractéristique.

Pour conclure, il suffit de remarquer que le contrôle des dérivées radiale  $\mathbf{r}\partial_{\mathbf{r}}$  et angulaire  $\partial_{\theta}$  suffit à contrôler les dérivées  $x_1 \partial_1$  et  $x_2 \partial_2$ . En effet, on a par définition :

$$\begin{cases} x_1 \partial_1 + x_2 \partial_2 = \mathbf{r}\partial_{\mathbf{r}}, \\ -x_2 \partial_1 + x_1 \partial_2 = \partial_{\theta}, \end{cases} \Leftrightarrow \begin{cases} \mathbf{r}^2 \partial_1 = x_1 \mathbf{r}\partial_{\mathbf{r}} - x_2 \partial_{\theta}, \\ \mathbf{r}^2 \partial_2 = x_2 \mathbf{r}\partial_{\mathbf{r}} + x_1 \partial_{\theta}, \end{cases} \Leftrightarrow \begin{cases} x_1 \partial_1 = \cos^2 \theta \mathbf{r}\partial_{\mathbf{r}} - \sin \theta \cos \theta \partial_{\theta}, \\ x_2 \partial_2 = \sin^2 \theta \mathbf{r}\partial_{\mathbf{r}} + \sin \theta \cos \theta \partial_{\theta}. \end{cases}$$

Cette méthode est ensuite généralisée aux ordres supérieurs.

<sup>35</sup>Cette idée d'utiliser les dérivées radiales et angulaires d'ordre élevé, afin de contrôler la norme de Sobolev dans les variables tangentielles cartésienne est peut-être classique, bien que nous ne l'ayons pas rencontrée dans la littérature.

Afin de décrire précisément le résultat de persistance de la régularité, que l'on a obtenu il nous faut définir quelques espaces fonctionnels. Pour  $s, p \in \mathbb{N}$  et  $\gamma > 0$ , on introduit l'espace de type Sobolev à poids anisotropiques  $\mathbf{H}_\gamma^{s,p}(\Omega)$  par

$$\mathbf{H}_\gamma^{s,p}(\Omega) := \left\{ u \in \mathcal{D}'(\Omega) \mid \text{pour } \alpha, \beta \in \mathbb{N}, \delta \in \mathbb{N}^{1+d-2}, \text{ t.q. } \alpha + \beta \leq p \text{ et } \alpha + \beta + |\delta| \leq s \right. \\ \left. \text{on ait } \forall k \in \llbracket 0, \alpha + \beta \rrbracket, x_1^k x_2^{\alpha+\beta-k} \partial_1^\alpha \partial_2^\beta \partial_{t,x'}^\delta u \in L_\gamma^2(\Omega) \right\}$$

où l'on rappelle que  $L_\gamma^2(\Omega)$  désigne l'espace  $L^2(\Omega)$  à poids  $e^{-\gamma t}$ . Cet espace est normé par

$$\| \cdot \|_{\mathbf{H}_\gamma^{s,p}(\Omega)}^2 := \sum_{\alpha+\beta \leq p, \alpha+\beta+|\delta| \leq s} \sum_{k=0}^{\alpha+\beta} \| x_1^k x_2^{\alpha+\beta-k} \partial_1^\alpha \partial_2^{\alpha+\beta-k} \partial_{t,x'}^\delta \cdot \|_{L_\gamma^2(\Omega)}^2.$$

On retiendra que dans cet espace  $p$  encode la régularité maximale par rapport aux dérivées tangentielles  $x_1 \partial_1$  et  $x_2 \partial_2$ , tandis que  $s$  décrit la régularité totale de la fonction.

On peut alors définir l'espace fonctionnel dans lequel on a établi une bonne théorie de la persistance de la régularité. Plus précisément, on considère l'espace  $\mathcal{H}_\gamma^s(\Omega)$  défini de la façon suivante : pour  $s \in \mathbb{N}$ ,

$$\mathcal{H}_\gamma^s(\Omega) := \left\{ u \in \mathcal{D}'(\Omega) \mid \forall \ell \in \llbracket 0, s \rrbracket, \mathbf{r}^\ell u \in \mathbf{H}_\gamma^{s,\ell}(\Omega) \right\}.$$

On le norme par

$$\| \cdot \|_{\mathcal{H}_\gamma^s(\Omega)}^2 := \sum_{\ell=0}^s \| \mathbf{r}^\ell \cdot \|_{\mathbf{H}_\gamma^{s,\ell}(\Omega)}^2.$$

On définit de même les espaces  $\mathbf{H}_\gamma^{s,p}$  et  $\mathcal{H}_\gamma^s$  pour les faces de la frontière,  $\partial\Omega_1$  et  $\partial\Omega_2$ .

Dans ce qui suit, on suppose que le problème vient avec une bonne théorie bien posée dans  $L^2$ . Le but est alors d'étendre le résultat à des espaces incluant une notion de régularité.

**Hypothèse II.4.4** *On considère le problème aux limites pur associé à (II.39), à savoir :*

$$\begin{cases} L(\partial)u = f & \text{pour } (t, x_1, x_2, x') \in \mathbb{R} \times \mathbb{R}_+^2 \times \mathbb{R}^{d-2} := \Omega, \\ B_1 u|_{x_1=0} = g_1 & \text{pour } (t, x_2, x') \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^{d-2} := \partial\Omega_1, \\ B_2 u|_{x_2=0} = g_2 & \text{pour } (t, x_1, x') \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^{d-2} := \partial\Omega_2, \end{cases} \quad (\text{II.54})$$

pour lequel on suppose le caractère fortement bien posé suivant. Soit  $\gamma > 0$ , alors pour tout  $f \in L_\gamma^2(\Omega)$ , pour tout  $(g_1, g_2) \in L_\gamma^2(\partial\Omega_1) \times L_\gamma^2(\partial\Omega_2)$ , le problème aux limites (pur) (II.54) admet une unique solution  $u \in L^2(\Omega)$ , à traces  $(u|_{x_1=0}, u|_{x_2=0}) \in L_\gamma^2(\partial\Omega_1) \times L_\gamma^2(\partial\Omega_2)$ , et vérifiant l'estimation d'énergie suivante. Il existe  $C > 0$  et  $\gamma_0 \geq 1$ , telles que pour tout  $\gamma \geq \gamma_0$ ,

$$\begin{aligned} \gamma \|u\|_{L_\gamma^2(\Omega)}^2 + \|u|_{x_1=0}\|_{L_\gamma^2(\partial\Omega_1)}^2 + \|u|_{x_2=0}\|_{L_\gamma^2(\partial\Omega_2)}^2 \\ \leq C \left( \frac{1}{\gamma} \|f\|_{L_\gamma^2(\Omega)}^2 + \|g_1\|_{L_\gamma^2(\partial\Omega_1)}^2 + \|g_2\|_{L_\gamma^2(\partial\Omega_2)}^2 \right), \end{aligned} \quad (\text{II.55})$$

où l'on insiste<sup>36</sup> sur le fait que  $C$  ne dépend pas de  $\gamma$ .

Le principal résultat de persistance de la régularité de [Benoit, 2024] est alors le suivant :

<sup>36</sup>La restriction de seuil dans l'estimation d'énergie (II.55) est cosmétique, son but est de simplifier autant que possible les facteurs  $\gamma^{-k}$  qui apparaissent dans les estimations d'énergie.

**Théorème II.4.1 ([Benoit, 2024]-Theorem 2.1)** *Sous les Hypothèses II.4.4 et II.4.1, soit  $s \in \mathbb{N}$ , il existe  $\gamma_0 := \gamma_0(s) \geq 1$  et  $C := C(s) > 0$ , telles que si  $\gamma > \gamma_0$ ,  $u \in \mathcal{H}_\gamma^s(\Omega)$ ,  $L(\partial)u \in \mathcal{H}_\gamma^s(\Omega)$  et  $(B_1u|_{x_1=0}, B_2u|_{x_2=0}) \in \mathcal{H}_\gamma^s(\partial\Omega_1) \times \mathcal{H}_\gamma^s(\partial\Omega_2)$ , alors on a l'estimation d'énergie a priori :*

$$\begin{aligned} \|u\|_{\mathcal{H}_\gamma^s(\Omega)}^2 + \|u|_{x_1=0}\|_{\mathcal{H}_\gamma^s(\partial\Omega_1)}^2 + \|u|_{x_2=0}\|_{\mathcal{H}_\gamma^s(\partial\Omega_2)}^2 \\ \leq C \left( \|L(\partial)u\|_{\mathcal{H}_\gamma^s(\Omega)}^2 + \|B_1u|_{x_1=0}\|_{\mathcal{H}_\gamma^s(\partial\Omega_1)}^2 + \|B_2u|_{x_2=0}\|_{\mathcal{H}_\gamma^s(\partial\Omega_2)}^2 \right). \end{aligned} \quad (\text{II.56})$$

Réciproquement, soit  $\gamma > 0$  assez grand, et un triplet de sources  $(f, g_1, g_2) \in \mathcal{H}_\gamma^s(\Omega) \times \mathcal{H}_\gamma^s(\partial\Omega_1) \times \mathcal{H}_\gamma^s(\partial\Omega_2)$ . On se place toujours sous les Hypothèses II.4.4 et II.4.1, et on suppose de surcroît que le problème aux limites dual de (II.54) vérifie l'estimation d'énergie a priori (II.56). Alors la solution  $u$  du problème aux limites (II.54) vérifie  $u \in \mathcal{H}_\gamma^s(\Omega)$ , et de plus on a l'estimation d'énergie : il existe  $\underline{\gamma} \geq 1$  et  $C > 0$  tels que

$$\forall \gamma \geq \underline{\gamma}, \|u\|_{\mathcal{H}_\gamma^s(\Omega)}^2 \leq C \left( \|f\|_{\mathcal{H}_\gamma^s(\Omega)}^2 + \|g_1\|_{\mathcal{H}_\gamma^s(\partial\Omega_1)}^2 + \|g_2\|_{\mathcal{H}_\gamma^s(\partial\Omega_2)}^2 \right).$$

On donne ici quelques éléments de preuve du Théorème II.4.1. La démonstration de l'estimation d'énergie a priori vient de la généralisation à tous les ordres des contrôles des dérivées mixtes (radiale et angulaire)  $(\mathbf{r}\partial_{\mathbf{r}})^k \partial_\theta^{s-k} u$ . Calculs que l'on a décrits dans le cas  $s = 1$  précédemment. La deuxième assertion du théorème est démontrée via une généralisation à la géométrie du quart d'espace de la méthode donnant une solution régulière sans passer par la régularisation (méthode de Lax-Philips régulière) de [Benoit, c]).

Concluons ce paragraphe en donnant quelques remarques à propos du Théorème II.4.1 :

- L'Hypothèse II.4.4, assurant le caractère fortement bien posé  $L^2$ , n'est pas anodine. Il en est de même pour celle d'estimation d'énergie a priori duale dans la réciproque du Théorème II.4.1. En effet, comme on l'a déjà mentionné une classification des problèmes fortement bien posés dans  $L^2$  n'est pas encore disponible, à l'heure actuelle. Toutefois, dans le cas (particulier) des opérateurs symétriques avec conditions de bord strictement dissipatives ce caractère fortement bien posé  $L^2$  peut-être démontré (on renvoie à [[Benoit, 2015]-Chapitre 4] ou bien à [[Benoit, 2024]] pour une extension aux coefficients variables). Il est alors assez aisé (en utilisant le fait que le problème dual se comporte bien pour les opérateurs symétriques à conditions de bord strictement dissipatives) de montrer que l'estimation d'énergie a priori duale est aussi satisfaite. En particulier, le Théorème II.4.1 s'applique, donc aux opérateurs symétriques avec conditions de bord strictement dissipatives. On a toutefois dans l'énoncé du Théorème II.4.1 préféré un formalisme plus général qui pourra s'appliquer lorsque l'on disposera d'une caractérisation complète des problèmes fortement bien posés  $L^2$ .
- Insistons sur le fait que le Théorème II.4.1 est très satisfaisant du point de vue propagation de la régularité car la solution  $u$  hérite exactement de la même régularité que les sources du problème. Ainsi, aucune perte de dérivées n'apparait ce qui n'était pas clair de prime abord, on renvoie au paragraphe suivant pour plus de commentaires.
- Dans [Benoit, 2024] on peut aussi trouver deux résultats de persistance de la régularité s'appliquant non plus au problème aux limites pur mais au problème aux limites avec condition initiale inhomogène (II.39). Ces deux résultats sont obtenus directement par la méthode d'obtention d'une solution régulière décrite au début de ce paragraphe. Cette approche repose sur l'application du théorème de Lummer-Phillips. En particulier, ces résultats établissent directement le caractère fortement bien posé  $L^2$  du problèmes aux limites (II.39) et donnent de nouveaux exemples de problèmes bien posés dans  $L^2$ . Mais, là où l'on peut espérer étendre le Théorème II.4.1 à des opérateurs (seulement) fortement bien posés dans  $L^2$ , ces résultats ne peuvent être appliqués que dans un cadre symétrique avec des conditions de bord dissipatives. Ce n'est toutefois pas une réelle obstruction car le problème étudié est linéaire.

- De plus, il est honnête de mentionner ici que les deux résultats pour le problème aux limites mixte (II.39) sont légèrement moins satisfaisants que le Théorème II.4.1 ne l'est pour le pur problème aux limites. On entend par là qu'ils nécessitent une certaine hypothèse assurant le caractère bien préparé de la donnée initiale, ou bien ils subissent une légère perte de régularité de la solution au voisinage du coin. Cette limite des résultats est liée au fait suivant : la donnée initiale  $u_0$  a vocation à être dans l'espace de Sobolev usuel  $H^s$  pour que l'on puisse facilement assurer que les itérées  $A^k u_0$  soient dans  $H^{s-k}$ . Or, les estimations dont nous disposons se font dans des espaces à poids (voir typiquement la définition de  $\mathbf{H}_\gamma^{s,p}(\Omega)$ ). Ces derniers sont difficiles à comparer avec l'espace de Sobolev usuel. Pour pallier ce problème, ou bien on a supposé que la donnée était bien préparée, ce qui assure que les itérées vivent dans le bon espace, ou bien on a dû rogner un petit peu sur la régularité de la solution. Nous pensons toutefois que ces résultats sont perfectibles et mentionnons que même en l'état ils constituent, à notre connaissance, les premiers résultats donnant le caractère fortement bien posé, en grande régularité, d'un problème aux limites de la forme (II.39). De plus, s'il est encore besoin de les motiver, notons que ces résultats ne se contentent pas d'établir le caractère fortement bien posé  $L^2$ , mais qu'ils donnent directement l'estimation du semi-groupe (à savoir le contrôle de la norme  $L^\infty$  en temps), ceci sans avoir à adapter au cas du quart d'espace le résultat de prolongation de Rauch de [Rauch, 1972]).
- Concluons ces remarques par un commentaire à propos du poids  $\mathbf{r}$  qui apparaît dans l'estimation d'énergie du problème établie au Théorème II.4.1. Nous pensions initialement que ce poids était une sorte de reliquat lié à la méthode employée. Or, il s'avère que, concernant la régularité des problèmes elliptiques posés dans des domaines anguleux (on renvoie par exemple aux travaux de Kondrat'ev, [Kondrat'ev, 1967] et [Kondrat'ev, 1977]), la distance à la singularité vient pondérer les estimations d'énergie. Ainsi, il se peut que finalement ce poids ait totalement sa place dans les estimations.

### Ce qui reste à entreprendre

Nous trouvons que les résultats de persistance de la régularité sont très satisfaisants puisqu'ils permettent de dépasser le cadre fortement bien posé  $L^2$  et qu'ils montrent que le problème est "vraiment" fortement bien posé.

Toutefois, au delà de leur intérêt intrinsèque, l'un des points d'applicabilité fondamentale des résultats de persistance de la régularité (du moins dans les espaces de Sobolev usuels) est, par injection de Sobolev, de pouvoir récupérer un contrôle  $L^\infty$  de la solution, contrôle qui est un pré-requis, si nous voulons ensuite traiter des problèmes non linéaires.

Ici, le résultat de persistance de la régularité que l'on a réussi à obtenir ne permet pas encore de récupérer un tel contrôle de la norme  $L^\infty$ , en raison du fait que nous ne contrôlons non pas les dérivées normales  $\partial_1$  et  $\partial_2$  mais les dérivées tangentielles  $x_1 \partial_1$  et  $x_2 \partial_2$ .

Cependant, le fait que nous disposions d'un tel résultat est toutefois, un résultat très positif qui n'était pas évident en soi. En effet, si l'on considère par exemple un système de dérivées partielles elliptique posé dans un domaine à coin, alors il se peut que, pour certaines valeurs particulières de l'angle (typiquement les multiples rationnels de  $\pi$ ), la solution soit  $H^s(\Omega)$  avec  $s < 2$ , là où elle serait  $H^2$  dans l'espace entier. Ce défaut de régularité est lié au fait que le phénomène de régularisation elliptique est moins bon dans un domaine à coin que dans un domaine régulier. On renvoie aux travaux de Grisvard (voir par exemple, [Grisvard, 2011]-[Grisvard, 1991] et [Grisvard, 1982]) pour une étude détaillée de ce phénomène.

Si une telle perte de régularité avait eu lieu pour le problème dans le quart d'espace (II.39) alors on aurait été amené, dans l'optique de pouvoir étudier des problèmes non linéaires, à devoir gérer ces pertes dans la méthode de point fixe. Ceci aurait probablement pu être fait par exemple *via* l'utilisation du point fixe de Nash-Moser. Cependant, les résultats de [Benoit, 2024] semblent, au contraire, indiquer



qu'un théorème de point fixe de Picard sera suffisant (ce dont on ne peut que se féliciter).

Nous aimerions, à partir du résultat de [Benoit, 2024], obtenir un moyen de récupérer le contrôle de la norme  $L^\infty$  afin de pouvoir traiter dans un proche avenir des problèmes non linéaires. Pour ce faire une approche possible serait de continuer l'analogie avec les problèmes caractéristiques et de suivre les analyses de Guès et Secchi (voir [Guès, 1990], [Secchi, 1995] ou [Secchi, 1996]). Pour de tels problèmes, on arrive finalement à contrôler une dérivée normale de la solution "en payant" deux dérivées tangentielles. Un tel résultat serait alors suffisant pour notre étude et on pourrait espérer réutiliser une partie du cadre fonctionnel utilisé par exemple dans [Secchi, 1995].

### II.4.3 Développements d'optique géométrique pour le quart d'espace

Ce Paragraphe inclut les contributions des articles [Benoit, 2016]-[Benoit, 2017a] et [Benoit, a]. Nous verrons dans ce qui suit comment les développements d'optique géométrique se voient être bouleversés par la géométrie du coin. En effet, on verra ce qui suit, quels sont les nouveaux phénomènes qui se déroulent dans la géométrie du quart d'espace.

Une fois de plus, nous insisterons dans le second paragraphe sur ce que peuvent nous apprendre les développements d'optique géométrique sur la condition d'Osher (voir Hypothèse II.4.2) qui semble régir le caractère fortement bien posé dans le cas du quart d'espace.

Enfin, on conclura par un dernier paragraphe, qui montre que, là encore, la théorie faiblement bien posée pour les problèmes aux limites hyperboliques posés dans une géométrie dont le bord admet un coin est bien plus riche que dans le cas du demi-espace.

#### Heuristique

Dans ce paragraphe, on expose la méthode heuristique de [Sarason and Smoller, 7475], ensuite rendue rigoureuse dans [Benoit, 2016], qui est à la base de la construction d'un développement d'optique géométrique dans la géométrie du quart d'espace. Dans un tel cadre, le phénomène d'auto-interaction, déjà rencontré dans le cas de la géométrie de la bande (voir Paragraphe II.3.2), se trouvera complexifié. Un autre point de complexification sera la plus grande variété de l'ensemble des phases qui doivent être citées dans le développement d'optique géométrique. Ceci est une conséquence de la possibilité de créations de nouvelles phases par rebonds successifs sur les deux faces du quart d'espace.

Ce paragraphe n'a pas vocation à donner un cadre rigoureux et formel au procédé de génération des phases. Pour cela on renvoie à [Benoit, 2016], à [[Benoit, 2015]-Chapitre 6] ou encore à [Benoit, a]. On se permet de signaler que [[Benoit, 2015]-Chapitre 6] peut être intéressant à consulter car des exemples explicites y sont traités.

On se contentera, ici, d'insister sur les nouveaux phénomènes qui nous semblent les plus pertinents et sur leur influence sur la résolution de la cascade d'équations.

Considérons le problème fortement oscillant suivant : soit  $0 < \varepsilon \ll 1$ ,

$$\begin{cases} \partial_t u^\varepsilon + A_1 \partial_1 u^\varepsilon + A_2 \partial_2 u^\varepsilon = 0 & \text{pour } (t, x) \in \mathbb{R} \times \mathbb{R}_+^2, \\ B_1 u^\varepsilon|_{x_1=0} = g^\varepsilon & \text{pour } (t, x_2) \in \mathbb{R} \times \mathbb{R}_+, \\ B_2 u^\varepsilon|_{x_2=0} = 0 & \text{pour } (t, x_1) \in \mathbb{R} \times \mathbb{R}_+, \\ u^\varepsilon|_{t \leq 0} = 0 & \text{pour } x \in \mathbb{R}_+^2. \end{cases} \quad (\text{II.57})$$

Système dans lequel, on prend un terme source  $g^\varepsilon$ , de la forme

$$g^\varepsilon(t, x_2) := e^{\frac{i}{\varepsilon} \psi(t, x_2)} g(t, x_2),$$

où  $g$  est une amplitude régulière, plate au coin, s'annulant pour les temps négatifs et où la fonction de phase  $\psi$  est plane, donnée par  $\psi(t, x_2) := \underline{\tau}t + \underline{\xi}_2 x_2$ , avec  $(i\underline{\tau}, \underline{\xi}_2) \in i\mathbb{R} \times \mathbb{R}$ , une fréquence disons hyperbolique, pour simplifier<sup>37</sup>, relativement à la matrice résolvante  $\mathcal{A}_1 := A_1^{-1}(\sigma I + i\underline{\xi}_2 A_2)$ .

Pour décrire les différentes phases attendues dans le développement d'optique géométrique ainsi que les possibles phénomènes d'auto-interaction, l'analyse reprend en adaptant les mêmes arguments que ceux développés dans la bande (voir Paragraphe II.3.2). Par vitesse finie de propagation, on peut, comme première approximation, considérer que le problème dans le quart d'espace ne voit pas (du moins initialement) la condition de bord sur la face  $\partial\Omega_2$ . On se ramène donc, dans un premier temps, à considérer le problème dans le demi-espace (de droite)

$$\begin{cases} \partial_t u^\varepsilon + A_1 \partial_1 u^\varepsilon + A_2 \partial_2 u^\varepsilon = 0 & \text{pour } (t, x) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}, \\ B_1 u^\varepsilon|_{x_1=0} = g^\varepsilon & \text{pour } (t, x_2) \in \mathbb{R} \times \mathbb{R}, \\ u^\varepsilon|_{t \leq 0} = 0 & \text{pour } x \in \mathbb{R}_+ \times \mathbb{R}, \end{cases} \quad (\text{II.58})$$

dont l'ansatz est donné par<sup>38</sup> :

$$u^\varepsilon \sim \sum_{n \geq 0} \varepsilon^n \sum_{k \in \mathfrak{K}} e^{i \varphi_k} u_{n,k}, \quad (\text{II.59})$$

où les phases  $\varphi_k$  sont données par

$$\varphi_k(t, x) := \psi(t, x_2) + \xi_1^k x_1. \quad (\text{II.60})$$

Les  $\xi_1^k$ , pour  $k \in \mathfrak{K}$ , sont les racines dans la variable  $\xi_1$  de la relation de dispersion  $\det \mathcal{L}(\underline{\tau}, \xi_1, \underline{\xi}_2) = 0$ .

Dans le cas du demi-espace, pour résoudre les équations de transport qui apparaissent après avoir appliqué le lemme de Lax [Lax, 1957] on serait amené à décomposer l'ensemble des phases  $\mathfrak{K}$  en deux ensembles, celui des phases entrantes et celui des phases sortantes (on renvoie, par exemple, au Paragraphe II.2.3). Dans le cas du quart d'espace, cette dichotomie doit être raffinée en considérant les quatre types de phases possibles. On donne la définition suivante :

**Définition II.4.2 (Types de vitesse de groupe)** Soit  $\mathbf{v} \in \mathbb{R}^2$ , ayant vocation à désigner une vitesse de groupe. On dit que  $\mathbf{v}$  est :

- *entrante-sortante (es)*, si  $\mathbf{v}_1 > 0$  et  $\mathbf{v}_2 < 0$ ;
- *sortante-entrante (se)*, si  $\mathbf{v}_1 < 0$  et  $\mathbf{v}_2 > 0$ ;
- *sortante-sortante (ss)*, si  $\mathbf{v}_1 < 0$  et  $\mathbf{v}_2 < 0$ ;
- *entrante-entrante (ee)*, si  $\mathbf{v}_1 > 0$  et  $\mathbf{v}_2 > 0$ .

Si  $\mathbf{v}_k$  est indexée par  $k \in \mathfrak{K}$ , on dira par extension que l'indice  $k$  est l'un des quatre types précédents si la vitesse de groupe  $\mathbf{v}_k$  est du type correspondant.

Pour  $a \in \{es, se, ss, ee\}$  on notera  $\mathcal{I}_a$  l'ensemble des indices  $k$  qui sont de type  $a$ .

Après application du lemme de Lax [Lax, 1957], on s'attend donc à devoir considérer les quatre cas possibles d'équation de transport :

- Si  $k \in \mathcal{I}_{ss}$ , le transport se fait de l'intérieur du domaine vers les bords. En l'absence de terme source à l'intérieur, on transporte zéro. Ainsi la phase associée est nulle et est donc exclue de l'ansatz (II.59).

<sup>37</sup>L'analyse de [Benoit, 2016] s'applique du moment que la fréquence n'est pas *glancing*. Cette restriction a ensuite été ôtée dans [Benoit, a], résultat que l'on décrira plus avant dans le Paragraphe II.4.4.

<sup>38</sup>On rappelle ici que la fréquence au bord est supposée hyperbolique.

- Si  $k \in \mathcal{I}_{se}$ , on peut venir transporter un terme source intérieur ou bien un terme source sur la face  $\partial\Omega_2$ . En l'absence de tels termes source dans (II.57), on transporte une fois de plus zéro. Cette phase peut-être **initialement** exclue de (II.59). Toutefois, comme on va le voir, cette phase ne peut-être exclue *a priori* qu'initialement de (II.59). En effet, l'auto-interaction peut forcer à devoir la reconsidérer par la suite.
- Si  $k \in \mathcal{I}_{ee}$ , on peut transporter n'importe quel terme source. Le transport s'échappe toutefois à l'infini. En particulier, l'information transportée ne rencontrera jamais la face  $\partial\Omega_2$  et elle n'y sera jamais réfléchi. Ainsi les indices entrant-entrants sont des conditions d'arrêt dans le processus de génération des phases.
- Si enfin  $k \in \mathcal{I}_{es}$ , on peut transporter le terme source (non trivial) sur la face  $\partial\Omega_1$ . L'information transportée frappe, après un certain temps de parcours, la face  $\partial\Omega_2$ . On doit, alors, décrire quelles sont les réflexions qui apparaissent durant ce rebond.

Pour déterminer ces nouvelles phases à ajouter dans l'ansatz (II.59), on fixe l'une des phases entrante-sortante  $\varphi_k$ . On reprend, alors, essentiellement les mêmes arguments que précédemment mais, cette fois, appliqués au problème dans le demi-espace (du haut) :

$$\begin{cases} \partial_t u^\varepsilon + A_1 \partial_1 u^\varepsilon + A_2 \partial_2 u^\varepsilon = 0 & \text{pour } (t, x) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+, \\ B_2 u^\varepsilon|_{x_2=0} = \tilde{g}_k^\varepsilon & \text{pour } (t, x_1) \in \mathbb{R} \times \mathbb{R}, \\ u^\varepsilon|_{t \leq 0} = 0 & \text{pour } x \in \mathbb{R} \times \mathbb{R}_+, \end{cases} \quad (\text{II.61})$$

où le nouveau terme source sur le bord  $\partial\Omega_2$  oscille selon la phase de bord  $\tilde{\psi}(t, x_1) := \tau t + \xi_1^k x_1$ , où  $\xi_1^k$  désigne la première coordonnée de la phase entrante-sortante  $\varphi_k$  que nous avons fixée. Pour déterminer les phases réfléchies, nous sommes donc amené à considérer les phases  $\varphi_k$ , toujours planes de la forme (II.60), où les secondes composantes de la fréquence (les  $\xi_2$ ) sont maintenant données par  $\xi_2^{k'}$  pour  $k' \in \mathcal{K}'$  (dépendant de  $k$ ) décrivant les racines de la relation de dispersion

$$\det \mathcal{L}(\tau, \xi_1^k, \xi_2) = 0, \quad (\text{II.62})$$

dans la variable  $\xi_2$ .

Notons que dans la géométrie de la bande, à cette étape, nous retombons exactement sur les phases sortantes du problème dans la bande. Dans le cas du quart d'espace, puisque  $\xi_1^k$  est autorisée à varier, les racines dans la variable  $\xi_2$  de (II.62) rajoutent de nouvelles phases dans l'ansatz.

Nous devons, en particulier, prendre garde au fait que si la première relation de dispersion était "bien-préparée" au sens où elle n'admettait, grâce au caractère hyperbolique de la fréquence  $(i\tau, \xi_2)$ , que des racines  $\xi_1$  réelles, ceci n'est absolument pas le cas pour les nouvelles relations de dispersion (II.62).

Ainsi, il se peut que les nouvelles relations de dispersion admettent des racines  $\xi_2$  (purement) complexes. Ces dernières donnent lieu à des couches limites, que l'on doit prendre en compte dans l'ansatz (II.59) ce qui est fait dans [Benoit, 2016]. Ces couches limites, qualifiées d'évanescences, ne donnent pas lieu à de nouvelles réflexions dans le processus de génération des phases. Donc comme les amplitudes entrante-entrantes, elles doivent être vues comme des conditions d'arrêt dans le procédé de génération des phases.

Plus problématique, il se peut aussi que l'une des nouvelles racines soit associée à une fréquence de *glancing* ayant une vitesse de groupe dégénérée (l'une des deux composantes de la vitesse de groupe, ici, typiquement,  $\mathbf{v}_2$  s'annule). Une telle possibilité était, par hypothèse, écartée dans [Benoit, 2016]. Dans

[Benoit, a] nous avons inclus de telles phases dans le développement d'optique géométrique. L'apparition de telles fréquences de *glancing* n'induit pas de changements très notables dans le développement d'optique géométrique. On renvoie au Paragraphe II.4.4 pour plus de détails.

On collecte donc dans l'ansatz (II.59), les nouvelles phases hyperboliques issues de la résolution des relations de dispersion (II.62). Parmi celles-ci, il y a les entrante-entrantes qui partent à l'infini, mais aussi les phases entrante-sortantes qui vont frapper la face  $\partial\Omega_1$ . Elles vont y être réfléchies et donner lieu à des nouvelles réflexions. Pour décrire ces dernières, on reprend le procédé décrit pour le système (II.58), pour lequel la fréquence de bord a été modifiée de façon idoine. Le processus s'arrête lorsque, durant tous les rebonds, nous n'obtenons plus que des phases entrante-entrantes et/ou évanescentes.

Par conséquent, le nombre de phases que l'on s'attend à générer pour un problème dans le quart d'espace est bien plus important que pour les problèmes dans les demi-espaces qui le composent. Pour représenter ce procédé, il est utile de considérer la variété caractéristique,  $\mathcal{V}$ , du symbole  $\mathcal{L}$ , à savoir, son lieu d'annulation :

$$\mathcal{V} := \{(\tau, \xi_1, \xi_2) \in \mathbb{R}^3 \mid \det \mathcal{L}(\tau, \xi_1, \xi_2) = 0\},$$

et plus particulièrement, la section  $\mathcal{W} := \mathcal{V} \cap \{\tau = \underline{\tau}\}$ .

Le processus décrit précédemment revient alors à considérer les points d'intersection de  $\mathcal{W}$  avec la verticale d'équation  $\{\xi_2 = \underline{\xi}_2\}$ , puis, pour chacun des points d'intersection correspondant à une phase entrante-sortante<sup>39</sup>, à considérer les points d'intersection de  $\mathcal{W}$  avec les verticales d'équation  $\{\xi_1 = \xi_1^k\}$ , enfin à réitérer le même procédé pour chaque point d'intersection sortant-entrant...

On renvoie à [[Benoit, 2015]-Chapitre 6] pour des exemples explicites ou encore à la Figure II.2 pour une application de cet algorithme à une équation de type ondes.

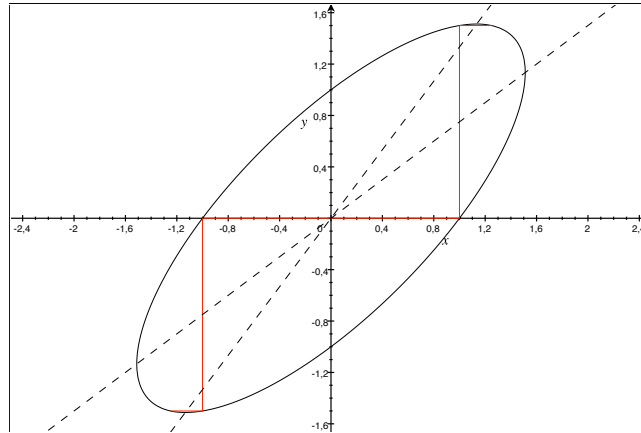


Figure II.2: Un exemple de procédé de génération des phases pour une équation de type ondes

Notons au passage que ce procédé est assez difficile à mettre en œuvre de façon effective. En effet, chaque étape nécessite alors la résolution (explicite) d'équations polynomiales de degré  $N$ . Il pourrait être intéressant de disposer d'une méthode numérique robuste qui permette de décrire toutes les phases obtenues par un tel procédé. Toutefois, pour être honnête, nous n'avons pas, à l'heure actuelle, d'idée permettant de débiter la mise en place d'un tel code, notamment en raison du fait que calculer de façon exacte un point d'intersection est délicat d'un point de vue purement numérique.

<sup>39</sup>Ces points peuvent être distingués graphiquement en considérant leur normale entrante.

Une question d'importance, notamment pour donner un sens précis à l'ansatz même, est celle de la finitude du nombre de phases obtenues. Ceci est placé comme hypothèse dans [Benoit, 2016]. Toutefois si l'on raffine quelque peu le procédé décrit précédemment, en tant que solution d'équation de transport on peut se rendre compte que l'apparition d'une nouvelle phases dans l'ensemble des phases coûte (ceci est totalement clair, si l'on considère un terme source à support loin du coin)<sup>40</sup>) un temps d'apparition strictement positif.

Ainsi, si l'on se restreint à un temps maximal de résolution  $T > 0$ , toutes les amplitudes dont les phases coûtent un temps d'apparition supérieur à  $T$  seront nulles et donc cette hypothèse sera vérifiée. On reviendra sur cette remarque dans le Paragraphe II.4.5.

Dans un cadre plus général, en particulier en temps infini, on se placera sous l'hypothèse de finitude du nombre de phases en présence.

Dans le procédé précédent la notion de phases auto-interagissantes a été gardée sous silence. En effet, dans le cas du quart d'espace, elle n'apparaît pas de façon aussi frontale que dans le cas de la bande (voir Paragraphe II.3.2). Elle est ici cachée dans des conditions bien particulières portant sur la géométrie de  $\mathcal{W}$ .

Pour s'en convaincre, considérons l'exemple suivant. On suppose que le terme source sur le bord génère une phase entrante-sortante,  $\varphi_1$ . On suppose cette dernière réfléchiée en une phase sortante-entrante,  $\varphi_2$ , contre la face  $\partial\Omega_2$ ; phase ensuite réfléchiée en une phase entrante-sortante,  $\varphi_3$ , contre la face  $\partial\Omega_1$ . Pour conclure notre illustration, supposons que la phase  $\varphi_3$ , qui voyage vers  $\partial\Omega_2$ , y soit réfléchiée en une phase sortante-entrante  $\varphi_4$ .

Alors, rien n'exclut dans l'analyse précédente que la phase  $\varphi_4$  vérifie la relation  $\xi_2 = \underline{\xi}_2$ , en d'autres termes, que cette phase fasse partie des phases sortante-entrantes que l'on a **initialement** exclues dans le procédé de génération des phases. Toutefois, dans une telle configuration, la phase  $\varphi_4$  (en tant que réflexion de la phase  $\varphi_3$ ) a maintenant parfaitement le droit de cité. Quand on étudie ses réflexions on retombe alors sur la phase initiale  $\varphi_1$ . Cette phase (et en fait les quatre phases  $\varphi_i$ ) se régénère donc elle-même par réflexions successives. On la qualifie d'auto-interagissante. On renvoie à la Figure II.3 pour une illustration.

Par conséquent, le phénomène d'auto-interaction est bien plus fin dans la géométrie du quart d'espace que dans le cas de la bande. En termes de la géométrie de  $\mathcal{W}$ , ce phénomène a lieu, si l'on est en mesure de trouver un rectangle dont les sommets sont des points de  $\mathcal{W}$  en lesquels les vitesses de groupe (ou de façon équivalente les normales) sont du bon type. On qualifiera ces quatre sommets de boucle dans la suite. Pour des exemples de systèmes hyperboliques admettant de telles boucles on renvoie par exemple à [[Benoit, 2015]-Chapitre 6]. Ces exemple tiennent toutefois plus du *toy-model*, que d'exemples ayant un réel sens physique.

L'existence de telles phases auto-interagissantes a de grandes conséquences sur la construction d'un développement d'optique géométrique. Comme on l'a déjà mentionné au Paragraphe II.2.3, construire un tel développement revient à déterminer une équation que l'on est en mesure de résoudre avant les autres, puis un ordre itératif de résolution.

Ici, puisque la première phase de l'ansatz  $\varphi_1$  est régénérée plusieurs fois, il n'est pas clair, à première vue, qu'elle soit une bonne candidate pour jouer ce rôle de phase initiatrice. De même, comme les phases auto-interagissantes, mais aussi leurs réflexions, sont régénérées périodiquement en temps, il n'est pas non plus clair qu'un ordre de résolution basé (directement) sur les réflexions soit un bon ordre<sup>41</sup>.

<sup>40</sup>On ne sait pas si se restreindre à de tels termes sources est toutefois bien moral.

<sup>41</sup>On entend par là un ordre basé sur la détermination d'une phase dont on déduit dans un second temps la valeur de ses phases réfléchies.

L'une des grandes étapes pour construire rigoureusement le développement d'optique géométrique a donc été, voir [Benoit, 2016], de trouver un ordre de résolution qui permette de résoudre la cascade d'équations d'optique géométrique. On donnera plus de détails quant à son initialisation dans le sous-paragraphe suivant.

Nous ne rentrerons ici pas spécifiquement dans les détails de la construction de [Benoit, 2016]. L'idée principale est que, sous des hypothèses de structure sur la géométrie de la section de la variété caractéristique  $\mathcal{W}$  (on demande essentiellement dans ce cadre, l'unicité de la boucle d'auto-interaction), on peut venir partitionner l'ensemble des indices de la façon suivante :

- d'une part les quatre éléments de la boucle et des éléments générés directement par rebonds des éléments de la boucle ; on parle alors des voisins de la boucle.
- D'autre part, des arbres dont les racines sont des voisins de la boucle.

On illustre cette partition sur la Figure II.3 où les  $n_i$  désignent les quatre éléments de la boucle, les  $a_i$ ,  $b_i$ ,  $c_i$  et  $d_i$  les éléments dans le voisinage direct de la boucle et enfin les  $A_{a_i}$ ,  $B_{b_i}$ ,  $C_{c_i}$  et  $D_{d_i}$  désignent les arbres issus de ces voisins.

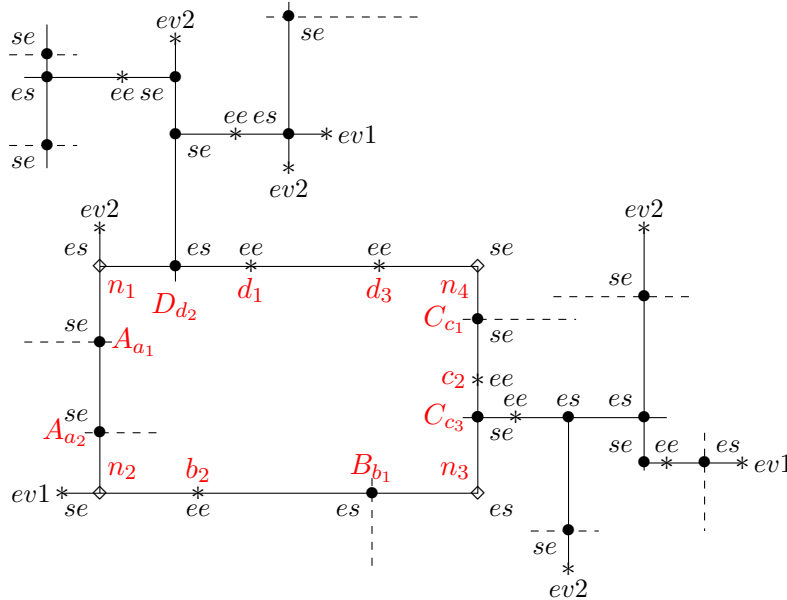


Figure II.3: Structure arborée de l'ensemble des fréquences  $\mathcal{F}$ .

On montre alors, assez facilement, que la détermination des amplitudes associées aux indices dans le voisinage direct de la boucle ne nécessite que la connaissance des éléments de la boucle. Puis, comme on a une partition de l'ensemble des indices (ce qui est fortement lié au fait que l'on se place sous l'hypothèse d'unicité de la boucle d'auto-interaction), on montre que dans chaque arbre on peut résoudre les équations de la cascade d'optique géométrique en déployant un ordre de résolution basé sur l'ordre naturel donné par celui des réflexions successives. En effet, par construction, on définit les arbres comme les paquets d'indices qui sont reliés directement les uns aux autres.

Il est honnête de mentionner ici que les hypothèses sous lesquelles on a été à même de construire le développement d'optique géométrique dans [Benoit, 2016] sont assez contraignantes. Nous verrons dans le Paragraphe II.4.4, comment on peut venir affaiblir les hypothèses sur la géométrie de la section de la variété caractéristique. Parmi les principales obstructions techniques, nous nous devons de mentionner :

- L'hypothèse d'unicité de la boucle dans  $\mathscr{W}$ . Cette obstruction est toutefois à tempérer. En effet, de façon "générique", on a de fortes raisons de penser que les problèmes dans le quart d'espace n'admettent pas de boucle d'auto-interaction. On entend par cela que bien qu'il soit toujours possible d'imaginer et de donner des exemples d'opérateurs admettant une (et même plusieurs) boucle(s) d'auto-interaction(s), ces derniers n'ont pas de sens physique et sont construits exclusivement de façon à générer plusieurs boucles.
- Le fait que la boucle considérée dans  $\mathscr{W}$  soit nécessairement une boucle d'auto-interaction. Par boucle non auto-interagissante on peut considérer la situation décrite dans la Figure II.3, mais dans laquelle l'indice  $n_2$  est maintenant entrant-entrant. Il est donc généré deux fois. L'une lors du rebond de  $n_1$  contre  $\partial\Omega_2$ , l'autre, lors du rebond de  $n_3$  contre  $\partial\Omega_1$ . Là encore, il est possible de construire de tels exemples mais ces derniers sont des *toy-models* dont le sens physique est à tempérer.
- L'hypothèse permettant d'éviter les modes de *glancing*. Il est ici honnête d'avouer que cette restriction est vraiment dure à vérifier en pratique. En effet, en raison de la difficulté que l'on a à calculer effectivement les phases obtenues dans le procédé de génération, il est vraiment compliqué de vérifier que des phases de *glancing* n'apparaissent pas durant la totalité du procédé. D'où l'idée de [Benoit, a] d'inclure ces phases pour avoir à éviter de telles considérations.

### Ce que nous apprend le développement d'optique géométrique

Ici encore, construire le développement d'optique géométrique pour la solution du problème dans le quart d'espace (II.57) a été très instructif pour la théorie globale. En effet, dans [Benoit, 2016], nous avons identifié une condition nécessaire et suffisante pour construire le développement d'optique géométrique pour le problème à coin (II.57).

L'analyse, bien qu'elle la précède historiquement, reprend ici essentiellement les mêmes techniques que celle exposées dans le cas de la bande (voir Paragraphe II.3.2). On se place sous les hypothèses de [Benoit, 2016] et l'on considère une unique boucle d'auto-interaction comprenant quatre éléments et allumée par le terme source  $g^\varepsilon$ .

On note  $s_1, s_3 \in \mathcal{I}_{es}$ , et  $s_2, s_4 \in \mathcal{I}_{se}$ , les quatre indices de la boucle. Pour déterminer les amplitudes dominantes  $u_{0,s_j}$  associées, on doit résoudre les équations de transport (on renvoie aux calculs du Paragraphe II.2.3):

$$\text{pour } j = 1, 3, \begin{cases} (\partial_t + \mathbf{v}_{s_j} \cdot \nabla_x)u_{0,s_j} = 0 & \text{dans } \Omega, \\ \text{condition de bord sur } \partial\Omega_1, \\ u_{0,s_j}|_{t \leq 0} = 0 & \text{sur } \Gamma, \end{cases}$$

et

$$\text{pour } j = 2, 4, \begin{cases} (\partial_t + \mathbf{v}_{s_j} \cdot \nabla_x)u_{0,s_j} = 0 & \text{dans } \Omega, \\ \text{condition de bord sur } \partial\Omega_2, \\ u_{0,s_j}|_{t \leq 0} = 0 & \text{sur } \Gamma, \end{cases}$$

où  $\mathbf{v}_{s_j}$  désigne la vitesse de groupe associée à  $s_j$ .

Le couplage entre ces équations se fait, justement, *via* les conditions de bord. En injectant l'ansatz

dans les conditions de bord, on est amené à résoudre :

$$\begin{cases} B_1 \begin{bmatrix} u_{0,s_1|x_1=0} + e_1^{s_1} \\ u_{0,s_2|x_2=0} + e_2^{s_2} \end{bmatrix} = g - B_1 u_{0,s_2|x_1=0}, \\ B_2 \begin{bmatrix} u_{0,s_2|x_2=0} + e_2^{s_2} \\ u_{0,s_3|x_1=0} + e_1^{s_3} \end{bmatrix} = -B_2 u_{0,s_3|x_2=0}, \\ B_1 \begin{bmatrix} u_{0,s_3|x_1=0} + e_1^{s_3} \\ u_{0,s_4|x_2=0} + e_2^{s_4} \end{bmatrix} = -B_1 u_{0,s_4|x_1=0}, \\ B_2 \begin{bmatrix} u_{0,s_4|x_2=0} + e_2^{s_4} \\ u_{0,s_1|x_1=0} + e_1^{s_1} \end{bmatrix} = -B_2 u_{0,s_1|x_2=0}, \end{cases} \quad (\text{II.63})$$

où pour  $j = 1, 3$ ,  $e_1^{s_j} \in \mathbf{E}_1^s(i\mathcal{T}, \xi_2^{s_j})$  et pour  $j = 2, 4$ ,  $e_2^{s_j} \in \mathbf{E}_2^s(i\mathcal{T}, \xi_1^{s_j})$ . Ces termes viennent du fait que l'amplitude liée à  $s_j$  n'est pas forcément la seule amplitude entrante. Toutefois, on utilise la structure de l'ensemble des indices<sup>42</sup> (voir Figure II.3), pour assurer que dans les équations précédentes, il n'y a qu'une amplitude sortante dans les membres de droite.

Appliquer la condition de Kreiss-Lopatinskii uniforme permet alors d'inverser les matrices apparaissant dans les membres de gauche de (II.63). Pour conclure il suffit de remarquer que, si l'on connaît l'une des quatre traces apparaissant dans l'un des quatre membres de droite de (II.63), alors on peut résoudre les quatre équations de transport en intégrant le long des caractéristiques. En effet, supposons, par exemple, que l'on connaisse  $u_{0,s_1|x_2=0}$ . Alors, on peut résoudre explicitement l'équation de transport

$$\begin{cases} (\partial_t + \mathbf{v}_{s_4} \cdot \nabla_x) u_{0,s_4} = 0 & \text{dans } \Omega, \\ u_{0,s_4|x_2=0} = -\Pi_2^{s_4} \phi_2^{s_4} B_2 u_{0,s_1|x_2=0} & \text{sur } \partial\Omega_2, \\ u_{0,s_4|t \leq 0} = 0 & \text{sur } \Gamma, \end{cases}$$

où  $\Pi_2^{s_4}$  est le projecteur sélectionnant la composante qui nous intéresse de  $\mathbf{E}_2^s(i\mathcal{T}, \xi_1^{s_4})$  et où  $\phi_2^{s_4}$  est l'inverse donné par la condition de Kreiss-Lopatinskii uniforme. Il suffit d'intégrer le long des caractéristiques pour obtenir

$$u_{0,s_4}(t, x) := -\Pi_2^{s_4} \phi_2^{s_4} B_2 u_{0,s_1|x_2=0} \left( t - \frac{1}{\mathbf{v}_{s_4,2}} x_2, x_1 - \frac{\mathbf{v}_{s_4,1}}{\mathbf{v}_{s_4,2}} x_2 \right).$$

Equation dont on déduit la valeur de la trace apparaissant dans le membre de droite de la troisième équation de (II.63). Réitérer la procédure, pour chaque équation de (II.63), donne alors la condition d'initialisation

$$(I - \mathbf{T})u_{0,s_1|x_2=0} = \Pi_1^{s_1} \phi_1^{s_1} g \left( t - \frac{1}{\mathbf{v}_{s_1,1}} x_1, -\frac{\mathbf{v}_{s_1,2}}{\mathbf{v}_{s_1,1}} x_1 \right), \quad (\text{II.64})$$

où l'opérateur  $\mathbf{T}$  est défini par  $(\mathbf{T}u)(t, x_1) := \mathbf{S}u(t - \alpha x_1, \beta x_1)$  avec

$$\begin{aligned} \mathbf{S} &:= \Pi_1^{s_1} \phi_1^{s_1} B_1 \Pi_2^{s_2} \phi_2^{s_2} B_2 \Pi_1^{s_3} \phi_1^{s_3} B_1 \Pi_2^{s_4} \phi_2^{s_4} B_2, \\ \alpha &:= \frac{1}{\mathbf{v}_{s_1,1}} \left[ 1 - \frac{\mathbf{v}_{s_1,2}}{\mathbf{v}_{s_2,2}} + \frac{\mathbf{v}_{s_1,2} \mathbf{v}_{s_2,1}}{\mathbf{v}_{s_2,2} \mathbf{v}_{s_3,1}} - \frac{\mathbf{v}_{s_1,2} \mathbf{v}_{s_2,1} \mathbf{v}_{s_3,2}}{\mathbf{v}_{s_2,2} \mathbf{v}_{s_3,1} \mathbf{v}_{s_4,2}} \right] > 0, \\ \beta &:= \frac{\mathbf{v}_{s_4,1} \mathbf{v}_{s_3,2} \mathbf{v}_{s_2,1} \mathbf{v}_{s_1,2}}{\mathbf{v}_{s_4,2} \mathbf{v}_{s_3,1} \mathbf{v}_{s_2,2} \mathbf{v}_{s_1,1}} > 0. \end{aligned}$$

Là encore, il est bon de comparer l'opérateur  $I - \mathbf{T}_{1 \rightarrow 1}$  donnant la valeur de la trace en  $\{x_1 = 0\}$  dans la condition d'Osher (voir Hypothèse II.4.2) et l'opérateur  $I - \mathbf{T}$  que l'on a besoin d'inverser pour initialiser la construction du développement d'optique géométrique. Ces deux opérateurs se ressemblent fortement au sens où  $\mathbf{T}$  constitue une version microlocalisée de l'opérateur  $\mathbf{T}_{1 \rightarrow 1}$  d'Osher. Ce phénomène se déroulant déjà pour les géométries du demi-espace et de la bande.

<sup>42</sup>Cette dernière implique, en effet, que le seul élément dans  $\mathbf{E}_1^s(i\mathcal{T}, \xi_2^{s_1})$  est la phase sortante-entrante  $s_2$  et que l'on a des relations analogues pour chacune des quatre conditions de bord.



Une fois que l'on suppose  $I - \mathbf{T}$  inversible (dans un espace fonctionnel idoine) la trace inconnue  $u_{0,s_1|x_2=0}$  est déterminée. Par résolution explicite d'équations de transport, on connaît alors les amplitudes auto-interagissantes.

