



HAL
open science

Méthodes itératives de décomposition de domaine sans recouvrement avec convergence géométrique pour l'équation de Helmholtz

Matthieu Lecouvez

► **To cite this version:**

Matthieu Lecouvez. Méthodes itératives de décomposition de domaine sans recouvrement avec convergence géométrique pour l'équation de Helmholtz. Equations aux dérivées partielles [math.AP]. Ecole Polytechnique, 2015. Français. NNT : 2015EPXX0020 . tel-01444540

HAL Id: tel-01444540

<https://hal.science/tel-01444540>

Submitted on 24 Jan 2017

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.



Méthodes itératives de décomposition de domaine sans recouvrement avec convergence géométrique pour l'équation de Helmholtz

Matthieu Lecouvez

► **To cite this version:**

Matthieu Lecouvez. Méthodes itératives de décomposition de domaine sans recouvrement avec convergence géométrique pour l'équation de Helmholtz. Analyse numérique [math.NA]. Ecole Polytechnique, 2015. Français. <tel-01229546>

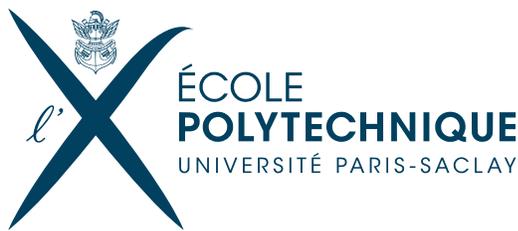
HAL Id: tel-01229546

<https://pastel.archives-ouvertes.fr/tel-01229546>

Submitted on 16 Nov 2015

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.



THÈSE

PRÉSENTÉE À

L'UNIVERSITÉ DE PARIS-SACLAY
ÉCOLE DOCTORALE DE L'ÉCOLE POLYTECHNIQUE

Par **Matthieu Lecouvez**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : mathématiques appliquées

Méthodes itératives de décomposition de domaine sans recouvrement avec convergence géométrique pour l'équation de Helmholtz

Présentée et soutenue publiquement le **02 Juillet 2015** devant le jury composé de

X. Antoine	Professeur	Institut Elie Cartan de Lorraine	Rapporteur
X. Claeys	Maître de conférence	Laboratoire Jacques-Louis Lions	Examineur
F. Collino	Ingénieur de Recherche		Examineur
B. Després	Professeur	Laboratoire Jacques-Louis Lions	Examineur
V. Dolean	Assistant professeur	Laboratoire J.A. Dieudonné	Rapporteur
P. Joly	Directeur de Recherche	INRIA - Equipe POEMS	Directeur
B. Stupfel	Ingénieur de Recherche	CEA/Cesta	Examineur



Remerciements

Une thèse de doctorat est l'aboutissement d'(au moins) trois années d'un travail que l'on pourrait penser solitaire. Bien au contraire ! Cette thèse ne serait pas sans l'apport et le soutien significatifs de nombreuses personnes. Il est grand temps de les en remercier dans ces quelques lignes...

Mes premiers remerciements vont à mon directeur de thèse Patrick Joly. Tous ses thésards peuvent en témoigner, moi le premier, Patrick est un excellent directeur de thèse, chaleureux, attentif et surtout pédagogue. Tes nombreux conseils et idées ont été très stimulants intellectuellement. Je souhaite également remercier particulièrement Bruno Stupfel, qui a été à l'initiative de cette thèse. Les mathématiciens et physiciens ont chacun leur propre langage, et même si ça n'a pas toujours été évident, nous avons su trouver, je pense, un langage commun. Et pour quelqu'un qui, je cite, *n'y connaît rien*, tes remarques, corrections et idées sont bien plus pertinentes que tu ne veux l'admettre !

Bien évidemment, je tiens ensuite à remercier l'ensemble du jury pour l'intérêt et l'attention qu'il a portés à mes travaux. Un grand merci à Bruno Després de m'avoir fait l'honneur d'être président du jury. Je remercie également Xavier Antoine et Victorita Dolean d'avoir rapporter mon manuscrit de thèse et Xavier Claeys d'avoir participer à ce jury. Les remarques et questions que chacun d'entre vous a formulées montrent l'intérêt que vous avez porté à la lecture de ma thèse.

Je me dois de remercier ensuite très chaleureusement Francis Collino. Francis est un grand habitué des encadrements officieux de thésards du CEA, et ma thèse ne déroge pas à la règle. Ton apport substantiel a été plus qu'apprécié et travailler avec toi a été très enrichissant. J'espère sincèrement avoir de nouveau cette opportunité. Dans la même optique, je remercie également Marc Duruflé. Tes compétences impressionnantes en programmation m'ont été d'un immense secours et j'espère que nous aurons l'occasion de travailler de nouveau ensemble.

Un thésard serait bien peu de chose sans tout un laboratoire comme soutien logis-

tique et moral, je n'aurais pu rêver mieux que mon laboratoire au CEA. Pour ça, j'exprime toute ma gratitude à mes chefs de service et laboratoire, et leur adjoint : Xavier, Nicolas, Michel, Pierre-Henri et David, pour leur chaleureux accueil. Un merci particulier pour David et tout ce qu'il a pu faire pour moi, avant, pendant et après ma thèse... Je ne remercierai probablement jamais assez Anne-Pascale, notre secrétaire ultra sympathique et performante.

Toujours au CEA, je tiens à remercier l'ensemble de mes collègues, que j'ai eu plaisir à côtoyer durant ces trois années. Dans le désordre et de manière succincte, au risque d'oublier quelqu'un : un grand merci à Jean-Marc de m'avoir accueilli dans son bureau (si tu veux ta revanche aux Sudoku, c'est quand tu veux!), Jean pour l'odeur de tes soupes à 10h, Pierre pour te souvenir de mon prénom, Olivier pour ta recette d'endive, Agnès pour m'avoir coaché à l'escalade (mes pauvres pseudo-muscles s'en souviennent encore), David H. pour m'avoir remplacé en tant que coaché à l'escalade, Isabelle pour ta gentillesse et ta bonne humeur (pose ce téléphone!), Julien, Céline, Murielle et tous ceux que j'ai oubliés, pour m'avoir supporté et soutenu tout ce temps. Merci!

Dans la catégorie thésards, je remercie très chaleureusement Cyril pour son humour et sa répartie sans égaux, Pascal, grand partenaire de pauses café et soutien sans faille pendant les derniers longs mois, Sébastien et Romain pour tous nos souvenirs de voyage, que je ne suis pas prêt d'oublier. Pour le paraphraser, une *mention spéciale* à Gaby. Ma thèse n'aurait pas eu la même saveur sans toi pour râler à mes côtés. Je m'ennuie un peu depuis que je ne peux plus te battre au squash, mais nos chemins se recroiseront, après tout, nous avons une agence de voyage à faire tourner!

Mes remerciements ne seraient pas complets si j'oubliais l'équipe de l'INRIA, qui m'a recueilli pre- et post-CEA. Pour cela, tout ma gratitude va à Jean Roman, Luc Giraud et Hélène Barucq. Je remercie également Pierre Ramet pour son accueil, Juliette Chabassier pour m'avoir prêté son bureau et tous les autres (Abdou, Matthieu, Manu...) pour leur amitié.

Pour finir, je remercie ma famille et mes amis de longue date, qui ont cru en moi, m'ont soutenu et encouragé toutes ces années. Qu'ils aient pu faire le déplacement pour ma soutenance à Bordeaux ou non, je les savais avec moi durant cette étape importante. Un grand merci à tous.

Résumé

Cette thèse s'intéresse aux aspects mathématiques des méthodes itératives de résolution basées sur la décomposition de domaine et appliquées à la simulation numérique de propagation d'ondes harmoniques. Plus précisément, nous nous sommes intéressés à l'élaboration de conditions de transmission optimisées garantissant la convergence exponentielle de ce type de méthodes. Une telle convergence requiert l'utilisation d'opérateurs de transmission non locaux puisqu'ils doivent correspondre formellement à un opérateur pseudo-différentiel d'ordre 1. Une méthode de localisation des opérateurs est proposée pour réduire le coût engendré par ces opérateurs tout en conservant leurs propriétés et donc la convergence exponentielle de ces méthodes itératives. Dans un cadre général, la convergence des méthodes de décomposition de domaine est établie pour toute une classe d'opérateurs vérifiant certaines conditions de positivité et d'isomorphisme entre espaces de Sobolev. Nous proposons ensuite plusieurs opérateurs différents, dépendants de paramètres, qui vérifient les conditions nécessaires à la convergence exponentielle de la méthode. Un premier type d'opérateur se base sur les normes des espaces de Sobolev d'ordre fractionnaire, tandis qu'un second type d'opérateur découle des potentiels de Riesz (puissance fractionnaire de l'opérateur de Laplace-Beltrami). Nous proposons ensuite un schéma numérique permettant d'appliquer la théorie développée à une méthode d'éléments finis. Une analyse modale dans le cas de géométries simples vient tout d'abord valider les conclusions théoriques de convergence exponentielle, puis plusieurs expériences numériques mettent en évidence les avantages des conditions de transmission proposées, et particulièrement dans le cas où une précision très fine sur la solution est demandée.



Abstract

In this thesis, we are concerned by the mathematical aspects of iterative methods based on domain decomposition and applied to the numerical simulation of wave propagation in frequency domain. More specifically, we are interested in developing optimized transmission conditions that guarantee the exponential convergence of the iterative process. Such a convergence requires non local transmission operators since they should correspond, at least formally, to pseudo differential operators of order 1. A localization method is proposed to reduce the cost caused by these operators, while keeping their properties and thus the exponential convergence of the iterative method. In a general framework, the convergence of the domain decomposition methods is established for a class of operators verifying some properties such as positiveness and isomorphism between Sobolev spaces. Then, we propose several operators, which depend on parameters, that verify the required properties to achieve exponential convergence. A first kind of operator is based on norms of Sobolev spaces of fractional orders, while a second kind of operator is derived from Riesz potential (fractional powers of Laplace-Beltrami operator). Finally, we propose a numerical scheme that allows us to apply the developed theory on a finite elements method. A modal analysis of simple geometries is used to validate the theoretical conclusions of exponential convergence, and then several numerical experiments highlight the advantages of the proposed transmission conditions, especially when high precision is needed.



Table des matières

I	Aspects théoriques	1
1	ÉTAT DE L'ART	3
1.1	Algorithmes avec recouvrement	5
1.2	Algorithmes sans recouvrement	7
1.2.1	Méthodes algébriques	7
1.2.1.1	Méthode FETI	8
1.2.1.2	Méthode FETI-DP	8
1.2.1.3	Autres méthodes algébriques	9
1.2.2	Méthodes variationnelles	10
1.2.2.1	Conditions de transmission de Després	10
1.2.2.2	Conditions d'ordre supérieur	12
1.2.2.3	Approximation de Padé	13
1.3	Conditions de transmission idéales	14
2	THÉORIE GÉNÉRALE	19
2.1	Problème modèle et notations	20
2.1.1	Problème modèle	21
2.1.2	Problème décomposé	21
2.1.3	Processus itératif	23
2.1.4	Espaces fonctionnels	26
2.2	Classe d'opérateurs de transmission robustes	27
2.3	Étude théorique	30
2.3.1	Caractère bien posé des problèmes	30
2.3.2	Reformulation à l'interface pour l'erreur	34
2.3.3	Reformulation à l'interface pour la solution	36
2.4	Résultats de convergence	37
2.5	Extension à d'autres méthodes itératives	42

3	ANALYSES MODALES	45
3.1	Problème circulaire	46
3.1.1	Solutions du problème de diffraction	46
3.1.2	Propriétés des coefficients de diffraction	49
3.1.3	Étude du taux de convergence	52
3.2	Cas plus complexes, optimisation	61
3.2.1	Géométrie plane	62
3.2.1.1	Condition de transmission dans le matériau	64
3.2.1.2	Condition de transmission à l'interface du matériau	67
3.2.1.3	Condition de transmission dans le vide	70
3.2.2	Géométrie circulaire	71
3.2.3	Optimisation	75
3.2.3.1	Remarques générales	75
3.2.3.2	Fonctions objectives et résultats	76
3.2.3.3	Comparaison des géométries plane et circulaire	81
3.2.3.4	Montée en fréquence	85
II	Réalisation des opérateurs	89
4	POTENTIELS DE RIESZ	91
4.1	Définition de l'opérateur Λ et notations	92
4.2	Opérateur de référence Λ_{ref}	93
4.3	Étude de l'opérateur Λ	94
4.3.1	Continuité de $\Lambda - \Lambda_{ref}$ dans $L^2(\Sigma)$	94
4.3.2	Propriété d'isomorphisme	98
4.4	Opérateur tronqué	99
4.5	Extension 3D	100
5	SEMI-NORMES DE GAGLIARDO	101
5.1	Courbes et espaces de fonctions	102
5.1.1	Courbes höldériennes	102
5.1.2	Espaces de fonctions sur une courbe	105
5.1.2.1	Espaces $L^2(\Sigma)$ et $H^1(\Sigma)$	105
5.1.2.2	Les espaces $H^\alpha(\Sigma)$ et leur différentes définitions	106
5.1.3	Comparaison des normes sur le cercle	108
5.2	Opérateur T	111
5.3	Opérateur Λ	111
5.3.1	Cas du cercle	112

5.3.2	Cas d'une courbe $C^{1,r}$	115
5.3.2.1	Transposition du problème sur le cercle	116
5.3.2.2	Isomorphisme pour une courbe höldériennes	118
5.4	Troncature	122
5.5	Extension 3D	124
5.6	Conclusions	125

III Aspects numériques 127

6	DISCRÉTISATION	129
6.1	Espaces des solutions	130
6.2	Discrétisation du problème global	131
6.3	Problèmes locaux, Construction par T	132
6.3.1	Formulation faible et unicité de la solution	132
6.3.2	Opérateurs discrets de <i>scattering</i>	135
6.3.3	Processus itératif	136
6.3.4	À la convergence	139
6.3.5	Forme matricielle du problème	140
6.4	Problèmes locaux, Construction par Λ	142
6.4.1	Formulation faible et unicité de la solution	142
6.4.2	Opérateurs discrets de <i>scattering</i>	146
6.4.3	Processus itératif	147
6.4.4	À la convergence	150
6.4.5	Forme matricielle du problème	151
7	MISE EN ŒUVRE NUMÉRIQUE	155
7.1	Généralités	156
7.1.1	Les maillages	156
7.1.2	Codes numériques et choix des éléments	162
7.1.3	Algorithmes	162
7.1.3.1	Algorithme de Jacobi	166
7.1.3.2	Algorithme GMRES	167
7.2	Spécificités de l'opérateur Λ	168
7.2.1	Assemblage	168
7.2.1.1	Validation sur le cercle	171
7.2.1.2	Validation des propriétés de Λ	173
7.2.2	Réalisation de la troncature	177
7.2.2.1	Approche naïve	180

7.2.2.2	Approche par intégration par parties	181
8	RÉSULTATS NUMÉRIQUES	185
8.1	Cas de la géométrie circulaire	186
8.1.1	Étude de l'évolution du résidu pour plusieurs raffinements	187
8.1.1.1	$kR_1 = 10$	187
8.1.1.2	$kR_1 = 30$	195
8.1.2	Comparaison avec les valeurs propres analytiques	201
8.1.2.1	Calculs des valeurs propres numériques	201
8.1.2.2	Comparaison	201
8.1.3	Influence de la troncature	208
8.2	Cas du carré	210
8.3	Diverses géométries	216
8.3.1	Domaine étoilé	216
8.3.2	Cavité résonnante	220
8.4	Multi sous-domaines	222
9	CONCLUSIONS ET PERSPECTIVES	225
9.1	Conclusion	225
9.2	Sur les jonctions	227
9.3	Sur les conditions de transmission	228
9.4	Retour sur l'analyse modale	230
A	ANNEXES	233
A.1	Développement limité du noyau $B(s, t)$ en $s = t$	233
A.2	Calcul des coefficients	236
A.2.1	Un argument d'analytité	236
A.2.2	Calcul par déformation de contour	238

Table des figures

1.1	Représentation de la décomposition de domaine, telle qu'elle apparaît dans l'article de Schwarz en 1870	5
1.2	Exemples de décomposition de domaine avec recouvrement.	6
1.3	Exemple de décomposition de domaine sans recouvrement.	7
1.4	Taux de convergence modaux usuels pour les conditions de transmission d'ordre 0 (Després) et d'ordre 2, cas de deux demi-espaces infinis, interface plane.	13
2.1	Schéma représentatif du problème de diffraction étudié, avec décomposition du domaine en pelures d'oignon.	20
2.2	Schéma représentatif du problème de diffraction étudié, avec décomposition du domaine en pelures d'oignon.	22
3.1	Géométrie du problème circulaire simple. $\Omega_1 : R_1 < r < R_\Sigma$ et $\Omega_2 : R_\Sigma < r < R_2$	47
3.2	Représentation dans le plan complexe du produit $s_{1,m}s_{2,m}$ pour un opérateur de symbole t_m et pour l'identité (Després : $t_m = 1$). Nous avons utilisé $x = 5e^{-2}$ et $\beta = 15$	53
3.3	Représentation dans le plan complexe de $\tau_{1,m,\pm} = \pm\sqrt{s_{1,m}s_{2,m}}$ pour un opérateur de symbole t_m et pour l'identité (Després : $t_m = 1$). Nous avons utilisé $x = 5e^{-2}$ et $\beta = 15$	56
3.4	Action de la relaxation sur les valeurs propres du problème non relaxé (cas où $r = \frac{1}{2}$). En bleu, les valeurs propres non relaxées, en rouge les valeurs propres relaxées.	57
3.5	Représentation dans le plan complexe de $\tau_{r,m,\pm} = 1+r\pm r\sqrt{s_{1,m}s_{2,m}}$ pour un opérateur de symbole t_m et pour l'identité (Després : $t_m = 1$). Nous avons utilisé $x = 5e^{-2}$, $\beta = 15$ et $r = 0.58$	60
3.6	Géométrie du problème plan.	62
3.7	Géométrie du problème circulaire avec matériau.	72

3.8	Comparaison des taux de convergence modaux avec et sans optimisation pour des conditions de type Després (en haut) ou avec opérateur non local (en bas). L'optimisation est réalisée sur l'intervalle $[0, s_M]$, avec prise en compte de la limite lorsque $s \rightarrow \infty$ pour l'opérateur, et le graphique est tracé sur l'intervalle $[0, 2s_M]$	80
3.9	La section courbe (figure de gauche) et approchée par son plan tangent (figure de droite), et les épaisseurs de couche sont conservées	81
3.10	Comparaison des taux de convergence modaux du plan et du cercle, pour des conditions de transmission de type Després ou avec opérateur non local. Les conditions de transmission ne sont pas optimisées.	83
3.11	Comparaison des taux de convergence modaux du plan et du cercle, pour des conditions de transmission de type Després ou avec opérateur non local. Les conditions de transmission sont optimisées sur le plan et appliquées sur le cercle.	84
3.12	Représentation du taux de convergence maximal $\tau = \tau(k)$	86
3.13	Évolution des paramètres optimaux en fonction du nombre d'onde k . Les graphiques représentent respectivement $ z $, $\arg(z)$ (en degrés), β et le paramètre de relaxation r	87
5.1	Représentation graphique de $\alpha \mapsto G(\alpha)$, définie dans (5.53)	110
7.1	Exemple de deux types de triangulation autour de l'interface Σ	157
7.2	Exemples de ratio σ pour différents triangles, allant du ratio optimal (triangle équilatéral) au pire ratio trouvé dans les maillages étudiés	158
7.3	Répartition de la taille des arêtes d'un maillage relative à la taille objective h_0 pour différents raffinements	159
7.4	Répartition cumulée de la taille des arêtes d'un maillage relative à la taille objective h_0 pour différents raffinements. Par exemple, dans (f), 97% des arêtes ont une taille inférieure à $1.2h_0$	160
7.5	Répartition des ratios σ d'un maillage pour différents raffinements. Tous les ratios sont inférieurs à 7.5.	161
7.6	Valeurs propres $\lambda_N(n)$ de l'opérateur discrétisé Λ_0 avec condensation de la masse versus le mode n . Nous avons également représenté la fonction racine carrée. On utilise $N = 401$ points de discrétisation sur le cercle.	172

7.7	Valeurs propres $\lambda_N(n)$ de l'opérateur discrétisé Λ_0 sans condensation de la masse versus le mode n . Nous avons également représenté la fonction racine carrée. On utilise $N = 401$ points de discrétisation sur le cercle.	173
7.8	Représentation des valeurs propres pour différents maillages du cercle. On représente $\frac{n}{N} \mapsto \lambda_n^N + a$. $a = 0, 1, 2, 3$ correspond à $N = 200, 400, 800, 1600$, N est le nombre de points du maillage. La valeur propre nulle est envoyée sur l'ordonnée a	176
7.9	Représentation des valeurs propres pour différents maillages du cerf-volant. On représente $\frac{n}{N} \mapsto \lambda_n^N + a$. $a = 0, 1, 2, 3$ correspond à $N = 200, 400, 800, 1600$, N est le nombre de points du maillage. La valeur propre nulle est envoyée sur l'ordonnée a	176
7.10	Représentation des valeurs propres pour différents maillages de la goutte d'eau. On représente $\frac{n}{N} \mapsto \lambda_n^N + a$. $a = 0, 1, 2$ correspond à $N = 400, 800, 1600$, N est le nombre de points du maillage. La valeur propre nulle est envoyée sur l'ordonnée a	177
7.11	Représentation de la fonction de troncature utilisée. La fonction est de classe c^2	178
7.12	Caractéristiques des matrices des opérateurs Λ_0 et $\Lambda_{0,\chi}$. En bleu, pour la matrice \mathbb{L}_0 non tronquée. En rouge, pour la matrice $\mathbb{L}_{0,\chi}$ tronquée.	179
7.13	Caractéristiques des matrices des opérateurs Λ_0 et $\tilde{\Lambda}_{0,\chi}$. En bleu, pour la matrice \mathbb{L}_0 non tronquée. En rouge, pour la matrice $\tilde{\mathbb{L}}_{0,\chi}$ tronquée.	181
7.14	Caractéristiques des matrices des opérateurs Λ_0 et $\hat{\Lambda}_{0,\chi}$. En bleu, pour la matrice \mathbb{L}_0 non tronquée. En rouge, pour la matrice $\hat{\mathbb{L}}_{0,\chi}$ tronquée.	183
8.1	Représentation de la solution aux problèmes de diffraction du cercle pour différentes valeurs de k . En pointillé, est représentée l'interface de transmission Σ	188
8.2	Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme de Jacobi	190
8.3	Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme de Gauss-Seidel	192
8.4	Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme GMRES	194

8.5	Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme de Jacobi	196
8.6	Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme de Gauss-Seidel	198
8.7	Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme GMRES	199
8.8	Comparaison des valeurs propres numériques et analytiques pour le cas de la géométrie circulaire ($kR_1 = 10$). Les valeurs propres analytiques sont représentées par des cercles noirs, les valeurs propres numériques par des points de couleurs : En bleu 10 points par longueur d'onde, en vert 20 points, en rouge 30 points, en noir 60 points, en cyan 90 points, en magenta 180 points et en jaune 270 points.	203
8.9	Représentation des taux de convergence modaux analytiques et numériques pour différents raffinement, avec conditions de transmission de type Després. Le trait plein noir représente les taux modaux analytiques	204
8.10	Représentation des taux de convergence modaux analytiques et numériques pour différents raffinement, avec conditions de transmission avec opérateur non local. Le trait plein noir représente les taux modaux analytiques	205
8.11	Représentation des taux de convergence modaux analytiques et numériques pour différents raffinement, avec conditions de transmission de type Després. Le trait plein noir représente les taux modaux analytiques	206
8.12	Représentation des taux de convergence modaux analytiques et numériques pour différents raffinement, avec conditions de transmission avec opérateur non local. Le trait plein noir représente les taux modaux analytiques. Le trait mixte point-pointillé symbolise le taux modal analytique limite (≈ 0.86), tandis que le trait pointillé correspond au maximum de tous les taux numériques (≈ 0.96)	207
8.13	Évolution du nombre d'itérations nécessaire à la convergence de l'algorithme pour différents raffinements. À gauche, le nombre d'onde vaut $k = 10$, à droite $k = 30$. En haut, on utilise l'algorithme de Jacobi, et en bas l'algorithme GMRES. L'abscisse est normalisée et vaut $x = k\delta$ où δ est la longueur de la troncature.	209

8.14	Représentation de la solution aux problèmes de diffraction du carré pour différentes valeurs de k . En pointillé, est représentée l'interface de transmission Σ	211
8.15	Évolution des résidus pour différents raffinements et différents types de condition de transmission. Cas du nombre d'onde $k = 10$. . .	212
8.16	Représentation des taux de convergence modaux pour plusieurs raffinements. Cas du carré avec $k = 10$. Le trait mixte point-pointillé correspond à la limite théorique des taux modaux analytiques (sur le cercle). Le trait pointillé simple correspond au maximum de tous les taux modaux numériques.	214
8.17	Évolution des résidus pour différents raffinements et différents types de condition de transmission. Cas du nombre d'onde $k = 30$. . .	215
8.18	Solution au problème de diffraction d'une étoile à 5 branches. Les pointillés représentent l'interface de transmission	216
8.19	Exemples de maillages utilisés pour le problème du domaine étoilé. Le sous-domaine Ω_1 est en rouge, et Ω_2 en vert.	217
8.20	Historique de convergence des algorithmes de Jacobi et GMRES utilisant les conditions avec opérateur non local (8.20a et 8.20c) ou les conditions de type Padé (8.20b et 8.20d).	219
8.21	Présentation du problème de cavité résonnante. En haut, les domaines de calcul Ω (gris), Ω_1 (rouge) et Ω_2 (vert). En bas, la partie réelle et la valeur absolue de la solution (champ total).	220
8.22	Historique de convergence des algorithmes de Jacobi et GMRES pour différentes conditions de transmission, sur le cas de la cavité résonnante.	222
8.23	Évolution des résidus au cours des itérations des algorithmes de Jacobi et GMRES. À gauche, les conditions de transmissions non locales sont utilisées. À droite, nous utilisons les conditions de type Padé.	223
9.1	Présentation des découpages étudiés. Dans le cas de quatre sous-domaines, le carré est troué en son centre : les interfaces de transmission ne se touchent pas.	228
9.2	Erreur par rapport à la solution après 1000 itérations de l'algorithme de Jacobi (la convergence n'est pas atteinte). Le rayon du trou dans 9.2b est de $r = 0.001$	229
A.1	Schéma représentant le contour d'intégration.	239

Première partie
Aspects théoriques

1

État de l'art

Sommaire

1.1	Algorithmes avec recouvrement	5
1.2	Algorithmes sans recouvrement	7
1.2.1	Méthodes algébriques	7
1.2.1.1	Méthode FETI	8
1.2.1.2	Méthode FETI-DP	8
1.2.1.3	Autres méthodes algébriques	9
1.2.2	Méthodes variationnelles	10
1.2.2.1	Conditions de transmission de Després . . .	10
1.2.2.2	Conditions d'ordre supérieur	12
1.2.2.3	Approximation de Padé	13
1.3	Conditions de transmission idéales	14

Historiquement, la méthode de décomposition de domaine apparaît en 1870 dans [37]. Cet algorithme est avant tout une ingénieuse astuce utilisée par Schwarz dans le but de prouver l'unicité de la solution au problème de Dirichlet dans un domaine aux formes géométriques complexes. À l'époque, la preuve a déjà été établie pour

des domaines ayant des formes bien particulières. Notamment, Poisson résout le problème pour des domaines circulaires et Fourier sur des domaines rectangulaires. Schwarz développe alors une méthode dans laquelle il décompose un domaine en deux domaines dont les solutions sont connues (voir son dessin extrait de l'article original [37] figure 1.1) et imagine un processus itératif permettant, à partir de ces deux solutions locales uniquement, d'obtenir la solution sur le domaine entier.

Le problème de Dirichlet peut se présenter sous la forme suivante : Existe-t-il une unique solution u vérifiant

$$\begin{aligned} -\Delta u &= 0 \quad \text{dans } T = T_1 \cup T_2 \\ u &= g \quad \text{sur } \partial T \end{aligned} \tag{1.1}$$

Pour répondre à cette question, Schwarz construit une suite de solutions (u_1^n, u_2^n) vérifiant dans chacun des sous-domaines correspondant les équations suivantes

$$\begin{aligned} (a) \quad & -\Delta u_1^n = 0 \quad \text{dans } T_1 & & -\Delta u_2^n = 0 \quad \text{dans } T_2 \\ (b) \quad & u_1^n = g \quad \text{sur } L_0 = \partial T_1 \cap \partial T & & u_2^n = g \quad \text{sur } L_3 = \partial T_2 \cap \partial T \\ (c) \quad & u_1^n = u_2^{n-1} \quad \text{sur } L_2 = \partial T_1 \cap T_2 & & u_2^n = u_1^n \quad \text{sur } L_1 = \partial T_2 \cap T_1 \end{aligned} \tag{1.2}$$

Dans les équations (1.2), l'équation (c) correspond aux conditions de transmission. Elle relie les solutions aux problèmes locaux entre eux. Tout l'enjeu des méthodes de décomposition de domaine réside dans la manière d'échanger de l'information à travers l'interface entre deux sous-domaines. Cet échange est régi par les conditions de transmission et de ces conditions dépend la vitesse de convergence de l'algorithme, c'est à dire le nombre d'itérations nécessaire pour obtenir la solution.

Aujourd'hui, les méthodes de type décomposition de domaine sont utilisées principalement pour rendre parallèle un code de simulation numérique. La résolution du problème sur chaque sous-domaine peut en effet être réalisée indépendamment des autres à chaque itération, avant d'échanger des informations à travers les conditions de transmission.

Dans l'article de Schwarz, les conditions de transmission utilisées décrivent simplement la continuité de la solution u sur l'ensemble du domaine. Ces conditions permettent la convergence des méthodes de décomposition de domaine dans de nombreux cas. Cependant, dans le cas de problèmes indéfinis, de type Helmholtz

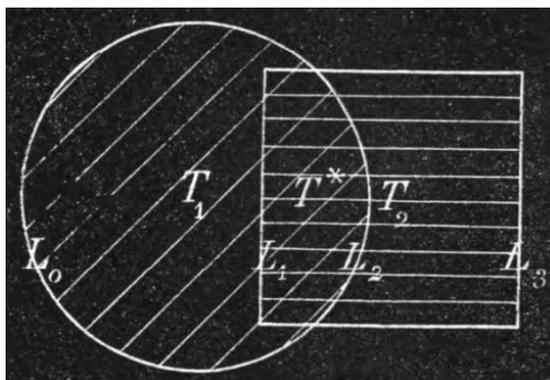


FIGURE 1.1 – Représentation de la décomposition de domaine, telle qu'elle apparaît dans l'article de Schwarz en 1870

par exemple, les algorithmes ne convergent pas. D'autres conditions sont donc nécessaires. Dans ce chapitre, nous passons en revue différentes techniques et conditions de transmission qui ont été développées pour améliorer la convergence des méthodes de décomposition de domaine dans différents cas. Pour un historique un peu plus détaillé sur les méthodes de Schwarz, nous renvoyons le lecteur à l'excellent article de Martin Gander [13] qui explique les similarités et différences des méthodes présentées ici à travers de nombreuses citations des contributeurs majeurs du sujet.

1.1 ALGORITHMES AVEC RECOUVREMENT

L'algorithme original de Schwarz fait partie de ces méthodes de décomposition de domaine dites *avec recouvrement*. Dans ces méthodes, l'intersection de deux sous-domaines voisins n'est pas réduite à leur interface, et les sous-domaines se chevauchent. La figure 1.1 montre une telle situation ; le domaine T_1 et le domaine T_2 se chevauchent et $T_1 \cap T_2 = T^*$. Intuitivement, l'existence d'une zone de recouvrement permet l'échange de davantage d'information à chaque itération, et donc suggère une meilleure convergence de la méthode. L'inconvénient principal est bien évidemment le surcoût engendré par ce recouvrement. Chaque domaine étant plus important, la résolution des problèmes locaux devient plus coûteuse.

Le recouvrement a longtemps été nécessaire pour obtenir des preuves de convergence. Pierre-Louis Lions a réalisé en 1988, dans [25], une étude théorique des méthodes de décomposition de domaine avec recouvrement, et leurs applications

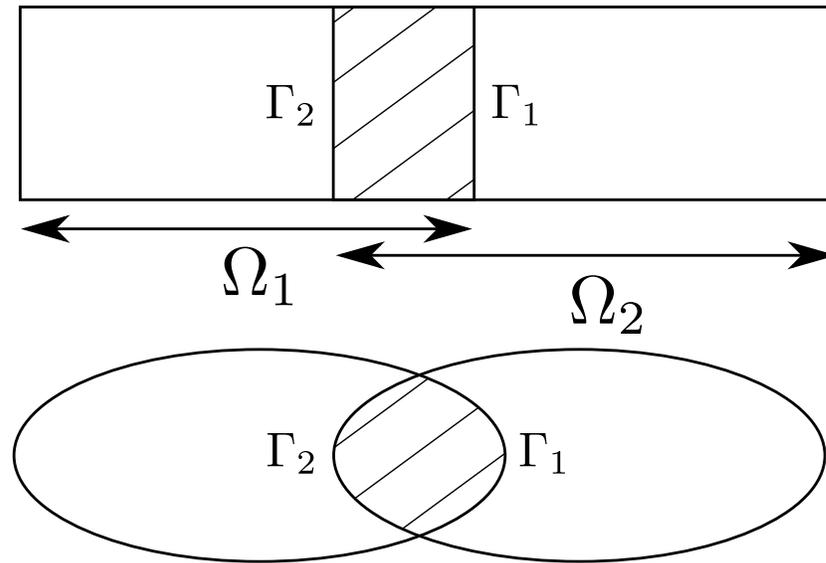


FIGURE 1.2 – Exemples de décomposition de domaine avec recouvrement.

à différents problèmes physiques. Lions démontre ainsi la convergence du processus itératif de décomposition de domaine lorsqu'il est appliqué à l'équation de la chaleur, ou à l'équation de Stokes. La convergence peut même être une convergence exponentielle, si certaines conditions de régularité sur le recouvrement sont respectées. La figure 1.2 présente deux exemples de décomposition de domaine avec recouvrement. Des conditions de transmission sont ensuite appliquées au niveau des interfaces Γ_1 et Γ_2 . Dans l'article de Lions, de simples conditions de Dirichlet sont imposées : ce sont les mêmes que celles utilisées par Schwarz. Ces conditions ne sont en générales pas suffisantes pour obtenir la convergence de méthodes de décomposition de domaine dans le cas de problèmes non elliptiques, comme l'équation de Helmholtz.

D'autres auteurs se sont donc intéressés aux conditions de transmission pour les méthodes avec recouvrement pour ce type d'équations. Par exemple, Xiao Cai, Mario Casarin, Frank Elliott et Olof Widlund [8] utilisent des conditions de transmission de Després (voir Section 1.2.2.1), tandis que Christiaan Stolk [39] utilise les bonnes propriétés des PML (*Perfectly Match Layers*) comme conditions de transmission.

Même si les méthodes avec recouvrement présentent plusieurs avantages, nous

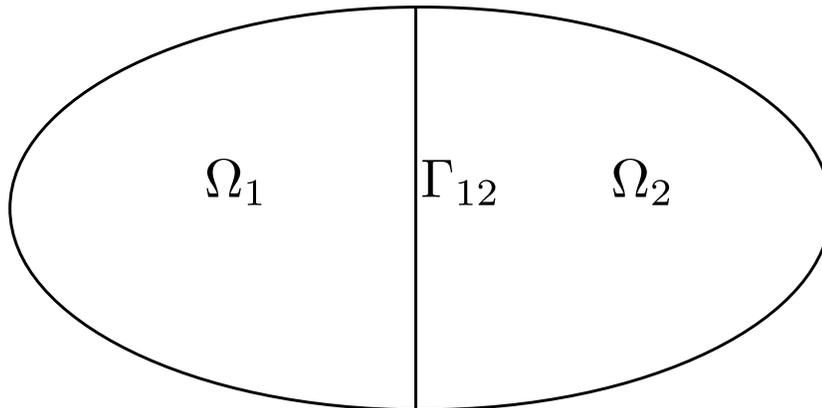


FIGURE 1.3 – Exemple de décomposition de domaine sans recouvrement.

nous focaliserons sur les méthodes sans recouvrement. En effet, le surcoût engendré par la zone de recouvrement peut devenir prohibitif, notamment dans le cas d'une décomposition de domaine en couches fines successives. Plus les couches sont fines, plus le recouvrement devient prépondérant dans le temps de calcul. Les méthodes sans recouvrement nous paraissent donc plus adaptées à nos applications.

1.2 ALGORITHMES SANS RECOUVREMENT

Dans les méthodes de décomposition de domaine sans recouvrement, l'intersection de deux sous-domaines est soit vide si les sous-domaines ne sont pas voisins, soit réduite à une interface. L'exemple présenté figure 1.3 montre une telle situation. Parmi ces méthodes sans recouvrement, deux classes d'algorithmes ressortent. On trouve ainsi les méthodes dites *algébriques* et les méthodes *variationnelles*.

1.2.1 MÉTHODES ALGÈBRIQUES

Les méthodes de décomposition de domaine algébriques appliquent la décomposition directement sur le problème global discrétisé. Après discrétisation, le problème considéré s'écrit sous la forme d'un système linéaire

$$\mathbb{Z}u = f. \tag{1.3}$$

1.2.1.1 MÉTHODE FETI

La méthode FETI, pour *Finite Element Tearing and Interconnecting*, est une méthode algébrique introduite par Charbel Farhat et François-Xavier Roux [12]. Dans cette méthode, les inconnues de chacun des domaines sont reliées par l'intermédiaire de multiplicateurs de Lagrange λ correspondant à la contrainte $u_1 = u_2$ sur l'interface, et le problème devient

$$\begin{aligned} \mathbb{Z}_1 u_1 &= f_1 + \mathbb{B}_1^T \lambda \\ \mathbb{Z}_2 u_2 &= f_2 - \mathbb{B}_2^T \lambda, \\ \mathbb{B}_1 u_1 &= \mathbb{B}_2 u_2 \end{aligned} \tag{1.4}$$

où les matrices \mathbb{B}_j sont des matrices booléennes permettant de récupérer les inconnues de chaque domaine respectivement, tandis que les matrices \mathbb{Z}_1 et \mathbb{Z}_2 représentent les problèmes locaux. Si ces problèmes locaux sont inversibles, le système (1.4) devient

$$\begin{aligned} (\mathbb{B}_1 \mathbb{Z}_1^{-1} \mathbb{B}_1^T + \mathbb{B}_2 \mathbb{Z}_2^{-1} \mathbb{B}_2^T) \lambda &= \mathbb{B}_2 \mathbb{Z}_2^{-1} f_2 - \mathbb{B}_1 \mathbb{Z}_1^{-1} f_1 \\ u_1 &= \mathbb{Z}_1^{-1} (f_1 + \mathbb{B}_1^T \lambda) \\ u_2 &= \mathbb{Z}_2^{-1} (f_2 - \mathbb{B}_2^T \lambda) \end{aligned} \tag{1.5}$$

et le problème revient à résoudre l'équation pour les multiplicateurs de Lagrange. Cependant, il n'existe aucune garantie que les problèmes locaux \mathbb{Z}_1 et \mathbb{Z}_2 soient inversibles. Il est donc nécessaire d'effectuer un traitement spécifique si le cas se produit, notamment calculer les pseudos-inverses, et ajouter une condition sur le second membre pour expliciter le fait qu'il ne possède pas de composante dans le noyau de la matrice non inversible.

1.2.1.2 MÉTHODE FETI-DP

Une extension de la méthode FETI a été développée pour améliorer son comportement dans le cas de nombreux sous-domaines, elle est proposée par Charbel Farhat, Michel Lesoinne, Patrick LeTallec, Kendall Pierson et Daniel Rixen [11]. Dans cette méthode, un traitement spécifique est appliqué aux inconnues placées sur les coin des sous-domaines, c'est-à-dire placées sur plusieurs interface. Dans chaque sous-domaine, les inconnues u_i sont séparées en $u_{i,r}$ et $u_{i,c}$ où $u_{i,c}$ sont les inconnues aux coins, et $u_{i,r}$ sont les inconnues restantes. Ensuite, pour résoudre le système (1.4), seules les inconnues $u_{i,r}$ sont éliminées. Il reste donc les

multiplicateurs de Lagrange ainsi que les inconnues aux coins. L'élimination des inconnues aux coins conduit à résoudre un problème de taille très réduite, mais sur l'ensemble du domaine global : c'est un problème auxiliaire dit *grossier*. L'utilisation de ce problème auxiliaire permet d'accélérer la convergence dans le cas de nombreux sous-domaines : au lieu de propager de l'information sous-domaine par sous-domaine, en ne faisant intervenir que les voisins, ce problème grossier relie l'ensemble des sous-domaines.

Le terme FETI-DP, pour *Dual-Primal*, vient du fait que la résolution du problème fait intervenir, contrairement à la méthode FETI, des inconnues duales (les multiplicateurs de Lagrange) ainsi que des inconnues primales (les degrés de liberté aux coins). Notons là encore qu'aucune garantie n'existe concernant l'inversibilité des problèmes locaux.

Cette méthode, bien que très intéressante, semble peu adaptée à notre situation où les interfaces ne s'intersectent pas.

1.2.1.3 AUTRES MÉTHODES ALGÈBRIQUES

D'autres méthodes algébriques ont été proposées par Frédéric Magoulès, François-Xavier Roux et Stéphanie Salmon [28] pour l'équation de Helmholtz. En partant d'une décomposition des inconnues u en u_1 , u_2 et u_Γ , où u_1 correspond aux inconnues intérieures du domaine Ω_1 , u_2 celles du domaine Ω_2 et u_Γ les inconnues de l'interface Γ , le système linéaire initial se réécrit

$$\begin{bmatrix} Z_1 & 0 & Z_{1\Gamma} \\ 0 & Z_2 & Z_{2\Gamma} \\ Z_{\Gamma 1} & Z_{\Gamma 2} & Z_{\Gamma\Gamma} \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ u_\Gamma \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_\Gamma \end{bmatrix}. \quad (1.6)$$

De manière équivalente, il est possible de séparer le système en deux systèmes

$$\begin{bmatrix} Z_1 & Z_{1\Gamma} \\ Z_{\Gamma 1} & Z_{1\Gamma\Gamma} \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_\Gamma \end{bmatrix} = \begin{bmatrix} f_1 \\ f_\Gamma - Z_{\Gamma 2}u_2 - Z_{2\Gamma\Gamma}u_\Gamma \end{bmatrix}, \quad (1.7)$$

$$\begin{bmatrix} Z_2 & Z_{2\Gamma} \\ Z_{\Gamma 2} & Z_{2\Gamma\Gamma} \end{bmatrix} \cdot \begin{bmatrix} u_2 \\ u_\Gamma \end{bmatrix} = \begin{bmatrix} f_2 \\ f_\Gamma - Z_{\Gamma 1}u_1 - Z_{1\Gamma\Gamma}u_\Gamma \end{bmatrix}, \quad (1.8)$$

où $Z_{1\Gamma} + Z_{2\Gamma} = Z_{\Gamma}$. Les auteurs montrent alors que le système augmenté suivant (1.9) est bien posé et équivalent au problème initial

$$\begin{aligned} \begin{bmatrix} Z_1 & Z_{1\Gamma} \\ Z_{\Gamma 1} & Z_{1\Gamma\Gamma} + A_1 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_{1\Gamma} \end{bmatrix} &= \begin{bmatrix} f_1 \\ f_{1\Gamma} + \lambda_1 \end{bmatrix} \\ \begin{bmatrix} Z_2 & Z_{2\Gamma} \\ Z_{\Gamma 2} & Z_{2\Gamma\Gamma} + A_2 \end{bmatrix} \cdot \begin{bmatrix} u_2 \\ u_{2\Gamma} \end{bmatrix} &= \begin{bmatrix} f_2 \\ f_{2\Gamma} + \lambda_2 \end{bmatrix} \\ u_{1\Gamma} - u_{2\Gamma} &= 0 \\ \lambda_1 + \lambda_2 - A_1 u_{1\Gamma} - A_2 u_{2\Gamma} &= 0 \end{aligned} \tag{1.9}$$

pour toutes matrices A_1 et A_2 , et où $u_{1\Gamma} = u_{2\Gamma} = u_{\Gamma}$ et $f_{1\Gamma} + f_{2\Gamma} = f_{\Gamma}$. Partant de ce système (1.9), les auteurs dérivent plusieurs méthodes de résolution en utilisant différentes combinaisons pour les matrices A_1 et A_2 :

- Cas de matrices augmentées nulles $A_1 = A_2 = 0$: se réduit au complément de Schur dual (matrice condensée à l'interface)
- Cas de matrices augmentées opposées $A_1 = -A_2 = A$: se réduit à une méthode de type FETI
- Cas général : agit comme un préconditionneur pour le système à l'interface.

L'optimum est alors de prendre pour A_j le complément de Schur du problème extérieur au domaine Ω_j considéré : la méthode converge alors en 2 itérations. Ce complément de Schur étant coûteux à calculer, plusieurs méthodes sont proposées par les auteurs pour en calculer une approximation.

1.2.2 MÉTHODES VARIATIONNELLES

1.2.2.1 CONDITIONS DE TRANSMISSION DE DESPRÉS

Bruno Després, en 1991, a été l'un des premiers à obtenir des résultats théoriques de convergence dans le cas général pour les méthodes de décomposition de domaine appliquées au problème indéfini de Helmholtz dans le vide

$$\begin{cases} -\Delta u - k^2 u = f & \text{dans } \Omega \\ \partial_n u + iku = g & \text{sur } \partial\Omega \end{cases} . \tag{1.10}$$

Notons que dans le système précédent, nous avons adopté pour la dépendance en temps la convention $e^{i\omega t}$. Nous conserverons cette convention dans l'ensemble du document. Dans sa thèse [10], Després développe des nouvelles conditions de transmission, de type conditions impédantes, pour lesquelles une preuve de convergence peut être écrite. L'idée est de réécrire les conditions de transmissions usuelles, qui correspondent à la continuité de la solution et de sa dérivée normale,

$$u_1 = u_2, \quad \partial_n u_1 = \partial_n u_2, \quad \text{sur } \partial\Omega_1 \cap \partial\Omega_2 \quad (1.11)$$

en les recombinaut en conditions de type Robin, à l'aide d'un coefficient particulier ik

$$\begin{aligned} \partial_n u_1 + ik u_1 &= \partial_n u_2 + ik u_2, \\ \partial_n u_1 - ik u_1 &= \partial_n u_2 - ik u_2 \end{aligned}, \quad \text{sur } \partial\Omega_1 \cap \partial\Omega_2. \quad (1.12)$$

Ces équations (1.11) et (1.12) sont équivalentes. Després montre alors que tous les problèmes locaux sont bien posés, c'est-à-dire qu'ils admettent tous une unique solution, et de plus, la convergence de la méthode vers la solution initiale est assurée.

Ces nouvelles conditions de transmission ont été très largement utilisées dans la littérature, nous citerons par exemple les travaux de Jin-Fa Lee et Marinos Vouvakis ([24], [43] et [42]). Une version optimisée des conditions de transmission de Després est également apparue, avec l'utilisation d'un autre coefficient pour recombinaut les conditions usuelles

$$\begin{aligned} \partial_n u_1 + ik\alpha_1 u_1 &= \partial_n u_2 + ik\alpha_1 u_2, \\ \partial_n u_1 - ik\alpha_2 u_1 &= \partial_n u_2 - ik\alpha_2 u_2 \end{aligned}, \quad \text{sur } \partial\Omega_1 \cap \partial\Omega_2. \quad (1.13)$$

où les paramètres α_1 et α_2 sont complexes. Ils servent de paramètres d'optimisation, et diverses techniques ont été développées pour trouver les coefficients permettant la meilleure convergence. Dans la plupart des cas, une analyse modale d'un cas simple (cercle ou plan) est utilisée pour déterminer ces paramètres. On notera par exemple les travaux de Martin Gander, Laurence Halpern et Frédéric Magoulès [14], ceux de Yassine Boubendir [4] ou encore ceux de Yvon Maday [26]. Cependant, peu de théorie a été développée pour l'utilisation de ces paramètres optimaux, notamment sur des conditions d'existence et d'unicité des solutions. De plus, la convergence de la méthode de décomposition de domaine reste lente dans certains cas, et c'est pourquoi des conditions de transmission d'ordre supérieur ont été développées.

1.2.2.2 CONDITIONS D'ORDRE SUPÉRIEUR

L'idée ici est d'ajouter aux conditions de transmission de Després optimisées (1.13) un opérateur de dérivation tangentielle d'ordre 2. Les conditions de transmission s'écrivent alors

$$\begin{aligned} \partial_n u_1 + ik\alpha_1 u_1 + \beta_1 \partial_{tt} u_1 &= \partial_n u_2 + ik\alpha_1 u_2 + \beta_1 \partial_{tt} u_2 \\ \partial_n u_1 - ik\alpha_2 u_1 - \beta_2 \partial_{tt} u_1 &= \partial_n u_2 - ik\alpha_2 u_2 - \beta_2 \partial_{tt} u_2 \end{aligned}, \quad \text{sur } \partial\Omega_1 \cap \partial\Omega_2. \quad (1.14)$$

Là encore, de nombreux auteurs se sont intéressés à l'amélioration des conditions de transmission de Després et ont testé différentes combinaisons pour les coefficients α_j , β_j . L'analyse repose toujours sur une décomposition en ondes planes (ou circulaires) de la solution

$$u_j(x) = \int \hat{u}_j(ks) e^{iksx} ds, \quad (1.15)$$

où s est le paramètre qui caractérise le mode de Fourier. Par définition, $|s| < 1$ correspond aux modes dits *propagatifs*, $|s| > 1$ correspond aux modes *évanescents*. Dans cette direction, mentionnons par exemple l'article de Martin Gander, Frédéric Magoulès et Frédéric Nataf [15] dans lequel les auteurs précisent des conditions d'optimalité pour différentes configurations. Ainsi, l'utilisation des paramètres

$$\alpha_1 = \alpha_2 = ik, \quad \beta_1 = \beta_2 = \frac{1}{2ik} \quad (1.16)$$

est optimal pour la convergence des modes propagatifs et correspond à l'extension de la méthode de Després. D'autres expressions analytiques peuvent être dérivées pour l'optimisation sur une bande de modes plus large.

La figure 1.4 représente les taux de convergence modaux pour des méthodes de décomposition de domaine utilisant d'une part les conditions de transmission de Després et d'autre part des conditions d'ordre 2. On voit clairement que les conditions de Després sont particulièrement efficaces pour les modes les plus bas (modes propagatifs, $s < 1$), mais non pour les modes élevés (modes évanescents, $s > 1$). L'ajout de l'ordre 2 dans les conditions de transmission permet, au moins sur ce modèle très simplifié, de pallier ce problème. D'autres auteurs ont développé des méthodes d'optimisation pour les conditions de transmission d'ordre 2, appliquées à des problèmes spécifiques. Citons par exemple Bruno Stupfel [41], Zhen Peng et Jin-Fa Lee ([32], [31]), et Frédéric Magoulès, Peter Iványi et Barry Topping [27].

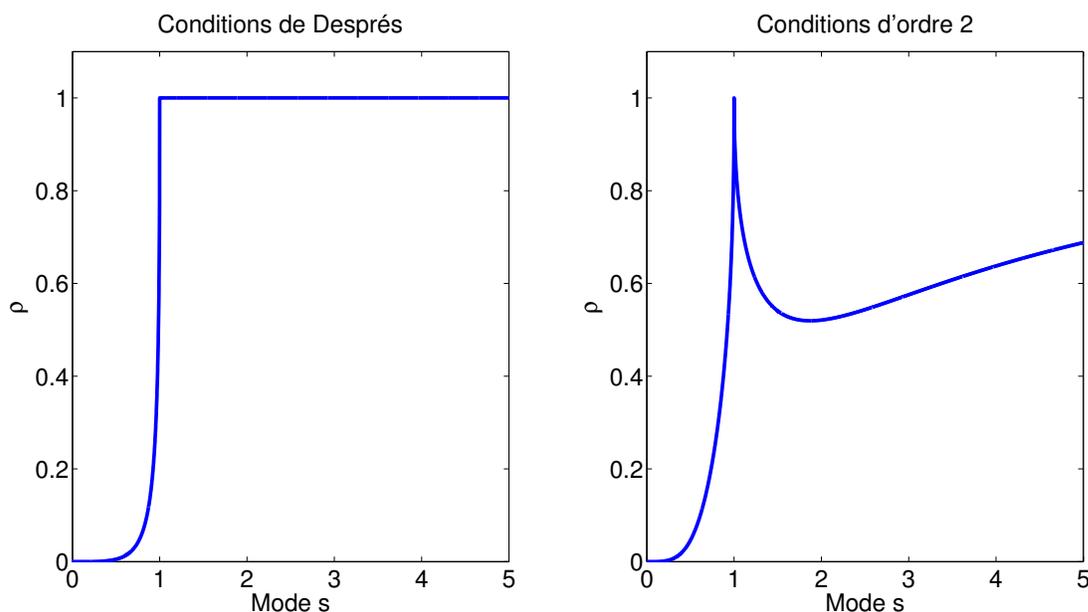


FIGURE 1.4 – Taux de convergence modaux usuels pour les conditions de transmission d'ordre 0 (Després) et d'ordre 2, cas de deux demi-espaces infinis, interface plane.

Pour résumer, l'obtention de conditions optimisées repose sur des analyses dites *modales*. On choisit une famille paramétrée d'opérateurs et on détermine les paramètres en essayant de contrôler les taux de convergence. Nous reprendrons cette méthodologie dans le chapitre 3 pour optimiser la classe d'opérateur que nous serons amenés à construire. Pour finir, remarquons que les conditions de transmission proposées dans la littérature sont fortement liées aux opérateurs apparaissant dans les conditions dites absorbantes (voir par exemple l'article de Bruno Stupfel [40] ou le chapitre 5 écrit par Dan Givoli dans [18]).

1.2.2.3 APPROXIMATION DE PADÉ

L'un des travaux qui nous a paru les plus aboutis est l'article de Yassine Boubendir, Xavier Antoine et Christophe Geuzaine [5]. Ils ont récemment utilisé des travaux réalisés sur l'approximation de conditions absorbantes pour des problèmes haute fréquence [21]. Dans le cas d'une interface plane, la condition absorbante

s'écrit

$$\partial_n u + ik \underbrace{\sqrt{1 + \partial_t \frac{1}{k^2} \partial_t u}}_{\mathcal{S}} = 0 \quad (1.17)$$

où ∂_t représente la dérivée tangentielle et l'opérateur \mathcal{S} est un opérateur pseudo-différentiel. L'opérateur \mathcal{S} étant très coûteux à calculer, il est approché par une somme d'opérateurs s'écrivant sous la forme de fractions rationnelles. Pour cela, les auteurs utilisent l'approximation de Padé complexe suivante

$$\sqrt{1+z} \approx C_0 + \sum_{j=1}^N \frac{A_j z}{1+B_j z}. \quad (1.18)$$

Les valeurs explicites des coefficients C_0 , A_j et B_j peuvent être trouvées dans [5] et font intervenir les valeurs des coefficients de l'approximation de Padé réelle. L'utilisation de fractions rationnelles permet de rendre l'opérateur de transmission complètement non local, contrairement au cas des conditions de Després ou d'ordre supérieur.