On en déduit alors, les amplitudes des voisins de la boucle et de proche en proche, les amplitudes dans les arbres illustrés dans la Figure II.3. Cela conclut la détermination du terme principal dans le développement d'optique géométrique telle que proposée dans [Benoit, 2016], sous des hypothèses restrictives sur la variété  $\mathscr{W}$ . Plus précisément, le théorème que l'on démontre est alors le suivant :

**Théorème II.4.2** *Sous hypothèse d'hyperbolicité, de non-caractérisité des bords et sous la condition de Kreiss-Lopatinskiï uniforme sur chaque face. Sous des hypothèses de structure sur la section de la variété caractéristique  $\mathscr{W}$ , voir [Benoit, 2016] pour plus de détails, on suppose en particulier, qu'il existe une unique boucle d'auto-interaction de taille quatre. On suppose que l'opérateur  $I - \mathbf{T}$  de (II.64) est inversible dans l'espace des fonctions régulières, plates au coin. On suppose que la phase au bord  $\psi$  n'est pas de glancing, qu'elle est associée à une amplitude plate au coin,  $g$ . Enfin, on suppose que le procédé de génération des phases en induit un nombre fini et qu'aucune de ces phases n'est de glancing.*

*Alors il existe une unique solution de la cascade d'optique géométrique associée au problème (II.57),  $(u_{n,j}, U_{n,ev,k,1}, U_{n,l,ev_2})_{n \geq 0, j \in \mathcal{J}, k \in \mathcal{K}, l \in \mathcal{L}}$ , où les ensembles  $\mathcal{J}$ ,  $\mathcal{K}$  et  $\mathcal{L}$  sont explicites et regroupent respectivement les phases hyperboliques, évanescences pour la face  $\partial\Omega_1$  et les phases évanescences pour la face  $\partial\Omega_2$ . Les profils sont construits respectivement dans  $H^\infty(\Omega)$  et  $H^\infty(\Omega, \mathcal{S}(\mathbb{R}))$ .*

*De plus, si l'on suppose que le problème dans le quart d'espace (II.39) est fortement bien posé dans  $L^2$ , alors le développement d'optique géométrique est effectivement une solution approchée de la solution  $u^\varepsilon$  de (II.57). Plus précisément pour  $N_0 \in \mathbb{N}$ , si  $u_{N_0,app}^\varepsilon$  désigne le développement d'optique géométrique tronqué à l'ordre  $N_0$ , alors on a l'estimation d'erreur*

$$\|u^\varepsilon - u_{N_0,app}^\varepsilon\|_{L^2(\Omega)} \leq C\varepsilon^{N_0+1}.$$

### Un dernier exemple étonnant de problème faiblement bien posé

Ce paragraphe résume les résultats de [Benoit, 2017a]. Il y est, une fois de plus, question de théorie faiblement bien posée, mais cette fois, pour des problèmes définis dans le quart d'espace. Nous avons, en utilisant les développements d'optique géométrique, essayé de comprendre, si de nouveaux phénomènes d'instabilités pouvaient apparaître. Nous avons vu que dans le cas de la bande, ceci se produisait et qu'une instabilité faible pouvait être répétée, périodiquement au cours du temps, grâce au phénomène d'auto-interaction, donnant ainsi naissance à un nouveau type d'instabilité.

Sans surprise, un tel procédé peut aussi se produire dans la géométrie du quart d'espace. Mais, il est à noter que ce dernier peut conduire à des résultats d'instabilités bien plus forts que dans le cas de la bande.

Plaçons nous dans le cadre d'un problème aux limites admettant une condition de bord dans la classe  $WR$  de [Benzoni-Gavage et al., 2002], disons, pour fixer les esprits, celle sur  $\partial\Omega_1$ . La seconde condition est, quant à elle, choisie comme vérifiant la condition de Kreiss-Lopatinskiï uniforme, ou même mieux elle peut être choisie strictement dissipative. On s'attend, comme dans le cas de la bande, à venir répéter l'instabilité faible à chaque fois que l'on vient réitérer une réflexion contre la face  $\partial\Omega_1$ , de l'amplitude liée à la dégénérescence de la condition de Kreiss-Lopatinskiï uniforme. Le problème que l'on étudie s'écrit

$$\begin{cases} L(\partial)u^\varepsilon = 0 & \text{dans } \Omega_T, \\ B_1 u^\varepsilon|_{x_1=0} = \varepsilon^M e^{\frac{i}{\varepsilon}(\tau t + \xi_2 x_2)} g & \text{sur } \partial\Omega_{1,T}, \\ B_2 u^\varepsilon|_{x_2=0} = 0 & \text{sur } \partial\Omega_{2,T}, \\ u^\varepsilon|_{t \leq 0} = 0 & \text{sur } \Gamma, \end{cases} \quad (\text{II.65})$$

où  $(i_{\mathcal{T}}, \underline{\xi}_2) \in \Xi_0$ , est la fréquence responsable de la dégénérescence de la condition de Kreiss-Lopatinskii uniforme, où  $g$  est supposée régulière, typiquement  $\mathcal{D}(\partial\Omega_{1,T})$ , nulle pour les temps négatifs. Dans (II.65), l'entier  $M > 0$ , qui prend en compte les pertes de régularité pourra être choisi quelconque et ce indépendamment du temps final  $T > 0$ .

On postule alors un ansatz, incluant toutes les phases décrites dans le procédé de génération de phases du Paragraphe II.4.3 et toutes à l'échelle  $O(1)$ . Nous ne décrivons pas dans le détail toutes les étapes dans la construction. Le point crucial dans la résolution de la cascade d'équations d'optique géométrique est que, comme dans le cas de la bande, l'on commence par déterminer les amplitudes d'ordre zéro, autrement dit les plus irrégulières. Ceci peut paraître quelque peu contre-intuitif, mais cela ne devrait pas. En effet, la seule façon de générer une amplitude à une telle échelle est de considérer l'amplitude initiale, d'ordre  $M$ , qui a fait le nombre de tours idoïne, pour venir amplifier exactement  $M$  fois la singularité produite par la dégénérescence de la condition de Kreiss-Lopatinskii uniforme.

Un tour de boucle et donc une régénération de l'instabilité, coûte un temps  $t_* > 0$ . Ce dernier dépend explicitement des vitesses de groupes associées aux phases de la boucle, mais aussi, et c'est une nouveauté par rapport à la bande, il dépend de la distance entre le bas du support de  $g$  et le coin. Dans la géométrie du quart d'espace nous pouvons toujours venir choisir le support de  $g$  assez proche du coin pour assurer qu'avant le temps de résolution total  $T > 0$ , l'instabilité ait été régénérée exactement  $M$  fois. Là où, dans la bande, nous avons besoin d'imposer  $T$  assez grand, ici un choix pertinent du terme source assure l'accumulation de  $M$  instabilités faibles.

Par conséquent, on a montré que pour tout  $T > 0$ , il existe un choix de terme source  $g$  tel que le développement d'optique géométrique associé à (II.65) soit  $O(1)$  en  $\varepsilon$ . En termes de pertes de dérivées, ceci se traduit par le théorème suivant :

**Théorème II.4.3 ([Benoit, 2017a])** *Sous hypothèse d'hyperbolicité, de bords non caractéristique, on suppose qu'il existe une fréquence  $(i_{\mathcal{T}}, \underline{\xi}_2) \in \Xi_0$  plongeant le problème dans la classe  $WR$  et qui est simultanément une fréquence d'auto-interaction. Alors le problème dans le quart d'espace (II.39) ne peut pas être faiblement bien posé<sup>43</sup>.*

*Autrement dit, pour tout  $K, K_2 \in \mathbb{N}$ , pour tout  $f \in H^K(\Omega_T)$  et  $g_2 \in H^{K_2}(\partial\Omega_{2,T})$ , et pour tout  $K_1 \in \mathbb{N}^*$ , on peut trouver  $g_1 \in H^{K_1}(\partial\Omega_{1,T})$ , telle que l'estimation d'énergie*

$$\begin{aligned} \|u\|_{L^2(\Omega_T)}^2 + \|u_{|x_1=0}\|_{L^2(\partial\Omega_{1,T})}^2 + \|u_{|x_2=0}\|_{L^2(\partial\Omega_{2,T})}^2 \\ \leq C_T \left( \|f\|_{H^K(\Omega_T)}^2 + \|g_1\|_{H^{K_1}(\partial\Omega_{1,T})}^2 + \|g_2\|_{H^{K_2}(\partial\Omega_{2,T})}^2 \right) \end{aligned}$$

*soit violée.*

En d'autres termes, un problème dans le quart d'espace ayant une condition de bord sur l'une de ses faces dans la classe  $WR$  peut ne peut pas être (même) faiblement bien posé. Car, il développe un nombre de pertes de dérivées arbitrairement grand avant le temps  $T$ . Il est toutefois remarquable de noter qu'ici le phénomène qui génère l'instabilité forte d'Hadamard n'est pas, comme c'est le plus souvent le cas un problème de nature spectrale (on pense par exemple à la dégénérescence de la condition de Kreiss-Lopatinskii (faible)). Mais, il est vraiment constitué d'une accumulation arbitrairement grande d'instabilités faibles. A notre connaissance, il s'agit du premier exemple d'un tel type d'instabilité. Nous la trouvons particulièrement esthétique même si elle implique que l'étude du problème d'équations aux dérivées partielles sous-jacente est une impasse.

En particulier, ce résultat tend à montrer que l'interaction de deux chocs lorsque l'un des deux ne vérifie pas la condition de stabilité de Lax peut-être particulièrement pathologique.

Il est toutefois à noter qu'un comportement si pathologique n'est possible que pour des dégénérescences de la condition de Kreiss-Lopatinskii uniforme pour une fréquence hyperbolique et auto-interagissante.

<sup>43</sup>Ceci est vrai indépendamment de la nature de la condition de bord sur la seconde face.

Nous n'avons pas soumis les résultats à publication, mais lors d'une dégénérescence hyperbolique non auto-interagissante ou elliptique (voir même *glancing*), on peut montrer que l'instabilité ne peut pas être amplifiée. Par conséquent, on retrouvera la classification du demi-espace (voir Théorème II.2.4).

#### II.4.4 Vers des développements d'optique géométriques en géométrie bornée

Dans ce paragraphe, nous nous proposons de donner des éléments de démonstration qui nous paraissent marquants concernant une thématique de recherche où nous venons d'entrer, à savoir la construction rigoureuse de développements d'optique géométrique dans un domaine borné.

Ce paragraphe est donc exploratoire et encore assez parcellaire, mais nous pensons néanmoins qu'en l'état actuel, il permet déjà de poser et de dégager quelques questions intéressantes.

Avoir un algorithme de détermination d'un développement d'optique géométrique en géométrie bornée nous semble être une thématique assez ambitieuse et nouvelle (même si l'on peut retrouver ces questions dans les travaux de [Ivanovici et al., 2014] et [Ivanovici et al., 2021]). Il n'est en effet pas besoin d'insister sur l'importance physique que peut revêtir un tel domaine de résolution. Nous aimerions aussi, à terme, étudier l'influence de la géométrie du bord sur la solution *via* son développement d'optique géométrique. Nous aimerions pouvoir traiter des géométries quelconques sans avoir à se ramener à un demi-espace (ou un quart d'espace) par changement de variables.

Puisque jusqu'ici nous avons considéré des frontières à coin, nous proposons dans un premier temps de considérer les bornées à coins les plus simples, à savoir les triangles et les rectangles. Dans le premier paragraphe, nous décrirons les résultats du *preprint* [Benoit, b], qui sera bientôt soumis à publication, établissant un tel développement dans le cadre bien particulier de l'équation des ondes dans un triangle. Comme on le verra, cet exemple, pourtant simple, nous a amené à nous poser de nouvelles questions concernant les développements d'optique géométrique dans le cas du quart d'espace. On exposera la "mise à jour" de l'article [Benoit, 2016] par le *preprint* [Benoit, a] dans le second paragraphe. Enfin, dans le dernier paragraphe, illustre de façon exhaustive, du moins avec notre connaissance actuelle, les difficultés que l'on va être amené à surmonter pour construire de tels développements dans un cadre plus général.

#### Développements d'optique géométrique dans un borné : le cas du triangle

Dans le *preprint* [Benoit, b], nous donnons un exemple de construction de développement d'optique géométrique pour l'équation des ondes posée dans un triangle isocèle. Plus précisément, nous avons considéré le problème (obtenu après mise sous la forme d'un système d'ordre un de l'équation des ondes)

$$\begin{cases} \partial_t u^\varepsilon + A_1 \partial_1 u^\varepsilon + A_2 \partial_2 u^\varepsilon = 0 & \text{dans } \mathbb{R} \times \mathcal{T}, \\ B_S u^\varepsilon|_{x \in \mathcal{T}_S} = e^{\frac{i}{\varepsilon} \psi(t, x_2)} g, \\ B_E u^\varepsilon|_{x \in \mathcal{T}_E} = 0, \\ B_W u^\varepsilon|_{x \in \mathcal{T}_W} = 0, \\ u^\varepsilon|_{t \leq 0} = 0 & \text{sur } \mathcal{T}, \end{cases} \quad (\text{II.66})$$

où l'amplitude  $g \in H^\infty(]0, 1[)$  vérifie pour des questions de compatibilités  $g|_{t \leq 0} = 0$ , et pour tout  $\alpha \in \mathbb{N}$ ,  $(\partial_1 g)|_{x_2=0} = 0 = (\partial_1 g)|_{x_2=1}$ . La phase de bord  $\psi(t, x_1)$  sera décrite plus en détails dans la suite.

Dans (II.66), les coefficients  $A_1$  et  $A_2$  sont donnés par

$$A_1 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ et } A_2 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

et  $\mathcal{T}$  désigne le triangle isocèle centré

$$\mathcal{T} := \{x \in \mathbb{R}^2 \mid -1 < x_1 < 1 \text{ et } x_2 < 1 - |x_1|\}.$$

La notation  $\mathcal{T}_S$  (resp.  $\mathcal{T}_W$  ; resp.  $\mathcal{T}_E$ ) renvoie alors à la frontière sud (resp. ouest ; resp. est) du triangle. On peut alors vérifier que les matrices  $A_1$ ,  $A_1 - A_2$  et  $-A_1 + A_2$  qui correspondent aux trois coefficients normaux pour chacune des faces de la frontière sont inversibles. Elles ont aussi une unique valeur propre positive, ainsi les matrices  $B_S$ ,  $B_E$  et  $B_W$  sont trois matrices lignes quelconques. On supposera, dans la suite, qu'elles vérifient la condition de Kreiss-Lopatinskii uniforme.

Une première étape dans la construction du développement d'optique géométrique est de comprendre quelles sont les différentes phases que l'on doit inclure dans l'ansatz. Ainsi, il nous faut comprendre comment le processus de génération des phases décrit dans le Paragraphe II.4.3 se trouve être modifié.

Une remarque capitale pour la suite est d'abord que puisque nous sommes en géométrie fermée aucune phase ne peut être exclue de prime abord. En effet, après avoir défini de façon idoine les différents types de phases (entrantes pour la face  $S$ , sortantes pour la face  $S, \dots$ ), on ne peut pas, comme dans le cas du quart d'espace, négliger un type de phase particulier comme on le faisait pour les phases sortante-sortantes. De plus, aucun type de phase (on pense aux phases hyperboliques entrante-entrantes dans le quart d'espace) ne joue ici le rôle d'une condition d'arrêt puisqu'une information transportée frappe toujours l'une des faces et est donc toujours réfléchi.

Ainsi, comme on n'a plus de "condition d'arrêt" dans le procédé de génération des phases, on a donc *a priori* deux situations qui peuvent se produire : ou bien, on génère un nombre infini de phases, ou bien, "on a de la chance" et dans le procédé, on revient à certaines étapes à des phases déjà considérées et on boucle. Le nombre de phases peut alors être fini, mais on est en présence d'une auto-interaction ou plus.

Concernant le procédé de génération des phases en tant que tel, il est inchangé par rapport à celui décrit dans le quart d'espace. Si ce n'est, et comme on va le voir on pense que cela sera d'importance à l'avenir, que l'on considère maintenant les intersections de la section de la variété caractéristique avec des droites dont le vecteur directeur est le vecteur normal à la face du bord où l'on se trouve. Vu de façon plus analytique, on peut aussi utiliser le fait que les réflexions sont explicitement calculables. En effet, dans le cas bien particulier de l'équation des ondes, on a les relations de Snell-Descartes et donc les réflexions se calculent explicitement. On peut montrer la proposition suivante :

**Proposition II.4.1** *Soit  $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$ , une notation pour une vitesse de groupe. On considère  $y$  le rayon de direction  $v$  alors (en cas de réflexion)*

- *le rayon réfléchi de  $y$  contre la face  $\mathcal{T}_W$  a pour direction  $(v_2, v_1)$ ;*
- *le rayon réfléchi de  $y$  contre la face  $\mathcal{T}_E$  a pour direction  $-(v_2, v_1)$ ;*
- *le rayon réfléchi de  $y$  contre la face  $\mathcal{T}_S$  a pour direction  $(v_1, -v_2)$ .*

Cette proposition donne alors un moyen de calculer toutes les phases que l'on va rencontrer dans le développement d'optique géométrique. Dans ce formalisme, il est plus commode de travailler avec les vitesses de groupe.

On fixe donc la phase au bord  $\psi$  de manière à ce que cette dernière soit associée à deux vitesses de groupe  $(v_1, v_2)$  et  $(-v_1, v_2)$  et sans perte de généralité, on impose  $v_1 < 0$ ,  $v_2 > 0$  et  $-v_1 < v_2$  (pour éviter le cas *glancing*  $-v_1 = v_2$ ). On obtient alors huit phases suivantes que l'on caractérise par leur vitesse de groupe. Ces dernières ainsi que leur ordre d'apparition sont illustrés dans la Figure II.4 dans laquelle la notation  $\begin{bmatrix} x \\ y \end{bmatrix}^A$  désigne le devenir d'une vitesse de groupe  $\begin{bmatrix} x \\ y \end{bmatrix}$  contre la face  $A$ .

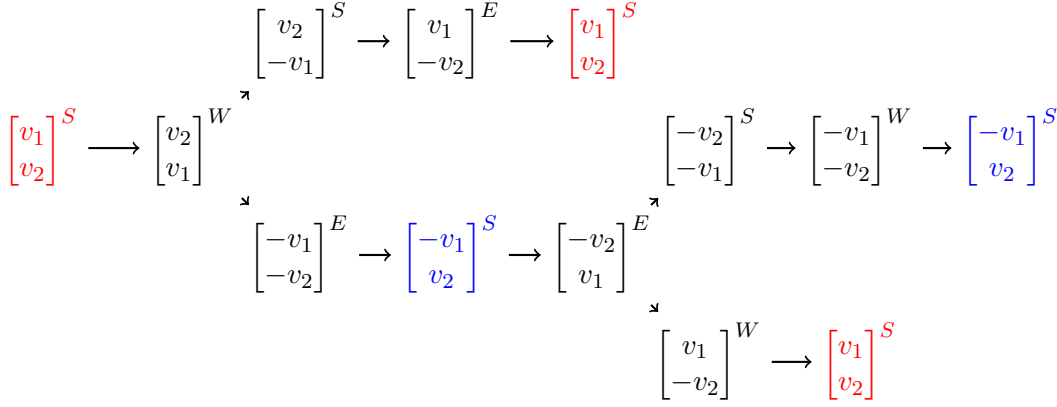


Figure II.4: La famille des vitesses de groupe g en er es dans le triangle

On remarque en particulier que ce proc ed e de g en eration des phases comprend plusieurs boucles et que ces derni eres n'ont pas forc ement quatre  el ements, comme celles pr ec edemment consid er es dans le cas du quart d'espace. Cette observation a en partie motiv e les r esultats du paragraphe suivant. Il est aussi  a noter que l'on obtient exactement le m eme ensemble de phases si on suit l'approche g eom etrique consistant  a faire des intersections successives de la section de la vari et e caract eristique avec des droites dont les directions sont les normales aux faces de  $\mathcal{T}$ .

On postule donc un ansatz  a huit  el ements :

$$\begin{aligned}
u^\varepsilon(t, x) &\sim \sum_{n \geq 0} \varepsilon^n \left( e^{\frac{i}{\varepsilon} \varphi_{+, \pm}^1} u_{n, +, \pm}^1 + e^{\frac{i}{\varepsilon} \varphi_{-, \pm}^1} u_{n, -, \pm}^1 + e^{\frac{i}{\varepsilon} \varphi_{+, \pm}^2} u_{n, +, \pm}^2 + e^{\frac{i}{\varepsilon} \varphi_{-, \pm}^2} u_{n, -, \pm}^2 \right). \quad (\text{II.67}) \\
&:= \sum_{n \geq 0} \sum_{k \in \mathcal{K}} \varepsilon^n \left( e^{\frac{i}{\varepsilon} \varphi_k^1} u_{n, k}^1 + e^{\frac{i}{\varepsilon} \varphi_k^2} u_{n, k}^2 \right),
\end{aligned}$$

dans lequel l'indexage par  $\pm$  se fait sur le signe de la composante de la vitesse de groupe associ ee. On doit alors r esoudre des  equations de transport avec des conditions sur une ou deux des faces de  $\mathcal{T}$ , ce nombre de conditions de bord  tant fonction de la nature de la vitesse de groupe  $\mathbf{v}_{\pm, \pm}$  consid eree. Par exemple, l'amplitude  $u_{0, -, -}$  est associ ee  a la vitesse de groupe  $(v_1, -v_2)$  qui est sortante pour les faces  $\mathcal{T}_S$  et  $\mathcal{T}_W$ . Sa seule condition de bord est donc sur la face  $\mathcal{T}_E$ .

Une fois de plus, les  equations internes sont d ecoupl ees. Ce n'est pas ici que se situe la difficult e. Le couplage  a lieu au sein des conditions de bord. En injectant l'ansatz dans les conditions de bord sur la face sud, on obtient les conditions : pour tout  $n \geq 0$ ,

$$\begin{aligned}
B_S \left[ e^{\frac{i}{\varepsilon} \varphi_{-, +, S}^1} u_{n, -, +}^1 + e^{\frac{i}{\varepsilon} \varphi_{-, +, S}^2} u_{n, -, +}^2 + e^{\frac{i}{\varepsilon} \varphi_{+, +, S}^1} u_{n, +, +}^1 + e^{\frac{i}{\varepsilon} \varphi_{+, +, S}^2} u_{n, +, +}^2 \right]_{|_S} \quad (\text{II.68}) \\
= \delta_{n, 0} g - B_S \left[ e^{\frac{i}{\varepsilon} \varphi_{-, -, S}^1} u_{n, -, -}^1 + e^{\frac{i}{\varepsilon} \varphi_{-, -, S}^2} u_{n, -, -}^2 + e^{\frac{i}{\varepsilon} \varphi_{+, -, S}^1} u_{n, +, -}^1 + e^{\frac{i}{\varepsilon} \varphi_{+, -, S}^2} u_{n, +, -}^2 \right]_{|_S},
\end{aligned}$$

o u les  $\varphi_{\cdot, \cdot, S}$  d esignent les traces des fonctions de phases  $\varphi_{\cdot, \cdot}$  sur  $\{x_2 = 0\}$ . Il s'agit de fonctions des variables  $(t, x_1)$  qui peuvent  tre rendues pr ecises. Toutefois comme ces derni eres induisent juste des facteurs de phases, elles peuvent  tre n eglig ees dans ce qui suit.

En proc edant de la m eme fa on pour les autres conditions de bord on obtient les deux conditions

$$\begin{aligned}
B_W \left[ e^{\frac{i}{\varepsilon} \varphi_{+, +, W}^2} u_{n, +, +}^2 + e^{\frac{i}{\varepsilon} \varphi_{-, -, W}^2} u_{n, -, -}^2 + e^{\frac{i}{\varepsilon} \varphi_{+, -, W}^1} u_{n, +, -}^1 + e^{\frac{i}{\varepsilon} \varphi_{+, -, W}^2} u_{n, +, -}^2 \right]_W = \quad (\text{II.69}) \\
- B_W \left[ e^{\frac{i}{\varepsilon} \varphi_{-, +, W}^1} u_{n, -, +}^1 + e^{\frac{i}{\varepsilon} \varphi_{-, +, W}^2} u_{n, -, +}^2 + e^{\frac{i}{\varepsilon} \varphi_{+, +, W}^1} u_{n, +, +}^1 + e^{\frac{i}{\varepsilon} \varphi_{-, -, W}^1} u_{n, -, -}^1 \right]_W,
\end{aligned}$$

et

$$\begin{aligned}
B_E \left[ e^{\frac{i}{\varepsilon}\varphi^{1,-,+},E} u_{n,-,+}^1 + e^{\frac{i}{\varepsilon}\varphi^{1,-,-},E} u_{n,-,-}^1 + e^{\frac{i}{\varepsilon}\varphi^{2,-,-},E} u_{n,-,-}^2 + e^{\frac{i}{\varepsilon}\varphi^{2,+,-},E} u_{n,+,-}^2 \right]_E = \\
- B_E \left[ e^{\frac{i}{\varepsilon}\varphi^{2,-,+},E} u_{n,-,+}^2 + e^{\frac{i}{\varepsilon}\varphi^{1,+,+},E} u_{n,+,+}^1 + e^{\frac{i}{\varepsilon}\varphi^{2,+,+},E} u_{n,+,+}^2 + e^{\frac{i}{\varepsilon}\varphi^{1,+,-},E} u_{n,+,-}^1 \right]_E.
\end{aligned} \tag{II.70}$$

On a écrit les conditions de bord précédentes de façon à ce que les membres de gauche de (II.68)-(II.69) et (II.70) regroupent les éléments du sous-espace stable, pour lesquels on peut inverser la matrice de bord grâce à la condition de Kreiss-Lopatinskii uniforme.

On voit donc que comparé au problème dans le quart d'espace, bien que l'on ait une structure des équations assez similaire à celle des conditions de bord dans (II.63), le phénomène d'auto-interaction et donc de couplage des équations commence à être nettement plus riche.

Dans ce qui suit, on ne décrira pas en détails la suite de l'analyse. Pour être honnête cette dernière est assez lourde d'un point de vue calculatoire, bien que nous soyons seulement sur un exemple explicite...

Le point remarquable de l'étude est que pour initialiser la construction du développement d'optique géométrique et déterminer les amplitudes principales  $u_{0,\pm,\pm}^j$ , on a besoin de supposer six conditions d'inversibilité. Elles sont toutes de la forme  $I - \mathbf{T}_j$ ,  $j \in \llbracket 1, 6 \rrbracket$ . Là, où l'analyse devient particulièrement complexe (et disons le assez désagréable), est que le membre de droite auquel on veut appliquer l'inverse de l'opérateur  $I - \mathbf{T}_j$  dépend des inverses des opérateurs précédents.

Contentons nous d'expliquer pourquoi il est clair que six conditions sont néanmoins à l'œuvre. On suppose que les quatre termes (rouges) dans le membre de droite de (II.68) sont connus. On peut alors intégrer le long des caractéristiques. Comme les équations de transport pour  $u_{-,+}^2$  et  $u_{+,+}^1$  nécessitent seulement la condition de bord en  $\mathcal{T}_S$ , on peut les résoudre. Ceci donne la valeur de ces deux fonctions sur les faces  $\mathcal{T}_E$  et  $\mathcal{T}_W$ . On a représenté ces termes en bleu dans les équations (II.69) et (II.70). Si l'on suppose que les deux termes restants (en violet) dans le membre de droite de (II.69) sont connus alors le membre de droite de cette condition de bord est totalement déterminé. Intégrant le long des caractéristiques, on peut maintenant déterminer  $u_{+,+}^2$  et  $u_{+,-}^1$ . On obtient, en particulier, leurs traces sur  $\mathcal{T}_S$  et sur  $\mathcal{T}_E$  (en orange), ce qui a deux conséquences :

- Le membre de droite de la condition de bord (II.70) est maintenant connu. Il dépend des six termes que l'on a supposé connus.
- La valeur de la trace de  $u_{0,+,-}^1$  sur  $\mathcal{T}_S$ , que l'on a supposé initialement connue, vérifie une équation de la forme :

$$(I - \mathbf{T}_1)u_{0,+,-|_S}^1 = F,$$

où  $F$  dépend de la source  $g$  et des cinq autres traces que l'on a supposé connues. L'opérateur  $\mathbf{T}_1$  dépend explicitement des vitesses de groupe et des inverses donnés par la condition de Kreiss-Lopatinskii uniforme. On obtient alors l'une de nos conditions d'inversibilité.

Résoudre les équations de transport les unes après les autres, nous donne les cinq conditions restantes. Il reste alors à étudier de façon précise chacune de ces conditions afin qu'elle assure l'existence d'une trace dans un espace fonctionnel idoine. Ceci sera facilité si l'on se place dans le cas d'un problème en temps fini pour lequel on doit pouvoir inverser les opérateurs  $I - \mathbf{T}_j$  en prenant les développements en série de Neumann.

Ainsi, on a vu, sur cet exemple qui bien que totalement explicite s'avère déjà assez désagréable à manipuler au vu des expressions en présence, que le phénomène d'auto-interaction, bien que proche de celui du quart d'espace, était beaucoup plus riche. Il y a plusieurs phénomènes d'auto-interaction qui coexistent les uns avec les autres. Ceci a motivé le retour sur la construction de développements d'optique

géométrique dans le quart d'espace, mais pour lesquels plusieurs phénomènes d'auto-interaction peuvent coexister que l'on va décrire dans le paragraphe suivant.

Notons en conclusion que nous avons été assez étonné par la présence de six conditions d'inversibilité pour le problème dans le triangle. En effet, nous avons débuté cette recherche par l'équation des ondes dans un rectangle (on n'a alors que quatre phases générées) et on obtient<sup>44</sup> cinq conditions d'inversibilité. Ceci semble faire sens, une condition pour chaque coin et une condition "globale" qui assure que ces quatre conditions locales interagissent "bien" les unes avec les autres.

Toutefois, ces six conditions dans le triangle ne peuvent pas être expliquées par le pressentiment précédent. Comprendre exactement pourquoi six conditions apparaissent dans le triangle est, bien sûr, un pré-requis avant de quitter le cadre rassurant de l'équation des ondes.

Si nous devons donner un pronostic, il se pourrait que le nombre de conditions ne soit pas vraiment lié au nombre de coins dans la frontière, mais, plutôt au nombre de boucles dans la section de la variété caractéristique. En effet, la Figure II.4 met en évidence trois boucles :

$$\begin{aligned}\mathcal{B}_1 &:= ((\mathbf{v}_1, \mathbf{v}_2); (\mathbf{v}_2, \mathbf{v}_1); (\mathbf{v}_2, -\mathbf{v}_1); (\mathbf{v}_1, -\mathbf{v}_2)), \\ \mathcal{B}_2 &:= ((-\mathbf{v}_1, \mathbf{v}_2); (-\mathbf{v}_2, \mathbf{v}_1); (-\mathbf{v}_2, -\mathbf{v}_1); (-\mathbf{v}_1, -\mathbf{v}_2)), \\ \mathcal{B}_3 &:= ((\mathbf{v}_1, \mathbf{v}_2); (\mathbf{v}_2, \mathbf{v}_1); (-\mathbf{v}_1, -\mathbf{v}_2); (-\mathbf{v}_1, \mathbf{v}_2); (-\mathbf{v}_2, -\mathbf{v}_1); (\mathbf{v}_1, -\mathbf{v}_2)).\end{aligned}$$

Boucles que l'on peut venir combiner les unes aux autres pour obtenir un total de six boucles. On a ainsi six boucles, ce qui serait cohérent avec les six conditions d'inversibilité. Toutefois, obtenir une telle explication pour les cinq conditions qui apparaissent dans la géométrie du rectangle nous échappe pour le moment.

## Retour sur les développements dans le cas du quart d'espace

Ce paragraphe expose succinctement les résultats du *preprint* [Benoit, a]. Comme on l'a vu sur l'exemple du paragraphe précédent, la construction rigoureuse de développements d'optique géométrique dans le cas d'un domaine borné va nécessiter de prendre en compte des phénomènes d'auto-interaction complexes.

Pour débiter une telle étude nous avons choisi de commencer par la mener à bien dans la géométrie du quart d'espace dans laquelle nous avons plus d'intuition. Ainsi les résultats de l'article [Benoit, 2016] ont été étendus dans [Benoit, a] pour prendre en compte les phénomènes suivants :

1. Comme on ne peut pas empêcher qu'une phase de *glancing* apparaisse à une certaine étape du procédé de génération des phases, on a décidé d'inclure ces phases dans le développement d'optique géométrique. L'essentiel de l'analyse se fait en suivant les travaux de Williams [Williams, 1996] et [Williams, 2000] qui traitent ces phases en ajoutant une nouvelle échelle dans le développement, *i.e.* une échelle en  $\sqrt{\varepsilon}$ . On montre alors que les amplitudes des modes de *glancing* sont régies par des équations de type Schrödinger. Comme les modes évanescents, les modes de *glancing* sont responsables de l'apparition de couches limites (à l'échelle  $\sqrt{\varepsilon}$  et non plus  $\varepsilon$ ). Notre principale contribution est ici d'avoir compris et justifié que ces couches limites, exponentiellement décroissantes dans la variable normale au bord auquel elles sont *glancing*, ne donnaient pas lieu à de nouvelles réflexions contre l'autre face du bord. Ainsi, comme les phases entrante-entrantes et les phases évanescents, les phases de *glancing* constituent des conditions d'arrêt dans le procédé de génération des phases. Par conséquent, l'ensemble des phases générées ne se trouve pas modifié outre mesure.

---

<sup>44</sup>La géométrie du rectangle est légèrement plus simple, moins calculatoire, mais aussi (nous le pensons) moins riche en nouveaux phénomènes et questions d'étude que celle du triangle. Le but n'étant pas ici (et ne sera pas) de décrire un bestiaire de toutes les géométries possibles. C'est pourquoi, nous avons préféré ici ne pas décrire la géométrie du rectangle.

2. C'est une extension directe de [Benoit, 2016] mais dans [Benoit, a], nous avons permis aux boucles d'avoir plus de quatre éléments. En effet, comme on peut le voir dans la Figure II.4 de telles boucles peuvent apparaître en géométrie bornée.
3. Enfin, et c'est peut-être la majeure contribution de [Benoit, a] par rapport à [Benoit, 2016], nous avons étendu la construction du développement d'optique géométrique à des opérateurs hyperboliques qui présentent plusieurs boucles dans la section de leur variété caractéristique, ces dernières boucles pouvant être auto-interagissantes ou non. Ceci dans le but d'avoir la plus grande généralité possible pour les phénomènes d'auto-interaction en présence.

Ainsi comparé à [Benoit, 2016], le résultat de [Benoit, a] a un plus grand domaine d'applicabilité et il permet maintenant de considérer des phénomènes qui apparaîtront en géométrie bornée.

A proprement parler, dans le quart d'espace, les généralisations données par les points 2. et 3. peuvent sembler un peu cosmétiques car, bien qu'il soit toujours possible de créer de toutes pièces des opérateurs générant plusieurs boucles, la "majorité" des opérateurs et du moins ceux ayant un sens physique ne développent pas de tels comportements. On insiste donc ici sur le fait que cette étude a été avant tout motivée comme étant un travail préparatoire à la compréhension des techniques que l'on sera (probablement) amené à développer pour traiter le cas d'une géométrie bornée.

De plus, les points qui sous-tendent la construction dans une configuration donnée par 2. et 3. sont, nous le pensons, intéressants en eux-mêmes et méritent quelques commentaires. En effet dans [Benoit, 2016], sous l'hypothèse d'unicité de la boucle d'auto-interaction, nous avons vu que l'ensemble des phases avait une structure arborée (voir Figure II.3). Cette dernière nous donnait alors un ordre naturel de résolution des équations de la cascade. La structure arborée de la Figure II.3 n'est plus vérifiée lorsque l'on a plusieurs boucles car alors certaines branches des arbres peuvent venir se mêler les unes aux autres. On renvoie à la Figure II.5 pour un exemple d'une telle intersection. L'ordre de résolution des équations n'est alors plus si clair que dans [Benoit, 2016]. En effet, pour sa résolution, une amplitude peut nécessiter la connaissance d'éléments qui la "précèdent" dans plusieurs arbres.

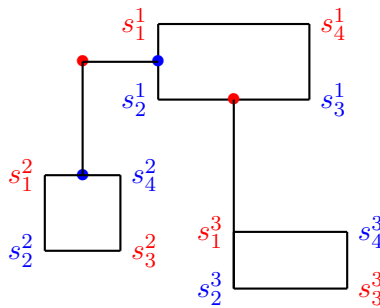


Figure II.5: Un exemple avec plusieurs boucles

Toutefois on peut, sans trop de difficultés, montrer qu'un ordre de résolution basé sur le nombre de façons de générer une phase considérée à partir de la racine  $s_1^1$ , ordre qui raffine donc l'ordre de [Benoit, 2016], est alors un bon choix. Ici  $s_1^1$  désigne l'indice de la phase auto-interagissante allumée par le terme source au bord. On commence par déterminer toutes les amplitudes générées de façon unique à partir de  $s_1^1$ .

Pour ces dernières, l'analyse de [Benoit, 2016] s'applique. Puis, on détermine celles qui sont générées d'exactly deux façons distinctes à partir de  $s_1^1$ . Plus précisément, on suppose que l'on peut dans le



procédé de génération des phases obtenir l'amplitude  $\underline{i}$  considérée à partir de deux séquences distinctes de phases  $\ell$  et  $\ell'$ . Alors on montre assez facilement que les éléments de  $\ell$  et  $\ell'$  sont nécessairement obtenus à partir de  $s_1^1$  de façon unique, faute de quoi il y aurait au moins trois séquences donnant  $\underline{i}$ . Donc, l'analyse précédente s'applique aux différents éléments de  $\ell$  et  $\ell'$ . En particulier, on connaît les éléments de leurs queues, nommément  $\ell_p$  et  $\ell'_p$ , qui permettent alors de déterminer l'amplitude  $u_{0,\underline{i}}$  recherchée. On conclut alors par itération.

Il est bon ici de préciser que l'ordre précédemment décrit ne peut pas non plus être mis en place si l'on ne fait aucune hypothèse de structure sur les boucles. En effet, dans [Benoit, a] nous avons introduit la notion de boucle simple qui empêche l'existence d'indices d'auto-interaction qui se trouvent appartenir à deux boucles simultanément. Nous avons de fortes raisons de penser que cette hypothèse n'est pas nécessaire, mais elle a l'avantage de simplifier nettement les démonstrations. Les Figures II.6 et II.7 illustrent les différences entre ces deux types d'auto-interaction simple et combinée.

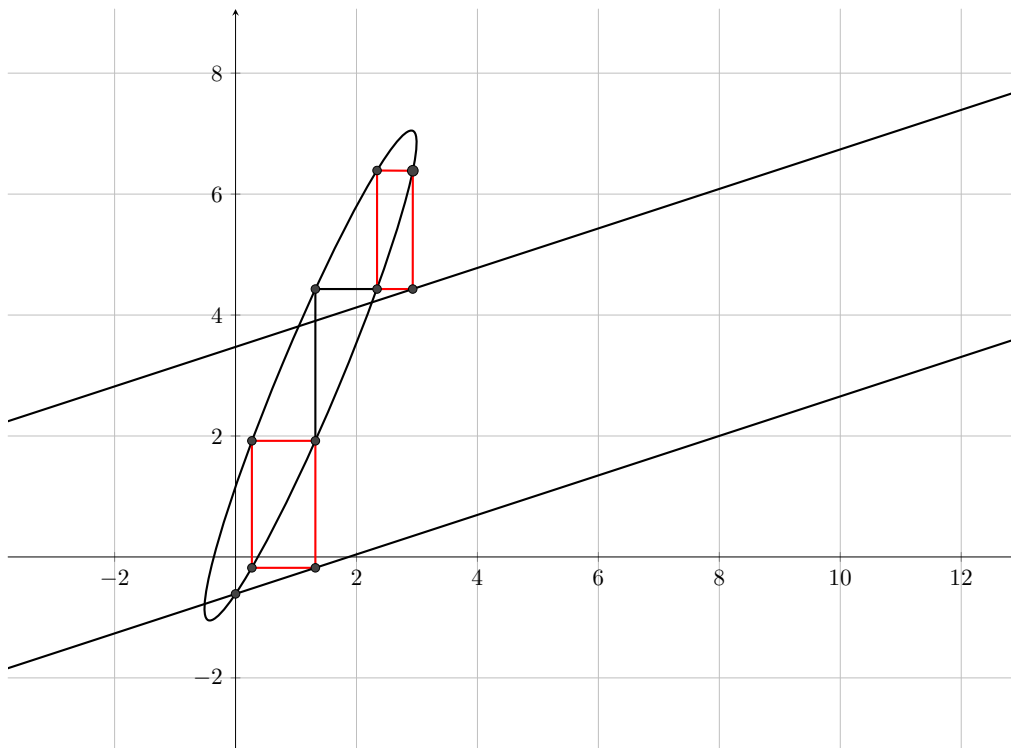


Figure II.6: Une variété caractéristique avec deux boucles simples.

Il y a fort à parier qu'un tel ordre raffiné de résolution de la cascade d'équations doive être mis en œuvre pour traiter des problèmes dans un borné. En particulier, on peut aussi remarquer que l'exemple de l'équation des ondes dans le triangle décrit au Paragraphe II.4.4 admet une boucle d'auto-interaction combinée car l'indice associé à la vitesse de groupe  $(-v_1, v_2)$  appartient aux boucles  $\mathcal{B}_2$  et  $\mathcal{B}_3$ .

Ainsi, la méthode de [Benoit, a], bien que raffinée, ne semble pas encore tout à fait suffisante pour traiter des problèmes quelconques dans un borné<sup>45</sup>.

Notons enfin, pour conclure quant à l'applicabilité effective du théorème donnant l'existence du développement d'optique géométrique, qu'une hypothèse sur laquelle nous n'avons malheureusement pas

<sup>45</sup>Nous avons bon espoir de pouvoir encore peaufiner notre ordre de résolution pour aboutir à un tel cadre. Nous ne l'avons pas fait dans [Benoit, a] car ce premier raffinement nous semblait déjà assez technique comme cela.

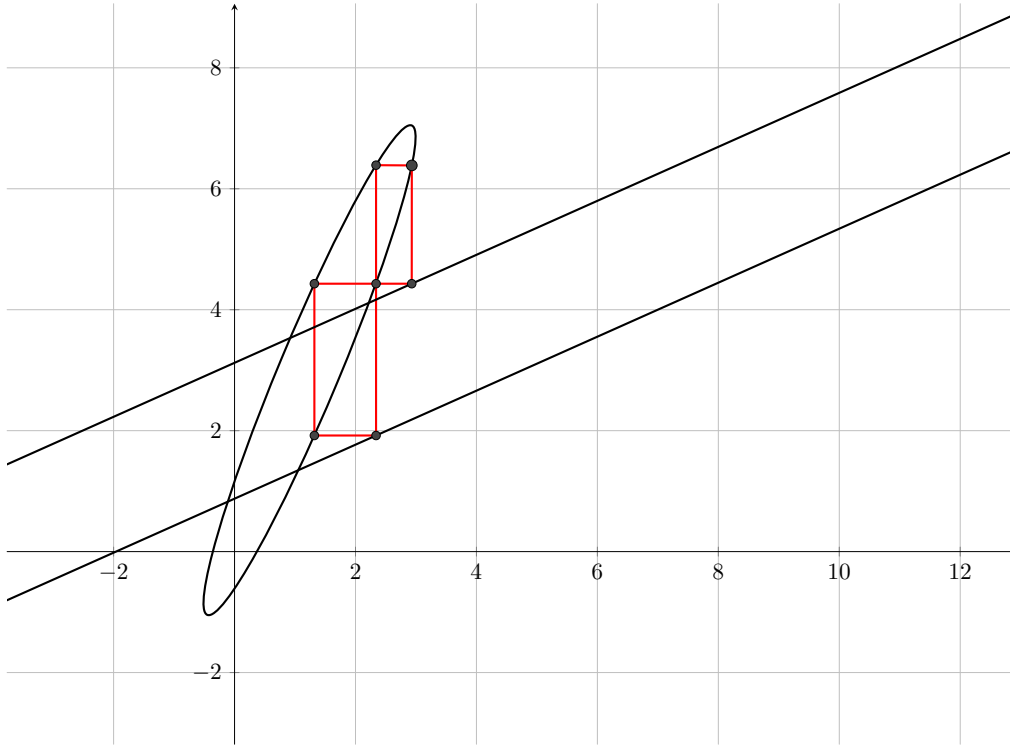


Figure II.7: Une variété caractéristique avec une auto-interaction combinée.

réussi à rogner, à l'heure actuelle, est celle assurant que le nombre de phases générées par le procédé décrit au Paragraphe II.4.3 soit fini. Cette hypothèse n'est toutefois pas utilisée pour construire les différentes amplitudes dans le développement d'optique géométrique, qui peuvent être manipulées comme le terme général de séries formelles, mais "seulement" pour donner un sens à l'ansatz dans sa partie justification. Cette hypothèse permet alors de donner un sens à l'ansatz en tant que série finie.

Il se peut que cette hypothèse ne soit pourtant pas nécessaire. En effet dans [[Benoit, 2015]-Section 6.6], nous avons considéré un exemple de développement d'optique géométrique avec une infinité de phases et sur cet exemple explicite, nous avons montré que pour (presque toutes) les conditions de bord strictement dissipatives, on pouvait faire converger les séries définissant cet ansatz. En cas d'un nombre infini de phases dans le développement, les amplitudes associées vont s'écrire (en tant que solution d'équations de transport) sous la forme

$$u_{0,k} := \mathbf{R}^k g(\mathbf{t}^k(t, x_1, x_2), \mathbf{x}_2^k(x_1, x_2)),$$

où  $\mathbf{R}^k$  est une matrice qui encode les différents coefficients de réflexion que l'on a rencontrés pour obtenir la phase  $k$  et où  $\mathbf{t}^k$  et  $\mathbf{x}_2^k$  sont des fonctions affines dépendant des vitesses de groupe que l'on a rencontrées. Ainsi, pour faire converger une série (en  $k$ ) ayant un terme général de la forme précédente, il faut, peu ou prou, que les coefficients de réflexion  $\mathbf{R}^k$  dissipent suffisamment l'énergie. Il ne nous semble toutefois pas totalement immédiat que dans le cas général on puisse espérer caractériser facilement<sup>46</sup> les conditions de bord permettant la convergence de la série. Mais, ceci est une piste intéressante pour pouvoir se passer de l'hypothèse de finitude de l'ensemble des phases.

On conclut ce paragraphe sur une conséquence du point 1. qui, nous le pensons, pourrait être d'une certaine importance. Cette dernière est que les opérateurs (on a maintenant plusieurs boucles) que l'on

<sup>46</sup>Nous aimerions typiquement que cela soit vérifié pour les conditions de bord strictement dissipatives

doit inverser pour initialiser la résolution de la cascade d'équations ne font pas intervenir les différents modes de *glancing* que l'on peut être amené à devoir considérer. Ceci était déjà le cas dans la géométrie de la bande, pour laquelle dans [Benoit, 2020b] la condition portait aussi uniquement sur les modes hyperboliques et où l'on donnait toutefois un correcteur suffisant pour les modes de *glancing* assurant un résultat de convergence comme dans [Williams, 1996].

Si l'on pense que la condition nécessaire (et suffisante) pour construire le développement d'optique géométrique est une version microlocalisée de la condition d'Osher [Osher, 1973], qui semble donner le caractère fortement bien posé, alors cette remarque semble indiquer que la bonne condition pour le caractère fortement bien posé serait en fait une version raffinée de celle d'Osher, étendue aux fréquences de  $\Xi_0$  mais dans laquelle on ne fait pas intervenir les modes de *glancing*.

Ce raffinement pourrait, en particulier, avoir son intérêt dans la démonstration du fait que les opérateurs  $\mathbf{T}_{1 \rightarrow 2}$  et  $\mathbf{T}_{2 \rightarrow 1}$  sont uniformément bornés, point qui, on le rappelle, conclurait la caractérisation des problèmes fortement bien posés. En effet, dans les avancées sur cette question apportées dans notre thèse [Benoit, 2015], nous arrivions à obtenir des résultats (certes ne suffisant pas à conclure) pour les modes elliptiques et les modes hyperboliques. Cependant, nous ne disposions pas d'un début de démonstration pour les modes de *glancing*. Il serait donc bon que ces derniers puissent être, pour ainsi dire, ôtés de l'analyse.

Toutefois, une formulation d'une version prolongée de l'hypothèse d'Osher ne prenant pas en compte les modes de *glancing*, mais permettant toujours de construire un *symétriseur* est toujours une question largement ouverte.

## Les principales difficultés à surmonter et les potentielles extensions

Dans ce paragraphe nous revenons sur l'idée de construire les développements d'optique géométrique dans une géométrie bornée en mettant en lumière quelques (possibles) difficultés que nous pensons (peut-être) rencontrer. Ce paragraphe est toutefois assez hypothétique et nous pensons que les réponses aux questions posées ci-dessous feront probablement l'objet de travaux assez conséquents.

Comme on l'a déjà mentionné dans le Paragraphe II.4.4, le fait que dans une géométrie bornée nous n'ayons pas de phases induisant un arrêt du procédé de génération nous fait nous tenir face à *Charybde ou Scylla*. Ou bien, on a une infinité de phases générées, ou bien on a des phases auto-interagissantes. Honnêtement, nous ne savons pas trop lequel de ces deux maux serait le moindre car comme indiqué précédemment dans le cas d'une infinité de phases donner un sens même à l'ansatz n'est pas clair de prime abord. Tandis que dans le cas d'une auto-interaction, c'est l'initialisation même de la résolution de la cascade d'optique géométrique qui va être très technique, comme on l'a vu dans le Paragraphe II.4.4.

Les différents travaux que nous avons menés jusqu'ici sur l'auto-interaction nous invitent à être peut-être plus optimiste dans la seconde configuration. Toutefois, une question que nous pensons importante de se poser est la suivante : "L'auto-interaction est-elle vraiment si fréquente que cela et ce même en géométrie bornée ?".

En effet, nous avons vu que cette dernière était automatique dans la bande, "rare" (car assez contraignante) dans le cas du quart d'espace et qu'elle semblait (du moins intuitivement) être aussi automatique dans le cas d'une géométrie bornée : en effet, qu'il s'agisse de l'exemple décrit dans le triangle décrit au Paragraphe II.4.4 ou bien que l'on considère l'équation des ondes dans un rectangle, on voit apparaître ces phénomènes d'auto-interaction.

Il se peut toutefois que nos exemples précédents soient un peu biaisés. Il se pourrait même que l'auto-interaction que l'on a indiquée dans les deux exemples précédents en géométrie bornée ne soit pas systémique. Le point commun de nos deux exemples est que les normales aux différentes faces du bord ont des coefficients on ne peut plus rationnels. On pense que cela influe sur la possibilité même d'avoir

une boucle dans la section de la variété caractéristique. En effet, si l'on considère toujours l'équation des ondes, dont la variété caractéristique est le cercle, posée dans un domaine angulaire différent de  $\frac{\pi}{2}$ , alors il est clair que cette dernière ne peut pas admettre de boucle ayant seulement quatre éléments. Il semble aussi que des boucles avec plus d'éléments peinent à exister lorsque l'angle est "assez réel".

On a donc l'impression floue que si les coefficients des normales sont des nombres "suffisamment réels" le phénomène d'auto-interaction va avoir beaucoup plus de mal à se dérouler. Il se peut que cette intuition soit infirmée à l'avenir et elle nécessite *a minima* d'être rendue rigoureuse en utilisant des notions de théorie des nombres ou bien encore de géométrie algébrique. Nous pensons toutefois que cela constitue une question intéressante que nous aimerions étudier dans un avenir à long terme. Un premier point à étudier serait de déterminer si dans un secteur angulaire de taille  $\theta$ , l'auto-interaction est effectivement possible pour toutes les valeurs de  $\theta$ , ou bien si elle est cantonnée, par exemple à des multiples rationnels de  $\pi$ .

Une autre thématique que l'on aimerait étudier à longue échéance serait l'établissement de tels développements d'optique géométrique dans un borné et d'étudier l'influence de la géométrie du bord sur ce développement. Ceci de façon directe, sans passer par des arguments de partition de l'unité et ou de redressement du bord. Cette méthode directe aurait alors l'avantage de pouvoir construire une solution approchée de l'équation facilement codable sur machine.

#### II.4.5 Quelques questions ouvertes

Ce paragraphe regroupe quelques perspectives de recherche que nous pensons intéressantes et que nous espérons avoir l'opportunité d'approfondir dans le futur. On a décomposé ces perspectives en trois groupes principaux, les premières, celles liées à l'analyse d'Osher, sont décrites dans le premier paragraphe, les secondes traitent de possibles extensions à des systèmes physiques non linéaires. Elles sont détaillées dans le second paragraphe. Enfin, les perspectives du dernier paragraphe II.4.5 sont liées à l'approximation visqueuse de problèmes hyperboliques posés dans le quart d'espace et des questions qui y sont liées.