L'analyse de cette condition de transmission reste cependant limitée à une analyse modale de l'opérateur non approché \mathcal{S} sur une géométrie simple (cercle), et aucune preuve de convergence ou même d'unicité n'est donnée. En pratique, appliquée à l'équation de Helmholtz avec des coefficients constants, cette méthode donne de très bons résultats, comme nous pourrons le voir dans le chapitre 8 concernant les résultats numériques.

1.3 CONDITIONS DE TRANSMISSION IDÉALES

Il est intéressant de noter que, pour un problème donné, il existe des conditions de transmission dites *idéales*, c'est-à-dire permettant la convergence de la méthode de décomposition de domaine en 2 itérations. Prenons le problème modèle de Helmholtz suivant sur le domaine Ω , borné intérieurement par Γ_0 et extérieurement par Γ

$$\begin{aligned} -\operatorname{div}(\mu^{-1} \nabla u) - k^2 \epsilon u &= f \quad \text{dans } \Omega \\ \mu^{-1} \partial_\nu u + ik \eta^{-1} u &= 0 \quad \text{sur } \Gamma \quad . \\ u &= 0 \quad \text{sur } \Gamma_0 \end{aligned} \quad (1.19)$$

Par simplicité, dans cette section, nous supposons que les paramètres physiques μ , ϵ et η sont strictement positifs et suffisamment réguliers pour garantir que le

problème est bien posé. Dans l'équation (1.19), la dérivée normale ∂_ν est classiquement définie par $\nu \cdot \nabla$, où ν est la normale sortante à Ω . Le domaine Ω est décomposé en deux domaines Ω_1 et Ω_2 par l'ajout d'une interface Σ . Toutes les quantités relatives au domaine Ω_1 (respectivement Ω_2) seront notées avec l'indice 1 (respectivement l'indice 2). Une définition plus précise des notations utilisées ici, ainsi que dans le reste du document, sera l'objet de la section 2.1. Les solutions locales dans chacun des sous-domaines vérifient les équations

$$\begin{aligned} -\operatorname{div}(\mu^{-1}\nabla u_1) - k^2\epsilon u_1 &= f_1 \text{ dans } \Omega_1 & -\operatorname{div}(\mu^{-1}\nabla u_2) - k^2\epsilon u_2 &= f_2 \text{ dans } \Omega_2 \\ u &= 0 \text{ sur } \Gamma_0 & \mu^{-1}\partial_\nu u_2 + ik\eta^{-1}u_2 &= 0 \text{ sur } \Gamma \end{aligned} \quad (1.20)$$

auxquelles il est nécessaire d'ajouter des conditions de transmission. Pour les définir, nous introduisons les opérateurs T_1 et T_2 définis sur l'interface Σ par

$$T_1 \varphi = \frac{1}{ik\mu} \partial_{\nu_1} v_1(\varphi), \quad T_2 \varphi = \frac{1}{ik\mu} \partial_{\nu_2} v_2(\varphi), \quad (1.21)$$

où les fonctions $v_1 = v_1(\varphi)$ et $v_2 = v_2(\varphi)$ sont solutions des problèmes dans Ω_1 et Ω_2 respectivement

$$T_1 : \begin{cases} -\operatorname{div}(\mu^{-1}\nabla v_1) - k^2\epsilon v_1 = 0 & \text{dans } \Omega_1 \\ v_1 = 0 & \text{sur } \Gamma_0 \\ v_1 = \varphi & \text{sur } \Sigma \end{cases} \quad (1.22)$$

$$T_2 : \begin{cases} -\operatorname{div}(\mu^{-1}\nabla v_2) - k^2\epsilon v_2 = 0 & \text{dans } \Omega_2 \\ \mu^{-1}\partial_\nu v_2 + ik\eta^{-1}v_2 = 0 & \text{sur } \Gamma \\ v_2 = \varphi & \text{sur } \Sigma \end{cases} \quad (1.23)$$

À l'aide de ces opérateurs, nous définissons les conditions de transmission par

$$\begin{aligned} \mu^{-1}\partial_{\nu_1} u_1 + ik T_2 u_1 &= -\mu^{-1}\partial_{\nu_2} u_2 + ik T_2 u_2, \\ \mu^{-1}\partial_{\nu_2} u_2 + ik T_1 u_2 &= -\mu^{-1}\partial_{\nu_1} u_1 + ik T_1 u_1. \end{aligned} \quad (1.24)$$

Alors, le processus itératif, tel que décrit par Schwarz, produit une suite de couple de solutions (u_1^n, u_2^n) vérifiant chacune l'équation 1.20 ainsi que les conditions sur

le bord Σ

$$\begin{aligned} \mu^{-1} \partial_{\nu_1} u_1^n + ik T_2 u_1^n &= -\mu^{-1} \partial_{\nu_2} u_2^{n-1} + ik T_2 u_2^{n-1}, \\ \mu^{-1} \partial_{\nu_2} u_2^n + ik T_1 u_2^n &= -\mu^{-1} \partial_{\nu_1} u_1^{n-1} + ik T_1 u_1^{n-1}. \end{aligned} \quad (1.25)$$

Ce processus itératif converge en deux itérations. Pour prouver cette convergence, introduisons les erreurs (e_1^n, e_2^n) sur chaque domaine

$$e_1^n = u_1^n - u_1, \quad e_2^n = u_2^n - u_2. \quad (1.26)$$

Alors, pour tout $n \geq 1$ et pour toute condition initiale (e_1^0, e_2^0) , le couple d'erreur (e_1^n, e_2^n) satisfait les équations suivantes

$$\begin{cases} -\operatorname{div}(\mu^{-1} \nabla e_1^n) - k^2 \epsilon e_1^n &= 0 \text{ dans } \Omega_1 \\ e_1^n &= 0 \text{ sur } \Gamma_0 \\ \mu^{-1} \partial_{\nu_1} e_1^n + ik T_2 e_1^n &= -\mu^{-1} \partial_{\nu_2} e_2^{n-1} + ik T_2 e_2^{n-1}, \end{cases} \quad (1.27)$$

$$\begin{cases} -\operatorname{div}(\mu^{-1} \nabla e_2^n) - k^2 \epsilon e_2^n &= 0 \text{ dans } \Omega_2 \\ \mu^{-1} \partial_{\nu} e_2^n + ik \eta^{-1} e_2^n &= 0 \text{ sur } \Gamma \\ \mu^{-1} \partial_{\nu_2} e_2^n + ik T_1 e_2^n &= -\mu^{-1} \partial_{\nu_1} e_1^{n-1} + ik T_1 e_1^{n-1}. \end{cases} \quad (1.28)$$

Or, par définition de l'opérateur T_1 (1.22) et parce que e_1^n vérifie les deux premières équations de (1.27), on a pour tout $n \geq 1$

$$T_1 e_1^n = \frac{1}{ik\mu} \partial_{\nu_1} e_1^n \Leftrightarrow -\mu^{-1} \partial_{\nu_1} e_1^n + ik T_1 e_1^n = 0 \quad (1.29)$$

De la même manière, par définition de l'opérateur T_2 (1.23), et grâce à (1.28), e_2^{n-1} vérifie pour tout $n \geq 1$

$$T_2 e_2^n = \frac{1}{ik\mu} \partial_{\nu_2} e_2^n \Leftrightarrow -\mu^{-1} \partial_{\nu_2} e_2^n + ik T_2 e_2^n = 0 \quad (1.30)$$

En particulier, ces égalités sont vraies pour $n = 1$, et alors, le couple d'erreur vérifie à l'itération $n = 2$

$$\begin{cases} -\operatorname{div}(\mu^{-1} \nabla e_1^2) - k^2 \epsilon e_1^2 &= 0 \text{ dans } \Omega_1 \\ e_1^2 &= 0 \text{ sur } \Gamma_0 \\ \mu^{-1} \partial_{\nu_1} e_1^2 + ik T_2 e_1^2 &= -\mu^{-1} \partial_{\nu_2} e_2^1 + ik T_2 e_2^1 = 0, \end{cases} \quad (1.31)$$

$$\left\{ \begin{array}{l} -\operatorname{div}(\mu^{-1}\nabla e_2^2) - k^2 \epsilon e_2^2 = 0 \text{ dans } \Omega_2 \\ \mu^{-1}\partial_\nu e_2^2 + ik \eta^{-1} e_2^2 = 0 \text{ sur } \Gamma \\ \mu^{-1}\partial_{\nu_2} e_2^2 + ik T_1 e_2^2 = -\mu^{-1}\partial_{\nu_1} e_1^1 + ik T_1 e_1^1 = 0. \end{array} \right. \quad (1.32)$$

Ces équations conduisent à la solution $e_1^2 = e_2^2 = 0$. L'algorithme converge bien en deux itérations. Évidemment, cette remarque n'est que d'un intérêt purement théorique et n'a pas de conséquences pratiques. En effet, déterminer les opérateurs T_{int} et T_{ext} est plus difficile que de résoudre le problème initial. Cependant, ces opérateurs peuvent donner une indication pour construire de bons opérateurs de transmission et c'est l'idée derrière la construction d'opérateurs d'ordre élevé (Section 1.2.2.2) ou d'approximation par des fractions rationnelles de type Padé (Section 1.2.2.3).

Pour chacune des méthodes proposées dans la littérature et exposées dans ce chapitre, plusieurs interrogations communes se posent, notamment :

- À chaque itération, les problèmes locaux sont-ils bien posés ?
- Est-on assuré de la convergence de la méthode ?
- Y a-t-il des conditions spécifiques pour assurer la convergence (conditions sur la régularité du domaine, des sous-domaines, des paramètres physiques, sur les coefficients des conditions de transmission) ?
- La convergence est-elle toujours garantie au niveau discret (si oui, quelle est la qualité de l'approximation obtenue) ?
- Quelle est la vitesse de convergence : est-elle géométrique, algébrique ?

Toutes ces questions définissent la qualité, l'utilité et la robustesse des conditions de transmission et nous répondrons dans ce document à toutes ces questions concernant les conditions de transmission optimisées que nous proposons.

2

Théorie générale

Sommaire

2.1	Problème modèle et notations	20
2.1.1	Problème modèle	21
2.1.2	Problème décomposé	21
2.1.3	Processus itératif	23
2.1.4	Espaces fonctionnels	26
2.2	Classe d'opérateurs de transmission robustes	27
2.3	Étude théorique	30
2.3.1	Caractère bien posé des problèmes	30
2.3.2	Reformulation à l'interface pour l'erreur	34
2.3.3	Reformulation à l'interface pour la solution	36
2.4	Résultats de convergence	37
2.5	Extension à d'autres méthodes itératives	42

Dans ce chapitre, nous exposerons les nouvelles conditions de transmission que nous avons développées, et en ferons une étude théorique détaillée, dans un cadre général. Basées sur des conditions d'impédance, elles font intervenir des opérateurs

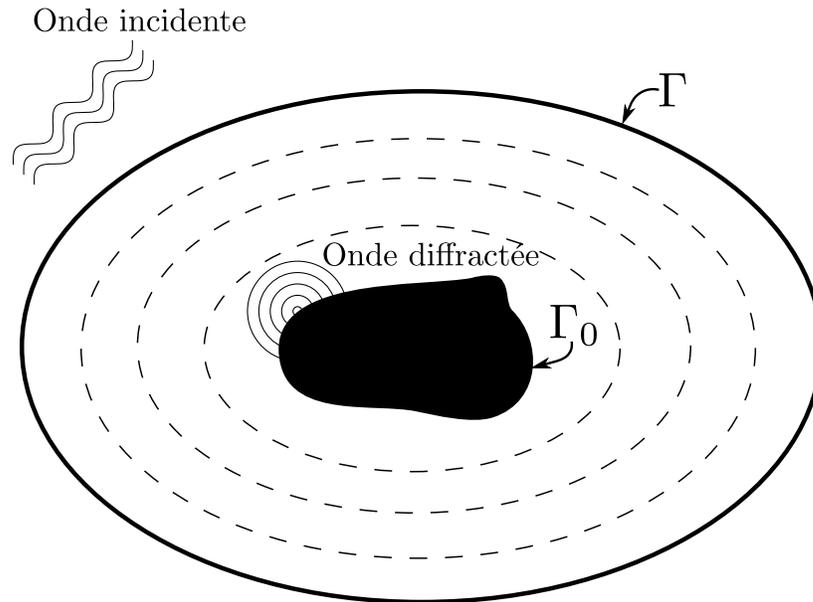


FIGURE 2.1 – Schéma représentatif du problème de diffraction étudié, avec décomposition du domaine en pelures d'oignon.

de transmission aux propriétés bien spécifiques. Compte tenu du caractère parfois technique des démonstrations, une première section détaillera les notations que l'on tentera de garder dans l'ensemble du document et explicitera le cheminement du chapitre.

2.1 PROBLÈME MODÈLE ET NOTATIONS

Nous nous sommes intéressés à la décomposition de domaine appliquée à la propagation d'onde en régime harmonique. Plus précisément, la finalité de ces travaux est la simulation numérique de la diffraction électromagnétique d'objets recouverts de couches de matériaux. Compte tenu du caractère spécifique du problème, nous nous intéresserons particulièrement à la décomposition du domaine en pelures d'oignon : chaque sous-domaine est en contact avec au plus deux sous-domaines voisins, et les interfaces sont fermées et non intersectantes (voir un exemple de décomposition présenté figure 2.1). Une discussion sur ces hypothèses sera effectuée à la fin du document.

2.1.1 PROBLÈME MODÈLE

Pour analyser la méthode de décomposition de domaine, nous utiliserons le problème modèle suivant. Soit un domaine Ω de \mathbb{R}^{d+1} fermé par une interface Γ . Le domaine Ω possède une frontière intérieure Γ_0 sur laquelle une condition de type conducteur parfait est appliquée. Le problème est alors régi par les équations

$$\begin{aligned} -\operatorname{div}(\mu^{-1}\nabla u) - k^2\epsilon u &= f \quad \text{dans } \Omega & (a) \\ \mu^{-1}\partial_\nu u + ik\eta^{-1}u &= 0 \quad \text{sur } \Gamma & (b) \\ u &= 0 \quad \text{sur } \Gamma_0, & (c) \end{aligned} \tag{2.1}$$

où la solution u est cherchée dans

$$H_{\Gamma_0}^1(\Omega) = \left\{ u \in H^1(\Omega) \text{ tel que } u|_{\Gamma_0} = 0 \right\}. \tag{2.2}$$

Dans (2.1), l'équation (a) correspond à l'équation de Helmholtz scalaire à coefficients variables. Le nombre d'onde dans le vide k est strictement positif, le terme source f est supposé dans $L^2(\Omega)$ et ϵ et μ ($\eta = \sqrt{\mu/\epsilon}$) sont des fonctions de l'espace supposées réelles et bornées :

$$0 < \epsilon_- \leq \epsilon(x) \leq \epsilon_+, \quad 0 < \mu_- \leq \mu(x) \leq \mu_+, \quad \text{pour tout } x \in \Omega. \tag{2.3}$$

L'équation (b) de (2.1) est une condition absorbante d'ordre 1 sur Γ , qui ferme le domaine Ω . La dérivée normale $\partial_\nu u$ est définie par le produit scalaire $(\nu \cdot \nabla u)|_\Gamma$, où le vecteur unitaire ν correspond à la normale sortante au domaine Ω . Enfin, l'équation (c) de (2.1) correspond à la condition de type conducteur parfait sur l'objet diffractant. Dans le contexte de l'électromagnétisme, les paramètres $\mu(x)$ et $\epsilon(x)$ représentent respectivement la perméabilité et la permittivité du matériau au point x du domaine Ω , et (2.1) correspond à la diffraction d'une onde Transverse Magnétique (le champ électrique est perpendiculaire au plan et $u = E_z$, où z est la normale au plan (x, y)). Ce problème peut également modéliser la propagation d'ondes acoustiques et dans ce cas, $\mu = \rho_0$ correspond à la densité du matériau et $\epsilon = \frac{1}{\rho_0 c^2}$, où c est la vitesse de propagation de l'onde dans le milieu. Nous supposons que ces fonctions μ et ϵ sont Lipschitz-continues par morceaux sur Ω . Cela signifie qu'il existe une partition disjointe finie de classe C^0 de Ω pour laquelle la restriction de chacun des paramètres μ et ϵ est Lipschitz-continue.

2.1.2 PROBLÈME DÉCOMPOSÉ

Nous limiterons notre étude au cas d'une décomposition en deux sous-domaines, mais l'ensemble de la théorie reste valide pour un nombre quelconque de sous-

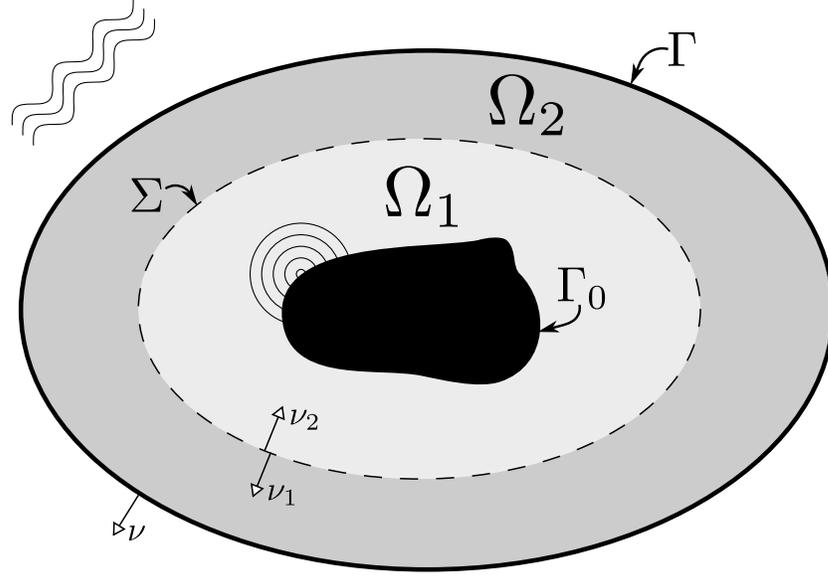


FIGURE 2.2 – Schéma représentatif du problème de diffraction étudié, avec décomposition du domaine en pelures d'oignon.

domaines, tant que les interfaces ne s'intersectent pas. L'interface séparant les deux sous-domaines est notée Σ (voir figure 2.2). Nous appellerons u la solution au problème (2.1), et nous noterons u_1 sa restriction au sous-domaine Ω_1 et u_2 celle au domaine Ω_2 (voir figure 2.2). De manière générale, les quantités relatives au sous-domaine Ω_1 , respectivement Ω_2 , seront notées avec l'indice 1, respectivement l'indice 2. Avec ces notations, les fonctions $u_1 \in H_{\Gamma_0}^1(\Omega_1)$ et $u_2 \in H^1(\Omega_2)$ sont solutions du problème décomposé suivant

$$\begin{aligned}
 -\operatorname{div}(\mu^{-1}\nabla u_1) - k^2\epsilon u_1 &= f_1 \quad \text{dans } \Omega_1 \\
 u_1 &= 0 \quad \text{sur } \Gamma_0 \\
 u_1 = u_2, \quad \partial_{\nu_1} u_1 &= -\partial_{\nu_2} u_2 \quad \text{sur } \Sigma \quad (*) \\
 -\operatorname{div}(\mu^{-1}\nabla u_2) - k^2\epsilon u_2 &= f_2 \quad \text{dans } \Omega_2 \\
 \mu^{-1}\partial_{\nu_2} u_2 + ik\eta^{-1} u_2 &= 0 \quad \text{sur } \Gamma
 \end{aligned} \tag{2.4}$$

L'équation (*) du système (2.4) correspond aux conditions de transmission exactes.

2.1.3 PROCESSUS ITÉRATIF

L'objectif de la méthode de décomposition de domaine est de construire une séquence

$$(u_1^n, u_2^n) \in H_{\Gamma_0}^1(\Omega_1) \times H^1(\Omega_2) \quad (2.5)$$

de telle sorte qu'à chaque itération n , (u_1^n, u_2^n) soit calculé à partir des solutions des itérations précédentes en résolvant des problèmes découplés, et telle que $(u_1^n, u_2^n) \rightarrow (u_1, u_2)$ lorsque $n \rightarrow +\infty$. Le couple (u_1^n, u_2^n) vérifie, à chaque itération,

$$\begin{cases} -\operatorname{div}(\mu^{-1}\nabla u_1^n) - k^2\epsilon u_1^n = f_1 & \text{dans } \Omega_1 \\ u_1^n = 0 & \text{sur } \Gamma_0 \end{cases} \quad (2.6)$$

$$\begin{cases} -\operatorname{div}(\mu^{-1}\nabla u_2^n) - k^2\epsilon u_2^n = f_2 & \text{dans } \Omega_2 \\ \mu^{-1}\partial_{\nu_2} u_2^n + ik\eta^{-1}u_2^n = 0 & \text{sur } \Gamma \end{cases} \quad (2.7)$$

Ces équations doivent être complétées par des conditions de bord sur Σ : les conditions de transmission du processus itératif. Dans un formalisme général, que nous développerons dans la suite de ce chapitre, nous introduisons des conditions de transmission basées sur des conditions d'impédance

$$\begin{aligned} \mu^{-1}\partial_{\nu_1} u_1 + ik T_1 u_1 &= -\mu^{-1}\partial_{\nu_2} u_2 + ik T_1 u_2 \\ \mu^{-1}\partial_{\nu_2} u_2 + ik T_2 u_2 &= -\mu^{-1}\partial_{\nu_1} u_1 + ik T_2 u_1, \end{aligned} \quad (2.8)$$

où T_1 et T_2 sont deux opérateurs¹ agissant sur des fonctions de frontière sur Σ . Deux autres opérateurs de bord sont introduits :

$$\mathcal{B}_1 = \mu^{-1}\partial_{\nu_1} + ik T_1, \quad \mathcal{B}_2 = \mu^{-1}\partial_{\nu_2} + ik T_2. \quad (2.9)$$

\mathcal{B}_1 doit être vu comme un opérateur de trace agissant sur des fonctions de Ω_1 , tandis que \mathcal{B}_2 est un opérateur de trace pour des fonctions de Ω_2 . Les conditions (2.8) peuvent alors être réécrites de la manière suivante

$$\begin{aligned} \mu^{-1}\partial_{\nu_1} u_1 + ik T_1 u_1 &= -(\mu^{-1}\partial_{\nu_2} u_2 + ik T_2 u_2) + ik(T_1 + T_2)u_2 \\ \mu^{-1}\partial_{\nu_2} u_2 + ik T_2 u_2 &= -(\mu^{-1}\partial_{\nu_1} u_1 + ik T_1 u_1) + ik(T_1 + T_2)u_1, \end{aligned} \quad (2.10)$$

1. Dans toute la suite du document, nous avons choisi la notation allégée : si u est une fonction définie sur les ouverts Ω , Ω_1 ou Ω_2 , et T un opérateur agissant sur des fonctions définies sur le bord Σ , on écrira Tu pour $T(u|_{\Sigma})$.

ou encore

$$\begin{aligned}\mathcal{B}_1 u_1 &= -\mathcal{B}_2 u_2 + ik(T_1 + T_2) u_2 \\ \mathcal{B}_2 u_2 &= -\mathcal{B}_1 u_1 + ik(T_1 + T_2) u_1\end{aligned}\tag{2.11}$$

Ainsi, le processus itératif s'écrit dans chaque sous-domaine

$$\left\{ \begin{array}{ll} -\operatorname{div}(\mu^{-1}\nabla u_1^n) - k^2\epsilon u_1^n = f_1 & \text{dans } \Omega_1 \\ u_1^n = 0 & \text{sur } \Gamma_0 \\ \mathcal{B}_1 u_1^n = g_1^{in,n} & \text{sur } \Sigma \end{array} \right.\tag{2.12}$$

$$\left\{ \begin{array}{ll} -\operatorname{div}(\mu^{-1}\nabla u_2^n) - k^2\epsilon u_2^n = f_2 & \text{dans } \Omega_2 \\ \mu^{-1}\partial_{\nu_2} u_2^n + ik\eta^{-1} u_2^n = 0 & \text{sur } \Gamma \\ \mathcal{B}_2 u_2^n = g_2^{in,n} & \text{sur } \Sigma \end{array} \right.\tag{2.13}$$

où les seconds membres sont donnés par

$$\begin{aligned}g_1^{in,n} &= -\mathcal{B}_2 u_2^{n-1} + ik(T_1 + T_2) u_2^{n-1} = g_2^{out,n-1} \\ g_2^{in,n} &= -\mathcal{B}_1 u_1^{n-1} + ik(T_1 + T_2) u_1^{n-1} = g_1^{out,n-1}.\end{aligned}\tag{2.14}$$

De cette manière, les conditions de transmission relient bien les *traces entrantes* dans chacun des sous-domaines aux *traces sortantes* du sous-domaine voisin de l'itération précédente. De par leur linéarité, les équations (2.12) et (2.13) définissent quatre opérateurs linéaires E_j et R_j , $j = 1, 2$ tels que

$$u_1^n = R_1 f_1 + E_1 g_1^{in,n}, \quad u_2^n = R_2 f_2 + E_2 g_2^{in,n}.\tag{2.15}$$

Plus explicitement, les opérateurs E_j correspondent à la résolution des problèmes (2.12) et (2.13) avec $f_1 = f_2 = 0$, c'est à dire

$$v_1^n = E_1 g_1^{in,n} \Leftrightarrow \left\{ \begin{array}{ll} -\operatorname{div}(\mu^{-1}\nabla v_1^n) - k^2\epsilon v_1^n = 0 & \text{dans } \Omega_1 \\ v_1^n = 0 & \text{sur } \Gamma_0 \\ \mathcal{B}_1 v_1^n = g_1^{in,n} & \text{sur } \Sigma \end{array} \right.\tag{2.16}$$

$$v_2^n = E_2 g_2^{in,n} \Leftrightarrow \left\{ \begin{array}{ll} -\operatorname{div}(\mu^{-1}\nabla v_2^n) - k^2\epsilon v_2^n = 0 & \text{dans } \Omega_2 \\ \mu^{-1}\partial_{\nu_2} v_2^n + ik\eta^{-1} v_2^n = 0 & \text{sur } \Gamma \\ \mathcal{B}_2 v_2^n = g_2^{in,n} & \text{sur } \Sigma \end{array} \right.\tag{2.17}$$

Bien que $(g_1^{in,n}, g_2^{in,n})$ soient les seconds membres, et donc les données des problèmes locaux (2.12) et (2.13), les équations (2.14) définissent un autre processus itératif construit à partir de $(g_1^{in,n}, g_2^{in,n})$ considérées comme inconnues. Il s'agit de regarder les traces à l'interface Σ comme les inconnues du problème, et la résolution des problèmes locaux comme une étape du calcul permettant d'obtenir les traces à l'itération suivante :

$$\begin{aligned} g_1^{in,n} &= -g_2^{in,n-1} + ik(T_1 + T_2)(R_2 f_2 + E_2 g_2^{in,n-1}) \\ g_2^{in,n} &= -g_1^{in,n-1} + ik(T_1 + T_1)(R_1 f_1 + E_1 g_1^{in,n-1}) \end{aligned} \quad (2.18)$$

ou encore, en notant $F_j = ik(T_1 + T_2)R_j f_j$,

$$\begin{aligned} g_1^{in,n} &= (-I + ik(T_1 + T_2)E_2)g_2^{in,n-1} + F_2 \\ g_2^{in,n} &= (-I + ik(T_1 + T_1)E_1)g_1^{in,n-1} + F_1. \end{aligned} \quad (2.19)$$

Le processus itératif apparaît plus clairement encore écrit sous forme matricielle. En notant S_1 et S_2 les opérateurs

$$S_1 = -I + ik(T_1 + T_2)E_1, \quad S_2 = -I + ik(T_1 + T_2)E_2, \quad (2.20)$$

nous obtenons

$$\begin{bmatrix} g_1^{in} \\ g_2^{in} \end{bmatrix}^n = \underbrace{\begin{bmatrix} 0 & S_2 \\ S_1 & 0 \end{bmatrix}}_{S_\pi} \cdot \begin{bmatrix} g_1^{in} \\ g_2^{in} \end{bmatrix}^{n-1} + \begin{bmatrix} F_2 \\ F_1 \end{bmatrix} \quad (2.21)$$

Les opérateurs S_1 et S_2 sont des *opérateurs de scattering*, qui calculent, à partir des traces entrantes dans un domaine, les traces sortantes associées (et qui seront utilisées à l'itération suivante dans le domaine voisin) par la relation

$$\begin{aligned} g_1^{in,n} &= g_2^{out,n-1} = S_2 g_2^{in,n-1} + F_2 \\ g_2^{in,n} &= g_1^{out,n-1} = S_1 g_1^{in,n-1} + F_1 \end{aligned} \quad (2.22)$$

Ainsi, les propriétés du processus itératif (est-il bien posé ? converge-t-il ? à quelle vitesse ?) sont directement liées aux propriétés de l'opérateur S_π . C'est pourquoi dans la suite de ce chapitre, nous étudierons précisément cet opérateur.

2.1.4 ESPACES FONCTIONNELS

Nous définissons ici quelques espaces fonctionnels qui nous seront utiles dans ce chapitre. Rappelons tout d'abord la définition des espaces de Sobolev. Soit m un entier positif et \mathcal{O} un ouvert de \mathbb{R}^d . Les espaces de Hilbert $H^m(\mathcal{O})$ sont définis par

$$H^m(\mathcal{O}) = \{v \in L^2(\mathcal{O}) \text{ telles que } \forall |\alpha| \leq m, D^\alpha v \in L^2(\mathcal{O})\}, \quad (2.23)$$

où α est un multi-indice et $D^\alpha v$ désigne une dérivée partielle de v au sens des distributions. Si $\mathcal{O} = \mathbb{R}^d$, alors les espaces de Sobolev peuvent être définis de manière équivalente à l'aide de la transformée de Fourier \hat{v} de v

$$H^m(\mathbb{R}^d) = \left\{ v \in L^2(\mathbb{R}^d) \text{ telles que } \int_{\mathbb{R}^d} (1 + \xi^2)^m |\hat{v}(\xi)|^2 d\xi < +\infty \right\}. \quad (2.24)$$

Notons que cette dernière définition peut être étendue à des valeurs de m non entières, et même non positives. Pour un réel s , les espaces de Sobolev sur \mathbb{R}^d sont alors définis par

$$H^s(\mathbb{R}^d) = \left\{ v \in L^2(\mathbb{R}^d) \text{ telles que } \int_{\mathbb{R}^d} (1 + \xi^2)^s |\hat{v}(\xi)|^2 d\xi < +\infty \right\}. \quad (2.25)$$

Dans le cas où $0 < s < 1$, et $\mathcal{O} \neq \mathbb{R}^d$, les espaces de Sobolev peuvent être définis de manière équivalente par

$$H^s(\mathcal{O}) = \left\{ v \in L^2(\mathcal{O}) \text{ telles que } \int_{\mathcal{O}} \int_{\mathcal{O}} \frac{|v(x) - v(y)|^2}{|x - y|^{d+2s}} dx dy < +\infty \right\}. \quad (2.26)$$

Enfin, si $s = m + \sigma$, avec $m \in \mathbb{N}$ et $0 < \sigma < 1$, alors $H^s(\mathcal{O})$ est défini par

$$H^s(\mathcal{O}) = \{v \in H^m(\mathcal{O}) \text{ telles que } \forall |\alpha| \leq m, D^\alpha v \in H^\sigma(\mathcal{O})\}. \quad (2.27)$$

En pratique, pour nos démonstrations, nous utiliserons principalement les espaces $H_{\Gamma_0}^1(\Omega)$ (fonctions de $H^1(\Omega)$ nulles sur le bord Γ_0), $H^1(\Omega_j)$, $j = 1, 2$, $H^{\frac{1}{2}}(\Sigma)$ et $H^{-\frac{1}{2}}(\Sigma)$ (espaces des traces). Enfin, définissons les espaces de recherche des solutions u_1 et u_2 :

$$\mathcal{H}_1 = \{v \in H_{\Gamma_0}^1(\Omega_1) \text{ telles que } \operatorname{div}(\mu^{-1} \nabla v) \in L^2(\Omega_1)\}, \quad (2.28)$$

$$\mathcal{H}_2 = \{v \in H^1(\Omega_2) \text{ telles que } \operatorname{div}(\mu^{-1} \nabla v) \in L^2(\Omega_2)\}. \quad (2.29)$$

Il nous reste à définir un certain nombre de normes que nous utiliserons régulièrement dans la suite du chapitre. L'espace $H^1(\mathcal{O})$ est muni de la norme

$$\|v\|_{H^1}^2 = \int_{\mathcal{O}} |v|^2 dx + \int_{\mathcal{O}} |\nabla v|^2 dx. \quad (2.30)$$

On en déduit les normes des espaces \mathcal{H}_j :

$$\|v\|_{\mathcal{H}_j}^2 = \|v\|_{H^1(\Omega_j)}^2 + \|\operatorname{div}(\mu^{-1}\nabla v)\|_{L^2(\Omega_j)}^2. \quad (2.31)$$

Munis de ces normes, les espaces \mathcal{H}_j sont des espaces de Hilbert. On peut alors donner un sens à la quantité $\mu^{-1}\partial_\nu u$ sur le bord Γ d'un ouvert \mathcal{O} de Ω de manière variationnelle. En effet, à toute fonction φ de $H^{\frac{1}{2}}(\Gamma)$, on lui associe la fonction $v(\varphi) \in H^1(\mathcal{O})$ définie comme étant le relèvement de φ dans $H^1(\mathcal{O})$ suivant

$$\begin{cases} -\operatorname{div}(\mu^{-1}\nabla v(\varphi)) + v(\varphi) = 0 & \text{dans } \mathcal{O} \\ v(\varphi) = \varphi & \text{sur } \Gamma \end{cases} \quad (2.32)$$

Alors, on a pour toute fonction u dans $H^1(\mathcal{O})$

$$\langle \mu^{-1}\partial_\nu u, \varphi \rangle = \int_{\mathcal{O}} (-\operatorname{div}(\mu^{-1}\nabla u) v(\varphi) + \nabla u \nabla v(\varphi)) dx. \quad (2.33)$$

2.2 CLASSE D'OPÉRATEURS DE TRANSMISSION ROBUSTES

Nous avons introduit précédemment des conditions de transmission de type conditions impédantes (voir équations (2.8)). Idéalement, ces conditions doivent satisfaire les propriétés suivantes :

- (i) Elles doivent relier les traces des solutions dans chaque sous-domaine à l'itération n aux traces des solutions à l'itération $n - 1$ du sous-domaine voisin. Ainsi, les problèmes sont découplés.
- (ii) Elles doivent être consistantes, à la convergence, avec les conditions de transmission exactes (2.4 - *).
- (iii) Elles doivent garantir que les problèmes locaux sont bien posés, ainsi que la convergence de la méthode
- (iv) Elles doivent être relativement faciles à implémenter numériquement.

L'approche naïve qui consiste à utiliser les conditions exactes

$$u_1^n = u_2^{n-1}, \quad \mu^{-1} \partial_{\nu_2} u_2^n = -\mu^{-1} \partial_{\nu_1} u_1^{n-1} \quad (2.34)$$

permet de satisfaire à (i), (ii) et (iv). Cependant, elle ne garantit pas que les problèmes locaux soient bien posés. Notons que la plupart des conditions de transmission de la littérature se retrouvent dans notre formalisme. En particulier :

- Si $T_1 = T_2 = I$, on retrouve les conditions de transmission de Després (voir Section 1.2.2.1), mais la méthode souffre d'une convergence lente.
- Si $T_1 = T_2 = \alpha I - \beta \partial_{tt}$, il s'agit des conditions de transmission d'ordre 2 (voir Section 1.2.2.2). Peu de théorie générale a été écrite sur le sujet, uniquement dans le cas d'un milieu homogène, pour des géométries simples.
- Si $T_1 = T_2 = C_0 + \sum A_j \partial_{tt} (1 - B_j \partial_{tt})^{-1}$, on retrouve l'approximation de l'opérateur racine par des fractions rationnelles de type Padé (voir Section 1.2.2.3).

Aucune de ces conditions de transmission à part celles de Després ne vérifie (iii). Pour satisfaire toutes les propriétés demandées, nous considérons une classe de conditions de transmission. Introduisons un opérateur $T = T_r + iT_i$ où les opérateurs T_r et T_i sont des opérateurs linéaires de $H^{\frac{1}{2}}(\Sigma)$ dans $H^{-\frac{1}{2}}(\Sigma)$, $H^{-\frac{1}{2}}(\Sigma)$ étant l'espace dual de $H^{\frac{1}{2}}(\Sigma)$.

Hypothèse 1 (Opérateurs de transmission).

Supposons que les propriétés suivantes sont vérifiées par T_r et T_i

- (a) T_r est positif et symétrique
- (b) T_i est symétrique et continu
- (c) T_r peut s'écrire comme $T_r = \Lambda^* \Lambda$ où $\Lambda \in \mathcal{L}(H^{\frac{1}{2}}(\Sigma), L^2(\Sigma))$ est un isomorphisme et $\Lambda^* \in \mathcal{L}(L^2(\Sigma), H^{-\frac{1}{2}}(\Sigma))$ est l'adjoint de Λ .

L'opérateur adjoint de T est alors $T^* = T_r - iT_i$, et nous utiliserons ces deux opérateurs comme opérateurs de transmission dans (2.8) ($T_1 = T$ et $T_2 = T^*$).

Proposition 1 (Équivalence des conditions de transmission).

Sous l'hypothèse 1, les conditions de transmission

$$\begin{aligned} \mu^{-1} \partial_{\nu_1} u_1 + ik T u_1 &= -\mu^{-1} \partial_{\nu_2} u_2 + ik T u_2 \\ \mu^{-1} \partial_{\nu_2} u_2 + ik T^* u_2 &= -\mu^{-1} \partial_{\nu_1} u_1 + ik T^* u_1 \end{aligned} \quad (2.35)$$

sont équivalentes aux conditions de transmission exactes

$$u_1 = u_2, \quad \mu^{-1} \partial_{\nu_1} u_1 = -\mu^{-1} \partial_{\nu_2} u_2 \quad \text{sur } \Sigma. \quad (2.36)$$

Démonstration. En soustrayant terme à terme les deux équations de (2.35), nous obtenons

$$(T + T^*) u_1 = (T + T^*) u_2 \iff T_r (u_1 - u_2) = 0. \quad (2.37)$$

Comme T_r est supposé injectif par hypothèse (c), on en déduit que $u_1 = u_2$ sur Σ . De même, en additionnant terme à terme les deux équations de (2.35), et grâce à la continuité de T_i , il est facile d'obtenir

$$\mu^{-1} \partial_{\nu_1} u_1 = -\mu^{-1} \partial_{\nu_2} u_2 \quad (2.38)$$

■

Nous noterons que seules l'injectivité de $T + T^*$ et la continuité de $T - T^*$ ont été utilisées pour montrer l'équivalence des conditions de transmission. En utilisant les opérateurs de bords \mathcal{B}_1 et \mathcal{B}_2 définis par $\mathcal{B}_1 = \mu^{-1} \partial_{\nu_1} + ik T$ et $\mathcal{B}_2 = \mu^{-1} \partial_{\nu_2} + ik T^*$, les conditions (2.35) peuvent être réécrites

$$\mathcal{B}_1 u_1 = -\mathcal{B}_2 u_2 + 2ik T_r u_2, \quad \mathcal{B}_2 u_2 = -\mathcal{B}_1 u_1 + 2ik T_r u_1. \quad (2.39)$$

Par définition, les membres de gauche des équations sont respectivement les *traces entrantes* de u_1 et u_2 sur Σ , tandis que les membres de droite sont les *traces sortantes* de u_2 et u_1 . En d'autres termes, (2.35) exprime le fait que les traces entrantes dans un sous-domaine sont les traces sortantes du sous-domaine voisin. Les conditions de transmission du processus itératif sont obtenues en écrivant des équations consistantes à la limite $n \rightarrow +\infty$ avec (2.39). La manière la plus simple pour cela est d'écrire

$$\mathcal{B}_1 u_1^n = -\mathcal{B}_2 u_2^{n-1} + 2ik T_r u_2^{n-1}, \quad \mathcal{B}_2 u_2^n = -\mathcal{B}_1 u_1^{n-1} + 2ik T_r u_1^{n-1}, \quad (2.40)$$

ce qui revient à imposer que, dans chaque sous-domaine, les traces entrantes de la solution à l'itération n correspondent aux traces sortantes de la solution à l'itération $n - 1$ dans le sous-domaine voisin. Nous utiliserons en fait une version un peu plus sophistiquée de cet algorithme, en remplaçant les traces sortantes de la solution à l'itération $n - 1$ par une combinaison linéaire convexe entre cette quantité, et la trace entrante à l'itération $n - 1$. Plus précisément, en introduisant un paramètre de relaxation $r \in]0, 1]$, nous écrivons les conditions de transmission relaxées

$$\begin{aligned}\mathcal{B}_1 u_1^n &= (1 - r) \mathcal{B}_1 u_1^{n-1} + r(-\mathcal{B}_2 u_2^{n-1} + 2ik T_r u_2^{n-1}) \\ \mathcal{B}_2 u_2^n &= (1 - r) \mathcal{B}_2 u_2^{n-1} + r(-\mathcal{B}_1 u_1^{n-1} + 2ik T_r u_1^{n-1})\end{aligned}\tag{2.41}$$

Ces conditions de transmission relaxées sont alors utilisées comme conditions de bord pour compléter les problèmes (2.6) et (2.7). Dans la section suivante, nous analyserons cet algorithme itératif plus en détail, et exhiberons un certain nombre de résultats théoriques.

2.3 ÉTUDE THÉORIQUE

2.3.1 CARACTÈRE BIEN POSÉ DES PROBLÈMES

L'objectif est de montrer que l'algorithme itératif (2.6) et (2.7), complété par les conditions de transmission relaxées (2.41) définit une unique séquence

$$(u_1^n, u_2^n) \in \mathcal{H}_1 \times \mathcal{H}_2,\tag{2.42}$$

où les espaces \mathcal{H}_1 et \mathcal{H}_2 sont les espaces de Hilbert définis Section 2.1.4. Ce résultat d'unicité découle des deux résultats suivants

- (i) Étant données $(v_1, v_2) \in \mathcal{H}_1 \times \mathcal{H}_2$, leurs traces entrantes $\mathcal{B}_1 v_1$ et $\mathcal{B}_2 v_2$ et sortantes sur Σ existent et sont des éléments de $H^{-\frac{1}{2}}(\Sigma)$. Ceci résulte directement du théorème de trace usuel dans $H(\text{div}; \Omega_j)$.
- (ii) Étant données $(g_1, g_2) \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$, les problèmes suivants sont bien posés dans $H_{\Gamma_0}^1(\Omega_1)$ pour (\mathcal{P}_1) et dans $H^1(\Omega_2)$ pour (\mathcal{P}_2) .

$$(\mathcal{P}_1) \left\{ \begin{array}{ll} -\text{div}(\mu^{-1} \nabla v_1) - k^2 \epsilon v_1 = f, & \text{dans } \Omega_1 \\ v_1 = 0, & \text{sur } \Gamma_0 \\ \mu^{-1} \partial_{\nu_1} v_1 + ik T v_1 = g_1 & \text{sur } \Sigma \end{array} \right. ,\tag{2.43}$$

$$(\mathcal{P}_2) \begin{cases} -\operatorname{div}(\mu^{-1}\nabla v_2) - k^2\epsilon v_2 = f, & \text{dans } \Omega_2 \\ \mu^{-1}\partial_\nu v_2 + ik\eta^{-1}v_2 = 0, & \text{sur } \Gamma \\ \mu^{-1}\partial_{\nu_2}v_2 + ikT^*v_2 = g_2 & \text{sur } \Sigma \end{cases} \quad (2.44)$$

Le point (ii) est l'objet du lemme suivant.

Lemme 1 (Caractère bien posé des problèmes locaux).

Si $f \in L^2(\Omega)$ et $(g_1, g_2) \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$, alors le problème (\mathcal{P}_1) (respectivement (\mathcal{P}_2)) est bien posé dans \mathcal{H}_1 (2.28) (respectivement dans \mathcal{H}_2 (2.29)).

Démonstration. Il est suffisant de prouver que (\mathcal{P}_1) et (\mathcal{P}_2) sont bien posés dans $H_{\Gamma_0}^1(\Omega_1)$ et dans $H^1(\Omega_2)$, puisque $(v_1, v_2) \in \mathcal{H}_1 \times \mathcal{H}_2$ vient des premières équations de (\mathcal{P}_1) et (\mathcal{P}_2) .

Nous n'effectuerons la démonstration que pour le problème (\mathcal{P}_2) , puisque la preuve pour (\mathcal{P}_1) est pratiquement identique. Tout d'abord, il est clair que le problème (\mathcal{P}_2) est équivalent au problème variationnel suivant

$$\text{Trouver } v \in H^1(\Omega_2) \text{ telle que } \forall w \in H^1(\Omega_2), a(v, w) = l(w) \quad (2.45)$$

où la forme bilinéaire a est donnée par

$$a(v, w) = \int_{\Omega_2} (\mu^{-1}\nabla v \nabla \bar{w} - k^2\epsilon v \bar{w}) dx + ik \int_{\Gamma} \eta^{-1} v \bar{w} d\sigma + ik \langle T^* v, w \rangle_{\Sigma} \quad (2.46)$$

et la forme linéaire l

$$l(w) = \int_{\Omega_2} f \bar{w} dx + \langle g, w \rangle_{\Sigma} \quad (2.47)$$

Dans les équations précédentes, $\langle \cdot, \cdot \rangle_{\Sigma}$ est le produit de dualité entre $H^{-\frac{1}{2}}(\Sigma)$ and $H^{\frac{1}{2}}(\Sigma)$

$$\langle v, w \rangle_{\Sigma} = \int_{\Sigma} v \bar{w} d\sigma. \quad (2.48)$$

De par les propriétés de T et d'après le théorème de trace usuel, il est clair que la forme bilinéaire $a(\cdot, \cdot)$ et la forme linéaire $l(\cdot)$ sont continues dans $H^1(\Omega_2)$.

Pour prouver l'existence et l'unicité de la solution, nous utilisons l'alternative de Fredholm. Dans ce but, nous décomposons la forme bilinéaire a de la manière suivante

$$\begin{aligned} a(v, w) &= a_*(v, w) + a_c(v, w) \\ a_*(v, w) &= \int_{\Omega_2} (\mu^{-1} \nabla v \nabla \bar{w} + \epsilon v \bar{w}) dx + ik \langle T^* v, w \rangle_{\Sigma} \\ a_c(v, w) &= - \int_{\Omega_2} (1 + k^2) \epsilon v \bar{w} dx + ik \int_{\Gamma} \eta^{-1} v \bar{w} d\sigma \end{aligned} \quad (2.49)$$

D'après le théorème de Riesz appliqué aux formes bilinéaires continues sur un espace de Hilbert, il existe trois opérateurs linéaires de $H^1(\Omega_2)$, notés A , A_* et A_c tels que

$$a(v, w) = \langle A v, w \rangle_{H^1}, \quad a_*(v, w) = \langle A_* v, w \rangle_{H^1}, \quad a_c(v, w) = \langle A_c v, w \rangle_{H^1} \quad (2.50)$$

Montrer que le problème (\mathcal{P}_2) est bien posé revient à montrer que l'opérateur $A = A_* + A_c$ est un isomorphisme dans $H^1(\Omega_2)$. La preuve de l'isomorphisme est réalisée en montrant successivement que A_* est un isomorphisme, que l'opérateur A_c est compact, et que A est injectif.

(i) L'opérateur A_* est un isomorphisme. Montrons tout d'abord que la forme bilinéaire a_* est coercive. Notons $a_{*,r}(v, v) := \operatorname{Re} a_*(v, v)$ et $a_{*,i}(v, v) := \operatorname{Im} a_*(v, v)$, et cherchons $(\lambda, \gamma) \in \mathbb{R} \times \mathbb{R}_*^+$ tels que

$$a_{*,r}(v, v) + \lambda a_{*,i}(v, v) \geq \gamma \|v\|_{H^1}^2, \quad \forall v \in H^1(\Omega_2) \quad (2.51)$$

Nous observons que

$$a_{*,r}(v, v) + \lambda a_{*,i}(v, v) = \int_{\Omega} (\mu^{-1} |\nabla v|^2 + \epsilon |v|^2) dx + k \langle T_i v, v \rangle_{\Sigma} + k \lambda \langle T_r v, v \rangle_{\Sigma} \quad (2.52)$$

Puisque $T_r = \Lambda^* \Lambda$, il existe un réel positif α_{Σ} tel que

$$\forall \varphi \in H^{\frac{1}{2}}(\Sigma), \quad \langle T_r \varphi, \varphi \rangle_{\Sigma} = \|\Lambda \varphi\|_{L^2}^2 \geq \alpha_{\Sigma} \|\varphi\|_{H^{\frac{1}{2}}(\Sigma)}^2. \quad (2.53)$$

De plus, en utilisant la continuité de T_i , qui s'écrit

$$|\langle T_i u, u \rangle_{\Sigma}| \leq M_i \|u\|_{H^{\frac{1}{2}}(\Sigma)}^2, \quad (2.54)$$

on obtient

$$a_{*,r}(v, v) + \lambda a_{*,i}(v, v) \geq \int_{\Omega} (\mu^{-1} |\nabla v|^2 + \epsilon |v|^2) dx + k(\lambda \alpha_{\Sigma} - M_i) \|v\|_{H^{\frac{1}{2}}(\Sigma)}^2. \quad (2.55)$$

Prenons $\lambda > \alpha_{\Sigma}^{-1} M_i$, et il vient

$$\gamma \|v\|_{H^1}^2 \leq a_{*,r}(v, v) + \lambda a_{*,i}(v, v) \leq \sqrt{1 + \lambda^2} |a_*(v, v)| \quad (2.56)$$

La coercivité de a_* est alors évidente. En utilisant le théorème de Lax-Milgram, nous déduisons que l'opérateur A_* est un isomorphisme.

(ii) L'opérateur A_c est compact. Ce résultat découle directement des propriétés de compacité de la forme bilinéaire a_c

$$(v_n, \tilde{v}_n) \rightharpoonup (v, \tilde{v}) \text{ faiblement dans } H^1(\Omega)^2 \implies a_c(v_n, \tilde{v}_n) \rightarrow a_c(v, \tilde{v}) \quad (2.57)$$

qui résulte de la compacité des inclusions $H^1(\Omega) \subset L^2(\Omega)$ et $H^{\frac{1}{2}}(\Sigma) \subset L^2(\Sigma)$.

(iii) L'opérateur A est injectif. Si v est dans le noyau de l'opérateur A , il vérifie en particulier

$$\begin{cases} -\operatorname{div}(\mu^{-1} \nabla v) - k^2 \epsilon v = 0, & \text{dans } \Omega_2 & (a) \\ \mu^{-1} \partial_{\nu_2} v + ikT^* v = 0 & \text{sur } \Sigma & (b) \end{cases} \quad (2.58)$$

et $a(v, v) = 0$ conduit à

$$\operatorname{Im} a(v, v) := k \langle T_r v, v \rangle_{\Sigma} + k \|\eta^{-\frac{1}{2}} v\|_{L^2(\Gamma)}^2 = k \|\Lambda v\|_{L^2(\Sigma)}^2 + k \|\eta^{-\frac{1}{2}} v\|_{L^2(\Gamma)}^2 = 0 \quad (2.59)$$

ceci implique que $v = 0$ sur Σ . En utilisant (2.58)(b), nous déduisons également que $\mu^{-1} \partial_{\nu} v = 0$ sur Σ . Grâce à (2.58)(a), nous pouvons en conclure que $v = 0$ dans Ω_2 , en utilisant un argument classique de continuation unique (voir [2] par exemple) et la connectivité de Ω_2 . Ceci conclut la preuve du lemme 1. ■

Comme nous l'avions fait remarqué rapidement dans la partie 2.1.3, le lemme 1 définit implicitement quatre opérateurs linéaires continus

$$R_j \in \mathcal{L}(L^2(\Omega_j), \mathcal{H}_j), \quad E_j \in \mathcal{L}(H^{\frac{1}{2}}(\Sigma), \mathcal{H}_j), \quad j = 1, 2 \quad (2.60)$$

tels que la solution v_j de (\mathcal{P}_j) satisfait

$$v_j = R_j f + E_j g_j. \quad (2.61)$$

De plus, il est évident, d'après le lemme 1 toujours, de déduire le corollaire suivant.

Corollaire 1 (Caractère bien posé du processus itératif).

Étant donné une solution initiale $(u_1^0, u_2^0) \in \mathcal{H}_1 \times \mathcal{H}_2$, le processus itératif décrit par les équations (2.6) et (2.7), complété par les conditions de transmission relaxées (2.41), définit une unique séquence

$$(u_1^n, u_2^n) \in \mathcal{H}_1 \times \mathcal{H}_2, \quad \forall n > 0 \quad (2.62)$$

2.3.2 REFORMULATION À L'INTERFACE POUR L'ERREUR

Maintenant que nous savons que le processus itératif est bien défini, il nous reste à prouver que le processus converge vers la bonne solution d'une part, et d'obtenir un estimé de la vitesse de convergence d'autre part. Pour cela nous allons étudier le comportement de l'erreur au cours des itérations. L'erreur à l'itération n est définie par

$$e^n = (e_1^n, e_2^n) \in \mathcal{H}_1 \times \mathcal{H}_2, \quad e_j^n = u_j^n - u_j, \quad j = 1, 2 \quad (2.63)$$

Ces erreurs satisfont de manière évidente les équations

$$\begin{cases} -\operatorname{div}(\mu^{-1}\nabla e_1^n) - k^2\epsilon e_1^n = 0 & \text{dans } \Omega_1 \\ e_1^n = 0 & \text{sur } \Gamma_0 \end{cases} \quad (2.64)$$

$$\begin{cases} -\operatorname{div}(\mu^{-1}\nabla e_2^n) - k^2\epsilon e_2^n = 0 & \text{dans } \Omega_2 \\ \mu^{-1}\partial_\nu e_2^n + ik\eta^{-1}e_2^n = 0 & \text{sur } \Gamma \end{cases} \quad (2.65)$$

À ces erreurs, que l'on appellera *volumiques* dans le sens où elles sont définies sur l'ensemble des sous-domaines, nous pouvons associer leur trace (ou *erreur à l'interface*) définie sur Σ par

$$\varepsilon^n = (\varepsilon_1^n, \varepsilon_2^n) = (\mathcal{B}_1 e_1^n, \mathcal{B}_2 e_2^n) \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma). \quad (2.66)$$

En se référant aux définitions des opérateurs E_j (2.61), on voit que les différentes erreurs sont reliées par

$$e_1^n = E_1 \varepsilon_1^n, \quad e_2^n = E_2 \varepsilon_2^n. \quad (2.67)$$

Cette relation montre que prouver la convergence de l'erreur volumique $e^n \rightarrow 0$ dans $\mathcal{H}_1 \times \mathcal{H}_2$ revient à montrer la convergence de l'erreur à l'interface $\varepsilon^n \rightarrow 0$ dans $H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$. Pour cette raison, dans la suite du chapitre, et comme nous l'avons entraperçu précédemment, nous réinterpréterons l'algorithme itératif décrit par (2.6), (2.7) et (2.41), comme un processus itératif abstrait pour ε^n défini à l'interface. En effet, les conditions de transmission (2.41) s'écrivent, pour (e_1^n, e_2^n)

$$\begin{cases} \mathcal{B}_1 e_1^n = r(-\mathcal{B}_2 e_2^{n-1} + 2ik T_r e_2^{n-1}) + (1-r) \mathcal{B}_1 e_1^{n-1} \\ \mathcal{B}_2 e_2^n = r(-\mathcal{B}_1 e_1^{n-1} + 2ik T_r e_1^{n-1}) + (1-r) \mathcal{B}_2 e_2^{n-1} \end{cases} \quad (2.68)$$

Nous introduisons alors les opérateurs S et Π , linéaires dans $H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$, et définis par

$$S(g_1, g_2) = (S_1 g_1, S_2 g_2) := (-g_1 + 2ik T_r E_1 g_1, -g_2 + 2ik T_r E_2 g_2) \quad (2.69)$$

$$\Pi(x_1, x_2) = (x_2, x_1). \quad (2.70)$$

On remarque tout d'abord que S est un opérateur diagonal, il découple les composantes des éléments de $H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$, tandis que l'opérateur d'échange Π les recouple. L'action de S_j est obtenu par la résolution du problème local dans Ω_j et transforme une trace entrante en trace sortante. Pour cette raison, nous l'appellerons *opérateur de diffraction* ou *scattering*.

Proposition 2 (Processus itératif pour l'erreur sur les traces).

Si S est l'opérateur de diffraction défini par (2.69) et Π l'opérateur d'échange de traces à l'interface, en notant S_π l'opérateur ΠS , l'erreur sur les traces à l'interface vérifie la relation

$$\varepsilon^n = (1-r) \varepsilon^{n-1} + r S_\pi \varepsilon^{n-1} \quad (2.71)$$

Démonstration. En utilisant le fait que $\mathcal{B}_j e_j^n = \varepsilon_j^n$, il s'agit d'une simple réécriture du système (2.68). ■

L'opérateur S_π est donc de manière évidente d'un intérêt crucial pour l'étude de la convergence, et la section suivante sera consacrée à cette étude.

2.3.3 REFORMULATION À L'INTERFACE POUR LA SOLUTION

Il est possible de formaliser le processus itératif pour la solution de la même manière que pour l'erreur, notamment en reformulant le problème à l'interface. Si on appelle $g^{in,n} = (g_1^{in,n}, g_2^{in,n}) \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$ les traces entrantes des solutions $(u_1^n, u_2^n) \in \mathcal{H}_1 \times \mathcal{H}_2$, alors elles vérifient les relations suivantes

$$\begin{aligned} g_1^{in,n} &= (1-r) g_1^{in,n-1} + r (-g_2^{in,n-1} + 2ik T_r u_2^{n-1}) \\ g_2^{in,n} &= (1-r) g_2^{in,n-1} + r (-g_1^{in,n-1} + 2ik T_r u_1^{n-1}). \end{aligned} \quad (2.72)$$

avec de plus

$$u_1^{n-1} = R_1 f_1 + E_1 g_1^{in,n-1}, \quad u_2^{n-1} = R_2 f_2 + E_2 g_2^{in,n-1}. \quad (2.73)$$

Soit F_1 et F_2 les traces des relèvements : $F_j = 2ik T_r R_j f_j$, $j = 1, 2$. Alors, en injectant (2.73) dans (2.72), on obtient

$$\begin{aligned} g_1^{in,n} &= (1-r) g_1^{in,n-1} + r (-g_2^{in,n-1} + 2ik T_r E_2 g_2^{in,n-1}) + r F_2 \\ g_2^{in,n} &= (1-r) g_2^{in,n-1} + r (-g_1^{in,n-1} + 2ik T_r E_1 g_1^{in,n-1}) + r F_1, \end{aligned} \quad (2.74)$$

et en utilisant l'opérateur S_π défini précédemment, on a directement la proposition suivante

Proposition 3 (Processus itératif pour les traces).

Le processus itératif pour les traces $g^{in,n} = (g_1^{in,n}, g_2^{in,n}) \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$ est

$$g^{in,n} = (1-r) g^{in,n-1} + r S_\pi g^{in,n-1} + r F, \quad (2.75)$$

avec $F = (2ik T_r R_2 f_2, 2ik T_r R_1 f_1)$. La suite de solutions locales $(u_1^n, u_2^n) \in \mathcal{H}_1 \times \mathcal{H}_2$ vérifie alors

$$u_1^n = R_1 f_1 + E_1 g_1^{in,n}, \quad u_2^n = R_2 f_2 + E_2 g_2^{in,n}. \quad (2.76)$$

Démonstration. Il s'agit d'une conséquence directe du système (2.74) est des définitions de S_π et des opérateurs R_j et E_j , $j = 1, 2$ (2.61). ■

Le processus (2.75) correspond à un algorithme (relaxé) de point fixe pour la résolution du problème

$$g^{in} = (1 - r) g^{in} + r S_\pi g^{in} + r F \Leftrightarrow (I - S_\pi) g^{in} = F. \quad (2.77)$$

La méthode de décomposition de domaine développée ici consiste en fait à calculer g^{in} par la série

$$g^{in} = (I - S_\pi)^{-1} F = \sum_{n=0}^{+\infty} S_\pi^n F. \quad (2.78)$$

C'est ce qu'on appelle l'algorithme de Jacobi. La convergence de l'algorithme est donc formellement reliée à celle de la série (2.78).

2.4 RÉSULTATS DE CONVERGENCE

L'analyse de la convergence de (2.71) repose sur deux propriétés fondamentales de l'opérateur S_π , toutes deux liées aux propriétés de l'opérateur de transmission utilisé. La première propriété de S_π est l'objet du lemme suivant

Lemme 2 (Isomorphisme).

L'opérateur $I - S_\pi$ est un isomorphisme dans $H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$.

Démonstration. Étant donné un quelconque $y = (y_1, y_2) \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$, nous souhaitons exhiber $x = (x_1, x_2) \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$ tel que $(I - S_\pi)x = y$. Alors, par définition de l'opérateur S_π

$$\begin{cases} x_1 - (-x_2 + 2ik T_r E_2 x_2) = y_1 \\ x_2 - (-x_1 + 2ik T_r E_1 x_1) = y_2 \end{cases}. \quad (2.79)$$

De plus, en notant $v_1 = E_1 x_1$ et $v_2 = E_2 x_2$ (voir définitions (2.16) et (2.17)), alors $x_1 = \mathcal{B}_1 v_1 = \mu^{-1} \partial_{\nu_1} v_1 + ik T v_1$ et $x_2 = \mathcal{B}_2 v_2 = \mu^{-1} \partial_{\nu_2} v_2 + ik T^* v_2$.

Introduisons maintenant $\sigma_n = \mu^{-1} \partial_{\nu_1} v_1 + \mu^{-1} \partial_{\nu_2} v_2$ le saut des $\mu^{-1} \partial_{\nu_j} v_j$ à l'interface Σ et $\sigma_d = v_1 - v_2$ le saut des v_j à l'interface Σ . Alors $(I - S_\pi)x = y$ est équivalent à chercher v_1 et v_2 tels que

$$\begin{cases} \sigma_n + ik T \sigma_d = y_1 \\ \sigma_n - ik T^* \sigma_d = y_2 \end{cases} \iff \begin{cases} \sigma_d = \frac{1}{ik} T_r^{-1} \left(\frac{y_1 - y_2}{2} \right) \\ \sigma_n = \frac{y_1 + y_2}{2} - i T_i T_r^{-1} \left(\frac{y_1 - y_2}{2} \right) \end{cases} \quad (2.80)$$

Pour cela, nous allons introduire un problème de saut particulier. Remarquons tout d'abord que, puisque $(y_1, y_2) \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$, et grâce aux propriétés de T_r et T_i ,

$$\sigma_d = \frac{1}{ik} T_r^{-1} \left(\frac{y_1 - y_2}{2} \right) \in H^{\frac{1}{2}}(\Sigma), \quad \sigma_n = \frac{y_1 + y_2}{2} - iT_i T_r^{-1} \left(\frac{y_1 - y_2}{2} \right) \in H^{-\frac{1}{2}}(\Sigma) \quad (2.81)$$

Ceci nous permet de considérer le problème de saut suivant : Trouver $v = (v_1, v_2) \in \mathcal{H}_1 \times \mathcal{H}_2$ telles que

$$\left\{ \begin{array}{ll} -\operatorname{div}(\mu^{-1} \nabla v_j) - k^2 \epsilon v_j = 0, & \text{dans } \Omega_j, \quad j = 1, 2 \\ v_1 = 0, & \text{sur } \Gamma_0 \\ \mu^{-1} \partial_\nu v_2 + ik \eta^{-1} v_2 = 0, & \text{sur } \Gamma \\ [v] = \sigma_d & \text{sur } \Sigma \\ [\mu^{-1} \partial_\nu v] = \sigma_n & \text{sur } \Sigma \end{array} \right. \quad (2.82)$$

où par définition $[v] = v_{1|\Sigma} - v_{2|\Sigma}$ and $[\mu^{-1} \partial_\nu v] = \mu^{-1} \partial_{\nu_1} v_{1|\Sigma} + \mu^{-1} \partial_{\nu_2} v_{2|\Sigma}$. Pour construire les solutions v_1 et v_2 solutions du problème de saut (2.82), nous allons procéder en deux étapes. Tout d'abord nous allons construire un relèvement du saut σ_d dans l'un des domaines, disons Ω_2 , puis construire la solution voulue comme une somme de ce relèvement et d'une fonction plus régulière ayant pour saut σ_n . Appelons $L \in \mathcal{L} \left(H^{\frac{1}{2}}(\Sigma), H^1(\Omega_2) \right)$ cet opérateur de relèvement, défini tel que $L(\sigma_d)|_\Sigma = \sigma_d$. Plus précisément, $u_{\sigma_d} = L(\sigma_d)$ est solution du problème aux limites elliptique

$$\left\{ \begin{array}{ll} -\operatorname{div}(\mu^{-1} \nabla u_{\sigma_d}) - k^2 \epsilon u_{\sigma_d} = 0, & \text{dans } \Omega_2 \\ \mu^{-1} \partial_\nu u_{\sigma_d} + ik \eta^{-1} u_{\sigma_d} = 0, & \text{sur } \Gamma \\ u_{\sigma_d} = \sigma_d, & \text{sur } \Sigma \end{array} \right. \quad (2.83)$$

Enfin, considérons l'espace affine $\mathcal{H}(\sigma_d)$ défini par

$$\mathcal{H}(\sigma_d) = H_{\Gamma_0}^1(\Omega) + \widetilde{L(\sigma_d)} \subset \mathcal{H}_1 \times \mathcal{H}_2, \quad (2.84)$$

où le mapping $v \in L^2(\Omega) \rightarrow \tilde{v} \in L^2(\Omega)$ est l'extension de v par 0 dans Ω_1 . Par construction, toute fonction de $\mathcal{H}(\sigma_d)$ vérifie la condition de saut $[v] = \sigma_d$. Nous pouvons alors reformuler le problème de saut initial (2.82) variationnellement de

la manière suivante : $v \in \mathcal{H}(\sigma_d)$ est solution du problème de saut (2.82) si et seulement si $w = v - \widetilde{L(\sigma_d)} \in H^1(\Omega)$ vérifie pour tout $w_t \in H^1(\Omega)$

$$\int_{\Omega} \mu^{-1} \nabla w \nabla \overline{w_t} - k^2 \int_{\Omega} \epsilon w \overline{w_t} - ik \int_{\Gamma} \eta^{-1} w \overline{w_t} = \langle \sigma_n, w_t \rangle_{\Sigma} - \mathcal{L}_{\sigma_d}(w_t) = \langle f, w_t \rangle \quad (2.85)$$

avec

$$\mathcal{L}_{\sigma_d}(w_t) = \int_{\Omega_2} \mu^{-1} \nabla L(\sigma_d) \nabla \overline{w_t} - k^2 \int_{\Omega} \epsilon L(\sigma_d) \overline{w_t} - ik \int_{\Gamma} \eta^{-1} L(\sigma_d) \overline{w_t} \quad (2.86)$$

Ces équations s'obtiennent en utilisant la décomposition du domaine Ω en Ω_1 et Ω_2 avant l'intégration par parties, pour faire apparaître les termes de bord sur l'interface Σ . Le membre de droite de l'équation (2.85) définit une forme linéaire continue sur $H^1(\Omega)$ (et donc dans $H^{-1}(\Omega)$). Le membre de gauche de (2.85) est la forme bilinéaire associée au problème global initial. Puisque ce problème admet une unique solution dans $H^1(\Omega)$ pour tout $f \in H^{-1}(\Omega)$, l'existence et l'unicité de w , et donc de v est prouvée. Ceci prouve la surjectivité de l'opérateur $I - S_{\pi}$. Il est intéressant de noter que si T_r^{-1} ne va pas dans $H^{\frac{1}{2}}(\Sigma)$, le relèvement $L(\sigma_d)$ n'est pas possible pour tout σ_d . Par exemple, si $T_r = I$, alors σ_d est dans $H^{-\frac{1}{2}}(\Sigma)$ et la preuve échoue.

Pour montrer l'injectivité, prenons $g = (g_1, g_2) \in H^{\frac{1}{2}}(\Sigma) \times H^{\frac{1}{2}}(\Sigma)$ tel que $(I - S_{\pi})g = z$, où $z = (z_1, z_2) \in H^{\frac{1}{2}}(\Sigma) \times H^{\frac{1}{2}}(\Sigma)$. Notons $e = (e_1, e_2)$ définis par $e_j = E_j g_j$, où les opérateurs E_j sont ceux de (2.16)-(2.17). Un simple calcul utilisant les définitions des E_j montre

$$\begin{aligned} z_1 &= \mu^{-1} \partial_{\nu_1} e_1 + \mu^{-1} \partial_{\nu_2} e_2 + ik T(e_1 - e_2) \\ z_2 &= \mu^{-1} \partial_{\nu_1} e_1 + \mu^{-1} \partial_{\nu_2} e_2 - ik T^*(e_1 - e_2) \end{aligned} \quad (2.87)$$

Alors, si $z = 0$, par l'injectivité de $T + T^*$, on obtient que $e_1 - e_2 = 0$ sur Σ et $\mu^{-1} \partial_{\nu_1} e_1 + \mu^{-1} \partial_{\nu_2} e_2 = 0$ sur Σ . e est alors solution du problème de Helmholtz initial sur tout Ω , sans terme source. Donc $e = 0$ et en particulier, $g = 0$. Ceci achève la démonstration. ■

Pour étudier la seconde propriété de S_{π} mentionnée auparavant, nous équipons l'espace $H^{-\frac{1}{2}}(\Sigma)$ de la norme

$$\|\varphi\|_{-\frac{1}{2}}^2 := \langle \varphi, T_r^{-1} \varphi \rangle_{\Sigma} = \left\| (\Lambda^*)^{-1} \varphi \right\|_{L^2(\Sigma)}^2 \quad (2.88)$$

et de manière similaire $H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$ avec la norme de Hilbert

$$\|x\|_X^2 = \|x_1\|_{-\frac{1}{2}}^2 + \|x_2\|_{-\frac{1}{2}}^2 \quad (2.89)$$

Nous avons alors le résultat suivant

Lemme 3 (Contractivité de S_π).

L'opérateur S_π est contractant dans $H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$ pour la norme $\|\cdot\|_X$.

Démonstration. Comme Π est unitaire (c'est un opérateur d'échange), le lemme 3 est une conséquence de

$$\forall \varphi \in H^{-\frac{1}{2}}(\Sigma), \quad \|S_1 \varphi\|_{-\frac{1}{2}} = \|\varphi\|_{-\frac{1}{2}}, \quad \|S_2 \varphi\|_{-\frac{1}{2}} \leq \|\varphi\|_{-\frac{1}{2}} \quad (2.90)$$

Nous prouverons le résultat pour S_2 , la démonstration pour S_1 étant pratiquement la même.

La fonction $v = E_2 \varphi \in \mathcal{H}_2$ vérifie

$$\begin{cases} -\operatorname{div}(\mu^{-1} \nabla v) - k^2 \epsilon v = 0, & \text{dans } \Omega_2 \\ \mu^{-1} \partial_\nu v + ik \eta^{-1} v = 0, & \text{sur } \Gamma \\ \mu^{-1} \partial_{\nu_2} v + ik T^* v = \varphi & \text{sur } \Sigma \end{cases} \quad (2.91)$$

et $S_2 \varphi = -\varphi + 2ik T_r v$; de par l'égalité $\|-\varphi + 2ik T_r v\|^2 = \|\varphi\|^2 + 4k^2 \|T_r v\|^2 - 4k \operatorname{Im}(\varphi, T_r v)$, nous avons

$$\|S_2 \varphi\|_{-\frac{1}{2}}^2 - \|\varphi\|_{-\frac{1}{2}}^2 = 4k^2 (T_r v, T_r v)_{-\frac{1}{2}} - 4k \operatorname{Im}(\varphi, T_r v)_{-\frac{1}{2}}. \quad (2.92)$$

Par définition du produit scalaire $(\cdot, \cdot)_{-\frac{1}{2}}$ (2.88), l'équation précédente devient

$$\|S_2 \varphi\|_{-\frac{1}{2}}^2 - \|\varphi\|_{-\frac{1}{2}}^2 = 4k^2 \langle T_r v, v \rangle_\Sigma - 4k \operatorname{Im} \langle \varphi, v \rangle_\Sigma. \quad (2.93)$$

Nous savons de plus que v satisfait l'équation

$$-k^2 \int_{\Omega_2} \epsilon |v|^2 + \int_{\Omega_2} \mu^{-1} |\nabla v|^2 + ik \int_{\Gamma} \eta^{-1} |v|^2 + ik \langle T v, v \rangle_\Sigma = \langle \varphi, v \rangle_\Sigma, \quad (2.94)$$

et en utilisant le fait que $T = T_r + iT_i$, avec T_i symétrique et donc $\langle T_i v, v \rangle_\Sigma$ réel, nous obtenons en prenant la partie imaginaire de l'équation précédente

$$\mathcal{I}m\langle \varphi, v \rangle_\Sigma = k \int_\Gamma \eta^{-1} |v|^2 d\sigma + k \langle T_r v, v \rangle_\Sigma. \quad (2.95)$$

Finalement,

$$\|S_2 \varphi\|_{-\frac{1}{2}}^2 - \|\varphi\|_{-\frac{1}{2}}^2 = -4k^2 \int_\Gamma \eta^{-1} |v|^2 d\sigma \leq 0. \quad (2.96)$$

La preuve pour S_1 suit le même principe de démonstration. La seule différence est qu'il n'y a pas de contribution de l'interface Γ_0 à cause de la condition de Dirichlet : l'inégalité devient une égalité et S_1 est unitaire. ■

Grâce à ces deux lemmes 2 et 3, nous pouvons établir notre principal résultat théorique concernant la convergence.

Théorème 1 (Convergence exponentielle).

Pour tout $r \in]0, 1[$, il existe une constante $C > 0$ et un réel $\tau_r \in]0, 1[$ tel que

$$\|u_1^n - u_1\|_{\mathcal{H}_1} + \|u_2^n - u_2\|_{\mathcal{H}_2} < C \tau_r^n \quad (2.97)$$

En conséquence, l'algorithme de Jacobi relaxé converge exponentiellement vite.

Démonstration. Pour tout $(x, y) \in V^2$ où V est un espace de Hilbert quelconque muni de la norme $\|\cdot\|_V$, nous avons

$$\|(1-r)x + ry\|_V^2 = (1-r)^2 \|x\|_V^2 + r^2 \|y\|_V^2 + 2r(1-r) \mathcal{R}e(x, y)_V \quad (2.98)$$

De plus

$$\|x - y\|_V^2 = \|x\|_V^2 + \|y\|_V^2 - 2\mathcal{R}e(x, y)_V \quad (2.99)$$

Il est alors facile d'obtenir, en utilisant ces deux précédentes relations, l'égalité suivante

$$\|(1-r)x + ry\|_V^2 = (1-r) \|x\|_V^2 + r \|y\|_V^2 - r(1-r) \|x - y\|_V^2. \quad (2.100)$$

En appliquant cette identité avec $x = \varepsilon^{n-1}$ défini en (2.67) et $y = S_\pi \varepsilon^{n-1}$, nous déduisons du processus itératif (2.71)

$$\|\varepsilon^n\|^2 = (1-r) \|\varepsilon^{n-1}\|^2 + r \|S_\pi \varepsilon^{n-1}\|^2 - r(1-r) \|(I - S_\pi)\varepsilon^{n-1}\|^2 \quad (2.101)$$

En utilisant le lemme de contraction 3, on obtient

$$\|\varepsilon^n\|^2 \leq \|\varepsilon^{n-1}\|^2 - r(1-r) \|(I - S_\pi)\varepsilon^{n-1}\|^2 \quad (2.102)$$

alors que d'après le lemme d'isomorphisme 2, il existe une constante $\delta > 0$ telle que

$$\forall g \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma), \quad \|(I - S_\pi)g\| \geq \delta \|g\| \quad (2.103)$$

avec, de plus, $\delta \leq 2$ car S_π est une contraction. Ceci conduit à l'inégalité

$$\|\varepsilon^n\|^2 \leq \|\varepsilon^{n-1}\|^2 - r(1-r)\delta^2 \|\varepsilon^{n-1}\|^2 \quad (2.104)$$

Il est alors facile de conclure, avec $\tau_r = \sqrt{1 - r(1-r)\delta^2}$ et $C = \max\{\|E_1\|, \|E_2\|\}$. ■

2.5 EXTENSION À D'AUTRES MÉTHODES ITÉRATIVES

D'autres méthodes itératives peuvent être utilisées pour résoudre le problème. Une première variante de l'algorithme de Jacobi est l'algorithme de Gauss-Seidel. Dans cette variante, les conditions de transmission du processus itératif (2.41) sont réécrites ainsi

$$\begin{aligned} \mathcal{B}_1 u_1^n &= (1-r) \mathcal{B}_1 u_1^{n-1} + r(-\mathcal{B}_2 u_2^{n-1} + 2ik T_r u_2^{n-1}) \\ \mathcal{B}_2 u_2^n &= -\mathcal{B}_1 u_1^n + 2ik T_r u_1^n. \end{aligned} \quad (2.105)$$

L'unique différence réside dans l'utilisation, pour la deuxième condition de transmission, de la solution u_1^n qui vient d'être calculée dans la même itération sans relaxation, à la place de la solution u_1^{n-1} à l'itération précédente que l'on relaxe. Dans cette configuration, le processus itératif (2.71) pour l'erreur sur les traces ne s'écrit plus

$$\text{Jacobi : } \varepsilon^n = \begin{bmatrix} (1-r)I & rS_2 \\ rS_1 & (1-r)I \end{bmatrix} \varepsilon^{n-1} \quad (2.106)$$

mais

$$\text{Gauss-Seidel : } \varepsilon^n = \begin{bmatrix} (1-r)I & rS_2 \\ (1-r)S_1 & rS_1S_2 \end{bmatrix} \varepsilon^{n-1}. \quad (2.107)$$

Il est tout à fait possible de développer une théorie similaire pour la méthode de Gauss-Seidel (voir pour cela l'article [9]) et d'obtenir les mêmes théorèmes de

convergence. En particulier, sous les mêmes hypothèses que pour la méthode de Jacobi, la convergence est également exponentielle. Pour les démonstrations, nous référerons le lecteur à l'article cité précédemment [9], car nous utilisons ici le même formalisme, et les démonstrations suivent les mêmes lignes.

Généralement, la convergence attendue avec l'algorithme de Gauss-Seidel est meilleure que celle pour Jacobi (on mettra en évidence cette propriété dans le chapitre suivant). Le revers de cette variante est sa difficulté de parallélisation. En effet, l'ordre de résolution des problèmes locaux importe, et le sous-domaine 1 doit être résolu avant de pouvoir résoudre le sous-domaine 2. Les problèmes locaux ne sont plus indépendants.

Enfin, les méthodes itératives de type Krylov peuvent aussi être utilisées pour résoudre le problème de décomposition de domaine. Pour un aperçu détaillé des méthodes de Krylov, nous renvoyons le lecteur au livre de Yousef Saad [34]. Au lieu d'écrire le problème sous la forme d'un problème de point fixe (2.77)

$$g^{in,n} = (1 - r) g^{in,n-1} + r S_\pi g^{in,n-1} + r F, \quad (2.108)$$

on peut considérer l'équation vérifiée par la solution g^{in} à la convergence

$$(I - S_\pi) g^{in} = F \quad (2.109)$$

et appliquer une méthode de Krylov (par exemple le GMRES) pour obtenir la solution. Les méthodes de Krylov sont connues pour avoir une convergence sensible à la position des valeurs propres de l'opérateur $I - S_\pi$ dans le plan complexe, mais estimer la vitesse de convergence pour ces méthodes est particulièrement ardu, surtout pour des problèmes non hermitiens tels que les problèmes de Helmholtz.

3

Analyses modales

Sommaire

3.1	Problème circulaire	46
3.1.1	Solutions du problème de diffraction	46
3.1.2	Propriétés des coefficients de diffraction	49
3.1.3	Étude du taux de convergence	52
3.2	Cas plus complexes, optimisation	61
3.2.1	Géométrie plane	62
3.2.1.1	Condition de transmission dans le matériau .	64
3.2.1.2	Condition de transmission à l'interface du matériau	67
3.2.1.3	Condition de transmission dans le vide . . .	70
3.2.2	Géométrie circulaire	71
3.2.3	Optimisation	75
3.2.3.1	Remarques générales	75
3.2.3.2	Fonctions objectives et résultats	76
3.2.3.3	Comparaison des géométries plane et circulaire	81
3.2.3.4	Montée en fréquence	85

Après avoir effectué l'analyse théorique et générale de la méthode de décomposition de domaine, nous allons étudier son comportement dans quelques cas simples. Nous nous intéresserons à des géométries pour lesquelles il existe une décomposition modale analytique. Ce chapitre comporte deux sections. Dans la première, le cas étudié est vraiment basique mais permettra d'explicitier toute la théorie vue dans le chapitre précédent et d'avoir une meilleure compréhension des phénomènes en jeu pour les grands modes. Il s'agit de la diffraction par un disque conducteur parfait, le domaine étant fermé par une condition absorbante. Aucun matériau n'est présent. Le problème étant à géométrie circulaire, une décomposition en fonction de Bessel de la solution est possible.

Dans la deuxième section, nous considérerons le cas où le conducteur parfait est recouvert d'une couche de matériau homogène. La géométrie sera soit circulaire (avec décomposition en fonction de Bessel) soit plane (et infinie, avec décomposition en exponentielles). Ce cas est plus d'intérêt pratique, et nous verrons comment nous pouvons l'utiliser pour optimiser la vitesse de convergence en pratique.

3.1 PROBLÈME CIRCULAIRE

Dans cette section, nous abordons le cas d'une géométrie circulaire dans le vide ($\epsilon = \mu = 1$). L'opérateur T est supposé diagonalisable dans la base $(e^{im\theta})_m$, c'est à dire

$$T e^{im\theta} = t_m e^{im\theta}, \quad T^* e^{im\theta} = \bar{t}_m e^{im\theta}, \quad (3.1)$$

avec $t_{-m} = t_m$, et m dénote le mode de l'onde considérée. Nous nous intéressons à l'évolution de l'erreur dans le processus itératif de Jacobi relaxé, donc aucune source n'est appliquée.

3.1.1 SOLUTIONS DU PROBLÈME DE DIFFRACTION

Le problème modèle, représenté figure 3.1, correspond au processus itératif décrit par les équations (2.12), (2.13) et (2.14) avec $T_1 = T$ et $T_2 = T^*$ dans (2.9). Il représente la diffraction par un objet circulaire de rayon R_1 , considéré comme un conducteur parfait. Le domaine est fermé par une condition aux limites absorbante en $r = R_2$. L'interface de décomposition de domaine Σ est placée en $r = R_\Sigma$. Dans toute la suite de la section, les figures ont été tracées avec les paramètres suivants : $R_1 = 1$, $R_2 = 2$, $R_\Sigma = 1.5$ et $k = 2\pi$. Si d'autres paramètres ont été nécessaires, ils seront précisés dans la légende.

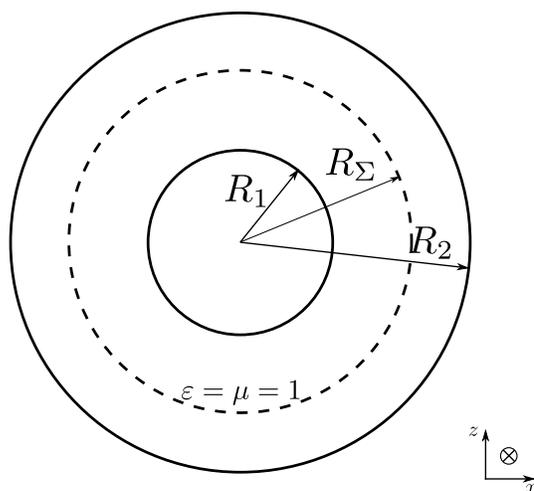


FIGURE 3.1 – Géométrie du problème circulaire simple. $\Omega_1 : R_1 < r < R_\Sigma$ et $\Omega_2 : R_\Sigma < r < R_2$

À chaque itération de l'algorithme de Jacobi, les solutions dans chaque sous-domaine, ainsi que les traces entrantes et sortantes sont décomposées modalement de la manière suivante

$$\begin{cases} u_j^n(r, \theta) = \sum_{m=-\infty}^{+\infty} u_{j,m}^n(r) e^{im\theta}, & j = 1, 2, n = 0, 1, \dots \\ g_j^{in,n}(\theta) = \sum_{m=-\infty}^{+\infty} g_{j,m}^{in,n} e^{im\theta}, & g_j^{out,n}(\theta) = \sum_{m=-\infty}^{+\infty} g_{j,m}^{out,n} e^{im\theta}. \end{cases} \quad (3.2)$$

Les données initiales correspondent à la collection de coefficients modaux $(g_{1,m}^{in,0}, g_{2,m}^{in,0})$, les coefficients des autres quantités pouvant être calculés à partir de ces seules données *via* l'algorithme de décomposition de domaine. Dans chaque sous-domaine Ω_j , la solution modale $u_{j,m}^n(r)$ peut être écrite comme

$$u_{j,m}^n(r) = A_{j,m}^n J_m(kr) + B_{j,m}^n Y_m(kr), \quad j = 1, 2 \quad (3.3)$$

où les fonctions $J_m(\xi)$ et $Y_m(\xi)$ sont les fonctions de Bessel de première et de deuxième espèce respectivement. Les coefficients $A_{j,m}^n$ et $B_{j,m}^n$ sont alors déterminés par les conditions de bords, faisant intervenir les $g_{j,m}^{in,n}$. Une première étape est de fournir une formule explicite pour ceux-ci. Dans ce but, et pour simplifier

les notations, nous définissons des coefficients modaux spécifiques à notre choix particulier de conditions de bord en $r = R_1$ et $r = R_2$, soit

$$\begin{aligned}\alpha_{1,m} &= J_m(kR_1), & \alpha_{2,m} &= J'_m(kR_2) + iJ_m(kR_2) \\ \beta_{1,m} &= Y_m(kR_1), & \beta_{2,m} &= Y'_m(kR_2) + iY_m(kR_2).\end{aligned}\tag{3.4}$$

Maintenant, dans Ω_1 , la condition de Dirichlet en $r = R_1$ nous donne une première équation

$$\alpha_{1,m} A_{1,m}^n + \beta_{1,m} B_{1,m}^n = 0.\tag{3.5}$$

Le système est complété par la condition de transmission $\partial_r u_{1,m}^n + ik t_m u_{1,m}^n = g_{1,m}^{in,n}$ en $r = R_\Sigma$, où le symbole t_m de l'opérateur T est donné par (3.1). Ces deux équations conduisent au système matriciel suivant

$$\left(\underbrace{\begin{bmatrix} \alpha_{1,m} & \beta_{1,m} \\ J'_m(kR_\Sigma) & Y'_m(kR_\Sigma) \end{bmatrix}}_{M_1} + it_m \underbrace{\begin{bmatrix} 0 & 0 \\ J_m(kR_\Sigma) & Y_m(kR_\Sigma) \end{bmatrix}}_{M_\Sigma} \right) \begin{bmatrix} A_{1,m}^n \\ B_{1,m}^n \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{g_{1,m}^{in,n}}{k}.\tag{3.6}$$

De même, dans le sous-domaine Ω_2 , la condition absorbante $\partial_r u_{2,m}^n + ik u_{2,m}^n = 0$ en $r = R_2$ est complétée par la condition de transmission en $r = R_\Sigma$, $-\partial_r u_{2,m}^n + ik \bar{t}_m u_{2,m}^n = g_{2,m}^{in}$. Le système matriciel correspondant est

$$\left(\underbrace{\begin{bmatrix} \alpha_{2,m} & \beta_{2,m} \\ -J'_m(kR_\Sigma) & -Y'_m(kR_\Sigma) \end{bmatrix}}_{M_2} + i\bar{t}_m \underbrace{\begin{bmatrix} 0 & 0 \\ J_m(kR_\Sigma) & Y_m(kR_\Sigma) \end{bmatrix}}_{M_\Sigma} \right) \begin{bmatrix} A_{2,m}^n \\ B_{2,m}^n \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{g_{2,m}^{in,n}}{k}.\tag{3.7}$$

Les traces sortantes peuvent être obtenues à partir des vecteurs solutions $\begin{bmatrix} A_{j,m}^n & B_{j,m}^n \end{bmatrix}^T$, $j = 1, 2$ grâce aux équations suivantes

$$\begin{aligned} \frac{g_{1,m}^{out,n}}{k} &= \left(\begin{bmatrix} -J'_m(kR_\Sigma) & -Y'_m(kR_\Sigma) \end{bmatrix} + i\bar{t}_m \begin{bmatrix} J_m(kR_\Sigma) & Y_m(kR_\Sigma) \end{bmatrix} \right) \begin{bmatrix} A_{1,m}^n \\ B_{1,m}^n \end{bmatrix} \\ \frac{g_{2,m}^{out,n}}{k} &= \left(\begin{bmatrix} J'_m(kR_\Sigma) & Y'_m(kR_\Sigma) \end{bmatrix} + it_m \begin{bmatrix} J_m(kR_\Sigma) & Y_m(kR_\Sigma) \end{bmatrix} \right) \begin{bmatrix} A_{2,m}^n \\ B_{2,m}^n \end{bmatrix}. \end{aligned} \quad (3.8)$$

En combinant ces quatre dernières équations, nous obtenons finalement

$$\begin{aligned} g_{1,m}^{out,n} &= \underbrace{\begin{bmatrix} 0 & 1 \end{bmatrix} (M_2 + i\bar{t}_m M_\Sigma) (M_1 + it_m M_\Sigma)^{-1}}_{s_{1,m}} \begin{bmatrix} 0 \\ 1 \end{bmatrix} g_{1,m}^{in,n} = s_{1,m} g_{1,m}^{in,n} \\ g_{2,m}^{out,n} &= \underbrace{\begin{bmatrix} 0 & 1 \end{bmatrix} (M_1 + it_m M_\Sigma) (M_2 + i\bar{t}_m M_\Sigma)^{-1}}_{s_{2,m}} \begin{bmatrix} 0 \\ 1 \end{bmatrix} g_{2,m}^{in,n} = s_{2,m} g_{2,m}^{in,n}, \end{aligned} \quad (3.9)$$

et l'assemblage de ces matrices à l'aide du logiciel **Maple** nous donne

$$s_{1,m} = -\frac{Q_{1,m} - i\bar{t}_m}{Q_{1,m} + it_m}, \quad s_{2,m} = -\frac{Q_{2,m} - it_m}{Q_{2,m} + i\bar{t}_m} \quad (3.10)$$

où les quantités $Q_{1,m}$ et $Q_{2,m}$ représentent en fait les opérateurs Dirichlet-to-Neumann pour l'interface Σ dans chacun des sous-domaines considérés. Ils sont donnés par

$$Q_{1,m} = \frac{\alpha_{1,m} Y'_m(kR_\Sigma) - \beta_{1,m} J'_m(kR_\Sigma)}{\alpha_{1,m} Y_m(kR_\Sigma) - \beta_{1,m} J_m(kR_\Sigma)}, \quad Q_{2,m} = -\frac{\alpha_{2,m} Y'_m(kR_\Sigma) - \beta_{2,m} J'_m(kR_\Sigma)}{\alpha_{2,m} Y_m(kR_\Sigma) - \beta_{2,m} J_m(kR_\Sigma)}. \quad (3.11)$$

Comme $Q_{j,-m} = Q_{j,m}$, on a $s_{j,-m} = s_{j,m}$.

3.1.2 PROPRIÉTÉS DES COEFFICIENTS DE DIFFRACTION

Les coefficients de diffraction $s_{j,m}$, $j = 1, 2$ sont très importants pour l'étude de la convergence. En effet, ils correspondent, modalement, aux opérateurs de diffraction S_j définis par (2.69) et permettent de faire correspondre à une trace entrante

dans un sous-domaine sa trace sortante associée. Les propriétés importantes de ces coefficients sont résumées dans le lemme suivant.

Lemme 4 (Propriétés des coefficients de diffraction).

Supposons que $\mathcal{R}e(t_m) > 0$ pour tout m , alors

(i) $|s_{1,m}| = 1$

(ii) $|s_{2,m}| < 1$.

De plus, s'il existe $z \in \mathbb{C}$, $\mathcal{R}e(z) > 0$ tel que $t_m \sim \frac{|m|}{kR_\Sigma} z$, alors il existe un réel $\psi(z)$ et une constante $\delta = \delta(kR_\Sigma, kR_1, kR_2, T)$ tels que

(iii) $\lim_{|m| \rightarrow +\infty} s_{1,m} s_{2,m} = e^{i\psi(z)} \neq 1$,

(iv) $\sup_m |s_{1,m} s_{2,m} - 1| > \delta$.

Démonstration. $Q_{1,m}$ étant toujours réel, il s'en suit que $|s_{1,m}| = 1$. Ensuite, d'après la définition de $s_{2,m}$, on a

$$|s_{2,m}| < 1 \Leftrightarrow \left| \frac{Q_{2,m} - it_m}{Q_{2,m} + i\bar{t}_m} \right| < 1 \Leftrightarrow \mathcal{I}m [(t_m + \bar{t}_m) Q_{2,m}] > 0, \quad (3.12)$$

donc pour prouver (ii), il nous suffit de montrer que $\mathcal{I}m Q_{2,m} > 0$. En utilisant la définition de $Q_{2,m}$, on peut voir que cette partie imaginaire vaut

$$\mathcal{I}m Q_{2,m} = \frac{W_2 \cdot W_\Sigma}{|\alpha_2 Y_m(kR_\Sigma) - \beta_2 J_m(kR_\Sigma)|^2}. \quad (3.13)$$

où W_2 et W_Σ correspondent aux Wronskiens suivants

$$\begin{aligned} W_2 &= J_m(kR_2)Y'_m(kR_2) - J'_m(kR_2)Y_m(kR_2) \\ W_\Sigma &= J_m(kR_\Sigma)Y'_m(kR_\Sigma) - J'_m(kR_\Sigma)Y_m(kR_\Sigma) \end{aligned} \quad (3.14)$$

En utilisant la propriété des Wronskiens des fonctions de Bessel $J_m(x)Y'_m(x) - J'_m(x)Y_m(x) = \frac{2}{\pi x} > 0$, (ii) devient évident. On retrouve ici le fait que l'opérateur de diffraction S_1 de la partie théorique est unitaire, et que l'opérateur de diffraction S_2 est une contraction.

Pour (iii), et comme $s_{j,-m} = s_{j,m}$, $j = 1, 2$, nous supposons que $m > 0$ et nous utilisons les asymptotiques suivantes

$$\begin{aligned} J_m(x) &\sim \frac{1}{m!} \left(\frac{x}{2}\right)^m & J'_m(x) &\sim \frac{1}{2(m-1)!} \left(\frac{x}{2}\right)^{m-1} \\ Y_m(x) &\sim -\frac{(m-1)!}{\pi} \left(\frac{2}{x}\right)^m & Y'_m(x) &\sim -\frac{m!}{\pi} \left(\frac{2}{x}\right)^{m+1} \end{aligned} \quad (3.15)$$

pour obtenir, après quelques manipulations algébriques

$$Q_{1,m} \sim Q_{2,m} \sim \frac{m}{kR_\Sigma}. \quad (3.16)$$

En utilisant l'hypothèse de l'asymptotique de t_m , il vient

$$s_{1,m}s_{2,m} \sim \frac{\frac{m}{kR_\Sigma} - i\bar{t}_m}{\frac{m}{kR_\Sigma} + it_m} \frac{\frac{m}{kR_\Sigma} - it_m}{\frac{m}{kR_\Sigma} + i\bar{t}_m} \sim \frac{1 - iy}{1 + iy} \quad (:= e^{i\psi(z)}), \quad \text{avec } y = \frac{z + \bar{z}}{1 - |z|^2}. \quad (3.17)$$

Finalement, comme $z + \bar{z} = 2\mathcal{R}e(z) > 0$, et si $|z| \neq 1$, y est réel et différent de 0. Si $|z| = 1$, on a simplement $s_{1,m}s_{2,m} \sim -1$. Il en découle que

$$\left| \frac{1 - iy}{1 + iy} \right| = 1 \text{ et } e^{i\psi(z)} \neq 1. \quad (3.18)$$

Ceci conclut la preuve de (iii). Enfin, pour (iv), une simple identité triangulaire nous donne

$$|s_{1,m}s_{2,m} - 1| \geq |e^{i\psi(z)} - 1| - |s_{1,m}s_{2,m} - e^{i\psi(z)}| \quad (3.19)$$

et comme $s_{1,m}s_{2,m} \sim e^{i\psi(z)}$, il existe un entier M tel que pour tout $m > M$, l'inégalité $-|s_{1,m}s_{2,m} - e^{i\psi(z)}| \geq -\frac{1}{2}|e^{i\psi(z)} - 1|$ est valide ; une constante convenable δ est alors donnée par

$$\delta = \min \left\{ \frac{1}{2}|e^{i\psi} - 1|, \min_{0 \leq m \leq M} |s_{1,m}s_{2,m} - 1| \right\} > 0. \quad (3.20)$$

■

Remarque 1. À partir de l'équation (3.17), il est facile de remarquer que toute autre asymptotique pour le symbole t_m de l'opérateur T conduirait à une limite de 1 pour $s_{1,m}s_{2,m}$. Dans ce cas, la propriété (iii) du lemme 4 (et donc la propriété (iv)) ne tiendrait plus.

Remarque 2. *Il est remarquable, bien qu'attendu, que $Q_{1,m}$ et $Q_{2,m}$, les opérateurs de Dirichlet-to-Neumann sur Σ associés à chaque sous-domaine, ont un comportement asymptotique qui ne dépend plus des interfaces en R_1 et R_2 ni des conditions imposées sur ces interfaces. Les modes élevés ne voient pas les conditions de bord physiques.*

La figure 3.2 illustre quelques résultats du lemme 4. Dans le premier cas (3.2a), nous avons utilisé pour l'opérateur T un symbole s'écrivant $t_m = x \left| 1 + \beta \frac{m}{kR_\Sigma} \right|$, $x, \beta > 0$, tandis que dans le second cas (3.2b), des conditions de Després sont appliquées ($t_m = 1$). Nous avons représenté dans le plan complexe les valeurs du produit $s_{1,m}s_{2,m}$. Notons que la forme de t_m pour l'opérateur permet de s'assurer que le symbole vérifie toutes les hypothèses du lemme (avec $z = x\beta$). Nous constatons trois propriétés

- Le produit $s_{1,m}s_{2,m}$, dont les valeurs sont représentées en bleu pour des valeurs de $m \in [0, 100]$, se trouve à l'intérieur du cercle unité (même s'il tend rapidement vers le cercle unité) dans les deux cas. Ceci découle directement des (i) et (ii) du lemme 4.
- La limite du produit $s_{1,m}s_{2,m}$, représentée par le point rouge, réside effectivement sur le cercle unité, mais une différence notable existe entre les deux cas. Pour l'opérateur, la limite est différente de 1. Ici $\lim_{m \rightarrow \infty} s_{1,m}s_{2,m} \approx -0.84 - 0.54i \approx e^{-3.7i}$ (résultat (iii) du lemme 4). Au contraire, dans le cas des conditions de Després, la limite vaut 1 et la proposition (iii) ne tient plus.
- Le cercle tracé en vert dans 3.2a représente un cercle centré en 1 et de rayon $\delta \approx 0.79$, et montre que pour tout m , le produit $s_{1,m}s_{2,m}$ reste éloigné d'une distance supérieure à δ , y compris à l'infini (résultat (iv) du lemme précédent). Dans le cas des conditions de Després, et comme le produit $s_{1,m}s_{2,m}$ tend vers 1, $\delta = 0$.

3.1.3 ÉTUDE DU TAUX DE CONVERGENCE

Nous retournons désormais à l'étude du processus itératif et de sa convergence. À chaque itération n , les traces entrantes sont calculées à partir des traces en-

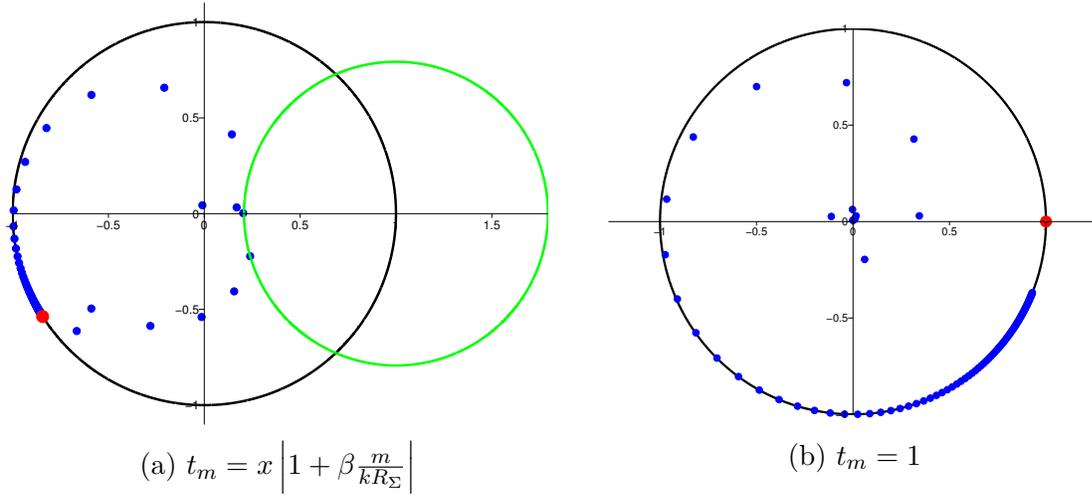


FIGURE 3.2 – Représentation dans le plan complexe du produit $s_{1,m}s_{2,m}$ pour un opérateur de symbole t_m et pour l'identité (Després : $t_m = 1$). Nous avons utilisé $x = 5e^{-2}$ et $\beta = 15$.

trantes et sortantes à l'itération précédente de la manière suivante

$$\begin{aligned} g_{1,m}^{in,n} &= (1-r) g_{1,m}^{in,n-1} + r g_{2,m}^{out,n-1} \\ g_{2,m}^{in,n} &= (1-r) g_{2,m}^{in,n-1} + r g_{1,m}^{out,n-1} \end{aligned} \quad (3.21)$$

En utilisant les calculs de la section précédente sur les coefficients de diffraction, il vient

$$\begin{aligned} \begin{bmatrix} g_{1,m}^{in} \\ g_{2,m}^{in} \end{bmatrix}^n &= \underbrace{\left((1-r) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + r \overbrace{\begin{bmatrix} \Pi & S \\ 0 & 1 \end{bmatrix}}^{\begin{bmatrix} \Pi & S \\ 0 & s_{2,m} \end{bmatrix}} \right)}_{K_m} \begin{bmatrix} g_{1,m}^{in} \\ g_{2,m}^{in} \end{bmatrix}^{n-1} = K_m \begin{bmatrix} g_{1,m}^{in} \\ g_{2,m}^{in} \end{bmatrix}^{n-1} \end{aligned} \quad (3.22)$$

L'équation précédente est l'exact équivalent modal de (2.71). S correspond à l'opérateur de diffraction et Π à l'opérateur d'échange. K_m est appelée *matrice d'itération* et nous verrons que la convergence de l'algorithme est intimement liée à ses valeurs propres.

Nous définissons la fonction racine sur l'ensemble des complexes \mathbb{C} de manière classique, l'axe de coupure étant l'axe des réels négatifs. Autrement dit, pour $z \in \mathbb{C}$,

on a $\mathcal{R}e(\sqrt{z}) \geq 0$, ou $\mathcal{I}m(\sqrt{z}) \geq 0$ si $\mathcal{R}e(\sqrt{z}) = 0$. Il est facile de voir que les valeurs propres de K_m sont

$$\tau_{r,m,\pm} = 1 - r \pm r\sqrt{s_{1,m}s_{2,m}}, \quad \text{avec } |\tau_{r,m,+}| \geq |\tau_{r,m,-}|. \quad (3.23)$$

De plus, quand $|s_{1,m}s_{2,m}| = |s_{2,m}| \neq 0$, la matrice K_m est diagonalisable. Il existe une matrice inversible P_m telle que

$$K_m = P_m \begin{bmatrix} \tau_{r,m,+} & 0 \\ 0 & \tau_{r,m,-} \end{bmatrix} P_m^{-1}, \quad \text{avec } P_m = \begin{bmatrix} \sqrt{s_{1,m}} & -\sqrt{s_{1,m}} \\ \sqrt{s_{2,m}} & \sqrt{s_{2,m}} \end{bmatrix} \quad (3.24)$$

Bien sûr, le comportement du processus itératif est commandé par le spectre de la matrice K_m . Cependant, la dérivation d'estimés simples est compliquée par le fait que $s_{2,m}$ peut s'annuler. La matrice K_m n'est alors plus diagonalisable mais s'écrit comme un bloc de Jordan. Le lemme suivant résume quelques résultats concernant ces estimés de convergence.

Lemme 5 (Résultats de convergence modale).

Pour tout r , $0 < r < 1$, soit $\tau_{r,m}$ le réel

$$\tau_{r,m} = \max_{\varepsilon=\pm 1} (|(1-r) + r\varepsilon\sqrt{s_{1,m}s_{2,m}}|) \quad (3.25)$$

où les coefficients de diffraction $s_{j,m}$ sont donnés par (3.10), (3.11) et (3.4). Alors le processus itératif (3.21) vérifie

$$\forall n \geq 0, \quad \|g_m^{in,n}\|_2 \leq (\tau_{r,m} + rn) \tau_{r,m}^{n-1} \|g_m^{in,0}\|_2. \quad (3.26)$$

De plus, quand $g_m^{in,0} = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$, nous avons

$$\|g_m^{in,n}\|_2 \geq \frac{1}{2} \min(\tau_{r,m}, r) \tau_{r,m}^{n-1} \|g_m^{in,0}\|_2 \quad (3.27)$$

Démonstration. Puisque $g_m^{in,n} = K_m g_m^{in,n-1} = K_m^n g_m^{in,0}$, nous allons étudier la matrice K_m^n . Lorsque la matrice K_m est diagonalisable (c'est-à-dire dès que $s_{2,m} \neq 0$), nous avons (pour alléger les notations, on enlève les indices m et r pour $\tau_{r,m,\pm}$)

$$K_m^n = P_m \begin{bmatrix} \tau_+^n & 0 \\ 0 & \tau_-^n \end{bmatrix} P_m^{-1} = \frac{\tau_+^n + \tau_-^n}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\tau_+^n - \tau_-^n}{2} \begin{bmatrix} 0 & \frac{s_{1,m}}{\sqrt{s_{1,m}s_{2,m}}} \\ \frac{s_{2,m}}{\sqrt{s_{1,m}s_{2,m}}} & 0 \end{bmatrix} \quad (3.28)$$

Il est alors possible de majorer la norme L^2 de K_m^n par

$$\|K_m^n\|_2 \leq \left| \frac{\tau_+^n + \tau_-^n}{2} \right| + \left| \frac{\tau_+^n - \tau_-^n}{2} \right| \left\| \begin{bmatrix} 0 & \frac{s_{1,m}}{\sqrt{s_{1,m}s_{2,m}}} \\ \frac{s_{2,m}}{\sqrt{s_{1,m}s_{2,m}}} & 0 \end{bmatrix} \right\|_2. \quad (3.29)$$

Comme $|s_{1,m}| = 1$, $|s_{2,m}| < 1$, et $|\tau_-| \leq |\tau_+|$, l'inégalité précédente devient

$$\|K_m^n\|_2 \leq |\tau_+^n| + \left| \frac{\tau_+^n - \tau_-^n}{2} \right| \left| \frac{s_{1,m}}{\sqrt{s_{1,m}s_{2,m}}} \right|. \quad (3.30)$$

En outre, comme

$$\tau_+^n - \tau_-^n = (\tau_+ - \tau_-) \sum_{k=0}^{n-1} \tau_-^k \tau_+^{n-1-k} = 2r \sqrt{s_{1,m}s_{2,m}} \sum_{k=0}^{n-1} \tau_-^k \tau_+^{n-1-k} \quad (3.31)$$

il est aisé d'obtenir la majoration suivante

$$\left| \frac{\tau_+^n - \tau_-^n}{2} \right| \left| \frac{s_{1,m}}{\sqrt{s_{1,m}s_{2,m}}} \right| = \left| r s_{1,m} \sum_{k=0}^{n-1} \tau_-^k \tau_+^{n-1-k} \right| \leq r n |\tau_+|^{n-1} \quad (3.32)$$

qui conduit directement à (3.26). Quand $s_{2,m} = 0$, la matrice d'itération K_m n'est plus diagonalisable, mais le résultat reste valide par un argument de continuité de $K_m^n(s_{2,m}) \rightarrow K_m^n(0)$.