#### Faire de l'analyse d'Osher une théorie aboutie

Comme on l'a déjà mentionné, l'apport de l'analyse d'Osher [Osher, 1973] dans l'étude des problèmes aux limites posés dans le quart d'espace est important, bien qu'à l'heure actuelle cette dernière ne soit pas totalement suffisante pour conclure à une caractérisation totale des conditions de bord donnant lieu à un problème fortement bien posé. A l'avenir, nous aimerions améliorer cette méthode de construction du *symétriseur* de façon à la rendre totalement opérante et conclure ainsi la théorie fortement bien posée précédemment exposée. Le fait que cette approche ait été appliquée avec succès dans le cas de la bande dans [Benoit, 2020a], nous autorise à un certain optimisme quant à la faisabilité de ce projet, même si comme on dit *a posse ad esse non valet consequentia*<sup>47</sup>.

Pour cela, le premier point à démontrer est le caractère uniformément borné des opérateurs de traces  $\mathbf{T}_{1 \rightarrow 2}(\zeta)$  et  $\mathbf{T}_{2 \rightarrow 1}(\zeta)$ . Ce point est toutefois délicat car il s'avère que ces opérateurs sont des objets mathématiques dont la nature n'est pas si claire, ils ressemblent le plus à des opérateurs intégraux de Fourier (à phase matricielle et complexe). Mais, à notre connaissance il n'est pas ici possible de s'appuyer directement sur une analyse existante dans la littérature.

Toutefois on a, à l'heure actuelle, plusieurs grandes pistes pour établir ces bornes :

- Nous pourrions reprendre, proprement et très calmement, la démonstration d'Osher [Osher, 1973] et voir si les bornes avec pertes de dérivées qu'Osher arrive à démontrer ne peuvent pas, en fait, être améliorées pour obtenir des bornes uniformes. Il est bon d'indiquer ici que l'article [Osher, 1973]

<sup>47</sup> Ce n'est pas de la possibilité d'une chose que l'on doit déduire son existence.

n'est pas ce que l'on pourrait qualifier d'article facile et plaisant à lire (il contient en effet beaucoup de ce que l'on pourrait qualifier *a minima* d'imprécisions) c'est d'autant plus vrai pour la dernière partie qui étudie ces bornes. Ainsi une telle piste nécessitant une relecture profonde de l'article [Osher, 1973] demanderait un peu de temps et son issue est incertaine.

- Une seconde idée, plus exploratoire, serait d'essayer comme dans les travaux de Le Rousseau [Le Rousseau, 2006] et [Le Rousseau and Hörmann, 2006], concernant la construction de *paramétrix*, de décomposer les opérateurs de traces comme une composition (probablement infinie) d'opérateurs plus simples. On pourrait, alors espérer démontrer le caractère uniformément borné de ces opérateurs et essayer de passer à la limite.

Sans entrer dans trop de détails techniques, l'idée centrale de Le Rousseau pour obtenir une *paramétrix* du problème de Cauchy :

$$\begin{cases} \partial_t u + a(t, x, D_x)u = 0 & \text{pour } 0 \leq t \leq T \text{ et } x \in \mathbb{R}^d, \\ u|_{t=0} = u_0, \end{cases} \quad (\text{II.71})$$

où  $a$  est un opérateur pseudodifférentiel est de considérer une subdivision de l'intervalle de résolution  $[0, T]$ ,  $(t_j)_j$  et de venir approcher  $u$  la solution de (II.71) itérativement en posant :

$$u_{j+1}(t, x) \simeq \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{i(x-y)\cdot\xi - (t-t_j)a(t_j, y, \xi)} u_j(t_j, y) dy d\xi,$$

où l'on a posé comme première itération

$$u_1(t, x) \simeq \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{i(x-y)\cdot\xi - (t-t_j)a(0, y, \xi)} u_0(y) dy d\xi.$$

On pourrait peut-être espérer adapter cette approche pour essayer d'obtenir des formulations approchées de nos opérateurs de trace sous forme de composées en procédant à une subdivision de l'intervalle  $[0, \infty[$ . Essayer d'adapter la méthode de Le Rousseau au problème dans le demi-espace, ou encore dans la bande serait un excellent travail préparatoire.

- A terme, il faudrait aussi voir quel est le comportement des opérateurs de trace et plus spécifiquement de la condition d'Osher pour les fréquences au bord  $\Xi_0$ . Si l'on fait l'analogie avec la condition de Kreiss-Lopatinskii uniforme, sa version posée dans la demi-sphère ouverte, n'est pas utilisable. En pratique, on lui préfère sa version compactifiée. Ainsi, il y a fort à parier qu'une version sur la demi-sphère compactifiée de l'hypothèse d'Osher, non contente de préciser les indications fournies par les développements d'optique géométrique, en donnerait une version, probablement, plus vérifiable. Cette étude pourrait aussi fournir le caractère uniformément borné des opérateurs de trace. Toutefois, le comportement précis de ces opérateurs au voisinage des fréquences de *glancing* devra alors être déterminé.

Même si nous avons des pistes à suivre, notre but semble être un problème délicat, assez ambitieux et est probablement un problème à longue échéance. Toutefois, nous aimerions, un jour établir ce résultat. Ceci nous permettrait de boucler notre sujet de doctorat.

Les autres points sur lesquels nous aimerions travailler sont directement en relation avec l'analyse d'Osher mais ne sont plus liés à l'aboutissement de cette méthode<sup>48</sup>. Ce sont plutôt des améliorations que nous aimerions apporter.

<sup>48</sup>A l'heure du bouclage de ce manuscrit, indiquons succinctement que nous pensons intéressant de voir si la construction du *symétriseur* d'Osher ne pourrait pas être simplifiée après transformée de Mellin [Mellin, 1897]. En effet, dans le système (II.46), l'opérateur radial pourrait alors être remplacé par un opérateur de multiplication

- Le premier point que nous pensons intéressant d'améliorer est que contrairement à ce qui a été déclaré jusqu'ici, le *symétriseur* d'Osher n'est pas un vrai *symétriseur*. Comme on l'a mentionné pour le problème dans la bande, nous avons employé cette terminologie pour faire écho à celle du *symétriseur* de Kreiss. Le *symétriseur* d'Osher n'est pas un *symétriseur* car pour opérer il nécessite la symétrie des coefficients. En effet, la construction nécessite une première réduction à un terme source à l'intérieur nul. Ceci est réalisable si le système est symétrique (posé comme hypothèse) avec des conditions de bord strictement dissipatives (dont on sait qu'elles existent toujours). Bien qu'en pratique, supposer la symétrie des coefficients  $A_j$  n'est pas choquant, car c'est une hypothèse vérifiée par la majeure partie des systèmes ayant un sens physique, nous aimerions, dans un souci esthétique, pouvoir nous passer de cette hypothèse. La principale différence serait alors que dans l'équation différentielle définissant le *symétriseur* à l'intérieur (voir iii) de II.4.1), on aurait à résoudre une équation inhomogène. Ceci doit pouvoir se faire assez aisément en utilisant la formule de Duhamel. Cependant, l'étude rigoureuse des contributions des nouveaux termes intégraux reste à entreprendre.
- Dans son état actuel, la construction du *symétriseur* d'Osher donnée dans [Osher, 1973] ne passe pas aux coefficients variables. Dans la bande nous avons quelque peu modéré l'intérêt d'une telle généralisation (voir Paragraphe II.3.4), cette généralisation prendrait toutefois toute sa place dans le cas du quart d'espace. En effet, avoir une théorie fortement bien posée permettant des coefficients variables est un premier pas dans l'étude de problèmes quasi-linéaires. Dans la même veine, nous pourrions aussi vouloir étendre la construction du *symétriseur* d'Osher à des problèmes dans le quart d'espace caractéristiques. Toutefois, tant que nous n'avons pas parfaitement compris le rôle (pourtant central) que semble jouer la matrice de Sarason  $A_1^{-1}A_2$  cette perspective semble être une mauvaise piste à suivre, car trop hasardeuse.

### Extension aux problèmes non linéaires

Comme indiqué dans le Paragraphe II.4.2, la principale motivation qui nous a poussé à nous intéresser à la persistance de la régularité pour la solution d'un problème hyperbolique posé dans le quart d'espace est les possibles utilisations de ce résultat pour des extensions de la théorie à des problèmes non linéaires. A notre connaissance, peu importe les estimations non linéaires utilisées (voir par exemple [[Rauch, 2012]-Chapter 6], [[Métivier, 2004]-Section 2.5]), l'un des "must have" pour étudier de tels problèmes est à un moment une borne  $H^s$  de la solution du problème. Borne qui est le plus ensuite complétée en une borne  $L^\infty$  obtenue par injection de Sobolev. Une théorie de la régularité de la solution en grande norme Sobolev est donc requise.

Par conséquent, établir, à partir de la théorie de persistance de la régularité décrite dans le Paragraphe II.4.2, une théorie permettant le contrôle de la norme  $L^\infty$  est une perspective de recherche que l'on étudiera dans un futur proche.

Les techniques utilisées pour passer du linéaire au non linéaire dépendront alors de l'espace fonctionnel dans lequel nous aurons notre théorie de persistance de la régularité. Comme déjà indiqué, notre problème a de nombreuses ressemblances avec le problème dans le demi-espace caractéristique. Une possibilité serait donc de récupérer le principe des deux dérivées tangentielles pour contrôler une dérivée normale comme c'est le cas dans les travaux de Secchi [Secchi, 1995]-[Secchi, 1996]. Nous aimerions ensuite pouvoir nous baser sur les techniques qui permettent alors de passer au non linéaire.

Pour être totalement exhaustif sur ce point, on se doit de mentionner que sous l'Hypothèse de Sarason II.4.1, si l'on suppose que les matrices  $A_1$  et  $A_2$  commutent, alors on sait démontrer, à moindre coût, une théorie de persistance de la régularité en norme  $H^s$ . Cela vient du fait que nous avons ici à résoudre des équations de transport entrante-sortantes ou sortante-entrantes avec des conditions de bords bien

déterminées<sup>4950</sup>.

Ceci ne constitue toutefois pas la seule extension à des problèmes non linéaires que nous avons à l'esprit en ce qui concerne le quart d'espace. Nous aimerions aussi établir, de façon rigoureuse, des développements d'optique géométrique pour des problèmes non linéaires posés dans le quart d'espace. Cette piste peut sembler assez ambitieuse, elle sera très technique, mais nous pensons que le plan de recherche suivant pourrait, si ce n'est la rendre totalement réalisable, du moins permettre d'obtenir quelques résultats intéressants.

Le phénomène que l'on pense intéressant à comprendre dans cette étude est comment l'auto-interaction des phases induite par la géométrie du coin détaillée au Paragraphe II.4.3 se trouve être influencée par la non-linéarité et, en particulier, par le phénomène de génération des harmoniques. On aimerait comprendre si de nouveaux phénomènes pourraient en résulter, ou si cela ne conduit qu'à une complexification technique de la construction des développements d'optique géométrique.

Pour éviter d'attaquer ce problème de façon trop frontale, un premier travail consisterait à étudier, comme mentionné dans le Paragraphe II.3.3, l'analogie de ce problème dans le cas de la bande. En effet, dans une telle géométrie, on conserve le phénomène d'auto-interaction que l'on veut étudier mais avec un processus de génération des phases bien plus simple que dans le cas du quart d'espace. Comprendre les interactions non linéaires de ce nombre restreint de phases semble donc être un premier travail préparatoire.

Concernant la non linéarité en tant que telle, nous pensons qu'il serait bon pour débiter cette entreprise, de considérer des non-linéarités semi-linéaires. On pourra déplorer le fait que de telles non-linéarités ont un sens moins physique que les non-linéarités quasi-linéaires. Cependant, en ce qui concerne le procédé de génération de phases, comme la non-linéarité ne voit pas les coefficients de l'opérateur hyperbolique, ce dernier devrait être inchangé par rapport au cas linéaire sous-jacent<sup>51</sup>. Tandis que dans le cadre quasi-linéaire, il faudrait *a minima* comprendre le procédé de génération des phases lorsque les coefficients  $A_j$  dépendent des variables de position  $(t, x)$ . Ensuite, il faudra aussi considérer le procédé de génération des phases où les coefficients sont de la forme  $A_j(t, x, \underline{u})$ ,  $\underline{u}$  étant l'état de linéarisation. Processus qui dépend alors fortement de l'état  $\underline{u}$ .

Ensuite, l'un des autres intérêts de considérer des non-linéarités semi-linéaires est que, grâce aux travaux de Williams [Williams, 1996] et [Williams, 2000] dans le demi-espace, le cadre fonctionnel à utiliser, celui des algèbres de Wiener, est bien compris. Pour simplifier l'analyse, on devrait aussi dans un premier temps faire l'hypothèse simplificatrice que la fréquence au bord n'est pas dans la zone de *glancing*. Là aussi, on voit l'avantage de se placer dans un premier temps dans la bande, où cette hypothèse est tout à fait licite tandis que même avec les résultats de [Benoit, a], cette dernière est plus aléatoire dans la géométrie du quart d'espace.

---

<sup>49</sup>A l'heure du bouclage de ce manuscrit, indiquons que nous pensons pouvoir venir nous passer de l'hypothèse de commutativité dans la méthode précédente du moins en ce qui concerne la construction de solutions fortes.

L'idée est alors de se baser sur la décomposition de l'espace  $\mathbb{R}^N = \mathbf{E}_1^s \oplus \mathbf{E}_2^s$ , supposée comme hypothèse mais souvent vérifiée par les systèmes physiques, où  $\mathbf{E}_j^s$  désigne le sous-espace propre associé aux valeurs propres négatives de  $A_j$ . Décomposition qui joue un rôle un peu analogue à celle utilisée dans le cas caractéristique.

Le défaut de commutativité induit alors un couplage interne entre les différents transports. En traitant ce couplage comme une source et en employant une méthode itérative, on résout alors avec une perte d'une dérivée à chaque étape. L'infinité de pertes nécessaire pour faire converger la méthode est alors compensée par la régularité infinie des termes sources que l'on est en droit d'imposer pour une solution forte.

<sup>50</sup>Encore à l'heure du bouclage de ce manuscrit, on se doit d'indiquer aussi que l'usage de la transformée de Mellin [Mellin, 1897] pourrait être de quelques avantages dans la théorie de persistance de la régularité. En effet, cette piste suggérée par Ingo Witt nous permettrait d'affiner les résultats de [Benoit, 2024]. Cette piste en étant qu'à ses débuts nous ne pouvons malheureusement pas en dire beaucoup plus même si les résultats sont jusqu'ici encourageants.

<sup>51</sup>Il n'est pas totalement clair que l'auto-interaction et la génération des harmoniques soient deux phénomènes qui se découpent totalement.

Une fois le problème dans la bande bien compris, il s'agira de voir si la méthode peut servir de socle et être étendue au cas du problème dans le quart d'espace. L'une des principales difficultés que nous y voyons est que, comme mentionné précédemment, le processus de génération des phases par rebonds successifs est bien plus complexe dans le cas du quart d'espace que de la bande. En particulier, la question de la finitude du nombre de phases générées par rebonds successifs est une question centrale, ne serait-ce que pour donner un sens rigoureux à l'ansatz.

Nous pensons toutefois qu'ici la non-linéarité pourrait être de quelques avantages. En effet, pour de tels problèmes, bien souvent, la non-linéarité induit la naissance d'une singularité qui nous invite à considérer un résultat d'existence local en temps de la solution.

Or, en temps fini, chaque phase coûtant un peu de temps pour apparaître dans le procédé de génération des phases, il s'avère que le nombre de phases générées est, pour ainsi dire, automatiquement fini. C'est vrai en particulier si le terme source a son support un peu loin du coin. De surcroît, en temps fini, le phénomène d'auto-interaction est aussi grandement simplifié, en utilisant les séries de Neumann.

Par conséquent, du moins si l'on se fixe comme objectif la construction d'un développement d'optique géométrique approchant la solution sur une échelle de temps finie, passer du problème linéaire au problème non linéaire semble être (peut-être) un objectif réalisable.

### L'approximation visqueuse et l'interaction des couches limites

Pour conclure ce chapitre mentionnons une dernière piste de recherche que nous étudierons dans un futur proche. Il s'agit de l'approximation visqueuse des problèmes dans le quart d'espace, *i.e.*, au lieu d'étudier le problème (purent) hyperbolique (II.39) on s'intéresse au système (linéaire dans un premier temps) suivant

$$\begin{cases} L(\partial)u^\varepsilon - \varepsilon\mathcal{E}u^\varepsilon = f & \text{dans } \Omega, \\ u^\varepsilon|_{x_1=0} = 0 & \text{sur } \partial\Omega_1, \\ u^\varepsilon|_{x_2=0} = 0 & \text{sur } \partial\Omega_2, \\ u^\varepsilon|_{t=0} = u_0 & \text{sur } \Gamma, \end{cases} \quad (\text{II.72})$$

où  $0 < \varepsilon \ll 1$ , encode la viscosité induite par l'opérateur elliptique  $\mathcal{E}$ . A  $\varepsilon$  fixé, l'opérateur  $L(\partial) - \varepsilon\mathcal{E}$  étant parabolique on doit bien dans (II.72) prescrire toutes les composantes des traces sur le bord de  $\Omega$ .

Certains diront, probablement avec raison, que les approximations visqueuses de problème hyperboliques sont des problèmes qui ont plus de sens physique que les problèmes hyperboliques. En effet, si l'on s'intéresse aux problèmes hyperboliques de la mécanique des fluides alors le fluide considéré vient toujours avec un certain (possiblement très faible) coefficient de viscosité ( $1.799 \cdot 10^{-5}$  pour l'air).

Mathématiquement, comme les approximations visqueuses permettent d'obtenir, par passage à la limite, des solutions de systèmes (éventuellement non linéaires) hyperboliques, l'étude est aussi entièrement justifiée.

L'étude des approximations visqueuses consiste alors à étudier la limite de la solution  $u^\varepsilon$  lorsque le paramètre  $\varepsilon \downarrow 0$ , et plus précisément, à en donner un développement asymptotique dans l'esprit des développements d'optique géométrique que nous avons déjà décrits. De telles questions sont bien comprises dans le cas du demi-espace, notamment grâce aux travaux de Guès [Guès, 1990], qui établissent de tels développements dans un cadre caractéristique et pour des opérateurs non linéaires. On se doit aussi de mentionner les travaux de Rousset [Rousset, 2005] et [Rousset, 2003] ou encore [Rousset, 2001] dans le cas de problèmes en dimension un.



Le fait remarquable dans cette analyse est que le problème limite naturel de l'équation interne de (II.72), à savoir le problème hyperbolique,  $L(\partial)u = f$ , change de type. Ainsi, pour être bien posé, ce dernier ne nécessite qu'un nombre de conditions de bord idoines.

On veut donc, passer du problème (II.72) (qui spécifie toutes les composantes des traces), au problème limite hyperbolique qui n'en spécifie qu'un certain nombre. Par conséquent, pour assurer le passage à la limite, des couches limites localisées au voisinage de la frontière du demi-espace apparaissent pour corriger les valeurs des traces. Dans les travaux de Guès on peut alors donner une solution approchée de (II.72) de la forme<sup>52</sup>

$$u^\varepsilon(t, x, \frac{x_d}{\varepsilon}) = \underline{u}(t, x) + \mathcal{U}(t, x, \frac{x_d}{\varepsilon})$$

où

- $\underline{u}$  est la solution de  $L(\partial)\underline{u} = f$  (avec des conditions de bord idoines et à déterminer).
- $\mathcal{U}$  est une fonction à décroissance rapide dans la dernière variable. Elle constitue donc une couche limite (correction localisée sur  $\{x_d \leq C\varepsilon\}$ ).

Nous aimerions dans une future recherche faire passer l'analyse, d'abord dans un cadre linéaire, du demi-espace vers le quart d'espace. On s'attend alors un développement de la solution de (II.72) de la forme

$$u^\varepsilon \sim \underline{u}(t, x) + \mathcal{U}_1(t, x, \frac{x_1}{\varepsilon}) + \mathcal{U}_2(t, x, \frac{x_2}{\varepsilon}),$$

où  $\underline{u}$  est la solution du problème à coin hyperbolique et où  $\mathcal{U}_i$  est une couche limite corrective localisée le long du bord  $\partial\Omega_i$ . Il s'agirait alors de comprendre comment les deux couches limites dialoguent l'une avec l'autre dans le voisinage du coin  $\{x_1, x_2 \leq C\varepsilon\}$ . De prime abord un tel dialogue ne peut être exclu au voisinage du coin. Toutefois, nous pensons pas qu'un tel dialogue ait effectivement lieu : en effet, nous pensons que pour des questions de compatibilités, la valeur au coin  $u|_{x_1=x_2=0} = 0$ . Or, c'est cette valeur qui détermine la valeur des couches limites au voisinage du coin.

L'analyse précédente nécessite cependant de donner un sens précis à la valeur de la double trace. On retrouve donc le besoin d'une théorie de régularité dans des espaces fonctionnels assez fins. Par exemple, avec une théorie de persistance de régularité dans les espaces de Sobolev usuels  $H^s$ , la définition d'une telle trace ne pose pas de problème<sup>53</sup>.

---

<sup>52</sup>On s'intéresse ici uniquement au cas non caractéristique. Le cas caractéristique semblant comme précédemment indiqué être un problème en dehors de notre portée actuelle. Dans le demi-espace caractéristique, on a deux couches limites l'une non caractéristique d'échelle  $\frac{1}{\varepsilon}$ , la seconde caractéristique d'échelle  $\frac{1}{\sqrt{\varepsilon}}$ .

<sup>53</sup>Notre résultat d'existence d'une solution forte (si il s'avère effectivement valide) précédemment mentionné en bas de page permettrait alors de définir cette double trace rigoureusement.

## Chapter III

# Schémas numériques de type différences finies

Ce chapitre expose nos contributions quant à l'étude de la stabilité des schémas d'approximation par différences finies pour la solution d'un problème hyperbolique posé dans un domaine à bord(s). Il reprend essentiellement le même plan que le Chapitre II. On considérera donc, une partie introductive formée des Sections III.1, III.2 et III.3 qui décrit les problèmes considérés, des motivations et les résultats les plus marquants de la littérature.

La Section III.4 se place dans un premier exemple de géométrie avec plusieurs bords, celle du segment. On y établit alors un résultat de stabilité. Plus précisément, on donne une caractérisation des conditions de bord numériques conduisant à la stabilité du schéma dans le segment.

La Section III.5 comprend des exemples de schémas stables dans la géométrie du quart d'espace. Ces exemples, en traitant des schémas particuliers et/ou des conditions de bord particulières établissent un premier jalon vers une caractérisation de la stabilité théorique, plus générale.

Enfin, la Section III.6 regroupe quelques questions ouvertes et perspectives de recherche que nous trouvons intéressantes et que nous nous proposons d'étudier à l'avenir.

### III.1 Introduction et description des schémas

Une fois un résultat d'existence et d'unicité de la solution d'un problème aux limites dans un domaine à bord établi, se pose alors très naturellement la question de l'obtention de méthodes numériques facilement implémentables permettant d'en calculer une valeur approchée. C'est cette question qui va nous intéresser dans ce chapitre.

Les méthodes numériques que nous étudierons ici sont celles obtenues par discrétisation de type différences finies. Ces méthodes, qui peuvent être considérées comme quelque peu naïves, sont toutefois parfaitement justifiées dans des géométries aussi simples que le demi-espace ou le quart d'espace. En effet, il est clair que ces deux dernières géométries peuvent être discrétisées en utilisant un maillage rectangulaire, ce dernier pouvant même être choisi régulier. On n'a pas ici d'intérêt particulier à utiliser des maillages élaborés<sup>1</sup>, qui inviteraient alors à travailler avec des discrétisations par éléments ou bien par volumes finis.

Ainsi, en un certain sens, les approximations par différences finies sont ici les plus simples que l'on peut imaginer dans ce contexte. Comprendre de telles méthodes "simples" est alors une première étape

---

<sup>1</sup>Étant particulièrement intéressé par le comportement de la solution au bord, on pourrait toutefois vouloir considérer des maillages plus fins au voisinage de la frontière. Ceci reste permis dans un cadre différences finies si l'on permet aux pas de discrétisation spatiaux de varier. Nous n'avons toutefois pas pris en compte cette possibilité dans le cadre d'étude que nous considérons ici.

dans la mise en œuvre de méthodes plus ”compliquées”.

Toutefois, comme on va le voir, l’analyse établissant la convergence de la suite des approximations vers la solution exacte de l’équation aux dérivées partielles n’est pas triviale, même si cette dernière partage, sans surprise, de nombreux points de ressemblance avec la théorie bien posée continue. Cette théorie de convergence comporte, toujours, même dans des géométries très simples, à l’heure actuelle, de nombreuses zones d’ombre, en particulier, on prendra garde à qualifier cette théorie de ”simple”.

Un exemple de ces lacunes que nous trouvons frappant est que dans tout ce qui suit (sauf dans la Section III.5, *i.e.* dans le quart d’espace) on se placera systématiquement dans le cas unidimensionnel :  $d = 1$ . En effet, à notre connaissance, à l’heure actuelle, on ne dispose pas d’une théorie de convergence vraiment convaincante et totalement aboutie dans le cas multidimensionnel, même dans une géométrie avec un seul bord. On tempèrera toutefois notre propos en mentionnant ici le travail de Michelson [Michelson, 1983] qui indique que la condition assurant la stabilité dans la demi-droite (la condition de Kreiss-Lopatinskii uniforme discrète) se généralise au cas du demi-espace. Toutefois ce travail se fait sous des hypothèses assez contraignantes sur les schémas considérés qui excluent de l’analyse de nombreux schémas usuels. On renvoie à la suite de ce mémoire pour plus de détails sur ce point.

On se donne donc deux pas  $\Delta t, \Delta x > 0$ , discrétisant respectivement les ensembles  $\mathbb{R}_{+,t}$  et  $\mathbb{R}_x$ . Ces derniers seront complétés d’un pas  $\Delta y > 0$  dans le cas du quart de plan. On notera dans ce qui suit  $(t^n)_{n \in \mathbb{N}}$ ,  $t^n := n\Delta t$  et  $(x_j)_{j \in \mathbb{Z}}$ ,  $x_j := j\Delta x$ , les nœuds de la discrétisation. Dans l’équation  $\partial_t u + A\partial_x u$ ,  $A \in \mathbf{M}_{N \times N}(\mathbb{R})$ , on vient alors approcher l’opérateur de dérivation spatiale par une combinaison linéaire de la fonction  $u$  évaluée aux  $\ell + r > 0$ , points voisin de  $\underline{x}$ , avec  $\ell, r \geq 0$ . C’est-à-dire :

$$A\partial_x u(t, \underline{x}) \sim \sum_{\mu=-\ell}^r A^\mu u(t, \underline{x} + \mu\Delta x),$$

où les coefficients  $A^\mu$  sont des matrices données dans  $\mathbf{M}_{N \times N}(\mathbb{R})$ .

On discrétisera toujours ici l’opérateur de dérivation temporelle par un schéma d’Euler explicite (donc en particulier d’ordre un). La théorie permet toutefois de considérer des schémas avec plus d’un pas en temps (on renvoie, par exemple, à [Coulombel, 2015]-[Coulombel, 2021] ou [Coulombel, 2012]). Nos résultats s’appliquant, pour une majeure partie, à des schémas à un pas en temps, on a choisi, dans cette partie introductive, de se restreindre à ce cadre simplifié dans un souci de simplification des notations.

Posant  $(U_j^n)_{n \in \mathbb{N}, j \in \mathbb{Z}}$ , le vecteur stockant les approximations  $U_j^n$ , qui ont vocation à approcher les  $u(t^n, x_j)$ , le schéma numérique que l’on considère dans ce qui suit s’écrira sous la forme générique :

$$\begin{cases} U_j^{n+1} = QU_j^n + \Delta t f_j^n & \text{pour } n \geq 0, j \in \mathbb{Z}, \\ U_j^0 = u_j & \text{pour } j \in \mathbb{Z}, \end{cases} \quad (\text{III.1})$$

où les termes sources  $(f_j^n)_{n \in \mathbb{N}, j \in \mathbb{Z}}$  et  $(u_j)_{j \in \mathbb{Z}}$  sont donnés et correspondent à des discrétisations du terme source intérieur  $f$  et de la donnée initiale  $u_0$ . On insiste sur le fait que, pour le moment, l’approximation précédente se fait sur toute la droite réelle  $\mathbb{R}_x$ .

On verra en détails dans le Paragraphe III.3.1, comment le fait de passer à une géométrie à bord influe sur la formulation générique (III.1).

Dans (III.1), l’opérateur  $Q$  est donné par

$$Q := \sum_{\mu=-\ell}^r A^\mu \mathbf{T}^\mu \text{ où } (\mathbf{T}U)_j := U_{j+1} \text{ est l’opérateur de translation,} \quad (\text{III.2})$$

et où les deux entiers  $\ell$  et  $r$  sont positifs. On a donc affaire à un schéma avec  $\ell$  pas vers la gauche et  $r$  pas vers la droite<sup>2</sup> et en vue d'obtenir des schémas ayant un ordre de convergence "élevé" on veut pouvoir prendre de "grandes" valeurs de  $\ell$  et de  $r$ . Dans l'équation (III.2) les coefficients  $A^\mu$ ,  $\mu \in \llbracket -\ell, r \rrbracket$ , sont donnés dans  $\mathbf{M}_{N \times N}(\mathbb{R})$ , ils sont indépendants de  $\Delta t$ , mais peuvent dépendre du ratio CFL (Courant-Friedrichs-Lewy) [Courant et al., 1928],  $\lambda := \frac{\Delta t}{\Delta x}$ .

Les deux questions qui se posent typiquement sur le schéma aux différences finies (III.1) sont celle de sa stabilité et celle de sa consistance. On espère obtenir la convergence directement par le (méta)théorème de Lax [Lax and Richtmyer, 1956], le sens des notions de stabilité et de consistance devant bien sûr être précisés dans notre cadre d'étude (voir Paragraphe III.3.2).

Dans nos travaux, nous avons uniquement considéré la question de la stabilité. Il s'agit, bien souvent, de la plus intéressante d'un point de vue mathématique. En effet, la question de la consistance se ramène bien souvent, dans ce cadre, à des développements limités. La théorie de la stabilité pour le problème dans la droite réelle est, pour ainsi dire, une question totalement comprise depuis les travaux de von Neumann. On renvoie, par exemple, à l'exposition complète de [[Coulombel, 2012]-Section 2]. Ainsi, dans ce qui suit nous nous intéresserons, à proprement parler, à essayer de comprendre en quoi considérer (III.1) dans une géométrie à bord bouleverse la théorie de stabilité sous-jacente.

La théorie de stabilité dans la droite réelle étant forte d'enseignements, on se permet d'en rappeler, ici, à titre de conclusion de cette section, quelques faits marquants.

Commençons par préciser la notion de stabilité utilisée. Ici, il s'agira de la notion la plus naturelle que l'on puisse imaginer. De façon générique, la stabilité revient à requérir un contrôle des itérées de la solution du schéma au cours du temps, par la donnée initiale :

$$\forall n \geq 0, \|U^n\| \leq C\|u\|.$$

Cette estimation peut être complétée dans le membre de droite par la contribution d'un éventuel terme source à l'intérieur. L'estimation précédente tient pour une certaine constante  $C > 0$ , indépendante des pas de discrétisation et dans une certaine norme  $\|\cdot\|$ , à préciser. Ici on sait que, sauf cas exceptionnel, la "bonne" norme à considérer au niveau continu est la norme  $L^2$  (voir Section II.2). Il est donc naturel de considérer la norme discrétisant la norme  $L^2$  et on définit alors pour la stabilité

$$\|U\| := \sqrt{\sum_{j \in \mathbb{Z}} \Delta x |U_j|^2},$$

**Définition III.1.1 (Stabilité de (III.1))** *Le schéma d'approximation par différences finies (III.1) est dit stable si il existe  $C > 0$  telle que pour tout  $\Delta t \in ]0, 1]$  on ait*

$$\forall n \in \mathbb{N}, \sqrt{\sum_{j \in \mathbb{Z}} |U_j^n|^2} \leq C \sqrt{\sum_{j \in \mathbb{Z}} |u_j|^2}. \quad (\text{III.3})$$

*On insiste sur le fait que la constante numérique  $C$  est indépendante des pas de discrétisations.*

Comme pour établir le caractère fortement bien posé d'un problème de Cauchy hyperbolique, caractériser les discrétisations stables se fait aisément par l'analyse de Fourier. En restant quelque peu formel, si l'on considère la transformée de Fourier discrète  $\widehat{U}^n(\xi)$  de  $(U_j^n)_{n \geq 0, j \in \mathbb{Z}}$  on remarque que cette dernière vérifie le problème

$$\forall \xi \in \mathbb{R}, \begin{cases} \widehat{U}^{n+1}(\xi) = \mathcal{A}(e^{i\xi}) \widehat{U}^n(\xi) & \text{pour } n \geq 0, \\ \widehat{U}^0(\xi) = \widehat{u}(\xi), \end{cases} \quad (\text{III.4})$$

<sup>2</sup>On pense d'ailleurs que ce choix (astucieux) de notations sur les pas, constitue, probablement, notre principale contribution à la théorie.

où la matrice, dite d'amplification,  $\mathcal{A}(\xi)$ , qui agit comme un symbole, est définie par

$$\forall z \in \mathbb{C}, \mathcal{A}(z) := \sum_{\mu=-\ell}^r z^\mu A^\mu.$$

La relation de récurrence (III.4) nous invite à définir

$$\widehat{U}^n(\xi) := \mathcal{A}^n(\xi) \widehat{u}(\xi), \quad (\text{III.5})$$

équation sur laquelle on se convainc assez facilement que le schéma sera stable au sens de la Définition III.1.1, si et seulement si la matrice d'amplification est de puissances bornées. Puisque l'on va aussi vouloir revenir dans l'indice de position  $j$ , par une transformée de Fourier inverse, on demande de surcroît que cette borne soit uniforme par rapport à la variable de fréquence  $\xi$ .

On a alors la proposition suivante :

**Proposition III.1.1 (Caractérisation de la stabilité)** *Le schéma d'approximation (III.1) est stable au sens de la Définition III.1.1, si et seulement si, il existe  $C > 0$  telle que*

$$\forall n \in \mathbb{N}, \forall \xi \in \mathbb{R}, |\mathcal{A}^n(e^{i\xi})| \leq C.$$

Caractériser les matrices de puissances bornées, et ce uniformément, n'est pas un exercice immédiat et facile. On se convainc assez aisément que l'on a la condition nécessaire suivante dite de von Neumann.

**Proposition III.1.2 (Condition de von Neumann)** *Si le schéma (III.1) est stable au sens de la Définition III.1.1, alors il vérifie la condition de rayon spectral dite de von Neumann :*

$$\forall \xi \in \mathbb{R}, \rho(\mathcal{A}(e^{i\xi})) \leq 1,$$

où  $\rho(\cdot)$  désigne le rayon spectral.

Pour une caractérisation totale des matrices de puissances bornées (et donc en particulier des matrices donnant lieu à des schémas stables) on renvoie au fameux théorème matriciel de Kreiss [Kreiss, 1962], qui, comme son analogue continu décrivant les opérateurs hyperboliques, caractérise les matrices de puissances bornées.

En conclusion les grands piliers sur lesquels repose la théorie de la stabilité que l'on vient de résumer sont les suivants : l'analyse de Fourier qui permet de ramener le schéma aux différences finies sous la forme d'une relation de récurrence de la forme (III.4). Relation qui est explicitement résoluble mais qui pour revenir dans l'indice  $j$  va nécessiter une borne uniforme par rapport au paramètre de fréquence.

Comme on va le voir, bien que l'on ne puisse pas directement se baser sur l'analyse de Fourier, cette approche est adaptable pour traiter la question de la stabilité dans la demi-droite. Par analogie avec le problème continu, la transformation à considérer sera alors la transformée de Laplace. On verra alors apparaître une version discrète de la condition de Kreiss-Lopatinskii uniforme qui caractérisera les conditions de bords numériques donnant lieu à un schéma stable.

## III.2 Motivations

Avant d'entrer dans le vif du sujet sur le problème dans la demi-droite, et plus particulièrement sur la détermination des conditions de bord donnant lieu à des schémas stables, on pense important de motiver et de justifier l'étude de la stabilité des schémas de différences finies dans des domaines à bord.

Bien (trop ?) souvent, les questions liées aux conditions de bords sont évitées lorsque l'on s'intéresse à des approximations numériques de solutions d'équations aux dérivées partielles. C'est d'autant plus vrai lorsque l'on considère ces dernières sous le prisme de l'analyse numérique "appliquée", dont la validation se base sur des tests numériques et que l'on s'écarte de l'analyse numérique "théorique" qui, elle, vise à démontrer la convergence des méthodes.

Nous pensons que cela ne devrait pas être le cas. En effet, la vocation première d'une méthode numérique est d'être implémentable, indépendamment de l'étude de sa convergence. Or, implémenter nécessite forcément de prendre en compte des conditions de bord. Le monde est (ici) mal fait, mais il faut rappeler que les ordinateurs, ces machines imparfaites qui disposent d'une mémoire finie, ne peuvent pas modéliser l'espace  $\mathbb{R}^d$  en entier. Ainsi, même si l'on s'intéresse à l'approximation de la solution d'un problème de Cauchy, qui semble bien peu se préoccuper des conditions de bord, lorsque l'on implémente le schéma numérique, ce dernier ne peut être posé que dans un (grand) hypercube de  $\mathbb{R}^d$ . Donc ce dernier va admettre des bord numériques. Si l'on fait tourner le schéma assez longtemps pour que l'onde commence à frapper le bord numérique, cette dernière va alors être réfléchi et induire des termes de réflexions parasites.

Comprendre comment minimiser voir annihiler ces termes parasites fait l'objet d'une littérature conséquente dont on va évoquer quelques unes des principales approches.

Avec la hausse de la mémoire machine et la diminution des temps de calcul, une méthode possible pour modéliser la propagation d'une onde dans l'espace  $\mathbb{R}^d$  est de se placer dans un très grand hypercube et d'arrêter la simulation (pour peu que l'on ne veuille pas d'une simulation en temps "trop long"), lorsque l'onde commence à venir interagir avec le bord numérique. C'est, en effet, une possibilité qui est souvent utilisée. Mais notons que cette dernière manque un petit peu de panache et est, à notre avis, assez insatisfaisante.

Une autre approche, qui nous paraît plus satisfaisante d'un point de vue mathématique, est celle des conditions de bords absorbantes ou transparentes. C'est-à-dire que l'on va venir chercher à imposer sur le bord de notre hypercube des conditions de bords non dictées par la physique de l'équation, dont la vocation est de venir minimiser (voir annihiler) ces termes de réflexions parasites. La littérature sur ce sujet est abondante et on renvoie, par exemple et de façon non exhaustive, à [Engquist and Majda, 1977] et [Ehrhardt, 2010], ou encore aux travaux de Coulombel et ses coauteurs qui recourent parfaitement le cadre d'étude dans lequel nous nous sommes placés [Coulombel, 2019] et [Besse et al., 2021]. Il est à noter toutefois que ces conditions sont le plus souvent définies dans le demi-espace. On n'est pas encore totalement dans le cadre d'implémentation d'un hypercube. On renverra toutefois à [Besse et al., 2021] pour une étude dans un quart d'espace, ou encore à [Halpern and Rauch, 2016] pour l'étude des *Perfectly Matched Layers* (PML) dans un coin à trois dimensions qui visent à cantonner l'influence des réflexions parasites au voisinage de la frontière.

Une autre raison qui, selon nous, penche vraiment en faveur du fait que les conditions de bord numériques devraient être étudiées plus en détails et qu'elles ne devraient pas être évacuées d'un tour de main est la nature assez arbitraire de leur l'origine. Pour illustrer ceci, considérons le cas très simple d'une équation de transport sortante:

$$\begin{cases} \partial_t u + a \partial_x u = 0 & \text{dans } \mathbb{R}_+^2, \\ u|_{t=0} = u_0 & \text{sur } \mathbb{R}_+, \end{cases}$$

où  $a < 0$  et  $u_0$  sont donnés.

Le phénomène de transport étant sortant, aucune condition de bord ne doit être prescrite sur le bord  $\{x = 0\}$ . Toutefois, si l'on veut approcher la solution  $u$  par un schéma ayant au moins un pas à gauche comme par exemple, le très naturel schéma de Lax-Friedrichs en première approximation. Alors, ce schéma

pour opérer nécessite une condition de bord numérique. On expliquera plus en détails la provenance de cette condition de bord au Paragraphe III.3.1.

L'utilisateur du schéma a donc un choix à faire pour définir cette condition de bord numérique. Ce choix est assez arbitraire dans la mesure où la condition de bord numérique n'est pas directement dictée par la physique du problème continu. Il n'est toutefois pas totalement arbitraire dans la mesure où prendre des conditions de bord numérique assurant la consistance du schéma (de préférence au plus grand ordre possible) sera, bien sûr, préféré. Mais, dès qu'il y a un choix, se pose alors naturellement la question de déterminer comment l'on peut assurer le fait que l'on a effectivement "bien" (on entend par là, dans ce cadre, que l'on peut assurer *a minima* la stabilité du schéma) choisi. C'est pourquoi, dans la suite de ce chapitre, nous nous intéresserons particulièrement à l'établissement d'une caractérisation des conditions de bords numériques qui donnent lieu à un schéma stable.

Dans une géométrie assez générale, c'est un programme ambitieux qui est rarement affronté dans la littérature. En effet, beaucoup de travaux sur la stabilité de schémas de types différences finies dans des discrétisations à bord(s) se "contentent" (et cela n'est déjà pas trivial) de présenter une ou plusieurs conditions de bord assurant la stabilité du schéma. En effet, d'un point de vue pratique, disposer d'une ou de quelques conditions assurant la stabilité est suffisant.

C'est d'ailleurs l'approche que nous avons choisie dans les deux articles [Benoit and Coulombel, b] et [Benoit, 2018] et dans le *preprint* [Benoit and Coulombel, a].

La première caractérisation complète des conditions de bord assurant la stabilité a été faite dans l'article pionnier de Gustafsson, Kreiss et Sundstrom [B. Gustafsson and Sundstrom., 1972], on renvoie aussi au gros travail de revue de Coulombel dans [Coulombel, 2012]. Nous allons décrire les éléments les plus marquants de cette caractérisation dans la section suivante.

### III.3 Le cas de la demi-droite

Cette section résume les résultats connus dans la littérature concernant la caractérisation des conditions de bord numériques dans le cas de la demi-droite. On reprend donc ici les points clés de l'analyse de [B. Gustafsson and Sundstrom., 1972].

On étudie une discrétisation du problème continu :

$$\begin{cases} \partial_t u + A \partial_x u = f & \text{pour } (t, x) \in \mathbb{R} \times \mathbb{R}_+, \\ B u|_{x=0} = g & \text{pour } t \in \mathbb{R}, \\ u|_{t \leq 0} = u_0 & \text{pour } x \in \mathbb{R}_+, \end{cases} \quad (\text{III.6})$$

où  $A \in \mathbf{M}_{N \times N}(\mathbb{R})$ . Par hyperbolicité de  $A$ , on remarque que le problème (III.6) peut se ramener à  $N$  équations de transports découplées à l'intérieur. Le couplage ayant lieu au sein de la condition de bord sur  $\{x = 0\}$ .

On discrétise alors l'équation interne de (III.6) par une approximation de type différences finies semblable à celle du problème dans la droite complète et on prend ici une condition initiale homogène<sup>3</sup> :

$$\begin{cases} U_j^{n+1} = Q U_j^n + \Delta t f_j^n & \text{pour } n \geq 0, j \geq 1, \\ \text{condition de bord numérique,} & \\ U_j^0 = 0 & \text{pour } j \geq 1, \end{cases} \quad (\text{III.7})$$

où l'on rappelle que  $Q$  est défini en (III.2) et dans le paragraphe suivant on va décrire plus en détails la condition de bord numérique.

<sup>3</sup>Certains des résultats que l'on va décrire dans ce qui suit peuvent se généraliser à des schémas avec un *stencil* en temps supérieur à un, noté  $s$ . On prescrit alors les valeurs des  $U_j^\sigma$  pour  $\sigma \leq s - 1, j \geq 1$

### III.3.1 Description des conditions de bord numériques

Pour décrire la condition de bord numérique, il est bon de décrire comment le schéma (III.7) opère. On suppose dans ce qui suit que l'on a calculé la solution du schéma à un certain niveau de temps  $n$ , ( $U^n$ ) et on explique comment (III.7) permet de calculer la valeur des  $U^{n+1}$ . Nous avons deux cas à considérer selon la valeur du paramètre de position  $j$  :

- Si  $j \leq \ell + 1$ , alors l'opérateur  $QU_j^n$  ne fait intervenir que des  $U_j^n$  avec  $j \geq 1$ . C'est-à-dire des coefficients de la suite des approximations ( $U_j$ ) à l'instant  $n$  et à l'intérieur. Ces derniers sont connus et le schéma permet de calculer  $U_j^{n+1}$ .
- Si  $1 \leq j \leq \ell$ , alors l'opérateur  $QU_j^n$  fait intervenir des valeurs de  $U_k^n$  pour  $k < 1$ . Ces derniers ne sont pas déterminés par l'équation interne. Pour que le schéma puisse alors opérer on doit prescrire les valeurs de ces termes. Le cas extrémal est atteint pour la valeur  $j = 1$  pour laquelle  $QU_1^n$  fait intervenir les  $U_k^n$  pour  $k \in \llbracket 1 - \ell, 1 + r \rrbracket$ . On introduit donc les mailles dites fantômes  $j \in \llbracket 1 - \ell, 0 \rrbracket$ . Donner un moyen de déterminer les valeurs des  $U_j^n$ , pour de tels indices  $j$ , à partir de celles des valeurs intérieures (*i.e.* celles pour  $j \geq 1$ ) est ce que l'on entend par une condition de bord numérique.

Dans ce qui suit, pour déterminer les valeurs aux mailles fantômes, on va imposer des conditions de bord numériques génériques s'écrivant sous la forme :

$$U_j^{n+1} = \sum_{\sigma=0}^1 B^{\sigma,j} U_1^{n+\sigma} + g_j^{n+1} \text{ pour } n \geq 0 \text{ et } j \in \llbracket 1 - \ell, 0 \rrbracket, \quad (\text{III.8})$$

où les opérateurs  $B^{\sigma,j}$  s'écrivent sous la forme

$$B^{\sigma,j} := \sum_{\mu=0}^b B^{\sigma,j,\mu} \mathbf{T}^\mu,$$

où  $b \geq 0$ , est le nombre de pas de la condition de bord. Par conséquent, le calcul des mailles fantômes se fait *via* celui des coefficients intérieurs  $U_j^m$  avec  $m = n$  ou  $m = n + 1$  et  $j \in \llbracket 1, 1 + b \rrbracket$ . Dans tout ce qui suit, on supposera, bien que cela ne soit pas nécessaire, que l'on prend un *stencil* sur le bord inférieur à celui de l'intérieur plus précisément  $b < r$ . On renvoie à [Coulombel, 2012] pour une analyse détaillée dans le cas où  $b \geq r$ .

Cette hypothèse est à notre sens très peu restrictive car il est quelque peu "immoral" (bien que cela puisse être fait) de considérer un schéma ayant un nombre de pas sur le bord supérieur à celui que l'on utilise à l'intérieur. En effet, pourquoi chercher à obtenir un ordre de convergence au bord supérieur à l'ordre à l'intérieur quand ce dernier comprend la majorité des mailles ?

On remarquera, en particulier, que déjà le formalisme de la condition de bord numérique (III.8) est plus riche que celui de la discrétisation intérieure. En effet, on permet, d'une part, aux coefficients de bord de dépendre de l'indice de position  $j$  (ce n'est pas le cas pour les coefficients intérieurs  $A^\mu$ ) et, d'autre part, on permet ici un calcul utilisant deux pas de temps. Ce formalisme plus riche est juste "motivé" par le fait que c'est celui qui est décrit dans [B. Gustafsson and Sundstrom., 1972] et qui a été ensuite repris dans les articles suivants voir [Coulombel, 2011]-[Coulombel, 2012].

Une analyse permettant aussi aux coefficients intérieurs de varier avec  $j$  pourrait aussi être intéressante et permettrait un gain dans les applications.



En regroupant (III.7) et (III.8), on obtient alors le schéma dans la demi-droite :

$$\begin{cases} U_j^{n+1} = QU_j^n + \Delta t f_j^n & \text{pour } n \geq 0, j \geq 1, \\ U_j^{n+1} = \sum_{\sigma=0}^1 B^{\sigma,j} U_1^{n+\sigma} + g_j^{n+1} & \text{pour } n \geq 0 \text{ et } j \in \llbracket 1 - \ell, 0 \rrbracket, \\ U_j^0 = u_{0,j} & \text{pour } j \geq 1 - \ell. \end{cases} \quad (\text{III.9})$$

### III.3.2 Définitions et commentaires sur la stabilité

Dans ce paragraphe on expose comment le concept de stabilité de la Définition III.1.1 peut être étendu à des problèmes posés dans la demi-droite. Pour ce faire, on rappelle que le cadre naturel dans lequel le problème (III.6) est bien posé est celui de l'espace  $L^2$  à poids exponentiel, et que l'on a alors une estimation d'énergie s'écrivant : il existe  $C > 0$ , telle que, pour tout  $\gamma > 0$ , on ait l'inégalité :

$$\gamma \|u\|_{L_\gamma^2(\Omega)}^2 + \|u|_{x=0}\|_{L_\gamma^2(\partial\Omega)}^2 \leq C \left( \frac{1}{\gamma} \|f\|_{L_\gamma^2(\Omega)}^2 + \|g\|_{L_\gamma^2(\partial\Omega)}^2 \right). \quad (\text{III.10})$$

Il est naturel d'imposer une notion de stabilité qui donne une version discrétisée de l'estimation d'énergie continue. Plus précisément, on prendra pour stabilité celle dite stabilité GKS (Gustafsson-Kreiss-Sundstrom) introduite par [Kreiss, 1968] et [B. Gustafsson and Sundstrom., 1972]

**Définition III.3.1** ([Kreiss, 1968]-[B. Gustafsson and Sundstrom., 1972]) *On dit que le schéma (III.9) à donnée initiale nulle ( $u_{0,j} \equiv 0$ ) est stable au sens GKS si il existe  $C > 0$ , telle que, pour tout  $\gamma > 0$ , pour tout  $\Delta t \in ]0, 1]$ , la solution  $(U_j^n)_{n \in \mathbb{N}, j \geq 1 - \ell}$  de (III.9) vérifie l'estimation :*

$$\begin{aligned} & \frac{\gamma}{\gamma \Delta t + 1} \sum_{n \geq 1} \sum_{j \leq 1 - \ell} \Delta t \Delta x e^{-2\gamma n \Delta t} |U_j^n|^2 + \sum_{n \geq 1} \sum_{j=1-\ell}^r \Delta t e^{-2\gamma n \Delta t} |U_j^n|^2 \\ & \leq C \left( \frac{\gamma \Delta t + 1}{\gamma} \sum_{n \geq 0} \sum_{j \geq 1} \Delta t \Delta x e^{-2\gamma(n+1)\Delta t} |F_j^n|^2 + \sum_{n \geq 1} \sum_{j=1-\ell}^0 \Delta t e^{-2\gamma n \Delta t} |g_j^n|^2 \right). \end{aligned} \quad (\text{III.11})$$

Comme précédemment indiqué l'estimation (discrète) (III.11) est une version dicrétisée de (III.10) dans la mesure où l'on peut (formellement) passer de (III.11) à (III.10) en prenant les limites  $\Delta t, \Delta x \downarrow 0$ .

On doit mentionner ici que l'estimation (III.11) n'est en fait pas exactement celle qui est donnée dans l'article fondateur [B. Gustafsson and Sundstrom., 1972]. En effet, elle diffère de l'estimation initiale par les points suivants :

- Dans l'estimation de [B. Gustafsson and Sundstrom., 1972], on accepte que l'estimation précédente soit vraie seulement au dessus d'un certain seuil  $\gamma \geq \gamma_0$ , avec  $\gamma_0 > 0$ . Ainsi, dans les travaux de [B. Gustafsson and Sundstrom., 1972], on permet à la solution numérique d'avoir une croissance exponentielle non triviale par rapport à  $n$  et donc en temps. Ce point a déjà été discuté dans la Section II.3, on verra que l'on sera amené à se poser la même question de caractérisation des schémas induisant ou non cette croissance lorsque l'on considérera le problème dans l'intervalle (voir Section III.4).
- Cela peut sembler anecdotique mais on n'impose pas exactement le même contrôle comme terme de trace dans (III.11) que dans l'estimation de [B. Gustafsson and Sundstrom., 1972]. Dans cette dernière, la trace est définie pour  $j \in \llbracket 1 - \ell, 0 \rrbracket$ . On utilise donc ici les traces étendues de Coulombel voir [Coulombel, 2012]. Ce point sera toutefois d'importance lorsque l'on sera amené à considérer les opérateurs de traces (versions discrètes des opérateurs,  $\mathbf{T}_{0 \rightarrow 1}$  et  $\mathbf{T}_{1 \rightarrow 0}$ , déjà rencontrés dans la géométrie de la bande).