Maintenant, si $g_m^{in,0} = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$, le calcul explicite de $g_m^{in,n} = K_m^n g_m^{in,0}$ donne, composante par composante

$$(g_m^{in,n})_1 = r s_{1,m} \tau_+^{n-1} \frac{1 - \chi^n}{1 - \chi}, \quad (g_m^{in,n})_2 = \tau_+^n \frac{1 + \chi^n}{2}, \quad (3.33)$$

avec $\chi = \frac{\tau_-}{\tau_+}$, $|\chi| \leq 1$. Comme la norme L^2 d'un vecteur est supérieure à la valeur absolue de chacune de ses composantes, on a

$$\|g_m^{in,n}\|_2 \geq \max(|(g_m^{in,n})_1|, |(g_m^{in,n})_2|) \quad (3.34)$$

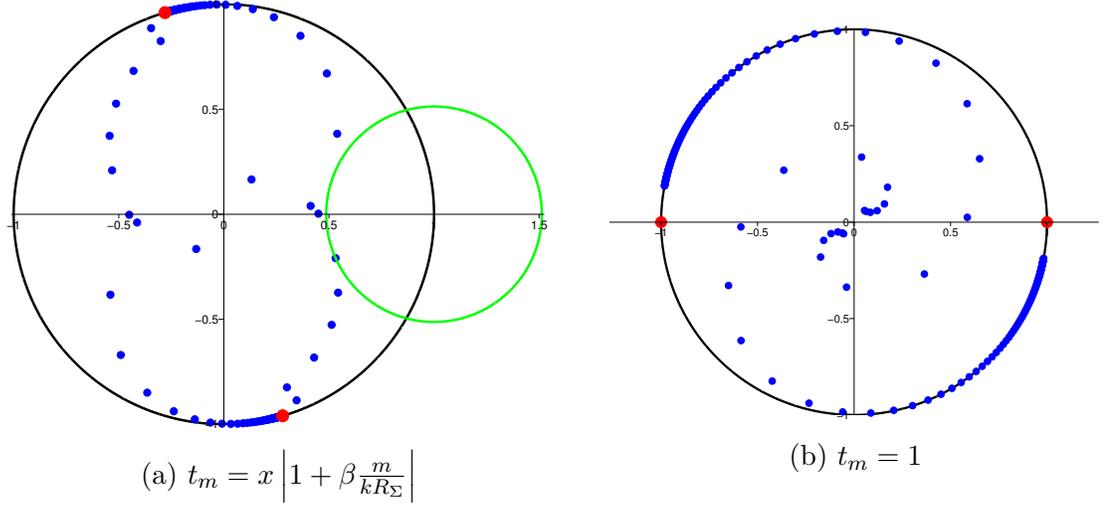


FIGURE 3.3 – Représentation dans le plan complexe de $\tau_{1,m,\pm} = \pm\sqrt{s_{1,m}s_{2,m}}$ pour un opérateur de symbole t_m et pour l'identité (Després : $t_m = 1$). Nous avons utilisé $x = 5e^{-2}$ et $\beta = 15$.

ce qui conduit, puisque $|s_{1,m}| = 1$, à

$$\|g_m^{in,n}\|_2 \geq |\tau_+|^{n-1} \min(|\tau_+|, r) \max_{|\chi| \leq 1} \left(\frac{|1 + \chi^n|}{2}, \frac{|1 - \chi^n|}{|1 - \chi|} \right). \quad (3.35)$$

Puisque $|\chi| \leq 1$, on a $\frac{1}{|1-\chi|} \geq \frac{1}{2}$ et l'estimé (3.27) est prouvé car

$$\min_{|\chi| \leq 1} \left(\max(|1 + \chi^n|, |1 - \chi^n|) \right) = 1. \quad (3.36)$$

Ceci montre que l'estimation du taux de convergence par $\tau_{r,m}$ est particulièrement bonne. ■

Pour illustrer les propriétés établies dans le lemme 5, nous reprenons l'exemple numérique de la figure 3.2. La figure 3.3 représente dans le plan complexe les valeurs de $\tau_{r,m,\pm}$ non relaxées ($r = 1$) et $\tau_{1,m,\pm} = \pm\sqrt{s_{1,m}s_{2,m}}$. Les remarques que nous avons pu faire sur le produit $s_{1,m}s_{2,m}$ s'appliquent également à $\pm\sqrt{s_{1,m}s_{2,m}}$. En particulier, dans le cas de l'opérateur non local utilisé figure (3.3a), il existe $\delta' > 0$ tel que $\sup |1 \pm \sqrt{s_{1,m}s_{2,m}}| > \delta'$.

Avant de poursuivre sur la convergence globale de l'algorithme itératif (et non plus modale), attardons nous sur l'action de la relaxation. Comme on peut le voir dans

leur expression analytique (3.23), les valeurs propres de K_m sont une combinaison linéaire convexe entre 1 et $\pm\sqrt{s_{1,m}s_{2,m}}$. Les valeurs $\pm\sqrt{s_{1,m}s_{2,m}}$ correspondent aux valeurs propres du problème non relaxé. De par les propriétés des coefficients de diffraction (voir lemme 4), ces valeurs propres non relaxées résident dans le cercle unité.

Définissons dans le plan complexe la fonction f_r par $f_r(z) = (1-r) + rz$. Comme $f_r(z) - 1 = r(z-1)$, le \mathbb{C} -mapping f_r est une homothétie de facteur $r < 1$ et de centre $z = 1$. Si \mathcal{D} dénote le disque unité ouvert centré en 0, il est clair que l'image de \mathcal{D} par f_r est un autre disque ouvert \mathcal{D}' centré en $f_r(0) = 1-r$ et de rayon r . On notera d'ailleurs que $\mathcal{D}' \subset \mathcal{D}$. Maintenant, comme $\tau_{r,m,\pm} = f_r(\pm\sqrt{s_{1,m}s_{2,m}})$, et que les valeurs propres non relaxées $\pm\sqrt{s_{1,m}s_{2,m}}$ appartiennent au disque \mathcal{D} , les valeurs propres relaxées $\tau_{r,m,\pm}$ appartiennent à \mathcal{D}' . La figure 3.4 présente schématiquement l'effet de la relaxation sur les valeurs propres (le cas représenté est pour un coefficient de relaxation de $1/2$).

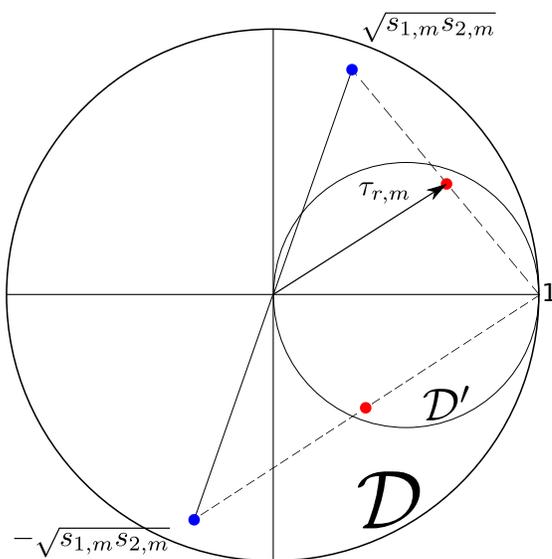


FIGURE 3.4 – Action de la relaxation sur les valeurs propres du problème non relaxé (cas où $r = \frac{1}{2}$). En bleu, les valeurs propres non relaxées, en rouge les valeurs propres relaxées.

Théorème 2 (Théorème de convergence exponentielle - bis).

Supposons que $T e^{im\theta} = t_m e^{im\theta}$ avec

$$t_{-m} = t_m, \quad \mathcal{R}e(t_m) > 0, \quad t_m \sim \frac{|m|}{kR_\Sigma} z, \quad |m| \rightarrow +\infty. \quad (3.37)$$

Alors, pour tout réel $0 < r < 1$, τ_r défini par

$$\tau_r = \sup_{m \geq 0} |\tau_{r,m,+}| = \sup_{m \geq 0} |(1-r) + r\sqrt{s_{1,m}s_{2,m}}| \quad (3.38)$$

où les coefficients $s_{j,m}$, $j = 1, 2$ sont donnés par (3.10), (3.11) et (3.4), vérifie $\tau_r < 1$, et la solution de l'algorithme itératif (3.21) satisfait

$$\|g^{in,n}\|_2 = \left(\sum_m \|g_m^{in,n}\|_2^2 \right)^{\frac{1}{2}} \leq (\tau_r + rn) \tau_r^{n-1} \left(\sum_m \|g_m^{in,0}\|_2^2 \right)^{\frac{1}{2}}. \quad (3.39)$$

En outre, soit $\tau < \tau_r$, et c une constante positive, alors il existe des données initiales $g_m^{in,0}$ et un entier $N > 0$ tels que

$$\left(\sum_m \|g_m^{in,n}\|_2^2 \right)^{\frac{1}{2}} > c \tau^n \left(\sum_m \|g_m^{in,0}\|_2^2 \right)^{\frac{1}{2}}, \quad \text{pour } n \geq N. \quad (3.40)$$

Démonstration. En appliquant l'identité remarquable (2.100) avec $x = 1$ et $y = \sqrt{s_{1,m}s_{2,m}}$, on a

$$|(1-r) + r\sqrt{s_{1,m}s_{2,m}}|^2 = (1-r) + r|\sqrt{s_{1,m}s_{2,m}}|^2 - r(1-r)|1 - \sqrt{s_{1,m}s_{2,m}}|^2. \quad (3.41)$$

Grâce au lemme 4, on sait d'une part que $|\sqrt{s_{1,m}s_{2,m}}| < 1$, donc $1 \leq \mathcal{R}e(1 + \sqrt{s_{1,m}s_{2,m}}) \leq 2$, et d'autre part, en utilisant (iv) de ce lemme on a

$$|1 - \sqrt{s_{1,m}s_{2,m}}| = \frac{|1 - s_{1,m}s_{2,m}|}{|1 + \sqrt{s_{1,m}s_{2,m}}|} \geq \frac{\delta}{\mathcal{R}e(1 + \sqrt{s_{1,m}s_{2,m}})} \geq \frac{\delta}{2},$$

et nous en déduisons, pour $0 < r < 1$

$$|(1-r) + r\sqrt{s_{1,m}s_{2,m}}|^2 \leq (1-r) + r - r(1-r)\frac{\delta^2}{4} = \delta_r^2 < 1 \quad (3.42)$$

Ceci implique

$$\sup_m \left| (1-r) + r\sqrt{s_{1,m}s_{2,m}} \right| \leq \delta_r < 1. \quad (3.43)$$

Prenons maintenant un réel positif $\tau < \tau_r$. Alors il existe un mode m_0 tel que

$$|\tau_{r,m_0}| \geq \frac{\tau + \tau_r}{2}. \quad (3.44)$$

Dans ce cas, si on prend $g_m^{in,0} = 0$ pour tout $m \neq m_0$ et $g_{m_0}^{in,0} = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$, d'après le lemme 5, on obtient

$$\|g_{m_0}^{in,n}\| \geq \frac{1}{2} \min(\tau_{r,m_0}, r) \tau_{r,m_0}^{n-1} \geq \frac{1}{2} \min\left(\frac{1}{2}(\tau + \tau_r), r\right) \left(\frac{\tau + \tau_r}{2}\right)^{n-1}. \quad (3.45)$$

et le dernier terme est minoré par $c\tau^n = c\tau^n \|g_{m_0}^{in,0}\|$ pour n assez grand. L'estimation de la convergence (3.39) est donc la meilleure possible, dans le sens où elle est atteinte pour certaines données initiales. ■

On retrouve ici la convergence exponentielle de l'algorithme itératif de Jacobi quand l'opérateur T vérifie quelques conditions modales. Il est intéressant de noter que l'équivalence $t_m \sim \frac{|m|}{kR_\Sigma} z$ signifie que T est un opérateur pseudo-différentiel d'ordre 1. En particulier, c'est un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $H^{-\frac{1}{2}}(\Sigma)$. En appliquant la relaxation à l'exemple numérique étudié dans cette section, nous obtenons les figures 3.5. L'effet homothétique de la relaxation est bien mise en évidence. De plus, dans le cas de l'opérateur (3.5a), nous observons que toutes les valeurs de $\tau_{r,m,\pm}$ se situent strictement à l'intérieur du cercle unité, ce qui conduit à la convergence exponentielle. Dans le cas de conditions de transmission locales (comme par exemple sur la figure 3.5b) la convergence exponentielle est empêchée par la limite à 1 du produit $s_{1,m}s_{2,m}$.

L'intérêt du théorème 2 est qu'il fournit une formule explicite du taux de convergence de la méthode de décomposition de domaine dans ce cas simple. Un autre résultat d'intérêt est le suivant

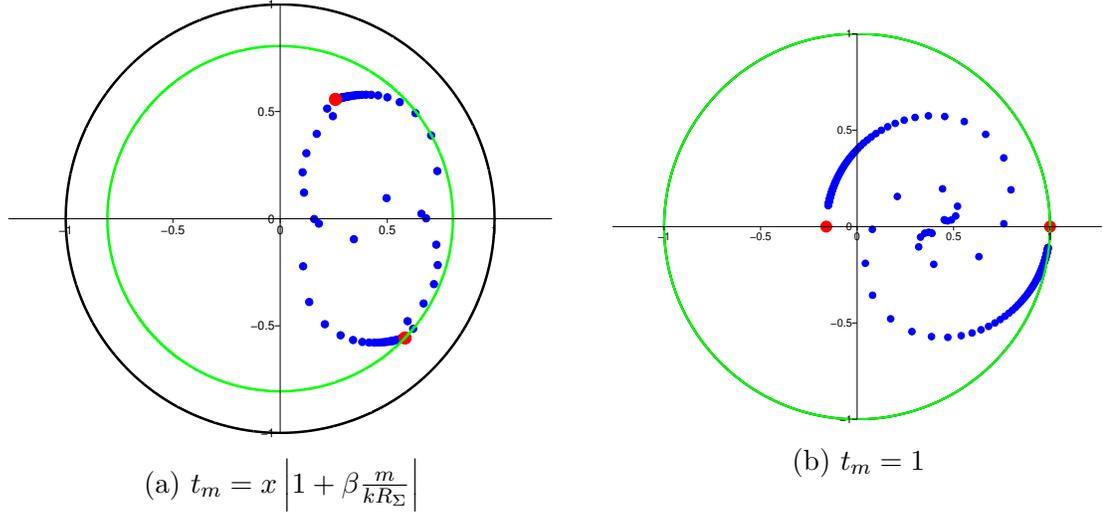


FIGURE 3.5 – Représentation dans le plan complexe de $\tau_{r,m,\pm} = 1 + r \pm r \sqrt{s_{1,m}s_{2,m}}$ pour un opérateur de symbole t_m et pour l'identité (Després : $t_m = 1$). Nous avons utilisé $x = 5e^{-2}$, $\beta = 15$ et $r = 0.58$.

Proposition 4 (Cas de non convergence exponentielle).

Supposons que $T e^{im\theta} = t_m e^{im\theta}$ avec

$$t_m = \frac{P_p(m^2)}{Q_q(m^2)} \quad (3.46)$$

avec P_p et Q_q deux polynômes quelconques de degré respectif p et q . Alors, pour tout $\tau < 1$ et $c > 0$, il existe des données initiales $g_m^{in,0}$ et un entier $N > 0$ tels que

$$\left(\sum_m \|g_m^{in,n}\|^2 \right)^{\frac{1}{2}} > c \tau^n \left(\sum_m \|g_m^{in,0}\|^2 \right)^{\frac{1}{2}}, \quad \text{pour } n \geq N \quad (3.47)$$

Démonstration. Pour montrer ce résultat, intéressons nous à l'équivalence (3.17)

$$s_{1,m}s_{2,m} \sim \frac{\frac{m}{kR_\Sigma} - i\bar{t}_m}{\frac{m}{kR_\Sigma} + it_m} \frac{\frac{m}{kR_\Sigma} - it_m}{\frac{m}{kR_\Sigma} + i\bar{t}_m}, \quad m \rightarrow +\infty. \quad (3.48)$$

Quand t_m est une fraction rationnelle de polynômes en m^2 , il existe une constante $\gamma \in \mathbb{C}$ telle que

$$t_m \sim \gamma m^{2(p-q)}, \quad m \rightarrow +\infty \quad (3.49)$$

et il est clair que, quelque soit le signe de $p - q$, ou même si $p = q$,

$$s_{1,m}s_{2,m} \sim 1, \quad m \rightarrow +\infty, \quad (3.50)$$

ce qui implique évidemment que

$$(1 - r) + r\sqrt{s_{1,m}s_{2,m}} \sim 1, \quad m \rightarrow +\infty. \quad (3.51)$$

Il suffit alors de prendre comme données initiales $g_m^{in,0} = 0$ pour $m \neq m_0$, et $g_{m_0}^{in,0} = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$ pour un m_0 suffisamment grand pour avoir

$$\left| (1 - r) + r\sqrt{s_{1,m_0}s_{2,m_0}} \right| \geq \frac{1}{2}(1 + \tau). \quad (3.52)$$

La preuve de (3.47) est immédiate pour N tel que

$$\frac{1}{2} \min \left(\frac{1}{2}(\tau + 1), r \right) \left(\frac{1 + \tau}{2} \right)^{n-1} \geq c\tau^n, \quad n \geq N \quad (3.53)$$

■

Ce résultat négatif montre qu'au moins pour cette géométrie particulière, il n'est pas possible d'obtenir une convergence exponentielle de l'algorithme de Jacobi relaxé quand l'opérateur de transmission T est une fraction rationnelle de polynômes appliquée à l'opérateur de Laplace-Beltrami sur Σ . Cette absence de convergence exponentielle est due aux modes les plus oscillants à l'interface.

3.2 CAS PLUS COMPLEXES, OPTIMISATION

Cette section est une extension de la section précédente concernant l'analyse modale de cas relativement simples. Ici, une couche de matériau est ajoutée et recouvre le conducteur parfait. Deux types de géométrie seront étudiées, une géométrie plane (infinie) et une géométrie circulaire (semblable à la géométrie de la section précédente). L'intérêt des analyses que nous présentons ici réside dans le lien qu'elles forment entre la partie très théorique que nous avons abordée dans le chapitre précédent, et les aspects numériques que nous aborderons dans une troisième partie. Compte tenu de la présence de matériaux, les calculs sont un

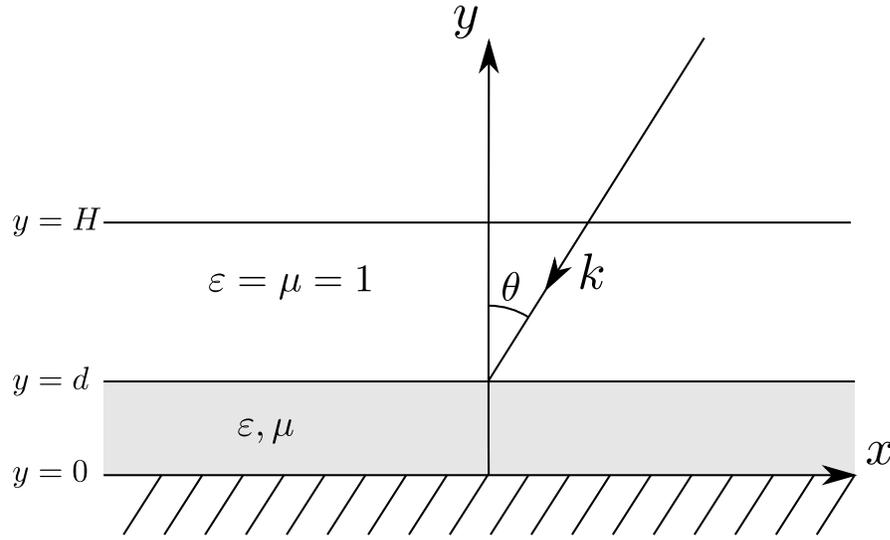


FIGURE 3.6 – Géométrie du problème plan.

peu plus techniques et surtout moins explicites. Si, dans la section précédente, nous avons suivi le chapitre 2 concernant l'étude théorique pour faire l'analyse modale, ici nous proposons une seconde méthode d'analyse modale, équivalente à la première, dans l'esprit de ce qui a pu être proposé dans la littérature, et notamment par Bruno Stupfel dans [41].

3.2.1 GÉOMÉTRIE PLANE

Nous considérons ici le problème présenté figure 3.6. Il s'agit d'un plan, placé en $y = 0$ sur lequel une condition impédante (d'impédance Z) est imposée, et recouvert d'une couche de matériau d'épaisseur d , les coefficients du matériau μ et ϵ étant supposés homogènes dans toute la couche. Une condition absorbante est placée en $y = H$. Vu comme un problème d'électromagnétisme en 2D, ici u représente la composante z du champ électrique \mathbf{E} , si z est la normale au plan d'incidence (x, y) . Du fait de l'homogénéité des coefficients et que la condition absorbante est dans le vide, le problème global devient

$$\begin{aligned}
 -\Delta u - k^2 \mu \epsilon u &= f \text{ dans } \Omega \\
 \partial_y u + ik u &= 0 \text{ sur } \Gamma \\
 -\mu^{-1} \partial_y u + \frac{ik}{Z} u &= 0 \text{ sur } \Gamma_0
 \end{aligned} \tag{3.54}$$

Notons que pour retrouver une condition de Dirichlet sur Γ_0 , il suffit de prendre $Z \rightarrow 0$ dans toute la suite. De manière similaire à la situation précédente, une décomposition du problème est réalisée grâce à la transformée de Fourier de la manière suivante

$$u(x, y) = \int_{-\infty}^{+\infty} u_{k_x}(y) e^{ik_x x} dk_x \quad (3.55)$$

On note $s = k_x/k$ le mode considéré. Dans chaque sous-domaine, la solution modale $u_s(y)$ peut s'écrire

$$u_s(y) = A_s(y) e^{ik_y(y)y} + B_s(y) e^{-ik_y(y)y}, \quad (3.56)$$

où $k_y(y)$ est donné par la relation d'onde $k_x^2 + k_y^2(y) = k^2 \mu \epsilon$, ou encore

$$k_y^2 = k^2 (\mu \epsilon - s^2). \quad (3.57)$$

La dépendance de $A_s(y)$, $B_s(y)$ et $k_y(y)$ en y est due à la dépendance de μ et ϵ en y . Dans la suite, ces dépendances seront explicités par l'ajout d'un exposant m dans le matériau et v dans le vide. On a ainsi quatre coefficients A_s^m , B_s^m , A_s^v et B_s^v . Sur l'interface Γ_0 , la condition d'impédance $-\mu^{-1} \partial_y u + \frac{ik}{Z} u = 0$ nous donne une première équation pour les coefficients A_s^m et B_s^m

$$A_s^m + \varphi_{Z,s} B_s^m = 0, \quad \varphi_{Z,s} = \frac{1 + Z \frac{k_y^m}{k\mu}}{1 - Z \frac{k_y^m}{k\mu}} = \frac{1 + \frac{Z}{\mu} \sqrt{\mu \epsilon - s^2}}{1 - \frac{Z}{\mu} \sqrt{\mu \epsilon - s^2}} \quad (3.58)$$

De même, sur Γ (située dans le vide), la condition absorbante $\partial_y u + ik u = 0$ conduit à la relation suivante pour les coefficients A_s^v et B_s^v

$$A_s^v + \phi_s B_s^v = 0, \quad \phi_s = \frac{1 - \frac{k_y^v}{k}}{1 + \frac{k_y^v}{k}} e^{-2ik_y^v H} = \frac{1 - \sqrt{1 - s^2}}{1 + \sqrt{1 - s^2}} e^{-2ik_y^v H} \quad (3.59)$$

Enfin, la continuité des quantités u et $\mu^{-1} \partial_y u$ à l'interface du matériau avec le vide donne deux dernières équations, reliant l'ensemble des inconnues.

$$A_s^m e^{ik_y^m d} + B_s^m e^{-ik_y^m d} = A_s^v e^{ik_y^v d} + B_s^v e^{-ik_y^v d} \quad (3.60)$$

$$\frac{ik_s^m}{\mu} (A_s^m e^{ik_y^m d} - B_s^m e^{-ik_y^m d}) = \frac{ik_s^v}{1} (A_s^v e^{ik_y^v d} - B_s^v e^{-ik_y^v d}) \quad (3.61)$$

Finalement, le problème global \mathcal{P} revient à résoudre le système

$$\mathcal{P} : \underbrace{\begin{bmatrix} \varphi_{Z,s} & 1 & 0 & 0 \\ 1 & \alpha & -\beta & -\gamma \\ 1 & \alpha' & -\beta' & -\gamma' \\ 0 & 0 & \phi_s & 1 \end{bmatrix}}_P \cdot \begin{bmatrix} B_s^m \\ A_s^m \\ B_s^v \\ A_s^v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (3.62)$$

où les coefficients $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ sont donnés par

$$\begin{cases} \alpha = e^{2ik_y^m d} & \alpha' = -e^{2ik_y^m d} \\ \beta = e^{i(-k_y^v + k_y^m)d} & \beta' = \mu \frac{k_s^v}{k_s^m} e^{i(-k_y^v + k_y^m)d} \\ \gamma = e^{i(k_y^v + k_y^m)d} & \gamma' = -\mu \frac{k_s^v}{k_s^m} e^{i(k_y^v + k_y^m)d}. \end{cases} \quad (3.63)$$

La résolution du système \mathcal{P} résout complètement le problème global. Le problème est inversible dès que la condition

$$\det P = \varphi_{Z,s} [\phi_s(\alpha\gamma' - \alpha'\gamma) - (\alpha\beta' - \alpha'\beta)] - (\phi_s(\gamma' - \gamma) - (\beta' - \beta)) \neq 0 \quad (3.64)$$

est satisfaite. On introduit maintenant la décomposition de domaine. Trois situations seront étudiées, selon la position y_Σ de l'interface Σ :

- $0 < y_\Sigma < d$, l'interface est dans le matériau.
- $y_\Sigma = d$, l'interface de transmission est confondue avec celle du matériau.
- $d < y_\Sigma < H$ l'interface de transmission est dans le vide.

3.2.1.1 CONDITION DE TRANSMISSION DANS LE MATÉRIAU

Dans cette situation, les deux inconnues A_s^m et B_s^m ne suffisent plus pour représenter la solution dans le matériau, mais quatre inconnues sont nécessaires. On les notera $(A_{1,s}^m, B_{1,s}^m)$ et $(A_{2,s}^m, B_{2,s}^m)$. Les conditions de transmission CT (2.35) relient ces quatre inconnues. En supposant l'opérateur T diagonalisable sur la base $(e^{ik_0sx})_s : T e^{ik_0sx} = t_s e^{ik_0sx}$, on obtient

$$\text{CT} : \begin{cases} A_{1,s}^m + \psi_1 B_{1,s}^m = A_{2,s}^m + \psi_1 B_{2,s}^m \\ A_{1,s}^m + \psi_2 B_{1,s}^m = A_{2,s}^m + \psi_2 B_{2,s}^m \end{cases} \quad (3.65)$$

avec les coefficients ψ_1 et ψ_2 donnés par

$$\begin{aligned}\psi_1 &= \frac{t_s - \frac{k_y^m}{k\mu}}{t_s + \frac{k_y^m}{k\mu}} e^{-2ik_y^m y_\Sigma} = \frac{t_s - \frac{1}{\mu} \sqrt{\mu\epsilon - s^2}}{t_s + \frac{1}{\mu} \sqrt{\mu\epsilon - s^2}} e^{-2ik_y^m y_\Sigma}, \\ \psi_2 &= \frac{\bar{t}_s + \frac{k_y^m}{k\mu}}{\bar{t}_s - \frac{k_y^m}{k\mu}} e^{-2ik_y^m y_\Sigma} = \frac{\bar{t}_s + \frac{1}{\mu} \sqrt{\mu\epsilon - s^2}}{\bar{t}_s - \frac{1}{\mu} \sqrt{\mu\epsilon - s^2}} e^{-2ik_y^m y_\Sigma}.\end{aligned}\tag{3.66}$$

Finalement, résoudre le problème décomposé revient à trouver la solution du système

$$\mathcal{P}_\Sigma : \underbrace{\begin{bmatrix} \varphi_{Z,s} & 1 & | & 0 & 0 & 0 & 0 \\ \psi_1 & 1 & | & -\psi_1 & -1 & 0 & 0 \\ \psi_2 & 1 & | & -\psi_2 & -1 & 0 & 0 \\ 0 & 0 & | & 1 & \alpha & -\beta & -\gamma \\ 0 & 0 & | & 1 & \alpha' & -\beta' & -\gamma' \\ 0 & 0 & | & 0 & 0 & \phi_s & 1 \end{bmatrix}}_{P_\Sigma} \cdot \begin{bmatrix} B_{1,s}^m \\ A_{1,s}^m \\ B_{2,s}^m \\ A_{2,s}^m \\ B_{2,s}^v \\ A_{2,s}^v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\tag{3.67}$$

Dans la matrice P_Σ , les lignes pointillées délimitent les inconnues en fonctions du sous-domaine associé. Les blocs diagonaux correspondent aux problème locaux, les blocs extra-diagonaux sont les conditions de transmission. Notons M_s la matrice des blocs diagonaux, et N_s la matrice telle que $P_\Sigma = M_s - N_s$

$$M_s = \begin{bmatrix} \varphi_{Z,s} & 1 & | & 0 & 0 & 0 & 0 \\ \psi_1 & 1 & | & 0 & 0 & 0 & 0 \\ 0 & 0 & | & -\psi_2 & -1 & 0 & 0 \\ 0 & 0 & | & 1 & \alpha & -\beta & -\gamma \\ 0 & 0 & | & 1 & \alpha' & -\beta' & -\gamma' \\ 0 & 0 & | & 0 & 0 & \phi_s & 1 \end{bmatrix}, \quad N_s = \begin{bmatrix} 0 & 0 & | & 0 & 0 & 0 & 0 \\ 0 & 0 & | & \psi_1 & 1 & 0 & 0 \\ -\psi_2 & -1 & | & 0 & 0 & 0 & 0 \\ 0 & 0 & | & 0 & 0 & 0 & 0 \\ 0 & 0 & | & 0 & 0 & 0 & 0 \\ 0 & 0 & | & 0 & 0 & 0 & 0 \end{bmatrix}\tag{3.68}$$

Le calcul du déterminant de la matrice M_s donne $\det M_s = (\psi_1 - \psi_2) \det P_\Sigma$. Sous la condition $\psi_1 \neq \psi_2$, la matrice M_s des problèmes locaux est donc inversible dès

que le problème \mathcal{P}_Σ est inversible ($\det P_\Sigma \neq 0$). De plus, il est facile de montrer que la condition $\psi_1 \neq \psi_2$ est équivalente à $t_s + \bar{t}_s \neq 0$. On retrouve la condition T_r injectif nécessaire pour avoir équivalence entre les conditions de transmission impédantes et le raccord des solutions et de leurs traces normales à l'interface.

L'algorithme de Jacobi non relaxé peut être vu comme une méthode itérative de résolution du système \mathcal{P}_Σ préconditionné par la matrice M_s des problèmes locaux. En notant x le vecteur solution, et x^n le vecteur solution à l'itération n , la méthode se dérive de la manière suivante :

$$P_\Sigma x = b \quad \Leftrightarrow \quad M_s x = N_s x + b \quad \Leftrightarrow \quad x = M_s^{-1} N_s x + M_s^{-1} b \quad (3.69)$$

et un algorithme de point fixe est utilisé. Ceci conduit au processus itératif

$$x^{n+1} = M_s^{-1} N_s x^n + M_s^{-1} b \quad (3.70)$$

Encore une fois, la convergence de l'algorithme et sa vitesse sont commandées par les valeurs propres de la matrice $M_s^{-1} N_s$.

Proposition 5 (Interface Σ dans le matériau).

Les valeurs propres non nulles de la matrice d'itération $M_s^{-1} N_s$ sont données par

$$\lambda_{s,\pm} = \pm \sqrt{Q_s}, \quad Q_s = \frac{(\psi_2 - \varphi_{Z,s})(\psi_1 - \Gamma_s)}{(\psi_1 - \varphi_{Z,s})(\psi_2 - \Gamma_s)} \quad (3.71)$$

avec

$$\Gamma_s = \frac{\phi_s(\gamma - \gamma') - (\beta - \beta')}{\phi_s(\gamma\alpha' - \gamma'\alpha) - (\beta\alpha' - \beta'\alpha)}. \quad (3.72)$$

Les valeurs propres de l'algorithme de Jacobi relaxé et son taux de convergence modal sont alors

$$\lambda_{r,s,\pm} = (1 - r) + r\lambda_{s,\pm}, \quad \tau_{r,s} = \left| (1 - r) + r\sqrt{Q_s} \right|. \quad (3.73)$$

Compte tenu des nombreux paramètres du problème, il est nettement plus difficile d'étudier analytiquement les valeurs propres, comme nous avons pu le faire dans le cas simple circulaire homogène. Cependant, si le problème global \mathcal{P} est bien posé (absence d'onde guidée, voir par exemple [22]), la théorie nous indique que toutes les valeurs propres résident dans le disque unité. Compte tenu de l'expression des

valeurs propres $\lambda_{r,s,\pm}$, il est relativement aisé de montrer la proposition suivante

Proposition 6 (Propriétés des valeurs propres).

Supposons $t_s + \bar{t}_s \neq 0$. Alors, pour tout r , $0 < r < 1$, on a les équivalences suivantes

$$\lambda_{r,s,\pm} \neq 1 \Leftrightarrow Q_s \neq 1 \Leftrightarrow \det P \neq 0. \quad (3.74)$$

De plus, supposons que $t_s \sim z s^\nu$, avec $\operatorname{Re} z > 0$ et ν réel quelconque lorsque $s \rightarrow +\infty$, alors

$$\lim_{s \rightarrow +\infty} Q_s = \begin{cases} \left(\frac{z\mu + i}{z\mu - i} \right) \left(\frac{\bar{z}\mu + i}{\bar{z}\mu - i} \right) \neq 1, & \text{si } \nu = 1 \\ 1, & \text{sinon} \end{cases} \quad (3.75)$$

Démonstration. D'après l'expression des valeurs propres relaxées, on a

$$\lambda_{r,s,\pm} \neq 1 \Leftrightarrow 1 - r \pm r\sqrt{Q_s} \neq 1 \Leftrightarrow Q_s \neq 1. \quad (3.76)$$

De plus, à partir de l'expression (3.71) de Q_s , on obtient

$$Q_s \neq 1 \Leftrightarrow (\varphi_{Z,s} - \Gamma_s)(\psi_2 - \psi_1) \neq 0 \Leftrightarrow \begin{cases} \det P \neq 0 \\ \text{et } t_s + \bar{t}_s \neq 0 \end{cases}. \quad (3.77)$$

Le calcul de la limite (3.75) est délicat et laissé au logiciel **Maple**. Il est cependant facile de voir, dans le cas $\nu = 1$, que la limite est différente de 1 dès que $\operatorname{Re} z \neq 0$. ■

Encore une fois, ces résultats sont tout à fait en adéquation avec les calculs effectués précédemment, pour la géométrie circulaire simple. Notamment, l'utilisation de conditions de transmission de Després, ou de fractions rationnelles de Laplacien-Beltrami, ne peut pas conduire à une convergence géométrique puisque la limite des valeurs propres (relaxées ou non), et donc du taux de convergence, vaut 1.

3.2.1.2 CONDITION DE TRANSMISSION À L'INTERFACE DU MATÉRIAU

Si l'interface Σ est placée à l'interface entre le matériau et le vide ($y_\Sigma = d$), les équations concernant la condition d'impédance (3.58) et la condition absorbante

(3.59) sont inchangées. Lors de l'application de la décomposition de domaine, les conditions (3.60) et (3.61) de continuité à l'interface du matériau sont simplement remplacées par les conditions de transmission. Il faut seulement faire attention à la discontinuité des coefficients μ et ϵ , et donc de k_y , à l'interface. On obtient alors les deux conditions de transmission suivantes

$$\begin{aligned} A_{1,s}^m + \tilde{\alpha} B_{1,s}^m &= \tilde{\beta} A_{2,s}^v + \tilde{\gamma} B_{2,s}^v \\ A_{1,s}^m + \tilde{\alpha}' B_{1,s}^m &= \tilde{\beta}' A_{2,s}^v + \tilde{\gamma}' B_{2,s}^v, \end{aligned} \quad (3.78)$$

où les coefficients sont donnés par

$$\begin{aligned} \tilde{\alpha} &= \frac{\frac{k_y^m}{k\mu} + t_s}{-\frac{k_y^m}{k\mu} + t_s} e^{2ik_y^m y_\Sigma} & \tilde{\alpha}' &= \frac{-\frac{k_y^m}{k\mu} + \bar{t}_s}{\frac{k_y^m}{k\mu} + \bar{t}_s} e^{2ik_y^m y_\Sigma} \\ \tilde{\beta} &= \frac{-\frac{k_y^v}{k} + t_s}{-\frac{k_y^m}{k\mu} + t_s} e^{i(-k_y^v + k_y^m) y_\Sigma} & \tilde{\beta}' &= \frac{\frac{k_y^v}{k} + \bar{t}_s}{\frac{k_y^m}{k\mu} + \bar{t}_s} e^{i(-k_y^v + k_y^m) y_\Sigma} \\ \tilde{\gamma} &= \frac{\frac{k_y^v}{k} + t_s}{-\frac{k_y^m}{k\mu} + t_s} e^{i(k_y^v + k_y^m) y_\Sigma} & \tilde{\gamma}' &= \frac{-\frac{k_y^v}{k} + \bar{t}_s}{\frac{k_y^m}{k\mu} + \bar{t}_s} e^{i(k_y^v + k_y^m) y_\Sigma}. \end{aligned} \quad (3.79)$$

En supposant $t_s + \bar{t}_s > 0$, les équations exprimant les conditions de transmission (3.78) sont équivalentes aux équations de continuité (3.60) et (3.61). Le problème décomposé s'écrit alors matriciellement sous la forme

$$\underbrace{\begin{bmatrix} \varphi_{Z,s} & 1 & | & 0 & 0 \\ 1 & \tilde{\alpha} & | & -\tilde{\beta} & -\tilde{\gamma} \\ \hline 1 & \tilde{\alpha}' & | & -\tilde{\beta}' & -\tilde{\gamma}' \\ 0 & 0 & | & \phi_s & 1 \end{bmatrix}}_{P_\Sigma} \cdot \begin{bmatrix} B_{1,s}^m \\ A_{1,s}^m \\ B_{2,s}^v \\ A_{2,s}^v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (3.80)$$

Encore une fois, les lignes pointillées délimitent l'appartenance des inconnues aux sous-domaines respectifs. Les blocs diagonaux correspondent aux problèmes intérieurs, et les blocs extra-diagonaux couplent les problèmes locaux. De la même manière que précédemment, la matrice P_Σ est décomposée sous la forme $P_\Sigma = M_s - N_s$

avec M_s et N_s les matrices

$$M_s = \begin{bmatrix} \varphi_{Z,s} & 1 & 0 & 0 \\ 1 & \tilde{\alpha} & 0 & 0 \\ 0 & 0 & -\tilde{\beta}' & -\tilde{\gamma}' \\ 0 & 0 & \phi_s & 1 \end{bmatrix}, \quad N_s = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \tilde{\beta} & \tilde{\gamma} \\ -1 & -\tilde{\alpha}' & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (3.81)$$

et le processus itératif de Jacobi (non relaxé) correspond à

$$x^{n+1} = M_s^{-1} N_s x^n + M_s^{-1} b. \quad (3.82)$$

Proposition 7 (Σ à l'interface du matériau et du vide).

Les valeurs propres de la matrice d'itération $M_s^{-1} N_s$ sont données par

$$\lambda_{s,\pm} = \pm \sqrt{Q_s}, \quad Q_s = \frac{(\tilde{\alpha}' \varphi_{Z,s} - 1)(\tilde{\gamma} \phi_s - \tilde{\beta})}{(\tilde{\alpha} \varphi_{Z,s} - 1)(\tilde{\gamma}' \phi_s - \tilde{\beta}')} \quad (3.83)$$

De même, les valeurs propres relaxées de l'algorithme de Jacobi, et son taux de convergence modal, sont alors

$$\lambda_{r,s,\pm} = (1-r) + r\lambda_{s,\pm}, \quad \tau_{r,s} = \left| (1-r) + r\sqrt{Q_s} \right|. \quad (3.84)$$

Si nous supposons de plus que $t_s + \bar{t}_s > 0$, alors, pour tout r , $0 < r < 1$, on a les équivalences suivantes

$$\lambda_{r,s,\pm} \neq 1 \Leftrightarrow Q_s \neq 1 \Leftrightarrow \det P \neq 0. \quad (3.85)$$

De plus, supposons que $t_s \sim z s^\nu$, avec $\operatorname{Re} z > 0$ et ν réel quelconque lorsque $s \rightarrow +\infty$, alors

$$\lim_{s \rightarrow +\infty} Q_s = \begin{cases} \left(\frac{z+i}{z\mu-i} \right) \left(\frac{\bar{z}\mu+i}{\bar{z}-i} \right) \neq 1, & \text{si } \nu = 1 \\ 1, & \text{sinon} \end{cases} \quad (3.86)$$

Les mêmes conclusions que dans le cas d'une interface de transmission dans le matériau s'appliquent au cas d'une interface de transmission entre le matériau et le vide. En effet, le calcul des valeurs propres limites montrent encore une fois que des conditions de transmission de type Després ou de type fractions rationnelles ne permettent pas une convergence exponentielle. L'utilisation de conditions de transmission faisant intervenir un opérateur pseudo-différentiel d'ordre 1, ayant donc un symbole équivalent à $t_s \sim z s$, est alors une condition nécessaire pour obtenir une convergence exponentielle dans le cas présenté.

3.2.1.3 CONDITION DE TRANSMISSION DANS LE VIDE

Nous passerons très rapidement sur le cas d'une interface de transmission placée dans le vide, car ce cas est très similaire au cas d'une interface dans le matériau : toutes les équations sont identiques, à l'exception des conditions de transmission, où les valeurs μ , ϵ et k_y sont les valeurs dans le vide et non dans le matériau, et ce sont les inconnues A_s^v et B_s^v qui sont décomposées en $(A_{1,s}^v, B_{1,s}^v)$ et $(A_{2,s}^v, B_{2,s}^v)$. Le problème se résume donc au système

$$\underbrace{\begin{bmatrix} \varphi_{Z,s} & 1 & 0 & 0 & | & 0 & 0 \\ 1 & \alpha & -\beta & -\gamma & | & 0 & 0 \\ 1 & \alpha' & -\beta' & -\gamma' & | & 0 & 0 \\ 0 & 0 & \psi_1 & 1 & | & -\psi_1 & -1 \\ \hline 0 & 0 & \psi_2 & 1 & | & -\psi_2 & -1 \\ 0 & 0 & 0 & 0 & | & \phi_s & 1 \end{bmatrix}}_{P_\Sigma} \cdot \begin{bmatrix} B_{1,s}^m \\ A_{1,s}^m \\ B_{1,s}^v \\ A_{1,s}^v \\ B_{2,s}^v \\ A_{2,s}^v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad (3.87)$$

où les coefficients $\varphi_{Z,s}$ et ϕ_s sont donnés par (3.58) et (3.59), tandis que les autres coefficients qui interviennent sont explicités dans (3.63). La même décomposition en $P_\Sigma = M_s - N_s$ est réalisée, avec M_s la matrice diagonale par bloc de P_Σ , et le processus itératif de Jacobi non relaxé est toujours

$$x^{n+1} = M_s^{-1} N_s x^n + M_s^{-1} b. \quad (3.88)$$

Proposition 8 (Interface Σ dans le vide).

Les valeurs propres non nulles de la matrice d'itération $M_s^{-1}N_s$ sont données par la formule

$$\lambda_{s,\pm} = \pm\sqrt{Q_s}, \quad Q_s = \frac{(\psi_2 - \Gamma_s)(\psi_1 - \phi_s)}{(\psi_1 - \Gamma_s)(\psi_2 - \phi_s)} \quad (3.89)$$

où le coefficient Γ_s est correspond à

$$\Gamma_s = \frac{\varphi_{Z,s}(\beta\alpha' - \beta'\alpha) - (\beta - \beta')}{\varphi_{Z,s}(\gamma\alpha' - \gamma'\alpha) - (\gamma - \gamma')}. \quad (3.90)$$

Les valeurs propres relaxées de l'algorithme de Jacobi, et son taux de convergence modal, sont alors

$$\lambda_{r,s,\pm} = (1 - r) + r\lambda_{s,\pm}, \quad \tau_{r,s} = \left| (1 - r) + r\sqrt{Q_s} \right|. \quad (3.91)$$

On a de plus, si $t_s \sim z s^\nu$, avec $\operatorname{Re} z > 0$ et ν réel quelconque lorsque $s \rightarrow +\infty$, alors

$$\lim_{s \rightarrow +\infty} Q_s = \begin{cases} \left(\frac{z+i}{z-i} \right) \left(\frac{\bar{z}+i}{\bar{z}-i} \right) \neq 1, & \text{si } \nu = 1 \\ 1, & \text{sinon} \end{cases} \quad (3.92)$$

3.2.2 GÉOMÉTRIE CIRCULAIRE

Dans cette section, nous nous intéressons de nouveau à une géométrie circulaire. Un disque de rayon R_0 , sur lequel une condition impédante est placée, est recouvert comme dans le cas précédent d'une couche de matériau d'épaisseur d , en $R_2 = R_0 + d$. Le matériau est supposé homogène, de paramètres μ et ϵ constants. Une condition absorbante placée en $R_3 = R_0 + H$ ferme le domaine (voir figure 3.7). Nous n'étudierons ici que le cas où la condition de transmission est placée dans le matériau, en R_1 , $R_0 < R_1 < R_2$. Les autres cas sont très similaires aux situations étudiées précédemment, et le développement des calculs serait immédiat. Comme dans le cas de la géométrie circulaire simple étudiée section 3.1, les solutions

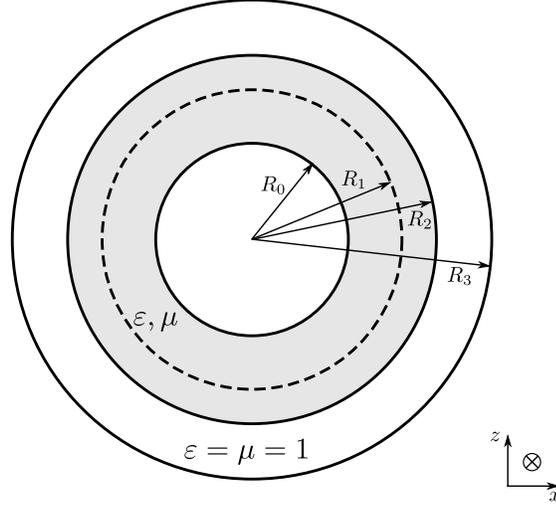


FIGURE 3.7 – Géométrie du problème circulaire avec matériau.

globale $u(r, \theta)$ et locales $u_1(r, \theta)$ et $u_2(r, \theta)$ sont décomposées sur la base $(e^{im\theta})_m$

$$u(r, \theta) = \sum_{m=-\infty}^{+\infty} u_m(r) e^{im\theta}, \quad u_j(r, \theta) = \sum_{m=-\infty}^{+\infty} u_{j,m}(r) e^{im\theta}, \quad j = 1, 2. \quad (3.93)$$

Du fait de la géométrie circulaire, dans le sous-domaine $\Omega_1 = \{(r, \theta) \text{ tel que } R_0 < r < R_1\}$, la solution locale peut s'écrire comme

$$u_{1,m}(r) = A_{1,m} J_m(k\sqrt{\mu\epsilon} r) + B_{1,m} Y_m(k\sqrt{\mu\epsilon} r). \quad (3.94)$$

Dans le sous-domaine Ω_2 , il faut différencier la partie avec matériau et la partie vide. On a

$$\begin{aligned} u_{2,m}(r) &= A_{2,m}^m J_m(k\sqrt{\mu\epsilon} r) + B_{2,m}^m Y_m(k\sqrt{\mu\epsilon} r), & R_1 < r < R_2, \\ u_{2,m}(r) &= A_{2,m}^v J_m(kr) + B_{2,m}^v Y_m(kr), & R_2 < r < R_3. \end{aligned} \quad (3.95)$$

Pour déterminer les six coefficients inconnus $A_{j,m}$ et $B_{j,m}$, nous utilisons les différentes conditions de bord et de transmission. La condition d'impédance $-\mu^{-1}\partial_r u_1 + \frac{ik}{Z} u_1 = 0$ en $r = R_0$ donne

$$A_{1,s}^m + \varphi_{Z,m} B_{1,s}^m = 0, \quad \varphi_{Z,m} = \frac{i Y_m(k\sqrt{\mu\epsilon} R_0) - Z \sqrt{\frac{\epsilon}{\mu}} Y_m'(k\sqrt{\mu\epsilon} R_0)}{i J_m(k\sqrt{\mu\epsilon} R_0) - Z \sqrt{\frac{\epsilon}{\mu}} J_m'(k\sqrt{\mu\epsilon} R_0)}, \quad (3.96)$$

tandis que la condition absorbante $\partial_r u_2 + ik u_2 = 0$ en $r = R_3$ conduit à l'équation

$$A_{2,s}^v + \phi_m B_{2,s}^v = 0, \quad \phi_m = \frac{i Y_m(kR_3) + Y_m'(kR_3)}{i J_m(kR_3) + J_m'(kR_3)}. \quad (3.97)$$

Deux autres équations sont données par les conditions de continuité à l'interface entre le matériau et le vide, en $r = R_2$. Pour cela, introduisons les coefficients suivants :

$$\begin{aligned} \alpha &= \frac{J_m(k\sqrt{\mu\epsilon}R_2)}{Y_m(k\sqrt{\mu\epsilon}R_2)} & \beta &= \frac{Y_m(kR_2)}{Y_m(k\sqrt{\mu\epsilon}R_2)} & \gamma &= \frac{J_m(kR_2)}{Y_m(k\sqrt{\mu\epsilon}R_2)} \\ \alpha' &= \frac{J_m'(k\sqrt{\mu\epsilon}R_2)}{Y_m'(k\sqrt{\mu\epsilon}R_2)} & \beta' &= \frac{Y_m'(kR_2)}{\sqrt{\frac{\epsilon}{\mu}} Y_m'(k\sqrt{\mu\epsilon}R_2)} & \gamma' &= \frac{J_m'(kR_2)}{\sqrt{\frac{\epsilon}{\mu}} Y_m'(k\sqrt{\mu\epsilon}R_2)}. \end{aligned} \quad (3.98)$$

Avec ces notations, les conditions de continuité s'écrivent

$$\begin{aligned} B_2^m + \alpha A_2^m &= \beta B_2^v + \gamma A_2^v \\ B_2^m + \alpha' A_2^m &= \beta' B_2^v + \gamma' A_2^v. \end{aligned} \quad (3.99)$$

Finalement, les conditions de transmission (2.35) nous donnent les deux dernières équations pour compléter le système. En supposant de nouveau que l'opérateur T est diagonalisable dans la base $(e^{im\theta})$, c'est à dire que $T e^{im\theta} = t_m e^{im\theta}$, nous introduisons les quantités suivantes

$$\begin{aligned} \psi_1 &= \frac{\sqrt{\frac{\epsilon}{\mu}} Y_m'(k\sqrt{\mu\epsilon}R_1) + i t_m Y_m(k\sqrt{\mu\epsilon}R_1)}{\sqrt{\frac{\epsilon}{\mu}} J_m'(k\sqrt{\mu\epsilon}R_1) + i t_m J_m(k\sqrt{\mu\epsilon}R_1)} \\ \psi_2 &= \frac{-\sqrt{\frac{\epsilon}{\mu}} Y_m'(k\sqrt{\mu\epsilon}R_1) + i \bar{t}_m Y_m(k\sqrt{\mu\epsilon}R_1)}{-\sqrt{\frac{\epsilon}{\mu}} J_m'(k\sqrt{\mu\epsilon}R_1) + i \bar{t}_m J_m(k\sqrt{\mu\epsilon}R_1)} \end{aligned} \quad (3.100)$$

alors, les conditions de transmission sont simplement

$$\begin{aligned} A_{1,m}^m + \psi_1 B_{1,m}^m &= A_{2,m}^m + \psi_1 B_{2,m}^m \\ A_{1,m}^m + \psi_2 B_{1,m}^m &= A_{2,m}^m + \psi_2 B_{2,m}^m \end{aligned} \quad (3.101)$$

et en regroupant l'ensemble de ces équations, on obtient, de manière très similaire au cas de la géométrie plane, le système suivant

$$\underbrace{\begin{bmatrix} \varphi_{Z,m} & 1 & 0 & 0 & 0 & 0 \\ \psi_1 & 1 & -\psi_1 & -1 & 0 & 0 \\ \psi_2 & 1 & -\psi_2 & -1 & 0 & 0 \\ 0 & 0 & 1 & \alpha & -\beta & -\gamma \\ 0 & 0 & 1 & \alpha' & -\beta' & -\gamma' \\ 0 & 0 & 0 & 0 & \phi_m & 1 \end{bmatrix}}_{P_\Sigma} \cdot \begin{bmatrix} B_{1,m}^m \\ A_{1,m}^m \\ B_{2,m}^m \\ A_{2,m}^m \\ B_{2,m}^v \\ A_{2,m}^v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ i^m \end{bmatrix}. \quad (3.102)$$

Étant donné que le système s'écrit formellement de la même manière que dans le cas de la géométrie plane (3.67), l'analyse de la méthode de Jacobi suit exactement le même modèle dans les deux cas. Nous regroupons les propriétés de ces valeurs propres de la méthode itérative dans la proposition suivante.

Proposition 9 (Valeurs propres - cas circulaire).

Les valeurs propres de la méthode de Jacobi relaxée ont pour expression (r étant toujours le coefficient de relaxation)

$$\lambda_{r,m,\pm} = (1-r) \pm r \sqrt{Q_m}, \quad Q_m = \frac{(\psi_2 - \varphi_{Z,m})(\psi_1 - \Gamma_m)}{(\psi_1 - \varphi_{Z,m})(\psi_2 - \Gamma_m)}, \quad (3.103)$$

avec

$$\Gamma_m = \frac{\phi_m(\gamma - \gamma') - (\beta - \beta')}{\phi_m(\gamma\alpha' - \gamma'\alpha) - (\beta\alpha' - \beta'\alpha)}. \quad (3.104)$$

Si, de plus, il existe un complexe z de partie réelle strictement positive et un réel ν tels que $t_m \sim z \left(\frac{m}{k\sqrt{\mu\epsilon} R_1} \right)^\nu$ quand $m \rightarrow +\infty$, alors

$$\lim_{m \rightarrow +\infty} Q_m = \begin{cases} \left(\frac{z\mu + i}{z\mu - i} \right) \left(\frac{\bar{z}\mu + i}{\bar{z}\mu - i} \right) \neq 1, & \text{si } \nu = 1 \\ 1, & \text{sinon} \end{cases} \quad (3.105)$$

3.2.3 OPTIMISATION

3.2.3.1 REMARQUES GÉNÉRALES

Nous présentons ici quelques remarques générales concernant l'expression des valeurs propres, leur limite, ou encore le processus itératif. Tout d'abord, on a pu voir que les valeurs propres relaxées de l'algorithme de Jacobi ont pour forme générale

$$\lambda_{r,\pm} = (1 - r) \pm r \sqrt{Q}, \quad (3.106)$$

où Q est un nombre complexe dépendant des conditions de bord physiques, de la géométrie et des conditions de transmission. Pour faire le lien avec les sections précédentes, concernant la théorie générale et l'analyse modale du cas circulaire très simple, le complexe Q correspond au produit des deux coefficients de *scattering* $s_{1,m}s_{2,m}$ (3.10)-(3.11). C'est pourquoi on retrouve toutes les propriétés de l'opérateur S_π et du symbole $s_{1,m}s_{2,m}$. En effet, l'opérateur S_π étant une contraction, on a $|s_{1,m}s_{2,m}| \leq 1$ et ici $|Q| \leq 1$. De plus, l'utilisation d'opérateurs de transmission pseudo-différentiels d'ordre 1 est une condition nécessaire pour que la limite quand $m \rightarrow \infty$ du produit $s_{1,m}s_{2,m}$ ou du complexe Q (et donc la limite des valeurs propres) soit différente de 1.

Enfin, nous avons mentionné précédemment d'autres méthodes itératives pour résoudre le problème de décomposition de domaine. Il est possible de réaliser la même analyse modale pour ces méthodes et nous donnons ici les principaux résultats. Là où les valeurs propres étaient données par $\lambda_{r,\pm} = (1 - r) \pm r \sqrt{Q}$ dans le cas de Jacobi, l'unique valeur propre non nulle pour Gauss-Seidel est donnée par

$$\lambda_{r,GS} = (1 - r) + r Q. \quad (3.107)$$

La forme de cette valeur propre suggère une meilleure convergence de la méthode de Gauss-Seidel comparée à la méthode de Jacobi. En effet, puisque le quotient Q réside dans le cercle unité, on a $|Q| \leq |\sqrt{Q}|$. L'effet de la relaxation complexifie l'analyse, mais on peut supposer néanmoins que, pour les modes les plus difficiles (proche de 1), le taux modal sera meilleur pour Gauss-Seidel. On l'a vu, ce gain de performance se fait au dépend de la parallélisation de l'algorithme. Comme la méthode de Gauss-Seidel est une méthode de type puissance (l'erreur à l'itération n s'écrit $\varepsilon^n = A^n \varepsilon^0$), cette valeur propre donne directement une excellente estimation du taux de convergence, comme la méthode de Jacobi.

Pour les méthodes de Krylov, comme GMRES par exemple, l'estimation du taux de convergence est plus compliqué. Si dans le cas d'une matrice A symétrique

définie positive, une estimation du taux de convergence τ de la méthode GMRES peut s'écrire

$$\tau = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}, \quad \kappa = \frac{\max \lambda(A)}{\min \lambda(A)} \quad (3.108)$$

où $\lambda(A)$ désigne le spectre de la matrice (A) et κ est par définition le conditionnement de la matrice A , cette estimation n'est plus valable pour des matrices générales. Bien que, dans notre cas, les problèmes considérés ne soient pas symétriques définis positifs, nous utiliserons dans la suite l'approximation suivante pour le calcul du conditionnement

$$\tilde{\kappa} = \frac{\max |\lambda(A)|}{\min |\lambda(A)|} \quad (3.109)$$

et ferons l'hypothèse que réduire le conditionnement de la matrice permet de réduire le nombre d'itérations nécessaire à la convergence de la méthode GMRES. Ici, les valeurs propres de la matrice $I - S_\pi$ qui est inversée par la méthode GMRES possède les valeurs propres

$$\lambda_{\pm, K} = 1 - \lambda_{1, \pm} = 1 \pm \sqrt{Q}. \quad (3.110)$$

3.2.3.2 FONCTIONS OBJECTIVES ET RÉSULTATS

Nous allons voir dans cette section comment utiliser les calculs analytiques des valeurs propres (et donc du taux de convergence) pour optimiser la vitesse de convergence des algorithmes. Jusqu'à maintenant, nous avons été très généraux concernant le symbole de l'opérateur T . Dans le cas des conditions de transmission de Després, l'opérateur est l'identité, donc son symbole vaut $t_s = 1$. Il est possible de généraliser les conditions de transmission de Després au cas où l'opérateur est zI avec $z \in \mathbb{C}$. Nous appellerons ces conditions des conditions de type Després généralisées. Pour rester dans le cadre théorique (unicité de la solution et convergence de l'algorithme assurées), il est nécessaire d'avoir $\mathcal{R}e z > 0$. Le symbole d'un tel opérateur est $t_s = z$. Pour obtenir une convergence exponentielle, nous avons introduit dans le chapitre de théorie générale des opérateurs de type $T = T_r + iT_i$ avec T_r et T_i vérifiant l'hypothèse 1 (page 28). Dans la suite du document, nous considérerons des opérateurs T tels que la partie T_r soit proportionnelle à la partie T_i . C'est à dire, il existe un complexe z tel que $T = zT_r = z\Lambda^*\Lambda$, avec $\mathcal{R}e z > 0$ et où Λ est un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$. Un tel opérateur T vérifie l'hypothèse 1. Avant d'étudier la réalisation pratique d'un tel opérateur Λ ,

nous utiliserons comme symbole pour celui-ci un symbole classique pour de tels opérateurs : $1 + i\sqrt{\beta|s|}$ avec $\beta \in \mathbb{R}_*^+$. L'opérateur T a alors pour symbole

$$t_s = z(1 + \beta|s|) \quad (3.111)$$

On notera que la limite $\beta \rightarrow 0$ conduit aux conditions de Després généralisées. La réalisation d'opérateurs ayant ce symbole, ou un symbole très proche, sera l'objet de la prochaine partie. Intéressons nous pour le moment à l'optimisation des conditions de transmission basées sur ce type d'opérateurs. Le processus itératif fait alors intervenir les paramètres suivants

- Paramètre $z \in \mathbb{C}$, $\Re z > 0$: ce coefficient détermine l'importance de T_i par rapport à T_r .
- Paramètre $\beta > 0$: ce coefficient détermine l'importance de l'opérateur pseudo-différentiel Λ d'ordre $\frac{1}{2}$.
- Paramètre $r \in]0, 1[$: le coefficient de relaxation de la méthode.

En dehors des conditions précitées, ces paramètres peuvent être fixés librement. Ils servent de variables d'ajustement (ou paramètres d'optimisation) pour réduire le taux de convergence. Nous avons vu que les valeurs propres de la matrice d'itération, qui gouvernent le taux de convergence, sont données sous la forme

$$\begin{aligned} \text{Jacobi : } \lambda_{s,J} &= 1 - r \pm r \sqrt{Q_s(z, \beta)} \\ \text{Gauss-Seidel : } \lambda_{s,GS} &= 1 - r + r Q_s(z, \beta) \\ \text{GMRES : } \lambda_{s,K} &= 1 \pm \sqrt{Q_s(z, \beta)} \end{aligned} \quad (3.112)$$

où $Q_s(z, \beta)$ a une expression donnée par (3.71) dans le cas de la géométrie plane avec condition de transmission dans le matériau, (3.83) si la condition de transmission est à l'interface du matériau avec le vide, et par (3.89) si la condition est dans le vide ; et par (3.103) dans le cas de la géométrie circulaire avec condition de transmission dans le matériau. Ce nombre complexe a bien évidemment une dépendance implicite en z et β puisque son expression fait intervenir le symbole de l'opérateur T . Définissons alors les fonctions objectives suivantes

$$\begin{aligned} \text{Jacobi : } f_J(z, \beta, r) &= \sup_{s \in \mathbb{R}} |\lambda_{s,J}| \\ \text{Gauss-Seidel : } f_{GS}(z, \beta, r) &= \sup_{s \in \mathbb{R}} |\lambda_{s,GS}| \\ \text{GMRES : } f_K(z, \beta) &= \inf_{s \in \mathbb{R}} |\lambda_{s,K}| \end{aligned} \quad (3.113)$$

L'objectif de l'optimisation sera alors de trouver les paramètres (z^*, β^*, r^*) qui minimisent (respectivement maximisent) les fonctions objectives f_J et f_{GS} (respectivement f_K). Si l'objectif de minimiser les fonctions f_J et f_{GS} pour réduire le taux de convergence repose effectivement sur des fondements théoriques solides (les valeurs propres étant directement reliées au taux de convergence), la maximisation de f_K pour les méthodes de Krylov (dont le GMRES) est plutôt une heuristique. En effet, l'estimation du taux de convergence de ces méthodes pour des matrices quelconques est particulièrement complexe. Cependant, puisque les valeurs propres $\lambda_{s,K}$ sont regroupées dans le cercle centré en 1 et de rayon 1 du plan complexe, nous supposons ici que c'est la plus petite valeur propre qui influence le plus le conditionnement de la matrice (et donc le taux de convergence des méthodes de Krylov). Maximiser la plus petite valeur propre réduirait donc le conditionnement. L'objectif de maximisation correspond, dans l'exemple de la figure 3.3, à maximiser le rayon du cercle vert.

En pratique, il n'est pas possible de calculer directement les fonctions objectives f_J , f_{GS} ou f_K puisqu'elles font intervenir des $\sup |\lambda_s|$ (ou $\inf |1 - \lambda_s|$) pour $s \in \mathbb{R}$. Cependant, puisque les λ_s possèdent une limite λ_∞ finie quand $s \rightarrow \infty$, nous utiliserons la propriété suivante (écrite ici pour Jacobi, mais qui s'écrit de la même manière pour les autres méthodes itératives)

$$\exists s_M > 0 \text{ tel que } f_J(z, \beta, r) = \sup_{s \in \mathbb{R}} |\lambda_{s,J}| = \max \left(\sup_{s \in [0, s_M]} |\lambda_{s,J}|, |\lambda_{\infty,J}| \right). \quad (3.114)$$

Cette propriété traduit simplement le fait qu'à partir d'un certain moment, $|\lambda_s|$ tend de manière monotone vers $|\lambda_\infty|$. Si l'on souhaite optimiser des conditions de type Després ($\beta = 0$), il faut utiliser des fonctions objectives différentes. En effet, si $\beta = 0$, on a $f_J(z, 0, r) = f_{GS}(z, 0, r) = 1$ et $f_K(z, 0) = 0$ pour tout z et r . Dans ce cas, l'optimisation n'est pas effectuée sur \mathbb{R} mais uniquement sur $[0, s_M]$.

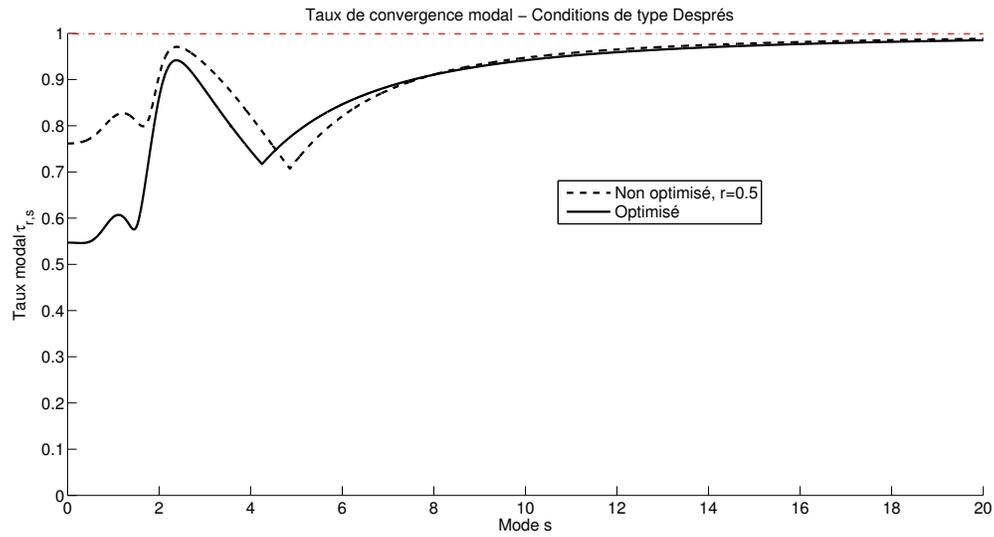
Au vu de l'expression du nombre complexe $Q_s(z, \beta)$, il n'est pas possible de résoudre analytiquement ce problème de minimisation (ou de maximisation) sous contraintes, même pour des cas très simplifiés. Nous avons donc développé un code `Matlab` pour approcher la solution de ce problème. Le code utilise la fonction `fminimax` de la toolbox *Optimization* qui permet la minimisation sous contrainte du maximum de la valeur absolue d'une fonction. Cette fonction reformule le problème sous la forme d'un problème d'optimisation non linéaire puis utilise une méthode itérative de type *Sequential Quadratic Programming* pour approcher la solution du problème. Pour plus d'informations sur les techniques d'optimisation de ce type, nous renverrons le lecteur aux articles [6], [19] et [33] par exemple.

Les exemples d'optimisation que nous allons présenter ont été réalisés sur la géométrie plane, dans le cas où l'interface de transmission est située dans le matériau, et nous prendrons dans toute la suite $s_M = 10$ dans (3.114). La condition absorbante d'ordre 1 est placée en $H = \frac{1}{2}$ (toutes les dimensions sont en m) et l'épaisseur du matériau (de paramètres $\mu = 3$ et $\epsilon = 2$) est $d = \frac{1}{3}$. Une condition impédante classique ($Z = 1$) est placée en $y = 0$. Enfin, l'interface de transmission est en $y_\Sigma = \frac{1}{4}$ de telle sorte qu'on ait bien $0 < y_\Sigma < d$. Le nombre d'onde vaut $k = 2\pi$. Les figures 3.8 présentent les taux de convergence modaux du problème, dans le cas des conditions de Després et dans le cas des conditions avec opérateur. Elles comparent les taux optimisés et non optimisés. Les paramètres et les taux maximum obtenus par l'optimisation sont donnés dans la table 3.1.

Comme attendu, nous pouvons tout d'abord constater que les taux de convergence modaux tendent vers 1 lorsque $s \rightarrow \infty$ lorsque les conditions de types Després sont utilisées. On retrouverait ce même comportement avec n'importe quelles conditions de transmission locales. Pour ces conditions, l'optimisation (qui est réalisée sur $[0, s_m]$ avec $s_M = 10$) ne réduit que peu le taux de convergence maximal, celui-ci diminuant de 0.97 à 0.94.

L'optimisation des conditions avec opérateur est plus intéressante. On voit clairement que le taux de convergence ne tend plus vers 1 quand $s \rightarrow \infty$ (bien que moins visible, cette remarque est vraie pour le cas non optimisé également). L'optimisation permet d'équilibrer, en un sens, les taux modaux des modes propagatifs ($s < 1$) et des modes évanescents ($s > 1$). Plus précisément, on s'aperçoit qu'après optimisation, le taux de convergence modal du mode 0, des modes rasants ($s \sim 1$) et des modes très évanescents ($s \rightarrow \infty$) sont égaux (voir la courbe rouge en trait mixte de la figure 3.8b). C'est une propriété que nous avons remarqué sur presque tous les cas d'optimisation que nous avons testés.

L'optimisation réduit significativement le maximum du taux de convergence modale, ce dernier passant de 0.99 à 0.85. Enfin, l'observation des coefficients optimisés est intéressante également (table 3.1). On voit tout d'abord que la partie imaginaire de z est assez faible. Nous verrons dans une section suivante que la partie imaginaire de z semble intervenir davantage lors de problème basse fréquence. On voit également que ce coefficient z est nettement plus faible lorsqu'on utilise un opérateur non local ($|z| \approx 0.1$) qu'avec des conditions de type Després ($|z| \approx 1.17$), alors que le coefficient β qui contrôle l'importance de l'opérateur non local est élevé. Ceci montre que sur cet exemple, l'opérateur non local joue un rôle important dans la diminution du taux de convergence.



(a) Conditions de type Després

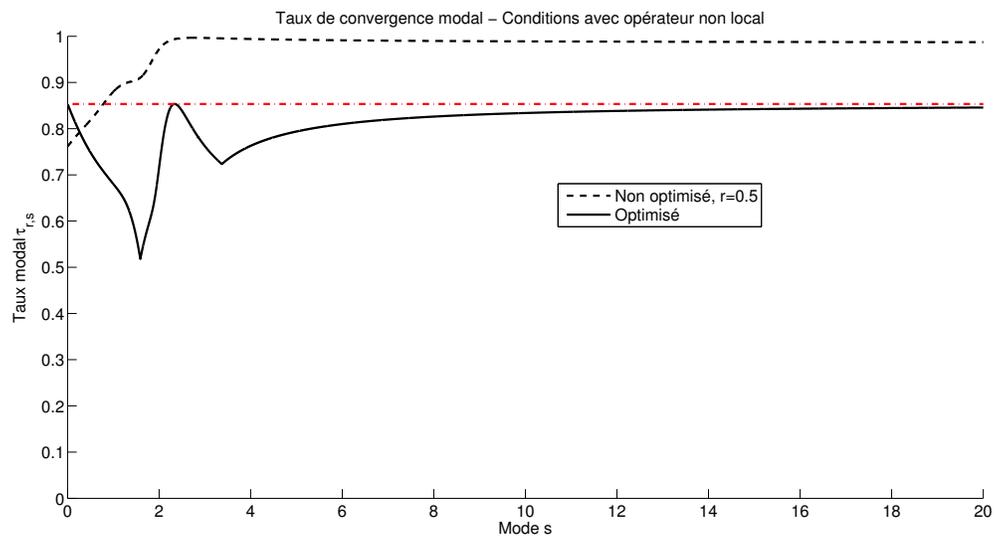
(b) Conditions avec opérateur non local $t_s = z|1 + \beta s|$

FIGURE 3.8 – Comparaison des taux de convergence modaux avec et sans optimisation pour des conditions de type Després (en haut) ou avec opérateur non local (en bas). L'optimisation est réalisée sur l'intervalle $[0, s_M]$, avec prise en compte de la limite lorsque $s \rightarrow \infty$ pour l'opérateur, et le graphique est tracé sur l'intervalle $[0, 2s_M]$.

TABLE 3.1 – Paramètres résultant de l’optimisation

	Després		Opérateur	
	Initial	Optimisé	Initial	Optimisé
z	1.0+1.0i	1.16-0.16i	1.0+1.0i	0.097-0.047i
β			1.0	1.97
r	0.5	0.58	0.5	0.61
$\max_{s \in [0, s_M]} \lambda_{s,J} $	0.97	0.94	0.99	0.85
$\sup_{s \in [0, \infty[} \lambda_{s,J} $	1.0	1.0	0.99	0.85

3.2.3.3 COMPARAISON DES GÉOMÉTRIES PLANE ET CIRCULAIRE

Comme nous l’avons vu, l’un des objectifs des calculs modaux précédents est de permettre d’optimiser nos conditions de transmission pour réduire le taux de convergence. Cependant, on peut légitimement se demander à quel point une optimisation réalisée sur un cas très simple comme le plan peut s’appliquer à un problème dont la géométrie est complexe. Il s’agit ici de justifier l’approximation du plan tangent, qui consiste à approcher une section courbe de la géométrie par le plan tangent à cette courbe, tout en gardant les mêmes épaisseurs de couche (voir figure 3.9). Nous comparerons ici les taux de convergence modaux (ou de manière

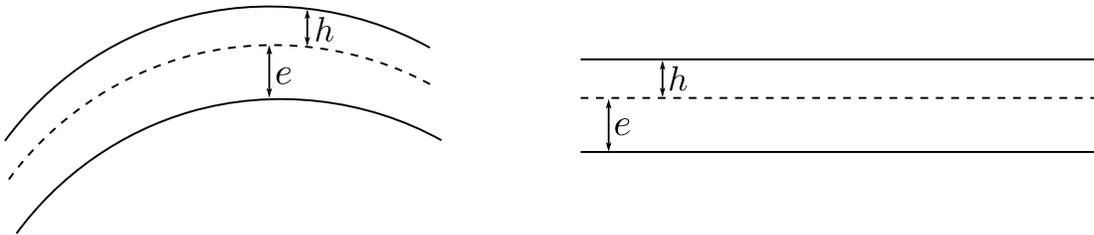
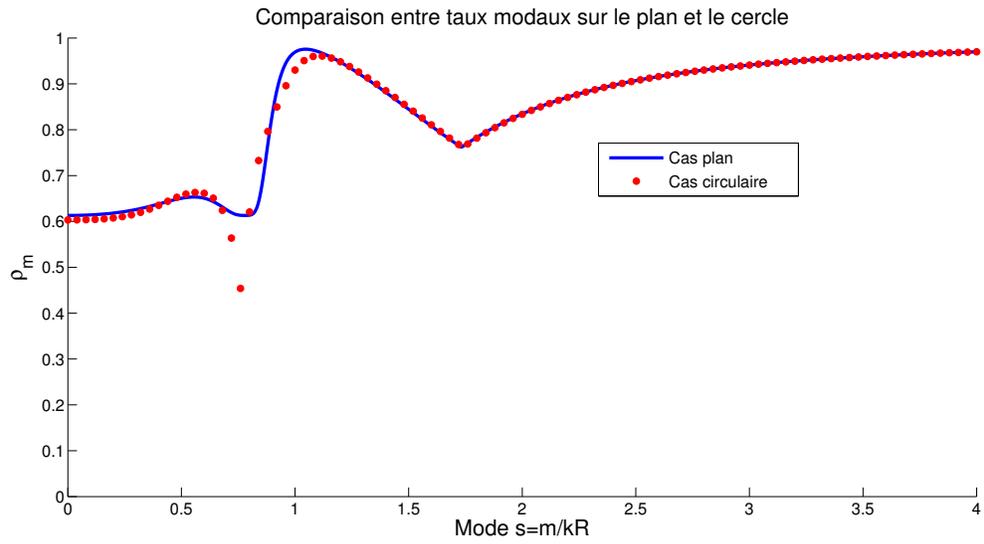


FIGURE 3.9 – La section courbe (figure de gauche) et approchée par son plan tangent (figure de droite), et les épaisseurs de couche sont conservées

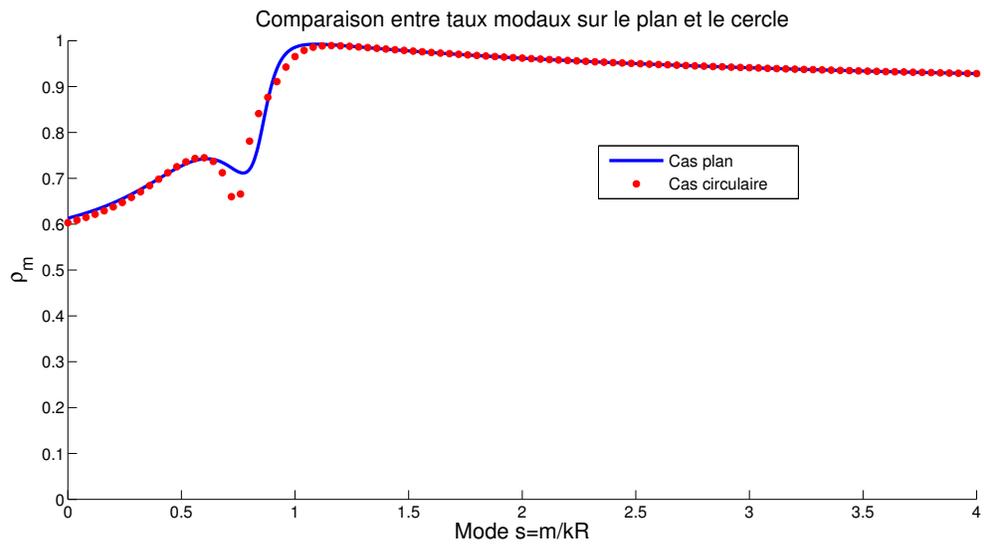
équivalente les valeurs propres) pour la géométrie circulaire et pour la géométrie plane. Étant donné la complexité des expressions des valeurs propres, l’objectif n’est pas de faire une étude comparative exhaustive, mais plutôt de montrer sur

quelques exemples que la courbure n'a que peu d'influence sur les taux modaux. Le rayon du cercle vaut $R_0 = 2$ et le nombre d'onde $k = 10$. La condition de transmission est placée à une distance $e = 0.5$ de l'objet et la condition absorbante est à une distance $h + e = 1.0$ de l'objet. Aucun matériau n'est présent ($\mu = \epsilon = 1$). Sur l'objet, une condition impédante est imposée ($Z = 1$).

À chaque fois, les mêmes paramètres sont utilisés pour les conditions de transmission sur le plan et sur le cercle. Les figures 3.10 et 3.11 présentent les résultats pour le cas non optimisé et optimisé respectivement. Dans chacune des situations, les taux de convergence modaux pour le plan et pour le cercle sont vraiment très proches. Notamment, les maxima des taux, les limites et les taux pour les faibles modes sont très similaires, et nous avons vu dans la section précédente que ces points sont particulièrement importants pour l'optimisation. On en conclut que le fait d'utiliser les paramètres optimisés du plan directement pour le cas de la géométrie circulaire est quasi-optimal. Ainsi, nous pourrions optimiser les problèmes que nous rencontrerons en utilisant les taux analytiques du plan, à condition de conserver les épaisseurs de couche. Évidemment, cette approximation du plan tangent est d'autant plus pertinente que kR_0 est grand.

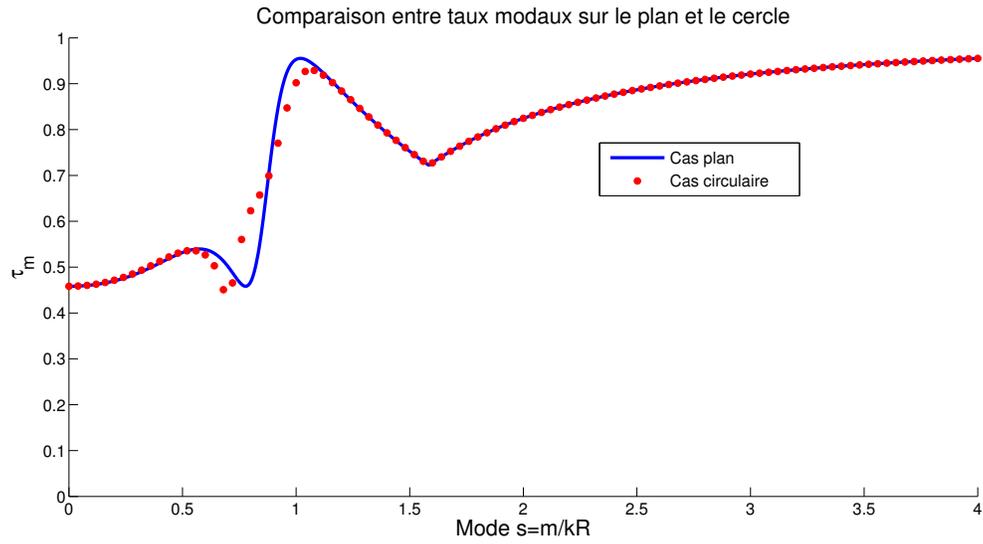


(a) Conditions de type Després

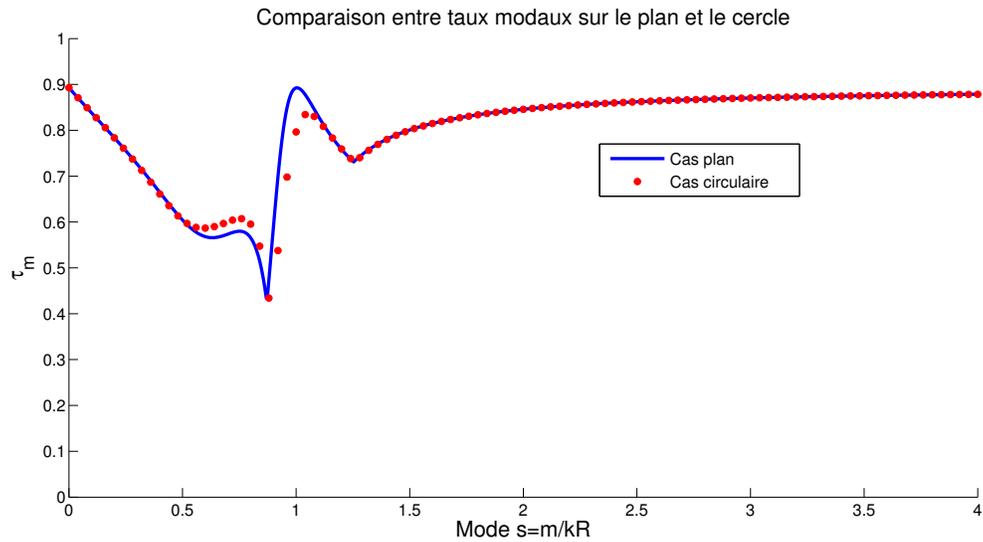


(b) Conditions avec opérateur non local

FIGURE 3.10 – Comparaison des taux de convergence modaux du plan et du cercle, pour des conditions de transmission de type Després ou avec opérateur non local. Les conditions de transmission ne sont pas optimisées.



(a) Conditions de type Després



(b) Conditions avec opérateur non local

FIGURE 3.11 – Comparaison des taux de convergence modaux du plan et du cercle, pour des conditions de transmission de type Després ou avec opérateur non local. Les conditions de transmission sont optimisées sur le plan et appliquées sur le cercle.

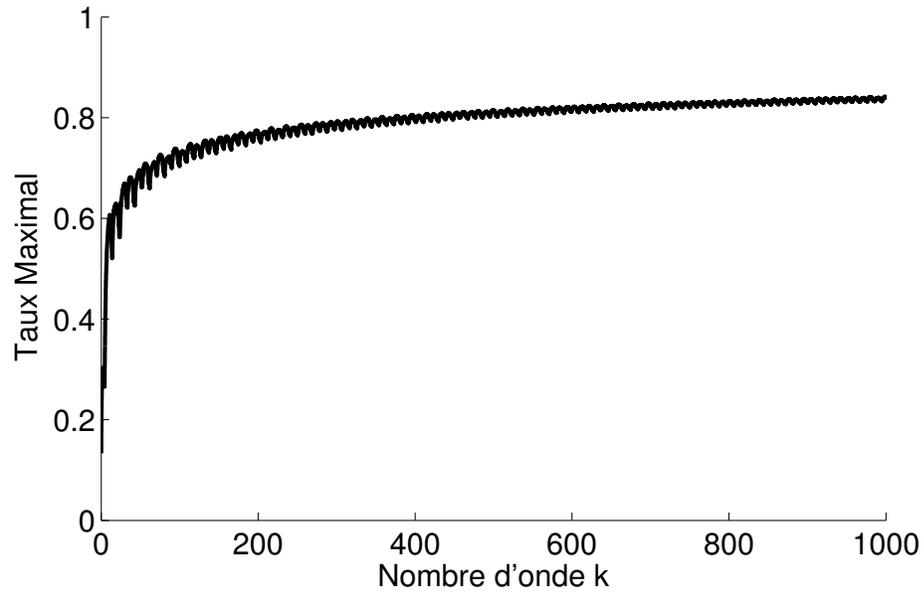
3.2.3.4 MONTÉE EN FRÉQUENCE

Dans cette section, nous allons nous intéresser à l'évolution du taux de convergence des méthodes itératives lorsque la fréquence augmente. Nous prendrons le cas de la géométrie circulaire simple (pas de matériau). Le rayon de l'objet est $R_0 = 1$. Une condition absorbante d'ordre 1 est placée en $R_2 = 2$ et la condition de transmission est située au milieu ($R_\Sigma = 1.5$). L'objet est supposé parfaitement conducteur (donc on impose une condition de Dirichlet sur l'objet). Le nombre d'onde varie de $k = 0.1$ à $k = 1000$, ce qui correspond à une fréquence f variant de $f = 4.7 \text{ MHz}$ à $f = 47.7 \text{ GHz}$. La condition de transmission imposée correspond aux conditions de transmission non locales, et l'opérateur a pour symbole $z |1 + \beta \frac{m}{kR_\Sigma}|$. Pour chaque nombre d'onde testé, nous effectuons une optimisation des coefficients z , β et du paramètre de relaxation r , et nous notons le taux de convergence maximal $\tau = \tau(k)$.

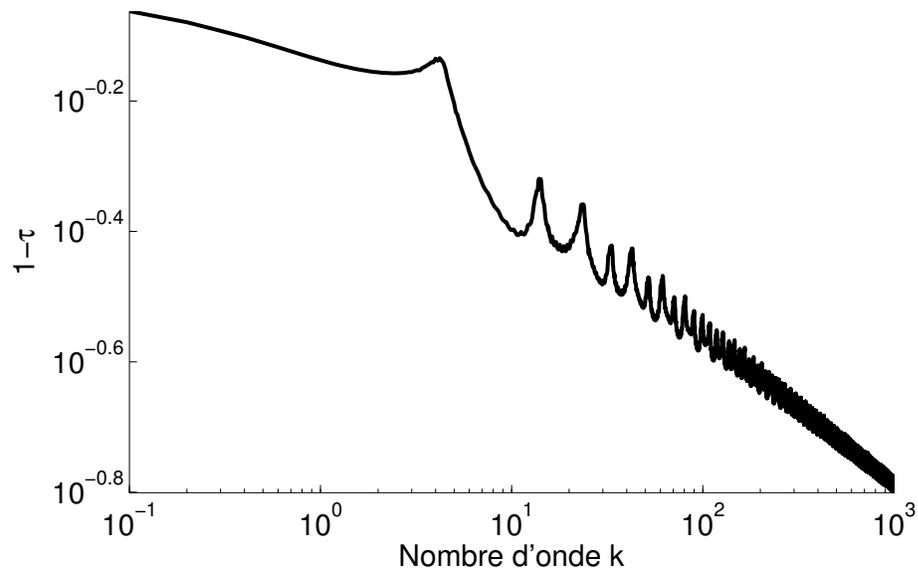
On peut ainsi observer que le taux de convergence maximal $\tau(k)$ augmente avec la fréquence. Cependant, cette augmentation est relativement limitée. La figure 3.12 représente d'une part le taux de convergence maximal $\tau(k)$ (3.12a), ainsi que la manière dont il tend vers 1 (3.12b). La figure 3.12b représente en échelle logarithmique $1 - \tau(k)$. S'il semble clair que le taux de convergence tend vers 1 avec la fréquence, cette tendance est assez lente. Une régression linéaire sur les données logarithmiques montrent que le comportement de $\tau(k)$ pour les grands nombres d'onde est

$$\tau(k) \sim 1 - \frac{0.78}{k^{0.226}}. \quad (3.115)$$

Enfin, nous avons représenté figure 3.13 l'évolution des paramètres optimaux en fonction du nombre d'onde. Le premier constat que nous pouvons faire est que pour des nombres d'onde suffisamment élevés, les paramètres semblent osciller autour d'une valeur moyenne et ne dépendent plus du nombre d'onde. On a par exemple le paramètre de relaxation qui vaut $r(k) \approx r^* = 0.65$ et le paramètre gouvernant l'opérateur $\beta(k) \approx \beta^* = 2.25$. On s'aperçoit également que l'argument de z oscille autour 0. Pour les hautes fréquences, il semble peu utile d'ajouter une partie complexe à z . Néanmoins, pour des nombres d'onde faibles $k < 10$, l'argument de z peut valoir jusqu'à 30° , ce qui donne une partie imaginaire non négligeable devant la partie réelle.



(a) Évolution du taux de convergence maximal τ en fonction du nombre d'onde k



(b) Évolution de $1 - \tau$ en fonction du nombre d'onde k (abscisses et ordonnées sont en échelle logarithmique).

FIGURE 3.12 – Représentation du taux de convergence maximal $\tau = \tau(k)$.

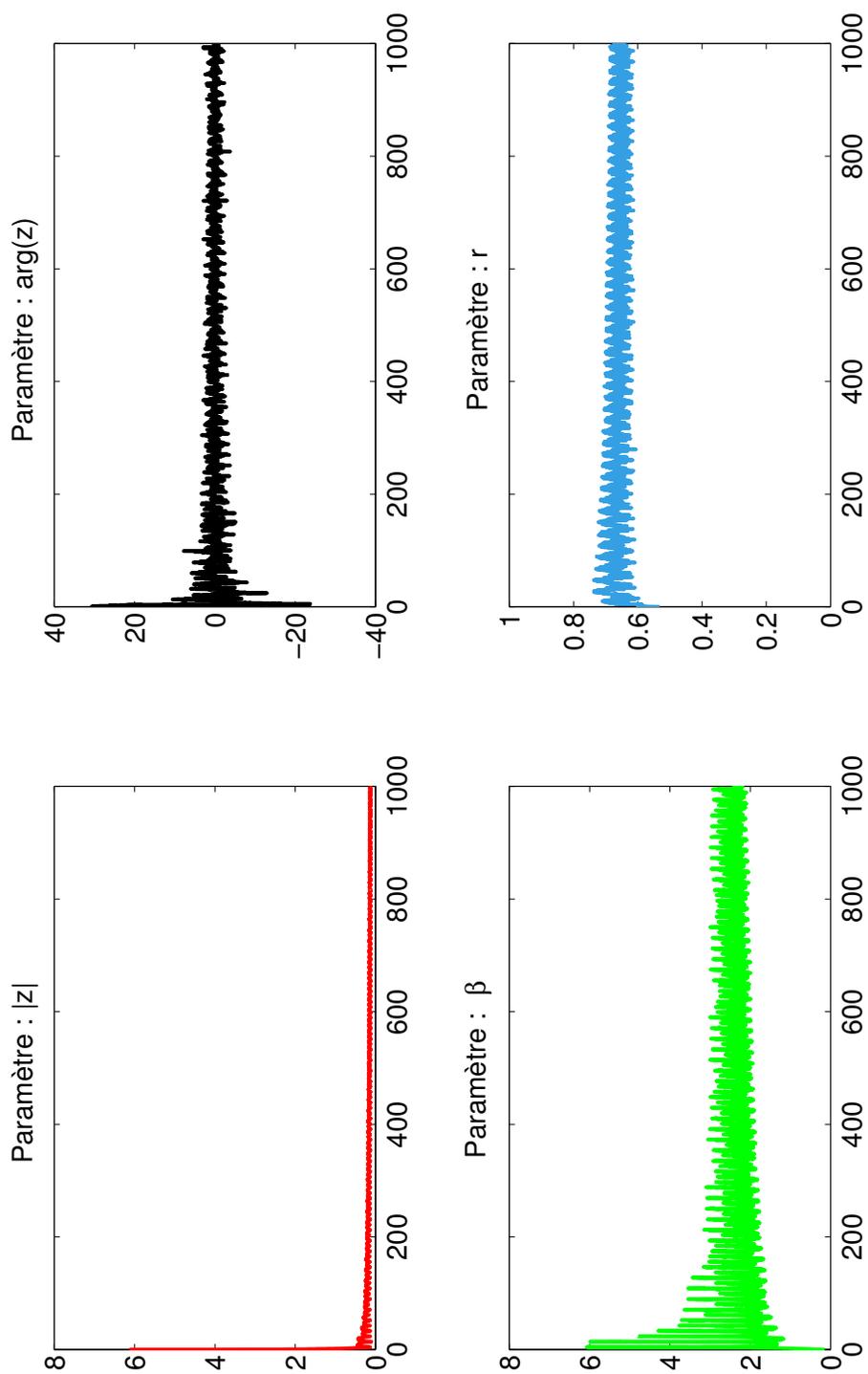


FIGURE 3.13 – Évolution des paramètres optimaux en fonction du nombre d'onde k . Les graphiques représentent respectivement $|z|$, $\arg(z)$ (en degrés), β et le paramètre de relaxation r .

Deuxième partie
Réalisation des opérateurs

4

Opérateurs induits par les potentiels de Riesz

Sommaire

4.1	Définition de l'opérateur Λ et notations	92
4.2	Opérateur de référence Λ_{ref}	93
4.3	Étude de l'opérateur Λ	94
4.3.1	Continuité de $\Lambda - \Lambda_{ref}$ dans $L^2(\Sigma)$	94
4.3.2	Propriété d'isomorphisme	98
4.4	Opérateur tronqué	99
4.5	Extension 3D	100

Après avoir étudié dans un cadre général les conditions de transmission basées sur notre classe d'opérateurs T , il est nécessaire maintenant de spécifier ces opérateurs. Ce chapitre sera consacré à l'étude d'opérateurs $T_r = \Lambda^* \Lambda$ pour lesquels la construction de Λ est basée sur la théorie des potentiels de Riesz [38]. Dans le chapitre suivant, nous verrons qu'il est possible de spécifier et construire d'autres types d'opérateurs.

4.1 DÉFINITION DE L'OPÉRATEUR Λ ET NOTATIONS

On se donne une courbe Σ de classe C^1 fermée, compacte et connexe du plan. On exclut les points doubles (courbe simple). Nous caractériserons cette courbe par son abscisse curviligne c'est à dire, la donnée d'une application

$$x : s \in \mathbb{R} \mapsto x(s) \in \Sigma, \quad (4.1)$$

avec $|\dot{x}(s)| = 1$ (où le point dénote la dérivation) et $x(s + L) = x(s)$ (x est L -périodique), L étant la longueur de la courbe. L'abscisse curviligne a une interprétation simple qui est la distance algébrique parcourue lorsque l'on tourne à vitesse constante le long de la courbe à partir d'un point de Σ dans le sens direct ($s > 0$) ou indirect ($s < 0$). Sur de telles courbes, on peut facilement construire une structure de variété de classe C^1 à l'aide par exemple des deux cartes :

$$\begin{aligned} & \left(U_1 = x^{-1}([0, L[), \quad]0, L[, \quad x_{|U_1}^{-1} \right), \\ & \left(U_2 = x^{-1} \left(\left] -\frac{L}{2}, \frac{L}{2} \right[\right), \quad \left] -\frac{L}{2}, \frac{L}{2} \right[, \quad x_{|U_2}^{-1} \right). \end{aligned} \quad (4.2)$$

Il est connu que toute variété de classe C^1 connexe et compacte est difféomorphe au cercle unité (ou bien sûr à tout autre cercle). Ce difféomorphisme Ψ_R est très simple à construire et utilise l'abscisse curviligne. Si \mathbb{S}_R est le cercle de rayon R , avec

$$R = \frac{L}{2\pi}, \quad (4.3)$$

on a

$$\Psi_R : x \in \Sigma \mapsto x_0 = \left(R \cos \left(\frac{s(x)}{R} \right), R \sin \left(\frac{s(x)}{R} \right) \right) \in \mathbb{S}_R, \text{ avec } x = x(s(x)). \quad (4.4)$$

Ce morphisme est bien défini car il ne dépend pas de la détermination choisie pour s : si $x(s) = x(t)$, on a $s - t \in 2\pi R\mathbb{Z}$ (la courbe est sans point double) et les valeurs des fonctions trigonométriques sont bien intrinsèquement définies. On remarquera que Ψ_R conserve les formes différentielles

$$dx = ds = dx_0. \quad (4.5)$$

Si on définit l'application θ_Σ qui à une fonction de x associe une fonction $2\pi R$ périodique *via*

$$\theta_\Sigma u(s) = u(x(s)) = \tilde{u}(s), \quad (4.6)$$

et si on pose

$$u_0(x_0) = (\psi_R^* u)(x_0) = \theta_\Sigma \circ \theta_{\mathbb{S}_R}^{-1} u(x_0), \quad (4.7)$$

on aura

$$\int_\Sigma u(x) dx = \int_{\mathbb{S}_R} u_0(x_0) dx_0. \quad (4.8)$$

Cette identification s'avère féconde car elle permet de transporter un problème défini sur Σ vers un autre problème défini sur le cercle, pour lequel on dispose d'outils puissants comme la transformation de Fourier.

L'objectif de ce chapitre est de démontrer que l'opérateur Λ défini par (4.9), possède les propriétés requises pour obtenir une convergence exponentielle dans le cadre de la DDM, notamment nous montrons que Λ est un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$.

$$\Lambda\varphi(x) = \varphi(x) - \beta \frac{\partial}{\partial \tau_x} \int_\Sigma |x - y|^{\frac{1}{2}} \frac{\partial \varphi}{\partial \tau_y}(y) d\sigma(y), \quad \beta \in \mathbb{C}, \quad (4.9)$$

où $\frac{\partial}{\partial \tau}$ est la dérivée tangentielle. Pour démontrer que cet opérateur est bien un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$, nous procéderons en deux étapes. Nous étudierons d'abord un opérateur de référence Λ_{ref} , qui coïnciderait avec notre opérateur Λ sur un cercle, puis nous montrerons que Λ est une perturbation compacte de l'opérateur de référence Λ_{ref} .

4.2 OPÉRATEUR DE RÉFÉRENCE Λ_{ref}

Introduisons l'opérateur de référence Λ_{ref} défini par

$$\Lambda_{ref}\varphi(s) = \varphi(s) - \beta \frac{\partial}{\partial s} \int_0^L \sqrt{\frac{L}{\pi}} \left| \sin \left[\frac{2\pi}{L} \frac{s-t}{2} \right] \right|^{\frac{1}{2}} \varphi'(t) dt \quad (4.10)$$

Nous allons étudier les valeurs propres de Λ_{ref} . Soit $(w_p)_{p \in \mathbb{N}}$ la base Fourier où

$$w_p(s) = e^{\frac{2i\pi}{L} ps}, \quad s \in [0, 2\pi]. \quad (4.11)$$

Alors, pour tout $p \in \mathbb{N}$,

$$\begin{aligned} \Lambda_{ref} w_p(s) &= w_p(s) - \beta \frac{\partial}{\partial s} \int_0^L \sqrt{\frac{L}{\pi}} \left| \sin \left[\frac{2\pi}{L} \frac{s-t}{2} \right] \right|^{\frac{1}{2}} \frac{2i\pi p}{L} w_p(t) dt \\ &= \left(1 + 2p^2 \beta \sqrt{\frac{\pi}{L}} \int_0^{2\pi} \left| \sin \frac{u}{2} \right|^{\frac{1}{2}} e^{ipu} du \right) w_p(s) \end{aligned} \quad (4.12)$$

Le calcul de l'intégral est détaillé dans la seconde section de l'annexe, dont le principal résultat est la formule

$$\hat{F}_p^{(\gamma)} = \int_{-\pi}^{\pi} \frac{e^{ip\theta}}{(2 - 2\cos\theta)^{\frac{1-\gamma}{2}}} d\theta = 2 \cos\left(\gamma\frac{\pi}{2}\right) \Gamma(\gamma) \frac{\Gamma\left(p + \frac{1-\gamma}{2}\right)}{\Gamma\left(p + \frac{1+\gamma}{2}\right)}. \quad (4.13)$$

Comme $\sin \frac{u}{2} = \frac{1}{2}\sqrt{2 - 2\cos u}$ pour $u \in [0, 2\pi]$, on obtient les valeurs propres ($\gamma = \frac{3}{2}$)

$$\lambda_p = 1 - p^2 \beta \frac{\pi}{\sqrt{L}} \frac{\Gamma\left(p - \frac{1}{4}\right)}{\Gamma\left(p + \frac{5}{4}\right)}. \quad (4.14)$$

Le calcul de l'asymptotique de λ_p pour $p \rightarrow +\infty$ montre que $\lambda_p \sim -\beta \frac{\pi}{\sqrt{L}} \sqrt{p}$. De plus, si l'on suppose que $\mathcal{I}m \beta \neq 0$, alors $\lambda_p \neq 0$ pour tout p et il est facile de montrer qu'il existe deux constantes c_1 et c_2 telles que

$$c_1 \leq \frac{|\lambda_p|}{(1 + p^2)^{\frac{1}{4}}} \leq c_2, \quad (4.15)$$

et Λ_{ref} est bien un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$.

4.3 ÉTUDE DE L'OPÉRATEUR Λ

Pour montrer que Λ est également un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$, nous montrerons que, sous réserve d'une certaine régularité sur la courbe Σ , l'opérateur Λ est une perturbation compacte de Λ_{ref} , ce qui donnera directement le résultat voulu. Pour cela, nous montrerons d'abord que l'opérateur $\Lambda - \Lambda_{ref}$ est continu dans $L^2(\Sigma)$, puis que Λ est injectif de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$. Ceci nous permettra de conclure. Dans la suite du chapitre, nous supposons que la courbe est de classe C^2 .

4.3.1 CONTINUITÉ DE $\Lambda - \Lambda_{ref}$ DANS $L^2(\Sigma)$.

Pour montrer la continuité de $\Lambda - \Lambda_{ref}$ dans L^2 , nous utiliserons le lemme suivant

Lemme 6 (Lemme de Schur).

Soit T un opérateur intégral : $T\varphi(s) = \int_0^L B(s,t)\varphi(t)dt$. On suppose les points suivants :

- $s \mapsto B(s,t)$ est dans $L^1(\Sigma)$ presque partout en t .
- $t \mapsto B(s,t)$ est dans $L^1(\Sigma)$ presque partout en s .
- $\sup_{t \in [0,L]} \|B(\cdot,t)\|_{L^1(\Sigma)} \leq c_1$ et $\sup_{s \in [0,L]} \|B(s,\cdot)\|_{L^1(\Sigma)} \leq c_2$

Alors T est borné de $L^2(\Sigma)$ dans $L^2(\Sigma)$: $\|T\varphi\|_{L^2(\Sigma)} \leq \sqrt{c_1 c_2} \|\varphi\|_{L^2(\Sigma)}$.

Démonstration. Soit T vérifiant les hypothèses du lemme. On a

$$|T\varphi(s)|^2 = \left| \int_0^L B(s,t)\varphi(t)dt \right|^2 \leq \left| \int_0^L |B(s,t)|^{\frac{1}{2}} |B(s,t)|^{\frac{1}{2}} \varphi(t)dt \right|^2 \quad (4.16)$$

Appliquons l'inégalité de Cauchy-Schwarz à cette dernière équation :

$$\begin{aligned} |T\varphi(s)|^2 &\leq \underbrace{\left(\int_0^L |B(s,t)| dt \right)}_{\|B(s,\cdot)\|_{L^1(\Sigma)}} \left(\int_0^L |B(s,t)| |\varphi(t)|^2 dt \right) \\ &\leq c_2 \int_0^L |B(s,t)| |\varphi(t)|^2 dt \end{aligned} \quad (4.17)$$

Dès lors, le calcul de la norme de $T\varphi$ donne

$$\|T\varphi\|_{L^2(\Sigma)}^2 = \int_0^L |T\varphi(s)|^2 ds \leq \int_0^L \left(c_2 \int_0^L |B(s,t)| |\varphi(t)|^2 dt \right) ds. \quad (4.18)$$

En inversant l'ordre des deux intégrales, on obtient

$$\begin{aligned} \|T\varphi\|_{L^2(\Sigma)}^2 &\leq c_2 \int_0^L \underbrace{\left(\int_0^L |B(s,t)| ds \right)}_{\|B(\cdot,t)\|_{L^1(\Sigma)}} |\varphi(t)|^2 dt \\ &\leq c_2 \int_0^L c_1 |\varphi(t)|^2 dt = c_1 c_2 \|\varphi\|_{L^2}^2 \end{aligned} \quad (4.19)$$

D'où le résultat voulu. ■

Par définition, on a, pour tout $\varphi \in L^2(\Sigma)$,

$$(\Lambda - \Lambda_{ref})\varphi(s) = \frac{\partial}{\partial s} \int_0^L N(s, t) \frac{\partial \varphi(t)}{\partial t} dt \quad (4.20)$$

où le noyau $N(s, t)$ est donné par

$$N(s, t) = |x(s) - x(t)|^{\frac{1}{2}} - \sqrt{\frac{L}{\pi}} \left| \sin \left[\frac{2\pi}{L} \frac{s-t}{2} \right] \right|^{\frac{1}{2}} \quad (4.21)$$

Après avoir inversé l'ordre de la dérivation et de l'intégration, puis intégration par parties, on trouve

$$(\Lambda - \Lambda_{ref})\varphi(s) = \int_0^L \frac{\partial^2 N}{\partial t \partial s}(s, t) \varphi(t) dt, \quad (4.22)$$

et on va chercher à appliquer le lemme 6 avec $B(s, t) = \frac{\partial^2 N}{\partial t \partial s}(s, t)$. Commençons par calculer les dérivées partielles de N . Notons N_Σ et N_C les fonctions

$$N_\Sigma = |x(s) - x(t)|^{\frac{1}{2}}, \quad N_C = \sqrt{\frac{L}{\pi}} \sin^{\frac{1}{2}} \left[\frac{2\pi}{L} \frac{|s-t|}{2} \right], \quad (4.23)$$

où x désigne l'abscisse curviligne de Σ . On a $N(s, t) = N_\Sigma(s, t) - N_C(s, t)$. On note de plus $R(s, t) = |x(s) - x(t)|$ et $F(R) = R^{\frac{1}{2}}$. On a tout d'abord

$$\frac{\partial N_\Sigma}{\partial s} = \frac{\partial R}{\partial s} F'(R) = (x(s) - x(t)) \cdot \dot{x}(s) \frac{F'(R)}{R}. \quad (4.24)$$

Puis, la seconde dérivation donne

$$\frac{\partial^2 N_\Sigma}{\partial t \partial s} = -\dot{x}(s) \cdot \dot{x}(t) \frac{F'(R)}{R} + \left(\tilde{R} \cdot \dot{x}(s) \right) \left(\tilde{R} \cdot \dot{x}(t) \right) R \frac{d}{dR} \left(\frac{F'(R)}{R} \right) \quad (4.25)$$

avec

$$\tilde{R} = \frac{x(s) - x(t)}{|x(s) - x(t)|}. \quad (4.26)$$

Ici $F(R) = \sqrt{R}$ donc $F'(R) = \frac{1}{2\sqrt{R}}$ et $\frac{F'(R)}{R} = \frac{1}{2R^{\frac{3}{2}}}$. Dès lors, $\frac{d}{dR} \left(\frac{F'(R)}{R} \right) = -\frac{3}{4} R^{-\frac{5}{2}}$.