### III.3.3 La théorie GKS en deux mots

Dans ce paragraphe, on décrit les grandes lignes de l'approche de [B. Gustafsson and Sundstrom., 1972] pour caractériser les conditions de bords numériques qui donnent lieu à un schéma stable dans la demi-droite. On se place sous la notion de la stabilité de la Définition III.3.1. Les deux grandes étapes de l'analyse sont :

- Une série de transformations préalables pour mettre le schéma sous une forme résolvente qui nous donne une relation de récurrence à résoudre.
- La démonstration de la stabilité à proprement parler *via* la construction d'un *symétriseur* discret jouant le même rôle que le *symétriseur* de Kreiss dans le cas continu.

#### Les différentes réductions et la condition de Kreiss-Lopatinskii discrète

Dans ce qui suit, on supposera que l'estimation discrète (III.11) (où l'on a posé  $\Delta t = 1$  pour simplifier autant que faire se peut les expressions) est satisfaite.

L'une des étapes-clés dans la démonstration de la stabilité du schéma (III.9) est de se ramener à une relation de récurrence qui va jouer un rôle analogue à celui que l'équation (III.4) jouait pour le problème dans la droite réelle. Il est à noter que nous ne pouvons pas ici prendre de transformée de Fourier discrète pour "tuer" l'indice de position  $j$ . Cependant, fort de la théorie du problème continu dans le demi-espace, un bon choix est de prendre la transformée de Laplace discrète en temps. Pour  $z \in \mathbb{C} \setminus \{0\}$ , on pose  $U_j^n := z^n W_j$ . Ainsi, du moins au niveau formel, la suite  $(W_j)_{j \geq 1-\ell}$  vérifie :

$$\begin{cases} W_j - z^{-1} Q W_j = F_j & \text{pour } j \geq 1, \\ W_j - z^{-1} B^{0,j} W_1 - B^{1,j} W_1 = G_j & \text{pour } j \in \llbracket 1 - \ell, 0 \rrbracket, \end{cases} \quad (\text{III.12})$$

où l'expression exacte des termes sources  $(F_j)_{j \geq 1-\ell}$  et  $(G_j)_{1-\ell \leq j \leq 0}$  n'est pas capitale pour ce qui suit.

Il est alors établi, on renvoie par exemple à [[Coulombel, 2012]-Section 3.2] pour une démonstration rigoureuse, que le caractère stable du schéma (III.9) peut se retranscrire sur le schéma (III.12) de la manière suivante :

**Proposition III.3.1** *Le schéma numérique (III.9) est GKS-stable au sens de la Définition III.3.1, si et seulement si, le schéma "mi-résolvant" (III.12) est stable au sens où il existe  $C > 0$ , telle que pour tout  $z \in \mathcal{U} := \{z \in \mathbb{C} \setminus |z| > 1\}$ , pour tout  $(F_j)_{j \geq 1-\ell} \in \ell^2(\llbracket 1 - \ell, \infty \rrbracket)$  et pour tout  $(G_j)_{1-\ell \leq j \leq 0} \in \mathbb{C}^\ell$ , le schéma (III.12) admet une unique solution  $(W_j)_{j \geq 1-\ell} \in \ell^2(\llbracket 1 - \ell, \infty \rrbracket)$  qui vérifie*

$$\frac{|z|-1}{|z|} \sum_{j \geq 1-\ell} |W_j|^2 + \sum_{j=1-\ell}^r |W_j|^2 \leq C \left( \frac{|z|}{|z|-1} \sum_{j \geq 1} |F_j|^2 + \sum_{j=1-\ell}^0 |G_j|^2 \right).$$

Nous sommes presque en position de trouver notre formulation récursive. Pour cela on reformule le schéma mi-résolvant (III.12) en termes d'un vecteur augmenté  $\mathcal{W}_j := (W_{j+r-1}, \dots, W_{j-\ell})^T \in \mathbb{C}^{(\ell+r)N}$ . On commence par introduire les matrices :

$$\forall \mu \in \llbracket -\ell, r \rrbracket, \mathbb{A}^\mu(z) := \delta_{\mu,0} I - z^{-1} A^\mu$$

$$\text{et } \forall \mu \in \llbracket 0, b \rrbracket, \forall j \in \llbracket 1 - \ell, 0 \rrbracket, \mathbb{B}^{\mu,j}(z) := \sum_{\sigma=-1}^0 z^{-\sigma-1} B^{\sigma,j,\mu}.$$

Ainsi le système mi-résolvant s'écrit sous la forme :

$$\begin{cases} \sum_{\mu=-\ell}^r \mathbb{A}^\mu(z) W_{j+\mu} = F_j & \text{pour } j \geq 1, \\ W_j - \sum_{\mu=0}^b \mathbb{B}^{\mu,j}(z) W_{1+\mu} = G_j & \text{pour } j \in \llbracket 1 - \ell, 0 \rrbracket. \end{cases} \quad (\text{III.13})$$

Un peu de calcul matriciel montre alors que le schéma précédent se réécrit sous la forme totalement résolvente :

$$\begin{cases} \mathscr{W}_j = \mathbb{M}(z) \mathscr{W}_j + \mathscr{F}_j & \text{pour } j \geq 1, \\ \mathbb{B}(z) \mathscr{W}_1 = \mathscr{G}, \end{cases} \quad (\text{III.14})$$

équation qui jouera dans ce qui suit le rôle de la relation de récurrence (III.4).

Dans (III.14), les nouveaux termes sources  $(\mathscr{F}_j)_{j \geq 1}$  et  $\mathscr{G} \in \mathbb{C}^{\ell N}$  dépendent respectivement (et explicitement) des sources  $(F_j)_{j \geq 1}$  et  $(G_j)_{1-\ell \leq j \leq 0}$  ; les matrices  $\mathbb{M}(z)$  et  $\mathbb{B}(z)$  sont explicitement calculables. On renvoie à [[Coulombel, 2012]-Section 3.3] pour une définition précise. Cette dernière ne sera toutefois pas capitale pour ce qui suit où l'on aura uniquement besoin de propriétés spectrales sur ces matrices (voir Propositions III.3.3-III.18).

La stabilité du schéma (III.9) se traduit de la façon suivante sur le schéma résolvant (III.14) :

**Proposition III.3.2** *Le schéma (III.9) est stable au sens de la Définition III.3.1, si et seulement si, le schéma résolvant (III.14) est stable au sens où il existe  $C > 0$ , telle que pour tout  $z \in \mathscr{U}$ , pour tout  $(\mathscr{F}_j)_{j \geq 1} \in \ell^2(\llbracket 1, \infty \rrbracket)$ , et pour tout  $\mathscr{G} \in \mathbb{C}^{\ell N}$ , le schéma (III.14) admet une unique solution  $(\mathscr{W}_j)_{j \geq 1} \in \ell^2(\llbracket 1, \infty \rrbracket)$  qui vérifie de plus l'estimation de stabilité :*

$$\frac{|z| - 1}{|z|} \sum_{j \geq 1} |\mathscr{W}_j|^2 + |\mathscr{W}_1|^2 \leq C \left( \frac{|z|}{|z| - 1} \sum_{j \geq 1} |\mathscr{F}_j|^2 + |\mathscr{G}|^2 \right). \quad (\text{III.15})$$

La question de la stabilité du schéma (III.9) se ramène donc à démontrer l'estimation de stabilité (III.15) pour le vecteur augmenté  $(\mathscr{W}_j)_{j \geq 1}$ .

On conclut ce paragraphe en donnant une condition nécessaire (dont un renforcement s'avérera aussi suffisante) pour démontrer la stabilité du schéma résolvant (III.14). Sans surprise, cette condition ressemblera fortement à la condition assurant le caractère fortement bien posé dans le cadre continu. On peut donc voir la condition suivante comme une version discrétisée de la condition continue.

La mise sous forme récursive (III.14) invite à poser

$$\forall j \geq 1, \mathscr{W}_j = \mathbb{M}^{j-1}(z) \mathscr{W}_1, \quad (\text{III.16})$$

équation sur laquelle on voit que le spectre de  $\mathbb{M}(z)$  va caractériser les solutions. On note dans ce qui suit  $\mathbb{E}^s(z)$  (resp.  $\mathbb{E}^u(z)$ ) le sous-espace propre associé aux valeurs propres de module (strictement) inférieur (resp. supérieur) à un. Nous avons alors l'équivalent d'un lemme de Hersh discret (voir [[Coulombel, 2012]-Lemma 3.7] ou [Kreiss, 1968])

**Proposition III.3.3** *[Propriétés spectrales de  $\mathbb{M}$ ] Pour tout  $z \in \mathscr{U}$ , la matrice  $\mathbb{M}(z)$  n'admet pas de valeur propre sur le cercle unité  $\mathbb{S}^1$ , elle en admet (comptées avec multiplicité)  $\ell N$  dans la boule (ouverte)  $B(0, 1)$  et  $rN$  dans  $\mathscr{U}$ . On a donc la décomposition*

$$\mathbb{C}^{(\ell+r)N} = \mathbb{E}^s(z) \oplus \mathbb{E}^u(z). \quad (\text{III.17})$$

Considérons le cas où les deux termes sources dans le schéma résolvant (III.14) s'annulent. La solution du schéma associée devrait alors être la solution nulle. Or, on voit que si l'on est en mesure de trouver un

élément non trivial  $\mathcal{W}_1 \in \mathbb{E}^s(z) \cap \ker \mathbb{B}(z)$ , alors la formule (III.16) nous donne une solution non triviale, dans  $\ell^2$ , par définition de  $\mathbb{E}^s(z)$ .

On a donc montré que pour être stable le schéma résolvant vérifiait nécessairement la condition de Kreiss-Lopatinskii discrète (faible):

**Proposition III.3.4 (Condition de Kreiss-Lopatinskii discrète)** *Si le schéma (III.9) est stable au sens de la Définition III.3.1, alors son schéma résolvant (III.14) vérifie la condition de Kreiss-Lopatinskii faible discrète :*

$$\forall z \in \mathcal{U}, \mathbb{E}^s(z) \cap \ker \mathbb{B}(z) = \{0\}. \quad (\text{III.18})$$

Comme dans le cas du demi-espace continu, la condition de Kreiss-Lopatinskii faible discrète ne suffit pas pour assurer la stabilité. Une fois de plus cela est essentiellement dû au manque d'uniformité par rapport à la variable de fréquence  $z$  qui est un obstacle pour revenir dans la variable temporelle  $n$ . Là encore, le moyen "le plus simple" d'obtenir cette uniformité est de compactifier l'espace des fréquences<sup>4</sup> en venant considérer les  $z \in \overline{\mathcal{U}}$ .

Mais alors, les propriétés spectrales décrites à la Proposition III.3.3 ne sont plus vérifiées, puisque certaines valeurs propres viennent "dégénérer" sur le cercle unité. Nous avons exactement le même phénomène dans le cas continu, mais pour la droite imaginaire. En particulier, les sous-espaces stables et instables ne sont pas trivialement bien définis pour  $z \in \overline{\mathcal{U}}$ .

On peut néanmoins montrer que ces derniers admettent un prolongement par continuité à  $z \in \overline{\mathcal{U}}$ . On renvoie par exemple à [[Coulombel, 2012]-Section 4.3], pour une démonstration rigoureuse. Une fois ce prolongement réalisé, on peut alors énoncer la condition qui caractérisera la stabilité du schéma (III.9), la condition de Kreiss-Lopatinskii uniforme discrète:

**Hypothèse III.3.1** [[B. Gustafsson and Sundstrom., 1972]] *On suppose que le schéma (III.9) vérifie la condition de Kreiss-Lopatinskii discrète uniforme, à savoir, que pour tout  $R \geq 2$ , pour tout  $z \in \overline{\mathcal{U}} \cap B(0, R)$ , on a*

$$\ker \mathbb{E}^s(z) \cap \mathbb{B}(z) = \{0\},$$

où  $\mathbb{B}$  désigne la condition de bord du problème résolvant (III.14) associé au schéma (III.9) et où  $\mathbb{E}^s$  désigne le prolongement par continuité<sup>5</sup>, à  $z \in \mathbb{S}^1$ , du sous-espace stable de (III.14).

Le théorème qui fonde l'étude de la stabilité des schémas aux différences finies dans des domaines à bords est alors dû au trio Gustafsson, Kreiss et Sundstrom [B. Gustafsson and Sundstrom., 1972].

**Théorème III.3.1** [[B. Gustafsson and Sundstrom., 1972],[Coulombel, 2011]] *Sous certaines hypothèses de structure<sup>6</sup> sur le schéma d'approximation (III.9) alors ce schéma est stable au sens de la Définition III.3.1, si et seulement, si il vérifie l'Hypothèse III.3.1.*

<sup>4</sup>On peut noter que dans la formulation (III.18), il est en fait suffisant que l'on ait cette propriété localement au voisinage du bord de  $\mathcal{U}$ .

<sup>5</sup>Ce sous-espace stable contient en particulier la contribution de valeurs propres de module un.

<sup>6</sup>Dans cet énoncé on se contentera de jeter un voile pudique sur ces hypothèses. Il est à noter que les hypothèses de [B. Gustafsson and Sundstrom., 1972] ne sont pas optimales par rapport à celles de [Coulombel, 2011]. Les hypothèses dont on a besoin pour ce théorème sont :

- La stabilité du problème de Cauchy sous-jacent à (III.9).
- Une propriété de non-caractérisité ( $\det \mathbb{A}_{-\ell}, \det \mathbb{A}_r \neq 0$ ).
- Des propriétés sur le comportement des valeurs propres de l'opérateur de discrétisation  $Q$ .

Les deux premiers points sont communs aux deux contributions. C'est sur le dernier point que le travail de Coulombel améliore grandement les hypothèses de [B. Gustafsson and Sundstrom., 1972] en se plaçant dans le cadre des opérateurs géométriquement réguliers au lieu d'opérateurs dissipatifs. Pour de tels opérateurs, les multiplicités des valeurs propres peuvent varier à condition que de telles variations viennent avec une base régulière de vecteurs propres. On renvoie à [[Coulombel, 2012]-Definition 2.3] pour une définition rigoureuse. Cette amélioration permet d'étendre l'analyse à des schémas bien connus dans la littérature comme par exemple le schéma saute-mouton.

Il est bon de préciser que la condition de Kreiss-Lopatinskii uniforme discrète est une condition qui est assez délicate à venir vérifier en pratique. Nous avons déjà du mal à vérifier la condition de Kreiss-Lopatinskii uniforme continue pour un opérateur et une condition de bord donnée. Ici, le problème est exactement de même nature, mais il est complexifié par le fait que l'on ne travaille plus dans l'espace  $\mathbb{C}^N$  mais dans l'espace  $\mathbb{C}^{(\ell+r)N}$ .

Par conséquent, même pour des schémas "raisonnables", avec de petits *stencils*, on pense typiquement au schéma de Lax-Friedrich, ou encore de Lax-Wendroff [Lax and Wendroff, 1964] pour lesquels  $\ell = r = 1$ , on a  $(\ell + r)N > N$ , le nombre de vecteurs propres à prolonger (à savoir  $\ell N$ ) est donc plus important dans  $\mathbb{E}^s(z)$  que dans son analogue continu. Par conséquent, on a souvent du mal, en dehors du cas scalaire, à mener à bien l'étude.

Cette difficulté intrinsèque a toutefois été surmontée (théoriquement et numériquement) dans la thèse de Le Barbenchon, déjà mentionnée au début de ce manuscrit. Dans cette thèse, la vérification de la condition de Kreiss-Lopatinskii uniforme discrète est ramenée à un calcul de nombre d'entrelacement.

Son analyse est toutefois restreinte au cadre unidimensionnel, principalement car une théorie de stabilité dans l'esprit du Théorème III.3.1 n'est pas totalement aboutie en dehors du cas unidimensionnel. Nous serions curieux et intéressé de voir cependant si, indépendamment de la théorie de stabilité sous-jacente, on ne pourrait pas venir étendre les techniques mises en œuvre par Le Barbenchon pour obtenir un algorithme de vérification de la condition de Kreiss-Lopatinskii uniforme discrète multidimensionnelle.

### III.3.4 Le "graal" de la stabilité

On conclut cette section introductive en décrivant la dernière notion de stabilité que nous avons rencontrée dans nos travaux. Cette dernière en donnant une estimation de la solution numérique en norme  $L_t^\infty(L_x^2)$  se trouve être le pendant discret du caractère fortement bien posé du semi-groupe.

Dans le cas de la demi-droite, les résultats les plus aboutis établissant cette stabilité sont les travaux de Coulombel [Coulombel, 2015] et [Coulombel, 2021], de Coulombel et Gloria [Coulombel and Gloria, 2011]. On citera aussi les résultats antérieurs de Wu [Wu, 1995]. On décrira, ici, plus avant les résultats de Coulombel et Gloria [Coulombel and Gloria, 2011]. Ces derniers s'appliquent sous l'hypothèse suivante :

**Hypothèse III.3.2** *On suppose que le schéma dans la droite réelle associé à (III.9) vérifie l'estimation*

$$\forall u \in \ell^2(\mathbb{Z}), \|Qu\|_{\ell^2(\mathbb{Z})} \leq \|u\|_{\ell^2(\mathbb{Z})}.$$

Cette hypothèse joue, en quelque sorte, le rôle de l'hypothèse de symétrie des coefficients dans le cas continu (au sens où elle permet de procéder à des intégrations par parties discrètes). Notons, au passage, que nous avons aussi rencontré cette hypothèse dans le cas du problème dans l'intervalle.

Le concept de stabilité associé est alors le suivant :

**Définition III.3.2** *Le schéma (III.9) est stable pour le semi-groupe si il est GKS-stable au sens de la Définition III.3.1 et si de plus on a la dernière estimation de stabilité : il existe  $C > 0$ , telle que, pour tout  $\Delta t \in ]0, 1]$ , on a l'inégalité*

$$\begin{aligned} \sup_{n \geq 0} e^{-2\gamma n \Delta t} \sum_{j \geq 1-\ell} \Delta x |U_j^n|^2 \leq C \left( \sum_{j \geq 1-\ell} \Delta x |u_{0,j}|^2 + \frac{\gamma \Delta t + 1}{\gamma} \sum_{n \geq 0} \sum_{j \geq 1} \Delta t \Delta x e^{-2\gamma(n+1)\Delta t} |F_j^n|^2 \right. \\ \left. + \sum_{n \geq 1} \sum_{j=1-\ell}^0 \Delta t e^{-2\gamma n \Delta t} |g_j^n|^2 \right). \end{aligned}$$

Le résultat principal de Coulombel et Gloria [Coulombel and Gloria, 2011] montre que l'Hypothèse III.3.2 est une condition suffisante pour la stabilité du semi-groupe au sens de la Définition III.3.2.

### III.4 Un premier pas vers un domaine de résolution borné, le problème dans le segment

Cette section expose les résultats de l'article [Benoit, 2022]. On s'intéressera dans la présente section à l'approximation numérique d'un problème posé dans l'intervalle. Le schéma dont on va étudier la stabilité s'écrit sous la forme suivante. On se donne une subdivision régulière de  $[0, 1]$  définie par  $x_j = \frac{j}{J}$ , pour  $J \geq 0$ ,  $j \in \llbracket 0, J \rrbracket$  correspondant au nombre de points dans la subdivision et on considère le schéma

$$\begin{cases} U_j^{n+1} = QU_j^n + \Delta t f_j^n & \text{pour } n \geq 0 \text{ et } j \in \llbracket 1, J \rrbracket, \\ U_j^{n+1} = \sum_{\sigma=0}^1 B_0^{\sigma,j} U_1^{n+\sigma} + g_{0,j}^{n+1} & \text{pour } n \geq 0 \text{ et } j \in \llbracket 1 - \ell, 0 \rrbracket, \\ U_j^{n+1} = \sum_{\sigma=0}^1 B_1^{\sigma,j} U_J^{n+\sigma} + g_{0,j}^{n+1} & \text{pour } n \geq 0 \text{ et } j \in \llbracket J + 1, J + r \rrbracket, \\ U_j^0 = u_{0,j} & \text{pour } j \in \llbracket 1 - \ell, J + r \rrbracket. \end{cases} \quad (\text{III.19})$$

Le but de ce qui suit est alors de comprendre quelle est l'influence du dialogue des deux conditions de bord numériques sur la stabilité du schéma.

#### III.4.1 Motivations

On décrit ici les motivations qui nous ont poussé à étudier le problème dans l'intervalle. Comme mentionné précédemment, l'une de nos principales motivations, bien qu'elle constitue un travail à longue échéance, est de pouvoir disposer de méthodes numériques convergentes dans des domaines bornés. Il est donc assez naturel de se demander si nous sommes en mesure de traiter le cas du borné le plus simple, l'intervalle.

En fait, comme déjà évoqué, l'absence de théorie de la stabilité totalement aboutie dans le cas d'un demi-espace multidimensionnel nous restreint fortement, pour l'instant, sur le choix du borné que l'on peut considérer. En un certain sens, cela nous oblige même à travailler dans l'intervalle.

Une autre motivation qui nous a aussi poussé à cette étude a été liée à la nature même de la théorie fortement bien posée que l'on cherche à discrétiser. Comme on l'a vu dans la Section II.2, concernant le problème continu dans la bande, ce dernier peut avoir deux types de solutions. Certaines présentent une croissance exponentielle de la solution par rapport à la variable temporelle, alors que pour les autres, cette croissance est exclue.

On s'est donc demandé si une telle dichotomie opérait aussi pour les schémas numériques. Pas pour le plaisir de travailler sur des problèmes discrets, mais pour des questions de pure modélisation. En effet, comme on l'a déjà indiqué les conditions de bords numériques proviennent de choix de la part de l'utilisateur.

Or, si l'on veut que le schéma approche bien la solution de l'équation aux dérivées partielles, il faut que la solution numérique ait le même comportement qualitatif par rapport au temps que la solution exacte. On doit donc pouvoir disposer d'un critère qui nous dit si telle condition de bord numérique donne un schéma dont la solution admet (ou non) une telle croissance exponentielle, afin d'être mieux à même de choisir une condition de bord idoine pour notre schéma numérique. C'est pourquoi, dans ce qui suit, nous nous sommes intéressé à une caractérisation complète des conditions de bord numériques qui empêchent une telle croissance exponentielle de la solution numérique par rapport au temps. De tels schémas seront alors parfaitement indiqués pour modéliser les problèmes continus étudiés dans la Section II.3.

#### III.4.2 Une discrétisation de l'analyse d'Osher

Ce paragraphe montre comment on peut venir adapter, pour obtenir une condition caractérisant la stabilité du schéma (III.19), l'analyse d'Osher mise en place dans le cadre continu à coin au cas discret dans l'intervalle.

Ce paragraphe est organisé de la façon suivante. Le premier paragraphe décrit succinctement les modifications à apporter au schéma (III.9) pour qu'il opère maintenant dans l'intervalle. Le second paragraphe expose la mise sous forme résolvente et la condition nécessaire de non croissance exponentielle de la solution. Enfin les deux derniers paragraphes se penchent sur son caractère suffisant. D'abord en introduisant le *symétriseur*, puis en esquissant sa construction dans le dernier paragraphe.

### Les modifications à apporter dans le cas de l'intervalle

On rappelle que l'intervalle d'étude,  $[0, 1]$  est discrétisé de façon régulière,  $x_j := \frac{j}{J}$ , avec  $J \gg 1$  et  $j \in \llbracket 0, J \rrbracket$ .

De nouvelles conditions numériques sur le bord droit apparaissent. L'apparition de ces dernières suit le même procédé et les mêmes raisons que ceux décrits dans le cas de la demi-droite. C'est pourquoi, on se permettra de ne pas les décrire en détails. On illustre ces conditions sur la Figure III.1. Plus précisément,

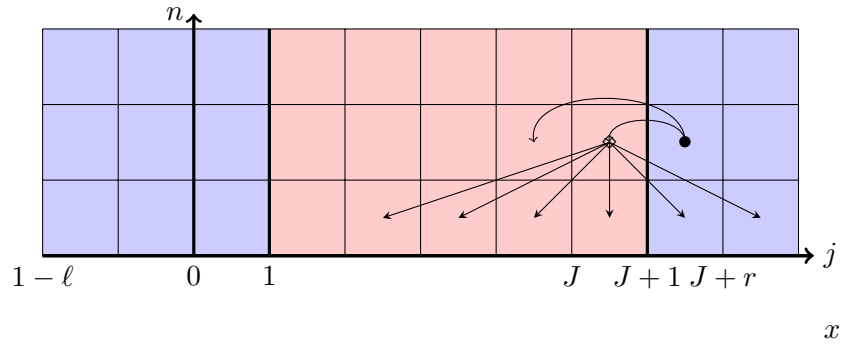


Figure III.1: Discretisation en espace-temps pour  $J = 5$ ,  $\ell = 3$  and  $r = 2$ . La zone rouge correspond à l'intérieur, les zones bleues aux frontières.

on précisera juste la forme des conditions de bord apparaissant dans (III.19). Dans les conditions de bord, les opérateurs aux bords se développent de la façon suivante

$$\sigma \in \{0, 1\}, j \in \llbracket 1 - \ell, 0 \rrbracket, B_0^{\sigma, j} := \sum_{\mu=0}^{b_0} B_0^{\sigma, j, \mu} \mathbf{T}^\mu \text{ et } j \in \llbracket J + 1, J + r \rrbracket, B_1^{\sigma, j} := \sum_{\mu=0}^{b_1} B_1^{\sigma, j, \mu} \mathbf{T}^{-\mu},$$

où les matrices  $B_0^{\sigma, j, \mu}$  et  $B_1^{\sigma, j, \mu}$  sont données dans  $\mathbf{M}_{N \times N}(\mathbb{R})$  et où  $b_0, b_1 \geq 0$  désignent respectivement les *stencils* de la condition de bord gauche et droite.

Les résultats de [Benoit, 2022] s'appliquent sous la condition, non nécessaire, mais simplificatrice et "non choquante" suivante sur les *stencils* des conditions de bord,  $b_0 < r$  et  $b_1 < \ell$ . Plus de commentaires seront faits dans ce qui suit sur cette hypothèse.

La définition de la stabilité a aussi besoin d'être modifiée de façon idoine. On demandera maintenant un contrôle des deux traces. Plus précisément, on utilisera la définition suivante de schéma stable :

#### Définition III.4.1 [Stabilité dans l'intervalle]

- On dit que le schéma dans l'intervalle (III.19) est stable au sens de GKS si pour tout  $J \geq 1$  et pour tous termes sources  $(f_j^n)_{n \geq 0, j \in \llbracket 1, J \rrbracket} \in \ell^2$ ,  $(g_{0, j}^n)_{n \geq 0, j \in \llbracket 1 - \ell, 0 \rrbracket} \in \ell^2$  et  $(g_{1, j}^n)_{n \geq 0, j \in \llbracket J + 1, J + r \rrbracket} \in \ell^2$ , l'unique solution  $(U_j^n)_{n \geq 0, j \in \llbracket 1 - \ell, J + r \rrbracket} \in \ell^2$  associée à la donnée initiale homogène vérifie l'estimation

de stabilité : il existe  $C > 0$ , telle que, pour tout  $\gamma > 0$ , on ait

$$\begin{aligned} & \frac{\gamma}{\gamma+1} \sum_{n \geq 0} e^{-2\gamma n} \sum_{j=1-\ell}^{J+r} |U_j^n|^2 + \sum_{n \geq 0} e^{-2\gamma n} \sum_{j=1-\ell}^r |U_j^n|^2 + \sum_{n \geq 0} e^{-2\gamma n} \sum_{j=J-\ell}^{J+r} |U_j^n|^2 \\ & \leq C \left( \frac{\gamma+1}{\gamma} \sum_{n \geq 0} e^{-2\gamma(n+1)} \sum_{j=1}^J |f_j^n|^2 + \sum_{n \geq 0} e^{-2\gamma n} \sum_{j=1-\ell}^0 |g_{0,j}^n|^2 + \sum_{n \geq 0} e^{-2\gamma n} \sum_{j=J+1}^{J+r} |g_{1,j}^n|^2 \right). \end{aligned} \quad (\text{III.20})$$

- On dit que le schéma dans l'intervalle (III.19) est stable au sens du semi-groupe (on peut alors avoir  $u_0 \neq 0$ ), si en plus de l'estimation de stabilité (III.20), on a aussi l'estimation :

$$\begin{aligned} \sup_{n \geq 0} e^{-2\gamma n} \sum_{j=1-\ell}^{J+r} |U_j^n|^2 & \leq C \left( \sum_{j=1-\ell}^{J+r} |u_{0,j}|^2 + \frac{\gamma+1}{\gamma} \sum_{n \geq 0} e^{-2\gamma(n+1)} \sum_{j=1}^J |f_j^n|^2 \right. \\ & \quad \left. + \sum_{n \geq 0} e^{-2\gamma n} \sum_{j=1-\ell}^0 |g_{0,j}^n|^2 + \sum_{n \geq 0} e^{-2\gamma n} \sum_{j=J+1}^{J+r} |g_{1,j}^n|^2 \right). \end{aligned}$$

On insiste ici sur le fait que dans les estimations précédentes, la constante  $C$  doit pouvoir être prise indépendante de  $J$ . En effet, à la fin des fins on veut vouloir prendre  $J \uparrow \infty$  pour discrétiser l'intervalle assez finement. Comme on va le voir, il est ici important de considérer pour les valeurs des traces, les traces étendues de Coulombel, à savoir celle sur  $\llbracket 1-\ell, r \rrbracket$  (resp.  $\llbracket J-\ell, J+r \rrbracket$ ) et non pas  $\llbracket 1-\ell, 0 \rrbracket$  (resp.  $\llbracket J+1, J+r \rrbracket$ ).

Puisque l'on ne permet pas à l'estimation de stabilité (III.20) d'être vraie seulement à partir d'un seuil  $\gamma_0 > 0$ , la solution du schéma ne peut pas développer de croissance exponentielle en temps. On est donc bien dans le cadre qui nous motivait dans le Paragraphe III.4.1.

Notre théorème final s'appliquera lorsque l'on se place sous l'Hypothèse III.3.2. Hypothèse donc légèrement plus contraignante que d'imposer uniquement la stabilité du schéma dans la droite réelle. On expliquera dans le cœur de la démonstration, l'endroit où l'on a été amené à faire cette hypothèse. Une fois que nous l'avons faite, il était alors tentant de l'utiliser pour obtenir la stabilité du semi-groupe en adaptant le travail de Coulombel & Gloria [Coulombel and Gloria, 2011].

C'est pourquoi, on arrivera *in fine* à un résultat de stabilité du semi-groupe. Dans ce qui suit on montre avec plus de détails comment établir la stabilité au sens de GKS (voir Définition III.4.1), ce qui constitue notre principale contribution. On n'indiquera pas comment on déduit de cette dernière, la stabilité du semi-groupe.

Il serait intéressant de trouver un moyen d'appliquer notre résultat sans l'Hypothèse III.3.2. On lui préférerait l'hypothèse de stabilité du schéma dans la droite réelle. Nous avons quelques idées pour débiter cette perspective. Nous les décrirons plus en détails au fil du texte.

## Le problème reformulé et la condition de stabilité

Dans tout ce qui suit, dans un souci d'allègement des notations on ne notera pas, sauf mention contraire, la variable fréquentielle  $z$ . On gardera toutefois à l'esprit que tous les objets définis dépendent de  $z$ .

On peut reproduire essentiellement les mêmes calculs que ceux décrits dans le cas de la demi-droite



(voir Paragraphe III.3.3). On se ramène alors, au problème sous forme totalement résolvente

$$\begin{cases} \mathscr{W}_{j+1} = \mathbb{M}\mathscr{W}_j & \text{pour } j \in \llbracket 1, J \rrbracket, \\ \mathbb{B}_0\mathscr{W}_1 = \mathscr{G}_0, \\ \mathbb{B}_1\mathscr{W}_{J+1} = \mathscr{G}_1, \end{cases} \quad (\text{III.21})$$

où l'on rappelle que l'on a posé  $\mathscr{W}_j := (W_{j+r-1}, \dots, W_{j-\ell})^T$  et  $U_j^n = z^n W_j$ ,  $z \in \mathbb{C} \setminus \{0\}$ . Le schéma (III.21) doit être comparé au schéma (III.14). Ce faisant, on s'aperçoit que le terme source à l'intérieur est nul dans (III.21), alors que l'on permettait un terme source non nul dans (III.14).

Ceci est dû au fait que l'étude de la *GKS*-stabilité du schéma dans l'intervalle requiert en premier lieu de se ramener à un terme source à l'intérieur homogène. Cela vient de la construction même du *symétriseur*. A l'heure actuelle, nous ne savons comment réaliser cette construction directement avec un terme source à l'intérieur non trivial. Ce point était déjà présent pour le problème continu pour lequel on se ramenait à un terme source homogène à l'intérieur en utilisant le caractère fortement bien posé des problèmes symétriques à conditions de bord strictement dissipatives.

Dans le cadre discret, on peut se ramener à un terme source nul à l'intérieur en montrant la stabilité (du semi-groupe) du schéma avec conditions de bords de Dirichlet sur les deux faces, *i.e.* :

$$\begin{cases} V_j^{n+1} = QV_j^n + \Delta t f_j^n & \text{pour } n \geq 0 \text{ et } j \in \llbracket 1, J \rrbracket, \\ V_j^{n+1} = 0 & \text{pour } n \geq 0 \text{ et } j \in \llbracket 1 - \ell, 0 \rrbracket, \\ V_j^{n+1} = 0 & \text{pour } n \geq 0 \text{ et } j \in \llbracket J + 1, J + r \rrbracket, \\ V_j^0 = 0 & \text{pour } j \in \llbracket 1 - \ell, J + r \rrbracket. \end{cases} \quad (\text{III.22})$$

Toutefois là où, dans le cadre continu, l'existence (automatique) de conditions de bords strictement dissipatives ne coûtait, pour ainsi dire, aucune hypothèse (sauf la symétrie des coefficients), dans le cas des schémas, on ne dispose pas d'une aussi bonne caractérisation de ce qui joue le rôle de conditions de bord dissipatives. Donc, pour montrer la stabilité de (III.22), on a ici utilisé le résultat de [Coulombel and Gloria, 2011] qui se base sur l'Hypothèse III.3.2.

Nous pensons intéressant de tâcher d'améliorer cette caractérisation des conditions numériques qui jouent le rôle des conditions de bord strictement dissipatives dans le cas continu. Ceci pourrait passer par une étude approfondie de l'article [Strang, 1969a] dans lequel Strang établit l'équivalence entre strictement dissipatif et vérifier la condition de Kreiss-Lopatinskii uniforme pour le problème continu à deux équations. Étendre le domaine d'application de nos résultats de stabilité dans le cas de l'intervalle serait alors une jolie application d'une telle caractérisation.

Pour conclure ce paragraphe on donne la condition qui permet d'assurer que la solution du schéma (III.19) ne développe pas une croissance exponentielle non triviale par rapport au temps. Comme on va le voir, cette analyse reprend, *mutatis mutandis*, celle que l'on a développée dans le cas du problème continu dans la bande.

On peut montrer que la stabilité du schéma (III.19) se ramène à étudier la stabilité du schéma résolvant (III.21). Plus précisément

**Proposition III.4.1** ([Benoit, 2022]) *Le schéma (III.19) est stable au sens de la Définition III.4.1, si et seulement si, le schéma résolvant (III.21) est stable au sens où pour tous termes sources  $\mathscr{G}_0 \in \mathbb{C}^{\ell N}$ ,  $\mathscr{G}_1 \in \mathbb{C}^{rN}$ , le schéma (III.21) admet une unique solution  $(\mathscr{W}_j)_{j \in \llbracket 1-\ell, J+r \rrbracket}$  qui vérifie de plus l'estimation ; il existe  $C > 0$ , telle que pour tout  $\gamma > 0$ , on ait l'inégalité*

$$|\mathscr{W}_1|^2 + |\mathscr{W}_J|^2 \leq C (|\mathscr{G}_0|^2 + |\mathscr{G}_1|^2). \quad (\text{III.23})$$

On suppose donc que le schéma (III.19) est stable au sens de la Définition III.4.1. Ainsi le schéma résolvant (III.21) admet une unique solution  $(\mathcal{W}_j)_{j \in \llbracket 1-\ell, J+r \rrbracket}$ . Soit  $(\widetilde{\mathcal{W}}_j)_{j \geq 1-\ell}$ , l'extension de  $(\mathcal{W}_j)_{j \in \llbracket 1-\ell, J+r \rrbracket}$  par zéro pour  $j > J+1$ . Un simple calcul montre que  $(\widetilde{\mathcal{W}}_j)_{j \geq 1-\ell}$  vérifie le schéma

$$\begin{cases} \mathcal{W}_{j+1} = \mathbb{M}\widetilde{\mathcal{W}}_j + \widetilde{\mathcal{F}}_j & \text{pour } j \geq 1, \\ \mathbb{B}_0(z)\mathcal{W}_1 = \mathcal{G}_0, \end{cases} \quad (\text{III.24})$$

où l'on a posé  $\widetilde{\mathcal{F}}_j := -\delta_{j, J+1}\mathbb{M}\mathcal{W}_j$ .

Par une formule de Duhamel discrète, on peut expliciter la solution de (III.24) en la décomposant, en utilisant la Proposition III.3.3, sous la forme  $\widetilde{\mathcal{W}}_j := \Pi^s \widetilde{\mathcal{W}}_j + \Pi^u \widetilde{\mathcal{W}}_j$ , où chacun des termes est explicitement donné par

$$\Pi^s \widetilde{\mathcal{W}}_j := \mathbb{M}^{j-1} \Pi^s \widetilde{\mathcal{W}}_1 + \sum_{k=1}^{j-1} \mathbb{M}^{j-1-k} \Pi^s \widetilde{\mathcal{F}}_k \quad \text{et} \quad \Pi^u \widetilde{\mathcal{W}}_j := - \sum_{k=j}^{\infty} \mathbb{M}^{j-1-k} \Pi^u \widetilde{\mathcal{F}}_k.$$

Si l'on suppose que la condition de Kreiss-Lopatinskii uniforme discrète est vérifiée pour le bord de gauche, on peut alors exprimer  $\mathcal{W}_1 = \widetilde{\mathcal{W}}_1 = \Phi_0 \mathcal{G}_0 - \Phi_0 \mathbb{B}_0 \Pi^u \widetilde{\mathcal{W}}_1$ , où  $\Phi_0$  désigne l'inverse de  $\mathbb{B}_0$ , restreinte à  $\mathbb{E}^s$ , donnée par la condition de Kreiss-Lopatinskii uniforme discrète. En utilisant les expressions explicites précédentes on obtient la donnée de la trace  $\mathcal{W}_1$  en fonction de la trace<sup>7</sup>  $\mathcal{W}_{J+1}$  :

$$\mathcal{W}_1 = \Phi_0 \mathcal{G}_0 + (I - \Phi_0 \mathbb{B}_0) \mathbb{M}^{-J} \Pi^u \mathcal{W}_{J+1}, \quad (\text{III.25})$$

où l'on a utilisé le fait que par définition de  $(\widetilde{\mathcal{F}}_k)_{k \geq 1}$ , ce terme s'annule sauf pour  $k = J+1$ .

On procède ensuite de même, mais en prolongeant par zéro pour  $j \leq 0$ . On obtient alors une relation donnant la trace de droite en fonction de la trace de gauche. Plus précisément

$$\mathcal{W}_{J+1} = (I - \Phi_1 \mathbb{B}_1) \mathbb{M}^J \Pi^s \mathcal{W}_1. \quad (\text{III.26})$$

Combinant les deux équations (III.25) et (III.26), on obtient alors la condition suivante sur la trace  $\mathcal{W}_1$  :

$$(I - (I - \Phi_0 \mathbb{B}_0) \mathbb{M}^{-J} \Pi^u (I - \Phi_1 \mathbb{B}_1) \mathbb{M}^J \Pi^s) \mathcal{W}_1 = \Phi_0 \mathcal{G}_0. \quad (\text{III.27})$$

Le membre de droite étant arbitraire, on s'attend donc à devoir imposer une condition d'inversibilité sur l'opérateur de gauche pour déterminer (de façon unique) la valeur de la trace  $\mathcal{W}_1$ . En fait, suivant des calculs algébriques exposés dans [Benoit, 2022], on peut affiner cette condition pour obtenir la condition suivante :

$$(I - \overline{\mathbb{T}}_{0 \rightarrow 0}(z)) \overline{\Pi}_{\ker \mathbb{B}_0} \mathcal{W}_1 := \mathcal{L} \mathcal{G}_0, \quad (\text{III.28})$$

où  $\mathcal{L}$  est un opérateur explicite à valeurs dans  $\ker \mathbb{B}_0$  et où l'on a posé

$$\overline{\mathbb{T}}_{0 \rightarrow 0} := \overline{\mathbb{T}}_{1 \rightarrow 0} \overline{\mathbb{T}}_{0 \rightarrow 1} \quad \text{avec} \quad \overline{\mathbb{T}}_{0 \rightarrow 1} := (I - \Phi_1 \mathbb{B}_1) \mathbb{M}^J \quad \text{et} \quad \overline{\mathbb{T}}_{1 \rightarrow 0} := (I - \Phi_0 \mathbb{B}_0) \mathbb{M}^{-J}.$$

On a donc obtenu la condition nécessaire pour la stabilité de (III.19) suivante :

**Proposition III.4.2 (Condition nécessaire de stabilité)** *Si le schéma (III.19) est stable au sens de la Définition III.4.1, alors, pour tout  $z \in \mathcal{U}$ , la restriction de l'opérateur  $I - \overline{\mathbb{T}}_{0 \rightarrow 0}(z)$  à  $\ker \mathbb{B}_0(z)$  est inversible.*

<sup>7</sup>C'est à cette étape qu'il est bon de travailler avec les traces étendues de Coulombel pour que la relation (III.25) fasse bien intervenir les mêmes valeurs de traces que celles dans le membre de gauche de (III.23).

Il est intéressant de remarquer ici que contrairement au cas continu nous n'avons qu'une seule condition. On rappelle que dans le cas continu, on imposait l'inversibilité de deux opérateurs  $I - \mathbf{T}_{0 \rightarrow 0}$  et  $I - \overline{\mathbf{T}}_{0 \rightarrow 0}$ . On se permet ici de rappeler que pour le problème continu les opérateurs de traces étaient définis par

$$\overline{\mathbf{T}}_{1 \rightarrow 0} = (\phi_0(\zeta)B_0 - I)e^{-\mathcal{A}(\zeta)} \text{ et } \overline{\mathbf{T}}_{0 \rightarrow 1} = (\phi_1(\zeta)B_1 - I)e^{\mathcal{A}(\zeta)},$$

et que l'on imposait l'inversibilité de  $I - \overline{\mathbf{T}}_{0 \rightarrow 0}$  sur  $\ker B_0$ . On voit donc bien sur ces formules que les opérateurs  $\overline{\mathbb{T}}_{0 \rightarrow 1}$  et  $\overline{\mathbb{T}}_{1 \rightarrow 0}$  sont des versions discrètes de leurs analogues continus.

Le fait que l'on soit dans un cadre discret explique aussi que pour étudier la stabilité de (III.19) nous n'ayons eu besoin de considérer qu'une seule condition d'inversibilité. En effet, la principale différence entre l'opérateur continu  $I - \overline{\mathbf{T}}_{0 \rightarrow 0}$  et son analogue discret  $I - \overline{\mathbb{T}}_{0 \rightarrow 0}$  est que l'opérateur continu agit sur un espace de dimension infinie (plus précisément  $L^2(\partial\Omega)$ ) là où l'opérateur discret agit sur un espace de dimension finie  $\mathbb{C}^{(\ell+r)N}$ . Dans le cas continu nous utilisons l'inverse de  $I - \overline{\mathbf{T}}_{0 \rightarrow 0}$  pour construire le *symétriseur* et l'inverse de  $I - \mathbb{T}_{0 \rightarrow 0}$  pour construire une solution faible.

Or, dans le cas de la dimension finie, l'existence d'une solution du schéma est automatique. Par conséquent nous n'avons pas ici besoin de condition d'inversibilité sur l'opérateur  $I - \mathbb{T}_{0 \rightarrow 0}$ . On pense que ce point est un fort argument en faveur du fait que les deux conditions d'inversibilités utilisées dans le cas continu sont en fait équivalentes. La démonstration d'une telle équivalence est toutefois laissée à de futurs travaux.

Dans la fin de ce paragraphe, nous allons montrer qu'une version renforcée de l'hypothèse nécessaire décrite à la Proposition III.4.2 est suffisante pour construire un *symétriseur*, et pour assurer la stabilité. Plus précisément, on se placera sous l'hypothèse suivante :

**Hypothèse III.4.1** *Pour tout  $z \in \mathcal{U}$ , l'opérateur  $I - \overline{\mathbb{T}}_{0 \rightarrow 0}(z)$  est uniformément inversible de  $\ker \mathbb{B}_0$  à valeurs dans lui-même. Plus précisément, il existe  $C > 0$ , telle que pour tout  $R > 0$ , pour tout  $z \in \mathcal{U} \cap B(0, R)$ , et pour tout  $\mathcal{V} \in \ker \mathbb{B}_0(z)$  on ait*

$$|\mathcal{V}| \leq C|(I - \overline{\mathbb{T}}_{0 \rightarrow 0}(z))\mathcal{V}|,$$

où l'on insiste sur le fait que la constante  $C > 0$  est indépendante de  $z$  et de  $J$ .

### Un *symétriseur* de type Osher discret

Comme dans le cas de la demi-droite, le *symétriseur* discret utilisé pour démontrer l'estimation de stabilité (III.23), ressemble beaucoup à celui qui a été construit dans le cadre continu pour étudier le caractère fortement bien posé (voir Définition II.3.2). Plus précisément on définit :

**Définition III.4.2** [*Symétriseur discret dans l'intervalle*] *On dit que  $(\mathbb{S}_j(z))_{j \in \llbracket 1, J+1 \rrbracket}$  est un *symétriseur* pour (III.21), si pour tout  $j \in \llbracket 1, J+1 \rrbracket$ , pour tout  $z \in \mathcal{U}$ , la matrice  $\mathbb{S}_j(z) \in \mathbf{M}_{(\ell+r)N \times (\ell+r)N}(\mathbb{C})$  vérifie les assertions suivantes :*

- i) *Pour tout  $j \in \llbracket 1, J+1 \rrbracket$ , pour tout  $z \in \mathcal{U}$ , on a  $\mathbb{S}_j^* = \mathbb{S}_j$ .*
- ii) *Pour tout  $R \geq 2$ , il existe  $C_{0,R}, C_{1,R} > 0$ , telles que, pour tout  $z \in \mathcal{U}$ , vérifiant  $|z| \leq R$ , et pour tout vecteurs  $\mathcal{V} \in \mathbb{E}^s$ ,  $\mathcal{W} \in \mathbb{E}^u(z)$ , on a les estimations*

$$|\mathbb{S}_1(z)\mathcal{V}| \leq C_{0,R}|\mathcal{V}| \text{ et } |\mathbb{S}_{J+1}(z)\mathcal{V}| \leq C_{1,R}|\mathcal{W}|.$$

*On insiste sur le fait que les constantes  $C_{0,R}, C_{1,R}$  peuvent dépendre de  $R$ , mais pas de  $z$  ou de  $J$ .*

- iii) *Pour tout  $z \in \mathcal{U}$ , et pour tout  $j \in \llbracket 1, J \rrbracket$ , la matrice  $\mathbb{S}_j(z)$  vérifie*

$$\mathbb{S}_j(z) = \mathbb{M}^*(z)\mathbb{S}_{j+1}(z)\mathbb{M}(z). \tag{III.29}$$

*iv)* Il existe  $\varepsilon_0 > 0$ , telle que pour tout  $R \geq 2$ , pour tout  $z \in \mathcal{U}$ , tels que  $|z| \leq R$  et enfin pour tout  $\mathcal{V} \in \ker \mathbb{B}_0(z)$ , on a

$$\mathcal{V}^* \mathbb{S}_1(z) \mathcal{V} \leq -\varepsilon_0 |\mathcal{V}|^2. \quad (\text{III.30})$$

*v)* Il existe  $\varepsilon_1 > 0$ , telle que pour tout  $R \geq 2$ , pour tout  $z \in \mathcal{U}$ , tels que  $|z| \leq R$  et enfin pour tout  $\mathcal{W} \in \ker \mathbb{B}_1(z)$ , on a

$$\mathcal{W}^* \mathbb{S}_{J+1}(z) \mathcal{W} \geq \varepsilon_1 |\mathcal{W}|^2. \quad (\text{III.31})$$

Discutons un peu les multiples demandes que l'on fait sur le *symétriseur* discret. Là encore elles sont à comparer avec celles faites sur le *symétriseur* continu (voir Définition II.3.2).

- Les points *iv)* et *v)* de la Définition III.4.2 assurent les contrôles idoines des traces. Ils sont à comparer avec les points *iv)* et *v)* de la Définition continue II.3.2. Notons ici, que les équations (III.30) et (III.31) sont ici des versions réduites des versions continues dans lesquelles les matrices de bords apparaissent. Ces versions réduites seront, toutefois, suffisantes pour démontrer l'estimation de stabilité (III.23) en décomposant, par la condition de Kreiss-Lopatinskii uniforme discrète,  $\mathcal{W}_1 := \mathcal{W}_1^b + \mathcal{W}_1^h$  où  $\mathcal{W}_1^b \in \ker \mathbb{B}_0$  est contrôlable par *iv)* et  $\mathcal{W}_1^h \in \mathbb{E}^s$  vérifie  $\mathbb{B}_0 \mathcal{W}_1^h = \mathcal{G}_0$ .
- Le point *iii)* assure qu'à l'intérieur le *symétriseur* vérifie une équation fonctionnelle, ce qui permet d'assurer que l'énergie interne est contrôlée. Là encore, nous avons la même demande pour le *symétriseur* continu. Dans le cas discret, l'équation différentielle du cas continu est remplacée par une équation aux différences.
- Le point *ii)* demande des propriétés de bornes sur le *symétriseur* discret. Ce dernier point peut sembler trivial car on est, ici, en dimension finie. Donc le *symétriseur* discret est matriciel. Toutefois, le point important est ici le caractère uniforme de ces bornes par rapport à la variable de fréquence et par rapport au nombre de discrétisation  $J$ . De telles bornes étaient aussi requises dans le cadre continu. On renvoie au point *ii)* de la Définition II.3.2.

Avec un *symétriseur* au sens de la Définition III.4.2 en main, la démonstration de l'estimation de stabilité au sens de la Définition III.4.1 est alors un simple bilan d'énergie discret laissé à titre d'exercice (on peut voir [[Benoit, 2022]-Proposition 5.2]). Par conséquent, pour établir la stabilité, il reste uniquement à construire un *symétriseur* discret au sens de la Définition III.4.2. C'est le sujet du prochain paragraphe.

### Quelques mots sur la construction d'un *symétriseur* discret

Dans tout ce mémoire, nous avons jusqu'à présent été assez avare en détails et laconique sur la construction des différents *symétriseurs*. Objets qui sont pourtant centraux dans la théorie. Dans ce paragraphe, on répare cet "oubli" en donnant les éléments clés de la construction d'un *symétriseur* au sens de la Définition III.4.2. Cette construction suit de près les idées de construction développées par Osher [Osher, 1973] pour traiter le problème continu dans le quart d'espace.

Toutefois, là où, dans le cas du quart d'espace on n'arrive pas, à l'heure actuelle, à démontrer le caractère uniformément borné du *symétriseur* (voir *ii)* de la Définition II.4.1), ce qui assurerait pourtant une estimation d'énergie sans pertes (voir [[Benoit, 2015]-Chapitre 5]), dans le cas de l'intervalle ou de la bande, le fait que l'on soit en dimension finie simplifie ce point et on est à même de conclure.

On va chercher un *symétriseur*  $\mathbb{S}_j$ , au sens de la Définition III.4.2, sous la forme

$$\mathbb{S}_j := \mathbb{S}_j(z) = \mathbb{J}_j(z) \mathbb{N}(z) \mathbb{J}_j^*(z), \quad (\text{III.32})$$

où l'on insiste sur le fait que  $\mathbb{N} = \mathbb{N}^*$  va être un poids, indépendant de  $j$ , choisi de telle manière que l'on ait *iv)* et *v)* de la Définition III.4.2. En d'autres termes, puisque toute la dépendance en  $j$  de  $\mathbb{S}_j$  se fait *via* le terme  $\mathbb{J}_j$ , c'est ce dernier qui sera choisi de manière à assurer le point *iii)*.

Un choix de candidat naturel pour la matrice  $\mathbb{J}_j$  fait intervenir les solutions du schéma dual du schéma (III.19). On introduit donc la notion de schéma dual<sup>8</sup>.

**Définition III.4.3 (Schéma dual)** *On dit que le schéma*

$$\begin{cases} \mathbb{L}_{dua}\mathcal{V} = \widetilde{\mathcal{F}}_j & \text{pour } j \in \llbracket 1, J \rrbracket, \\ \mathbb{C}_0(z)\mathcal{V}_1 = \widetilde{\mathcal{G}}_0, \\ \mathbb{C}_1(z)\mathcal{V}_{J+1} = \widetilde{\mathcal{G}}_1, \end{cases}$$

où  $\mathbb{C}_0 \in \mathbf{M}_{rN \times (\ell+r)N}(\mathbb{C})$  et  $\mathbb{C}_1 \in \mathbf{M}_{\ell N \times (\ell+r)N}(\mathbb{C})$ , est un schéma dual pour le schéma (primal) (III.21), si il existe des matrices  $\mathbb{N}_0(z), \mathbb{M}_1(z) \in \mathbf{M}_{rN \times (\ell+r)N}(\mathbb{C})$  et  $\mathbb{N}_1(z), \mathbb{M}_0(z) \in \mathbf{M}_{rN \times (\ell+r)N}(\mathbb{C})$ , telles que l'on ait la formule de dualité discrète : pour tout  $(\mathcal{V}_j)_{j \in \llbracket 1, J \rrbracket}$  et  $(\mathcal{W}_j)_{j \in \llbracket 1, J \rrbracket}$ ,

$$\begin{aligned} \sum_{j=1}^J (\mathbb{L}_{pri}\mathcal{W}_j)^* \mathcal{V}_{j+1} - \sum_{j=1}^J \mathcal{W}_j \mathbb{L}_{dua}\mathcal{V}_j = & - (\mathcal{W}_1^* \mathbb{N}_0^* \mathbb{C}_0 \mathcal{V}_1 + \mathcal{W}_1^* \mathbb{B}_0^* \mathbb{M}_0 \mathcal{V}_1) \\ & + (\mathcal{W}_{J+1}^* \mathbb{N}_1^* \mathbb{C}_1 \mathcal{V}_{J+1} + \mathcal{W}_{J+1}^* \mathbb{B}_1^* \mathbb{M}_1 \mathcal{V}_{J+1}). \end{aligned}$$

Suivant [[Benoit, 2022]-Proposition 5.3], on peut montrer qu'il existe un schéma dual associé à (III.21). Son équation à l'intérieur est alors caractérisée par

$$\mathbb{L}_{dua}\mathcal{V}_j := \mathcal{V}_j - \mathbb{M}^* \mathcal{V}_{j+1}.$$

On va donc chercher  $\mathbb{J}_j$  comme vérifiant

$$\mathbb{J}_j = \mathbb{M}^* \mathbb{J}_{j+1} \text{ en d'autres termes } \mathbb{J}_j \mathcal{V}_j \text{ est solution de l'équation duale homogène.}$$

Il est alors clair qu'avec un tel choix de  $\mathbb{J}_j$  un *symétriseur* sous la forme (III.32) vérifie le point *iii*) de la Définition III.4.2. Notons qu'à ce stade on a juste imposé que  $\mathbb{J}_j$  soit déterminé par l'équation duale intérieure mais que l'on n'a pas imposé de conditions de bord. Dans ce qui suit, on va utiliser ce degré de liberté pour décomposer  $\mathbb{J}_j$  sous la forme

$$\mathbb{J}_j := \mathbb{J}_{0,j} + \mathbb{J}_{1,j},$$

où l'on se permet de spécifier la valeur de la trace de  $\mathbb{J}_{0,j}$  (resp.  $\mathbb{J}_{1,j}$ ) à gauche (resp. à droite). Plus précisément, on définit les deux matrices précédentes comme les deux solutions des problèmes :

$$\begin{cases} \mathbb{J}_{0,j} = \mathbb{M}^* \mathbb{J}_{0,j+1} & \text{pour } j \geq 1, \\ \mathbb{J}_{0,1} = \widetilde{\Phi}_0 \mathbb{C}_0, \end{cases} \quad \text{et} \quad \begin{cases} \mathbb{J}_{1,j} = \mathbb{M}^* \mathbb{J}_{1,j+1} & \text{pour } j \leq J, \\ \mathbb{J}_{1,J+1} = \widetilde{\Phi}_1 \mathbb{C}_1, \end{cases} \quad (\text{III.33})$$

où l'on désigne par  $\widetilde{\Phi}_0$  et  $\widetilde{\Phi}_1$ , les inverses données par la condition de Kreiss-Lopatinskii uniforme discrète du problème dual. On montre en effet assez facilement (voir [[Benoit, 2022]-Proposition 5.5]) que, comme dans le cas du problème continu, cette condition passe bien à l'adjoint.

En d'autres termes,  $\mathbb{J}_{0,j}$  (resp.  $\mathbb{J}_{1,j}$ ) correspond à l'évolution libre de la solution du problème dual sur lequel on spécifie la valeur de trace à gauche (resp. droite) que l'on voudrait voir satisfaite.

On peut résoudre le duo de problèmes (III.33) explicitement et on obtient :

$$\forall j \geq 1, \mathbb{J}_{0,1} = \mathbb{M}^{-(j-1)*} \widetilde{\Phi}_0 \mathbb{C}_0 \text{ et } \forall j \leq J, \mathbb{J}_{1,j} = \mathbb{M}^{(J+1-j)*} \widetilde{\Phi}_1 \mathbb{C}_1. \quad (\text{III.34})$$

<sup>8</sup>Les problèmes duaux sont classiquement utilisés dans le cadre continu pour montrer l'existence de solutions. Dans le cas discret, cette existence étant automatique l'intérêt intrinsèque de tels objets n'est pas si clair.

Ce choix étant fait, on a alors fixé la dynamique en  $j$  dans le *symétriseur*. Il reste à choisir le poids  $\mathbb{N}$  pour assurer les points  $iv)$  et  $v)$  de la Définition III.4.2. On explicite ces conditions pour un *symétriseur* s'écrivant sous la forme (III.32). Elles deviennent

$$\begin{cases} [(\mathbb{J}_{0,1}^* + \mathbb{J}_{1,1}^*)\mathcal{V}]^* \mathbb{N} (\mathbb{J}_{0,1}^* + \mathbb{J}_{1,1}^*) \mathcal{V} \leq -\varepsilon_0 |\mathcal{V}|^2 & \forall \mathcal{V} \in \ker \mathbb{B}_0, \\ -[(\mathbb{J}_{0,J+1}^* + \mathbb{J}_{1,J+1}^*)\mathcal{W}]^* \mathbb{N} (\mathbb{J}_{0,J+1}^* + \mathbb{J}_{1,J+1}^*) \mathcal{W} \leq -\varepsilon_1 |\mathcal{W}|^2 & \forall \mathcal{W} \in \ker \mathbb{B}_1, \end{cases} \quad (\text{III.35})$$

équations sur lesquelles on voit que le choix de  $\mathbb{N}$  sera conditionné par les valeurs des traces  $\mathbb{J}_{0,\cdot}^* + \mathbb{J}_{1,\cdot}^*$ . En utilisant les formules explicites (III.34), on peut calculer ces valeurs et on obtient sans grande surprise<sup>9</sup>

$$\mathbb{J}_{0,1}^* + \mathbb{J}_{1,1}^* = (I - \Phi_0 \mathbb{B}_0) + \bar{\mathbb{T}}_{0 \rightarrow 1} \text{ et } \mathbb{J}_{0,J+1}^* + \mathbb{J}_{1,J+1}^* = \bar{\mathbb{T}}_{1 \rightarrow 0} + (I - \Phi_1 \mathbb{B}_1).$$

Ces deux résultats utilisent principalement les relations

$$(\tilde{\Phi}_0 \mathbb{C}_0)^* = I - \Phi_0 \mathbb{B}_0 \text{ et } (\tilde{\Phi}_1 \mathbb{C}_1)^* = I - \Phi_1 \mathbb{B}_1,$$

que l'on admettra ici. On renvoie à [[Benoit, 2022]-Proposition 5.6] pour une démonstration.

Dans ce qui suit, le but ne sera pas de donner explicitement un choix du poids  $\mathbb{N}$  assurant que (III.35) soit vérifiée. On renvoie à [[Benoit, 2022]-Equation (66)] pour une définition précise. De la même façon, on ne démontrera pas qu'un *symétriseur* de la forme (III.32) avec  $\mathbb{J}_j$  et  $\mathbb{N}$  définis comme fait précédemment vérifie le point  $ii)$  de la Définition III.4.2.

On pense qu'il est ici important et intéressant de pousser la construction jusqu'à l'apparition de l'opérateur discret  $(I - \bar{\mathbb{T}}_{0 \rightarrow 0})^{-1}$ . En particulier, comme on essaiera de le faire sentir, le choix de  $\mathbb{N}$  donné dans ce qui suit tient plus du "tour de force brute" que de "l'analyse fine et raffinée". Il y a probablement une façon plus astucieuse de choisir ce poids. Une telle construction améliorée pourrait conduire à un *symétriseur* avec une expression plus maniable et symétrique.

Qu'il s'agisse du problème dans la bande, ou bien du problème discret dans l'intervalle, cette expression raffinée serait plus intéressante, et ce pas seulement dans un souci de propreté esthétique de la démonstration. En effet, même avec les expressions actuelles, on a réussi à démontrer le point  $ii)$  assurant la propriété de borne uniforme. Toutefois, il se peut que dans le cas du quart d'espace (continu) une expression plus maniable du *symétriseur* soit utile pour finir la construction en établissant le point  $ii)$  manquant.

Revenons en à la construction du *symétriseur* discret. On va chercher le poids  $\mathbb{N}$  sous forme normale

$$\mathbb{N} := \Pi_{\ker \mathbb{B}_0}^* \mathbb{N}_0 \Pi_{\ker \mathbb{B}_0} + \Pi_{\ker \mathbb{B}_1}^* \mathbb{N}_1 \Pi_{\ker \mathbb{B}_1} + 2\text{Re}(\Pi_{\ker \mathbb{B}_0}^* \mathbb{N}_{01} \Pi_{\ker \mathbb{B}_1}). \quad (\text{III.36})$$

On note que pour ce faire, on utilise l'hypothèse tacite de [Osher, 1973] que les deux noyaux des conditions de bord sont supplémentaires. Plus précisément, dans le cas discret, cette hypothèse s'écrit :

**Hypothèse III.4.2** *On suppose que pour tout  $z \in \mathcal{U}$ , on a la somme directe*

$$\mathbb{C}^{(\ell+r)N} = \ker \mathbb{B}_0(z) \oplus \mathbb{B}_1(z),$$

où l'on rappelle que les matrices  $\mathbb{B}_0(z)$  et  $\mathbb{B}_1(z)$  encodent les conditions de bord du schéma mis sous forme résolvante (III.21). On note alors  $\Pi_{\ker \mathbb{B}_0}$  et  $\Pi_{\ker \mathbb{B}_1}$  les deux projecteurs associés à la décomposition précédente.

<sup>9</sup>On a vu que  $\mathbb{J}_{1,1}$  correspondait à la trace à gauche de l'évolution libre par le problème dual d'une trace à droite. Ainsi, ce terme joue le rôle d'un opérateur  $\bar{\mathbb{T}}_{1 \rightarrow 0}$  pour le problème dual qui, rappelons le, échange les conditions de bord. Il n'est donc pas vraiment étonnant que le dual de  $\mathbb{J}_{1,1}$  corresponde à la trace à droite de l'évolution d'une trace à gauche pour l'équation primale. On retrouve donc  $\bar{\mathbb{T}}_{0 \rightarrow 1}$  tel qu'il a été défini.

Le système d'inégalité (III.35) se réécrit donc en utilisant la forme normale (III.36) et les expressions explicites précédentes comme

$$\begin{cases} \mathcal{V}^* \mathbb{N}_0 \mathcal{V} + \mathcal{V}^* \bar{\mathbb{T}}_{0 \rightarrow 1}^* \mathbb{N}_1 \bar{\mathbb{T}}_{0 \rightarrow 1} \mathcal{V} + 2\mathcal{V}^* \operatorname{Re}(\mathbb{N}_{01} \bar{\mathbb{T}}_{0 \rightarrow 1}) \mathcal{V} \leq -\varepsilon_0 |\mathcal{V}|^2 & \forall \mathcal{V} \in \ker \mathbb{B}_0, \\ \mathcal{W}^* \bar{\mathbb{T}}_{1 \rightarrow 0}^* \mathbb{N}_0 \bar{\mathbb{T}}_{1 \rightarrow 0} \mathcal{W} - \mathcal{W}^* \mathbb{N}_1 \mathcal{W} - 2\mathcal{W}^* \operatorname{Re}(\mathbb{N}_{01}^* \bar{\mathbb{T}}_{1 \rightarrow 0}) \mathcal{W} \leq -\varepsilon_1 |\mathcal{W}|^2 & \forall \mathcal{W} \in \ker \mathbb{B}_1. \end{cases} \quad (\text{III.37})$$

On va chercher à résoudre les inégalités précédentes dans tout l'espace et non pas seulement sur les noyaux. C'est la première perte de précision dans l'analyse. On va donc chercher à résoudre les inégalités matricielles (ou dans un souci d'allègement des notations on a posé  $\varepsilon_0 = \varepsilon_1 = 1$ )

$$\begin{cases} \mathbb{N}_0 + \bar{\mathbb{T}}_{0 \rightarrow 1}^* \mathbb{N}_1 \bar{\mathbb{T}}_{0 \rightarrow 1} + 2\operatorname{Re}(\mathbb{N}_{01} \bar{\mathbb{T}}_{0 \rightarrow 1}) \leq -I, \\ \bar{\mathbb{T}}_{1 \rightarrow 0}^* \mathbb{N}_0 \bar{\mathbb{T}}_{1 \rightarrow 0} - \mathbb{N}_1 - 2\operatorname{Re}(\mathbb{N}_{01}^* \bar{\mathbb{T}}_{1 \rightarrow 0}) \leq -I. \end{cases} \quad (\text{III.38})$$

Le plan pour assurer les conditions (III.38) est alors le suivant :

1. Dans un premier temps on force l'apparition du terme  $I - \bar{\mathbb{T}}_{0 \rightarrow 0}$ . Ce point est jusqu'ici naturel, car notre hypothèse nécessaire pour la stabilité porte sur l'inverse de cet opérateur. Pour ce faire, on multiplie la seconde équation de (III.38) par  $\bar{\mathbb{T}}_{0 \rightarrow 1}^*$  à gauche et  $\bar{\mathbb{T}}_{0 \rightarrow 1}$  à droite. On somme avec la première équation de (III.38) pour obtenir :

$$\mathbb{N}_0 - \bar{\mathbb{T}}_{0 \rightarrow 0}^* \mathbb{N}_0 \bar{\mathbb{T}}_{0 \rightarrow 0} + 2\operatorname{Re}((I - \bar{\mathbb{T}}_{0 \rightarrow 0})^* \mathbb{N}_{01} \bar{\mathbb{T}}_{0 \rightarrow 1}) \leq -I - \bar{\mathbb{T}}_{0 \rightarrow 1}^* \bar{\mathbb{T}}_{0 \rightarrow 1}. \quad (\text{III.39})$$

2. On choisit alors les poids  $\mathbb{N}_0$  et  $\mathbb{N}_{01}$  de telle manière à avoir l'égalité dans (III.39). On note alors que, ce faisant, on perd un peu de généralité. Mais, plus "grave" de ce point de vue, le poids  $\mathbb{N}_0$  est choisi comme étant l'identité. On reviendra sur ce point dans la suite. De façon assez naturelle, le poids  $\mathbb{N}_{01}$  est choisi sous la forme  $\mathbb{N}_{01} := (I - \bar{\mathbb{T}}_{0 \rightarrow 0})^{-*} \mathbb{V}$  pour faciliter la résolution de (III.39) comme une égalité. On ne s'intéressera pas ici à la détermination rigoureuse de  $\mathbb{V}$ .
3. Une fois  $\mathbb{N}_0$  et  $\mathbb{N}_{01}$  déterminés, on revient au système (III.38). On vient choisir le poids  $\mathbb{N}_1$  de manière à ce que l'égalité soit vérifiée dans la seconde équation de (III.38). On perd ici encore un peu de généralité. Ce choix entraîne une expression "compliquée" de  $\mathbb{N}_1$  qui fait intervenir  $\mathbb{N}_{01}$ , et donc, en particulier l'inverse  $(I - \bar{\mathbb{T}}_{0 \rightarrow 0})^{-1}$ .
4. On vérifie alors par un calcul direct qu'avec un tel choix de  $\mathbb{N}_0$ ,  $\mathbb{N}_1$  et  $\mathbb{N}_{01}$ , la première équation de (III.38) est vérifiée.

La méthode précédemment décrite a un avantage certain, celui de donner un triplet de solutions du système (III.38). Toutefois, il est à noter qu'elle entraîne aussi une certaine disymétrie dans les coefficients. On rappelle que le poids  $\mathbb{N}$  est choisi sous la forme (III.36), équation dans laquelle on force un poids simple sur la partie de l'espace liée à  $\ker \mathbb{B}_0$ , et puis, ce choix étant fait, on fait "plus ou moins ce que l'on peut" pour obtenir un poids idoine sur la partie de l'espace liée à  $\ker \mathbb{B}_1$ .

Il nous semble que nous n'avons pourtant aucune raison de préférer la composante liée à  $\ker \mathbb{B}_0$ , en lui assurant un traitement simple à celle liée à  $\ker \mathbb{B}_1$ , où l'on impose un traitement compliqué.

Nous avons ce genre de traitements différents des composantes de l'espace dans la construction du *symétriseur* de Kreiss. En effet, considérons, par exemple un point elliptique  $\zeta \in \Xi_0$ , au voisinage duquel la propriété de Hersh sur le spectre reste vérifiée. La matrice résolvante s'écrit alors sous la forme (à changement de base près):

$$\mathcal{A}(\zeta) = \begin{bmatrix} \mathcal{A}_-(\zeta) & 0 \\ 0 & \mathcal{A}_+(\zeta) \end{bmatrix},$$

où l'on a  $\operatorname{Re}(sp(\mathcal{A}_\pm(\zeta))) \subset \pm [c, \infty[$ ,  $c > 0$ .

Par conséquent, si l'on veut assurer la négativité de  $\text{Re}(\mathcal{S}(\zeta)\mathcal{A}(\zeta))$ , où  $\mathcal{S}$  désigne le *symétriseur*, on peut venir normaliser la composante liée à  $\mathcal{A}_-(\zeta)$  (cette dernière fournissant forcément le bon signe) et laisser un degré de liberté sur celle liée à  $\mathcal{A}_+(\zeta)$ . Ainsi un choix naturel de *symétriseur* au voisinage d'un tel point est

$$\mathcal{S} := \begin{bmatrix} I & 0 \\ 0 & -\varrho I \end{bmatrix} \text{ où } \varrho > 0, \text{ est ensuite fixé de manière à avoir le contrôle au bord.}$$

Dans un tel cadre, le traitement différent des deux composantes de l'espace et, en particulier, la normalisation sont alors justifiés par des propriétés spectrales. Dans le cas du *symétriseur* discret (dans l'intervalle), nous ne comprenons pas pourquoi l'un des sous-espaces, en l'occurrence celui lié à  $\ker \mathbb{B}_0$ , devrait être normalisé plutôt que celui lié à  $\ker \mathbb{B}_1$ .

Nous pensons donc, qu'il serait préférable d'obtenir un poids de la forme (III.36) qui n'avantage pas l'un des deux sous-espaces, mais qui agit plutôt de façon comparable sur les deux sous-espaces. Un tel choix de *symétriseur* aurait probablement une expression plus claire, moins issue d'un "tour de force", sur laquelle il serait plus aisé de montrer que la propriété de borne uniforme (voir *ii*) de la Définition III.4.2) est satisfaite. Ceci pourrait être motivé (en plus de l'aspect purement esthétique) par le fait que dans le cas du quart d'espace, seule la propriété de borne uniforme du *symétriseur* est manquante.

Pour être totalement honnête, il nous faut avouer que nos tentatives pour obtenir un tel *symétriseur* à poids symétriques sont jusqu'ici restées lettres mortes. Toutefois, il faut quand même préciser que le système simplifié imposé sur les conditions de bord à savoir (III.38) est essentiellement un système de deux équations à trois inconnues. Par conséquent, on n'a pas de raison de penser que le choix de *symétriseur* donné par Osher dans [Osher, 1973] soit unique. Donc nous pensons qu'un *symétriseur* à poids symétriques existe (peut-être) quand même.

On conclut ce paragraphe par l'énoncé de caractérisation de la stabilité que nous avons obtenu dans [Benoit, 2022] :

**Théorème III.4.1 ([Benoit, 2022]-Theorem 3.1)** *Sous des hypothèses de structure<sup>10</sup> sur le schéma (III.19), on a alors les deux assertions suivantes :*

- *Si le schéma (III.19) est optimalement stable au sens de la Définition III.4.1 alors l'opérateur  $I - \overline{\mathbb{T}}_{0 \rightarrow 0}(z)$  (voir Proposition III.4.2) est inversible sur  $\ker \mathbb{B}_0(z)$ , pour tout  $z \in \mathcal{U}$ .*
- *Réciproquement, si l'on suppose que pour tout  $z \in \mathcal{U}$ , l'inverse de  $I - \overline{\mathbb{T}}_{0 \rightarrow 0}(z)$  est uniforme (voir Hypothèse III.4.1) alors le schéma (III.19) est optimalement stable au sens de la Définition III.4.1. Comme on a l'Hypothèse III.3.2, il est aussi automatiquement stable pour le semi-groupe au sens de la Définition III.3.2.*

---

<sup>10</sup>Plus précisément, on suppose que :

- Le schéma est géométriquement régulier.
- Les bords sont non caractéristiques.
- La condition de Kreiss-Lopatinskii uniforme discrète est vérifiée sur chaque face.
- Les *stencils* de bord vérifient  $b_0 < r$  et  $b_1 < \ell$ .
- On a l'hypothèse d'Osher tacite,  $\ker \mathbb{B}_0(z) \oplus \ker \mathbb{B}_1(z) = \mathbb{C}^{(\ell+r)N}$ .
- Enfin, on a l'hypothèse de stabilité pour le problème dans la droite réelle, Hypothèse III.3.2.

Les trois premières hypothèses sont donc parfaitement admissibles car elles sont déjà présentes et assez optimales dans la demi-droite. Les trois dernières hypothèses sont probablement améliorables. Elles ont déjà ou seront discutées dans la suite du texte.



## III.5 Vers un domaine de discrétisation borné, première étape le quart d'espace

Cette section évoque nos tentatives pour se rapprocher d'une théorie de la stabilité des schémas obtenus par approximations de types différences finies dans une géométrie bornée. En raison de la définition même des schémas de différences finies, la géométrie dans laquelle nous aimerions travailler, dans le futur, est typiquement celle d'un rectangle. Nous ne nous intéressons pas, pour l'instant, à des géométries bornées plus élaborées, pour lesquelles des schémas de type volumes ou éléments finis seraient tout aussi pertinents.

Il est à noter qu'à l'heure actuelle, nous sommes bien loin de pouvoir considérer la première géométrie bornée, le rectangle, nos résultats s'étant pour l'instant concentrés sur la géométrie du quart d'espace afin de mieux comprendre quelles peuvent être les interactions de deux conditions de bord numériques, plus particulièrement, au voisinage du coin.

Ceci semble être une première étape, plus ou moins incontournable, si nous voulons traiter le rectangle et ses quatre coins.

Les discussions de cette section se décomposent en deux parties :

- Les Paragraphes III.5.1 et III.5.2 décrivent des résultats que nous avons obtenus sur la stabilité dans la géométrie du quart d'espace. Ils traitent cependant de schémas particuliers, ou bien de conditions de bord numériques particulières. Ces études doivent être vues comme un premier pas vers une théorie plus générale.
- Le Paragraphe III.5.3 est de nature plus exploratoire. Il regroupe des pistes de recherche pour obtenir un résultat à propos des schémas et des conditions de bord numériques générales *via* une discrétisation de l'analyse d'Osher dans le quart d'espace [Osher, 1973].

### III.5.1 Stabilité du semi-groupe pour les conditions de Dirichlet

Ce paragraphe résume les résultats de l'article [Benoit, 2018]. Dans ce petit travail directement inspiré de [Coulombel and Gloria, 2011], on montre que, dans le cas du quart d'espace, si l'on dispose d'une bonne théorie de la stabilité (au sens GKS) on peut en déduire automatiquement la stabilité du semi-groupe, pour peu que l'on se place sous l'Hypothèse III.3.2.

Le résultat principal de [Benoit, 2018] s'applique à un schéma dans le quart d'espace de la forme<sup>11</sup> :

$$\begin{cases} U_j^{n+1} + QU_j^n = \Delta t f_j^{n+1} & \text{pour } j = (j_1, j_2, j') \in \mathcal{I} \times \mathbb{Z}^{d-2}, n \geq 0, \\ U_j^{n+1} + \sum_{\sigma=0}^1 B_1^{\sigma, j_1} U_{1, j_2}^{n+\sigma} = g_{1, j}^{n+1} & \text{pour } j \in \mathcal{B}_1 \times \mathbb{Z}^{d-2}, n \geq 0, \\ U_j^{n+1} + \sum_{\sigma=0}^1 B_2^{\sigma, j_2} U_{j_1, 1}^{n+\sigma} = g_{2, j}^{n+1} & \text{pour } j \in \mathcal{B}_2 \times \mathbb{Z}^{d-2}, n \geq 0, \\ U_j^{n+1} + \sum_{\sigma=0}^1 C^{\sigma, j} U_{1, 1}^{n+\sigma} = h_j^{n+1} & \text{pour } j \in \mathcal{C} \times \mathbb{Z}^{d-2}, n \geq 0, \\ U_j^0 = u_0 & \text{pour } j \in \llbracket 1 - \ell_1, \infty \rrbracket \times \llbracket 1 - \ell_2, \infty \rrbracket \times \mathbb{Z}^{d-2}, \end{cases} \quad (\text{III.40})$$

où l'on a posé

$$\begin{aligned} \mathcal{I} &:= \{(j_1, j_2) \in \mathbb{Z}^2 \setminus \{j_1, j_2 \geq 1\}\}, \\ \mathcal{C} &:= \{(j_1, j_2) \in \mathbb{Z}^2 \setminus \{j_1 \in \llbracket 1 - \ell_1, 0 \rrbracket \text{ et } j_2 \in \llbracket 1 - \ell_2, 0 \rrbracket\}\}, \\ \mathcal{B}_1 &:= \{(j_1, j_2) \in \mathbb{Z}^2 \setminus \{j_1 \in \llbracket 1 - \ell_1, 0 \rrbracket \text{ et } j_2 \geq 1\}\} \text{ et} \\ \mathcal{B}_2 &:= \{(j_1, j_2) \in \mathbb{Z}^2 \setminus \{j_2 \in \llbracket 1 - \ell_2, 0 \rrbracket \text{ et } j_1 \geq 1\}\}, \end{aligned}$$

<sup>11</sup>En fait dans [Benoit, 2018], encore en adaptant des travaux préexistants, ceux de [Coulombel, 2015] cette fois, on peut aussi considérer des schémas à plusieurs pas de temps, à condition que les schémas soient scalaires. On a choisi de ne pas décrire en détails ce résultat pour ne pas avoir à introduire un nouveau formalisme avec plusieurs pas de temps et pour garder les notations les plus consistantes possibles.

où  $l_1$  (resp.  $l_2$ ) désigne le *stencil* à gauche (resp. en bas) de l'opérateur de discrétisation interne. On renvoie à la Figure III.2 pour une illustration de la discrétisation.



Figure III.2: Une discrétisation du quart d'espace dans le cas  $\ell = r = 1$ .

Ainsi, pour des schémas dans le quart d'espace, on voit apparaître un nouveau type de mailles fantômes, celles au coin. De façon générique, elles sont indexées par les éléments de  $\mathcal{C}$ .

Dans ce qui suit, on ne décrira pas en détails les conditions numériques que l'on a imposées sur les mailles fantômes du coin. En effet, plusieurs possibilités s'offrent ici à nous. Par exemple, on peut imposer que les valeurs au coin soient calculées uniquement *via* les valeurs à l'intérieur ce qui découple, alors, la condition au coin des deux conditions de bord. Mais, on peut aussi imposer que les valeurs au coin soient calculées à partir de valeurs intérieures, mais aussi à partir de valeurs apparaissant dans  $\mathcal{B}_1$  et  $\mathcal{B}_2$ . Il faut alors calculer, d'abord les valeurs aux bords  $\mathcal{B}_1$  et  $\mathcal{B}_2$ , avant de se pencher sur la condition au coin. Une des contributions de [Benoit, 2018] est d'avoir formalisé les différents types de conditions de bord au coin que l'on peut être amené à rencontrer.

Le théorème principal de [Benoit, 2018] assure alors que si le problème (III.40) vient avec une bonne théorie de stabilité  $\ell^2$  au sens de GKS, alors, sous la même hypothèse que dans le cas du demi-espace sur l'opérateur de discrétisation interne  $Q$ , à savoir l'Hypothèse III.3.2, on peut automatiquement déduire de cette stabilité GKS, la stabilité du semi-groupe. On obtient alors la meilleure notion de stabilité possible.

L'analyse suit ici de près celle faite dans le cas du demi-espace dans [Coulombel and Gloria, 2011] qui vient elle-même généraliser celle de [Wu, 1995]. L'un des principaux ingrédients dans la démonstration est de démontrer au préalable la stabilité au sens du semi-groupe du problème avec conditions numériques

de bord de Dirichlet c'est-à-dire :

$$\begin{cases} V_j^{n+1} + QV_j^n = \Delta t f_j^{n+1} & \text{pour } j = (j_1, j_2, j') \times \mathcal{I} \times \mathbb{Z}^{d-2}, n \geq 0, \\ V_j^{n+1} = g_{1,j}^{n+1} & \text{pour } j \in \mathcal{B}_1 \times \mathbb{Z}^{d-2}, n \geq 0, \\ V_j^{n+1} = g_{2,j}^{n+1} & \text{pour } j \in \mathcal{B}_2 \times \mathbb{Z}^{d-2}, n \geq 0, \\ V_j^{n+1} = h_j^{n+1} & \text{pour } j \in \mathcal{C} \times \mathbb{Z}^{d-2}, n \geq 0, \\ V_j^0 = u_0 & j \in \llbracket 1 - \ell_1, \infty \rrbracket \times \llbracket 1 - \ell_2, \infty \rrbracket \times \mathbb{Z}^{d-2}. \end{cases} \quad (\text{III.41})$$

On obtient alors la stabilité du semi-groupe du schéma (III.40) en écrivant sa solution sous la forme  $U_j^n = V_j^n + W_j^n$ , où  $(V_j^n)$  est la solution du problème avec conditions de Dirichlet (III.41) et où  $(W_j^n)$  est solution d'un schéma de la même forme que (III.40), mais avec des termes sources idoines, et surtout avec une condition initiale homogène. On arrive alors à montrer la stabilité pour le semi-groupe de  $(U_j^n)$ , par analyse d'erreur, en utilisant celle de  $(V_j^n)$  et la stabilité au sens GKS de  $(W_j^n)$ .

On se permet ici de ne pas décrire précisément le résultat de [Benoit, 2018]. Un exposé détaillé en nécessite, en effet, de lourdes notations, peu intéressantes par rapport aux idées principales que l'on vient d'exposer. On renvoie le lecteur intéressé à [[Benoit, 2018]-Theorem 3.1] ou [Coulombel and Gloria, 2011].

A notre connaissance, l'exemple du schéma avec conditions de bord de Dirichlet constitue le premier exemple de schéma posé dans le quart d'espace qui soit stable (et ce même au sens du semi-groupe). Par conséquent, l'étude du schéma (III.41) qui est utilisée comme un lemme dans [Benoit, 2018] constitue tout de même un résultat intéressant pour lui-même. Ce résultat est toutefois à tempérer. En effet, il y a malheureusement peu de chances, qu'imposer la condition de Dirichlet sur les deux composantes du bord et au coin fournisse un résultat numériquement consistant. On pense, par exemple, à des phénomènes de transport sortants, phénomènes que nous étudions plus en détails dans le paragraphe suivant.

### III.5.2 Le schéma de Lax-Wendroff dans le quart d'espace

Ce paragraphe expose les résultats de l'article [Benoit and Coulombel, b], résultats ensuite affinés dans le *preprint* [Benoit and Coulombel, a]. Dans ces travaux, nous nous sommes intéressés à la stabilité du schéma de Lax-Wendroff [Lax and Wendroff, 1964] en deux dimensions d'espace. On s'est, plus précisément, efforcé d'obtenir une méthode simple, la moins technique possible, qui permette d'établir la stabilité du schéma de Lax-Wendroff, lorsque ce dernier est posé dans  $\mathbb{Z}^2$ ,  $\mathbb{N} \times \mathbb{Z}$ , ou enfin  $\mathbb{N}^2$ . Le schéma a vocation<sup>12</sup>, à approcher la solution de l'équation de transport sortant

$$\begin{cases} \partial_t u + a \partial_x u + b \partial_y u = 0 & t > 0, (x, y) \in \Gamma, \\ u|_{t=0} = u_0 & (x, y) \in \Gamma, \end{cases}$$

où  $\Gamma$  désigne selon le cas d'étude, l'espace  $\mathbb{R}^2$ , le demi-espace  $\mathbb{R}_+ \times \mathbb{R}$  (auquel cas on prescrit la condition de signe  $a < 0$ ), ou enfin le quart-d'espace  $\mathbb{R}_+^2$  (auquel cas on a  $a, b < 0$ ).

Le transport étant choisi sortant aucune condition de bord n'est à imposer dans l'équation de transport (III.5.2). Dans [Benoit and Coulombel, b] et [Benoit and Coulombel, a] respectivement, nous nous sommes intéressés à l'influence des conditions de bords de Neumann d'ordre un et deux respectivement

<sup>12</sup>On peut s'interroger sur une vocation consistant à modéliser numériquement une équation aux dérivées partielles pour laquelle on dispose d'une solution exacte explicite... L'intérêt est, en effet, assez discutable, mais s'il faut donner des motivations on se permet de donner les deux suivantes :

- Il faut voir le cas de l'équation de transport comme un "bébé-modèle" de système hyperbolique. Comprendre, précisément, les phénomènes de sortie du domaine pour l'équation de transport peut alors être vu comme un pré-requis avant de s'attaquer à de vrais systèmes.
- Et après tout, pourquoi pas ?

sur la stabilité de la discrétisation donnée par le schéma de Lax-Wendroff.

Avant de décrire plus avant les résultats obtenus, il est utile de préciser pourquoi nous avons étudié le schéma de Lax-Wendroff. Le schéma de Lax-Wendroff avec stabilisateur [Lax and Wendroff, 1964] que nous avons considéré est défini par :

$$\begin{aligned}
u_{j,k}^{n+1} = & u_{j,k}^n - \frac{\lambda a}{2}(u_{j+1,k}^n - u_{j-1,k}^n) - \frac{\mu b}{2}(u_{j,k+1}^n - u_{j,k-1}^n) \\
& + \frac{(\lambda a)^2}{2}(u_{j+1,k}^n - 2u_{j,k}^n + u_{j-1,k}^n) + \frac{(\mu b)^2}{2}(u_{j,k+1}^n - 2u_{j,k}^n + u_{j,k-1}^n) \\
& + \frac{\lambda a \mu b}{4}(u_{j+1,k+1}^n - u_{j+1,k-1}^n - u_{j-1,k+1}^n + u_{j-1,k-1}^n) \\
& - \frac{(\lambda a)^2 + (\mu b)^2}{8}(u_{j+1,k+1}^n - 2u_{j+1,k}^n + u_{j+1,k-1}^n \\
& \quad - 2u_{j,k+1}^n + 4u_{j,k}^n - 2u_{j,k-1}^n + u_{j-1,k+1}^n - 2u_{j-1,k}^n + u_{j-1,k-1}^n),
\end{aligned} \tag{III.42}$$

où l'on a posé  $\lambda := \frac{\Delta t}{\Delta x}$  et  $\mu := \frac{\Delta t}{\Delta y}$ , les deux ratios CFL (Courant-Friedrichs-Lewy) [Courant et al., 1928],  $\Delta t$ ,  $\Delta x$  et  $\Delta y$  désignent les pas de discrétisation en temps et dans les deux directions d'espace. En particulier, le schéma (III.42) s'écrit sous la forme (III.1) et on a un coefficient  $a_{-1,-1}$  non trivial lorsque l'on est en présence d'un maillage pour lequel  $\Delta x \neq \Delta y$ .

Ainsi, pour opérer dans le quart d'espace (qui a été notre objectif final de résolution), le schéma (III.42) nécessite :

- D'une part, l'ajout de conditions qui déterminent les mailles fantômes le long des bords  $\{-1\} \times \mathbb{N}$  et  $\mathbb{N} \times \{-1\}$ . Ceci est classique.
- D'autre part, l'ajout d'une condition déterminant la valeur de la maille fantôme du coin, à savoir  $(-1, -1)$ .

Le schéma de Lax-Wendroff a donc été choisi car il s'agit d'un exemple simple de schéma rencontré dans la littérature pour lequel une telle condition au coin apparaît. Cette condition est, par exemple, absente dans le schéma de Lax-Friedrichs.

Puisque l'on modélise un phénomène de sortie, il est alors assez naturel d'imposer les conditions de Neumann sur les bords. Dans [Benoit and Coulombel, b] nous nous sommes intéressés à une condition de Neumann d'ordre un, c'est-à-dire que l'on a imposé :

$$\begin{cases} u_{-1,k} = u_{0,k} & \forall k \in \mathbb{N}, \\ u_{j,-1} = u_{j,0} & \forall j \in \mathbb{N}. \end{cases} \tag{III.43}$$

Tandis que dans [Benoit and Coulombel, a] nous avons, pour obtenir une méthode d'ordre deux sur le bord (motivé par le fait que le schéma de Lax-Wendroff est d'ordre deux à l'intérieur), imposé des conditions de Neumann d'ordre deux :

$$\begin{cases} \Delta_1 u_{0,k} := u_{-1,k} - 2u_{0,k} + u_{1,k} = 0 & \forall k \in \mathbb{N}, \\ \Delta_2 u_{j,0} := u_{j,-1} - 2u_{j,0} + u_{j,1} = 0 & \forall j \in \mathbb{N}. \end{cases} \tag{III.44}$$

Dans les deux contributions [Benoit and Coulombel, b] et [Benoit and Coulombel, a] nous avons plus particulièrement cherché à décrire l'influence de la maille fantôme au coin (la maille  $(-1, -1)$ ) sur la stabilité globale de la méthode. Nous avons cherché à donner des conditions au coin explicites, pour lesquelles la stabilité du schéma peut-être obtenue. Ceci, de façon simple, sans utiliser l'analyse de Fourier.

Dans le cas de la condition de Neumann d'ordre un, il est alors moral de ne pas utiliser un *stencil* plus grand au coin que sur les bords. Nous pouvons donc choisir de définir  $u_{-1,-1}$ , grâce à l'un (ou bien grâce à une combinaison linéaire) de ses trois voisins  $u_{-1,0}$ ,  $u_{0,-1}$  ou  $u_{0,0}$ . Cependant, d'après les conditions de Neumann (III.43), ces trois voisins sont égaux. On a donc imposé la condition au coin :

$$u_{-1,-1} = u_{0,0}, \quad (\text{III.45})$$

qui est le choix probablement le plus naturel.

Dans [Benoit and Coulombel, b] nous avons aussi donné un exemple numérique montrant qu'imposer la condition "immorale"  $u_{-1,-1} = Cu_{0,0}$ , avec  $C > 0$  grand, conduit à l'instabilité de tout le schéma. Cela est cohérent avec l'intuition qu'un traitement "catastrophique" au coin peut venir détruire la stabilité globale du schéma.

Dans le cas de la condition de Neumann d'ordre deux, on peut alors considérer des conditions au coin avec un plus grand *stencil*. On a alors plus de liberté. Un choix qui nous a semblé naturel dans [Benoit and Coulombel, a] et qui conduit à la stabilité est de tâcher d'imposer la condition

$$\frac{1}{2}\Delta_1 u_{0,-1} + \frac{1}{2}\Delta_2 u_{-1,0} = 0,$$

de manière à ne pas privilégier l'une des directions d'espace par rapport à l'autre. En utilisant les conditions de bords (III.44), ce choix conduit alors à imposer

$$u_{-1,-1} = 4u_{0,0} - 2u_{0,1} - 2u_{1,0} + u_{1,1}. \quad (\text{III.46})$$

La stabilité du schéma de Lax-Wendroff dans le cas de l'espace entier est bien documentée et à trait à la théorie de von-Neumann. On obtient alors la stabilité, sous la condition CFL (optimale)

$$(a\lambda)^2 + (b\mu)^2 \leq \frac{1}{2}, \quad (\text{III.47})$$

où l'on rappelle que les paramètres  $a$  et  $b$  sont les vitesses du phénomène de transport et que  $\lambda$  et  $\mu$  désignent les ratios CFL. On peut, en utilisant la théorie GKS, voir par exemple [Coulombel, 2014], obtenir la même condition CFL assurant la stabilité du schéma dans le demi-espace, lorsque ce dernier est associé à la condition de Neumann d'ordre un.

Les deux méthodes précédentes se basent sur l'analyse de Fourier et viennent donc avec un certain niveau incompressible de technicité. De plus, l'analyse de Fourier est difficile à transposer dans le quart d'espace. En effet faire porter au temps le rôle d'une variable spatiale par transformée de Laplace n'est plus suffisant.

Dans [Benoit and Coulombel, b] et [Benoit and Coulombel, a] on a cherché une méthode, plus simple, conduisant à la stabilité de (III.42). On s'est donc appuyé sur une méthode d'énergie plus basique. Pour ce faire, nous aurons besoin d'introduire quelques notations usuelles. Dans ce qui suit, on notera

$$\begin{aligned} (D_{1,+}u)_{j,k} &= u_{j+1,k} - u_{j,k}, \quad (D_{1,-}u)_{j,k} = u_{j,k} - u_{j-1,k}, \quad D_{1,0} = \frac{1}{2}(D_{1,+} + D_{1,-}) \text{ et } \Delta_1 = D_{1,+}D_{1,-}. \\ (D_{2,+}u)_{j,k} &= u_{j,k+1} - u_{j,k}, \quad (D_{2,-}u)_{j,k} = u_{j,k} - u_{j,k-1}, \quad D_{2,0} = \frac{1}{2}(D_{2,+} + D_{2,-}) \text{ et } \Delta_2 = D_{2,+}D_{2,-}. \end{aligned}$$

Le résultat obtenu dans [Benoit and Coulombel, b] est alors le suivant. On donne ici son énoncé dans le cas peut-être le plus nouveau du quart d'espace :

**Théorème III.5.1** ([Benoit and Coulombel, b]-Theorem 4.1) *Soit  $a < 0$  et  $b < 0$  et soit  $\lambda$  et  $\mu$  vérifiant la condition CFL (III.47). Alors, le schéma de Lax-Wendroff (III.42) complété des conditions de bords de Neumann d'ordre un (III.43) et de la condition au coin (III.45), vérifie l'estimation de stabilité :*

$$\begin{aligned} \sup_{n \in \mathbb{N}} \|u^n\|_{\mathcal{H}}^2 + \frac{\lambda|a|}{8} \sum_{n \geq 0} \sum_{k \geq 0} (u_{0,k}^n)^2 + \frac{\mu|b|}{8} \sum_{n \geq 0} \sum_{j \geq 0} (u_{j,0}^n)^2 \\ + \frac{(\mu b)^2}{32} ((\lambda a)^2 + (\mu b)^2) \sum_{n \geq 0} \sum_{k \geq 0} (\Delta_2 u_{0,k}^n)^2 + \frac{(\lambda a)^2}{32} ((\lambda a)^2 + (\mu b)^2) \sum_{n \geq 0} \sum_{j \geq 0} (\Delta_1 u_{j,0}^n)^2 \leq 2\|u^0\|_{\mathcal{H}}^2, \end{aligned}$$

où  $(u^0) \in \ell^2(\llbracket -1, \infty \rrbracket^2)$  désigne la donnée initiale du schéma et où l'on a noté

$$\|u\|_{\mathcal{H}}^2 := \sum_{(j,k) \in \llbracket -1, \infty \rrbracket^2} u_{j,k}^2.$$

Le Théorème III.5.1 a aussi des énoncés analogues dans le cas du demi-espace et de l'espace entier, même si les résultats étaient déjà connus. Dans [Benoit and Coulombel, a] nous obtenons un énoncé assez analogue au Théorème III.5.1, mais dans le cas des conditions de bords de Neumann d'ordre deux (III.44) avec la condition au coin (III.46).

Toutefois, il est à pointer que dans [Benoit and Coulombel, a] comparé à [Benoit and Coulombel, b] nous n'arrivons pas à récupérer la condition de CFL optimale (III.47). En effet, pour boucler notre analyse nous devons nous restreindre à des paramètres CFL dans un secteur angulaire autour de la première bissectrice. Nous ne récupérons donc pas le quart de boule optimal de (III.47). Il est à noter que nous pensons que notre résultat n'est probablement pas optimal et qu'il pourrait peut-être être amélioré. Toutefois, comme on essayera de l'expliquer dans les éléments de preuve qui suivent, on espère faire sentir au lecteur qu'une telle amélioration n'est pas immédiate.

Dans ce qui suit, on va décrire les grandes idées de la démonstration du Théorème III.5.1 : la méthode d'énergie. Le schéma de démonstration étant indépendant de la géométrie, nous décrirons en détails les calculs lorsque l'on se place dans le cas, le plus simple, de l'espace entier  $\mathbb{Z}^2$ . Nous donnerons, dans un second temps, des éléments de démonstration qui indiquent quelles modifications sont à effectuer et quelles sont les difficultés qui apparaissent lorsque l'on se place dans le demi ou bien le quart d'espace.

## La démonstration dans $\mathbb{Z}^2$

Le premier point de la démonstration est d'écrire le schéma (III.42) sous la forme

$$u_{j,k}^{n+1} = u_{j,k}^n + v_{j,k}^n - w_{j,k}^n, \quad (\text{III.48})$$

où  $v_{j,k}^n$  (resp.  $w_{j,k}^n$ ) est la somme de tous les termes dans le membre de droite de (III.42) qui proviennent de l'application d'un opérateur anti-autoadjoint (resp. autoadjoint) à  $u^n$ .

On a alors le bilan d'énergie

$$\begin{aligned} \|u^{n+1}\|_{\ell^2(\mathbb{Z}^2)}^2 - \|u^n\|_{\ell^2(\mathbb{Z}^2)}^2 &= \langle v^n - w^n; 2u^n + v^n - w^n \rangle \\ &= \|v^n\|_{\ell^2(\mathbb{Z}^2)}^2 + \|w^n\|_{\ell^2(\mathbb{Z}^2)}^2 - 2\langle u^n; w^n \rangle + 2\langle u^n; v^n \rangle - 2\langle w^n; v^n \rangle. \end{aligned} \quad (\text{III.49})$$

On conclut alors en estimant tous les termes dans le membre de droite. Puisque  $v^n$  est associé à des termes anti-autoadjoints, le terme  $2\langle u^n; v^n \rangle$  apparaissant dans le membre de droite de (III.49) s'annule. Il en est de même du terme  $2\langle w^n; v^n \rangle$ . Le bilan d'énergie se simplifie, alors en

$$\|u^{n+1}\|_{\ell^2(\mathbb{Z}^2)}^2 - \|u^n\|_{\ell^2(\mathbb{Z}^2)}^2 = \|v^n\|_{\ell^2(\mathbb{Z}^2)}^2 - 2\langle u^n; w^n \rangle + \|w^n\|_{\ell^2(\mathbb{Z}^2)}^2$$

On conclut alors assez facilement en utilisant les deux lemmes suivants :

**Lemme III.5.1** Soit les trois suites  $(u_{j,k}^n)_{n \in \mathbb{N}, (j,k) \in \mathbb{Z}^2}$ ,  $(v_{j,k}^n)_{n \in \mathbb{N}, (j,k) \in \mathbb{Z}^2}$  et  $(w_{j,k}^n)_{n \in \mathbb{N}, (j,k) \in \mathbb{Z}^2}$  définies par la relation (III.48), alors on a l'égalité :

$$\begin{aligned} \|v^n\|_{\ell^2(\mathbb{Z}^2)}^2 - 2\langle u^n; w^n \rangle &= -\frac{(\lambda a)^2}{4} \|\Delta_1 u^n\|_{\ell^2(\mathbb{Z}^2)}^2 - \frac{(\mu b)^2}{4} \|\Delta_2 u^n\|_{\ell^2(\mathbb{Z}^2)}^2 \\ &\quad - \frac{1}{4} ((\lambda a)^2 + (\mu b)^2) \|D_{1,+} D_{2,+} u^n\|_{\ell^2(\mathbb{Z}^2)}^2. \end{aligned}$$

**Lemme III.5.2** Soit  $(w_{j,k}^n)_{n \in \mathbb{N}, (j,k) \in \mathbb{Z}^2}$  définie par la relation (III.48), alors on a l'estimation :

$$\begin{aligned} 4\|w^n\|_{\ell^2(\mathbb{Z}^2)}^2 &\leq 2((\lambda a)^2 + (\mu b)^2) \left[ (\lambda a)^2 \|\Delta_1 u^n\|_{\ell^2(\mathbb{Z}^2)}^2 + (\mu b)^2 \|\Delta_2 u^n\|_{\ell^2(\mathbb{Z}^2)}^2 \right. \\ &\quad \left. + ((\lambda a)^2 + (\mu b)^2) \|D_{1,+} D_{2,+} u^n\|_{\ell^2(\mathbb{Z}^2)}^2 \right]. \end{aligned}$$

Ainsi, on a bien la propriété de signe attendue si l'on impose la condition CFL (III.47) qui joue alors sur le terme bleu.

Le Lemme III.5.1 provient de calculs directs mais le Lemme III.5.2 demande lui des estimations un peu plus astucieuse si l'on veut à la fin obtenir la condition de stabilité optimale (III.47). On insistera, toutefois, sur le fait que le Lemme III.5.2 ne nécessite rien de plus technique, pour sa démonstration, que l'inégalité de Young. Ainsi la méthode d'énergie est beaucoup moins technique que celle basée sur l'analyse de Fourier. Mais, cette méthode se base sur le fait que l'on a des expressions explicites (et bien choisies) dans la discrétisation.

### Les modifications dans le cas du quart d'espace avec Neumann d'ordre un

Lorsque l'on n'est plus dans l'espace entier, mais dans le demi ou encore le quart d'espace, les calculs effectués précédemment vont aussi permettre de conclure et, essentiellement, le bilan d'énergie (III.49) se basera sur l'application de deux lemmes analogues aux Lemmes III.5.1 et III.5.2. On doit, cependant, être plus prudent dans tous nos calculs. En effet, lorsque l'on n'est plus dans l'espace entier, les termes qui composent  $v^n$  (resp.  $w^n$ ) ne sont plus anti-adjoints (resp. autoadjoints). Ils le sont *modulo* certains termes de bord qu'il faut venir prendre en compte.

Pour étayer ce propos donnons l'expression des termes  $v^n$  et  $w^n$ . Ils sont définis par

$$\begin{aligned} v^n &= -\lambda a D_{1,0} u^n - \mu b D_{2,0} u^n, \\ w^n &= -\frac{(\lambda a)^2}{2} \Delta_1 u^n - \frac{(\mu b)^2}{2} \Delta_2 u^n - \lambda \mu a b D_{1,0} D_{2,0} u^n + \frac{1}{8} ((\lambda a)^2 + (\mu b)^2) \Delta_1 \Delta_2 u^n. \end{aligned}$$

Ainsi, lorsque l'on fait le produit scalaire  $2\langle u^n; v^n \rangle$ , par intégration par parties discrète, des termes de bords apparaissent. On peut, toutefois assez facilement se convaincre que ces derniers ont le bon signe. Ils conduisent au contrôle des termes de bord dans l'estimation de stabilité du Théorème III.5.1.