On en déduit :

$$\frac{\partial^2 N_\Sigma}{\partial s \partial t} = \frac{1}{4R^{\frac{3}{2}}} \left(-2 \dot{x}(s) \cdot \dot{x}(t) + 3 \left(\tilde{R} \cdot \dot{x}(s) \right) \left(\tilde{R} \cdot \dot{x}(t) \right) \right) \quad (4.27)$$

Calculons maintenant les dérivées de N_C . En notant $a = \frac{L}{2\pi}$, on trouve

$$\frac{\partial N_C}{\partial s} = \frac{\partial}{\partial s} \left[\sqrt{a} \sin^{\frac{1}{2}} \frac{|s-t|}{a} \right] = \frac{1}{2\sqrt{a}} \frac{s-t}{|s-t|} \frac{\cos \frac{|s-t|}{a}}{\sin^{\frac{1}{2}} \frac{|s-t|}{a}}, \quad (4.28)$$

puis après plusieurs calculs que l'on ne détaillera pas

$$\frac{\partial^2 N_C}{\partial t \partial s} = \frac{2 - \cos^2 \frac{|s-t|}{a}}{4a^{\frac{3}{2}} \sin^{\frac{3}{2}} \frac{|s-t|}{a}} \quad (4.29)$$

De manière évidente, le noyau $B(s, t) = \frac{\partial^2 N_C}{\partial s \partial t} - \frac{\partial^2 N_C}{\partial s \partial t}$ est symétrique, il nous suffit donc d'étudier l'application $s \mapsto B(s, t)$ pour t fixé. Il est facile de voir que cette application est périodique de période L . On a donc

$$\int_0^L |B(s, t)| ds = \int_{t-\frac{L}{2}}^{t+\frac{L}{2}} |B(s, t)| ds \quad (4.30)$$

et on va montrer que cette intégrale est bornée. L'application $s \mapsto B(s, t)$ est continue donc bornée sur tout intervalle ne contenant pas $t + 2k\pi$, $k \in \mathbb{Z}$. Elle est donc bornée sur $[t - \frac{L}{2}, t - \epsilon] \cup [t + \epsilon, t + \frac{L}{2}]$ pour tout $\epsilon > 0$. De plus, le développement limité de $B(s, t)$ pour s au voisinage de t montre

$$B(s, t) = o\left(\frac{1}{|s-t|^{\frac{1}{2}}}\right). \quad (4.31)$$

Le lecteur trouvera le détail des calculs du développement limité en annexe. Notons que ce calcul nécessite que la courbe (et donc l'abscisse curviligne x) soit de classe C^2 . Donc, pour tout $M > 0$ il existe ϵ_M tel que pour tout s tel que $|s-t| < \epsilon_M$ on ait

$$B(s, t) \leq \frac{M}{|s-t|^{\frac{1}{2}}}. \quad (4.32)$$

Dès lors, l'intégrale donne

$$\int_{t-\frac{L}{2}}^{t+\frac{L}{2}} |B(s, t)| ds = \int_{t-\frac{L}{2}}^{t-\epsilon_M} |B(s, t)| ds + \int_{t-\epsilon_M}^{t+\epsilon_M} |B(s, t)| ds + \int_{t+\epsilon_M}^{t+\frac{L}{2}} |B(s, t)| ds. \quad (4.33)$$

Si l'on note

$$M_1 = \sup_{s \in [t-\frac{L}{2}, t-\epsilon_M]} |B(s, t)| \quad \text{et} \quad M_2 = \sup_{s \in [t+\epsilon_M, t+\frac{L}{2}]} |B(s, t)| \quad (4.34)$$

on trouve finalement

$$\int_0^L |B(s, t)| ds \leq \max(M_1, M_2) L + 4M \sqrt{\epsilon_M} < \infty. \quad (4.35)$$

Le lemme de Schur est donc applicable et ceci montre la continuité de l'opérateur $\Lambda - \Lambda_{ref}$ dans $L^2(\Sigma)$.

4.3.2 PROPRIÉTÉ D'ISOMORPHISME

Pour montrer que Λ est un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$, écrivons Λ sous la forme

$$\Lambda = \left(I + (\Lambda - \Lambda_{ref}) \Lambda_{ref}^{-1} \right) \Lambda_{ref}. \quad (4.36)$$

Puisque Λ_{ref} est un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$, il suffit de montrer que l'opérateur $I + (\Lambda - \Lambda_{ref}) \Lambda_{ref}^{-1}$ est un isomorphisme de $L^2(\Sigma)$. Notons tout d'abord que l'opérateur Λ est injectif. En effet, pour tout $\varphi \in H^{\frac{1}{2}}(\Sigma)$ tel que $\Lambda \varphi = 0$, on a

$$(\Lambda \varphi, \varphi) = \int_{\Sigma} \Lambda \varphi(x) \overline{\varphi(x)} dx = \|\varphi\|_{L^2(\Sigma)}^2 + \beta \int_{\Sigma} \int_{\Sigma} |x - y|^{\frac{1}{2}} \frac{\partial \varphi(y)}{\partial \tau} \frac{\partial \overline{\varphi(x)}}{\partial \tau} dy dx = 0 \quad (4.37)$$

Supposons maintenant que β soit imaginaire pur, alors en prenant la partie réelle de cette équation, on trouve $\varphi = 0$. Alors, l'opérateur $I + (\Lambda - \Lambda_{ref}) \Lambda_{ref}^{-1}$ est également injectif. Nous garderons dans la suite l'hypothèse $\beta \in i\mathbb{R}$.

Pour conclure *via* l'alternative de Fredholm que l'opérateur Λ est un isomorphisme, il nous suffit maintenant de montrer que $(\Lambda - \Lambda_{ref}) \Lambda_{ref}^{-1}$ est un opérateur compact dans $L^2(\Sigma)$. Pour cela, prenons ψ_n une suite bornée de $L^2(\Sigma)$. Alors la suite $\varphi_n = \Lambda_{ref}^{-1} \psi_n$ est bornée dans $H^{\frac{1}{2}}(\Sigma)$. Par compacité de l'inclusion de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$, on en déduit, à l'extraction d'une sous-suite près, que φ_n converge fortement dans $L^2(\Sigma)$. Alors, comme $\Lambda - \Lambda_{ref}$ est continu dans $L^2(\Sigma)$, $(\Lambda - \Lambda_{ref}) \varphi_n = (\Lambda - \Lambda_{ref}) \Lambda_{ref}^{-1} \psi_n$ converge fortement dans $L^2(\Sigma)$. D'où la compacité, et on vient de démontrer le lemme suivant

Lemme 7.

Si β est un imaginaire pur, l'opérateur Λ défini par

$$\Lambda\varphi(x) = \varphi(x) - \beta \frac{\partial}{\partial \tau_x} \int_{\Sigma} |x - y|^{\frac{1}{2}} \frac{\partial \varphi}{\partial \tau_y}(y) d\sigma(y) \quad (4.38)$$

est un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$.

4.4 OPÉRATEUR TRONQUÉ

Le principal inconvénient à l'utilisation d'opérateur intégraux tels que nous les avons construits ici est qu'ils conduisent, après discrétisation, à des matrices pleines, et donc à un surcoût prohibitif lors du processus itératif. Pour éviter cela, nous introduisons une fonction de troncature dans le noyau de l'opérateur. L'objectif de cette troncature est de localiser l'opérateur tout en conservant ses propriétés fonctionnelles. La troncature est introduite de la manière suivante

$$\Lambda_{\delta} \varphi(x) = \varphi(x) - \beta \frac{\partial}{\partial \tau_x} \int_{\Sigma} \chi_{\delta}(|x - y|) |x - y|^{\frac{1}{2}} \frac{\partial \varphi}{\partial \tau_y}(y) d\sigma(y) \quad (4.39)$$

où la fonction $\chi_{\delta}(r) = \chi\left(\frac{r}{\delta}\right)$ est une fonction suffisamment régulière telle que

$$\forall r, 0 \leq \chi(r) \leq 1, \quad \chi(r) = 1 \text{ pour } r \leq r_0, \quad \text{et } \chi(r) = 0 \text{ pour } r \geq r_1. \quad (4.40)$$

et δ est un paramètre gouvernant la localité de Λ_{δ} . Une telle fonction sera détaillée dans la section 7.2.2 du chapitre 7. La localité de l'opérateur ainsi défini apparaît grâce à la propriété suivante

$$\text{dist}(\text{supp } \varphi, \text{supp } \psi) \leq \delta \Rightarrow \langle \Lambda_{\delta} \varphi, \psi \rangle = 0. \quad (4.41)$$

Comme précédemment, pour montrer que Λ_{δ} reste un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$, nous allons montrer qu'il s'agit d'une perturbation compacte de l'opérateur Λ , qui est un isomorphisme. On écrit

$$\Lambda_{\delta} = \left(I + (\Lambda_{\delta} - \Lambda) \Lambda^{-1} \right) \Lambda. \quad (4.42)$$

Pour pouvoir appliquer la même démonstration que dans la section précédente, il suffit de montrer que Λ_{δ} est injectif et que $\Lambda_{\delta} - \Lambda$ est continu dans $L^2(\Sigma)$. Avec

ces arguments, l'opérateur $I + (\Lambda_\delta - \Lambda)$ sera injectif et $(\Lambda_\delta - \Lambda) \Lambda^{-1}$ compact et nous pourrons conclure *via* l'alternative de Fredholm. L'injectivité de Λ_δ est quasi immédiate puisque l'équation

$$(\Lambda_\delta \varphi, \varphi) = \|\varphi\|_{L^2(\Sigma)} + \beta \int_\Sigma \int_\Sigma \chi_\delta(|x-y|) |x-y|^{\frac{1}{2}} \frac{\partial \varphi(y)}{\partial \tau} \overline{\frac{\partial \varphi(x)}{\partial \tau}} dy dx = 0 \quad (4.43)$$

implique $\varphi = 0$ (il suffit de prendre la partie réelle de l'équation). De plus, comme la fonction $(x, y) \mapsto (1 - \chi_\delta(|x-y|)) |x-y|^{\frac{1}{2}}$ est bornée sur $\Sigma \times \Sigma$, (soit $C(\delta)$ cette borne), on a

$$|\langle (\Lambda_\delta - \Lambda) \varphi, \psi \rangle| \leq C(\delta) L \|\varphi\|_{L^2(\Sigma)} \|\psi\|_{L^2(\Sigma)}, \quad (4.44)$$

où L est la longueur de Σ . Alors, grâce à la compacité de l'inclusion $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$ et la continuité de $\Lambda_\delta - \Lambda$, on en déduit que $(\Lambda_\delta - \Lambda) \Lambda^{-1}$ est un opérateur compact. L'alternative de Fredholm s'applique bien et l'opérateur Λ_δ reste un isomorphisme.

4.5 EXTENSION 3D

L'opérateur Λ ayant été défini à l'aide de l'abscisse curviligne, son extension au cas 3D n'est pas évidente. Pour cette raison, nous admettrons le résultat suivant.

Lemme 8.

L'opérateur Λ défini par

$$\langle \Lambda u, v \rangle = \int_\Sigma u(x) \bar{v}(x) dx + \beta \int_\Sigma \int_\Sigma \chi_\delta(|x-y|) \frac{\text{rot}_\Sigma u(x) \cdot \text{rot}_\Sigma v(y)}{|x-y|^{d-\frac{3}{2}}} dx dy, \quad (4.45)$$

où d est la dimension de Σ ($d = 1$ en 2D, $d = 2$ en 3D) est un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$.

Quand $d = 2$, rot_Σ est l'opérateur rotationnel surfacique classique, tandis que si $d = 1$, il faut comprendre $\text{rot}_\Sigma = \partial_\tau$ comme la dérivée tangentielle 1D et on retrouve l'opérateur que nous avons défini en (4.9).

5

Opérateurs induits par les semi-normes de Gagliardo

Sommaire

5.1	Courbes et espaces de fonctions	102
5.1.1	Courbes höldériennes	102
5.1.2	Espaces de fonctions sur une courbe	105
5.1.2.1	Espaces $L^2(\Sigma)$ et $H^1(\Sigma)$	105
5.1.2.2	Les espaces $H^\alpha(\Sigma)$ et leur différentes défini- tions	106
5.1.3	Comparaison des normes sur le cercle	108
5.2	Opérateur T	111
5.3	Opérateur Λ	111
5.3.1	Cas du cercle	112
5.3.2	Cas d'une courbe $C^{1,r}$	115
5.3.2.1	Transposition du problème sur le cercle	116
5.3.2.2	Isomorphisme pour une courbe höldériennes	118
5.4	Troncature	122
5.5	Extension 3D	124
5.6	Conclusions	125

Dans le chapitre précédent, nous avons montré une première façon de construire un opérateur $T = \Lambda^* \Lambda$ (donné par la forme bilinéaire de Λ) vérifiant la propriété d'isomorphisme qui conduit à la convergence exponentielle des algorithmes itératifs. Cette construction était basée sur les potentiels de Riesz. Nous allons maintenant voir une seconde construction possible, cette fois à partir des semi-normes de Gagliardo des espaces de Sobolev d'ordre fractionnaire. Avant, nous allons introduire quelques notions qui vont nous être utiles dans la suite du chapitre.

5.1 COURBES ET ESPACES DE FONCTIONS

5.1.1 COURBES HÖLDÉRIENNES

Dans la suite on s'intéressera aux courbes un peu plus régulières que C^1 . On suppose qu'il existe une constante C_Σ et r un nombre positif non nul tels que $\dot{x}(s)$ satisfait à l'inégalité

$$|\dot{x}(s) - \dot{x}(t)| \leq C_\Sigma |t - s|^r. \quad (5.1)$$

On parlera de courbes höldériennes de classe $C^{1,r}$. Remarquons que plus r est petit et plus la contrainte est relâchée : une courbe de classe $C^{1,r'}$ est de classe $C^{1,r}$ si $r' > r$. L'analyse des opérateurs induits par les semi-normes de Gagliardo nous amènera à contrôler la singularité des fonctions

$$(s, t) \in [0, L] \times \left[-\frac{L}{2}, \frac{L}{2}\right] \mapsto |x(s+t) - x(s)|^\gamma, \quad (5.2)$$

avec γ un nombre réel donné. L'outil que nous utiliserons s'appuiera sur la comparaison des singularités entre la courbe Σ et le cercle $\mathbb{S}_R = \Psi_R(\Sigma)$. Pour le cercle, on a simplement

$$|x_0(s+t) - x_0(s)|^\gamma = R^\gamma \left(2 - 2 \cos \left(\frac{t}{R}\right)\right)^{\frac{\gamma}{2}}. \quad (5.3)$$

Pour les courbes höldériennes, on a la comparaison suivante,

Lemme 9.

Si $x(s)$ est höldérienne de classe $C^{1,r}$, pour tout γ réel, il existe une constante $C_{\Sigma, \gamma, r}$ telle que pour tout $(s, t) \in [0, 2\pi R] \times [-\pi R, \pi R]$

$$\left| |x(s+t) - x(s)|^\gamma - |x_0(s+t) - x_0(s)|^\gamma \right| \leq C_{\Sigma, \gamma, r} |x_0(s+t) - x_0(s)|^{\gamma+r}. \quad (5.4)$$

Ce lemme signifie que la différence entre les fonctions singulières attachées respectivement à la courbe et au cercle peut encore être une fonction singulière mais avec une singularité plus faible que la singularité commune initiale.

Démonstration. Par un simple changement d'échelle sur la courbe, il suffit de traiter le cas où $R = 1$. On peut supposer $\gamma \neq 0$ car le résultat est trivial dans ce cas. Nous allons montrer le résultat en deux étapes; nous commencerons par montrer que pour tout (s, t) dans $[0, 2\pi] \times [-\pi, \pi]$, si la courbe est $C^{1,r}$ alors il existe une constante $\tilde{C}_{\Sigma, \gamma, r}$ telle que

$$\left| |x(s+t) - x(s)|^\gamma - t^\gamma \right| \leq \tilde{C}_{\Sigma, \gamma, r} |t|^{\gamma+r}, \quad (5.5)$$

en appliquant cette estimation pour \mathbb{S}_1 et Σ on aura

$$\left| |x(s+t) - x(s)|^\gamma - |x_0(s+t) - x_0(s)|^\gamma \right| \leq \left(\tilde{C}_{\Sigma, \gamma, r} + \tilde{C}_{S_1, \gamma, r} \right) |t|^{\gamma+r}, \quad (5.6)$$

et comme

$$\frac{2|t|}{\pi} \leq \sqrt{2 - 2\cos t} \leq |t|, \quad |t| \leq \pi, \quad (5.7)$$

le résultat en découle en prenant

$$C_{\Sigma, \gamma, r} = \max \left(1, \left(\frac{\pi}{2} \right)^{\gamma+r} \right) \left(\tilde{C}_{\Sigma, \gamma, r} + \tilde{C}_{S_1, \gamma, r} \right). \quad (5.8)$$

Reste à montrer (5.5). On commence par remarquer que

$$|x(s+t) - x(s)| = \left| \int_s^{s+t} \dot{x}(\tau) d\tau \right| \leq \left| \int_s^{s+t} |\dot{x}(\tau)| d\tau \right| = |t|, \quad (5.9)$$

et donc

$$\frac{|x(s+t) - x(s)|}{|t|} \leq 1. \quad (5.10)$$

Maintenant, on a également

$$1 = |\dot{x}(s)| = \left| \frac{x(s+t) - x(s)}{t} - \frac{1}{t} \int_s^{s+t} (\dot{x}(\tau) - \dot{x}(s)) d\tau \right| \quad (5.11)$$

d'où il vient

$$\frac{|x(s+t) - x(s)|}{|t|} \geq 1 - \left| \frac{1}{t} \int_s^{s+t} (\dot{x}(\tau) - \dot{x}(s)) d\tau \right|. \quad (5.12)$$

Nous utilisons maintenant le caractère höldérien de la courbe pour obtenir la minoration, cf (5.1),

$$\frac{|x(s+t) - x(s)|}{|t|} \geq 1 - \frac{C_\Sigma}{|t|} \left| \int_s^{s+t} |\tau - s|^r d\tau \right| = 1 - \frac{C_\Sigma}{r+1} |t|^r. \quad (5.13)$$

On en déduit

- Si $\gamma > 0$ et $\frac{C_\Sigma}{r+1} |t|^r < 1$

$$\left(1 - \frac{C_\Sigma}{r+1} |t|^r \right)^{\frac{\gamma}{2}} \leq \frac{|x(s+t) - x(s)|^\gamma}{|t|^\gamma} \leq 1 \quad (5.14)$$

- Si $\gamma < 0$ et $\frac{C_\Sigma}{r+1} |t|^r < 1$

$$1 \leq \frac{|x(s+t) - x(s)|^\gamma}{|t|^\gamma} \leq \left(1 - \frac{C_\Sigma}{r+1} |t|^r \right)^{\frac{\gamma}{2}}, \quad (5.15)$$

que l'on résume en

$$\left| 1 - \frac{|x(s+t) - x(s)|^\gamma}{|t|^\gamma} \right| \leq \left| \left(1 - \frac{C_\Sigma}{r+1} |t|^r \right)^{\frac{\gamma}{2}} - 1 \right|. \quad (5.16)$$

On prend alors un nombre $\eta_\gamma < 1$ tel que pour tout η de $[0, \eta_\gamma]$, on ait

$$\left| \left(1 - \eta \right)^{\frac{\gamma}{2}} - 1 \right| \leq |\gamma| \eta, \quad (5.17)$$

(l'existence de η_γ se montre facilement à l'aide d'un développement limité) et on pose

$$t^* = \left(\eta_\gamma \frac{r+1}{C_\Sigma} \right)^{\frac{1}{r}}, \quad (5.18)$$

alors pour $0 < |t| < t^*$, on a

$$\left| 1 - \frac{|x(s+t) - x(s)|^\gamma}{|t|^\gamma} \right| \leq |\gamma| \frac{C_\Sigma}{r+1} |t|^r. \quad (5.19)$$

De plus, comme l'application

$$(s, t) \mapsto \left| 1 - \frac{|x(s+t) - x(s)|^\gamma}{|t|^\gamma} \right| \quad (5.20)$$

est continue sur $(s, t) \in K = [0, 2\pi] \times [-\pi, -t^*] \cup [0, 2\pi] \times [t^*, \pi]$, elle est bornée sur K . Soit $M(\Sigma, t^*, \gamma)$ cette borne, on aura pour (s, t) dans K

$$\left| 1 - \frac{|x(s+t) - x(s)|^\gamma}{|t|^\gamma} \right| \leq M(\Sigma, t^*, \gamma) \leq \frac{M(\Sigma, t^*, \gamma)}{(t^*)^r} |t|^r. \quad (5.21)$$

Posons $\tilde{C}(\Sigma, \gamma, r) = \max\left(\frac{M(\Sigma, t^*, \gamma)}{(t^*)^r}, |\gamma| \frac{C_\Sigma}{r+1}\right)$, alors (5.5) est satisfaite et par conséquence, (5.4) également. ■

5.1.2 ESPACES DE FONCTIONS SUR UNE COURBE

Dans cette section, nous allons nous intéresser à certains espaces de fonctions définies sur une courbe. On regardera en particulier la caractérisation des espaces de Sobolev d'indice fractionnaire.

5.1.2.1 ESPACES $L^2(\Sigma)$ ET $H^1(\Sigma)$

L'espace le plus simple est l'espace $L^2(\Sigma)$ des fonctions de carré intégrable muni de la norme

$$\|u\|_{L^2(\Sigma)} = \left(\int_\Sigma |u(x)|^2 dx \right)^{\frac{1}{2}} = \left(\int_0^{2\pi R} |\tilde{u}(s)|^2 ds \right)^{\frac{1}{2}} \quad (5.22)$$

où nous avons posé

$$\tilde{u}(s) = u(x(s)) := (\theta_\Sigma u)(s). \quad (5.23)$$

Il est intéressant de noter que la seconde intégrale est indépendante de Σ . Si on interprète cette intégrale comme une intégrale sur le cercle, on aura

$$\|u\|_{L^2(\Sigma)} = \|(\theta_{\mathbb{S}_R}^{-1} \circ \theta_\Sigma) u\|_{L^2(\mathbb{S}_R)}. \quad (5.24)$$

Ainsi, on peut identifier les espaces L^2 sur une variété de dimension un aux fonctions de carré intégrable $2\pi R$ périodiques. On peut par exemple utiliser la transformation de Fourier

$$(\mathcal{F}\tilde{u})_p = \frac{1}{\sqrt{2\pi R}} \int_0^{2\pi R} \tilde{u}(s) e^{-ip\frac{s}{R}} ds = \frac{1}{\sqrt{2\pi R}} \int_\Sigma u(x) e^{-ip\frac{s(x)}{R}} dx \quad (5.25)$$

pour caractériser cet ensemble. On a, grâce au théorème de Plancherel,

$$\|u\|_{L^2(\Sigma)}^2 = \sum_{p=-\infty}^{\infty} \left| (\mathcal{F}(\theta_\Sigma u))_p \right|^2. \quad (5.26)$$

On peut également s'intéresser aux espaces $H^1(\Sigma)$. Si on note

$$\vec{\nabla}_\Sigma u(x(s)) = \partial_s \tilde{u}(s) \dot{x}(s), \quad (5.27)$$

l'espace $H^1(\Sigma)$ est alors défini comme un sous espace de $L^2(\Sigma)$ tel que le module du gradient tangentiel soit de carré intégrable. $H^1(\Sigma)$ est muni de la norme

$$\|u\|_{H^1(\Sigma)}^2 = \|u\|_{L^2(\Sigma)}^2 + R^2 \|\vec{\nabla} u\|_{L^2(\Sigma)}^2 \quad (5.28)$$

(on a choisi de normaliser le gradient par $R = \frac{L}{2\pi}$). En remarquant que $|\vec{\nabla}_\Sigma u(x(s))| = |\partial_s \tilde{u}(s)|$ car $|\dot{x}(s)| = 1$, on a également

$$\|u\|_{H^1(\Sigma)}^2 = \int_0^{2\pi R} |\tilde{u}(s)|^2 ds + R^2 \int_0^{2\pi R} |\partial_s \tilde{u}(s)|^2 ds \quad (5.29)$$

et enfin on peut caractériser la norme sur $H^1(\Sigma)$ *via* les coefficients de Fourier

$$\|u\|_{H^1(\Sigma)}^2 = \sum_{p=-\infty}^{\infty} \left| (\mathcal{F}(\theta_\Sigma u))_p \right|^2 (1 + p^2). \quad (5.30)$$

Dans la suite, on s'intéressera aux fonctions de moyenne nulle sur Σ . On notera

$$L_{mo}^2(\Sigma) = \left\{ u \in L^2(\Sigma), \int_\Sigma u(s) ds = 0 \right\} \quad (5.31)$$

et

$$H_{mo}^1(\Sigma) = \left\{ u \in H^1(\Sigma), \int_\Sigma u(s) ds = 0 \right\} \quad (5.32)$$

Ces espaces sont également caractérisés par

$$(\mathcal{F}(\theta_\Sigma u))_0 = 0 \quad (5.33)$$

puisque le mode 0 n'est autre que la moyenne à une constante multiplicative près.

5.1.2.2 LES ESPACES $H^\alpha(\Sigma)$ ET LEUR DIFFÉRENTES DÉFINITIONS

Il existe de multiples manières de définir les espaces $H^\alpha(\Sigma)$ ou encore $H^\alpha(\mathbb{S}_R)$ où α est un nombre choisi dans $]0, 1[$. La plus abstraite utilise la théorie de l'interpolation

$$H^\alpha(\Sigma) = [L^2(\Sigma), H^1(\Sigma)]_\alpha, \quad (5.34)$$

la plus constructive les coefficients de Fourier (pour \mathbb{S}_R). Enfin, une alternative repose sur les semi-normes de Gagliardo. Dans le cas $\alpha = \frac{1}{2}$, on peut également définir $H^{\frac{1}{2}}(\Sigma)$ comme les traces des fonctions $H^1(\Omega)$ où Ω est le domaine intérieur à Σ . Bien sûr toutes ces définitions conduisent à la même structure topologique (ce qui n'est souvent pas facile à montrer) même si les normes ou encore les produits scalaires diffèrent d'un choix à l'autre (avec la propriété d'équivalence toutefois). Ici nous nous contenterons d'utiliser et de comparer deux définitions. On commence par définir pour α réel

$$\|u\|_{H_f^\alpha(\Sigma)}^2 = \sum_{p=-\infty}^{\infty} \left| (\mathcal{F}(\theta_\Sigma u))_p \right|^2 (1+p^2)^\alpha \quad (5.35)$$

Si α est positif, on définit l'espace

$$H_f^\alpha(\Sigma) = \left\{ u \in L^2(\Sigma), \|u\|_{H_f^\alpha(\Sigma)} < \infty \right\} \quad (5.36)$$

Cet espace a un dual facilement définissable, si $H_f^{-\alpha}(\Sigma) = (H_f^\alpha(\Sigma))'$ alors

$$H_f^{-\alpha}(\Sigma) = \left\{ u \in \mathcal{D}(\Sigma), \|u\|_{H_f^{-\alpha}(\Sigma)} < \infty \right\} \quad (5.37)$$

où $\mathcal{D}(\Sigma)$ est l'espace des distributions sur Σ . On a la dualité

$$\langle u, v \rangle_{H_f^{-\alpha}(\Sigma), H_f^\alpha(\Sigma)} = \sum_{p=-\infty}^{\infty} (\mathcal{F}(\theta_\Sigma u))_p \overline{(\mathcal{F}(\theta_\Sigma v))_p} \quad (5.38)$$

Maintenant, on va se tourner vers les semi-normes de Gagliardo. Si α est un nombre réel positif, on note

$$H_g^\alpha(\Sigma) = \left\{ u \in L^2(\Sigma), |u|_{H_g^\alpha(\Sigma)} < \infty \right\} \quad (5.39)$$

avec

$$|u|_{H_g^\alpha(\Sigma)}^2 = \frac{1}{2} \int_\Sigma \int_\Sigma \frac{(u(x) - u(y))^2}{|x - y|^{1+2\alpha}} dx dy \quad (5.40)$$

Si on utilise la paramétrisation de la courbe, on a

$$|u|_{H_g^\alpha(\Sigma)}^2 = \frac{1}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(\tilde{u}(s+t) - \tilde{u}(s))^2}{|x(s+t) - x(s)|^{1+2\alpha}} ds dt \quad (5.41)$$

Une fois les semi-normes définies, on construit une norme sur $H_g^\alpha(\Sigma)$ via

$$\|u\|_{H_g^\alpha(\Sigma)}^2 = \|u\|_{L^2(\Sigma)}^2 + R^{2\alpha} |u|_{H_g^\alpha(\Sigma)}^2 \quad (5.42)$$

5.1.3 COMPARAISON DES NORMES SUR LE CERCLE

On commence par comparer les normes lorsque la courbe est le cercle de rayon R . Le principal résultat de ce paragraphe est contenu dans la

Proposition 10.

Soit $\alpha \in [0, 1]$, alors il existe deux constantes $C^\pm(\alpha) > 0$ telles que

$$C^-(\alpha) \|u\|_{H_f^\alpha(\mathbb{S}_R)}^2 \leq \|u\|_{H_g^\alpha(\mathbb{S}_R)}^2 \leq C^+(\alpha) \|u\|_{H_f^\alpha(\mathbb{S}_R)}^2 \quad (5.43)$$

Cette proposition établit l'équivalence des normes précédemment définies pour le cercle. Ce résultat n'est pas nouveau : on retrouve la cohérence des différentes définitions des espaces $H^\alpha(\mathbb{S}_R)$. Toutefois, la dérivation précise des constantes intervenant dans les estimations est (à notre connaissance) nouvelle.

Démonstration. Par densité des polynômes trigonométriques dans $L^2(\mathbb{S}_R)$, il suffit de démontrer (5.43) pour u de la forme

$$u(x_0(s)) = \tilde{u}(s) = \sum_{|p| \leq P} \hat{u}_p e^{ip \frac{s}{R}} \quad (5.44)$$

où P est un entier positif donné. Le calcul des coefficients de Fourier de u donne

$$\|u\|_{H_f^\alpha(\mathbb{S}_R)}^2 = 2\pi R \sum_{|p| \leq P} |\hat{u}_p|^2 (1 + p^2)^\alpha \quad (5.45)$$

Par ailleurs, on a

$$\begin{aligned} R^{2\alpha} \|u\|_{H_g^\alpha(\mathbb{S}_R)}^2 &= \\ &= \frac{1}{2} R^{2\alpha} \sum_{|p_1| \leq P} \sum_{|p_2| \leq P} \hat{u}_{p_1} \overline{\hat{u}_{p_2}} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(e^{ip_1 \frac{s+t}{R}} - e^{ip_1 \frac{s}{R}})(e^{-ip_2 \frac{s+t}{R}} - e^{-ip_2 \frac{s}{R}})}{R^{1+2\alpha} (2 - 2 \cos t)^{\frac{1}{2} + \alpha}} ds dt \end{aligned} \quad (5.46)$$

soit, avec les changements de variables $s = R\varphi$ et $t = R\theta$

$$\begin{aligned}
 R^{2\alpha} |u|_{H_g^\alpha(\mathbb{S}_R)}^2 &= \\
 &\sum_{|p_1| \leq P} \sum_{|p_2| \leq P} \hat{u}_{p_1} \overline{\hat{u}_{p_2}} \frac{R}{2} \left(\int_0^{2\pi} e^{i(p_1 - p_2)\varphi} d\varphi \right) \left(\int_{-\pi}^{\pi} \frac{(e^{ip_1\theta} - 1)(e^{-ip_2\theta} - 1)}{(2 - 2\cos\theta)^{\frac{1}{2} + \alpha}} d\theta \right) \\
 &= \pi R \sum_{|p| \leq P} |\hat{u}_p|^2 \int_{-\pi}^{\pi} \frac{2 - 2\cos p\theta}{(2 - 2\cos\theta)^{\frac{1}{2} + \alpha}} d\theta \quad (5.47)
 \end{aligned}$$

et finalement,

$$R^{2\alpha} |u|_{H_g^\alpha(\mathbb{S}_R)}^2 = 2\pi R \sum_{|p| \leq P} |\hat{u}_p|^2 \hat{\Lambda}_p^{(\alpha)} \quad (5.48)$$

avec

$$\hat{\Lambda}_p^{(\alpha)} = \int_{-\pi}^{\pi} \frac{1 - \cos p\theta}{(2 - \cos\theta)^{\frac{1}{2} + \alpha}} d\theta. \quad (5.49)$$

Le calcul des coefficients de Fourier sont donnés en annexe. On a

$$\hat{\Lambda}_p^{(\alpha)} = -2 \cos(\pi\alpha) \Gamma(-2\alpha) \left(\frac{\Gamma(|p| + \frac{1}{2} + \alpha)}{\Gamma(|p| + \frac{1}{2} - \alpha)} - \frac{\Gamma(\frac{1}{2} + \alpha)}{\Gamma(\frac{1}{2} - \alpha)} \right) \quad (5.50)$$

et en particulier

$$\hat{\Lambda}_0^{(\alpha)} = 0, \quad \hat{\Lambda}_p^{(\alpha)} = \hat{\Lambda}_{-p}^{(\alpha)} > 0 \text{ si } p \neq 0. \quad (5.51)$$

Le coefficient $-2 \cos(\pi\alpha) \Gamma(-2\alpha)$ est toujours positif pour $\alpha \in]0, 1[$, il explose avec un pôle simple quand $\alpha \downarrow 0^+$ ou $\alpha \uparrow 1^-$. Le cas $\alpha = \frac{1}{2}$ est particulier, dans le sens où la formule doit être prise à la limite lorsque $\alpha \rightarrow \frac{1}{2}$. On trouve alors simplement

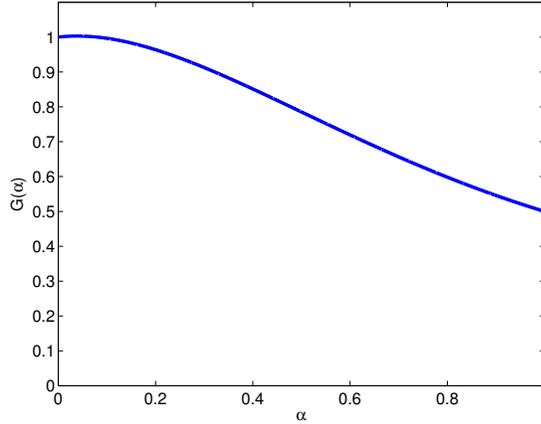
$$\hat{\Lambda}_p^{(\frac{1}{2})} = \pi|p| \quad (5.52)$$

Sur la figure suivante 5.1, on a représenté le graphe de l'application

$$\alpha \mapsto G(\alpha) = \alpha(1 - \alpha) \left(-2 \cos(\pi\alpha) \Gamma(-2\alpha) \right) \quad (5.53)$$

Sur $[0, 1]$, on voit clairement que le coefficient $G(\alpha)$ atteint un minimum en 1 valant $\frac{1}{2}$. Maintenant, on remarque que

$$\frac{\Gamma(|p| + \frac{1}{2} + \alpha)}{\Gamma(|p| + \frac{1}{2} - \alpha)} = \frac{\Gamma([|p| + \frac{1}{2} - \alpha] + 2\alpha)}{\Gamma([|p| + \frac{1}{2} - \alpha])}, \quad (5.54)$$

FIGURE 5.1 – Représentation graphique de $\alpha \mapsto G(\alpha)$, définie dans (5.53)

et comme

$$\frac{\Gamma(x+2\alpha)}{\Gamma(x)} \sim x^{2\alpha}, \quad x \rightarrow \infty \quad (5.55)$$

on en déduit que pour $p \neq 0$, la suite

$$\hat{\Lambda}_p^{(\alpha)} (1+p^2)^{-\alpha} \quad (5.56)$$

est positive, non nulle et tend vers $\frac{G(\alpha)}{\alpha(1-\alpha)} > \frac{1}{2\alpha(1-\alpha)} \geq 2$ lorsque p tend vers l'infini. Il existe donc deux constantes ne dépendant que de α telles que

$$C_\alpha^- \leq \hat{\Lambda}_p^{(\alpha)} (1+p^2)^{-\alpha} \leq C_\alpha^+, \quad \text{pour tout } p \neq 0 \quad (5.57)$$

On a donc

$$\begin{aligned} C_\alpha^- 2\pi R \sum_{p \neq 0, |p| \leq P} |\hat{u}_p|^2 (1+p^2)^\alpha &\leq 2\pi R \sum_{p \neq 0, |p| \leq P} |\hat{u}_p|^2 \hat{\Lambda}_p^{(\alpha)} \\ &\leq C_\alpha^+ 2\pi R \sum_{p \neq 0, |p| \leq P} |\hat{u}_p|^2 (1+p^2)^\alpha \end{aligned} \quad (5.58)$$

Si l'on introduit les espaces de fonctions à moyenne nulle

$$H_{mo}^s(\mathbb{S}_R) = H^s(\mathbb{S}_R) \cap L_{mo}^2(\mathbb{S}_R) \quad (5.59)$$

et si u est un polynôme trigonométrique dans $H_{mo}^s(\mathbb{S}_R)$, on a $\hat{u}_0 = 0$, et, d'après (5.45) et (5.48),

$$C_\alpha^- \|u\|_{H_f^\alpha(\mathbb{S}_R)}^2 \leq R^{2\theta} |u|_{H_g^\alpha(\mathbb{S}_R)}^2 \leq C_\alpha^+ \|u\|_{H_f^\alpha(\mathbb{S}_R)}^2 \quad (5.60)$$

comme les constantes C_α^\pm ne dépendent pas de P , la densité des fonctions harmoniques dans $L^2(\mathbb{S}_R)$ permet de conclure à l'équivalence de la semi-norme de Gagliardo avec la norme Fourier sur $H_f^\alpha(\mathbb{S}_R) \cap L_{mo}^2(\mathbb{S}_R)$.

Maintenant, si l'on rajoute la part L^2 de la norme, on a facilement

$$\min(C_\alpha^-, 1) \|u\|_{H_f^\alpha(\mathbb{S}_R)}^2 \leq \|u\|_{H_g^\alpha(\mathbb{S}_R)}^2 \leq (C_\alpha^+ + 1) \|u\|_{H_f^\alpha(\mathbb{S}_R)}^2, \quad (5.61)$$

ce qui termine la preuve. ■

5.2 OPÉRATEUR T

À partir des semi-normes de Gagliardo, il est naturel de construire un opérateur T sur Σ vérifiant l'hypothèse 1 (voir page 28). Nous avons le lemme suivant

Lemme 10 (Stratégie 1 : Opérateur $T_r^{(1)}$).

Étant donné un réel $\beta > 0$, l'opérateur $T_r^{(1)}$ défini par

$$\langle T_r^{(1)} u, v \rangle_\Sigma = \int_\Sigma u(x) \bar{v}(x) dx + \beta \int_\Sigma \int_\Sigma \frac{(u(x) - u(y))(\bar{v}(x) - \bar{v}(y))}{|x - y|^2} dx dy. \quad (5.62)$$

vérifie l'hypothèse 1.

Démonstration. Cet opérateur est clairement symétrique et positif ($\beta > 0$). De plus, il est évident que cette forme bilinéaire définit une norme équivalente à la norme de Gagliardo (5.42) sur $H^{\frac{1}{2}}(\Sigma)$. En effet, on a

$$\min\left(\frac{\beta}{R}, 1\right) \|u\|_{H_g^{\frac{1}{2}}(\Sigma)}^2 \leq \langle T_r^{(1)} u, u \rangle_\Sigma \leq \max\left(\frac{\beta}{R}, 1\right) \|u\|_{H_g^{\frac{1}{2}}(\Sigma)}^2. \quad (5.63)$$

En tant que norme sur $H^{\frac{1}{2}}(\Sigma)$, c'est un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans son dual $H^{-\frac{1}{2}}(\Sigma)$. ■

5.3 OPÉRATEUR Λ

Pour construire l'opérateur T_r , nous pouvons également le construire *via* un opérateur Λ tel que $T_r = \Lambda^* \Lambda$. Par construction, T_r ainsi défini est automatiquement

symétrique et positif. Nous devons simplement construire Λ de telle sorte que ce soit un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$. Encore une fois, les semi-normes de Gagliardo vont nous permettre de faire cela. Cependant, la démonstration n'est pas aussi évidente. Notons $\Lambda^{(2)}$ l'opérateur défini par

$$\langle \Lambda^{(2)} u, v \rangle_{\Sigma} = \int_{\Sigma} u(x) \bar{v}(x) dx + \beta \int_{\Sigma} \int_{\Sigma} \frac{(u(x) - u(y))(\bar{v}(x) - \bar{v}(y))}{|x - y|^{\frac{3}{2}}} dx dy. \quad (5.64)$$

Cet opérateur est de manière évidente un isomorphisme de $H^{\frac{1}{4}}(\Sigma)$ dans son dual $H^{-\frac{1}{4}}(\Sigma)$. Il s'agit d'un opérateur pseudo-différentiel d'ordre $\frac{1}{2}$ et nous pouvons donc nous attendre à ce qu'il soit également un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$. Nous n'avons pas trouvé de résultats explicites dans la littérature pour nous permettre de conclure, et dans cette section, nous allons montrer que c'est le cas sous certaines conditions de régularité de Σ .

5.3.1 CAS DU CERCLE

Si u et v sont deux fonctions régulières sur le cercle \mathbb{S}_R , et α un nombre dans $]0, 1[$, on considère la forme bilinéaire

$$b_R^{(\alpha)}(u, v) = R^{2\alpha} \frac{1}{2} \int_{\mathbb{S}_R} \int_{\mathbb{S}_R} \frac{(u(x_0) - u(y_0))(v(x_0) - v(y_0))}{|x_0 - y_0|^{1+2\alpha}} dx_0 dy_0 \quad (5.65)$$

qui est telle que

$$b_R^{\alpha}(u, u) = R^{2\alpha} |u|_{H_g^{\alpha}(\mathbb{S}_R)}^2. \quad (5.66)$$

À partir de cette forme bilinéaire, on construit l'opérateur linéaire associé

$$\langle \Lambda_R^{(\alpha)} u, v \rangle = b_R^{(\alpha)}(u, v). \quad (5.67)$$

Cet opérateur peut s'interpréter comme une valeur principale; on a

$$\Lambda_R^{(\alpha)} u(x_0) = \text{V.P.} \int_{\mathbb{S}_R} R^{2\alpha} \frac{u(y_0) - u(x_0)}{|x_0 - y_0|^{1+2\alpha}} dy_0 \quad (5.68)$$

Une simple inégalité de Cauchy-Schwarz donne

$$\left| \langle \Lambda_R^{(\alpha)} u, v \rangle \right| \leq R^{2\alpha} |u|_{H_g^{\alpha}(\mathbb{S}_R)} |v|_{H_g^{\alpha}(\mathbb{S}_R)} \leq C^+(\alpha) \|u\|_{H_f^{\alpha}(\mathbb{S}_R)} \|v\|_{H_f^{\alpha}(\mathbb{S}_R)} \quad (5.69)$$

et donc

$$\sup_{v \in H_f^\alpha(\mathbb{S}_R)} \frac{\left| \langle \Lambda_R^{(\alpha)} u, v \rangle \right|}{\|v\|_{H_f^\alpha(\mathbb{S}_R)}} \leq C^+(\alpha) \|u\|_{H_f^\alpha(\mathbb{S}_R)}. \quad (5.70)$$

Ceci montre que $\Lambda_R^{(\alpha)}$ est continu de $H_f^\alpha(\mathbb{S}_R)$ dans son dual $H_f^{-\alpha}(\mathbb{S}_R)$. Comme $\left| \langle \Lambda_R^{(\alpha)} u, u \rangle \right|$ est une norme sur $H_{mo}^\alpha(\mathbb{S}_R)$ c'est également un isomorphisme de l'espace sur son dual. On peut voir l'action de $\Lambda_R^{(\alpha)}$ comme créant une perte de régularité de 2α crans, au moins pour les fonctions de $H_f^\alpha(\mathbb{S}_R)$.

Sur le cercle, on peut expliciter l'opérateur en utilisant la synthèse de Fourier. Un calcul analogue à celui effectué précédemment montre que si u et v sont deux polynômes trigonométriques,

$$u(x_0) = \sum_{|p| \leq P} \hat{u}_p e^{ip \frac{x_0}{R}}, \quad v(x_0) = \sum_{|p| \leq P} \hat{v}_p e^{ip \frac{x_0}{R}} \quad (5.71)$$

alors on a explicitement

$$\langle \Lambda_R^{(\alpha)} u, v \rangle = 2\pi R \sum_p \hat{\Lambda}_p^{(\alpha)} \hat{u}_p \overline{\hat{v}_p} = \sum_p \hat{\Lambda}_p^{(\alpha)} (\mathcal{F}(\theta_{\mathbb{S}_R} u))_p \overline{(\mathcal{F}(\theta_{\mathbb{S}_R} v))_p} \quad (5.72)$$

On obtient alors le théorème suivant.

Théorème 3.

Pour tout β réel et α dans $]0, 1[$, l'opérateur $\Lambda_R^{(\alpha)}$ est un isomorphisme de $H_{mo}^\beta(\mathbb{S}_R)$ dans $H_{mo}^{\beta-2\alpha}(\mathbb{S}_R)$; il existe deux constantes ne dépendant que de α telles que

$$\left\| \Lambda_R^{(\alpha)} \right\|_{\mathcal{L}(H_{mo}^\beta(\mathbb{S}_R), H_{mo}^{\beta-2\alpha}(\mathbb{S}_R))} \leq C_\alpha^+ \quad (5.73)$$

$$\left\| (\Lambda_R^{(\alpha)})^{-1} \right\|_{\mathcal{L}(H_{mo}^{\beta-2\alpha}(\mathbb{S}_R), H_{mo}^\beta(\mathbb{S}_R))} \leq \frac{1}{C_\alpha^-} \quad (5.74)$$

De plus, l'opérateur "augmenté" de la moyenne

$$u \mapsto \frac{1}{2\pi R} \int_{\mathbb{S}_R} u(x_0) dx_0 + \Lambda_R^{(\alpha)} u := \Lambda_R^{(\alpha)+mo} u \quad (5.75)$$

est un isomorphisme de $H^\beta(\mathbb{S}_R)$ dans $H^{\beta-2\alpha}(\mathbb{S}_R)$.

Démonstration. La preuve pour (5.73) repose sur une simple inégalité de Cauchy-Schwarz. On a

$$\begin{aligned} \left| \left\langle \Lambda_R^{(\alpha)} u, v \right\rangle \right| &\leq \left(\sum_p |(\mathcal{F}(\theta_{\mathbb{S}_R} u))_p|^2 |\hat{\Lambda}_p^\alpha|^2 (1+p^2)^{-2\alpha+\beta} \right)^{\frac{1}{2}} \\ &\quad \left(\sum_p |(\mathcal{F}(\theta_{\mathbb{S}_R} v))_p|^2 (1+p^2)^{-\beta+2\alpha} \right)^{\frac{1}{2}} \quad (5.76) \\ &\leq C_\alpha^+ \|u\|_{H_f^\beta(\mathbb{S}_R)} \|v\|_{H_f^{-(\beta-2\alpha)}(\mathbb{S}_R)} \end{aligned}$$

où on a utilisé (5.57) ; le résultat s'obtient par dualité

$$\left\| \Lambda_R^{(\alpha)} u \right\|_{H_f^{(\beta-2\alpha)}(\mathbb{S}_R)} = \sup_{v \in H_f^{-(\beta-2\alpha)}(\mathbb{S}_R)} \frac{\left| \left\langle \Lambda_R^{(\alpha)} u, v \right\rangle \right|}{\|v\|_{H_f^{-(\beta-2\alpha)}(\mathbb{S}_R)}} \leq C_\alpha^+ \|u\|_{H_f^\beta(\mathbb{S}_R)} \quad (5.77)$$

Le résultat (5.74) pour l'inverse s'établit exactement de la même manière en utilisant la borne inférieure dans (5.57). Il faut bien sûr utiliser le fait que les fonctions sont à moyennes nulles car $\hat{\Lambda}_0^\alpha = 0$. Si l'on utilise l'opérateur augmenté de la moyenne, les coefficients $\hat{\Lambda}_p^\alpha$ deviennent $\delta_p^0 + \hat{\Lambda}_p^\alpha$ et le produit de $(\delta_p^0 + \hat{\Lambda}_p^\alpha)$ par $(1+p^2)^{-\alpha}$ est alors minoré par $\min(1, C_\alpha^-)$ pour tout p ; l'hypothèse de moyenne nulle n'est plus nécessaire. ■

On obtient alors facilement le corollaire

Corollaire 2.

Soit f une fonction de $L^2(\mathbb{S}_R)$, l'équation

$$f(x_0) = Cu(x_0) + \text{V.P.} \sqrt{R} \int_{\mathbb{S}_R} \frac{u(y_0) - u(x_0)}{|x_0 - y_0|^{\frac{3}{2}}} dy_0 \quad (5.78)$$

où C est une constante positive non nulle quelconque admet une solution unique $u \in H^{\frac{1}{2}}(\mathbb{S}_R)$ et on a pour une certaine constante M ,

$$\|u\|_{H^{\frac{1}{2}}(\mathbb{S}_R)} \leq M \|f\|_{L^2(\mathbb{S}_R)}. \quad (5.79)$$

Démonstration. En effet, l'équation s'écrit

$$f = \left(\Lambda_R^{(\frac{1}{4})+mo} + K \right) u \quad (5.80)$$

avec

$$Ku = Cu - \frac{1}{2\pi R} \int_{\mathbb{S}_R} u(x_0) dx_0 \quad (5.81)$$

Il est clair que K est compact de $H^{\frac{1}{2}}(\mathbb{S}_R)$ dans $L^2(\mathbb{S}_R)$ et comme $\Lambda_R^{(\frac{1}{4})+mo}$ est un isomorphisme de $H^{\frac{1}{2}}(\mathbb{S}_R)$ sur $L^2(\mathbb{S}_R) = H^0(\mathbb{S}_R)$, la théorie de Fredholm s'applique : si $\left(\Lambda_R^{(\frac{1}{4})+mo} + K \right)$ est injectif, il est inversible. Maintenant, si $\left(\Lambda_R^{(\frac{1}{4})+mo} + K \right) u = 0$, on a en multipliant par u l'équation et en intégrant sur \mathbb{S}_R

$$C \int_{\mathbb{S}_R} u^2 dx_0 + \frac{\sqrt{R}}{2} \int_{\mathbb{S}_R} \int_{\mathbb{S}_R} \frac{(u(y_0) - u(x_0))^2}{|x_0 - y_0|^{\frac{3}{2}}} dx_0 dy_0 = 0 \quad (5.82)$$

et donc $u = 0$, le résultat en découle. ■

Sur un cercle, nous avons donc bien le résultat voulu. Nous allons étudier un cas plus général.

5.3.2 CAS D'UNE COURBE $C^{1,r}$

Le but de ce qui va suivre est d'étendre le résultat établi pour le cercle au cas d'une courbe plus générale. Tout d'abord, et par mimétisme, on peut définir les analogues des opérateurs agissant sur des fonctions définies sur le cercle à une courbe Σ . En particulier, on s'intéressera à l'équation

$$f = Cu + \Lambda_{\Sigma}^{(\frac{1}{4})} u \quad (5.83)$$

avec C une constante réelle positive ou plus explicitement

$$f(x) = Cu(x) + \text{V.P.} \sqrt{R} \int_{\Sigma} \frac{u(y) - u(x)}{|x - y|^{\frac{3}{2}}} dy \quad (5.84)$$

où l'inconnue est u et la donnée la fonction f supposée de carré intégrable sur Σ ; cette équation peut également s'écrire variationnellement

$$\int_{\Sigma} f(x)v(x) dx = C \int_{\Sigma} u(x)v(x) dx + \frac{\sqrt{R}}{2} \int_{\Sigma} \int_{\Sigma} \frac{(u(y) - u(x))(v(y) - v(x))}{|x - y|^{\frac{3}{2}}} dx dy \quad (5.85)$$

Si l'on veut étendre le résultat valable sur le cercle, on doit montrer que ce problème admet une solution unique dans $H^{\frac{1}{2}}(\Sigma)$ lorsque $f \in L^2(\Sigma)$. Si l'on utilise une simple inégalité de Cauchy-Schwarz pour faire apparaître la norme de u dans cet espace, on obtient que pour tout v assez régulier

$$\left| \left\langle Cu + \Lambda_{\Sigma}^{(\frac{1}{4})} u, v \right\rangle \right| \leq \left(C \int_{\Sigma} u^2 dx + \frac{R}{2} \int_{\Sigma} \int_{\Sigma} \frac{(u(y) - u(x))^2}{|x - y|^2} dx dy \right)^{\frac{1}{2}} \\ \left(C \int_{\Sigma} v^2 dx + \int_{\Sigma} \int_{\Sigma} \frac{(v(y) - v(x))^2}{|x - y|} dx dy \right)^{\frac{1}{2}} \quad (5.86)$$

Malheureusement le terme portant sur v n'est pas contrôlé par la norme L^2 de v ; si on veut utiliser cette majoration, on est conduit à introduire l'espace des fonctions "0 höldériennes"

$$H_{log}(\Sigma) = \{v \in L^2(\Sigma), \|v\|_{H_{log}(\Sigma)} < \infty\} \quad (5.87)$$

avec

$$\|v\|_{H_{log}(\Sigma)}^2 = C \int_{\Sigma} v^2 dx + \int_{\Sigma} \int_{\Sigma} \frac{(v(y) - v(x))^2}{|x - y|} dx dy \quad (5.88)$$

Cet espace est un peu plus petit que $L^2(\Sigma)$ dans le sens où il contient tous les $H^{\alpha}(\Sigma)$, $\alpha > 0$. La majoration précédente (5.86) montre que

$$\sup_{v \in H_{log}(\Sigma)} \frac{\left| \left\langle Cu + \Lambda_{\Sigma}^{(\frac{1}{4})} u, v \right\rangle \right|}{\|v\|_{H_{log}(\Sigma)}} \leq \|u\|_{H_g^{\frac{1}{2}}(\Sigma)} \quad (5.89)$$

ce qui établit la continuité de l'opérateur $C \text{Id} + \Lambda_{\Sigma}^{(\frac{1}{4})}$ de $H^{\frac{1}{2}}(\Sigma)$ dans le dual de $H_{log}(\Sigma)$ qui est lui un peu plus gros que $L^2(\Sigma)$ (il est contenu dans tous les $H^{-\alpha}(\Sigma)$, $\alpha > 0$). Il n'est donc pas clair, avec ces simples arguments, que l'équation considérée avec second membre $L^2(\Sigma)$ admette une solution de régularité $H^{\frac{1}{2}}(\Sigma)$.