Les choses sont moins agréables lorsque l'on considère l'autre produit scalaire que l'on a négligé, à savoir  $2\langle w^n; v^n \rangle$ . Nous n'allons pas développer ici tout ce terme. Cela serait un peu fastidieux, bien que pas difficile. Considérons seulement le premier terme qui apparait dans le développement *i.e.* le terme  $\langle D_{1,0} u^n; \Delta_1 u^n \rangle$ . Alors, si l'on se place, pour fixer notre propos, dans le demi-espace  $\llbracket -1, \infty \rrbracket \times \mathbb{Z}$ , on a maintenant une annulation à l'intérieur du domaine de ce terme, mais *modulo* un terme de bord. Plus précisément, on a la formule de dualité :

$$\langle D_{1,0} u^n; \Delta_1 u^n \rangle = -\frac{1}{2} \sum_{k \geq 0} (u_{0,k}^n - u_{-1,k}^n)^2. \quad (\text{III.50})$$

Cet exemple est assez frappant, car on remarque que sous la condition de Neumann d'ordre un (III.43), le membre de droite de (III.50) s'annule. Donc, au final, on a toujours la même propriété de dualité que dans le cas de l'espace entier mais ceci uniquement grâce au fait que la condition de bord est bien choisie.

Le cas du quart d'espace est alors très semblable au cas du demi-espace. La seule nouveauté étant qu'en plus des termes de bord précédemment décrits, des termes faisant intervenir la valeur au coin apparaissent dans des formules de dualité de type (III.50).

Ainsi notre analyse dans l'espace entier se transpose dans les géométries du demi-espace ou du quart d'espace. Les seules modifications notables à apporter sont que, dans les calculs explicites donnant le Lemme III.5.1 et les majorations donnant lieu au Lemme III.5.2, on doit prêter un soin particulier aux termes de bord ou de coin qui proviennent de nos défauts de dualité.

Une analyse détaillée (voir [Benoit and Coulombel, b]) de ces différents termes montre, néanmoins, qu'ils ont, au final, le bon signe si l'on se place sous la condition CFL (III.47).

### Les modifications avec Neumann d'ordre deux

L'analyse précédente peut aussi être étendue au cas de conditions de bord de Neumann d'ordre deux. Ceci a été réalisé dans [Benoit and Coulombel, a]. Toutefois, ceci nécessite un peu "d'astuce", d'être prêt à rogner un peu sur les paramètres CFL, et enfin pour finir, plus de peine.

Le schéma de la preuve et les diverses quantités en présence sont les mêmes. Cependant, comme on l'a vu dans le paragraphe précédent, dans la géométrie du quart d'espace, on a des défauts de dualité de la forme (III.50). Ces derniers avaient le bon goût de venir s'annuler lorsque l'on considérait des conditions de Neumann d'ordre un. Mais, cela n'est clairement plus le cas, lorsque l'on considère des conditions de Neumann d'ordre deux (voir par exemple (III.50)).

Ainsi, la même démonstration opère, mais, à chaque étape, on a à gérer des termes de bord et ou de coin, termes qu'il faut bien réussir à signer, à la fin de l'analyse, pour boucler le bilan d'énergie (III.48). C'est dans cette étape de bouclage final que l'on a été amené à restreindre les paramètres CFL à un domaine angulaire autour de la première bissectrice pour assurer que la contribution de tous ces termes de bord soit bien négative.

Il est ici, très probable, que nous n'ayons pas utilisé les estimations optimales et donc que notre résultat soit perfectible. Toutefois, notons que dans [Benoit and Coulombel, a], nous avons tout de même été assez précautionneux. En particulier, pour annuler le "maximum" de termes de bords et de coin nous avons :

- Utilisé la condition au coin (III.46) qui en annule "beaucoup".
- Modifié les énergies de façon ponctuelle sur les mailles fantômes, de manière à annuler le maximum de termes de bord.

On donne, ici, plus de détails sur le deuxième point. Dans le cas de la demi-droite, la "bonne" énergie à considérer lorsque l'on s'intéresse à des conditions de Neumann d'ordre deux est l'énergie localement pondérée :

$$\|u\|^2 := \frac{1}{2}u_0^2 + \sum_{j \geq 1} u_j^2, \text{ à la place de la norme 2 usuelle.}$$

Des phénomènes d'annulation analogues se déroulent aussi dans le cas du quart d'espace. Dans ce dernier, nous avons été amené à définir comme énergie

$$\|u\|_{\mathcal{H}}^2 := \sum_{j \geq 1} \sum_{k \geq 1} u_{j,k}^2 + \frac{1}{2} \sum_{k \geq 1} u_{0,k}^2 + \frac{1}{2} \sum_{j \geq 1} u_{j,0}^2 + \frac{1}{4} u_{0,0}^2,$$



là encore, pour pouvoir simplifier le plus de termes de bords et au coin possible et pour avoir le bilan d'énergie final le plus simple possible.

Certains termes de bords n'en demeurent pas moins dans l'analyse et pour gérer la contribution de ces derniers dans l'estimation d'énergie finale, on doit faire ces deux hypothèses :

1. Les deux ratios CFL  $\lambda$  et  $\mu$  sont comparables. On se restreint donc à des paramètres CFL autour de la première bissectrice.
2. La quantité  $(a\lambda)^2 + (b\mu)^2$  est assez petite<sup>13</sup>. On n'obtient donc pas, même au voisinage de la première bissectrice, la condition CFL optimale (III.47).

Peut-être, qu'un jour, nous reprendrons nos calculs et que nous tacherons alors d'améliorer le domaine d'applicabilité du résultat de [Benoit and Coulombel, a]. A l'heure actuelle, nous pensons, qu'en l'état, ce résultat est suffisant<sup>14</sup>. En effet, il donne, pour un problème dans un quart d'espace, un ensemble de conditions de bord et au coin qui conduisent à la stabilité. Il y a même fort à parier, qu'un tel schéma est convergent à l'ordre deux. Démontrer sa convergence devrait alors pouvoir se faire assez facilement puisque nous avons démontré sa stabilité.

Il se peut toutefois que le fait d'avoir récupéré dans [Benoit and Coulombel, b], la condition CFL optimale tienne, en fait, du "miracle". Et que génériquement, une méthode d'énergie ne permette pas de récupérer cet optimum.

### III.5.3 Une discrétisation de l'analyse d'Osher ?

Dans les deux paragraphes précédents nous avons donné des éléments que l'on pense nouveaux et notables concernant la stabilité des schémas définis dans un quart d'espace. Toutefois, les résultats précédemment mentionnés s'appliquent à des schémas déterminés et à des conditions de bords numériques particulières.

Une perspective de recherche, que l'on pense intéressante, serait d'essayer d'obtenir une caractérisation plus formelle et systémique des conditions de bord numériques donnant lieu à des schémas stables dans le quart d'espace. Nous avons entrepris une telle caractérisation dans un *preprint* (non soumis à publication car incomplet) (voir [Benoit, 2017b]). Nous aimerions donc reprendre ce travail et nous espérons, cette fois, pouvoir aboutir à un résultat. Il est probablement réaliste de considérer une caractérisation complète des conditions de bord numériques donnant lieu à un schéma stable, comme un problème ambitieux. Il doit donc être traité comme un problème à longue échéance. Toutefois, on aimerait dans ce qui suit, partager quelques pistes qui nous semblent intéressantes et qui pourraient permettre d'initialiser une telle étude.

Fort de l'analyse que l'on a fait dans le cas du segment, il est très tentant de tâcher d'adapter la méthode du problème continu dans le quart d'espace de Osher [Osher, 1973] à un cadre discret.

Cela nécessite essentiellement trois grandes étapes :

1. Identifier ce qui jouera le rôle des opérateurs de traces discrets,  $\mathbb{T}_{1 \rightarrow 2}$  et  $\mathbb{T}_{2 \rightarrow 1}$ . C'est-à-dire, les opérateurs qui à la trace (étendue) sur l'un des bords fantôme associent la valeur de la trace sur l'autre bord. On obtiendrait, alors une condition nécessaire pour la stabilité. Une telle condition nécessaire serait déjà une avancée, car elle serait très probablement une version discrétisée de la condition d'Osher [Osher, 1973]. C'est cette étude que nous avons initiée (peut-être de façon trop

<sup>13</sup>Nous n'avons pas ici cherché à quantifier ce "assez petit", ni à l'optimiser.

<sup>14</sup>Mais, ne dit-on pas que "tout ce qui n'est que suffisant ne suffit jamais ?" Marivaux

frontale) dans le *preprint* [Benoit, 2017b]. Nous reviendrons sur cette condition dans la suite de ce paragraphe.

2. Identifier proprement ce que l'on doit imposer sur un *symétriseur* dans le cas du quart d'espace. Ceci pour avoir l'estimation de stabilité du problème sous la forme résolvente.
3. Adapter la construction du *symétriseur* dans le cas continu d'Osher [Osher, 1973], au cadre discret. On espérerait alors obtenir le caractère suffisant de la condition (ou plus exactement d'une version uniforme) décrite au point 1.

Le point qui semble le plus délicat dans ce plan de bataille est le point 3, non parce que l'on pense que la méthode de construction d'Osher décrit dans le cas discret (voir Paragraphe III.4.2), ne s'appliquera pas (du tout) mais parce que cette dernière va se compliquer de deux façons.

Premièrement, les objets que l'on va venir manipuler vont une fois de plus hériter de la dimension infinie. En effet, les calculs que nous avons débutés montrent que les opérateurs de traces  $\mathbb{T}_{1 \rightarrow 2}$  et  $\mathbb{T}_{2 \rightarrow 1}$  vont agir sur des suites de  $\ell^2(\mathbb{N})$ . Par conséquent, nous aurons à résoudre le même problème que dans le cas continu, à savoir, l'obtention de bornes **uniformes** de ces opérateurs (c'est-à-dire les points *ii*) des Définitions III.4.2 ou II.3.2). Bornes uniformes, que l'on ne sait pas encore établir dans le cadre continu, pourtant réputé plus favorable.

Le deuxième point est d'une toute autre nature. Il s'agit du fait que pour sa construction le *symétriseur* d'Osher continu nécessite, dans un premier temps, de se ramener à des conditions de bords qui dépendent de la même variable. Cette étape dans le cas continu est facilement réalisée en passant en coordonnées polaires. Toutefois, dans le cas discret nous ne pouvons clairement pas effectuer un tel changement de variables sans modifier drastiquement le problème. Ainsi, il faudrait reprendre la construction du *symétriseur* discret et revoir clairement les requêtes que l'on fait sur ce dernier dans le point 2. Il faudrait alors déterminer si une telle paramétrisation des conditions de bord par rapport à la même variable est vraiment nécessaire ou si cette dernière a juste l'avantage de la simplicité.

Une autre piste qui pourrait peut-être, ici, être utile, pour la construction du *symétriseur* serait de changer de point de vue, comme proposé par Coulombel. L'idée centrale serait de ne plus considérer la solution du schéma comme une suite, mais comme une fonction (discrétisée) sur laquelle agissent des opérateurs de translation. L'avantage étant donc que l'on peut opérer un changement de variables en coordonnées polaires à une telle fonction. On peut alors, espérer que la construction d'Osher passe au discret. Les détails d'un tel changement de paradigme seront étudiés prochainement.

Un premier pas dans la caractérisation des conditions de bords numériques donnant lieu à la stabilité pourrait être de commencer par un (ou deux) schéma(s) bien particulier(s). On pense ici au schéma de Lax-Wendroff que l'on a déjà étudié au Paragraphe III.5.2, mais aussi au schéma de Lax-Friedrichs. La différence fondamentale de ces deux schémas dans un quart d'espace étant que le schéma de Lax-Wendroff nécessite une condition au coin, tandis que le schéma de Lax-Friedrichs s'en dispense. Donc, en disposant de résultats indépendants pour ces deux schémas particuliers, on peut espérer bien voir à quel niveau la condition au coin intervient.

Expliciter les opérateurs  $\mathbb{T}_{1 \rightarrow 2}$  et  $\mathbb{T}_{2 \rightarrow 1}$  pour ces deux schémas particuliers, démontrer leurs bornes uniformes<sup>15</sup>, serait intéressant car cela permettrait de mieux comprendre le rôle de la condition discrète d'Osher dans la théorie de stabilité. Nous avons débuté une telle étude et nous espérons pouvoir, à moyen terme, obtenir des résultats aboutis. Il serait ensuite éclairant de montrer la stabilité de ces deux schémas non pas en utilisant une méthode d'énergie comme celle que l'on a décrite dans le Paragraphe III.5.2, mais en construisant à la main et le plus explicitement possible, les symétriseurs pour ces deux problèmes. Ceci serait un excellent travail préparatoire avant de revenir au problème général.

---

<sup>15</sup>On pourra, ici, pleinement utiliser le fait que ces deux schémas ne développent pas de modes de *glancing*. Fréquences au voisinage desquelles, nous n'arrivons pas dans [Benoit, 2015] à établir de bornes uniformes.

## III.6 Quelques autres questions ouvertes

Dans cette dernière section, nous exposerons quelques autres perspectives de recherche à propos de la stabilité des schémas aux différences finies, perspectives que nous n'avons pas mentionnées dans la discussion précédente. Ces perspectives, qu'elles soient dans la géométrie du demi-espace, de l'intervalle, ou encore du quart d'espace, nous semblent constituer des questions intéressantes qui gagneraient à avoir des réponses claires.

### III.6.1 Établir un résultat probant de stabilité du problème dans le demi-espace

A notre avis, l'un des grands manques dans la théorie de stabilité GKS depuis son introduction dans [B. Gustafsson and Sundstrom., 1972] est, qu'à l'heure actuelle, elle ne s'applique, de façon totalement satisfaisante, qu'au problème dans la demi-droite.

A notre connaissance, le seul résultat donnant une extension à un demi-espace en dimension quelconque du résultat de [B. Gustafsson and Sundstrom., 1972] est dû à Michelson dans [Michelson, 1983]. L'auteur y montre essentiellement que la condition de Kreiss-Lopatinskii uniforme discrète passe aux problèmes dans le demi-espace multidimensionnel. Toutefois, ce résultat qui, à l'avantage certain d'être une première analyse requiert des hypothèses assez restrictives sur le schéma aux différences finies. Plus précisément, il demande *a minima* que l'opérateur de discrétisation à l'intérieur soit dissipatif dans le sens où les valeurs propres de la matrice d'amplification de von Neumann sont de module inférieur à un (voir par exemple [[B. Gustafsson and Sundstrom., 1972]-Assumption 5.4]). Une telle hypothèse est restrictive car elle exclut, par exemple le schéma de saute-mouton. De plus, avec les travaux de Coulombel [Coulombel, 2011], on sait qu'elle n'est pas nécessaire dans le cas de la demi-droite.

Par conséquent, on aimerait à l'avenir donner une théorie plus aboutie concernant la stabilité dans le demi-espace. Ceci nécessiterait probablement, dans un premier temps, de reprendre les travaux de Michelson [Michelson, 1983] et de voir s'ils peuvent être modifiés pour venir s'appliquer à une classe de schémas moins restrictive. Ou bien si, *a contrario*, on a plutôt besoin d'adapter l'approche des *K-symétriseurs*, d'abord introduite dans le cas continu par [Métivier and Zumbrun, 2004] pour généraliser les *symétriseurs* de Kreiss, et ensuite adaptée au cas discret par Coulombel, pour traiter le cas des schémas géométriquement réguliers.

Une telle extension de la théorie à des schémas dans le demi-espace serait, bien sûr, en plus de ses applications pratiques, intéressante en elle-même d'un point de vue théorique. Mais, il serait aussi plaisant d'en disposer avant de débiter l'étude générale dans le cas du quart d'espace car cela éviterait de venir imposer la condition de Kreiss-Lopatinskii uniforme discrète sur chaque bord sans être sûr de son caractère ni nécessaire, ni suffisant.

### III.6.2 Améliorer les conditions d'applicabilité du résultat de [Benoit, 2022]

L'article [Benoit, 2022] a l'avantage d'être un premier travail sur le sujet, mais, par conséquent, il faut aussi noter que les hypothèses sous lesquelles s'appliquent ce résultat sont probablement loin d'être optimales. Nous pensons qu'il serait intéressant de reprendre l'analyse de [Benoit, 2022], en tentant d'améliorer les hypothèses sous lesquelles on se place. Nous avons deux grandes pistes d'amélioration :

1. La première hypothèse que l'on a été amené à faire, pour rendre l'analyse de [Benoit, 2022], la moins technique et la plus plaisante possible, est une restriction sur les *stencils* de bords. En effet, comme déjà évoqué, on a supposé que le *stencil* sur le bord gauche (resp. droit) vérifie la condition  $b_0 < r$  (resp.  $b_1 < \ell$ ). Cette hypothèse est utilisée de façon non triviale dans la reformulation du problème sous forme résolvente. Toutefois, il a été montré notamment dans les travaux de Coulombel

[Coulombel, 2012], que cette hypothèse n'était pas nécessaire dans le cas de la demi-droite. Sans cette restriction, on peut quand même, mettre le problème sous la forme résolvente

$$\begin{cases} \mathcal{W}_{j+1} = \tilde{\mathbb{M}}(z)\mathcal{W}_j & j \geq 1, \\ \tilde{\mathbb{B}}_0(z)\mathcal{W}_1 = \tilde{\mathcal{G}}, \end{cases}$$

pour de nouvelles matrices  $\tilde{\mathbb{M}}(z)$  et  $\tilde{\mathbb{B}}_0(z)$  adéquates. L'un des principaux points qui fait alors diverger cette nouvelle forme résolvente de celle que l'on a traitée est qu'alors, la nouvelle matrice résolvente  $\tilde{\mathbb{M}}(z)$  n'est plus inversible. Or, nous avons utilisé ce fait de façon non triviale dans l'analyse du Paragraphe III.4. Il serait donc assez intéressant, peut-être un peu plus pour la beauté de l'art qu'au vu des applications, de voir si l'analyse de [Benoit, 2022] peut se généraliser dans un tel cadre. En effet, nous avons, ici, quelques doutes sur l'utilité pratique de schémas pour lesquels on a un *stencil* au bord supérieur à celui à l'intérieur. Est-il vraiment moral de préférer une plus grande précision sur les quelques mailles fantômes au bord, que sur les nombreuses mailles à l'intérieur ?

Qu'à cela ne tienne, nous pensons que de tels schémas méritent, néanmoins, une analyse théorique. Pour cela, nous pensons que le manque d'inversibilité de la matrice résolvente  $\tilde{\mathbb{M}}(z)$  devrait pouvoir être surmonté. En effet, comparé au cadre précédent nous n'avons pas de perte d'informations car une analyse approfondie de l'expression du nouveau couple résolvant  $(\tilde{\mathbb{M}}, \tilde{\mathbb{B}}_0)$  montre que les informations manquantes pour l'inversibilité de  $\tilde{\mathbb{M}}$  sont maintenant stockées dans la matrice de bord numérique  $\tilde{\mathbb{B}}_0$ . Il s'agirait donc de trouver une décomposition de l'espace  $\mathbb{C}^N$  qui permette d'utiliser simultanément la structure de  $\tilde{\mathbb{M}}$  et celle de  $\tilde{\mathbb{B}}_0$ .

2. Une deuxième amélioration que l'on pourrait apporter quant aux hypothèses de [Benoit, 2022] porte sur l'hypothèse que l'on a été amené à faire sur la norme  $\ell^2(\mathbb{Z})$  de l'opérateur de discrétisation interne  $Q$ , *i.e.*, l'Hypothèse III.3.2. On aimerait pouvoir se passer de cette hypothèse qui est assez restrictive. Puisque cette hypothèse est utilisée (en ce qui concerne uniquement la stabilité GKS, pas celle du semi-groupe) pour se ramener à un terme source nul à l'intérieur nul, la méthode la plus simple (mais au combien insatisfaisante) est de se restreindre à un terme source à l'intérieur nul.

Nous pensons que cette hypothèse pourrait toutefois être écartée de façon plus satisfaisante en suivant l'une des deux méthodes décrites ci-dessous :

- La conséquence de l'Hypothèse III.3.2 que l'on utilise, pour la stabilité *GKS*, est la stabilité du schéma dans l'intervalle avec conditions de bord de Dirichlet. Comme précédemment mentionné, il serait bon de pouvoir déterminer, de façon *ad hoc*, la stabilité de schémas par des arguments directs d'intégration par parties pour une classe de conditions de bords numériques plus grande que juste la condition de Dirichlet. Nous voudrions en fait avoir un véritable équivalent discret des conditions de bord de type strictement dissipatives. Un premier point dans l'étude d'une telle caractérisation serait de reprendre le travail de Strang [Strang, 1969b]. Pour les systèmes (continus) à deux équations, ce qui doit se traduire par des schémas à "petits" *stencils*, Strang établit l'équivalence entre le fait pour une condition de bord d'être strictement dissipative et de vérifier la condition de Kreiss-Lopatinskii uniforme. On aimerait donc essayer de voir si cette caractérisation passe au niveau discret et, si oui, quelles sont exactement les conditions de bord que l'on obtient. Il est toutefois à noter que l'on ne pense pas que la méthode décrite précédemment permette de se passer totalement de conditions sur l'opérateur de discrétisation interne. En effet, si l'on fait l'analogie avec le cas continu, bien que l'on n'ait pas besoin de faire d'hypothèse pour avoir une condition strictement dissipative, l'hypothèse de symétrie des coefficients demeure.

- Une deuxième idée qui, elle, permettrait de se libérer totalement de l’Hypothèse III.3.2, rejoint une des perspectives dans la bande, à savoir celle d’éviter totalement l’étape de restriction à un terme source intérieur nul. C’est cette perspective que l’on a qualifiée de construction d’un *symétriseur* qui ”symétrise vraiment”.

Pour ce faire, on aura alors besoin d’un *symétriseur* qui peut gérer des termes sources à l’intérieur autrement dit d’assurer le point *iii*) de la Définition III.4.2 avec un terme source non nul dans le membre de droite. Un tel *symétriseur* pourrait être construit, d’abord dans le cas continu, puis dans le cas discret, même si l’on pense que la transposition d’un cadre à l’autre devrait être assez aisée.

### III.6.3 Une méthode numérique de vérification de la condition de stabilité dans l’intervalle à le Barbenchon ?

Concluons ce chapitre par une perspective assez exploratoire mais que l’on pense intéressante, notamment au vu de ses applications pratiques. Elle est directement inspirée des résultats de vérification numérique de la condition de Kreiss-Lopatinskii uniforme discrète initiée dans les articles [Le Barbenchon et al., a] et [Le Barbenchon et al., b].

On a déjà mentionné dans le Paragraphe II.2.2, avec en vue des extensions dans le cadre continu, que cette méthode consiste en l’introduction d’une fonction caractérisant la stabilité. La quantité introduite par le Barbenchon est un déterminant de Lopatinskii classique si ce n’est que, ce dernier a le bon goût d’être holomorphe par rapport à la variable fréquentielle,  $z$  dans le cas discret.

Ainsi, la question de la stabilité revient à déterminer les zéros de cette fonction holomorphe, zéros qui peuvent donc être déterminés par des nombres d’entrelacements. Ces derniers sont alors, assez faciles et robustes à déterminer de façon numérique.

Une extension possible de cette jolie idée, nous semble être la détermination numérique de la stabilité des schémas, mais dans l’intervalle cette fois. En effet, comme on l’a vu dans la Section III.4, la stabilité est alors caractérisée par :

- Avoir la condition de Kreiss-Lopatinskii discrète uniforme à chaque bord de l’intervalle. Ici la méthode de [Le Barbenchon et al., b] s’applique directement.
- Avoir l’inversibilité d’une nouvelle matrice, à savoir la matrice  $I - \bar{\mathbb{T}}_{0 \rightarrow 0}(z)$ , qui est en jeu dans la Proposition III.4.2. Puisqu’ici, contrairement au problème dans la bande, cette condition fait intervenir une matrice, on pense qu’il est intéressant d’essayer d’appliquer une méthode basée sur celle de [Le Barbenchon et al., b] pour étudier l’inversibilité de cette matrice. Plus précisément, on aimerait trouver une matrice  $I - \tilde{\mathbb{T}}_{0 \rightarrow 0}(z)$  dont les (éventuels) zéros du déterminant seraient les mêmes que ceux de  $I - \bar{\mathbb{T}}_{0 \rightarrow 0}(z)$ . La seule différence étant ici que l’on voudrait que  $I - \tilde{\mathbb{T}}_{0 \rightarrow 0}(z)$  dispose de surcroît de propriétés d’holomorphie par rapport à  $z$ . Ceci ramènerait alors la question de l’inversibilité à un calcul d’un nombre d’entrelacements que l’on pourrait venir calculer avec les méthodes numériques développées dans [Le Barbenchon et al., b].

Une question qui demeure toutefois dans ce programme est de comprendre comment venir traiter de façon numérique le fait que cette inversibilité doit avoir lieu de façon uniforme par rapport au nombre de mailles dans la discrétisation de l’intervalle utilisée.

Une autre extension possible de ce travail, qui devrait nécessiter plus d’efforts, serait une extension de la méthode à un cadre multidimensionnel. Mais alors, il y a plus de variables fréquentielles. On ne peut espérer aucune propriété d’holomorphie par rapport aux nouvelles variables. Se ramener à un calcul de nombre d’entrelacement n’est donc pas immédiat. Nous pensons néanmoins que disposer d’un code numérique vérifiant la condition de Kreiss-Lopatinskii discrète uniforme mérite bien de se pencher

sur ce problème. En effet, nous pensons qu'un tel code de vérification numérique pourrait permettre de populariser une théorie peut-être un peu impopulaire.

# Chapter IV

## Autres résultats notables

Ce chapitre regroupe des résultats que nous avons obtenus (ou qui sont en cours de rédaction), lors de différentes collaborations scientifiques. Bien qu'il s'agisse de résultats portant sur des équations aux dérivées partielles (le plus souvent toujours hyperboliques), ils n'entrent pas directement dans le cadre de la théorie que l'on a décrite dans les deux chapitres précédents. Par conséquent, on préfère les séparer des précédents dans un souci de cohérence.

### IV.1 Homogénéisation en temps long de l'équation des ondes

Ce paragraphe décrit les résultats obtenus durant mon postdoctorat sous la direction d'Antoine Gloria. Ils ont été publiés dans [Benoit and Gloria, 2019].

La question qui nous a intéressé est celle de l'homogénéisation, principalement pour l'équation des ondes, même si les résultats de [Benoit and Gloria, 2019] couvrent aussi le problème elliptique archétype de la théorie de l'homogénéisation  $-\nabla \cdot (\mathbf{a}\nabla)$ . Nous nous sommes penchés sur l'équation des ondes posée dans un milieu où les lois de propagation dépendent d'un certain champs  $\mathbf{a} : \mathbb{R}^d \rightarrow \mathbf{M}_{d \times d}(\mathbb{R})$ , et où cette dépendance se fait de façon rapide à une échelle  $0 < \varepsilon \ll 1$ . Plus précisément, l'équation que l'on a considéré est la suivante : pour  $T > 0$

$$\begin{cases} \partial_{tt}^2 u^\varepsilon - \nabla \cdot (\mathbf{a}(\frac{x}{\varepsilon}) \nabla u^\varepsilon) = 0 & \text{pour } (t, x) \in ]0, T[ \times \mathbb{R}^d, \\ u^\varepsilon|_{t=0} = u_0 \text{ et } (\partial_t u^\varepsilon)|_{t=0} = 0 & \text{pour } x \in \mathbb{R}^d, \end{cases} \quad (\text{IV.1})$$

la donnée initiale  $u_0$  étant supposée régulière. On a typiquement,  $u_0 \in \mathcal{S}(\mathbb{R}^d)$ .

Le but de la théorie de l'homogénéisation est, alors de caractériser le comportement de la limite de la solution,  $u^0$ , lorsque le petit paramètre  $\varepsilon \downarrow 0$ . Plus précisément, on cherche à déterminer  $u^0$  comme la solution d'une équation limite "moyennée".

De tels résultats, que cela soit pour l'équation des ondes ou bien pour l'équation elliptique archétype  $-\nabla \cdot \mathbf{a}(\frac{\cdot}{\varepsilon}) \nabla u$ , sont bien connus et font l'objet d'une littérature très conséquente. On renvoie par exemple aux deux ouvrages [Bensoussan et al., 1978] et [Tartar, 2009]. Ces résultats dépendent principalement de la loi de propagation qui est autorisée par le champs  $\mathbf{a}$ . Historiquement, les premiers cas considérés ont été des dépendances déterministes périodiques, qui ont ensuite été étendues à des cadres quasi-périodiques, ou bien presque périodiques. Plus récemment des dépendances aléatoires sont aussi considérées.

Concernant l'homogénéisation de l'équation des ondes (IV.1), pour fixer les idées, intéressons nous au cas périodique, il est bien connu que la limite de la solution  $u^0$  vérifie l'équation des ondes :

$$\begin{cases} \partial_{tt}^2 u^0 - \mathbf{a}^0 \cdot \nabla^2 u^0 = 0 & \text{pour } (t, x) \in ]0, T[ \times \mathbb{R}^d, \\ u^0|_{t=0} = u_0 \text{ et } (\partial_t u^0)|_{t=0} = 0 & \text{pour } x \in \mathbb{R}^d, \end{cases} \quad (\text{IV.2})$$

où le coefficient  $\mathbf{a}^0$ , dit coefficient homogénéisé, est défini comme une version moyennée du coefficient oscillant  $\mathbf{a}(\cdot/\varepsilon)$ . On renvoie, par exemple, à [Francfort and Murat, 1992], à [Brahim-Otsmane et al., 1992] où la question de la construction d’un correcteur est aussi adressée. Ce coefficient est le même que celui de l’équation elliptique stationnaire  $-\nabla \cdot \mathbf{a}(x/\varepsilon)\nabla u = f$ . Dans les cadres quasi-périodique, presque-périodique ou bien aléatoires, on obtient alors des résultats analogues en modifiant de façon idoine la façon de moyenner  $\mathbf{a}$ .

Il est toutefois à noter que le résultat d’homogénéisation précédent ne s’applique que sur une fenêtre de temps  $T > 0$ , fini et indépendant de  $\varepsilon$ . La question que nous nous sommes donc posé dans l’article [Benoit and Gloria, 2019] est d’essayer de décrire quel peut être le comportement de la limite de la solution de (IV.1), mais sur des échelles de temps plus longues, c’est-à-dire typiquement de la forme  $T = O(\varepsilon^{-\ell})$  avec  $\ell \geq 1$ .

Le travail pionnier sur cette question est celui de Santosa-Symes [Santosa and Symes, 1991], ensuite rendu rigoureux dans la série de travaux (qui nous a motivés à nous plonger dans ce sujet) de [Dohnal et al., 2015], [Dohnal et al., 2014] et [Lamacz, 2011]. Tous ces contributions montrent que l’apparition d’effets dispersifs sur des échelles de temps de taille  $T = O(1/\varepsilon^2)$  est attendue et donc qu’elle doit donc être prise en compte. Plus précisément, dans le cas où le champs  $\mathbf{a}$  est périodique, il est démontré dans les articles précédents que, si  $u^\varepsilon$  désigne la solution de (IV.1) et si  $w^\varepsilon$  désigne la solution de l’équation d’onde dispersive, qui précise l’équation homogénéisée, suivante :

$$\begin{cases} \partial_{tt}^2 w^\varepsilon - \mathbf{a}^0 \cdot \nabla^2 w^\varepsilon - \varepsilon^2 \mathbf{b} \cdot \nabla^2 \partial_{tt}^2 w^\varepsilon + \varepsilon^2 \mathbf{c} \cdot \nabla^4 w^\varepsilon = 0 & \text{pour } (t, x) \in ]0, T[ \times \mathbb{R}^d, \\ w^\varepsilon|_{t=0} = u_0 \text{ et } (\partial_t w^\varepsilon)|_{t=0} = 0 & \text{pour } x \in \mathbb{R}^d, \end{cases} \quad (\text{IV.3})$$

où  $\mathbf{b}$  est une matrice définie positive et où  $\mathbf{c}$  est un tenseur d’ordre quatre défini positif. Ces deux objets pouvant être explicitement calculés à partir du champs  $\mathbf{a}$ . Alors, on a l’asymptotique suivante :

$$\sup_{0 \leq t \leq T} \|u^\varepsilon - w^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 \leq C\varepsilon^3 T, \quad (\text{IV.4})$$

qui tient en particulier sur des échelles de temps de la forme  $[1/\varepsilon^2, 1/\varepsilon^3[$  et précise donc le résultat d’homogénéisation standard.

Dans [Benoit and Gloria, 2019] (voir aussi l’article [Allaire et al., 2022]), nous avons étendu ce genre de résultats d’homogénéisation pour des échelles de temps de la forme  $1/\varepsilon^\ell$  avec  $\ell$  quelconque, en fournissant une équation homogénéisée qui présente un nombre de termes dispersifs correctifs croissant avec  $\ell$ . Notre résultat se base sur une théorie de Bloch-Floquet approchée (dite de Taylor-Bloch), que l’on exposera avec plus de détails dans ce qui suit.

On doit ici aussi mentionner le travail indépendant et simultané de [Allaire et al., 2022] qui se base lui sur l’expansion à double échelles de [Allaire, 1992]. Dans ce travail, les auteurs se placent dans un cadre périodique et fournissent (en temps long) des équations homogénéisées ”criminelles” (au sens où ces dernières nécessitent un gros travail de reformulation pour aboutir à des équations bien posées). Comme on le verra, ce point n’a pas été spécialement problématique dans [Benoit and Gloria, 2019] où les équations homogénéisées en temps long peuvent être facilement rendues bien posées. Toutefois, il est à noter que l’ordre de convergence obtenu dans [Allaire et al., 2022] est meilleur que celui de [Benoit and Gloria, 2019]. Ces deux approches semblent maintenant avoir été réconciliées dans le récent *preprint* [Duerinckx et al., ], où les auteurs obtiennent des résultats en temps longs, des équations homogénéisées bien posées, avec les ordres de convergence optimaux.

L’idée de la démonstration que l’on a choisie de suivre dans [Benoit and Gloria, 2019] est d’essayer d’adapter l’analyse de Bloch-Floquet qui permet, dans le cas d’un coefficient  $\mathbf{a}$  périodique, de diagonaliser



l'opérateur  $-\nabla \cdot (\mathbf{a}\nabla)$ . Cette théorie généralise donc l'analyse de Fourier qui diagonalise le Laplacien. Sans entrer dans une foule de détails (on renvoie par exemple à [[Dohnal et al., 2015]-Section 2.1]), dans le cas périodique, il est un résultat établi que l'on peut venir exprimer la solution,  $u^\varepsilon$ , de (IV.1) sous la forme :

$$u^\varepsilon(t, x) := \sum_{m \geq 0} \int_{Z/\varepsilon} \widehat{u}_0(k) e^{ik \cdot x} \psi_m\left(\frac{x}{\varepsilon}, \varepsilon k\right) \cos(\varepsilon^{-1} \sqrt{\lambda_m(\varepsilon k)} t) dk, \quad (\text{IV.5})$$

où  $Z := [-\frac{1}{2}, \frac{1}{2}]^d$  et où le vecteur propre  $\psi_m$  et la valeur propre  $\lambda_m$  sont solutions du problème spectral :

$$-(\nabla_y + ik) \cdot \mathbf{a}(\nabla_y + ik) \psi_m(y, k) = \lambda_m(k) \psi_m(y, k) \text{ dans le cube.} \quad (\text{IV.6})$$

Une première étape dans [Benoit and Gloria, 2019] a été de définir la notion d'ondes de Bloch-Floquet approchées. C'est-à-dire qu'au lieu de résoudre l'équation spectrale (IV.6) de façon exacte, on permet à cette dernière d'être résolue de façon approchée. On a donc une propriété de diagonalisation avec un défaut de diagonalisation,  $\mathfrak{d}$ . Plus précisément, nous avons montré que

**Proposition IV.1.1 ([Benoit and Gloria, 2019]-Proposition 2)** *Soit  $k := \kappa e$ , une fréquence avec  $\kappa \in \mathbb{R}_+$  (ce dernier ayant vocation à être rescalé en  $\varepsilon$ ) et  $e \in \mathbb{S}^{d-1}$ , soit  $\ell \geq 1$ , alors il existe une onde de Taylor-Bloch,  $\psi_{k,\ell}$ , une valeur propre de Taylor-Bloch,  $\tilde{\lambda}_{k,\ell}$ , et un défaut de diagonalisation de Taylor-Bloch,  $\mathfrak{d}_{k,\ell}$ , tels que l'on ait*

$$-(\nabla + ik) \cdot \mathbf{a}(\nabla + ik) \psi_{k,\ell} = \tilde{\lambda}_{k,\ell} \psi_{k,\ell} - (i\kappa)^{\ell+1} \mathfrak{d}_{k,\ell}. \quad (\text{IV.7})$$

Les trois fonctions précédentes s'expriment de façon explicite en termes d'un quadruplet de fonctions  $(\varphi_j, \lambda_j, \sigma_j, \chi_j)_{0 \leq j \leq \ell}$ , qui s'avèrent être les correcteurs étendus usuels en homogénéisation.

La définition des correcteurs étendus  $(\varphi_j, \lambda_j, \sigma_j, \chi_j)_{0 \leq j \leq \ell}$  est assez lourde et n'est pas au cœur de notre propos. On renvoie à [[Benoit and Gloria, 2019]-Definition 2.1] pour une définition précise. Expliquons cependant, pourquoi de tels objets sont ici attendus. Il est assez bien connu que lorsque l'on utilise la théorie de Bloch-Floquet pour l'homogénéisation (voir [Dohnal et al., 2014]), alors le coefficient homogénéisé  $\mathbf{a}^0$  s'obtient comme la dérivée fréquentielle, en zéro, de la valeur propre de Bloch. De la même façon, le premier correcteur  $\varphi_1$  est obtenu comme dérivée, en zéro, du vecteur propre de Bloch. Ainsi si l'on définit

$$\psi_{k,\ell} := \sum_{j=0}^{\ell} (i\kappa)^j \varphi_j \text{ et } \tilde{\lambda}_{k,\ell} := \kappa^2 \sum_{j=0}^{\ell} (i\kappa)^j \lambda_j,$$

où les  $\varphi_j$  sont les correcteurs étendus et où les réels  $\lambda_j$  proviennent de l'évaluation dans la direction  $e$  des coefficients homogénéisés étendus  $\mathbf{a}^j$  (tenseur d'ordre  $j+2$ ), alors, on généralise juste à tous les ordres la remarque précédente.

L'avantage d'une telle théorie approchée est que l'on ne demande pas l'existence d'une onde de Bloch exacte. Par conséquent, on peut se passer de l'hypothèse de périodicité du champs  $\mathbf{a}$ . Ceci nous a permis de considérer, par exemple, des coefficients presque périodiques ou même aléatoires.

Cela ne se fait toutefois pas sans prix à payer. En effet, pour définir nos ondes de Taylor-Bloch approchées, on a besoin de définir les correcteurs étendus. Cela peut se faire assez naturellement dans le cas périodique (on utilise ici essentiellement l'inégalité de Poincaré), mais dans les autres cadres nous avons eu besoin du travail (très) conséquent de Otto et Gloria et de leurs coauteurs pour montrer l'existence de tels objets. Ce travail a fait l'objet de nombreuses publications, on renverra à [Armstrong et al., 2016] en ce qui concerne le cas presque périodique et à [Duerinckx and Gloria, 2020a], [Duerinckx and Gloria, 2020b] et [Gloria et al., 2021] en ce qui concerne le cas aléatoire.

Pour être totalement précis, plus que l'existence de ces correcteurs, nous avons aussi besoin de pouvoir les contrôler. On renvoie à l'Hypothèse 1 de [Benoit and Gloria, 2019] pour une description

détaillée. On se contentera ici d'une version un peu brute de cette hypothèse qui affirme que les  $\ell - 1$  premiers correcteurs étendus sont bornés, tandis que le  $\ell^{\text{ième}}$  peut être contrôlé par le poids  $\mu_\alpha(|x|)$  défini par  $\mu_\alpha(t) := (1 + t)^{\alpha_1} \log^{\alpha_2}(2 + t)$ , où  $\alpha := (\alpha_1, \alpha_2) \in \mathbb{R}^2$ . Cette hypothèse est motivée par le fait que ces contrôles sont typiquement ceux que l'on va obtenir dans le cadre aléatoire. On renvoie à [Gloria et al., 2021] pour plus de détails.

Ainsi, si l'on définit

$$v_{\varepsilon, \ell} := \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \omega_\ell(\varepsilon|k|) \widehat{u}_0(k) e^{ik \cdot x} \cos(\varepsilon^{-1} \sqrt{\widetilde{\lambda}_{k, \ell} t}) \psi_{k, \ell} \left( \frac{x}{\varepsilon}, \varepsilon k \right) dk,$$

où la fonction  $\omega_\ell$  désigne une fonction de *cut-off* qui n'est pas au cœur de nos préoccupations, on obtient une solution approchée basée sur les ondes de Taylor-Bloch et les valeurs propres de Taylor-Bloch. Cette solution approchée doit être vue comme une généralisation de la vraie solution donnée par la théorie de Bloch-Floquet, voir équation (IV.5).

Le résultat principal de [Benoit and Gloria, 2019] donne alors une solution approchée de  $u^\varepsilon$  sur des échelles de temps longues grâce aux ondes de Taylor-Bloch et aux valeurs propres de Taylor-Bloch. On a plus précisément :

**Théorème IV.1.1 ([Benoit and Gloria, 2019]-Theorem 1)** *Soit  $\ell \geq 1$ , sous l'hypothèse de contrôle et d'existence des  $\ell$  premiers correcteurs étendus alors, pour tout  $T > 0$ , et pour tout  $0 < \varepsilon \ll 1$ , si on définit la représentation donnée par l'onde de Taylor-Bloch tronquée<sup>1</sup>,  $u_{\varepsilon, \ell}$ , par*

$$u_{\varepsilon, \ell} := \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \omega_\ell(\varepsilon|k|) \widehat{u}_0(k) e^{ik \cdot x} \cos(\varepsilon^{-1} \sqrt{\widetilde{\lambda}_{k, \ell} t}) dk$$

on a l'estimation

$$\sup_{0 \leq t \leq T} \mathbb{E} \left[ \|u^\varepsilon - u_{\varepsilon, \ell}\|_{L^2(\mathbb{R}^d)}^2 \right]^{\frac{1}{2}} \leq C_\ell (\max \{ \varepsilon, \varepsilon^\ell \mu_\alpha(\varepsilon^{-1}) \} + \varepsilon^\ell T \mu_\alpha(\varepsilon^{-1} T)), \quad (\text{IV.8})$$

où  $\mu_\alpha$  désigne le poids contrôlant le  $\ell^{\text{ième}}$  correcteur et où  $\mathbb{E}[\cdot]$  désigne l'espérance mathématique qui rentre en compte dans un cadre aléatoire.

En particulier, on obtient donc une solution approchée de l'équation des ondes sur des temps de l'ordre de  $O(\varepsilon^{-\ell})$ .

La démonstration de ce théorème se fait alors en trois étapes qui apparaissent déjà dans le travail de [Dohnal et al., 2014] :

1. On se ramène à étudier l'équation des ondes (IV.1) pour laquelle les données initiales sont bien préparées à la décomposition de Taylor-Bloch. Soit  $\widetilde{v}_{\varepsilon, \ell}$  sa solution on montre alors que l'écart entre  $u^\varepsilon$  et  $\widetilde{v}_{\varepsilon, \ell}$  est contrôlable par un terme compatible avec le membre de droite de (IV.8).
2. On montre ensuite que l'écart entre  $\widetilde{v}_{\varepsilon, \ell}$  et la propagation "naturelle" d'une donnée initiale bien préparée, à savoir  $v_{\varepsilon, \ell}$ , est contrôlable par le terme de droite de (IV.8). Ce point se fait principalement en appliquant des estimations d'énergie pour l'équation des ondes au défaut de diagonalisation<sup>2</sup>  $\mathfrak{D}_{k, \ell}$ .

<sup>1</sup>On utilise ici le fait que  $\varphi_0 \equiv 1$ .

<sup>2</sup>On utilise ici, en particulier, le fait que ce défaut de diagonalisation s'écrit sous une forme très particulière (voir [Benoit and Gloria, 2019]-Definition 2.2) dont l'un des termes est sous forme divergence. C'est important pour obtenir des estimations d'erreur assez fine.

3. On conclut en montrant que dans l'expression de  $v_{\varepsilon,\ell}$ , on peut venir oublier la contribution des correcteurs d'ordre supérieur dans le vecteur propre et ne conserver que le premier pour obtenir la fonction  $u_{\varepsilon,\ell}$  du Théorème IV.1.1.

Une fois l'estimation (IV.8) établie, le résultat d'homogénéisation en temps long pour l'équation des ondes (IV.1) suit assez naturellement. En effet, on a, au vu de sa définition, que

$$\partial_{tt}^2 u_{\varepsilon,\ell} = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \omega_\ell(\varepsilon|k|) \widehat{u_0}(k) e^{ik \cdot x} \left[ \sum_{j=0}^{\ell-1} (i\varepsilon)^j |k|^{j+2} \lambda_j \right] \cos(\varepsilon^{-1} \sqrt{\widetilde{\lambda}_{k,\ell} t}) dk.$$

On rappelle ici, que les scalaires  $\lambda_j$  proviennent de l'évaluation dans la direction  $e$  d'un tenseur d'ordre  $j+2$ ,  $\mathbf{a}^j$  défini à l'aide des correcteurs étendus. Après quelques calculs sur les tenseurs, on obtient que  $u_{\varepsilon,\ell}$  est solution de l'équation homogénéisée

$$\begin{cases} \partial_{tt}^2 u_{\varepsilon,\ell} - \underbrace{\sum_{j=0}^{\ell-1} \varepsilon^j \mathbf{a}^j \cdot \nabla^{j+2}}_{:= \mathcal{L}_{hom,\varepsilon,\ell}} u_{\varepsilon,\ell} = 0 & \text{dans } ]0, T[ \times \mathbb{R}^d, \\ u_{\varepsilon,\ell}|_{t=0} = u_0 \text{ et } (\partial_t u_{\varepsilon,\ell})|_{t=0} = 0. \end{cases} \quad (\text{IV.9})$$

où l'on a posé :

$$\mathbf{a}^j \cdot \nabla^{j+2} v := \prod_{h=1}^{j+2} \sum_{i_h=1}^d [\mathbf{a}^j]_{i_1, \dots, i_{j+2}} \nabla_{i_1} \cdots \nabla_{i_{j+2}} v.$$

Définition dans laquelle on peut remarquer que tous les tenseurs impairs  $\mathbf{a}^{2k+1}$  s'annulent (on renvoie à [Benoit and Gloria, 2019]-Proposition 1).

Il n'est toutefois pas clair que le dernier tenseur de (IV.9) soit signé et donc, que cette équation soit bien posée. Ce problème avait déjà été rencontré dans [Dohnal et al., 2014] concernant le deuxième tenseur d'homogénéisation  $\mathbf{a}^2$ . Les auteurs utilisaient alors *l'astuce de Boussinesq* pour l'exprimer en fonction de  $\mathbf{a}^0$  et des deux tenseurs (signés)  $\mathbf{b}$  et  $\mathbf{c}$  apparaissant dans l'équation (IV.2).

Dans [Benoit and Gloria, 2019], pour assurer que notre équation homogénéisée est bien posée et ainsi éviter les équations "criminelles" de [Allaire et al., 2022], nous avons juste procédé par régularisation elliptique en considérant, en place de (IV.9), l'équation régularisée :

$$\begin{cases} \partial_{tt}^2 w_{\varepsilon,\ell} - \mathcal{L}_{hom,\varepsilon,\ell} w_{\varepsilon,\ell} - \gamma_\ell (i\varepsilon)^{2\lceil \frac{\ell-1}{2} \rceil + 1} I_d \cdot \nabla^{2(\lceil \frac{\ell-1}{2} \rceil + 2)} w_{\varepsilon,\ell} = 0 & \text{dans } ]0, T[ \times \mathbb{R}^d, \\ w_{\varepsilon,\ell}|_{t=0} = u_0 \text{ et } (\partial_t w_{\varepsilon,\ell})|_{t=0} = 0, \end{cases} \quad (\text{IV.10})$$

où  $\gamma_\ell$  peut être choisi assez grand pour que l'équation (IV.10) soit bien posée (on pourra consulter [[Benoit and Gloria, 2019]-Lemma 4.2]).

Le théorème d'homogénéisation final dans [Benoit and Gloria, 2019] est alors le suivant

**Théorème IV.1.2** *Soit  $\ell \geq 1$ , tel que l'on ait la propriété de contrôle sur les  $\ell$  premiers correcteurs étendus. Soit  $\varepsilon > 0$ , alors si  $u^\varepsilon$  désigne la solution de (IV.1) et si  $w_{\varepsilon,\ell}$  désigne celle de (IV.10), on a l'estimation*

$$\sup_{0 \leq t \leq T} \mathbb{E} \left[ \|u^\varepsilon - w_{\varepsilon,\ell}\|_{L^2(\mathbb{R}^d)}^2 \right]^{\frac{1}{2}} \leq C_\ell (\varepsilon + \varepsilon^\ell T \mu_\alpha(\varepsilon^{-1} T)), \quad (\text{IV.11})$$

où l'on rappelle que  $\mu_\alpha$  est le poids contrôlant le  $\ell^{\text{ième}}$  correcteur.

En particulier, on a donc bien le résultat d’homogénéisation sur les échelles de temps en  $O(\varepsilon^{-\ell})$  que l’on recherchait.

On conclura ici en mentionnant que la méthode précédente de diagonalisation approchée de Taylor-Bloch permet aussi d’obtenir des résultats d’homogénéisation pour l’équation elliptique stationnaire ainsi qu’un résultat de transport balistique (ou de délocalisation) pour l’équation des ondes en milieu aléatoire. On ne décrira pas précisément ces résultats ici, et l’on renvoie à [[Benoit and Gloria, 2019]-Theorems 3 and 4] pour plus de détails.

## IV.2 Contrôlabilité

Cette section résume des résultats de théorie du contrôle obtenus dans le cadre du coencadrement de la thèse de Romain Loyer avec Lionel Rosier au sein de l’ULCO. Il y est ici question de méthodes de contrôles directes c’est-à-dire qui évitent l’introduction d’un problème dual et les inégalités d’observabilités qui en résultent. Nous ne décrirons pas ici la littérature très conséquente concernant la théorie du contrôle et nous n’insisterons pas non plus sur l’importance que de tels résultats peuvent revêtir au vu de leurs nombreuses applications. Indiquons juste que de telles méthodes directes sont intéressantes car elles fournissent une expression explicite du terme de contrôle à appliquer. En particulier, elles sont alors souvent plus facile à implémenter d’un point de vue numérique.

### IV.2.1 Un résultat sur des équations paraboliques (dégénérées)

Ce paragraphe expose les résultats de [Benoit et al., 2023]. Nous y avons considéré l’équation parabolique suivante<sup>3</sup>

$$\begin{cases} \varrho(x)\partial_t u - \partial_x(a(x)\partial_x u) = 0 & \text{pour } t \in ]0, T[ \text{ et } x \in ]0, 1[, \\ (a\partial_x u)|_{x=0} = 0, \\ \alpha u|_{x=1} + \beta(a\partial_x u)|_{x=1} = h, \\ u|_{t=0} = u_0, \end{cases} \quad (\text{IV.12})$$

dans laquelle les coefficients  $a$  et  $\varrho$  sont donnés dans des espaces fonctionnels que l’on précisera et où les réels  $\alpha, \beta > 0$ .

Le but est alors de choisir un contrôle  $h$  tel qu’en temps  $T > 0$ , la solution de l’équation (IV.12) soit ramenée à zéro. On a donc un problème de contrôlabilité par le bord *via* une condition de Robin, à zéro en temps fixe. Là encore, la question de la contrôlabilité pour une équation parabolique n’est pas nouvelle et elle constitue même une problématique assez classique. On peut par exemple renvoyer à [Zuazua, 2006] pour des résultats sur ce sujet.

Ce qui a motivé notre étude ici a été de pouvoir permettre au coefficient de diffusion de (IV.12) de venir dégénérer. De tels résultats de contrôle, même dans le cas dégénéré, ont aussi été étudiés dans la littérature. On distingue deux types de dégénérescences : la dégénérescence faible pour laquelle on a typiquement  $a(x) \sim x^{2-\varepsilon}$ , avec  $\varepsilon \in ]1, 2[$ , et la dégénérescence forte pour laquelle  $a(x) \sim x^{2-\varepsilon}$ , pour  $\varepsilon \in ]0, 1[$ .

Concernant les travaux notables sur de tels résultats de contrôle, on renvoie, par exemple (de façon non exhaustive), à [Cannarsa et al., 2021], [Cannarsa et al., 2017], [Cannarsa et al., 2008], et aussi aux travaux de [Fagnelli and Mugnai, 2020], [Fagnelli and Mugnai, 2016]. Le point commun de ces résultats est qu’ils se basent sur des estimations d’observabilité pour le problème dual de (IV.12), *via* des estimations

---

<sup>3</sup>Suivant [Benoit et al., 2023] nous pourrions aussi considérer, au prix de quelques hypothèses supplémentaires pour se ramener à l’équation (IV.12), ajouter dans (IV.12) un terme potentiel. On a, néanmoins, choisi de ne pas considérer ce terme pour simplifier autant que possible les hypothèses.

de Carleman. Le contrôle est alors obtenu par un résultat d'existence (abstrait), à partir de la solution du problème dual de (IV.12). Cette méthode, très satisfaisante d'un point de vue théorique, est toutefois difficile à venir implémenter numériquement, en raison de l'argument abstrait dont est issu le contrôle et elle nécessite de plus, des calculs approchés de la solution du problème dual, ce qui est coûteux numériquement.

On notera aussi que les résultats précédents utilisent bien souvent le fait que l'on ait explicitement une dégénérescence de la forme  $a(x) \sim x^{2-\varepsilon}$  et que les résultats de Fragnelli et Mugnai requièrent aussi des conditions de monotonie et de régularité sur les coefficients. On a voulu [Benoit et al., 2023] pouvoir considérer des coefficients plus génériques, à savoir seulement mesurables.

Dans [Benoit et al., 2023], nous avons de plus essayé de donner une méthode constructiviste du contrôle qui a l'avantage de pouvoir être plus facile à implémenter. Le schéma de démonstration, à savoir l'approche de platitude, que nous avons utilisée dans [Benoit et al., 2023] est dans la lignée directe de celui de [Martin et al., 2016], qui opère sous hypothèse de dégénérescence faible  $\frac{1}{a} \in L^1(]0, 1[)$ . On citera aussi l'article [Moyano, 2016], dans lequel la même méthode est utilisée avec le coefficient de diffusion explicite  $a(x) = x^{2-\varepsilon}$ , ou bien les articles [Martin et al., 2018] et [Martin et al., 2019] pour des exemples d'utilisation de cette méthode à d'autres équations (Schrödinger et Korteweg de Vries).

Dans la suite nous nous placerons sous l'hypothèse suivante, qui caractérise les coefficients de (IV.12) pour lesquels notre résultat de contrôle s'appliquera.

**Hypothèse IV.2.1** *Il existe deux réels  $p \in ]1, \infty[$  et  $r \in ]\frac{p}{p-1}, \infty[$ , tels que les coefficients  $a$  et  $\varrho$  de l'équation d'évolution de (IV.12) vérifient :*

- i) Pour presque tous  $x \in ]0, 1[$ , on a  $a(x), \varrho(x) > 0$ .*
- ii) Le coefficient  $a \in L^1_{loc}(]0, 1[)$  et la fonction  $x \mapsto \frac{x}{a(x)} \in L^p(]0, 1[)$ .*
- iii) Le coefficient  $\varrho \in L^r(]0, 1[)$  et  $\overline{\lim}_{x \rightarrow 0^+} \varrho(x) < \infty$ .*
- iv) On a  $\lim_{x \rightarrow 0^+} \frac{1}{a(x)} \left( \int_x^1 \frac{dy}{a(y)} \right)^{-2} = +\infty$ .*

On remarque, en particulier, que l'Hypothèse IV.2.1 ne requiert ni propriété de monotonie, ni propriété de régularité, ni asymptotique précise de la dégénérescence. On peut aussi vérifier que les points *ii)* et *iv)* sont vérifiés par les fonctions mesurables  $a$  satisfaisant

$$C_1 x^{2-\varepsilon_1} \leq a(x) \leq C_2 x^{2-\varepsilon_2},$$

avec quatre constantes  $C_1, C_2, \varepsilon_1, \varepsilon_2 > 0$ , qui vérifient  $\varepsilon_1 \leq \varepsilon_2 < 2\varepsilon_1 \leq 4$ . Ainsi, on peut effectivement considérer des dégénérescences fortes au sens précédemment mentionné.

L'idée centrale de la méthode de la platitude est de chercher une solution de l'équation parabolique (IV.12) sous la forme découplée (au vu des variables) :

$$u(t, x) \sim \sum_{i=0}^{\infty} y^{(i)}(t) g_i(x). \quad (\text{IV.13})$$

Injecter l'ansatz (IV.13) dans l'équation d'évolution de (IV.12), nous invite à imposer pour les fonctions génératrices,  $g_i$ , les contraintes suivantes<sup>4</sup> :

$$\begin{cases} (ag'_0)' = 0 & \text{d'où } g_0 \equiv 1, \\ (ag'_i)' = \varrho g_{i-1} \quad \forall i \geq 1, \end{cases} \quad (\text{IV.14})$$

---

<sup>4</sup>Pour une fois dans ce mémoire nous ne décrirons pas en détails les conditions de bord.

équations qui peuvent être résolues explicitement par intégrations successives en espace (on renvoie à [[Benoit et al., 2023]-Lemma 3.2]) et on s'attend donc à avoir une borne des  $|g_i|$  en  $\frac{1}{i!R^i}$ . Pour contrebalancer cette borne, il est alors naturel, en vue de donner un sens rigoureux à la série formelle (IV.13), de chercher un terme  $y$  dans la classe de Gevrey,  $G^s$ . Ceci expliquera l'ensemble d'appartenance de notre terme de contrôle.

A ce stade, nous avons donc fixé la dynamique en  $x$  de l'équation dans (IV.13) et on peut toujours jouer sur celle en temps *via* la fonction  $y$ . De plus, l'équation (IV.14) nous invite à étudier en détail le problème elliptique :

$$\begin{cases} (au')' = f & \text{dans } ]0, 1[, \\ (au')(0) = 0 \text{ et } \alpha u(1) + \beta(au')(1) = 0. \end{cases} \quad (\text{IV.15})$$

Le caractère fortement bien posé de ce problème s'obtient en utilisant des inégalités de Hardy dites généralisées (voir [[Benoit et al., 2023]-Lemma 2.1]) qui sont des conséquences directes des travaux de [Trudinger, 1973]. Ce résultat s'applique, car les hypothèses IV.2.1 (plus particulièrement *iii*) font tout pour que l'on puisse l'appliquer. Les autres arguments (théorème de Lax-Milgram essentiellement) sont alors très classiques. Ainsi, notre contribution sur ce point n'est pas fondamentale.

Cependant, on va avoir besoin, pour donner un sens rigoureux à (IV.13), de connaître assez précisément l'asymptotique des couples de valeurs-vecteurs propres<sup>5</sup>  $(\lambda_n, e_n)_{n \in \mathbb{N}}$  du problème spectral associé à (IV.15). On renvoie au Théorème 2.9 de [Benoit et al., 2023], pour des estimations précises. On utilisera dans la suite les deux estimations suivantes : il existe deux constantes  $C_1, C_2 > 0$  telles que

- $\|e_n\|_{L^\infty([0,1])} \leq C_1 \lambda_n^{\kappa_1}$ ,
- pour tout  $n \geq 0$ ,  $\lambda_n \geq C_2 n^{\kappa_2}$ ,

pour un couple  $(\kappa_1, \kappa_2) \in \mathbb{R}_+^2$  dépendant explicitement des paramètres de l'Hypothèse IV.2.1.

C'est sur cette analyse spectrale que l'approche de [Martin et al., 2016] a du être raffinée notamment en modifiant de façon idoine la substitution de Prüfer (on renvoie, par exemple à [Birkhoff and Rota, 1962]) qui donnait ces asymptotiques dans [Martin et al., 2016].

Avec une base de vecteur propre associée à (IV.15), nous aimerions définir, après avoir décomposé la donnée initiale dans cette base  $u_0 := \sum_{n \geq 0} c_n e_n$  ;  $u$  comme l'évolution naturelle de la donnée initiale préparée à la diagonalisation. Plus précisément, on veut poser :

$$u(t, x) := \sum_{n \geq 0} c_n e^{-\lambda_n t} e_n.$$

Pour que cette expression soit compatible avec la forme (IV.13), on a besoin d'exprimer les vecteurs propres  $e_n$  en termes des fonctions génératrices. Puisque ces deux objets vérifient, essentiellement, les mêmes problèmes, une telle relation est attendue et on a plus précisément (voir [[Benoit et al., 2023]-Proposition 3.3])  $e_n = e_n(0) \sum_{i \geq 0} (-\lambda_i)^i g_i$ .

Ainsi, poser

$$\tilde{y}(t) := \sum_{n \geq 0} e_n(0) e^{-\lambda_n t},$$

définition où l'on voit apparaître l'importance de la borne  $\lambda_n \geq C_2 n^{\kappa_2}$ , pour assurer la convergence de la série, donne bien une solution  $u$  de l'équation (IV.12), écrite sous la forme (IV.13). On rappelle, pour conclure, que nous voulons un résultat de contrôlabilité à zéro. On peut modifier la fonction  $\tilde{y}$  précédente, en la multipliant par une fonction de *cut-off*,  $\chi$ , dans la classe de Gevrey vérifiant :

$$\chi(t) := \begin{cases} 1 & \text{pour } t \leq \frac{T}{3}, \\ 0 & \text{pour } t \geq \frac{2T}{3}. \end{cases}$$

<sup>5</sup>Ces derniers existent par le théorème spectral.

On pose alors,

$$y(t) := \chi(t)\tilde{y}(t),$$

c'est-à-dire que l'on laisse l'équation évoluer librement sur l'intervalle  $[0, \frac{T}{3}]$ , avant d'accompagner la solution vers zéro sur les temps  $]\frac{T}{3}, \frac{2T}{3}[$ . Par construction, une fonction  $u$  de la forme (IV.13), vérifie la propriété de contrôlabilité à zéro en temps  $T > 0$ . Le terme de contrôle au bord est dans la classe de Gevrey, il agit sur l'intervalle  $]\frac{T}{3}, \frac{2T}{3}[$  et est "par définition" donné par

$$h(t) := \sum_{i=0}^{\infty} y^{(i)}(t) [\alpha g_i(1) + \beta a g_i'(1)].$$

On conclut alors avec le théorème suivant :

**Théorème IV.2.1** ([Benoit et al., 2023]) *Soit  $a, \rho$  vérifiant l'Hypothèse IV.2.1, pour deux réels  $p \in ]1, \infty[$  et  $r \in ]\frac{p}{p-1}, \infty[$ . Pour toute condition initiale  $u_0 \in L^2_\rho(]0, 1[)$ , pour tout paramètre de Robin,  $\alpha, \beta$ , et pour tout  $s \in ]1, 1 + \frac{p-1}{p} - \frac{1}{r}[$ , il existe un contrôle  $h \in G^s(]0, T[)$ , tel que la solution,  $u$ , de (IV.12) vérifie*

$$u \in \mathcal{C}^0([0, T]; L^2_\rho(]0, 1[)) \cap \mathcal{C}^1([0, T[; W^{1,1}(]0, 1[)) \text{ et } a\partial_x u \in \mathcal{C}^1([0, T[; W^{1,r}(]0, 1[)),$$

ainsi que

$$u(x, T) = 0, \forall x \in [0, 1].$$

Dans ce théorème, on rappelle que  $G^s$  désigne l'espace de Gevrey d'ordre  $s > 0$ , et on a noté  $L^2_\rho$  l'espace  $L^2$  à poids  $\rho$ .

## IV.2.2 Contrôle pour l'équation des ondes

Le paragraphe suivant expose un second résultat de contrôle, résultat que nous avons obtenus lors d'une collaboration avec Lionel Rosier et Romain Loyer lors du co-encadrement de la thèse de ce dernier. Ce travail fait actuellement l'objet d'un *preprint* et sera très prochainement soumis à publication [Be et al., ].

Comme au paragraphe précédent, nous nous sommes ici intéressé à un résultat de contrôle à zéro en temps fini *via* le bord du domaine de résolution. Là encore, nous avons cherché à obtenir un résultat de contrôle basé sur une méthode constructiviste directe et non pas sur une méthode reposant sur l'observabilité. Comme on va le voir ceci permettra en particulier de faciliter l'implémentation numérique du contrôle.

Toutefois contrairement au paragraphe précédent, on ne considère plus cette fois une équation de type parabolique, mais on revient au cadre hyperbolique en venant considérer l'équation des ondes.

Les questions qui vont nous intéresser en détails dans ce qui suit vont être, l'existence théorique d'un terme de contrôle au bord en géométrie "quelconque" et l'implémentation numérique d'un tel contrôle. Plus précisément, le problème de contrôle considéré est le suivant : on se donne  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 1$  un ouvert borné et on fixe  $\Gamma_0 \subset \partial\Omega$  une partie du bord sur laquelle on va se permettre d'agir. On cherche alors un terme de contrôle  $h \in L^2(]0, T[ \times \Omega)$  tel que la solution de l'équation des ondes :

$$\begin{cases} \square u := \partial_t^2 u - \Delta u = 0 & \text{dans } ]0, T[ \times \Omega, \\ u|_{\partial\Omega} := h \mathbf{1}_{\Gamma_0} & \text{sur } ]0, T[ \times \partial\Omega, \\ (u, \partial_t u)|_{t=0} = (u_0, u_1) & \text{sur } \Omega, \end{cases} \quad (\text{IV.16})$$

où les sources  $(u_0, u_1) \in L^2(\Omega) \times H^{-1}(\Omega)$  sont données, vérifie qu'il existe  $T > 0$  tel que

$$\forall x \in \Omega, (u, \partial_t u)|_{t=T} = (0, 0).$$

Une telle question est, on peut le dire, assez classique et fait l'objet d'une littérature très abondante. De façon loin d'être exhaustive, on renverra à l'article pionnier [Bardos et al., 1992] dans lequel les auteurs introduisent la condition de contrôle dite géométrique. Condition ensuite reprise et largement étudiée par de nombreux auteurs, comme par exemple [Burq and Gérard, 1997] et [Le Rousseau et al., 2017]. Cette approche qui utilise fortement l'analyse microlocale, se base sur l'observabilité du problème dual et est donc une méthode indirecte.

L'approche qui a été proposée de suivre à Romain Loyer dans sa thèse était une méthode directe, qui insistons sur ce point est constructiviste et est fortement inspirée de l'article [Rosier and Rosier, 2022]. L'idée centrale de la méthode de Rosier et Rosier est de se baser sur le principe de Huygens dont on rappelle un énoncé classique dans l'espace entier  $\mathbb{R}^{1+3}$ :

**Proposition IV.2.1 (Principe de Huygens)** *Si  $v \in \mathcal{C}^\infty(\mathbb{R}^{1+3})$  vérifie l'équation des ondes*

$$\begin{cases} \square v = 0 & \text{dans } \mathbb{R}^{1+3}, \\ (v, \partial_t v)|_{t=0} = (v_0, v_1) & \text{sur } \mathbb{R}^3, \end{cases}$$

et si on note

$$\Theta(t) := \text{supp } v(t, \cdot) \cup \text{supp } \partial_t v(t, \cdot),$$

l'ensemble qui contient l'information non triviale liée à  $v$ . Alors l'information initiale se propage le long des rayons d'optique géométrique. Plus précisément on a l'inclusion

$$\Theta(t) \subset \{x \in \mathbb{R}^3 \mid \text{dist}(x, \Theta(0)) = t\}.$$

En particulier, si les sources  $v_0$  et  $v_1$  sont à support compact dans  $K \subset \mathbb{R}^3$ , et si on prend  $T > \text{diam}(K)$  alors

$$\forall x \in K, t \geq T, v(t, x) = \partial_t v(t, x) = 0.$$

Ce principe, propre aux dimensions impaires, peut-être résumé informellement de la façon suivante, l'onde se propage en suivant les rayons de l'optique géométrique et elle passe une seule et unique fois par un point  $x$  donné. En particulier, si on se place sur une fenêtre d'étude fixée et que l'on attend assez longtemps, l'onde sera passée puis sortie de notre fenêtre d'observation et elle n'y reviendra plus.

Une conséquence directe d'un tel principe est alors le résultat de contrôle suivant :

**Théorème IV.2.2** *On se place sous des hypothèses sur l'ouvert  $\Omega$  assurant que le principe de Huygens de la Proposition IV.2.1 soit vérifié. Soit  $\Omega' \subset \Omega$  un ouvert tel que  $\partial\Omega' = \Gamma := \Gamma_0 \cup (\Gamma \setminus \Gamma_0)$  avec  $\Gamma \setminus \Gamma_0 \subset \partial\Omega$  et*

- *il contient l'information initiale non triviale, i.e.  $\text{supp } u_0 \cup \text{supp } u_1 \subset \overline{\Omega'}$ .*
- *il existe un temps  $T > 0$  pour lequel les rayons d'optique géométrique quittent  $\Omega'$ , i.e.*

$$\forall x \in \Omega', \forall z \in \mathbb{S}^2 \text{ on a } \mathbf{x}(T, x, z) \notin \overline{\Omega'},$$

où  $\mathbf{x}(t, x, z)$  désigne le rayon d'optique géométrique issu du point  $x$ , partant selon la direction  $z$  au temps  $t$ .

Alors la restriction  $v$  de  $u$  la solution de (IV.16) à  $\Omega'$  vérifie d'une part

$$\begin{cases} \square v = 0 & \text{dans } ]0, T[ \times \Omega', \\ v|_{\partial\Omega'} = v \mathbf{1}_{\Gamma_0} & \text{sur } ]0, T[ \times \Gamma, \\ (v, \partial_t v)|_{t=0} = (u_{0|\Omega'}, u_{1|\Omega'}) & \text{sur } \Omega', \end{cases}$$

et la condition terminale

$$(v, \partial_t v)|_{t=T} = (0, 0).$$



La question est alors ramenée à établir des conditions sur la géométrie de  $\Omega$  assurant que l'on ait le principe de Huygens et donc le résultat de contrôle. Cette question a d'abord été adressée dans [Rosier and Rosier, 2022] lorsque  $\Omega$  est un demi-espace, on a alors le principe de Huygens et le résultat de contrôle idoine. Résultat que nous avons ensuite généralisé dans le *preprint* [Be et al., ] au cas où  $\Omega$  est borné, convexe de classe  $\mathcal{C}^\infty$  et qui vérifie l'Hypothèse IV.2.2, que nous discuterons plus en détails dans ce qui suit.

Le résultat exposé ici se trouve dans la lignée directe de celui de Rosier et Rosier décrit dans [Rosier and Rosier, 2022]. En effet, il suit le même schéma de démonstration. Dans cet article, les auteurs déduisent le principe de Huygens d'une formule de représentation de Kirchhoff : si  $u$  désigne la solution de l'équation des ondes à données initiales  $(u_0, u_1)$  alors on peut écrire (on renvoie à l'article [[Rosier and Rosier, 2022]-Paragraph 2.2])

$$u(t, x) = \frac{t}{4\pi} \int_{\mathbb{S}^2} u_1(t, \mathbf{x}(t, x, z)) \varepsilon(t, x, z) d\sigma(z) + \partial_t \left( \frac{t}{4\pi} \int_{\mathbb{S}^2} u_0(t, \mathbf{x}(t, x, z)) \varepsilon(t, x, z) d\sigma(z) \right), \quad (\text{IV.17})$$

où la fonction  $\varepsilon$  est une fonction de comptage qui vaut un sur le front incident et moins un sur le front réfléchi lors d'une réflexion contre le bord du demi-espace. On a alors bien une généralisation de la formule de Kirchhoff ( qui correspond au cas  $\varepsilon \equiv 1$ ).

Dans le cas d'un ouvert convexe borné quelconque, la formule (IV.17) se généralise. Toutefois le devenir d'un rayon d'optique géométrique issu d'un point  $x$  et suivant une direction  $z$  est plus riche que dans le cas du demi-espace. En effet, ce rayon peut (et même va) subir plus d'une réflexion sur le bord du domaine de résolution.

On démontre toutefois dans [Be et al., ] que la formule de Kirchhoff (IV.17) se généralise dans une telle géométrie si l'on modifie la fonction de comptage  $\varepsilon$  en posant  $\varepsilon(t, x, z) := (-1)^{N(t, x, z)}$ , où  $N(t, x, z)$  désigne le nombre de réflexions contre le bord du domaine, subies par le rayon d'optique géométrique  $\mathbf{x}(\cdot, x, z)$  sur la fenêtre de temps  $[0, t]$ .

En particulier, pour assurer que la nouvelle fonction de comptage  $\varepsilon$  soit bien-définie, nous nous sommes placés sous l'hypothèse suivante (cette dernière est automatiquement vérifiée dans le cas du demi-espace)

**Hypothèse IV.2.2** *On suppose qu'il existe  $T > 0$  tel que les rayons d'optique géométrique admettent un nombre fini de réflexions sur la fenêtre de temps  $[0, T]$ . Plus précisément on impose*

$$\sup_{z \in \mathbb{S}^2} \#\{t \in [0, T], x \in \Omega \setminus \mathbf{x}(t, x, z) \in \partial\Omega\} < \infty.$$

Le point noir de l'Hypothèse IV.2.2 est qu'il ne s'agit pas directement d'une hypothèse sur la géométrie du domaine de résolution<sup>6</sup>. Donner une caractérisation plus géométrique de l'Hypothèse IV.2.2 constitue une perspective de ce travail.

Une fois un tel résultat d'existence théorique d'un contrôle démontré, on propose alors dans [Be et al., ] une méthode donnant des résultats numériques qualitativement assez satisfaisants. Nous n'entrerons pas dans des détails trop techniques, mais les idées de cette approche sont les suivantes (on suppose ici sans grande perte de généralité qu'il existe une fonction régulière  $f$  telle que  $\Omega \subset \{x_3 > f(x_1, x_2)\}$ ):

1. On ramène l'équation des ondes (IV.16) à la géométrie de base  $\tilde{\Omega} := ]-1, 1[^3$  par changement de

---

<sup>6</sup>On montre toutefois dans [Be et al., ] que l'Hypothèse IV.2.2 est une conséquence de la stricte convexité en dimension deux (mais on n'a alors plus le principe de Huygens) et qu'elle est vérifiée pour la sphère unité  $\mathbb{S}^2$ .

variable (dépendant explicitement de  $f$ ). On a donc à résoudre numériquement l'équation des ondes:

$$\begin{cases} \partial_{tt}^2 u - (\partial_{y_1 y_1}^2 + \partial_{y_2 y_2}^2 + a \partial_{y_3 y_3}^2 - b \partial_{y_1 y_3}^2 - c \partial_{y_2 y_3}^2 - d \partial_{y_3}^2) u = 0 & \text{pour } t \in ]0, T[ \text{ et } y \in \tilde{\Omega}, \\ u|_{\partial \tilde{\Omega}} = 0 \\ (u, \partial_t u)|_{t=0} = (u_0, u_1) \end{cases} \quad \text{pour } y \in \tilde{\Omega}, \quad (\text{IV.18})$$

où les coefficients  $a$ ,  $b$ ,  $c$  et  $d$  sont explicites en termes de  $f$  et de ses dérivées.

2. On approche alors numériquement la solution de (IV.18) par un schéma. Dans [Be et al., ] nous avons discrétisé la dérivée temporelle en utilisant un schéma de saute-mouton et une discrétisation de Chebyshev pour les dérivées spatiales. Bien sûr, le choix des méthodes de discrétisation lors de cette seconde étape n'est pas ici unique. Le choix fait ici s'avère être motivé par le fait qu'il généralise assez naturellement l'analyse de Rosier et Rosier [Rosier and Rosier, 2022] qui s'appuie sur des expressions exactes données par des séries de Fourier tronquées.

On notera ici que contrairement à [Rosier and Rosier, 2022], l'analyse théorique de la convergence et celle de l'erreur n'ont pas été entreprises. Toutefois, comme la méthode de contrôle est directe, on obtient (du moins qualitativement sur des exemples numériques) de bons résultats de contrôle. Ceci sans avoir à mettre en œuvre des méthodes numériques assez élaborées.

### IV.3 Un dernier exemple de développement d'optique géométrique pour conclure

Dans ce paragraphe, nous exposerons un travail de collaboration actuellement en cours avec Franck Sueur. A l'heure de la rédaction de ce mémoire, les résultats décrits dans ce paragraphe ne constituent donc pas encore un théorème à proprement parler et des points techniques restent à être clarifiés dans les démonstrations. Toutefois, nous avons bon espoir que ces derniers le soient rapidement et ainsi, de pouvoir soumettre à publication ce travail prochainement.

Nous avons tenu à mentionner ces travaux, car ils nous semblent une belle illustration de ce que peuvent être les utilisations des développements d'optique géométrique dont on a (amplement) parlé dans le Chapitre II. De plus, ce travail en cours donne aussi lieu à quelques perspectives intéressantes.

Pour varier un peu des problèmes linéaires, nous avons considéré cette fois un problème hyperbolique quasi-linéaire. Ce dernier étant posé dans l'espace entier  $\mathbb{R}^d$ . Plus précisément, on a considéré le problème de Cauchy : pour  $T > 0$ ,  $d \in \mathbb{N}^*$ ,

$$\begin{cases} \mathcal{H}(u)u := \mathcal{H}(t, x, u)u := A_0(t, x, u)\partial_t u + \sum_{j=1}^d A_j(t, x, u)\partial_j u = 0, & \text{pour } (t, x) \in ]0, T[ \times \mathbb{R}^d, \\ u|_{t=0} = u_0, \end{cases} \quad \text{pour } x \in \mathbb{R}^d, \quad (\text{IV.19})$$

où l'on suppose que les coefficients  $A_j$  dépendent de façon régulière de leurs variables, sont à valeurs dans l'ensemble des matrices  $\mathbf{M}_{N \times N}(\mathbb{R})$ ,  $N \geq 1$ , et sont symétriques. Enfin dans le problème précédent, on suppose de surcroît que la matrice  $A_0$  est définie positive, pour que la variable  $t$  joue bien le rôle d'un temps. La donnée initiale  $u_0$  est régulière et donnée.

On introduit la notation suivante pour le symbole de  $\mathcal{H}$  : pour  $(t, x, u, \xi) \in ]0, T[ \times \mathbb{R}^d \times \mathbb{R}^N \times \mathbb{R}^{1+d}$

$$\mathbf{A}(u, \xi) := \mathbf{A}(t, x, u, \xi) := \xi_0 A_0(t, x, u) + \underbrace{\sum_{j=1}^d \xi_j A_j(t, x, u)}_{:= \mathbf{A}^\dagger(t, x, u, \zeta)} \quad \text{où } \xi := (\xi_0, \zeta) \quad (\text{IV.20})$$

Pour avoir une chance que le problème soit bien posé, on se place l'hypothèse d'hyperbolicité à multiplicité constante suivante :

**Hypothèse IV.3.1** *Les valeurs propres de  $A_0^{-1}(t, x, u)\mathbf{A}^\dagger(t, x, u, \zeta)$  sont de multiplicités constantes ; ces multiplicités sont indépendantes des variables  $(t, x, u, \xi)$ .*

L'existence de solutions pour des problèmes hyperboliques quasi-linéaires a fait l'objet d'une littérature abondante. On arrive bien souvent à construire des solutions de (IV.19) qui sont régulières de chaque côté d'une interface de transition  $\Gamma \subset [0, T] \times \mathbb{R}^d$ . La solution peut alors, par exemple, présenter une discontinuité le long de  $\Gamma$ , c'est le cas d'un choc on renvoie par exemple à [Guès et al., 2010] ou [Guès et al., 2005b].

Dans ce qui suit, nous nous sommes intéressé à des singularités plus faibles, les ondes soniques de Métivier [Métivier, 1991]. De telles solutions sont régulières le long de  $\Gamma$ , mais c'est leur gradient qui présente un saut le long de  $\Gamma$ . Plus précisément, le théorème d'existence est le suivant :

**Théorème IV.3.1 (Métivier [Métivier, 1991])** *On se place sous l'Hypothèse IV.3.1. Dans ce qui suit, on note  $\lambda$  l'une des valeurs propres de  $A_0^{-1}\mathbf{A}^\dagger$ . Alors il existe  $T > 0$ , une hypersurface régulière  $\Gamma$  de  $[0, T] \times \mathbb{R}^d$  et une solution faible  $u^0$  de (IV.19) tels que :*

- i) *il existe une fonction de phase  $\varphi$  dont le lieu d'annulation décrit  $\Gamma$ . De plus, elle vérifie  $\nabla\varphi \neq 0$  sur  $\Gamma$  et est solution de l'équation eikonale*

$$\partial_t\varphi + \lambda(u^0, \nabla_x\varphi) = 0. \quad (\text{IV.21})$$

- ii) *La solution  $u^0$  est continue sur  $\mathbb{R}^d$ . En particulier, elle l'est à travers  $\Gamma$ , mais  $\partial_n u^0$ , la dérivée normale à  $\Gamma$ , présente un saut,  $[\partial_n u^0]$ , non trivial, à travers  $\Gamma$ . Ce dernier vérifie la condition de polarisation*

$$\mathbf{A}(u^0, \nabla_{t,x}\varphi) [\partial_n u^0] = 0. \quad (\text{IV.22})$$

Avec Franck Sueur, nous nous sommes intéressés à l'obtention de la solution onde sonique de Métivier comme la limite visqueuse de solutions d'équations paraboliques régularisées, dont la solution a de meilleures propriétés de régularité sur le gradient. De telles méthodes de régularisation ont déjà été mises en œuvre avec succès dans [Guès et al., 2010], [Guès et al., 2005b] ou bien [Blochas and Rodrigues, 2022] à propos des chocs, ou encore dans [Sueur, 2006] pour l'analogie semi-linéaire de (IV.19). On citera aussi le travail de [Guès, 1995] qui établit le développement asymptotique de la solution en faisant intervenir des couches limites pour pallier le changement de nature du problème dans la limite visqueuse.

On se donne alors un opérateur elliptique,  $\alpha$ -coercif,  $\alpha > 0$  :

$$\mathcal{E}(u)u := \mathcal{E}(t, x, u)u := \sum_{i,j=1}^d \partial_i(E_{i,j}(t, x, u)\partial_j u),$$

où les coefficients  $E_{i,j}$  sont réguliers à valeurs dans l'ensemble des matrices  $\mathbf{M}_{N \times N}(\mathbb{R})$ . Pour  $0 < \varepsilon \ll 1$ , on considère, alors, les solutions de l'équation parabolique

$$\begin{cases} \mathcal{H}(u^\varepsilon)u^\varepsilon - \varepsilon\mathcal{E}(u^\varepsilon)u^\varepsilon = 0, & \text{pour } (t, x) \in ]0, T^\varepsilon[ \times \mathbb{R}^d, \\ u^\varepsilon|_{t=0} = u_0, & \text{pour } x \in \mathbb{R}^d. \end{cases} \quad (\text{IV.23})$$

On cherche, alors, des solutions approchées de (IV.23) sous la forme d'un développement de type optique géométrique, en termes de puissance de  $\varepsilon$ , dont le premier terme est  $u^0$ , la solution du Théorème IV.3.1. On va de plus, chercher à imposer des propriétés de régularité de la solution approchée et de sa dérivée normale, le long de l'interface  $\Gamma$ .

Dans ce qui suit, on précisera les principales difficultés avec lesquelles on va devoir composer et on prendra un soin particulier à décrire les différentes amplitudes qui composent notre ansatz. Les constructions rigoureuses de certaines des amplitudes comme l'analyse d'erreur sont encore à l'étude et ne seront pas abordées. Ainsi, les paragraphes suivants doivent être vus comme une justification formelle et heuristique.

### IV.3.1 Première approximation "naturelle"

On pense que l'amplitude principale déterminant le comportement de  $u^\varepsilon$  la solution de (IV.23) doit être donnée par une fonction des variables lentes  $(t, x)$ . Puisque nous sommes dans un cadre caractéristique, on s'attend à un correcteur de type couche limite à l'échelle  $\sqrt{\varepsilon}$  comme dans [Guès, 1995], [Rousset, 2005] et [Sueur, 2006]. On postule donc comme première approximation

$$u^\varepsilon(t, x) := \underline{u}(t, x) + \sqrt{\varepsilon} \mathcal{U}(t, x; \underbrace{\frac{\varphi(t, x)}{\sqrt{\varepsilon}}}_{:=X}), \quad (\text{IV.24})$$

où les profils  $\underline{u}$  et  $\mathcal{U}$ , et la fonction de phase  $\varphi$  sont à déterminer.

Il est à noter qu'ici nous avons choisi de ne pas redresser le front  $\Gamma$  pour rendre l'analyse la plus directe possible. En injectant cet ansatz dans (IV.23), en identifiant les différentes puissances de  $\varepsilon$ , on obtient la cascade d'équations :

$$\begin{aligned} \mathcal{H}(u^\varepsilon)u^\varepsilon - \varepsilon \mathcal{E}(u^\varepsilon)u^\varepsilon & \quad (\text{IV.25}) \\ = \varepsilon^0 (\mathcal{H}(\underline{u})\underline{u} + \mathbf{A}(\underline{u}, \nabla_{t,x}\varphi)\partial_X \mathcal{U}) + \sqrt{\varepsilon} (L(\partial)\mathcal{U} + L_0\mathcal{U} + \mathcal{N}\mathcal{L}(\mathcal{U})\mathcal{U} - \mathbf{E}\partial_X^2 \mathcal{U}) + \text{h.o.t} \end{aligned}$$

où l'on a défini les opérateurs

$$\begin{aligned} L(\partial)\mathcal{U} & := L(t, x, \underline{u}, \partial)\mathcal{U} := \sum_{j=0}^d A_j(t, x, \underline{u})\partial_j \mathcal{U} \\ \mathcal{N}\mathcal{L}(\mathcal{U})\mathcal{U} & := \mathcal{N}\mathcal{L}(t, x, \underline{u}, \mathcal{U})\mathcal{U} = \sum_{j=0}^d (\partial_j \varphi) d_u A_j(t, x, \underline{u}) \cdot \mathcal{U} \partial_X \mathcal{U} \end{aligned}$$

et les matrices

$$L_0\mathcal{U} := L_0(t, x, \underline{u})\mathcal{U} := \sum_{j=0}^d d_u A_j(t, x, \underline{u}) \cdot \mathcal{U} \partial_j \underline{u}, \text{ et } \mathbf{E} := \mathbf{E}(t, x, \underline{u}) := \sum_{i,j=1}^d (\partial_i \varphi)(\partial_j \varphi) E_{i,j}(t, x, \underline{u})$$

Ainsi, sans surprise, pour annuler le terme en  $\varepsilon^0$  dans le membre de droite de (IV.25), on pose pour  $\underline{u}$  la solution du problème hyperbolique donnée par le Théorème IV.3.1 de Métivier. De la même manière, on prend pour la fonction de phase  $\varphi$  celle donnée par le Théorème IV.3.1, pour avoir  $\det \mathbf{A}(u^0, \nabla_{t,x}\varphi) = 0$ , d'après (IV.21). Par conséquent, la couche limite vérifie la condition de polarisation

$$\partial_X \mathcal{U} \in \ker \mathbf{A}(u^0, -\lambda, \nabla_x \varphi).$$

On note  $\Pi$  le projecteur<sup>7</sup> sur  $\ker \mathbf{A}(u^0, -\lambda, \nabla_x \varphi)$ . En appliquant  $\Pi$  au terme d'ordre  $\varepsilon^{1/2}$  de (IV.25), on veut donc trouver  $\mathcal{U} = \Pi \mathcal{U}$ , polarisée, telle que

$$\Pi L(\partial)\Pi \mathcal{U} + \Pi L_0 \Pi \mathcal{U} + \Pi \mathcal{N}\mathcal{L}(\Pi \mathcal{U})\Pi \mathcal{U} - \Pi \mathbf{E} \partial_X^2 \Pi \mathcal{U} = 0 \text{ dans } \Omega_T \times \mathbb{R}. \quad (\text{IV.26})$$

<sup>7</sup>Cette partie étant formelle, on se sent libre ici de ne pas donner plus de détails sur la définition de ce projecteur.

On applique alors le lemme de Lax classique [Lax, 1957] à  $\Pi L(\partial)\Pi$ . Le premier terme devient un opérateur de transport. Ce lemme admet aussi une variante pour le terme non linéaire  $\Pi \mathcal{N} \mathcal{L}(\Pi \mathcal{U})$ , qui permet de le ramener à une non-linéarité<sup>8</sup> de type Burgers.

On rappelle que notre but est ici de construire une solution approchée  $u^\varepsilon$  qui vérifie des conditions de raccord le long du front  $\Gamma$ . La continuité de  $u^\varepsilon$  à travers  $\Gamma$ , nous pousse à imposer

$$0 = [u^\varepsilon] = \underbrace{[u^0]}_{:=0} + \varepsilon^{1/2} [\mathcal{U}].$$

Imposer la continuité de la dérivée normale donne

$$[\partial_n u^\varepsilon] = \underbrace{[\partial_n u^0]}_{\neq 0} + |(\nabla_{t,x} \varphi)|_\Gamma|^2 [\partial_X \mathcal{U}] + \varepsilon^{1/2} [\partial_n \mathcal{U}],$$

où  $f_\Gamma$  désigne l'évaluation de la fonction  $f$  en  $\Gamma$ .

Pour avoir une régularisation de la dérivée normale, on a alors envie d'imposer l'égalité pour venir annuler le terme en  $\varepsilon^0$ . On note que ceci est possible, car le saut  $[\partial_n u^0]$  se trouve polarisé d'après (IV.22). Donc, on peut bien imposer une telle condition de bord au terme polarisé  $\partial_X \mathcal{U}$  et on obtient les conditions de transition suivantes à l'interface (on remarque que cette dernière s'exprime *via*  $X = 0$ )

$$\begin{cases} [\Pi \mathcal{U}]|_{X=0} = 0, \\ [\Pi \partial_X \mathcal{U}]|_{X=0} = -\frac{1}{|(\nabla_{t,x} \varphi)|_\Gamma|^2} [\partial_n u^0], \end{cases} \quad (\text{IV.27})$$

conditions compatibles avec le fait que l'équation régissant  $\mathcal{U}$ , à savoir (IV.26), est d'ordre deux en  $X$ . Ainsi, on cherchera dans ce qui suit une solution de (IV.26) sous la forme de deux états  $\mathcal{U}_\pm$ , définis pour  $\{\pm X > 0\}$  et reliés par les conditions à l'interface. Le fait que l'on ait un saut non trivial de la dérivée normale de  $u^0$  allume alors un terme source pour ces équations et les  $\mathcal{U}_\pm$  associées sont non nulles.

L'approximation (IV.24) semble donc être une solution approchée qui admet un saut de la dérivée normale en  $O(\varepsilon^{1/2})$ . Elle est donc plus régulière.

### IV.3.2 Quand continuer sur une bonne lancée conduit à l'accident

Fort des résultats du paragraphe précédent, un choix naturel pour poursuivre le développement et de tâcher d'améliorer l'ordre d'erreur et la régularisation du saut de la dérivée normale, est de poser

$$u^\varepsilon(t, x) := u^0(t, x) + \sqrt{\varepsilon} \mathcal{U}(t, x; \frac{\varphi(t, x)}{\sqrt{\varepsilon}}) + \varepsilon \left( u^2(t, x) + \mathcal{U}^2(t, x; \frac{\varphi(t, x)}{\sqrt{\varepsilon}}) \right) \quad (\text{IV.28})$$

où les deux premières amplitudes ont été déterminées au paragraphe précédent. Les amplitudes correctrices  $u^2$  et  $\mathcal{U}^2$  sont maintenant à déterminer. En injectant cet ansatz dans l'équation d'évolution de (IV.23), on obtient le développement:

$$\begin{aligned} & \mathcal{H}(u^\varepsilon)u^\varepsilon - \varepsilon \mathcal{E}(u^\varepsilon)u^\varepsilon \\ &= \varepsilon^{1/2} (\mathbf{A}(\nabla_{t,x} \varphi) \mathcal{U}^2 + L(\partial) \mathcal{U} + L_0 \mathcal{U} + \mathcal{N} \mathcal{L}(\mathcal{U}) \mathcal{U} - \mathbf{E} \partial_X^2 \mathcal{U}) \\ & \quad + \varepsilon \left( L(\partial)(u^2 + \mathcal{U}^2) + \mathcal{N} \mathcal{L}(\mathcal{U}^1) \partial_X \mathcal{U}^2 + \tilde{L}_0(u^2 + \mathcal{U}^2) - \mathbf{E} \partial_X^2 \mathcal{U}^2 - \mathbf{f}(u^0, \mathcal{U}^1) \right) + \text{h.o.t} \end{aligned} \quad (\text{IV.29})$$

<sup>8</sup>Il est intéressant de remarquer ici que dans le cas où l'on travaille avec une valeur propre linéairement dégénérée (voir [[Serre, 1996a]-Definition 2.3.3]), alors cette non-linéarité s'annule. On pourra, alors, résoudre l'équation (IV.26) en tous temps.

où  $\tilde{L}_0 v := \tilde{L}_0(u^0)v := \sum_{j=0}^d d_u A_j(u^0) \cdot v (\partial_j u^0 + \partial_j \partial_X \mathcal{U})$  et où  $\mathbf{f}(\cdot, \cdot)$  dépend de façon explicite de ses variables.

On choisit alors la partie non polarisée de  $\mathcal{U}^2$  de manière à ce que cette dernière vienne compenser la partie non polarisée du reste du terme en  $\varepsilon^{1/2}$ . Ceci se fait en introduisant un pseudo-inverse pour  $\mathbf{A}(u^0, -\lambda, \nabla_x \varphi)$  et détermine de façon unique la partie non polarisée  $(I - \Pi)\mathcal{U}^2$ . Le point important est ici que, puisque l'on a déterminé  $\mathcal{U}$  de manière à venir annuler la partie polarisée du terme précédemment mentionné, on annule ainsi tout le terme en  $\varepsilon^{1/2}$ .

A ce stade, on peut encore jouer sur la partie polarisée  $\Pi\mathcal{U}^2$  et sur  $u^2$ . Pour les déterminer, découpler les variables lentes et rapides, dans le terme d'ordre  $\varepsilon$ , donne alors les deux systèmes linéaires:

$$L(\partial)u^2 + \tilde{L}_0 u^2 = \mathbf{f}_\#(u^0)$$

et

$$\Pi L(\partial)\Pi\mathcal{U}^2 + \Pi \mathcal{N} \mathcal{L}(\mathcal{U}^1)\partial_X \Pi\mathcal{U}^2 + \Pi \tilde{L}_0 \Pi\mathcal{U}^2 - \Pi \mathbf{E} \partial_X^2 \Pi\mathcal{U}^2 = \mathbf{f}_\#(u^0, \mathcal{U}) + \mathbf{g}((I - \Pi)\mathcal{U}^2),$$

où  $\mathbf{f}_\#$  et  $\mathbf{f}_\#$  correspondent à la décomposition de  $\mathbf{f}$  selon les variables lentes et rapides et où  $\mathbf{g}$  est explicite.

Ainsi, le schéma itératif se met en place. En effet, on remarque que la partie polarisée de  $\mathcal{U}^2$  vérifie le même système que  $\mathcal{U}^1$ . La seule différence est que ce système est maintenant linéaire, on n'aura donc pas de restriction en temps pour déterminer les correcteurs.

On étudie maintenant les conditions le long du front. On a pour la condition de continuité :

$$0 = [u^\varepsilon] = \varepsilon [u^2] + \varepsilon [\mathcal{U}^2],$$

que l'on décompose en

$$\begin{cases} [(I - \Pi)u^2] = - [(I - \Pi)\mathcal{U}^2], \\ [\Pi\mathcal{U}^2] = - [\Pi u^2], \end{cases} \quad (\text{IV.30})$$

où l'on rappelle que le membre de droite de la première équation est connu en termes de  $u^0$  et  $\mathcal{U}$ . On peut donc résoudre en premier  $u^2$  en l'exprimant comme la juxtaposition de deux états  $u_\pm^2$  chacun solution du problème aux limites linéaire dans le "demi-espace":

$$\begin{cases} L(\partial)u_\pm^2 + \tilde{L}_0 u_\pm^2 = \mathbf{f}_\#(u^0)_\pm & \text{dans } \{\pm\varphi > 0\}, \\ \Pi_\mp u_\pm^2|_{\varphi=0} = -\Pi_\mp g & \text{sur } \{\varphi = 0\}, \\ u_{|t=0}^2 = 0, & \text{sur } \Omega_T, \end{cases} \quad (\text{IV.31})$$

problème qui a été amplement étudié dans le Chapitre II.

Dans l'équation précédente, on a utilisé le fait que  $I - \Pi$  se décompose sous la forme  $\Pi_+ + \Pi_-$ , où  $\Pi_\pm$  est le projecteur sur le sous-espace propre associé aux valeurs propres  $\lambda_k$  telles que  $\pm(\lambda_k - \lambda) > 0$ . En particulier, on a écrit  $(I - \Pi)\mathcal{U} := \Pi_+ g + \Pi_- g$ .

On remarque alors que le problème aux limites (IV.31) est alors bien posé, car il tombe dans le cadre des systèmes à conditions de bord strictement dissipatives. En effet, on propage "dans le bon sens" la partie de la trace "bien choisie" (plus précisément des modes liés à des valeurs propres négatives (resp. positives) pour  $\varphi > 0$  (resp.  $\varphi < 0$ )).

Ceci détermine en particulier  $[\Pi u^2]$  et par conséquent le membre de droite de la seconde équation de (IV.30). Ceci donne la première condition de bord pour déterminer  $\Pi\mathcal{U}^2$ .

Pour déterminer la seconde, on tente de résoudre la condition de continuité sur la dérivée normale. On a

$$[\partial_n u^\varepsilon] = \varepsilon^{1/2} ([\partial_n \mathcal{U}] + |(\nabla_{t,x} \varphi)_{|\Gamma}|^2 [\partial_X \mathcal{U}^2]) + \varepsilon ([\partial_n u^2] + [\partial_n \mathcal{U}^2])$$

et en décomposant selon les parties polarisée et non polarisée dans le coefficient en  $\varepsilon^{1/2}$ , on voit les limites de notre analyse.

En effet,  $\mathcal{U}^2$  n'étant plus polarisée, on ne peut pas venir annuler le terme  $(I - \Pi)\partial_X \mathcal{U}^2$  même si l'on impose la condition sur la partie polarisée

$$[\Pi \partial_X \mathcal{U}^2] = -\frac{1}{|(\nabla_{t,x} \varphi)_{|\Gamma}|^2} [\Pi \partial_n \mathcal{U}],$$

qui nous donne la condition manquante pour résoudre  $\Pi \mathcal{U}^2$ . En particulier,  $\mathcal{U}^2$ , est alors complètement déterminée.

Donc l'ansatz (IV.28) donne une solution approchée avec un saut de la dérivée normale en  $O(\varepsilon^{1/2})$ , ce qui ne constitue pas une amélioration par rapport à l'ansatz (IV.24). Dans le paragraphe qui suit, on va voir comment la prise en compte des couches limites non caractéristiques permet de venir améliorer cette régularité de la dérivée normale.

### IV.3.3 Apparition des correcteurs non caractéristiques

Comme indiqué précédemment, l'ansatz (IV.28) n'est pas assez fin pour obtenir une bonne régularité de la dérivée normale. Une façon naturelle d'obtenir une meilleure régularité est de venir prendre en compte des profils de couche limite non caractéristique. En effet, on cherche à corriger (IV.28) de la façon suivante

$$u^\varepsilon(t, x) := u^0(t, x) + \varepsilon^{1/2} \mathcal{U}(t, x; \frac{\varphi}{\sqrt{\varepsilon}}) + \varepsilon \left( u^2(t, x) + \mathcal{U}^2(t, x; \frac{\varphi}{\sqrt{\varepsilon}}) \right) + \varepsilon^\alpha \mathcal{V}(t, x; \frac{\varphi}{\varepsilon^\beta}), \quad (\text{IV.32})$$

où  $\mathcal{V}$  est une nouvelle amplitude correctrice, dont le saut normal à vocation à être non polarisé, et où les exposants  $\alpha$  et  $\beta$  sont choisis de telle sorte que

1. L'ajout du terme  $\mathcal{V}$  n'influe pas sur les conditions de continuité de  $u^\varepsilon$  que l'on a déjà établies. Ceci impose  $\alpha \geq \frac{3}{2}$ . Le choix naturel consiste à poser  $\alpha = \frac{3}{2}$ .
2. Dans la condition de régularité de la dérivée normale, le terme  $\mathcal{V}$  contribue dans le terme en  $\varepsilon^{1/2}$ , on a donc  $\beta = 1$ . On retrouve donc la taille de la couche limite non caractéristique (voir par exemple [Guès, 1995]).

Avec un tel choix de paramètres, l'ajout de  $\mathcal{V}$  dans (IV.28) introduit les termes suivants dans la cascade d'équation interne :

- Le terme  $\mathbf{A}(\nabla_{t,x} \varphi) \partial_Y \mathcal{V} - \mathbf{E} \partial_Y^2 \mathcal{V}$  en  $\varepsilon^{1/2}$ , où  $Y$  désigne la seconde variable rapide liée aux couches limites non caractéristiques définie par  $Y := \frac{\varphi}{\varepsilon}$ .
- Puis des termes d'ordre supérieur en  $\varepsilon$ . Ces derniers ne nous importent pas, car l'on rappelle que les correcteurs  $u^2$  et  $\mathcal{U}^2$  n'annulent, de toutes façons, que la partie polarisée du terme en  $\varepsilon$ .

Pour simplifier<sup>9</sup> ce qui suit on suppose ici que  $\mathbf{E} = I$ . On pose alors  $\mathcal{W} = \partial_Y \mathcal{V}$  on résout alors l'équation d'ordre un en  $Y$  vérifiée par  $\mathcal{W}$  et on obtient  $\mathcal{V}$  par intégration. Pour déterminer la condition

<sup>9</sup>Ceci n'est pas du tout nécessaire, mais cette hypothèse aura l'avantage, dans ce qui suit, de continuer à travailler avec le projecteurs  $\Pi$  et  $I - \Pi$ . Sans cette hypothèse, on introduit juste un nouveau jeu de projecteurs adapté à  $\mathbf{E}$ .

de bord de cette équation, on étudie la condition d'annulation du saut de la dérivée normale. Elle s'écrit maintenant:

$$\begin{aligned} [\partial_n u^\varepsilon] &= \varepsilon^{1/2} ([\partial_n \mathcal{U}] + |(\nabla_{t,x} \varphi)|_\Gamma|^2 [\partial_X \mathcal{U}^2] + |(\nabla_{t,x} \varphi)|_\Gamma|^2 [\partial_Y \mathcal{V}]) \\ &\quad + \varepsilon ([\partial_n u^2] + [\partial_n \mathcal{U}^2]) + \varepsilon^{3/2} [\partial_n \mathcal{V}], \end{aligned}$$

on impose donc la condition de bord sur le saut

$$[\mathcal{W}] = [\partial_Y \mathcal{V}] = -\frac{1}{|(\nabla_{t,x} \varphi)|_\Gamma|^2} [(I - \Pi) \partial_n \mathcal{U}^2].$$

On rappelle que, puisque la partie non polarisée de  $\mathcal{U}^2$  ne dépend que des termes précédents, le membre de droite de l'équation précédente ne dépend pas de  $\Pi \mathcal{U}^2$ . Par conséquent, la détermination de  $\mathcal{W}$  peut se faire parallèlement à celles décrites dans le paragraphe précédent.

On construira  $\mathcal{W}$  comme la juxtaposition de deux états  $\mathcal{W}_\pm$  définis dans  $\{\pm Y > 0\}$  et chacun vérifiant

$$\begin{cases} \mathbf{A}(\nabla_{t,x} \varphi) \mathcal{W}_\pm - \partial_Y \mathcal{W}_\pm = 0, & \text{dans } \Omega_T \times \{\pm Y > 0\}, \\ \mathcal{W}_\pm|_{Y=0} = \pm \Pi_\pm \partial_X \mathcal{U}|_{X=0}^2. \end{cases} \quad (\text{IV.33})$$

Equation différentielle ordinaire, que l'on peut aisément résoudre *via* la formule de Duhamel. On renvoie typiquement à [Guès, 1995].

Ainsi avec ce terme correctif la solution approchée (IV.32) vérifie la condition de continuité en  $O(\varepsilon^{3/2})$  et la condition de continuité de la dérivée normale en  $O(\varepsilon)$ . On a alors bien la régularisation espérée du saut normal.