5.3.2.1 TRANSPOSITION DU PROBLÈME SUR LE CERCLE

Si on utilise l'abscisse curviligne, on a de façon équivalente

$$\int_0^{2\pi R} \tilde{f}(s) \tilde{v}(s) ds = C \int_0^{2\pi R} \tilde{u}(s) \tilde{v}(s) ds \\ + \frac{\sqrt{R}}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(\tilde{u}(s+t) - \tilde{u}(s)) (\tilde{v}(s+t) - \tilde{v}(s))}{|x(x+t) - y(s)|^{\frac{3}{2}}} ds dt \quad (5.90)$$

avec $R = \frac{L}{2\pi}$, L la longueur de Σ . Cette dernière forme est intéressante car on peut l'interpréter comme une équation sur le cercle toujours *via* l'identification

$$u(x(s)) = \tilde{u}(s) = u_0(x_0(s)) \quad (5.91)$$

Maintenant, on peut réécrire, l'équation sous la forme

$$\begin{aligned} \int_0^{2\pi R} \tilde{f}(s)\tilde{v}(s) ds &= C \int_0^{2\pi R} \tilde{u}(s)\tilde{v}(s) ds \\ &+ \frac{\sqrt{R}}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(\tilde{u}(s+t) - \tilde{u}(s)) (\tilde{v}(s+t) - \tilde{v}(s))}{|x_0(x+t) - y_0(s)|^{\frac{3}{2}}} ds dt \\ &+ \frac{\sqrt{R}}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} N(s,t) (\tilde{u}(s+t) - \tilde{u}(s)) (\tilde{v}(s+t) - \tilde{v}(s)) ds dt \end{aligned} \quad (5.92)$$

avec

$$N(s,t) = \frac{1}{|x(x+t) - y(s)|^{\frac{3}{2}}} - \frac{1}{|x_0(x+t) - y_0(s)|^{\frac{3}{2}}} \quad (5.93)$$

ce qui s'écrit sous la forme ($f_0(x_0(s)) = \tilde{f}(s) = f(x(s))$)

$$f_0 = (C \text{Id} + \Lambda_R^{(\frac{1}{4})}) u_0 + K_N u_0 \quad (5.94)$$

Si l'on parvient à montrer que K_N est compact de $H^{\frac{1}{2}}(\mathbb{S}_R)$ dans $L^2(\mathbb{S}_R)$, l'opérateur $(C \text{Id} + \Lambda_R^{(\frac{1}{4})}) + K_N$ sera Fredholm (isomorphisme + compact), et cet opérateur sera lui même un isomorphisme s'il est injectif. Mais comme on a la propriété de positivité

$$\left\langle (C \text{Id} + \Lambda_R^{(\frac{1}{4})} + K_N) u_0, u_0 \right\rangle \geq C \|u_0\|_{L^2(\mathbb{S}_R)}^2, \quad (5.95)$$

l'injectivité est évidente. Toujours sous cette hypothèse de K_N compact, l'unique solution $u_0(x_0)$ vérifiera pour une certaine constante M

$$\|u_0\|_{H_g^{\frac{1}{2}}(\mathbb{S}_R)} \leq M \|f_0\|_{L^2(\mathbb{S}_R)} = M \|f\|_{L^2(\Sigma)} \quad (5.96)$$

grâce à la propriété

$$\|f_0\|_{L^2(\mathbb{S}_R)}^2 = \int_0^{2\pi R} \tilde{f}^2(s) ds = \|f\|_{L^2(\Sigma)}^2. \quad (5.97)$$

En posant ensuite

$$u(x(s)) = \tilde{u}(s) = u_0(x_0(s)), \quad (5.98)$$

u ainsi définie vérifiera (5.85). Il faudra ensuite montrer la régularité $H_g^{\frac{1}{2}}(\Sigma)$ de la fonction ainsi construite. Pour cela, il suffit d'établir que pour tout u de $H^{\frac{1}{2}}(\Sigma)$, il existe une constante C_Σ^+ telle que l'on ait

$$\|u\|_{H_g^{\frac{1}{2}}(\Sigma)} \leq C_\Sigma^+ \|u_0\|_{H_g^{\frac{1}{2}}(\mathbb{S}_R)} \quad (5.99)$$

et on aura bien le contrôle :

$$\|u\|_{H_g^{\frac{1}{2}}(\Sigma)} \leq C_\Sigma^+ M \|f\|_{L^2(\Sigma)}. \quad (5.100)$$

Une dernière chose que l'on doit regarder est la continuité de $\Lambda_\Sigma^{(\frac{1}{4})}$ de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$. Si K_N est compact, $A = (C \text{Id} + \Lambda_R^{(\frac{1}{4})}) + K_N$ sera *a fortiori* continu de $H^{\frac{1}{2}}(\mathbb{S}_R)$ dans $L^2(\mathbb{S}_R)$, et on aura

$$\|\Lambda_\Sigma u\|_{L^2(\Sigma)} = \|Au_0\|_{L^2(\mathbb{S}_R)} \leq C_A \|u_0\|_{H^{\frac{1}{2}}(\mathbb{S}_R)}. \quad (5.101)$$

Alors, si on parvient à obtenir l'inégalité

$$\|u\|_{H^{\frac{1}{2}}(\Sigma)} \geq C_\Sigma^- \|u_0\|_{H^{\frac{1}{2}}(\mathbb{S}_R)}, \quad (5.102)$$

on obtiendra la continuité puisque la norme $L^2(\Sigma)$ de $\Lambda_\Sigma u$ vérifiera

$$\|\Lambda_\Sigma u\|_{L^2(\Sigma)} \leq \frac{C_A}{C_\Sigma^-} \|u\|_{H^{\frac{1}{2}}(\Sigma)}. \quad (5.103)$$

5.3.2.2 ISOMORPHISME POUR UNE COURBE HÖLDÉRIENNES

Si l'on synthétise le raisonnement du paragraphe précédent, on a deux propriétés non triviales à établir

- (i) La compacité de K_N de $H^{\frac{1}{2}}(\mathbb{S}_R)$ dans $L^2(\mathbb{S}_R)$
- (ii) le contrôle (5.99) entre normes $H^{\frac{1}{2}}$ lors du passage de Σ vers le cercle, i.e. l'existence de deux constantes telles que

$$C_\Sigma^- \|u_0\|_{H^{\frac{1}{2}}(\mathbb{S}_R)} \leq \|u\|_{H^{\frac{1}{2}}(\Sigma)} \leq C_\Sigma^+ \|u_0\|_{H^{\frac{1}{2}}(\mathbb{S}_R)} \quad (5.104)$$

Si celles-ci sont satisfaites, on obtient l'extension voulue. De fait, ces deux propriétés ne peuvent être vérifiées sans hypothèse de régularité pour Σ . Par exemple, nous pensons qu'il n'y a pas la compacité lorsque Σ a des coins. C'est là que la régularité $C^{1,r}$ de la courbe va venir à notre aide. On a le théorème suivant.

Théorème 4.

Soit Σ une courbe de classe $C^{1,r}$ (i.e. vérifiant (5.1)); soit f une fonction de $L^2(\Sigma)$ et C une constante positive, l'équation

$$f(x) = Cu(x) + \text{V.P.} \sqrt{R} \int_{\Sigma} \frac{u(y) - u(x)}{|x - y|^{\frac{3}{2}}} dy, \quad (5.105)$$

admet une solution unique $u \in H^{\frac{1}{2}}(\Sigma)$ et il existe $M > 0$ telle que,

$$\|u\|_{H^{\frac{1}{2}}(\Sigma)} \leq M \|f\|_{L^2(\Sigma)}. \quad (5.106)$$

Démonstration. D'après ce que nous venons de voir, il suffit d'établir les points (i) et (ii) de l'énumération ci-dessus. Commençons par le point (ii). L'équivalence des normes utilise les propriétés des fonctions singulières que nous avons étudiées auparavant. En effet, prenons $\gamma = -2$ dans l'estimation (5.4) du lemme 9

$$\left| |x(s+t) - x(s)|^{-2} - |x_0(s+t) - x_0(s)|^{-2} \right| \leq C |x_0(s+t) - x_0(s)|^{-2+r} \quad (5.107)$$

avec $C = C_{\Sigma, -2, r}$. En appliquant une inégalité de Cauchy-Schwarz, on trouve d'abord

$$|x(s+t) - x(s)|^{-2} \leq |x_0(s+t) - x_0(s)|^{-2} (1 + C |x_0(s+t) - x_0(s)|^r), \quad (5.108)$$

puis, comme $r > 0$ et $|x_0(s+t) - x_0(s)| \leq 2R$ obtient finalement

$$|x(s+t) - x(s)|^{-2} \leq C' |x_0(s+t) - x_0(s)|^{-2}. \quad (5.109)$$

avec $C' = 1 + (2R)^r$. En découle alors les inégalités

$$\begin{aligned} |u|_{H_g^\alpha(\Sigma)}^2 &= \frac{1}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(\tilde{u}(s+t) - \tilde{u}(s))^2}{|x(s+t) - x(s)|^2} ds dt \\ &\leq C' \frac{1}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(\tilde{u}(s+t) - \tilde{u}(s))^2}{|x_0(s+t) - x_0(s)|^2} ds dt \\ &= C' |u_0|_{H_g^{\frac{1}{2}}(\mathbb{S}_R)}^2 \end{aligned} \quad (5.110)$$

et comme

$$\|u\|_{H_g^{\frac{1}{2}}(\Sigma)}^2 = \|u\|_{L^2(\Sigma)}^2 + R |u|_{H_g^{\frac{1}{2}}(\Sigma)}^2 \quad (5.111)$$

et les normes L^2 se conservant, on obtient

$$\|u\|_{H_g^{\frac{1}{2}}(\Sigma)}^2 \leq C' \left(\|u_0\|_{L^2(\mathbb{S}_R)}^2 + R^2 |u_0|_{H_g^{\frac{1}{2}}(\mathbb{S}_R)}^2 \right) \quad (5.112)$$

soit encore

$$\|u\|_{H_g^{\frac{1}{2}}(\Sigma)} \leq C'^{\frac{1}{2}} \|u_0\|_{H_g^{\frac{1}{2}}(\mathbb{S}_R)}. \quad (5.113)$$

Regardons maintenant l'inégalité inverse. On a également

$$\begin{aligned} |u|_{H_g^\alpha(\Sigma)}^2 &= \frac{1}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(\tilde{u}(s+t) - \tilde{u}(s))^2}{|x(s+t) - x(s)|^2} ds dt \\ &\geq \frac{1}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(\tilde{u}(s+t) - \tilde{u}(s))^2}{|x_0(s+t) - x_0(s)|^2} ds dt \\ &\quad - \frac{C}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(\tilde{u}(s+t) - \tilde{u}(s))^2}{|x_0(s+t) - x_0(s)|^{2-r}} ds dt \\ &\geq \frac{1}{4} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(\tilde{u}(s+t) - \tilde{u}(s))^2}{|x_0(s+t) - x_0(s)|^2} ds dt \\ &\quad - 2\pi RMC \int_0^{2\pi R} u^2(x) dx \end{aligned} \quad (5.114)$$

avec

$$M = \sup_{s \in [0, 2\pi], |t| \in [t^*, \pi]} |x_0(s+t) - x_0(s)|^{-2} \quad (5.115)$$

où t^* est défini par

$$CR^r |2 - 2 \cos t^*|^{\frac{r}{2}} = \frac{1}{4} \quad (5.116)$$

de telle sorte que $\|u\|_{H_g^\alpha(\Sigma)}^2 \geq C' \|u_0\|_{H_g^\alpha(\mathbb{S}_R)}^2$, avec une constante $0 < C' < \frac{1}{2}$ et le point (ii) est établi.

Venons en au point (i). Quitte à diminuer r , on peut toujours supposer $r < \frac{1}{2}$ car une courbe $C^{1,r}$ est une courbe $C^{1,r'}$ si $r' < r$. La compacité va découler de la continuité de K_N de $H^{\frac{1}{2}-\frac{r}{2}}(\mathbb{S}_R)$ dans $L^2(\mathbb{S}_R)$; comme l'injection $H^{\frac{1}{2}}(\mathbb{S}_R) \mapsto H^{\frac{1}{2}-\frac{r}{2}}(\mathbb{S}_R)$ est compacte, K_N est la composée d'un opérateur compact et d'un opérateur continu, il est donc compact. Pour cela, on se donne un u assez régulier par exemple $C^1(\Sigma)$ de telle sorte que

$$K_N(u) \in L^2(\Sigma). \quad (5.117)$$

Puis on prend un polynôme trigonométrique v ,

$$\tilde{v}(s) = \sum_{|p| \leq P} \hat{v}_p e^{ipsR^{-1}} \quad (5.118)$$

et on forme

$$\langle K_N u, v \rangle = \frac{\sqrt{R}}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} N(s, t) (\tilde{u}(s+t) - \tilde{u}(s)) (\bar{\tilde{v}}(s+t) - \bar{\tilde{v}}(s)) ds dt \quad (5.119)$$

avec

$$N(s, t) = |x(s+t) - x(s)|^{-\frac{3}{2}} - |x_0(s+t) - x_0(s)|^{-\frac{3}{2}} \quad (5.120)$$

On a (avec cette fois $C = C_{\Sigma, -\frac{3}{2}, r}$),

$$|N(s, t)| \leq C |x_0(s+t) - x_0(s)|^{-\frac{3}{2}+r} \quad (5.121)$$

et donc

$$\begin{aligned} |\langle K_N u, v \rangle| &\leq \frac{C\sqrt{R}}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{|\tilde{u}(s+t) - \tilde{u}(s)| |\bar{\tilde{v}}(s+t) - \bar{\tilde{v}}(s)|}{|x_0(s+t) - x_0(s)|^{1+\frac{1}{2}-r}} ds dt \\ &\leq C \left(\frac{R^{1-r}}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{(\tilde{u}(s+t) - \tilde{u}(s))^2}{|x_0(s+t) - x_0(s)|^{1+2(\frac{1}{2}-\frac{r}{2})}} ds dt \right)^{\frac{1}{2}} \\ &\quad \left(\frac{R^r}{2} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{|\tilde{v}(s+t) - \tilde{v}(s)|^2}{|x_0(s+t) - x_0(s)|^{1-2\frac{r}{2}}} ds dt \right)^{\frac{1}{2}} \end{aligned} \quad (5.122)$$

Notons que pour appliquer l'inégalité de Cauchy-Schwarz dans l'expression précédente, nous avons distribué la puissance r en mettant un peu de poids sur l'intégrale sur v pour ne pas retomber sur la semi-norme de $H_{\log}(\mathbb{S}_R)$ et le reste sur l'intégrale en u pour obtenir une norme un peu plus faible que celle de $H^{\frac{1}{2}}(\mathbb{S}_R)$.

On a donc

$$|\langle K_N u, v \rangle| \leq CR^r \|u\|_{H_g^{\frac{1}{2}-\frac{r}{2}}(\mathbb{S}_R)} N_r(v) \quad (5.123)$$

avec

$$N_r^2(v) = R^{-r} \int_0^{2\pi R} \int_{-\pi R}^{\pi R} \frac{|\tilde{v}(s+t) - \tilde{v}(s)|^2}{|x_0(s+t) - x_0(s)|^{1-r}} ds dt \quad (5.124)$$

Le calcul de $N_r^2(v)$ peut s'effectuer à l'aide de la transformation de Fourier

$$N_r^2(v) = 2\pi R \sum_p |\hat{v}_p|^2 \hat{\Lambda}_p^{(-\frac{r}{2})} \quad (5.125)$$

avec

$$\begin{aligned}
\hat{\Lambda}_p^{(-\frac{r}{2})} &= \int_{-\pi}^{\pi} \frac{1 - \cos p\theta}{(2 - 2\cos\theta)^{\frac{1}{2}-\frac{r}{2}}} d\theta \\
&= 2\cos\left(\frac{\pi r}{2}\right) \Gamma(r) \left(\frac{\Gamma(\frac{1}{2}-\frac{r}{2})}{\Gamma(\frac{1}{2}+\frac{r}{2})} - \frac{\Gamma(|p|-\frac{1}{2}-\frac{r}{2})}{\Gamma(|p|+\frac{1}{2}+\frac{r}{2})} \right) \\
&\leq 2\Gamma(r)
\end{aligned} \tag{5.126}$$

et donc

$$N_r(v) \leq \sqrt{2\Gamma(r)} \|v\|_{L^2(\mathbb{S}_R)}. \tag{5.127}$$

Nous en déduisons que si V est l'espace des polynômes trigonométriques

$$\sup_{v \in V} \frac{|\langle K_N u, v \rangle|}{\|v\|_{L^2(\mathbb{S}_R)}} \leq CR^r \sqrt{2\Gamma(r)} \|u\|_{H_g^{\frac{1}{2}-\frac{r}{2}}(\mathbb{S}_R)} \tag{5.128}$$

mais comme V est dense dans $L^2(\mathbb{S}_R)$, on a

$$\|K_N u\|_{L^2(\mathbb{S}_R)} \leq CR^r \sqrt{2\Gamma(r)} \|u\|_{H_g^{\frac{1}{2}-\frac{r}{2}}(\mathbb{S}_R)}. \tag{5.129}$$

Ce résultat, vrai pour u régulière, s'étend encore par un argument de densité aux fonctions u de $H_g^{\frac{1}{2}-\frac{r}{2}}(\mathbb{S}_R)$ et on a la continuité voulue. ■

Grâce au théorème précédent, on a directement le corollaire

Corollaire 3.

L'opérateur $\Lambda^{(2)}$ défini en (5.64) est un isomorphisme continu de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$ et l'opérateur $T_r^{(2)} = \Lambda^{(2),*} \Lambda^{(2)}$ vérifie l'hypothèse 1.

5.4 TRONCATURE

Les opérateurs que nous avons utilisés, définis par (5.62) pour T et par (5.64) pour Λ sont complètement non locaux. Comme pour les opérateurs définis par les potentiels de Riesz du chapitre précédent, il est nécessaire d'introduire une troncature pour réduire le coût qu'ils engendrent (lors de leur assemblage et à

la résolution des problèmes locaux). Nous écrivons les opérateurs tronqués de la manière suivante

$$\langle T_\delta u, v \rangle_\Sigma = \int_\Sigma u(x) \bar{v}(x) dx + \beta \int_\Sigma \int_\Sigma \chi_\delta(|x-y|) \frac{(u(x) - u(y))(\bar{v}(x) - \bar{v}(y))}{|x-y|^2} dx dy. \quad (5.130)$$

et

$$\langle \Lambda_\delta u, v \rangle_\Sigma = \int_\Sigma u(x) \bar{v}(x) dx + \beta \int_\Sigma \int_\Sigma \chi_\delta(|x-y|) \frac{(u(x) - u(y))(\bar{v}(x) - \bar{v}(y))}{|x-y|^{\frac{3}{2}}} dx dy. \quad (5.131)$$

où la fonction de troncature χ est définie de la même manière que dans le chapitre précédent par (4.40)

$$\forall r, 0 \leq \chi(r) \leq 1, \quad \chi(r) = 1 \text{ pour } r \leq r_0, \quad \text{et } \chi(r) = 0 \text{ pour } r \geq r_1. \quad (5.132)$$

Nous avons alors le théorème suivant

Théorème 5 (Opérateurs tronqués).

Les opérateurs tronqués définis ci-dessus conservent les mêmes propriétés que les opérateurs non tronqués, à savoir T_δ est positif symétrique et un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $H^{-\frac{1}{2}}(\Sigma)$, et Λ_δ est symétrique et un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$.

Démonstration. Nous présenterons la preuve pour T_δ . La preuve pour Λ_δ est presque identique à celle pour T_δ ainsi qu'à celle de l'opérateur Λ défini grâce aux opérateurs de Riesz (voir section 4.4).

Comme $\chi_\delta(r)$ est positif pour tout r et $\beta > 0$, la positivité de T_δ est évidente, de même que sa symétrie. il reste à montrer que T_δ est l'isomorphisme voulu. Écrivons

$$T_\delta = \left(I + (T_\delta - T_r^{(1)}) T_r^{(1),-1} \right) T_r^{(1)}. \quad (5.133)$$

Alors, comme $T_r^{(1)}$ est un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $H^{-\frac{1}{2}}(\Sigma)$, il suffit de montrer que l'opérateur $I + (T_\delta - T_r^{(1)}) T_r^{(1),-1}$ est un isomorphisme de $H^{-\frac{1}{2}}(\Sigma)$ dans lui-même. L'injectivité de cet opérateur découle directement de l'injectivité de $T_r^{(1)}$. En effet, on a

$$\|\varphi\|_{L^2(\Sigma)} \leq \langle T_\delta \varphi, \varphi \rangle \leq \langle T_r^{(1)} \varphi, \varphi \rangle, \quad (5.134)$$

donc T_δ est injectif et $I + (T_\delta - T_r^{(1)}) T_r^{(1),-1}$ aussi. Pour conclure *via* l'alternative de Fredholm, il nous suffit de montrer que l'opérateur $(T_\delta - T_r^{(1)}) T_r^{(1),-1}$ est un opérateur compact dans $H^{-\frac{1}{2}}(\Sigma)$. Comme la fonction

$$(x, y) \mapsto (1 - \chi_\delta(|x - y|)) |x - y|^{-\frac{3}{2}} \quad (5.135)$$

est bornée sur $\sigma \times \Sigma$ (soit $C(\delta)$ cette borne), on a

$$|\langle (T_\delta - T_r^{(1)}) \varphi, \psi \rangle| \leq C(\delta) L \|\varphi\|_{L^2(\Sigma)} \|\psi\|_{L^2(\Sigma)}, \quad (5.136)$$

où L est la longueur de Σ . Ceci montre que l'opérateur $T_\delta - T_r^{(1)}$ est continu dans $L^2(\Sigma)$ et

$$\forall \varphi \in H^{\frac{1}{2}}(\Sigma), \quad \|(T_\delta - T_r^{(1)}) \varphi\|_{L^2(\Sigma)} \leq C(\delta) L \|\varphi\|_{L^2(\Sigma)}. \quad (5.137)$$

Ensuite, prenons une suite bornée ψ_n dans $H^{-\frac{1}{2}}(\Sigma)$. Alors $\varphi_n := T_r^{(1),-1} \psi_n$ est bornée dans $H^{\frac{1}{2}}(\Sigma)$. Par compacité de l'inclusion de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$, et à l'extraction d'une sous-suite près, φ_n converge fortement dans $L^2(\Sigma)$. Comme $T_\delta - T_r^{(1)}$ est continu dans $L^2(\Sigma)$, $(T_\delta - T_r^{(1)}) \varphi_n$ converge aussi fortement dans $L^2(\Sigma)$, et *a fortiori* dans $H^{-\frac{1}{2}}(\Sigma)$. Ceci montre la compacité de l'opérateur $(T_\delta - T_r^{(1)}) T_r^{(1),-1}$ dans $H^{-\frac{1}{2}}(\Sigma)$ et conclut la preuve. ■

5.5 EXTENSION 3D

L'extension des opérateurs définis par les semi-normes de Gagliardo au cas 3D est plus aisée que pour les opérateurs basés sur les potentiels de Riesz. En effet, les semi-normes sont définies de la même manière pour toutes les dimensions. Notons d la dimension de l'interface Σ ($d = 1$ en 2D et $d = 2$ en 3D). Ainsi, l'opérateur T défini par

$$\langle T u, v \rangle_\Sigma = \int_\Sigma u(x) \bar{v}(x) dx + \beta \int_\Sigma \int_\Sigma \frac{(u(x) - u(y))(\bar{v}(x) - \bar{v}(y))}{|x - y|^{d+1}} dx dy \quad (5.138)$$

est immédiatement un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $H^{-\frac{1}{2}}(\Sigma)$ car cette forme bilinéaire définit une norme dans $H^{\frac{1}{2}}(\Sigma)$ (par définition des semi-normes de Gagliardo). De même, l'opérateur Λ défini par

$$\langle \Lambda u, v \rangle_\Sigma = \int_\Sigma u(x) \bar{v}(x) dx + \beta \int_\Sigma \int_\Sigma \frac{(u(x) - u(y))(\bar{v}(x) - \bar{v}(y))}{|x - y|^{d+\frac{1}{2}}} dx dy \quad (5.139)$$

est un isomorphisme de $H^{\frac{1}{4}}(\Sigma)$ dans $H^{-\frac{1}{4}}(\Sigma)$. Malheureusement, nous n'avons pas réussi à étendre notre démonstration de cas 2D pour montrer qu'il s'agissait également d'un isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$ dans le cas 3D.

5.6 CONCLUSIONS

Ainsi, plusieurs possibilités sont offertes pour choisir l'opérateur T (ou Λ). D'un point de vue théorique, T donné par l'opérateur norme $H^{\frac{1}{2}}(\Sigma)$ (5.62) permet de s'affranchir de l'opérateur intermédiaire Λ et de donner immédiatement la propriété d'isomorphisme entre $H^{\frac{1}{2}}(\Sigma)$ et $H^{-\frac{1}{2}}(\Sigma)$ tout en restant positif (même après troncature) et cela pour des géométries d'interfaces très générales (Lipschitz par morceaux par exemple). L'inconvénient apparaît lors de l'implémentation d'une version discrétisée. En effet, le calcul des matrices élémentaires du type

$$\iint_{K \times L} \frac{(\varphi_i(x) - \varphi_i(y))(\varphi_j(x) - \varphi_j(y))}{|x - y|^{d+a}}, \quad a = 1 \quad (5.140)$$

où K et L sont une paire d'éléments du maillage de Σ et φ_i et φ_j les fonctions de base, conduit à évaluer numériquement des intégrales double en 2D et quadruple en 3D avec des intégrandes qui sont L^∞ mais discontinues (en certains points en 2D, et selon des arêtes en 3D), ce qui n'est pas chose facile. Dans le cas des éléments finis d'ordre peu élevé (P1 par exemple), il est néanmoins possible de s'appuyer sur les techniques d'intégrations de fonctions homogènes [36], pour les éléments non disjoints ; mais cette technique est vite très laborieuse en 3D ou pour des éléments d'ordre élevé en 2D. Le même problème d'intégration numérique se pose pour l'opérateur Λ donné par (5.64) (prendre $a = \frac{1}{2}$ dans (5.140)), toutefois l'intégrande est cette fois continue, *a priori* plus facile à traiter. Toutes ces difficultés qui, soulignons-le, ne sont pas insurmontables, sont liées au caractère singulier des intégrales. Définies de manière forte, elles correspondent à des valeurs principales, qui ne sont pas aisées à manipuler.

Au contraire, les opérateurs intégraux définis grâce aux potentiels de Riesz (4.9) font intervenir des intégrandes moins singulières (voire même continues dans le cas 2D). L'assemblage pour ces opérateurs est alors grandement facilité et peut être réalisé traditionnellement, élément de frontière par élément de frontière, même s'ils font intervenir les dérivées des fonctions de base.

C'est pourquoi dans la suite (partie numérique notamment), nous utiliserons les opérateurs définis grâce aux potentiels de Riesz.

Troisième partie
Aspects numériques

6

Discrétisation

Sommaire

6.1	Espaces des solutions	130
6.2	Discrétisation du problème global	131
6.3	Problèmes locaux, Construction par T	132
6.3.1	Formulation faible et unicité de la solution	132
6.3.2	Opérateurs discrets de <i>scattering</i>	135
6.3.3	Processus itératif	136
6.3.4	À la convergence	139
6.3.5	Forme matricielle du problème	140
6.4	Problèmes locaux, Construction par Λ	142
6.4.1	Formulation faible et unicité de la solution	142
6.4.2	Opérateurs discrets de <i>scattering</i>	146
6.4.3	Processus itératif	147
6.4.4	À la convergence	150
6.4.5	Forme matricielle du problème	151

Ce chapitre est dédié à la construction et à l'analyse de schémas numériques pour la décomposition de domaine. Pour cela, nous présentons ici une discrétisation de type *éléments finis*. Dans toute la suite du document, nous considérerons un maillage possédant les propriétés suivantes

- Les maillages de Ω_1 et de Ω_2 sont conformes à l'interface Σ .
- Le maillage de Ω correspond à l'union des maillages de Ω_1 et Ω_2

Par simplification, et sans perte de généralité, nous supposons les maillages comme étant des triangulations \mathcal{T}_h des domaines considérés, h étant une taille caractéristique des éléments du maillage. Toutes les quantités définies sur ce maillage seront notées avec l'indice h , notamment $\mathcal{T}_h(\Omega)$ la triangulation du domaine Ω , u_h la solution globale sur $\mathcal{T}_h(\Omega)$, $u_{h,j}$, $j = 1, 2$ la solution restreinte à $\mathcal{T}_h(\Omega_j)$, etc.

6.1 ESPACES DES SOLUTIONS

La solution u du problème global (2.1) est recherchée dans l'espace

$$V = H_{\Gamma_0}^1(\Omega) = \left\{ u \in H^1(\Omega) \text{ tel que } u|_{\Gamma_0} = 0 \right\}. \quad (6.1)$$

De même, à chaque itération n , les solutions locales u_1^n et u_2^n appartiennent respectivement aux espaces $V_1 = H_{\Gamma_0}^1(\Omega_1)$ et $V_2 = H^1(\Omega_2)$. Introduisons les espaces d'approximation

$$\begin{aligned} V_h &= \left\{ v_h \in V \text{ tel que } \forall K \in \mathcal{T}_h(\Omega), \quad v_h|_K \in P_l(K) \right\}, \\ V_{h,1} &= \left\{ v_h \in V_1 \text{ tel que } \forall K \in \mathcal{T}_h(\Omega_1), \quad v_h|_K \in P_l(K) \right\}, \\ V_{h,2} &= \left\{ v_h \in V_2 \text{ tel que } \forall K \in \mathcal{T}_h(\Omega_2), \quad v_h|_K \in P_l(K) \right\}. \end{aligned} \quad (6.2)$$

où $P_l(K)$ est l'espace des fonctions polynomiales d'ordre $l \geq 1$ sur l'élément K de $\mathcal{T}_h(\Omega)$. Il est clair que $V_h \subset V$. Notons que les éléments de $V_{h,j}$ sont les restrictions à Ω_j des éléments de V_h , $j = 1, 2$. Les espaces $V_{h,1}$ et $V_{h,2}$ seront donc les espaces d'approximation pour les problèmes locaux. Enfin, il nous sera utile d'introduire l'espace $V_{h,\Sigma}$

$$V_{h,\Sigma} = \left\{ \phi_h \in L^2(\Sigma) \text{ tel que } \forall K \in \mathcal{T}_h(\Sigma), \quad \phi_h|_K \in P_l(K) \right\}. \quad (6.3)$$

Nous avons ici choisi de prendre des éléments finis d'ordre identique dans la définition de $V_{h,\Sigma}$ et dans celle de $V_{h,1}$ et $V_{h,2}$. Ceci implique, grâce à la conformité

des maillages sur Σ , que

$$V_{h,\Sigma} = \{ \phi_h \in L^2(\Sigma); \exists v_{h,i} \in V_{h,i} \text{ et } \phi_h = v_{h,i}|_{\Sigma} \}. \quad (6.4)$$

Dit plus simplement, $V_{h,\Sigma}$ est l'espace commun des traces de $V_{h,1}$ et $V_{h,2}$. Dans la suite, nous choisirons cet espace pour approcher indifféremment des fonctions de $H^{\frac{1}{2}}(\Sigma)$ ou de $L^2(\Sigma)$.

6.2 DISCRÉTISATION DU PROBLÈME GLOBAL

Le problème (2.1) traité ici s'écrit variationnellement de la manière suivante

$$\begin{aligned} & \text{Trouver } u \in V = \mathcal{H}_{\Gamma_0}^1(\Omega) \text{ tel que } \forall v \in V, \\ & a(u, v) = \int_{\Omega} (\mu^{-1} \nabla u \nabla \bar{v} - k^2 \epsilon u \bar{v}) dx + ik \int_{\Gamma} \eta^{-1} u \bar{v} d\sigma = \int_{\Omega} f \bar{v} dx \end{aligned} \quad (6.5)$$

La solution approchée u_h dans V_h de u vérifie pour tout $v_h \in V_h$

$$\int_{\Omega} (\mu^{-1} \nabla u_h \nabla \bar{v}_h - k^2 \epsilon u_h \bar{v}_h) dx + ik \int_{\Gamma} \eta^{-1} u_h \bar{v}_h d\sigma = \int_{\Omega} f \bar{v}_h dx. \quad (6.6)$$

On notera n_h le nombre de degrés de liberté associés à cette discrétisation ($\mathcal{T}_h(\Omega)$ et éléments d'ordre l); les fonctions de base associées à ces degrés de liberté seront notées $(\Psi_{h,i})_{i=1..n_h}$. Dans la suite, pour ne pas alourdir les notations, nous oublierons régulièrement l'indice h des fonctions de bases : $\Psi_{h,i} = \Psi_i$. Les composantes de u_h dans cette base seront notées \mathbf{u}_h , et celles de la projection de f dans V_h seront notées \mathbf{f}_h . Alors, il est facile de voir que l'équation précédente se réduit au système linéaire

$$\mathbb{K}_h \mathbf{u}_h = \mathbf{f}_h, \quad (6.7)$$

où \mathbb{K}_h est la matrice de taille $n_h \times n_h$ donnée par

$$(\mathbb{K}_h)_{i,j} = \int_{\Omega} (\mu^{-1} \nabla \Psi_i \nabla \Psi_j - k^2 \epsilon \Psi_i \Psi_j) dx + ik \int_{\Gamma} \eta^{-1} \Psi_i \Psi_j d\sigma. \quad (6.8)$$

Intéressons-nous au caractère bien posé du problème global discret. Remarquons tout d'abord que (6.6) est équivalent à : trouver $u_h \in V_h$ tel que pour tout $v_h \in V_h$, $a(u_h, v_h) = \int_{\Omega} f \bar{v}_h$. On écrit

$$a(u, v) = a_*(u, v) - \int_{\Omega} 2k^2 \epsilon u \bar{v} dx = a_*(u, v) + a_c(u, v). \quad (6.9)$$

La forme bilinéaire a_* est de manière évidente coercive sur V , tandis que a_c est compacte (l'injection de $L^2(\Omega)$ dans V est compacte). Les résultats de [7] et notamment les proposition 1.38 et théorème 1.40 (pages 29-30) montrent que l'approximation par éléments finis des problèmes de type coercif + compact est bien posée uniquement pour des maillages suffisamment fins. On a la proposition suivante.

Proposition 11 (Caractère bien posé du problème global discret).

Il existe $h_0 > 0$ et $C > 0$ tel que pour tout $h < h_0$ le problème discret (6.6) (ou de manière équivalente la matrice \mathbb{K}) est inversible et

$$\|u_h - u\|_V \leq C \inf_{v_h \in V_h} \|u - v_h\|_V \quad (6.10)$$

Nous allons maintenant présenter deux discrétisations possibles, selon que l'opérateur T est construit directement à partir d'une forme bilinéaire (comme par exemple dans le cas des semi-normes de Gagliardo (5.62)), ou si l'opérateur T est construit à partir d'une forme bilinéaire pour Λ (comme par exemple pour une construction basée sur les potentiels de Riesz (4.9)).

6.3 PROBLÈMES LOCAUX, CONSTRUCTION PAR T

Nous ferons ici l'hypothèse que l'opérateur T est construit directement.

6.3.1 FORMULATION FAIBLE ET UNICITÉ DE LA SOLUTION

Les formulations faibles de ces problèmes locaux (2.43) et (2.44) sont, pour le sous-domaine Ω_1

Trouver $u_1 \in V_1$ tel que $\forall v_1 \in V_1$,

$$\int_{\Omega_1} (\mu^{-1} \nabla u_1 \nabla \bar{v}_1 - k^2 \epsilon u_1 \bar{v}_1) dx + ik \langle T u_1, v_1 \rangle_{\Sigma} = \int_{\Omega_1} f \bar{v}_1 dx + \langle g_1, v_1 \rangle_{\Sigma}, \quad (6.11)$$

et pour le sous-domaine Ω_2

Trouver $u_2 \in V_2$ tel que $\forall v_2 \in V_2$,

$$\begin{aligned} \int_{\Omega_2} (\mu^{-1} \nabla u_2 \nabla \bar{v}_2 - k^2 \epsilon u_2 \bar{v}_2) dx + ik \int_{\Gamma} \eta^{-1} u_2 \bar{v}_2 d\sigma + ik \langle T^* u_2, v_2 \rangle_{\Sigma} & \quad (6.12) \\ & = \int_{\Omega_2} f \bar{v}_2 dx + \langle g_2, v_2 \rangle_{\Sigma}. \end{aligned}$$

Alors, en utilisant les espaces de discrétisation $V_{h,1}$ et $V_{h,2}$ introduits précédemment, les problèmes discrets approchant les problèmes (6.11) et (6.12) s'écrivent faiblement de la manière suivante pour le sous-domaine Ω_1

$$\left\{ \begin{array}{l} \text{Trouver } u_{h,1} \in V_{h,1} \text{ tel que } \forall v_{h,1} \in V_{h,1}, \\ \int_{\Omega_1} (\mu^{-1} \nabla u_{h,1} \nabla \bar{v}_{h,1} - k^2 \epsilon u_{h,1} \bar{v}_{h,1}) dx + ik \langle T u_{h,1}, v_{h,1} \rangle_{\Sigma} \\ \qquad \qquad \qquad = \int_{\Omega_1} f \bar{v}_{h,1} dx + \langle g_1, v_{h,1} \rangle_{\Sigma}, \end{array} \right. \quad (6.13)$$

et pour le sous-domaine Ω_2

$$\left\{ \begin{array}{l} \text{Trouver } u_{h,2} \in V_{h,2} \text{ tel que } \forall v_{h,2} \in V_{h,2}, \\ \int_{\Omega_2} (\mu^{-1} \nabla u_{h,2} \nabla \bar{v}_{h,2} - k^2 \epsilon u_{h,2} \bar{v}_{h,2}) dx + ik \int_{\Gamma} \eta^{-1} u_{h,2} \bar{v}_{h,2} d\sigma \\ \qquad \qquad \qquad + ik \langle T^* u_{h,2}, v_{h,2} \rangle_{\Sigma} = \int_{\Omega_2} f \bar{v}_{h,2} dx + \langle g_2, v_{h,2} \rangle_{\Sigma}, \end{array} \right. \quad (6.14)$$

Nous noterons les formes bilinéaires associées à chacun des problèmes locaux a_1 et a_2 et leur forme linéaire du second membre l_1 et l_2 , pour tout $(u_{h,j}, v_{h,j}) \in$

$V_{h,j} \times V_{h,j}$, tout $f \in H^1(\Omega)$ et $(g_1, g_2) \in H^{-\frac{1}{2}}(\Sigma)$

$$\begin{aligned}
a_1(u_{h,1}, v_{h,1}) &= \int_{\Omega_1} (\mu^{-1} \nabla u_{h,1} \nabla \bar{v}_{h,1} - k^2 \epsilon u_{h,1} \bar{v}_{h,1}) dx + ik \langle T u_{h,1}, v_{h,1} \rangle_{\Sigma} \\
l_1(f, g_1, v_{h,1}) &= \int_{\Omega_1} f \bar{v}_{h,1} dx + \langle g_1, v_{h,1} \rangle_{\Sigma} \\
a_2(u_{h,2}, v_{h,2}) &= \int_{\Omega_2} (\mu^{-1} \nabla u_{h,2} \nabla \bar{v}_{h,2} - k^2 \epsilon u_{h,2} \bar{v}_{h,2}) dx + ik \int_{\Gamma} \eta^{-1} u_{h,2} \bar{v}_{h,2} d\sigma \\
&\quad + ik \langle T^* u_{h,2}, v_{h,2} \rangle_{\Sigma} \\
l_2(f, g_2, v_{h,2}) &= \int_{\Omega_2} f \bar{v}_{h,2} dx + \langle g_2, v_{h,2} \rangle_{\Sigma}.
\end{aligned} \tag{6.15}$$

Nous avons alors

Lemme 11 (Caractère bien posé des problèmes locaux discrets).

Il existe $H > 0$ tel que pour tout $h < H$, $f \in H^1(\Omega)$ et $(g_1, g_2) \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$, les problèmes (6.13) et (6.14) admettent une unique solution $u_{h,1}$ dans $V_{h,1}$ et $u_{h,2}$ dans $V_{h,2}$.

Démonstration. La démonstration sera faite uniquement pour le problème (6.14), la preuve pour (6.13) étant pratiquement identique. Puisque le problème est de dimension finie, il suffit de montrer l'unicité de la solution. On cherche alors à montrer l'implication suivante

$$\text{Pour } u_{h,2} \in V_{h,2}, \quad a_2(u_{h,2}, v_{h,2}) = 0 \quad \forall v_{h,2} \in V_{h,2} \quad \Rightarrow \quad u_{h,2} = 0. \tag{6.16}$$

Nous avons vu dans la partie de théorie générale (chapitre 2), que la forme bilinéaire a_2 pouvait s'écrire sous la forme d'une somme d'une forme bilinéaire coercive a_* et d'une partie compacte a_c (voir (2.49)). Nous pouvons donc appliquer une nouvelle fois les résultats de [7], et il existe un $h_2 > 0$ tel que pour tout $h < h_2$ le problème qui consiste à trouver $u_{h,2} \in V_{h,2}$ tel que pour tout $v_{h,2} \in V_{h,2}$, $a_2(u_{h,2}, v_{h,2}) = 0$ soit bien posé. Ceci montre que pour tout $h < h_2$, $u_{h,2} = 0$.

La même démonstration pour le problème posé dans le sous-domaine 1 conduirait à l'existence d'un $h_1 > 0$ tel que pour tout $h < h_1$ le problème soit bien posé également. Pour conclure, il suffit alors de prendre $H = \min(h_1, h_2)$. ■

6.3.2 OPÉRATEURS DISCRETS DE *scattering*

Dans la suite, et même si cela n'est pas précisé, nous supposons que $h < H$. Puisque les problèmes locaux discrets (6.13) et (6.14) sont bien posés, et comme dans le cas continu, ils définissent implicitement quatre opérateurs $E_{h,j}$ et $R_{h,j}$, $j = 1, 2$. Les opérateurs $E_{h,j}$ sont définis comme opérateurs de $H^{-\frac{1}{2}}(\Sigma)$ dans $V_{h,j}$ pour tout $g_j \in H^{-\frac{1}{2}}(\Sigma)$ par

$$e_{h,j} = E_{h,j} g_j \Leftrightarrow \forall v_{h,j} \in V_{h,j}, a_j(e_{h,j}, v_{h,j}) = l_j(0, g_j, v_{h,j}). \quad (6.17)$$

Les opérateurs $R_{h,j}$ sont définis comme opérateurs de $H^1(\Omega_j)$ dans $V_{h,j}$ pour tout $f \in H^1(\Omega_j)$ par ($j = 1, 2$)

$$r_{h,j} = R_{h,j} f \Leftrightarrow \forall v_{h,j} \in V_{h,j}, a_j(r_{h,j}, v_{h,j}) = l_j(f, 0, v_{h,j}). \quad (6.18)$$

Avec ces définitions, les solutions $u_{h,1}$ et $u_{h,2}$ des problèmes locaux s'écrivent

$$u_{h,1} = E_{h,1} g_1 + R_{h,1} f, \quad u_{h,2} = E_{h,2} g_2 + R_{h,2} f. \quad (6.19)$$

Nous allons maintenant écrire le processus itératif discret de la décomposition de domaine. Dans le cas continu, le processus itératif est écrit sur les traces entrantes et sortantes (voir la proposition 3). Nous allons faire de même pour le processus discret. Cela est rendu possible grâce à la conformité des maillages à l'interface Σ , mais aussi parce que l'espace de discrétisation $V_{h,\Sigma}$ des espaces $H^{\frac{1}{2}}(\Sigma)$, $L^2(\Sigma)$ et $H^{-\frac{1}{2}}(\Sigma)$ correspond effectivement à l'espace des traces des fonctions de $V_{h,1}$ et de $V_{h,2}$. Comme nous l'avons fait pour le processus itératif du problème continu, il nous est alors possible de définir pour des traces entrantes $g_{h,j}^{in} \in V_{h,\Sigma}$ les traces sortantes associées $g_{h,j}^{out} \in V_{h,\Sigma}$ grâce aux opérateurs de *scattering discrets* $S_h = (S_{h,1}, S_{h,2})$ pour tout $\phi_h \in V_{h,\Sigma}$ par ($v_{h,j}$ sont telles que $v_{h,j}|_{\Sigma} = \phi_h$, pour $j = 1, 2$)

$$\begin{aligned} \forall \phi_h \in V_{h,\Sigma}, \quad \langle g_{h,j}^{out}, \phi_h \rangle_{\Sigma} &= \langle S_{h,j} g_{h,j}^{in}, \phi_h \rangle_{\Sigma} \\ &= -\langle g_{h,j}^{in}, \phi_h \rangle_{\Sigma} + 2ik \langle T_r E_{h,j} g_{h,j}^{in}, v_{h,j} \rangle_{\Sigma}. \end{aligned} \quad (6.20)$$

Enfin, nous définissons l'opérateur d'échange de traces discret Π_h de $V_{h,\Sigma} \times V_{h,\Sigma}$ dans $V_{h,\Sigma} \times V_{h,\Sigma}$ par

$$\Pi_h(g_{h,1}, g_{h,2}) = (\Pi_{h,21} g_{h,2}, \Pi_{h,12} g_{h,1}), \quad (6.21)$$

avec pour tout $\phi_h \in V_{h,\Sigma}$

$$\begin{aligned} \langle \Pi_{h,12} g_{h,1}, \phi_h \rangle_{\Sigma} &= \langle g_{h,1}, \phi_h \rangle_{\Sigma} \\ \langle \Pi_{h,21} g_{h,2}, \phi_h \rangle_{\Sigma} &= \langle g_{h,2}, \phi_h \rangle_{\Sigma}. \end{aligned} \quad (6.22)$$

Π_h n'est autre que la restriction de l'opérateur d'échange continu Π à l'espace d'approximation $V_{h,\Sigma} \times V_{h,\Sigma}$.

6.3.3 PROCESSUS ITÉRATIF

Nous pouvons alors écrire le processus itératif pour les traces entrantes $g_{h,1}^{in,n}$ et $g_{h,2}^{in,n}$ de $V_{h,\Sigma}$

$$\begin{aligned} \langle g_{h,1}^{in,n+1}, \phi_h \rangle_\Sigma &= (1-r) \langle g_{h,1}^{in,n}, \phi_h \rangle_\Sigma + r \left(\langle -g_{h,2}^{in,n}, \phi_h \rangle_\Sigma + 2ik \langle T_r u_{h,2}^n, v_{h,2} \rangle_\Sigma \right) \\ &= (1-r) \langle g_{h,1}^{in,n}, \phi_h \rangle_\Sigma + r \langle S_{h,2} g_{h,2}^{in,n}, \phi_h \rangle_\Sigma + 2r ik \langle T_r R_{h,2} f, v_{h,2} \rangle_\Sigma \end{aligned} \quad (6.23)$$

$$\begin{aligned} \langle g_{h,2}^{in,n+1}, \phi_h \rangle_\Sigma &= (1-r) \langle g_{h,2}^{in,n}, \phi_h \rangle_\Sigma + r \left(\langle -g_{h,1}^{in,n}, \phi_h \rangle_\Sigma + 2ik \langle T_r u_{h,1}^n, v_{h,1} \rangle_\Sigma \right) \\ &= (1-r) \langle g_{h,2}^{in,n}, \phi_h \rangle_\Sigma + r \langle S_{h,1} g_{h,1}^{in,n}, \phi_h \rangle_\Sigma + 2r ik \langle T_r R_{h,1} f, v_{h,1} \rangle_\Sigma \end{aligned} \quad (6.24)$$

Pour étudier la convergence du processus itératif discret, nous prendrons pour la suite de cette section $f = 0$, c'est à dire que nous nous intéresserons à l'erreur sur les traces. Rappelons en effet que la convergence de l'erreur $(\varepsilon_{h,1}^n, \varepsilon_{h,2}^n)$ sur les traces $(g_{h,1}^{in,n}, g_{h,2}^{in,n})$ vers 0 implique la convergence de l'erreur sur les solutions $(u_{h,1}^n, u_{h,2}^n)$ grâce à la relation $u_{h,j}^n = E_{h,j} g_{h,j}^{in,n} + R_{h,j} f$ où les opérateurs $E_{h,j}$ et $R_{h,j}$ sont donnés par (6.17) et (6.18). Les erreurs $(\varepsilon_{h,1}^n, \varepsilon_{h,2}^n)$ vérifient les relations

$$\begin{aligned} \langle \varepsilon_{h,1}^{n+1}, \phi_h \rangle_\Sigma &= (1-r) \langle \varepsilon_{h,1}^n, \phi_h \rangle_\Sigma + r \langle S_{h,2} \varepsilon_{h,2}^n, \phi_h \rangle_\Sigma \\ \langle \varepsilon_{h,2}^{n+1}, \phi_h \rangle_\Sigma &= (1-r) \langle \varepsilon_{h,2}^n, \phi_h \rangle_\Sigma + r \langle S_{h,1} \varepsilon_{h,1}^n, \phi_h \rangle_\Sigma \end{aligned} \quad (6.25)$$

et on est ramené, de manière similaire au cas continu, à l'étude de l'opérateur $I - S_{h,\pi} = I - \Pi_h S_h$.

Lemme 12 (Opérateurs discrets).

Pour tout $h < H$, l'opérateur $I - S_{h,\pi}$ est un isomorphisme de $V_{h,\Sigma}^2$ dans $V_{h,\Sigma}^2$ et l'opérateur $S_{h,\pi}$ est contractant dans $V_{h,\Sigma}^2$ pour la norme $\|\cdot\|_X$ (2.88-2.89).

Démonstration. Les preuves de ce lemme s'inspirent fortement des preuves des lemmes 2 et 3 concernant les opérateurs continus S_π et $I - S_\pi$. Puisque $V_{h,\Sigma}$ est de dimension finie, seule l'injectivité de $I - S_{h,\pi}$ doit être montrée pour obtenir l'isomorphisme. Prenons $g_h = (g_{h,1}, g_{h,2}) \in V_{h,\Sigma}^2$ tel que $(I - S_{h,\pi})g_h = 0$ et notons $e_h = (e_{h,1}, e_{h,2}) \in V_{h,1} \times V_{h,2}$ définis par $e_{h,j} = E_{h,j}g_{h,j}$. Par définition des opérateurs $S_{h,j}$ (6.20), pour tout $\phi_h \in V_{h,\Sigma}$ on a

$$\begin{aligned} \langle g_{h,1} - S_{h,2}g_{h,2}, \phi_h \rangle_\Sigma &= \langle g_{h,1} + g_{h,2} - 2ik T_r e_{h,2}, \phi_h \rangle_\Sigma = 0 \\ \langle g_{h,2} - S_{h,1}g_{h,1}, \phi_h \rangle_\Sigma &= \langle g_{h,1} + g_{h,2} - 2ik T_r e_{h,1}, \phi_h \rangle_\Sigma = 0 \end{aligned} \quad (6.26)$$

En soustrayant ces deux équations, on obtient que pour tout $\phi_h \in V_{h,\Sigma}$

$$\langle T_r(e_{h,1} - e_{h,2}), \phi_h \rangle_\Sigma = 0 \quad (6.27)$$

et on en déduit, grâce à la positivité de T_r que $e_{h,1}|_\Sigma = e_{h,2}|_\Sigma$. Puisque les solutions $e_{h,1}$ et $e_{h,2}$ se recollent à l'interface, nous pouvons définir une solution sur le domaine entier e_h par $e_{h|\Omega_1} = e_{h,1}$ et $e_{h|\Omega_2} = e_{h,2}$. De plus, pour tout $v_h \in V_h$, notons $v_{h,1} = v_{h|\Omega_1} \in V_{h,1}$, $v_{h,2} = v_{h|\Omega_2} \in V_{h,2}$ et $\phi_h \in V_{h,\Sigma}$ la trace commune de $v_{h,1}$ et $v_{h,2}$ sur Σ . Alors, puisque $e_{h,j} = E_{h,j}g_{h,j}$, $j = 1, 2$, on a

$$a_1(e_{h,1}, v_{h,1}) + a_2(e_{h,1}, v_{h,1}) = l_1(0, g_{h,1}, v_{h,1}) + l_2(0, g_{h,2}, v_{h,2}) \quad (6.28)$$

ou encore, en sommant les intégrales sur Ω_1 et Ω_2 ,

$$a(e_h, v_h) + 2ik \langle T_r e_h, v_h \rangle_\Sigma = \langle g_{h,1} + g_{h,2}, \phi_h \rangle_\Sigma \quad (6.29)$$

En utilisant le fait que $\langle g_{h,1} + g_{h,2}, \phi_h \rangle_\Sigma = 2ik \langle T_r e_{h,j}, \phi_h \rangle_\Sigma$ pour tout ϕ_h , on trouve que $a(e_h, v_h) = 0$ et donc $e_h = 0$ dans tout le domaine Ω . Ceci conduit en particulier à $g_{h,1} = g_{h,2} = 0$. D'où l'isomorphisme.

Pour montrer que $S_{h,\pi}$ est une contraction sur $V_{h,\Sigma}^2$ et comme Π_h est unitaire, nous nous intéresserons aux opérateurs $S_{h,1}$ et $S_{h,2}$. Comme pour le problème continu, le résultat sera prouvé pour $S_{h,2}$, la démonstration étant pratiquement identique pour $S_{h,1}$. Pour tout $g_h \in V_{h,\Sigma}$, nous allons montrer que $\|S_{h,2}g_h\| \leq \|g_h\|$. Notons $e_{h,2} = E_{h,2}g_h$. Nous avons

$$\|S_{h,2}g_h\|^2 = \|-g_h + 2ik T_r e_{h,2}\|^2 = \|g_h\|^2 + 4k^2 \|T_r e_{h,2}\|^2 - 4k \mathcal{I}m(g_h, T_r e_{h,2}). \quad (6.30)$$

et il vient alors

$$\|S_{h,2}g_h\|_{-\frac{1}{2}}^2 - \|g_h\|_{-\frac{1}{2}}^2 = 4k^2 (T_r e_{h,2}, T_r e_{h,2})_{-\frac{1}{2}} - 4k \mathcal{I}m(g_h, T_r e_{h,2})_{-\frac{1}{2}}, \quad (6.31)$$

puis par définition du produit scalaire $(\cdot, \cdot)_{-\frac{1}{2}}$ (2.88)

$$\|S_{h,2} g_h\|_{-\frac{1}{2}}^2 - \|g_h\|_{-\frac{1}{2}}^2 = 4k^2 \langle T_r e_{h,2}, e_{h,2} \rangle_\Sigma - 4k \mathcal{I}m \langle g_h, e_{h,2} \rangle_\Sigma. \quad (6.32)$$

En reprenant la preuve du lemme continu, et en utilisant l'équation vérifiée par $e_{h,2}$, il vient

$$\mathcal{I}m \langle g_h, e_{h,2} \rangle_\Sigma = k \int_\Gamma \eta^{-1} |e_{h,2}|^2 d\sigma + k \langle T_r e_{h,2}, e_{h,2} \rangle_\Sigma. \quad (6.33)$$

Finalement

$$\|S_{h,2} g_h\|_{-\frac{1}{2}}^2 - \|g_h\|_{-\frac{1}{2}}^2 = -4k^2 \int_\Gamma \eta^{-1} |e_{h,2}|^2 d\sigma \leq 0. \quad (6.34)$$

■

À partir de ce lemme, il est aisé d'obtenir le théorème suivant.

Théorème 6 (Convergence de la DDM discrétisée – Opérateur T).

Pour tout $h < H$ et pour tout $r \in]0, 1[$, il existe une constante $C_h > 0$ et un réel $\tau_{h,r} \in]0, 1[$ tel que

$$\|u_{h,1}^n - u_{h,1}\|_{V_{h,1}} + \|u_{h,2}^n - u_{h,2}\|_{V_{h,2}} < C_h \tau_{h,r}^n, \quad (6.35)$$

et la méthode de décomposition de domaine discrétisée converge.