On a donc, grâce à l'analyse formelle précédente, identifié toutes les amplitudes que l'on s'attend à voir apparaître dans le développement d'une solution approchée de (IV.23). A l'heure actuelle, nous avons démontré que les différents problèmes qui apparaissent dans la cascade d'équations étaient bien posés. Pour aboutir à un théorème propre et définitif, il nous reste à vérifier que notre solution approchée construite s'avère effectivement être une solution approchée. Ceci se fera par une analyse d'erreur adaptée et reprendra les mêmes idées que celles développées dans le cas semi-linéaire dans [Sueur, 2006] devrait s'appliquer. On pense alors arriver sous peu à un théorème de la forme :

**Théorème IV.3.2** *Sous les hypothèses de symétrie des coefficients et d'hyperbolicité à multiplicité constante. On suppose donnée un temps  $T^0 > 0$ , une hypersurface  $\Gamma \subset [0, T] \times \mathbb{R}^d$  et  $u^0$  une solution faible de l'équation hyperbolique (IV.19) donnée par le Théorème IV.3.1. Alors il existe  $T^* \in ]0, T^0[$ , des profils hyperboliques  $(u^n)_{n \geq 0}$  ; des profils de transition caractéristique  $(\mathcal{U}^n)_{n \geq 1}$  ; des profils de transition non caractéristique  $(\mathcal{V}^n)_{n \geq 3}$  et enfin un reste<sup>10</sup>  $(u_r^\varepsilon)_{\varepsilon \in (0,1)}$  tels que la famille  $(u^\varepsilon)_{\varepsilon \in (0,1)}$  définie, pour tout  $\varepsilon \in ]0, 1[$ , par*

$$\begin{aligned} u^\varepsilon(t, x) &= u^0(t, x) + \sqrt{\varepsilon} \mathcal{U}^1(t, x; \frac{\varphi(t, x)}{\sqrt{\varepsilon}}) + \varepsilon \left( u^2(t, x) + \mathcal{U}^2(t, x; \frac{\varphi(t, x)}{\sqrt{\varepsilon}}) \right) \\ &\quad + \sum_{n=3}^N \sqrt{\varepsilon}^n \left( u^n(t, x) + \mathcal{U}^n(t, x; \frac{\varphi(t, x)}{\sqrt{\varepsilon}}) + \mathcal{V}^n(t, x, \frac{\varphi(t, x)}{\varepsilon}) \right) \\ &\quad + \sqrt{\varepsilon}^{N+1} u_r^\varepsilon(t, x), \end{aligned} \quad (\text{IV.34})$$

soit solution de (IV.23) sur l'intervalle de temps  $[0, T^*]$ .

<sup>10</sup>Le sens du terme reste sera à être précisé. On pense obtenir un reste analogue aux termes de reste considérés dans [Guès, 1995] ou [Sueur, 2006].



# Chapter V

## Conclusion

Tâchons de dresser, à titre de conclusion provisoire, un résumé de nos principales contributions sur les problèmes aux limites et leurs approximations numériques. On reprendra aussi ici les perspectives de recherche qui nous semble les plus prometteuses parmi celles que l'on a déjà mentionnées.

Nous pensons que l'une de nos contributions sur les problèmes aux limites, que ces derniers soient posés dans le demi-espace, la bande, ou encore le quart d'espace, est d'avoir précisé les résultats traitant des problèmes faiblement bien posés. Nous avons en particulier, (presque) terminé la classification dans la géométrie canonique du demi-espace. Mais, nous avons aussi (et surtout) montré que lorsque la frontière de l'espace de base admettait plusieurs composantes, la classification des problèmes faiblement bien posés pouvait être beaucoup plus riche.

Bien sûr, des cas de dégénérescence restent à étudier, notamment les dégénérescences en zone de *glancing*, par exemple, dans une bande ou un quart d'espace, mais aussi pour un mode de *glancing* d'ordre supérieur à deux. Toutefois, il s'agit ici de conclure une théorie bien comprise et non pas de développer des nouvelles techniques. On pense donc que ces sujets pourraient être pertinents pour un début de travail doctoral.

Dans la géométrie du demi-espace nous avons proposé une nouvelle méthode d'obtention de solutions régulières. Cette méthode est directe et elle n'utilise pas de sous-étape de régularisation. De plus, elle a l'avantage certain de passer, sans modifications notables, à la géométrie du quart d'espace.

Dans le futur, nous poursuivrons l'étude de cette méthode. Le but n'est pas ici de donner de nouveaux résultats, mais plutôt de retrouver ceux de la littérature en permettant, par exemple, aux coefficients de l'opérateur de varier avec le temps. Puisque cette méthode de non régularisation est très peu sujette à la géométrie de l'espace, on espère pouvoir ensuite transposer ces nouveaux résultats dans la géométrie du quart d'espace.

Dans la géométrie de la bande nos contributions ont été d'avoir identifié une condition "nécessaire et suffisante" pour assurer le caractère fortement bien posé *optimal* (c'est-à-dire celui excluant la croissance exponentielle en temps de la solution) et d'avoir construit les développements d'optique géométrique associés.

Comme on l'a déjà mentionné, il est honnête de garder à l'esprit qu'une telle théorie est optimale mais "coûteuse". Un simple argument de localisation donnerait une théorie tout à fait satisfaisante bien qu'elle ne soit pas optimale.

On pense toutefois que l'un des intérêts du développement de cette méthode est d'avoir démontré une bonne fois pour toutes que les techniques de construction du *symétriseur* d'Osher [Osher, 1973] pouvaient être utilisées de façon concluante pour établir des estimations d'énergie *a priori*. Ceci est de bon augure pour établir un jour une caractérisation des problèmes fortement bien posés dans une géométrie à coin.

Disposant maintenant d'une théorie optimale, nous aimerions l'utiliser pour voir si cette dernière ne permettrait pas d'obtenir des résultats plus fins que ceux basés sur un simple argument de localisation. On pense par exemple à la question de l'amélioration du temps d'existence pour les problèmes non linéaires. C'est une question qui nous attire et que nous étudierons. Toutefois cette perspective peut (peut-être) être plus considérée comme des "mathématiques pour le plaisir". Nous ne pensons pas que les apports de connaissance qu'elle apporterait à la théorie générale soient vraiment, ici, une motivation importante.

Nous pensons aussi qu'obtenir les développements d'optique géométrique pour des problèmes non linéaires dans la bande est une perspective intéressante. Cette étude nous amènera à comprendre précisément comment l'auto-interaction de phases induite par la géométrie vient communiquer avec celle induite par la non-linéarité, renforçant ainsi nos connaissances sur l'auto-interaction de la géométrie.

Si l'on se place maintenant dans le cas du quart d'espace, alors les perspectives sont (presque) toujours aussi nombreuses que lorsque nous avons débuté notre thèse de doctorat. Ce n'est pas que nous pensions que les nouveaux résultats décrits dans le Paragraphe II.4.1, ou bien nos contributions soient à négliger, mais, il reste encore beaucoup de choses intéressantes, belles et motivantes à comprendre.

Nous pensons que notre résultat de persistance de la régularité est un théorème intéressant, rassurant et qu'il à l'avantage d'initier une discussion. Toutefois il faudra, à l'avenir, reprendre ce résultat, l'affiner et l'améliorer afin de disposer d'une théorie de persistance de la régularité à même de traiter des problèmes non linéaires. En effet, nous avons effectué le premier pas, mais de nombreux suivront.

Obtenir une caractérisation complète des conditions de bord donnant lieu à un problème dans le quart d'espace fortement bien posé est un objectif ambitieux, mais capital, pour une meilleure compréhension de la théorie générale. Fort de notre étude dans la bande, nous pensons que l'analyse d'Osher pourrait permettre de conclure et en se plaçant sous le prisme de l'analyse microlocale. Des arguments doivent être raffinés pour les rendre vraiment concluants. Cette perspective semble toutefois être un objectif à long terme.

En effet, pour conclure, nous pensons que nous avons besoin de comprendre précisément le rôle des opérateurs de trace d'Osher et en particulier leur nature mathématique. Ces derniers semblent rentrer dans la classe des opérateurs intégraux de Fourier à phase matricielle complexe, objets pour lesquels la théorie (des espaces fonctionnels où ces opérateurs sont bornés) n'est, du moins pas à notre connaissance, bien établie. Par conséquent, une étude abstraite approfondie dépassant le cadre de la boîte à outils semble nécessaire. Nous pensons toutefois qu'une telle étude, menée peut-être en se basant sur les techniques microlocales développées par Le Rousseau (voir [Le Rousseau, 2009]) qui visent à décomposer de tels opérateurs comme un produit d'opérateurs plus simples, devrait être très intéressante.

Enfin, concernant l'analyse d'Osher, il "faudra" aussi à l'avenir comprendre si l'analyse donnée dans [Osher, 1973] est optimale. En particulier, nous pensons qu'elle gagnerait à être affinée pour améliorer le *symétriseur* obtenu, pour la rendre à même de traiter des cas non symétriques ou encore à coefficients variables.

Nous pensons aussi que la question de l'extension de la condition d'Osher au bord de l'espace des fréquences, pour en obtenir "gratuitement" une version uniforme, devra être posée.

Une autre de nos contributions concernant les problèmes dans le quart d'espace est d'avoir construit des développements d'optique géométrique rigoureux et ceci même lorsque les divers phénomènes d'auto-interaction en présence sont élaborés. Nous aurons toutefois besoin de pousser l'analyse un peu plus loin pour pouvoir traiter des auto-interactions composites.

Ces développements ont deux avantages certains. Ils permettent d'illustrer, au niveau microlocal, les conditions que l'on voit apparaître lors de l'étude du caractère fortement bien posé du problème. Ceci justifie, en un certain sens, l'apparition de telles conditions. Et d'autre part, ils fournissent des solutions

approchées explicitement calculables et donc implémentables.

Motivé par le fait que cette construction donne une solution approchée clé en main, nous nous proposons, à l'avenir, de venir étendre cette analyse dans des géométries bornées. Dans de telles géométries l'auto-interaction devient alors générique, du moins, elle a plus de chances de l'être. Le procédé de génération des phases devient lui aussi extrêmement complexe. Cependant, nous pensons que l'obtention d'une solution approchée calculable mérite bien de devoir faire face à un peu de technique.

Là encore, il s'agit d'une perspective à long terme qui doit être menée graduellement. On pourrait, par exemple, finir de considérer l'équation des ondes dans des géométries bien particulières (rectangle ou triangle), puis passer aux systèmes à deux équations, avant d'essayer d'attaquer frontalement les systèmes de taille  $N$ . Une telle approche graduelle serait, nous le pensons, adaptée à des travaux doctoraux.

Poursuivons maintenant du côté discret des problèmes aux limites. Cet aspect des choses étant l'un des derniers apparus dans notre recherche, il est normal que nos contributions soient moins nombreuses. Nous continuerons, néanmoins, à l'avenir de contribuer dans ce domaine.

Si nous ne devons garder qu'une seule contribution, nous aimerions mettre en lumière notre résultat caractérisant les conditions de bord conduisant à la stabilité pour des problèmes dans l'intervalle. Ce travail, qui doit être vu comme un premier pas dans l'obtention de résultats dans de "vraies" géométries bornées, est intéressant car il démontre que l'analyse d'Osher admet une version discrète exactement comme cela est le cas de la condition de Kreiss-Lopatinskii uniforme. Il permet, de plus, de mieux comprendre les relations que plusieurs conditions de bord numériques peuvent entretenir les unes avec les autres.

A l'avenir, nous pensons qu'il serait très intéressant pour la théorie générale de voir si le schéma de démonstration utilisé pour étudier le cas de l'intervalle ne serait pas assez robuste pour traiter le cas d'un schéma dans le quart d'espace. Ceci permettrait alors de faire un grand pas vers des géométries plus proches de celles qui sont réellement implémentées, notamment en levant le voile sur le rôle et l'influence des conditions fantômes au coin.

Cette étude pourrait se faire de façon graduelle, d'abord en considérant des schémas bien particuliers (on pense ici aux schémas de Lax-Friedrichs et de Lax-Wendroff). En effet, sur un schéma particulier, on peut espérer pouvoir étudier avec assez de précision les opérateurs de traces et en déduire les conditions idoines sur les mailles fantômes. Puis, une fois cette analyse d'exemples bien rodée, s'attaquer à un cadre plus général.

A propos de nos contributions dans le cas discret, nous aimerions aussi revenir sur l'obtention de la stabilité pour le schéma de Lax-Wendroff par la directe méthode d'énergie. Nous trouvons cette méthode particulièrement satisfaisante. En effet, elle évite totalement la technique induite par l'analyse de Fourier-Laplace<sup>1</sup> et permet d'obtenir la stabilité du schéma de façon directe. Toutefois, cette méthode a un prix : elle utilise fortement la structure et l'expression explicite du schéma. Ainsi, elle semble difficile à transposer dans un cadre général et bien que nous pensons que cette méthode soit transposable à d'autres schémas, son application se fera alors, au cas par cas. Nous n'avons pas pour but de nourrir un bestiaire des schémas pour lesquels la méthode d'énergie s'applique, cependant il serait peut-être bon, à l'avenir, de donner quelques autres exemples de schémas pour lesquels cette méthode d'énergie s'applique, afin de se convaincre que le domaine d'applicabilité de l'étude n'est pas si restreint.

Pour conclure sur les perspectives de recherche concernant les problèmes discrets nous aimerions revenir sur l'obtention d'une caractérisation pour des schémas stables posés dans des géométries multidimensionnelles (avec toutefois en première intention le cas du demi-espace): ce sont clairement des géométries qui nous intéressent d'un point de vue implémentation. A notre avis, le fait que la théorie de

---

<sup>1</sup>Analyse qui, on le pense, induit un frein dans la popularisation de la théorie.

stabilité générale soit essentiellement cantonnée au cas unidimensionnel est l'un des grands points noirs de cette théorie. Les travaux de [Michelson, 1983] semblent pourtant indiquer que la condition de stabilité de [B. Gustafsson and Sundstrom., 1972] s'étend (du moins dans le cas particulier des opérateurs dissipatifs) en plus grande dimension. Cette extension, on ne peut plus naturelle, gagnerait à être rendue plus systématique et c'est un point sur lequel nous aimerions apporter des précisions.

Enfin pour conclure, nous nous permettons de revenir quelque peu sur les autres résultats, que nous avons obtenus ces dernières années, hors de la théorie pure et dure des systèmes hyperboliques. Ces résultats traitent de "sous-théories" des équations aux dérivées partielles variées, qu'il s'agisse de la théorie de l'homogénéisation ou bien de la théorie du contrôle. Nous pensons continuer de travailler sur ces questions connexes qui n'en sont pas moins intéressantes et ont l'avantage certain de varier les plaisirs.

Le dernier travail de recherche actuellement en cours que nous avons mentionné dans ce texte, à savoir l'approximation régulière des ondes soniques de Métivier [Métivier, 1991] en est un bon exemple. En effet, nous sommes enchanté de venir appliquer ici des techniques proches de celles rencontrées au Chapitre II, mais avec, cette fois, comme objectif final la meilleure compréhension d'un phénomène physique existant.

Il ne nous reste plus qu'à remercier le lecteur de ce manuscrit. Nous espérons que sa lecture n'en a pas été (trop) pénible et qu'elle a été de quelques enrichissements mathématiques.

# Bibliography

- [Allaire, 1992] Allaire, G. (1992). Homogenization and two-scale convergence. *SIAM J. Math. Anal.*, 23(6):1482–1518.
- [Allaire et al., 2022] Allaire, G., Lamacz-Keymliing, A., and Rauch, J. (2022). Crime pays; homogenized wave equations for long times. *Asymptotic Anal.*, 128(3):295–336.
- [Armstrong et al., 2016] Armstrong, S., Gloria, A., and Kuusi, T. (2016). Bounded correctors in almost periodic homogenization. *Arch. Ration. Mech. Anal.*, 222(1):393–426.
- [Artola and Majda, 1987] Artola, M. and Majda, A. (1987). Nonlinear development of instabilities in supersonic vortex sheets. i. the basic kink modes. *Physica D: Nonlinear Phenomena*, 28(3):253–281.
- [B. Gustafsson and Sundstrom., 1972] B. Gustafsson, H.-O. K. and Sundstrom., A. (1972). Stability theory of difference approximations for mixed initial boundary value problems. ii. *Math. Comp.*, 26(119):649–686.
- [Bardos et al., 1992] Bardos, C., Lebeau, G., and Rauch, J. (1992). Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary. *SIAM J. Control Optim.*, 30(5):1024–1065.
- [Be et al., ] Be, oit, A., Loyer, R., and Rosier, L. Kirchhoff’s formula and numerical control of the wave equation (en cours de soumission).
- [Benoit, a] Benoit, A. Geometric optics expansions for corner problem iii: glancing modes and complex selfinteraction (soumis aux annales de la faculté de toulouse).
- [Benoit, b] Benoit, A. Geometric optics expansions for the wave equation in bounded domains with corners (preprint).
- [Benoit, c] Benoit, A. Regular solution to hyperbolic boundary value problems without regularization (*accepté dans Differential and Integral Equations*).
- [Benoit, 2014] Benoit, A. (2014). Geometric optics expansions for linear hyperbolic boundary value problems and optimality of energy estimates for surface waves. *Differential Integral Equations*, 27(5-6):531–562.
- [Benoit, 2015] Benoit, A. (2015). *Problèmes aux limites, optique géométrique et singularités*. PhD thesis, Université de Nantes. <https://hal.archives-ouvertes.fr/tel-01180449v1>.
- [Benoit, 2016] Benoit, A. (2016). Geometric optics expansions for hyperbolic corner problems, I: Self-interaction phenomenon. *Anal. PDE*, 9(6):1359–1418.
- [Benoit, 2017a] Benoit, A. (2017a). Geometric optics expansions for hyperbolic corner problems. II: From weak stability to violent instability. *SIAM J. Math. Anal.*, 49(5):3335–3395.
- [Benoit, 2017b] Benoit, A. (2017b). A necessary condition for the strong stability of finite difference scheme approximations for hyperbolic corner domains. <https://hal.science/hal-01490903> (non soumis).
- [Benoit, 2018] Benoit, A. (2018). Semi-group stability of finite difference schemes in corner domains. *Numer. Math., Theory Methods Appl.*, 11(3):618–654.
- [Benoit, 2020a] Benoit, A. (2020a). Lower exponential strong well-posedness of hyperbolic boundary value problems in a strip. *Indiana Univ. Math. J.*, 69(7):2267–2323.
- [Benoit, 2020b] Benoit, A. (2020b). Weak well-posedness of hyperbolic boundary value problems in a strip: when instabilities do not reflect the geometry. *Commun. Pure Appl. Anal.*, 19(12):5475–5486.

- [Benoit, 2020c] Benoit, A. (2020c). WKB expansions for hyperbolic boundary value problems in a strip: selfinteraction meets strong well-posedness. *J. Inst. Math. Jussieu*, 19(5):1629–1675.
- [Benoit, 2021] Benoit, A. (2021). WKB expansions for weakly well-posed hyperbolic boundary value problems in a strip: time depending loss of derivatives. *J. Hyperbolic Differ. Equ.*, 18(3):557–608.
- [Benoit, 2022] Benoit, A. (2022). Stability of finite difference schemes approximation for hyperbolic boundary value problems in an interval. *Math. Comput.*, 91(335):1171–1212.
- [Benoit, 2024] Benoit, A. (2024). Persistence of regularity of the solution to a hyperbolic boundary value problem in domain with corner, *à paraître dans Journal of Differential Equations* <https://hal.science/hal-03580984>
- [Benoit and Coulombel, a] Benoit, A. and Coulombel, J.-F. The neumann boundary condition for the two-dimensional lax-wendroff scheme ii, *en préparation*,
- [Benoit and Coulombel, b] Benoit, A. and Coulombel, J.-F. The neumann boundary condition for the two-dimensional lax-wendroff scheme (to appear in cms).
- [Benoit and Gloria, 2019] Benoit, A. and Gloria, A. (2019). Long-time homogenization and asymptotic ballistic transport of classical waves. *Ann. Sci. Éc. Norm. Supér. (4)*, 52(3):703–759.
- [Benoit and Loyer, 2023] Benoit, A. and Loyer, R. (2023). Geometric optics expansion for weakly well-posed hyperbolic boundary value problems: the glancing degeneracy. *prochainement dans Asymptotic Analysis*.
- [Benoit et al., 2023] Benoit, A., Loyer, R., and Rosier, L. (2023). Null controllability of strongly degenerate parabolic equations. *ESAIM, Control Optim. Calc. Var.*, 29:36. Id/No 48.
- [Bensoussan et al., 1978] Bensoussan, A., Lions, J.-L., and Papanicolaou, G. (1978). Asymptotic analysis for periodic structures. *Studies in Mathematics and its Applications*. Vol. 5. Amsterdam - New York - Oxford: North-Holland Publ. Company. XXIV, 700 p. \$ 49.00; Dfl. 110.00 (1978).
- [Benzoni-Gavage et al., 2012] Benzoni-Gavage, S., Chupin, L., Jamet, D., and Vovelle, J. (2012). On a phase field model for solid-liquid phase transitions. *Discrete Contin. Dyn. Syst.*, 32(6):1997–2025.
- [Benzoni-Gavage and Freistühler, 2004] Benzoni-Gavage, S. and Freistühler, H. (2004). Effects of surface tension on the stability of dynamical liquid-vapor interfaces. *Arch. Ration. Mech. Anal.*, 174(1):111–150.
- [Benzoni-Gavage et al., 2002] Benzoni-Gavage, S., Rousset, F., Serre, D., and Zumbrun, K. (2002). Generic types and transitions in hyperbolic initial-boundary-value problems. *Proc. Roy. Soc. Edinburgh Sect. A*, 132(5):1073–1104.
- [Benzoni-Gavage, 2007] Benzoni-Gavage, D. Serre, S. (2007). *Multidimensional hyperbolic partial differential equations*. Oxford Mathematical Monographs. Oxford University Press.
- [Besse et al., 2021] Besse, C., Coulombel, J.-F., and Noble, P. (2021). Discrete transparent boundary conditions for the two-dimensional leap-frog scheme: approximation and fast implementation. *ESAIM, Math. Model. Numer. Anal.*, 55:535–571.
- [Birkhoff and Rota, 1962] Birkhoff, G. and Rota, G.-C. (1962). Ordinary differential equations. *Introductions to Higher Mathematics*. Boston-New York-Chicago: Ginn and Company, VII, 318 p. (1962).
- [Blochas and Rodrigues, 2022] Blochas, P. and Rodrigues, L. M. (2022). Uniform asymptotic stability for convection-reaction-diffusion equations in the inviscid limit towards riemann shocks.
- [Blokhnin and Tkachev, 1996a] Blokhin, A. M. and Tkachev, D. L. (1996a). A mixed problem for the wave equation in coordinate domains. I: Mixed problem for the wave equation in a quadrant. *Vychisl. Tekhnol.*, 1(1):13–37.
- [Blokhnin and Tkachev, 1996b] Blokhin, A. M. and Tkachev, D. L. (1996b). A mixed problem for the wave equation in coordinate domains. II. Obtaining a priori estimates. *Vychisl. Tekhnol.*, 1(2):26–46.
- [Blokhnin and Tkachev, 1996c] Blokhin, A. M. and Tkachev, D. L. (1996c). Mixed problems for the wave equation in the domain with a corner. In *Differential equation theory*, pages 167–197. New York, NY: Nova Science Publishers, Inc.

- [Brahim-Otsmane et al., 1992] Brahim-Otsmane, S., Francfort, G. A., and Murat, F. (1992). Correctors for the homogenization of the wave and heat equations. *J. Math. Pures Appl. (9)*, 71(3):197–231.
- [Brenner, 1966] Brenner, P. (1966). The Cauchy problem for symmetric hyperbolic systems in  $L_p$ . *Math. Scand.*, 19:27–37.
- [Brenner, 1973] Brenner, P. (1973). The Cauchy problem for systems in  $L_p$  and  $l_p, \alpha$ . *Ark. Mat.*, 11:75–101.
- [Burq and Gérard, 1997] Burq, N. and Gérard, P. (1997). A necessary and sufficient condition for the exact controllability of the wave equation. *C. R. Acad. Sci., Paris, Sér. I, Math.*, 325(7):749–752.
- [Cannarsa et al., 2008] Cannarsa, P., Martinez, P., and Vancostenoble, J. (2008). Carleman estimates for a class of degenerate parabolic operators. *SIAM J. Control Optim.*, 47(1):1–19.
- [Cannarsa et al., 2017] Cannarsa, P., Martinez, P., and Vancostenoble, J. (2017). The cost of controlling weakly degenerate parabolic equations by boundary controls. *Math. Control Relat. Fields*, 7(2):171–211.
- [Cannarsa et al., 2021] Cannarsa, P., Martinez, P., and Vancostenoble, J. (2021). Sharp estimate of the cost of controllability for a degenerate parabolic equation with interior degeneracy. *Minimax Theory Appl.*, 6(2):251–280.
- [Casella et al., 2006] Casella, E., Secchi, P., and Trebeschi, P. (2006). Non-homogeneous linear symmetric hyperbolic systems with characteristic boundary. *Differ. Integral Equ.*, 19(1):51–74.
- [Chazarain and Piriou, 1981] Chazarain, J. and Piriou, A. (1981). *Introduction à la théorie des équations aux dérivées partielles linéaires*. Gauthier-Villars, Paris.
- [Colombini et al., 2013a] Colombini, F., Del Santo, D., Fanelli, F., and Métivier, G. (2013a). Time-dependent loss of derivatives for hyperbolic operators with non regular coefficients. *Commun. Partial Differ. Equations*, 38(10):1791–1817.
- [Colombini et al., 2013b] Colombini, F., Del Santo, D., Fanelli, F., and Métivier, G. (2013b). A well-posedness result for hyperbolic operators with Zygmund coefficients. *J. Math. Pures Appl. (9)*, 100(4):455–475.
- [Colombini et al., 2015] Colombini, F., Del Santo, D., Fanelli, F., and Métivier, G. (2015). The well-posedness issue in Sobolev spaces for hyperbolic systems with Zygmund-type coefficients. *Commun. Partial Differ. Equations*, 40(11):2082–2121.
- [Correa, ] Correa, S. Energy estimate for the wr class (preprint <https://arxiv.org/abs/2307.12431>).
- [Coulombel, 2002a] Coulombel, J.-F. (2002a). *Stabilité multidimensionnelle d’interfaces dynamiques. Application aux transitions de phase liquide-vapeur*. PhD thesis, ENS Lyon.
- [Coulombel, 2002b] Coulombel, J.-F. (2002b). Weak stability of nonuniformly stable multidimensional shocks. *SIAM J. Math. Anal.*, 34(1):142–172.
- [Coulombel, 2011] Coulombel, J.-F. (2011). Stability of finite difference schemes for hyperbolic initial boundary value problems II. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 10(1):37–98.
- [Coulombel, 2012] Coulombel, J.-F. (2012). Stability of finite difference schemes for hyperbolic initial boundary value problems. In *Hyperbolic problems. Theory, numerics and applications. Vol. 1. Proceedings of the 13th international conference on hyperbolic problems, HYP 2010, Beijing, China, June 15–19, 2010*, pages 184–195. Hackensack, NJ: World Scientific; Beijing: Higher Education Press.
- [Coulombel, 2014] Coulombel, J.-F. (2014). On the strong stability of finite difference schemes for hyperbolic systems in two space dimensions. *Calcolo*, 51(1):97–108.
- [Coulombel, 2015] Coulombel, J.-F. (2015). The Leray-Gårding method for finite difference schemes. *J. Éc. Polytech., Math.*, 2:297–331.
- [Coulombel, 2019] Coulombel, J.-F. (2019). Transparent numerical boundary conditions for evolution equations: derivation and stability analysis. *Ann. Fac. Sci. Toulouse, Math. (6)*, 28(2):259–327.
- [Coulombel, 2021] Coulombel, J.-F. (2021). The Leray-Gårding method for finite difference schemes. II. Smooth crossing modes. *North-West. Eur. J. Math.*, 7:151–184.

- [Coulombel and Gloria, 2011] Coulombel, J.-F. and Gloria, A. (2011). Semigroup stability of finite difference schemes for multidimensional hyperbolic initial-boundary value problems. *Math. Comput.*, 80(273):165–203.
- [Coulombel and Guès, 2010] Coulombel, J.-F. and Guès, O. (2010). Geometric optics expansions with amplification for hyperbolic boundary value problems: linear problems. *Ann. Inst. Fourier (Grenoble)*, 60(6):2183–2233.
- [Coulombel et al., 2011] Coulombel, J.-F., Gues, O., and Williams, M. (2011). Resonant leading order geometric optics expansions for quasilinear hyperbolic fixed and free boundary problems. *Commun. Partial Differ. Equations*, 36(10-12):1797–1859.
- [Coulombel et al., 2014] Coulombel, J.-F., Guès, O., and Williams, M. (2014). Semilinear geometric optics with boundary amplification. *Anal. PDE*, 7(3):551–625.
- [Coulombel and Williams, 2020] Coulombel, J.-F. and Williams, M. (2020). *Geometric optics for surface waves in nonlinear elasticity*, volume 1271 of *Mem. Am. Math. Soc.* Providence, RI: American Mathematical Society (AMS).
- [Courant et al., 1928] Courant, R., Friedrichs, K., and Lewy, H. (1928). Über die partiellen Differenzgleichungen der mathematischen Physik. *Math. Ann.*, 100:32–74.
- [Dohnal et al., 2014] Dohnal, T., Lamacz, A., and Schweizer, B. (2014). Bloch-wave homogenization on large time scales and dispersive effective wave equations. *Multiscale Model. Simul.*, 12(2):488–513.
- [Dohnal et al., 2015] Dohnal, T., Lamacz, A., and Schweizer, B. (2015). Dispersive homogenized models and coefficient formulas for waves in general periodic media. *Asymptotic Anal.*, 93(1-2):21–49.
- [Duerinckx and Gloria, 2020a] Duerinckx, M. and Gloria, A. (2020a). Multiscale functional inequalities in probability: concentration properties. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 17(1):133–157.
- [Duerinckx and Gloria, 2020b] Duerinckx, M. and Gloria, A. (2020b). Multiscale functional inequalities in probability: constructive approach. *Ann. Henri Lebesgue*, 3:825–872.
- [Duerinckx et al., ] Duerinckx, M., Gloria, A., and Ruf, M. A spectral ansatz for the long-time homogenization of the wave equation. *preprint*, <https://arxiv.org/pdf/2303.07684.pdf>.
- [Ehrhardt, 2010] Ehrhardt, M. (2010). Absorbing boundary conditions for hyperbolic systems. *Numer. Math., Theory Methods Appl.*, 3(3):295–337.
- [Eller, 2012] Eller, M. (2012). On symmetric hyperbolic boundary problems with nonhomogeneous conservative boundary conditions. *SIAM J. Math. Anal.*, 44(3):1925–1949.
- [Eller, 2018] Eller, M. (2018). Loss of derivatives for hyperbolic boundary problems with constant coefficients. *Discrete Contin. Dyn. Syst., Ser. B*, 23(3):1347–1361.
- [Engquist and Majda, 1977] Engquist, B. and Majda, A. (1977). Absorbing boundary conditions for the numerical simulation of waves. *Math. Comput.*, 31:629–651.
- [Faye and Rodrigues, 2023] Faye, G. and Rodrigues, L. M. (2023). Exponential asymptotic stability of Riemann shocks of hyperbolic systems of balance laws. *SIAM J. Math. Anal.*, 55(6):6425–6456.
- [Fagnelli and Mugnai, 2016] Fagnelli, G. and Mugnai, D. (2016). *Carleman estimates, observability inequalities and null controllability for interior degenerate non smooth parabolic equations*, volume 1146 of *Mem. Am. Math. Soc.* Providence, RI: American Mathematical Society (AMS).
- [Fagnelli and Mugnai, 2020] Fagnelli, G. and Mugnai, D. (2020). Singular parabolic equations with interior degeneracy and non smooth coefficients: the Neumann case. *Discrete Contin. Dyn. Syst., Ser. S*, 13(5):1495–1511.
- [Francfort and Murat, 1992] Francfort, G. A. and Murat, F. (1992). Oscillations and energy densities in the wave equation. *Commun. Partial Differ. Equations*, 17(11-12):1785–1865.
- [Glimm, 1965] Glimm, J. (1965). Solutions in the large for nonlinear hyperbolic systems of equations. *Commun. Pure Appl. Math.*, 18:697–715.
- [Gloria et al., 2021] Gloria, A., Neukamm, S., and Otto, F. (2021). Quantitative estimates in stochastic homogenization for correlated coefficient fields. *Anal. PDE*, 14(8):2497–2537.



- [Grisvard, 1982] Grisvard, P. (1982). Singularities des problèmes aux limites dans des polyèdres. Sémin. Goulaouic-Meyer-Schwartz 1981-1982, Équat. dériv. part., Exposé No. 8, 19 p. (1982).
- [Grisvard, 1991] Grisvard, P. (1991). Some known results that fail on non-convex domains. Nonlinear partial differential equations and their applications, Lect. Coll. de France Semin., Paris/Fr. 1987-88, Vol. X, Pitman Res. Notes Math. Ser. 220, 209-215 (1991).
- [Grisvard, 2011] Grisvard, P. (2011). *Elliptic problems in nonsmooth domains*, volume 69 of *Class. Appl. Math.* Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), reprint of the 1985 hardback ed. edition.
- [Guès, 1990] Guès, O. (1990). Problème mixte hyperbolique quasi-linéaire caractéristique. (A mixed characteristic quasilinear hyperbolic problem). *Commun. Partial Differ. Equations*, 15(5):595–645.
- [Guès, 1995] Guès, O. (1995). Viscous perturbations of hyperbolic mixed problems and boundary layers. *Ann. Inst. Fourier*, 45(4):973–1006.
- [Guès et al., 2005a] Guès, O., Métivier, G., Williams, M., and Zumbrun, K. (2005a). Existence and stability of multidimensional shock fronts in the vanishing viscosity limit. *Arch. Ration. Mech. Anal.*, 175(2):151–244.
- [Guès et al., 2005b] Guès, O., Métivier, G., Williams, M., and Zumbrun, K. (2005b). Multidimensional viscous shocks. I: Degenerate symmetrizers and long time stability. *J. Am. Math. Soc.*, 18(1):61–120.
- [Guès et al., 2010] Guès, O., Métivier, G., Williams, M., and Zumbrun, K. (2010). Existence and stability of noncharacteristic boundary layers for the compressible Navier-Stokes and viscous MHD equations. *Arch. Ration. Mech. Anal.*, 197(1):1–87.
- [Halpern and Rauch, 2016] Halpern, L. and Rauch, J. (2016). Hyperbolic boundary value problems with trihedral corners. *Discrete Contin. Dyn. Syst.*, 36(8):4403–4450.
- [Hernandez, 2015] Hernandez, M. (2015). Resonant leading term geometric optics expansions with boundary layers for quasilinear hyperbolic boundary problems. *Commun. Partial Differ. Equations*, 40(3):387–437.
- [Hersh, 1963] Hersh, R. (1963). Mixed problems in several variables. *J. Math. Mech.*, 12:317–334.
- [Huang, 2015] Huang, A. (2015). Existence of solutions for linear hyperbolic initial-boundary value problems in a rectangle. *Appl. Anal.*, 94(9):1897–1925.
- [Huang and Temam, 2014] Huang, A. and Temam, R. (2014). The linear hyperbolic initial and boundary value problems in a domain with corners. *Discrete Contin. Dyn. Syst., Ser. B*, 19(6):1627–1665.
- [Ivanovici et al., 2014] Ivanovici, O., Lebeau, G., and Planchon, F. (2014). Dispersion for the wave equation inside strictly convex domains. I: The Friedlander model case. *Ann. Math. (2)*, 180(1):323–380.
- [Ivanovici et al., 2021] Ivanovici, O., Lebeau, G., and Planchon, F. (2021). Strichartz estimates for the wave equation on a 2d model convex domain. *J. Differ. Equations*, 300:830–880.
- [Joly et al., 1993] Joly, J.-L., Métivier, G., and Rauch, J. (1993). Generic rigorous asymptotic expansions for weakly nonlinear multidimensional oscillatory waves. *Duke Math. J.*, 70(2):373–404.
- [Joly et al., 1995a] Joly, J.-L., Métivier, G., and Rauch, J. (1995a). Coherent and focusing multidimensional nonlinear geometric optics. *Ann. Sci. Éc. Norm. Supér. (4)*, 28(1):51–113.
- [Joly et al., 1995b] Joly, J. L., Métivier, G., and Rauch, J. (1995b). Trilinear compensated compactness and nonlinear geometric optics. *Ann. Math. (2)*, 142(1):121–169.
- [Joly et al., 2000] Joly, J.-L., Métivier, G., and Rauch, J. (2000). *Caustics for dissipative semilinear oscillations*, volume 685 of *Mem. Am. Math. Soc.* Providence, RI: American Mathematical Society (AMS).
- [Kato, 1956] Kato, T. (1956). On linear differential equations in Banach spaces. *Commun. Pure Appl. Math.*, 9:479–486.
- [Kilque, 2022] Kilque, C. (2022). Weakly nonlinear multiphase geometric optics for hyperbolic quasilinear boundary value problems: construction of a leading profile. *SIAM J. Math. Anal.*, 54(2):2413–2507.
- [Kojima and Taniguchi, 1980] Kojima, K. and Taniguchi, M. (1980). Mixed problem for hyperbolic equations in a domain with a corner. *Funkc. Ekvacioj, Ser. Int.*, 23:171–195.

- [Kondrat'ev, 1967] Kondrat'ev, V. A. (1967). Boundary problems for elliptic equations in domains with conical or angular points. *Trans. Mosc. Math. Soc.*, 16:227–313.
- [Kondrat'ev, 1977] Kondrat'ev, V. A. (1977). Singularities of a solution of Dirichlet's problem for a second-order elliptic equation in the neighborhood of an edge. *Differ. Equations*, 13:1411–1415.
- [Kreiss, 1959] Kreiss, H.-O. (1959). Über Matrizen, die beschränkte Halbgruppen erzeugen. *Math. Scand.*, 7:71–80.
- [Kreiss, 1962] Kreiss, H.-O. (1962). Über die Stabilitätsdefinition für Differenzgleichungen, die partielle Differentialgleichungen approximieren. *BIT, Nord. Tidskr. Inf.-behandl.*, 2:153–181.
- [Kreiss, 1968] Kreiss, H.-O. (1968). Stability theory for difference approximations of mixed initial boundary value problems. I. *Math. Comput.*, 22:703–714.
- [Kreiss, 1970] Kreiss, H.-O. (1970). Initial boundary value problems for hyperbolic systems. *Comm. Pure Appl. Math.*, 23:277–298.
- [Kruzhkov, 1969] Kruzhkov, S. N. (1969). Generalized solutions of the Cauchy problem in the large for nonlinear equations of first order. *Dokl. Akad. Nauk SSSR*, 187:29–32.
- [Kruzhkov, 1970] Kruzhkov, S. N. (1970). First order quasilinear equations with several independent variables. *Mat. Sb., Nov. Ser.*, 81:228–255.
- [Kupka and Osher, 1971] Kupka, I. A. K. and Osher, S. J. (1971). On the wave equation in a multi-dimensional corner. *Commun. Pure Appl. Math.*, 24:381–393.
- [Lamacz, 2011] Lamacz, A. (2011). Dispersive effective models for waves in heterogeneous media. *Math. Models Methods Appl. Sci.*, 21(9):1871–1899.
- [Laurent et al., 2010] Laurent, C., Rosier, L., and Zhang, B.-Y. (2010). Control and stabilization of the Korteweg-de Vries equation on a periodic domain. *Commun. Partial Differ. Equations*, 35(4):707–744.
- [Lax, 1957] Lax, P. D. (1957). Asymptotic solutions of oscillatory initial value problems. *Duke Math. J.*, 24:627–646.
- [Lax, 1973] Lax, P. D. (1973). *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*, volume 11 of *CBMS-NSF Reg. Conf. Ser. Appl. Math.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.
- [Lax and Phillips, 1960] Lax, P. D. and Phillips, R. S. (1960). Local boundary conditions for dissipative symmetric linear differential operators. *Comm. Pure Appl. Math.*, 13:427–455.
- [Lax and Richtmyer, 1956] Lax, P. D. and Richtmyer, R. D. (1956). Survey of the stability of linear finite difference equations. *Commun. Pure Appl. Math.*, 9:267–293.
- [Lax and Wendroff, 1964] Lax, P. D. and Wendroff, B. (1964). Difference schemes for hyperbolic equations with high order of accuracy. *Commun. Pure Appl. Math.*, 17:381–398.
- [Le Barbenchon et al., a] Le Barbenchon, P., Boutin, B., and Seguin, N. On the stability of totally upwind schemes for the hyperbolic initial boundary value problem, (preprint <https://hal.science/hal-03732720v1>).
- [Le Barbenchon et al., b] Le Barbenchon, P., Boutin, B., and Seguin, N. Stability of finite difference schemes for the hyperbolic initial boundary value problem by winding number computations, (preprint <https://arxiv.org/abs/2302.01410>).
- [Le Rousseau, 2006] Le Rousseau, J. (2006). Fourier-integral-operator approximation of solutions to first-order hyperbolic pseudodifferential equations I: Convergence in Sobolev spaces. *Commun. Partial Differ. Equations*, 31(4-6):867–906.
- [Le Rousseau, 2009] Le Rousseau, J. (2009). Fourier-integral-operator product representation of solutions to first-order symmetrizable hyperbolic systems. *J. Anal. Math.*, 109:81–162.
- [Le Rousseau and Hörmann, 2006] Le Rousseau, J. and Hörmann, G. (2006). Fourier-integral-operator approximation of solutions to first-order hyperbolic pseudodifferential equations. II: Microlocal analysis. *J. Math. Pures Appl. (9)*, 86(5):403–426.
- [Le Rousseau et al., 2017] Le Rousseau, J., Lebeau, G., Terpolilli, P., and Trélat, E. (2017). Geometric control condition for the wave equation with a time-dependent observation domain. *Anal. PDE*, 10(4):983–1015.

- [Lebeau, 1996] Lebeau, G. (1996). Damped wave equation. In *Algebraic and geometric methods in mathematical physics. Proceedings of the 1st Ukrainian-French-Romanian summer school, Kaciveli, Ukraine, September 1-14, 1993*, pages 73–109. Dordrecht: Kluwer Academic Publishers.
- [Lescarret, 2007] Lescarret, V. (2007). Wave transmission in dispersive media. *Math. Models Methods Appl. Sci.*, 17(4):485–535.
- [Lumer and Phillips, 1961] Lumer, G. and Phillips, R. S. (1961). Dissipative operators in a Banach space. *Pac. J. Math.*, 11:679–698.
- [Majda, 1988] Majda, M. Artola, A. (1988). Nonlinear geometric optics for hyperbolic mixed problems. In *Analyse mathématique et applications*, pages 319–356. Gauthier-Villars.
- [Majda, 1983] Majda, R. Rosales, A. (1983). A theory for spontaneous Mach stem formation in reacting shock fronts. I. The basic perturbation analysis. *SIAM J. Appl. Math.*, 43(6):1310–1334.
- [Marcou, 2010] Marcou, A. (2010). Rigorous weakly nonlinear geometric optics for surface waves. *Asymptot. Anal.*, 69(3-4):125–174.
- [Martin et al., 2019] Martin, P., Rivas, I., Rosier, L., and Rouchon, P. (2019). Exact controllability of a linear Korteweg-de Vries equation by the flatness approach. *SIAM J. Control Optim.*, 57(4):2467–2486.
- [Martin et al., 2016] Martin, P., Rosier, L., and Rouchon, P. (2016). Null controllability of one-dimensional parabolic equations by the flatness approach. *SIAM J. Control Optim.*, 54(1):198–220.
- [Martin et al., 2018] Martin, P., Rosier, L., and Rouchon, P. (2018). Controllability of the 1d Schrödinger equation using flatness. *Automatica*, 91:208–216.
- [Massey, 1972] Massey, F. J. I. (1972). Abstract evolution equations and the mixed problem for symmetric hyperbolic system. *Trans. Am. Math. Soc.*, 168:165–188.
- [Mellin, 1897] Mellin, H. (1897). Zur theorie zweier allgemeinen klassen bestimmter integrale. *Acta Societatis scientiarum Fennicae : Series A Fennica*, XXII 2:1–75.
- [Métivier, 1986] Métivier, G. (1986). Interaction de deux chocs pour un système de deux lois de conservation, en dimension deux d’espace. (Interaction of two shock waves for a system with two conservation laws in two dimensional space). *Trans. Am. Math. Soc.*, 296:431–479.
- [Métivier, 1991] Métivier, G. (1991). Ondes soniques. (Sonic waves). *J. Math. Pures Appl. (9)*, 70(2):197–268.
- [Métivier, 2000] Métivier, G. (2000). The block structure condition for symmetric hyperbolic systems. *Bull. London Math. Soc.*, 32(6):689–702.
- [Métivier, 2004] Métivier, G. (2004). *Small viscosity and boundary layer methods. Theory, stability analysis, and applications*. Model. Simul. Sci. Eng. Technol. Boston, MA: Birkhäuser.
- [Métivier and Rauch, 2017] Métivier, G. and Rauch, J. (2017). Strictly dissipative nonuniqueness with corners. In *Shocks, singularities and oscillations in nonlinear optics and fluid mechanics. Papers based on the workshop, Rome, Italy, September 2015*, pages 141–149. Cham: Springer.
- [Métivier and Zumbrun, 2004] Métivier, G. and Zumbrun, K. (2004). Symmetrizers and continuity of stable subspaces for parabolic-hyperbolic boundary value problems. *Discrete Contin. Dyn. Syst.*, 11(1):205–220.
- [Métivier and Zumbrun, 2005] Métivier, G. and Zumbrun, K. (2005). Hyperbolic boundary value problems for symmetric systems with variable multiplicities. *J. Differential Equations*, 211(1):61–134.
- [Michelson, 1983] Michelson, D. (1983). Stability theory of difference approximations for multidimensional initial-boundary value problems. *Math. Comput.*, 40:1–45.
- [Morisse, 2021] Morisse, B. (2021). On hyperbolicity and Gevrey well-posedness. III: A model of weakly hyperbolic systems. *Indiana Univ. Math. J.*, 70(2):743–780.
- [Moyano, 2016] Moyano, I. (2016). Flatness for a strongly degenerate 1-D parabolic equation. *Math. Control Signals Syst.*, 28(4):22. Id/No 28.
- [Olivier, 2017] Olivier, P. (2017). *Nappes de tourbillon-courant en magnétohydrodynamique*. PhD thesis, Université de Nantes.

- [Osher, 1973] Osher, S. (1973). Initial-boundary value problems for hyperbolic systems in regions with corners. I. *Trans. Amer. Math. Soc.*, 176:141–164.
- [Osher, 1974a] Osher, S. (1974a). An ill-posed problem for a strictly hyperbolic equation in two unknowns near a corner. *Bull. Amer. Math. Soc.*, 80:705–708.
- [Osher, 1974b] Osher, S. (1974b). Initial-boundary value problems for hyperbolic systems in regions with corners. II. *Trans. Amer. Math. Soc.*, 198:155–175.
- [Peysner, 1965] Peysner, G. (1965). On the identity of weak and strong solutions of differential equations with local boundary conditions. *Am. J. Math.*, 87:267–277.
- [Peysner, 1975] Peysner, G. (1975). Symmetric positive systems in corner domains. *J. Differ. Equations*, 18:135–157.
- [Rauch, 1972] Rauch, J. (1972).  $\mathcal{L}_2$  is a continuable initial condition for Kreiss’ mixed problems. *Commun. Pure Appl. Math.*, 25:265–285.
- [Rauch, 1985] Rauch, J. (1985). Symmetric positive systems with boundary characteristic of constant multiplicity. *Trans. Amer. Math. Soc.*, 291(1):167–187.
- [Rauch, 1999] Rauch, J. (1999). Lectures on geometric optics. volume 5 of *IAS/Park City Math. Ser.*, pages 383–466. Amer. Math. Soc.
- [Rauch, 2012] Rauch, J. (2012). *Hyperbolic partial differential equations and geometric optics*, volume 133 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS).
- [Rauch and Massey, 1974] Rauch, J. B. and Massey, F. J. I. (1974). Differentiability of solutions to hyperbolic initial-boundary value problems. *Trans. Am. Math. Soc.*, 189:303–318.
- [Rosier and Rosier, 2022] Rosier, C. and Rosier, L. (2022). Numerical control of the wave equation and Huygens’ principle. *Math. Rep. (Bucur.)*, 24(74).
- [Rousset, 2001] Rousset, F. (2001). Inviscid boundary conditions and stability of viscous boundary layers. *Asymptotic Anal.*, 26(3-4):285–306.
- [Rousset, 2003] Rousset, F. (2003). Stability of small amplitude boundary layers for mixed hyperbolic-parabolic systems. *Trans. Am. Math. Soc.*, 355(7):2991–3008.
- [Rousset, 2005] Rousset, F. (2005). Characteristic boundary layers in real vanishing viscosity limits. *J. Differ. Equations*, 210(1):25–64.
- [Sablé-Tougeron, 1988] Sablé-Tougeron, M. (1988). Existence pour un problème de l’élastodynamique Neumann non linéaire en dimension 2. *Arch. Rational Mech. Anal.*, 101(3):261–292.
- [Sakamoto, 1972] Sakamoto, R. (1972).  $L^2$ -well-posedness for hyperbolic mixed problems. *Publ. Res. Inst. Math. Sci.*, 8:265–293.
- [Santosa and Symes, 1991] Santosa, F. and Symes, W. W. (1991). A dispersive effective medium for wave propagation in periodic composites. *SIAM J. Appl. Math.*, 51(4):984–1005.
- [Sarason, 1962] Sarason, L. (1962). On weak and strong solutions of boundary value problems. *Comm. Pure Appl. Math.*, 15:237–288.
- [Sarason, 1967] Sarason, L. (1967). Elliptic regularization for symmetric positive systems. *J. Math. Mech.*, 16:807–827.
- [Sarason, 6970] Sarason, L. (1969/70). Symmetrizable systems in regions with corners and edges. *J. Math. Mech.*, 19:601–607.
- [Sarason, 1977] Sarason, L. (1977). Hyperbolic and other symmetrizable systems in regions with corners and edges. *Indiana Univ. Math. J.*, 26(1):1–39.
- [Sarason and Smoller, 7475] Sarason, L. and Smoller, J. A. (1974/75). Geometrical optics and the corner problem. *Arch. Rational Mech. Anal.*, 56:34–69.
- [Secchi, 1995] Secchi, P. (1995). Linear symmetric hyperbolic systems with characteristic boundary. *Math. Methods Appl. Sci.*, 18(11):855–870.

- [Secchi, 1996] Secchi, P. (1996). Well-posedness of characteristic symmetric hyperbolic systems. *Arch. Ration. Mech. Anal.*, 134(2):155–197.
- [Secchi, 1997] Secchi, P. (1997). Characteristic symmetric hyperbolic systems with dissipation: Global existence and asymptotics. *Math. Methods Appl. Sci.*, 20(7):583–597.
- [Serre, 1996a] Serre, D. (1996a). *Systèmes de lois de conservation. I: Hyperbolicité, entropies, ondes de choc*. Paris: Diderot Editeur.
- [Serre, 1996b] Serre, D. (1996b). *Systèmes de lois de conservation. II: Structures géométriques, oscillations et problèmes mixtes*. Paris: Diderot Editeur.
- [Strang, 1969a] Strang, G. (1969a). Hyperbolic initial-boundary value problems in two unknowns. *J. Differ. Equations*, 6:161–171.
- [Strang, 1969b] Strang, G. (1969b). Hyperbolic initial-boundary value problems in two unknowns. *J. Differential Equations*, 6:161–171.
- [Sueur, 2006] Sueur, F. (2006). Viscous approach of discontinuous solutions of semilinear symmetric hyperbolic systems. *Ann. Inst. Fourier*, 56(1):183–245.
- [Taniguchi, 1978] Taniguchi, M. (1978). Mixed problem for wave equation in the domain with a corner. *Funkc. Ekvacioj, Ser. Int.*, 21:249–259.
- [Tartar, 2009] Tartar, L. (2009). *The general theory of homogenization. A personalized introduction*, volume 7 of *Lect. Notes Unione Mat. Ital.* Berlin: Springer.
- [Trudinger, 1973] Trudinger, N. S. (1973). Linear elliptic operators with measurable coefficients. *Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat., III. Ser.*, 27:265–308.
- [Williams, 1996] Williams, M. (1996). Nonlinear geometric optics for hyperbolic boundary problems. *Comm. Partial Differential Equations*, 21(11-12):1829–1895.
- [Williams, 2000] Williams, M. (2000). Boundary layers and glancing blow-up in nonlinear geometric optics. *Ann. Sci. École Norm. Sup. (4)*, 33(3):383–432.
- [Wu, 1995] Wu, L. (1995). The semigroup stability of the difference approximations for initial- boundary value problems. *Math. Comput.*, 64(209):71–88.
- [Zuazua, 2006] Zuazua, E. (2006). Control and numerical approximation of the wave and heat equations. In *Proceedings of the international congress of mathematicians (ICM), Madrid, Spain, August 22–30, 2006. Volume III: Invited lectures*, pages 1389–1417. Zürich: European Mathematical Society (EMS).