Démonstration. En utilisant les identités remarquables et les inégalités de la preuve du théorème 1 du cas continu, il est facile de montrer que l'erreur $\varepsilon_h^n = (\varepsilon_{h,1}^n, \varepsilon_{h,2}^n)$ sur les traces entrantes vérifie

$$\|\varepsilon_h^n\|^2 \leq \|\varepsilon_h^{n-1}\|^2 - r(1-r) \|(I - S_{h,\pi}) \varepsilon_h^{n-1}\|^2. \quad (6.36)$$

Puisque $(I - S_{h,\pi})$ est un isomorphisme pour tout $h < H$, il existe une constante δ_h telle que

$$\forall g \in V_{h,\Sigma} \times V_{h,\Sigma}, \quad \|(I - S_{h,\pi})g\| \geq \delta_h \|g\| \quad (6.37)$$

Il est une nouvelle fois facile de conclure avec $\tau_h = \sqrt{1 - r(1-r)} \delta_h^2$ et $C_h = \max(\|E_{h,1}\|, \|E_{h,2}\|)$. ■

Remarque 3. Bien que nous intuitions ce résultat, nous ne sommes malheureusement pas parvenu à montrer l'indépendance de la convergence en fonction du pas du maillage h , c'est à dire qu'il existerait un réel $\tau \in]0, 1[$ tel que pour tout $h < H$, $\tau_h < \tau$.

6.3.4 À LA CONVERGENCE

Nous avons vu que la manière de discrétiser la méthode que nous avons développée permet d'obtenir très facilement des propriétés de convergence analogues au cas non discrétisé. Une autre propriété très intéressante peut être déduite concernant la solution de la méthode à la convergence. Notons $u_{h,1}^\infty$ et $u_{h,2}^\infty$ les limites, lorsque $n \rightarrow \infty$ de $u_{h,1}^n$ et $u_{h,2}^n$. Nous avons alors le lemme suivant.

Lemme 13 (Propriété de recollement).

La solution obtenue à la convergence de la méthode de décomposition de domaine correspond exactement à la solution du problème discret non décomposé, c'est à dire

$$u_{h,1}^\infty = u_h|_{\Omega_1}, \quad \text{et} \quad u_{h,2}^\infty = u_h|_{\Omega_2} \quad (6.38)$$

Démonstration. Notons $g_{h,1}^{in,\infty}$ et $g_{h,2}^{in,\infty}$ les traces entrantes des solutions $u_{h,1}^\infty$ et $u_{h,2}^\infty$. Par définition, nous avons

$$u_{h,1}^\infty = E_{h,1} g_{h,1}^{in,\infty} + R_{h,1} f, \quad u_{h,2}^\infty = E_{h,2} g_{h,2}^{in,\infty} + R_{h,2} f. \quad (6.39)$$

De plus, les équations (6.23) et (6.24) décrivant le processus itératif sur les traces entrantes conduisent aux égalités suivantes, pour tout $\phi_h \in V_{h,\sigma}$

$$\begin{aligned} \langle g_{h,1}^{in,\infty} + g_{h,2}^{in,\infty}, \phi_h \rangle_\Sigma &= 2ik \langle T_r u_{h,1}^\infty, \phi_h \rangle_\Sigma \\ \langle g_{h,1}^{in,\infty} + g_{h,2}^{in,\infty}, \phi_h \rangle_\Sigma &= 2ik \langle T_r u_{h,2}^\infty, \phi_h \rangle_\Sigma \end{aligned} \quad (6.40)$$

et donc, en utilisant une nouvelle fois la positivité de T_r , on en déduit que $u_{h,1}^\infty|_{\Sigma} = u_{h,2}^\infty|_{\Sigma}$. Puisque $u_{h,1}^\infty$ et $u_{h,2}^\infty$ sont continues à l'interface Σ (et les maillages conformes), définissons $u_h^\infty \in V_h$ telle que

$$u_{h|\Omega_1}^\infty = u_{h,1}^\infty \quad \text{et} \quad u_{h|\Omega_2}^\infty = u_{h,2}^\infty. \quad (6.41)$$

De plus pour tout $v_h \in V_h$, nous noterons $v_{h,1} = v_h|_{\Omega_1} \in V_{h,1}$, $v_{h,2} = v_h|_{\Omega_2} \in V_{h,2}$ et $\phi_h = v_{h,1}|_\Sigma = v_{h,2}|_\Sigma$. Alors, pour tout $v_h \in V_h$, on a

$$a_1(u_{h,1}^\infty, v_{h,1}) = l_1(f, g_{h,1}^{in,\infty}, v_{h,1}), \quad \text{et} \quad a_2(u_{h,2}^\infty, v_{h,2}) = l_2(f, g_{h,2}^{in,\infty}, v_{h,2}) \quad (6.42)$$

En sommant ces deux équations, et en sommant les intégrales sur Ω_1 et Ω_2 , on obtient

$$a(u_h^\infty, v_h) + 2ik \langle T_r u_h^\infty, v_h \rangle_\Sigma = \int_\Omega f \bar{v}_h dx + \langle g_{h,1}^{in,\infty} + g_{h,2}^{in,\infty}, v_h \rangle_\Sigma. \quad (6.43)$$

Finalement, en utilisant (6.40), on trouve directement

$$\forall v_h \in V_h, \quad a(u_h^\infty, v_h) = \int_\Omega f \bar{v}_h dx \quad (6.44)$$

et u_h^∞ , construit en recollant les solutions $u_{h,1}^\infty$ et $u_{h,2}^\infty$ des problèmes locaux à la convergence, correspond en fait à la solution du problème global discret non décomposé. ■

6.3.5 FORME MATRICIELLE DU PROBLÈME

On notera $n_{h,j}$ le nombre de degrés de liberté associés à la discrétisation ($\mathcal{T}(\Omega_j)$ et éléments d'ordre l), $j = 1, 2$; les fonctions de base associées à ces degrés de liberté seront notées $(\Psi_{h,j,i})_{i=1..n_{h,j}}$ pour $j = 1, 2$. Pour alléger les notations, l'indice h des fonctions de base sera régulièrement oublié. Les composantes de $u_{h,j}^n$ dans cette base seront notées $\mathbf{u}_{h,j}^n$.

Enfin, les fonctions de base de $V_{h,\Sigma}$ (associées à la discrétisation de l'interface Σ) seront notées $(\Phi_{h,\Sigma,i})_{i=1..n_{h,\Sigma}}$, où $n_{h,\Sigma}$ est le nombre de degrés de liberté pour $V_{h,\Sigma}$. Les composantes des traces entrantes $g_{h,j}^{in,n}$ (respectivement sortantes $g_{h,j}^{out,n}$) dans cette base seront notées $\mathbf{g}_{h,j}^{in,n}$ (respectivement $\mathbf{g}_{h,j}^{out,n}$).

Il nous sera utile d'introduire les matrices \mathbb{B}_j discrétisant les opérateurs $V_{h,j} \ni u_{h,j} \mapsto u_{h,j}|_\Sigma \in V_{h,\Sigma}$, $j = 1, 2$. Puisque $V_{h,\Sigma}$ est exactement l'espace des traces sur Σ des espaces $V_{h,j}$, les matrices \mathbb{B}_j sont des matrices de taille $n_{h,\Sigma} \times n_{h,j}$ ($j = 1, 2$), creuses et telles que

$$(\mathbb{B}_j)_{k,i} = 1 \Leftrightarrow \Psi_{h,j,i}|_\Sigma = \Phi_{h,\Sigma,k}, \quad j = 1, 2. \quad (6.45)$$

\mathbb{B}_j sont simplement des matrices de sélection de degrés de liberté. Alors, à chaque itération de l'algorithme de Jacobi, les problèmes locaux discrétisés s'écrivent

$$\mathbb{K}_{h,1} \mathbf{u}_{h,1}^n = \mathbf{f}_1 + \mathbb{B}_1^T \mathbf{g}_{h,1}^{in,n}, \quad \mathbb{K}_{h,2} \mathbf{u}_{h,2}^n = \mathbf{f}_2 + \mathbb{B}_2^T \mathbf{g}_{h,2}^{in,n} \quad (6.46)$$

où les matrices $\mathbb{K}_{h,1}$ et $\mathbb{K}_{h,2}$, de taille respective $n_{h,1} \times n_{h,1}$ et $n_{h,2} \times n_{h,2}$, sont données par

$$\begin{aligned} (\mathbb{K}_{h,1})_{i,j} &= \int_{\Omega_1} (\mu^{-1} \nabla \Psi_{1,i} \nabla \bar{\Psi}_{1,j} - k^2 \epsilon \Psi_{1,i} \bar{\Psi}_{1,j}) dx + ik \langle T \Psi_{1,i}, \Psi_{1,j} \rangle_{\Sigma} \\ (\mathbb{K}_{h,2})_{i,j} &= \int_{\Omega_2} (\mu^{-1} \nabla \Psi_{2,i} \nabla \bar{\Psi}_{2,j} - k^2 \epsilon \Psi_{2,i} \bar{\Psi}_{2,j}) dx + ik \int_{\Gamma} \eta^{-1} \Psi_{2,i} \bar{\Psi}_{2,j} d\sigma \\ &\quad + ik \langle T^* \Psi_{2,i}, \Psi_{2,j} \rangle_{\Sigma} \end{aligned} \quad (6.47)$$

et les seconds membres \mathbf{f}_j correspondent aux composantes de $f|_{\Omega_j}$ dans $V_{h,j}$. Le lemme 11 nous indique que les matrices $\mathbb{K}_{h,j}$ sont inversibles.

Remarque 4 (Notes sur l'opérateur T). *Dans (6.47), le produit scalaire $\langle \cdot, \cdot \rangle_{\Sigma}$ est un produit de dualité entre $H^{\frac{1}{2}}(\Sigma)$ et son dual $H^{-\frac{1}{2}}(\Sigma)$. Les notations peuvent être trompeuses ici, et il est intéressant de rappeler que le produit scalaire agit sur des traces. En toute rigueur, il faudrait écrire (pour le sous-domaine Ω_1)*

$$\left\langle T \Psi_{1,i|\Sigma}, \Psi_{1,j|\Sigma} \right\rangle_{\Sigma} \quad (6.48)$$

et si les fonctions de bases $\Psi_{1,i}$, $\Psi_{1,j}$ sont non nulles sur Σ , il existe des indices l_1, k_1 tels que

$$\Psi_{1,i|\Sigma} = \Phi_{\Sigma,l_1}, \quad \Psi_{1,j|\Sigma} = \Phi_{\Sigma,k_1}. \quad (6.49)$$

Les matrices $\mathbb{K}_{h,j}$ peuvent alors s'écrire

$$\mathbb{K}_{h,1} = \tilde{\mathbb{K}}_{h,1} + ik \mathbb{B}_1^T \mathbb{T} \mathbb{B}_1, \quad \mathbb{K}_{h,2} = \tilde{\mathbb{K}}_{h,2} + ik \mathbb{B}_2^T \mathbb{T}^* \mathbb{B}_2 \quad (6.50)$$

avec les matrices $\tilde{\mathbb{K}}_{h,1}$, $\tilde{\mathbb{K}}_{h,2}$ et \mathbb{T} données par

$$\begin{aligned} (\tilde{\mathbb{K}}_{h,1})_{i,j} &= \int_{\Omega_1} (\mu^{-1} \nabla \Psi_{1,i} \nabla \bar{\Psi}_{1,j} - k^2 \epsilon \Psi_{1,i} \bar{\Psi}_{1,j}) dx \\ (\tilde{\mathbb{K}}_{h,2})_{i,j} &= \int_{\Omega_2} (\mu^{-1} \nabla \Psi_{2,i} \nabla \bar{\Psi}_{2,j} - k^2 \epsilon \Psi_{2,i} \bar{\Psi}_{2,j}) dx + ik \int_{\Gamma} \eta^{-1} \Psi_{2,i} \bar{\Psi}_{2,j} d\sigma \\ \mathbb{T}_{k,l} &= \langle T \Phi_{\Sigma,k}, \Phi_{\Sigma,l} \rangle_{\Sigma} \end{aligned} \quad (6.51)$$

La matrice \mathbb{T} , de taille $n_{h,\Sigma} \times n_{h,\Sigma}$ est particulièrement importante pour le processus itératif, comme nous allons le voir.

Nous l'avons vu précédemment, les traces entrantes et sortantes dans chacun des sous-domaines sont liées par la relation

$$\forall \phi_h \in V_{h,\Sigma}, \quad \langle g_{h,j}^{out,n}, \phi_h \rangle_\Sigma = -\langle g_{h,j}^{in,n}, \phi_h \rangle_\Sigma + 2ik \langle T_r u_{h,j}^n, \phi_h \rangle_\Sigma \quad (6.52)$$

où $u_{h,j}^n$ est solution de l'équation $a_j(u_{h,j}^n, v_{h,j}) = l_j(f, g_{h,j}^{in,n}, v_{h,j})$ pour tout $v_{h,j} \in V_{h,j}$ (6.15). En notant \mathbb{T}_r la matrice $(\mathbb{T} + \mathbb{T}^*)/2$, l'équation (6.52) s'écrit matriciellement

$$\mathbf{g}_{h,j}^{out,n} = -\mathbf{g}_{h,j}^{in,n} + 2ik \mathbb{T}_r \mathbb{B}_j \mathbf{u}_{h,j}^n, \quad \text{avec} \quad \mathbb{K}_{h,j} \mathbf{u}_{h,j}^n = \mathbf{f}_j + \mathbb{B}_j^T \mathbf{g}_{h,j}^{in,n}. \quad (6.53)$$

En éliminant les inconnues $\mathbf{u}_{h,j}^n$, on obtient finalement

$$\mathbf{g}_{h,j}^{out,n} = \underbrace{\left(-\mathbb{I}_\Sigma + 2ik \mathbb{T}_r \mathbb{B}_j \mathbb{K}_{h,j}^{-1} \mathbb{B}_j^T \right)}_{\mathbb{S}_{h,j}} \mathbf{g}_{h,j}^{in,n} + \underbrace{2ik \mathbb{T}_r \mathbb{B}_j \mathbb{K}_{h,j}^{-1} \mathbf{f}_j}_{\mathbf{F}_j} \quad (6.54)$$

où la matrice \mathbb{I}_Σ est la matrice identité de taille $n_{h,\Sigma} \times n_{h,\Sigma}$. Les matrices \mathbb{S}_j , $j = 1, 2$, définies dans (6.54) correspondent aux opérateurs de *scattering* discrets $\mathbb{S}_{h,j}$ définis en (6.20). Enfin, la formule de récurrence vérifiée par les traces entrantes à l'itération $n + 1$ s'écrit, grâce au paramètre de relaxation $r \in]0, 1[$

$$\begin{aligned} \mathbf{g}_{h,1}^{in,n+1} &= (1-r) \mathbf{g}_{h,1}^{in,n} + r \mathbf{g}_{h,2}^{out,n} \\ \mathbf{g}_{h,2}^{in,n+1} &= (1-r) \mathbf{g}_{h,2}^{in,n} + r \mathbf{g}_{h,1}^{out,n} \end{aligned} \quad (6.55)$$

6.4 PROBLÈMES LOCAUX, CONSTRUCTION PAR Λ

Nous présentons dans cette section une deuxième discrétisation possible, lorsque l'opérateur Λ nous est donné explicitement. Nous allons montrer qu'avec cette discrétisation et sous une hypothèse additionnelle, toutes les propriétés sont conservées, notamment le caractère bien posé des problèmes locaux discrets et la convergence de l'algorithme.

6.4.1 FORMULATION FAIBLE ET UNICITÉ DE LA SOLUTION

Nous introduisons des variables intermédiaires qui sont des fonctions de $L^2(\Sigma)$, définies par

$$\varphi_1 = \Lambda u_{1|\Sigma} \text{ et } \varphi_2 = \Lambda u_{2|\Sigma}. \quad (6.56)$$

Variationnellement, ces égalités s'écrivent $((x, y) = \int_{\Sigma} x \bar{y})$

$$\begin{aligned} (\Lambda u_1, \phi_1) - (\varphi_1, \phi_1) &= 0, \quad \forall \phi_1 \in L^2(\Sigma) \\ (\Lambda u_2, \phi_2) - (\varphi_2, \phi_2) &= 0, \quad \forall \phi_2 \in L^2(\Sigma). \end{aligned} \quad (6.57)$$

Dans les formulations faibles des problèmes locaux, on peut alors remplacer les produits scalaires faisant intervenir l'opérateur T par

$$\begin{aligned} \langle T u_1, v_1 \rangle_{\Sigma} &= z \langle \Lambda^* \varphi_1, v_1 \rangle_{\Sigma} \\ \langle T^* u_2, v_2 \rangle_{\Sigma} &= \bar{z} \langle \Lambda^* \varphi_2, v_2 \rangle_{\Sigma} \end{aligned} \quad (6.58)$$

Alors, en utilisant les espaces de discrétisation $V_{h,1}$, $V_{h,2}$ et $V_{h,\Sigma}$ introduits précédemment, les problèmes discrets approchant les problèmes (6.11) et (6.12) s'écrivent faiblement de la manière suivante pour le sous-domaine Ω_1

$$\left\{ \begin{array}{l} \text{Trouver } (u_{h,1}, \varphi_{h,1}) \in V_{h,1} \times V_{h,\Sigma} \text{ tel que } \forall (v_{h,1}, \phi_{h,1}) \in V_{h,1} \times V_{h,\Sigma}, \\ \int_{\Omega_1} (\mu^{-1} \nabla u_{h,1} \nabla \bar{v}_{h,1} - k^2 \epsilon u_{h,1} \bar{v}_{h,1}) dx + ik z \langle \Lambda^* \varphi_{h,1}, v_{h,1} \rangle_{\Sigma} \\ \qquad \qquad \qquad = \int_{\Omega_1} f \bar{v}_{h,1} dx + \langle g_1, v_{h,1} \rangle_{\Sigma}, \\ \text{et } (\Lambda u_{h,1}, \phi_{h,1}) - (\varphi_{h,1}, \phi_{h,1}) = 0 \end{array} \right. \quad (6.59)$$

et pour le sous-domaine Ω_2

$$\left\{ \begin{array}{l} \text{Trouver } (u_{h,2}, \varphi_{h,2}) \in V_{h,2} \times V_{h,\Sigma} \text{ tel que } \forall (v_{h,2}, \phi_{h,2}) \in V_{h,2} \times V_{h,\Sigma}, \\ \int_{\Omega_2} (\mu^{-1} \nabla u_{h,2} \nabla \bar{v}_{h,2} - k^2 \epsilon u_{h,2} \bar{v}_{h,2}) dx + ik \int_{\Gamma} \eta^{-1} u_{h,2} \bar{v}_{h,2} d\sigma \\ \qquad \qquad \qquad + ik \bar{z} \langle \Lambda^* \varphi_{h,2}, v_{h,2} \rangle_{\Sigma} = \int_{\Omega_2} f \bar{v}_{h,2} dx + \langle g_2, v_{h,2} \rangle_{\Sigma}, \\ \text{et } (\Lambda u_{h,2}, \phi_{h,2}) - (\varphi_{h,2}, \phi_{h,2}) = 0. \end{array} \right. \quad (6.60)$$

Pour simplifier les notations, notons les espaces

$$\tilde{V}_{h,1} = V_{h,1} \times V_{h,\Sigma}, \text{ et } \tilde{V}_{h,2} = V_{h,2} \times V_{h,\Sigma}. \quad (6.61)$$

De plus, définissons les formes suivantes, pour tout $(u_{h,j}, \varphi_{h,j}) \in \tilde{V}_{h,j}$ et tout $v_{h,j} \in V_{h,j}$, $j = 1, 2$

$$\begin{aligned}
\tilde{a}_1(u_{h,1}, \varphi_{h,1}, v_{h,1}) &= \int_{\Omega_1} (\mu^{-1} \nabla u_{h,1} \nabla \bar{v}_{h,1} - k^2 \epsilon u_{h,1} \bar{v}_{h,1}) dx \\
&\quad + ik z \langle \Lambda^* \varphi_{h,1}, v_{h,1} \rangle_{\Sigma} \\
\tilde{b}_1(u_{h,1}, \varphi_{h,1}, \phi_{h,1}) &= (\Lambda u_{h,1}, \phi_{h,1}) - (\varphi_{h,1}, \phi_{h,1}) \\
\tilde{a}_2(u_{h,2}, \varphi_{h,2}, v_{h,2}) &= \int_{\Omega_2} (\mu^{-1} \nabla u_{h,2} \nabla \bar{v}_{h,2} - k^2 \epsilon u_{h,2} \bar{v}_{h,2}) dx \\
&\quad + ik \int_{\Gamma} \eta^{-1} u_{h,2} \bar{v}_{h,2} d\sigma + ik \bar{z} \langle \Lambda^* \varphi_{h,2}, v_{h,2} \rangle_{\Sigma} \\
\tilde{b}_2(u_{h,2}, \varphi_{h,2}, \phi_{h,2}) &= (\Lambda u_{h,2}, \phi_{h,2}) - (\varphi_{h,2}, \phi_{h,2})
\end{aligned} \tag{6.62}$$

Les seconds membres l_1 et l_2 restent identiques au cas précédent et sont donnés par (6.15). Enfin, nous supposerons dans la suite de la section que l'hypothèse suivante est vérifiée.

Hypothèse 2 (Injectivité faible de Λ).

Pour tout élément de $x \in V_{h,\Sigma}$, alors, l'implication suivante est vraie

$$\forall y \in V_{h,\Sigma}, (\Lambda x, y) = 0 \Rightarrow x = 0 \tag{6.63}$$

Remarque 5. Notons que cette hypothèse est relativement faible et est vérifiée pour les opérateurs Λ que nous avons construits dans la Partie 2.

Nous avons alors le lemme suivant.

Lemme 14 (Caractère bien posé des problèmes locaux discrets).

Il existe $H > 0$ tel que pour tout $h < H$, $f \in H^1(\Omega)$ et $(g_1, g_2) \in H^{-\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$, les problèmes (6.59) et (6.60) admettent une unique solution $u_{h,1}$ dans $V_{h,1}$ et $u_{h,2}$ dans $V_{h,2}$.

Démonstration. Encore une fois, la preuve sera faite uniquement pour le problème (6.60) puisque la preuve pour (6.59) est pratiquement identique. De même, le

problème étant de dimension finie, seule l'unicité de la solution est requise. On cherche alors à montrer, pour $(u_{h,2}, \varphi_{h,2}) \in \tilde{V}_{h,2}$

$$\begin{cases} \tilde{a}_2(u_{h,2}, \varphi_{h,2}, v_{h,2}) = 0, \quad \forall v_{h,2} \in V_{h,2} \\ (\Lambda u_{h,2}, \phi_{h,2}) - (\varphi_{h,2}, \phi_{h,2}) = 0, \quad \forall \phi_{h,2} \in V_{h,\Sigma} \end{cases} \Rightarrow \begin{cases} u_{h,2} = 0 \\ \varphi_{h,2} = 0 \end{cases} \quad (6.64)$$

Malheureusement, nous n'avons pas réussi à écrire directement ce problème sous la forme coercif + compact, comme nous avons pu le faire à plusieurs occasions précédemment. Nous allons utiliser une astuce de calcul. Tout d'abord, appliquons les équations de (6.64) avec $v_{h,2} = u_{h,2}$ et $\phi_{h,2} = \varphi_{h,2}$. La seconde équation nous donne

$$(\Lambda u_{h,2}, \varphi_{h,2}) = (\varphi_{h,2}, \varphi_{h,2}) = (u_{h,2}, \Lambda^* \varphi_{h,2}). \quad (6.65)$$

De plus, en prenant la partie imaginaire de $\tilde{a}_2(u_{h,2}, \varphi_{h,2}, u_{h,2}) = 0$ on obtient

$$k \int_{\Gamma} \eta^{-1} |u_{h,2}|^2 d\sigma + k \operatorname{Re}(z) (\Lambda^* \varphi_{h,2}, u_{h,2}) = 0. \quad (6.66)$$

En utilisant (6.65) dans (6.66) (et $\operatorname{Re}(z) > 0$), on obtient $u_{h,2} = 0$ sur Γ et $\varphi_{h,2} = 0$. Puisque $\varphi_{h,2} = 0$, alors la seconde équation de (6.64) donne

$$(\Lambda u_{h,2}, \phi_{h,2}) = 0, \quad \forall \phi_{h,2} \in V_{h,\Sigma} \quad (6.67)$$

et d'après l'hypothèse 2, on a $u_{h,2} = 0$ sur Σ . Nous pouvons alors utiliser l'astuce de calcul suivante. Puisque $u_{h,2} = 0$ sur Σ et $\varphi_{h,2} = 0$, alors $u_{h,2}$ est également solution du problème

$$a_2(u_{h,2}, v_{h,2}) = 0 \quad (6.68)$$

où la forme bilinéaire a_2 est définie par (6.15) et correspond au problème local discret dans Ω_2 où l'opérateur T est donné directement par sa forme bilinéaire. Alors, en utilisant le lemme 11 concernant le caractère bien posé des problèmes locaux discrets, il existe H tel que pour tout $h < H$, le problème donné par la forme bilinéaire a_2 est bien posé. Donc pour tout $h < H$, $u_{h,2} = 0$. Ceci conclut la preuve. \blacksquare

Encore une fois, nous supposons dans la suite que $h < H$, de telle sorte que les problèmes locaux discrets sont bien posés.

6.4.2 OPÉRATEURS DISCRETS DE *scattering*

Nous allons définir ici les opérateurs de relèvement et de scattering que nous avons eu l'occasion d'introduire sur le problème continu et dans le cas des problèmes discrets définis avec l'opérateur T . Commençons par les opérateurs $E_{h,j}$ et $R_{h,j}$, $j = 1, 2$. Puisque les problèmes locaux discrets (6.59) et (6.60) sont bien posés, ils définissent une nouvelle fois quatre opérateurs. Les opérateurs $E_{h,j}$, $j = 1, 2$, sont définis comme opérateurs de $H^{-\frac{1}{2}}(\Sigma)$ dans $\tilde{V}_{h,j}$ pour tout $g_j \in H^{-\frac{1}{2}}(\Sigma)$ par

$$(e_{h,j}, \varphi_{E,h,j}) = E_{h,j} g_j \Leftrightarrow \begin{cases} \forall v_{h,j} \in V_{h,j}, \tilde{a}_j(e_{h,j}, \varphi_{E,h,j}, v_{h,j}) = l_j(0, g_j, v_{h,j}) \\ \forall \phi_{h,j} \in V_{h,\Sigma}, \tilde{b}_j(e_{h,j}, \varphi_{E,h,j}, \phi_{h,j}) = 0 \end{cases} \quad (6.69)$$

où les formes \tilde{a}_j , \tilde{b}_j , sont données par (6.62) et les second membres l_j par (6.15). Les opérateurs R_j sont définis de manière similaire comme opérateurs de $H^1(\Omega)$ dans $\tilde{V}_{h,j}$ pour tout $f \in H^1(\Omega)$ par

$$(r_{h,j}, \varphi_{R,h,j}) = R_{h,j} f \Leftrightarrow \begin{cases} \forall v_{h,j} \in V_{h,j}, \tilde{a}_j(r_{h,j}, \varphi_{R,h,j}, v_{h,j}) = l_j(f, 0, v_{h,j}) \\ \forall \phi_{h,j} \in V_{h,\Sigma}, \tilde{b}_j(r_{h,j}, \varphi_{R,h,j}, \phi_{h,j}) = 0 \end{cases} \quad (6.70)$$

Ainsi définis, les solutions $(u_{h,1}, \varphi_{h,1})$ et $(u_{h,2}, \varphi_{h,2})$ des problèmes locaux s'écrivent ($j = 1, 2$)

$$u_{h,j} = e_{h,j} + r_{h,j}, \quad \varphi_{h,j} = \varphi_{E,h,j} + \varphi_{R,h,j} \quad (6.71)$$

Ces opérateurs permettent une nouvelle fois de définir les opérateurs de scattering $S_h = (S_{h,1}, S_{h,2})$ qui permettent de calculer, à partir des traces entrantes $g_j^{in} \in V_{h,\Sigma}$ (paramètres des seconds membres l_j) les traces sortantes $g_j^{out} \in V_{h,\Sigma}$ ($j = 1, 2$). Pour cela, notons $(e_{h,j}, \varphi_{E,h,j}) = E_{h,j} g_j^{in}$, et pour tout $\phi_{h,j} \in V_{h,\Sigma}$. Alors, pour tout $\phi_{h,j} \in V_{h,\Sigma}$, on a

$$\begin{aligned} \langle g_{h,j}^{out}, \phi_{h,j} \rangle_{\Sigma} &= \langle S_{h,j} g_{h,j}^{in}, \phi_{h,j} \rangle_{\Sigma} \\ &= - \langle g_{h,j}^{in}, \phi_{h,j} \rangle_{\Sigma} + 2ik \langle T_r e_{h,j}, \phi_{h,j} \rangle_{\Sigma} \\ &= - \langle g_{h,j}^{in}, \phi_{h,j} \rangle_{\Sigma} + 2ik \operatorname{Re}(z) \langle \Lambda^* \varphi_{E,h,j}, \phi_{h,j} \rangle_{\Sigma}. \end{aligned} \quad (6.72)$$

Finalement, il nous reste à définir l'opérateur d'échange de traces discret Π_h . Cette fois, il est exactement identique au cas où l'opérateur T est donnée directement par sa forme bilinéaire. Il est donc défini par les équations (6.21) et (6.22).

6.4.3 PROCESSUS ITÉRATIF

Nous allons ici étudier les propriétés du processus itératif que nous avons déjà mis en évidence dans le cas continu, ainsi que lors de la première discrétisation étudiée. Il s'agit simplement d'adapter les démonstrations en utilisant les variables intermédiaires $\varphi_{h,j}$. En reprenant les notations suivantes

$$\begin{aligned} (e_{h,j}^n, \varphi_{E,h,j}^n) &= E_{h,j} g_j^{in,n} \\ (r_{h,j}, \varphi_{R,h,j}) &= R_{h,j} f \\ u_{h,j}^n &= e_{h,j}^n + r_{h,j} \\ \varphi_{h,j}^n &= \varphi_{E,h,j}^n + \varphi_{R,h,j} \end{aligned} \tag{6.73}$$

le processus itératif s'écrit

$$\begin{aligned} \langle g_{h,1}^{in,n+1}, \phi_h \rangle_\Sigma &= (1-r) \langle g_{h,1}^{in,n}, \phi_h \rangle_\Sigma + r \left(\langle -g_{h,2}^{in,n}, \phi_h \rangle_\Sigma + 2ik \langle T_r u_{h,2}^n, \phi_h \rangle_\Sigma \right) \\ &= (1-r) \langle g_{h,1}^{in,n}, \phi_h \rangle_\Sigma + r \langle S_{h,2} g_{h,2}^{in,n}, \phi_h \rangle_\Sigma + 2r ik \operatorname{Re}(z) \langle \Lambda^* \varphi_{R,h,2}, \phi_h \rangle_\Sigma \end{aligned} \tag{6.74}$$

$$\begin{aligned} \langle g_{h,2}^{in,n+1}, \phi_h \rangle_\Sigma &= (1-r) \langle g_{h,2}^{in,n}, \phi_h \rangle_\Sigma + r \left(\langle -g_{h,1}^{in,n}, \phi_h \rangle_\Sigma + 2ik \langle T_r u_{h,1}^n, \phi_h \rangle_\Sigma \right) \\ &= (1-r) \langle g_{h,2}^{in,n}, \phi_h \rangle_\Sigma + r \langle S_{h,1} g_{h,1}^{in,n}, \phi_h \rangle_\Sigma + 2r ik \operatorname{Re}(z) \langle \Lambda^* \varphi_{R,h,1}, \phi_h \rangle_\Sigma \end{aligned} \tag{6.75}$$

Dans la suite de cette section, et afin d'étudier la convergence du processus, nous prendrons $f = 0$. Nous nous intéressons à l'erreur $(\varepsilon_{h,1}^n, \varepsilon_{h,2}^n)$ sur les traces $(g_{h,1}^{in,n}, g_{h,2}^{in,n})$. Ces erreurs vérifient les relations

$$\begin{aligned} \langle \varepsilon_{h,1}^{n+1}, \phi_h \rangle_\Sigma &= (1-r) \langle \varepsilon_{h,1}^n, \phi_h \rangle_\Sigma + r \langle S_{h,2} \varepsilon_{h,2}^n, \phi_h \rangle_\Sigma, \\ \langle \varepsilon_{h,2}^{n+1}, \phi_h \rangle_\Sigma &= (1-r) \langle \varepsilon_{h,2}^n, \phi_h \rangle_\Sigma + r \langle S_{h,1} \varepsilon_{h,1}^n, \phi_h \rangle_\Sigma \end{aligned} \tag{6.76}$$

et nous allons étudier l'opérateur $S_{h,\pi}$. Comme précédemment, nous avons le lemme suivant.

Lemme 15 (Opérateurs discrets).

Sous l'hypothèse 2, et pour tout $h < H$, l'opérateur $I - S_{h,\pi}$ est un isomorphisme de $V_{h,\Sigma}^2$ dans $V_{h,\Sigma}^2$ et l'opérateur $S_{h,\pi}$ est contractant dans $V_{h,\Sigma}^2$ pour la norme $\|\cdot\|_X$ (2.88-2.89).

Démonstration. Pour montrer que $I - S_{h,\pi}$ est un isomorphisme, seule son injectivité doit être prouvée. Pour cela, prenons $g_h = (g_{h,1}, g_{h,2})$ tel que $(I - S_{h,\pi}) g_h = 0$. Alors, en notant $(e_{h,j}, \varphi_{E,h,j}) = E_{h,j} g_{h,j}$, et par définitions des opérateurs $S_{h,j}$ (6.72), on a

$$\begin{aligned} \langle g_{h,1} - S_{h,2} g_{h,2}, \phi_h \rangle_\Sigma &= \langle g_{h,1} + g_{h,2} - 2ik \operatorname{Re}(z) \Lambda^* \varphi_{E,h,2}, \phi_h \rangle_\Sigma = 0 \\ \langle g_{h,2} - S_{h,1} g_{h,1}, \phi_h \rangle_\Sigma &= \langle g_{h,1} + g_{h,2} - 2ik \operatorname{Re}(z) \Lambda^* \varphi_{E,h,1}, \phi_h \rangle_\Sigma = 0 \end{aligned} \quad (6.77)$$

En soustrayant ces deux équations, et puisque $\operatorname{Re}(z) > 0$, on obtient que pour tout $\phi_h \in V_{h,\Sigma}$

$$\langle \Lambda^* (\varphi_{E,h,2} - \varphi_{E,h,1}), \phi_h \rangle_\Sigma = 0. \quad (6.78)$$

En utilisant l'hypothèse 2 sur l'opérateur Λ , on en déduit que $\varphi_{E,h,2} = \varphi_{E,h,1}$. Puis, par définition des variables intermédiaires $\varphi_{h,j}$ et en utilisant une nouvelle fois l'hypothèse 2, on trouve que $e_{h,1} = e_{h,2}$ sur Σ . Une nouvelle fois, les solutions $e_{h,1}$ et $e_{h,2}$ se recollent à l'interface Σ . Nous pouvons définir une solution sur le domaine entier Ω par $e_{h|\Omega_1} = e_{h,1}$ et $e_{h|\Omega_2} = e_{h,2}$. De plus, pour tout $v_h \in V_h$, notons $v_{h,1} = v_{h|\Omega_1} \in V_{h,1}$, $v_{h,2} = v_{h|\Omega_2} \in V_{h,2}$ et $\phi_h \in V_{h,\Sigma}$ la trace commune de $v_{h,1}$ et $v_{h,2}$ sur Σ . Comme $(e_{h,j}, \varphi_{E,h,j}) = E_{h,j} g_{h,j}$,

$$\tilde{a}_1(e_{h,1}, \varphi_{E,h,1}, v_{h,1}) + \tilde{a}_2(e_{h,2}, \varphi_{E,h,2}, v_{h,2}) = l_1(0, g_{h,1}, v_{h,1}) + l_2(0, g_{h,2}, v_{h,2}) \quad (6.79)$$

En procédant comme précédemment, nous pouvons sommer les intégrales sur Ω_1 et Ω_2 en une intégrale sur Ω pour obtenir $(\varphi_{E,h} = \varphi_{E,h,1} = \varphi_{E,h,2})$

$$a(e_h, v_h) + 2ik \operatorname{Re}(z) \langle \Lambda^* \varphi_{E,h}, \phi_h \rangle_\Sigma = \langle g_{h,1} + g_{h,2}, \phi_h \rangle_\Sigma \quad (6.80)$$

où la forme bilinéaire a est donnée par 6.5 et correspond à la forme bilinéaire du problème global. En utilisant le fait que (6.77)

$$\langle g_{h,1} + g_{h,2}, \phi_h \rangle_\Sigma = 2ik \operatorname{Re}(z) \langle \Lambda^* \varphi_{E,h}, \phi_h \rangle_\Sigma, \quad (6.81)$$

il vient $a(e_h, v_h) = 0$ pour tout $v_h \in V_h$ et donc $e_h = 0$ dans Ω . Ceci conduit en particulier à $g_{h,1} = g_{h,2} = 0$. D'où l'isomorphisme.

Pour montrer que $S_{h,\pi}$ est une contraction dans $V_{h,\Sigma} \times V_{h,\Sigma}$ pour la norme $\|\cdot\|_X$, nous nous contenterons encore une fois d'étudier $S_{h,2}$ car Π_h est unitaire et la démonstration pour $S_{h,1}$ est presque identique. Par définition des opérateurs $S_{h,j}$, nous avons pour tout $\phi_h \in V_{h,\Sigma}$

$$\langle S_{h,2} g_{h,2}, \phi_h \rangle_\Sigma = \langle -g_{h,2} + 2ik \operatorname{Re}(z) \Lambda^* \varphi_{h,2}, \phi_h \rangle_\Sigma \quad (6.82)$$

Donc

$$\|S_{h,2} g_{h,2}\|_{-\frac{1}{2}}^2 - \|g_{h,2}\|_{-\frac{1}{2}}^2 = 4k^2 \mathcal{R}e(z)^2 \|\Lambda^* \varphi_{h,2}\|_{-\frac{1}{2}}^2 - 4k \mathcal{R}e(z) \mathcal{I}m(g_{h,2}, \Lambda^* \varphi_{h,2})_{-\frac{1}{2}} \quad (6.83)$$

Par définition du produit scalaire $(\cdot, \cdot)_{-\frac{1}{2}}$ (2.88), et comme $T_r^{-1} = (\mathcal{R}e(z) \Lambda^* \Lambda)^{-1}$

$$\|\Lambda^* \varphi_{h,2}\|_{-\frac{1}{2}}^2 = \left\langle \Lambda^* \varphi_{h,2}, \frac{1}{\mathcal{R}e(z)} \Lambda^{-1} \varphi_{h,2} \right\rangle_{\Sigma} = \frac{1}{\mathcal{R}e(z)} \langle \Lambda^* \varphi_{h,2}, e_{h,2} \rangle_{\Sigma}. \quad (6.84)$$

De même, on a

$$(g_{h,2}, \Lambda^* \varphi_{h,2})_{-\frac{1}{2}} = \frac{1}{\mathcal{R}e(z)} \langle g_{h,2}, e_{h,2} \rangle_{\Sigma}. \quad (6.85)$$

On sait de plus que les quantités $e_{h,2}$, $\varphi_{h,2}$ et $g_{h,2}$ vérifient pour tout $v_{h,2} \in V_{h,2}$

$$\tilde{a}_2(e_{h,2}, \varphi_{h,2}, v_{h,2}) = l_2(0, g_{h,2}, v_{h,2}). \quad (6.86)$$

En prenant la partie imaginaire de cette équation pour $v_{h,2} = e_{h,2}$, on obtient

$$\mathcal{I}m \langle g_{h,2}, e_{h,2} \rangle_{\Sigma} = k \int_{\Gamma} \eta^{-1} |e_{h,2}|^2 d\sigma + k \mathcal{R}e(z) \langle \Lambda^* \varphi_{h,2}, e_{h,2} \rangle_{\Sigma}. \quad (6.87)$$

Alors, en combinant les équations (6.84), (6.85) et (6.87) dans (6.83), on trouve

$$\|S_{h,2} g_{h,2}\|_{-\frac{1}{2}}^2 - \|g_{h,2}\|_{-\frac{1}{2}}^2 = -4k^2 \int_{\Gamma} \eta^{-1} |e_{h,2}|^2 d\sigma \leq 0 \quad (6.88)$$

$S_{h,2}$ est bien une contraction. ■

Le théorème suivant découle des propriétés de l'opérateur de scattering $S_{h,\pi}$.

Théorème 7 (Convergence de la DDM discrétisée – Opérateur Λ).

Pour tout $h < H$ et pour tout $r \in]0, 1[$, il existe une constante $C_h > 0$ et un réel $\tau_{h,r} \in]0, 1[$ tel que

$$\|u_{h,1}^n - u_{h,1}\|_{V_{h,1}} + \|u_{h,2}^n - u_{h,2}\|_{V_{h,2}} < C_h \tau_{h,r}^n, \quad (6.89)$$

et la méthode de décomposition de domaine discrétisée converge.

Démonstration. La preuve est identique à celle du théorème 6 ■

6.4.4 À LA CONVERGENCE

Bien que les démonstrations soient un peu plus techniques, la seconde discrétisation que nous proposons ici permet également de conserver les propriétés importantes du cas continu, à savoir le caractère bien posé des problèmes locaux, l'aspect isomorphique de $I - S_{h,\pi}$ et la contraction de $S_{h,\pi}$. Toutes ces propriétés assurent la convergence de la méthode de décomposition de domaine, y compris au niveau discret. Encore une fois, la preuve du lemme 15 laisse entrevoir une autre propriété spécifique à notre discrétisation. En notant $u_{h,1}^\infty$ et $u_{h,2}^\infty$ les limites de $u_{h,1}^n$ et $u_{h,2}^n$ lorsque $n \rightarrow \infty$, nous avons le lemme suivant

Lemme 16 (Propriété de recollement).

La solution obtenue à la convergence de la méthode de décomposition de domaine correspond exactement à la solution du problème discret non décomposé, c'est à dire

$$u_{h,1}^\infty = u_h|_{\Omega_1}, \quad \text{et} \quad u_{h,2}^\infty = u_h|_{\Omega_2} \quad (6.90)$$

Démonstration. Notons $g_{h,1}^{in,\infty}$ et $g_{h,2}^{in,\infty}$ les traces entrantes des solutions $u_{h,1}^\infty$ et $u_{h,2}^\infty$. En s'inspirant des notations définies en (6.73), nous définissons

$$\begin{aligned} (e_{h,j}^\infty, \varphi_{E,h,j}^\infty) &= E_{h,j} g_{h,j}^{in,n} \\ (r_{h,j}, \varphi_{R,h,j}) &= R_{h,j} f \\ u_{h,j}^\infty &= e_{h,j}^\infty + r_{h,j} \\ \varphi_{h,j}^\infty &= \varphi_{E,h,j}^\infty + \varphi_{R,h,j} \end{aligned} \quad (6.91)$$

En écrivant les équations du processus itératif (6.74) et (6.75) pour $n \rightarrow \infty$, on trouve

$$\begin{aligned} \langle g_{h,1}^{in,\infty} + g_{h,2}^{in,\infty}, \phi_h \rangle_\Sigma &= 2ik \mathcal{R}e(z) \langle \Lambda^* \varphi_{h,1}^\infty, \phi_h \rangle_\Sigma \\ \langle g_{h,1}^{in,\infty} + g_{h,2}^{in,\infty}, \phi_h \rangle_\Sigma &= 2ik \mathcal{R}e(z) \langle \Lambda^* \varphi_{h,2}^\infty, \phi_h \rangle_\Sigma \end{aligned} \quad (6.92)$$

En utilisant l'hypothèse 2, on voit que $\varphi_{h,1}^\infty = \varphi_{h,2}^\infty$. Par définition des inconnues $\varphi_{h,j}$ et en utilisant une nouvelle fois l'hypothèse 2, il vient que $u_{h,1|_\Sigma}^\infty = u_{h,2|_\Sigma}^\infty$. Nous pouvons dès lors construire $u_h^\infty \in V_h$ tel que

$$u_{h|_{\Omega_1}}^\infty = u_{h,1}^\infty \quad \text{et} \quad u_{h|_{\Omega_2}}^\infty = u_{h,2}^\infty. \quad (6.93)$$

De plus pour tout $v_h \in V_h$, nous noterons $v_{h,1} = v_{h|\Omega_1} \in V_{h,1}$, $v_{h,2} = v_{h|\Omega_2} \in V_{h,2}$ et $\phi_h = v_{h,1|\Sigma} = v_{h,2|\Sigma}$. Alors, pour tout $v_h \in V_h$, on a

$$\tilde{a}_1(u_{h,1}^\infty, v_{h,1}) = l_1(f, g_{h,1}^{in,\infty}, v_{h,1}), \quad \text{et} \quad \tilde{a}_2(u_{h,2}^\infty, v_{h,2}) = l_2(f, g_{h,2}^{in,\infty}, v_{h,2}) \quad (6.94)$$

En sommant ces deux équations, et en sommant les intégrales sur Ω_1 et Ω_2 , on obtient

$$a(u_h^\infty, v_h) + 2ik \operatorname{Re}(z) \langle \Lambda^* \varphi_h^\infty, v_h \rangle_\Sigma = \int_\Omega f \bar{v}_h dx + \langle g_{h,1}^{in,\infty} + g_{h,2}^{in,\infty}, v_h \rangle_\Sigma. \quad (6.95)$$

Finalement, en utilisant (6.92), on trouve directement

$$\forall v_h \in V_h, \quad a(u_h^\infty, v_h) = \int_\Omega f \bar{v}_h dx \quad (6.96)$$

et u_h^∞ , construit en recollant les solutions $u_{h,1}^\infty$ et $u_{h,2}^\infty$ des problèmes locaux à la convergence, correspond en fait à la solution du problème global discret non décomposé. ■

6.4.5 FORME MATRICIELLE DU PROBLÈME

L'utilisation des inconnues intermédiaires $\varphi_{h,j}^n = \Lambda u_{h,j}^n$ lorsque l'opérateur de transmission T n'est donné que par la forme bilinéaire de Λ modifie légèrement les formes matricielles des problèmes étudiés, et du processus itératif. Nous reprendrons ici les notations introduites section 6.3.5, notamment concernant les fonctions de base des espaces de discrétisation $V_{h,1}$, $V_{h,2}$ et $V_{h,\Sigma}$. De même, les matrices \mathbb{B}_j de sélection des degrés de liberté sur Σ restent définies par (6.45). Ajoutons seulement les vecteurs $\varphi_{h,j}^n$ des composantes de $\varphi_{h,j}^n$ dans la base de $V_{h,\Sigma}$.

Comme précédemment pour l'opérateur T , l'opérateur Λ est un opérateur agissant sur les traces des fonctions de $V_{h,j}$, $j = 1, 2$, qui sont de fait des fonctions de $V_{h,\Sigma}$. Notons \mathbb{L} sa représentation matricielle dans la base de $V_{h,\Sigma}$

$$\mathbb{L}_{k,l} = (\Lambda \Phi_{h,k}, \Phi_{h,l}) = \int_\Sigma \Lambda \Phi_{h,k} \bar{\Phi}_{h,l} d\sigma. \quad (6.97)$$

Enfin, définissons \mathbb{M}_Σ la matrice de masse des fonctions de bases de $V_{h,\Sigma}$

$$(\mathbb{M}_\Sigma)_{k,l} = \int_\Sigma \Phi_{h,k} \bar{\Phi}_{h,l} d\sigma. \quad (6.98)$$

Avec toutes ces notations, les équations vérifiées par $(\mathbf{u}_{h,j}^n, \boldsymbol{\varphi}_{h,j}^n, \mathbf{g}_{h,j}^{in,n})$ sont

$$\underbrace{\begin{bmatrix} \tilde{\mathbb{K}}_{h,1} & ik z \mathbb{B}_1^T \mathbb{L}^* \\ -\mathbb{L} \mathbb{B}_1 & \mathbb{M}_\Sigma \end{bmatrix}}_{\mathbb{K}_{h,1}} \cdot \begin{bmatrix} \mathbf{u}_{h,1}^n \\ \boldsymbol{\varphi}_{h,1}^n \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1 + \mathbb{B}_1^T \mathbf{g}_{h,1}^{in,n} \\ 0 \end{bmatrix}, \quad (6.99)$$

$$\underbrace{\begin{bmatrix} \tilde{\mathbb{K}}_{h,2} & ik \bar{z} \mathbb{B}_2^T \mathbb{L}^* \\ -\mathbb{L} \mathbb{B}_2 & \mathbb{M}_\Sigma \end{bmatrix}}_{\mathbb{K}_{h,2}} \cdot \begin{bmatrix} \mathbf{u}_{h,2}^n \\ \boldsymbol{\varphi}_{h,2}^n \end{bmatrix} = \begin{bmatrix} \mathbf{f}_2 + \mathbb{B}_2^T \mathbf{g}_{h,2}^{in,n} \\ 0 \end{bmatrix}, \quad (6.100)$$

où les matrices $\tilde{\mathbb{K}}_{h,1}$ et $\tilde{\mathbb{K}}_{h,2}$ sont les mêmes que dans le cas précédent (6.51). Les matrices $\mathbb{K}_{h,1}$ et $\mathbb{K}_{h,2}$ représentant les problèmes locaux sont, dans ce cas, augmentées et de tailles respectives $(n_{h,1} + n_{h,\Sigma}) \times (n_{h,1} + n_{h,\Sigma})$ et $(n_{h,2} + n_{h,\Sigma}) \times (n_{h,2} + n_{h,\Sigma})$. Obtenir une expression équivalente à (6.54) donnant $\mathbf{g}_{h,j}^{out,n}$ à partir des $\mathbf{g}_{h,j}^{in,n}$ est un peu plus compliqué dans ce cas. Cela nécessite l'introduction de nouvelles matrices de sélection d'inconnues. Soient $\mathbb{B}_{j,\Sigma}$ les matrices de taille $n_{h,\Sigma} \times (n_{h,1} + n_{h,\Sigma})$ et $n_{h,\Sigma} \times (n_{h,2} + n_{h,\Sigma})$ telles que

$$\mathbb{B}_{j,\Sigma} \cdot \begin{bmatrix} \mathbf{u}_{h,j}^n \\ \boldsymbol{\varphi}_{h,j}^n \end{bmatrix} = \boldsymbol{\varphi}_{h,j}^n, \quad \mathbb{B}_{j,\Sigma} = \begin{bmatrix} 0 & \mathbb{I}_\Sigma \end{bmatrix} \quad (6.101)$$

et \mathbb{B}_{j,Ω_j} les matrices de taille $n_{h,1} \times (n_{h,1} + n_{h,\Sigma})$ et $n_{h,2} \times (n_{h,2} + n_{h,\Sigma})$ telles que

$$\mathbb{B}_{j,\Omega_j} \cdot \begin{bmatrix} \mathbf{u}_{h,j}^n \\ \boldsymbol{\varphi}_{h,j}^n \end{bmatrix} = \mathbf{u}_{h,j}^n, \quad \mathbb{B}_{j,\Omega_j} = \begin{bmatrix} \mathbb{I}_{\Omega_j} & 0 \end{bmatrix}. \quad (6.102)$$

Le calcul des traces sortantes, donné variationnellement par (6.72) s'écrit matriciellement

$$\mathbf{g}_{h,j}^{out,n} = -\mathbf{g}_{h,j}^{in,n} + 2ik \operatorname{Re}(z) \mathbb{L}^* \boldsymbol{\varphi}_{h,j}^n, \quad (6.103)$$

où $\boldsymbol{\varphi}_{h,1}^n$ est solution de (6.99) pour le sous-domaine Ω_1 et $\boldsymbol{\varphi}_{h,2}^n$ est solution de (6.100) pour le sous-domaine Ω_2 . Puisque les matrices locales $\mathbb{K}_{h,1}$ et $\mathbb{K}_{h,2}$ sont

inversibles d'après le lemme 14, il est possible d'éliminer les inconnues $\varphi_{h,j}^n$ dans (6.103) pour obtenir

$$\mathbf{g}_{h,j}^{out,n} = \underbrace{\left(-\mathbb{I}_\Sigma + 2ik \operatorname{Re}(z) \mathbb{L}^* \mathbb{B}_{j,\Sigma} \mathbb{K}_{h,j}^{-1} \mathbb{B}_{j,\Omega_j}^T \mathbb{B}_j^T \right)}_{\mathbb{S}_{h,j}} \mathbf{g}_{h,j}^{in,n} + \underbrace{2ik \operatorname{Re}(z) \mathbb{L}^* \mathbb{B}_{j,\Sigma} \mathbb{K}_{h,j}^{-1} \mathbb{B}_{j,\Omega_j}^T}_{\mathbf{F}_j} \mathbf{f}_j \quad (6.104)$$

Même si les expressions (6.104) font intervenir de multiples produits matrice-matrice, il faut bien voir que la multiplication par l'une des matrices \mathbb{B} correspond simplement à une sélection d'inconnues, dont le coût est négligeable. De plus, les coefficients \mathbf{F}_j sont indépendants des itérations et ne nécessitent d'être calculés qu'une seule fois. Enfin, le calcul des inverses $\mathbb{K}_{h,j}^{-1}$, certes coûteux, n'a également besoin d'être réalisé qu'une seule fois, les matrices $\mathbb{K}_{h,j}$ étant elles aussi invariantes au cours des itérations.

Enfin, la relation de récurrence vérifiée par les traces entrantes s'écrit

$$\begin{aligned} \mathbf{g}_{h,1}^{in,n+1} &= (1-r) \mathbf{g}_{h,1}^{in,n} + r \mathbf{g}_{h,2}^{out,n} \\ \mathbf{g}_{h,2}^{in,n+1} &= (1-r) \mathbf{g}_{h,2}^{in,n} + r \mathbf{g}_{h,1}^{out,n} \end{aligned} \quad (6.105)$$

Comme dans le cas où T est directement donné par sa forme bilinéaire, la matrice \mathbb{L} est dense si aucune troncature n'est appliquée. Le surcoût induit par l'utilisation d'un opérateur intégral sur Σ peut donc être relativement important en terme de temps de calcul par itération. Cependant, si la troncature est ajoutée, la matrice \mathbb{L} est nettement plus creuse car elle ne relie plus que les points suffisamment proches. Le profil creux de la matrice \mathbb{L} est directement gouverné par la troncature. Nous verrons dans les chapitres suivants que la troncature permet d'avoir un surcoût presque négligeable, sans dégrader la vitesse de convergence. Une autre approche, non implémentée et non testée, consisterait à ne pas appliquer de troncature à l'opérateur Λ , mais à la place, d'utiliser des algorithmes de compression de matrice, de type *Adaptive Cross-Approximation (ACA)* ou *\mathcal{H} -matrix* (voir pour plus de détail [3]), car la structure de la matrice \mathbb{L} s'y prête bien. Puisqu'elle représente un opérateur intégral surfacique, il s'agit en effet d'une matrice très peu creuse, mais de rang faible. Nous intuitions que ce type d'algorithme permettrait une réduction substantielle du surcoût induit par l'opérateur sans troncature.

7

Mise en œuvre numérique

Sommaire

7.1	Généralités	156
7.1.1	Les maillages	156
7.1.2	Codes numériques et choix des éléments	162
7.1.3	Algorithmes	162
7.1.3.1	Algorithme de Jacobi	166
7.1.3.2	Algorithme GMRES	167
7.2	Spécificités de l'opérateur Λ	168
7.2.1	Assemblage	168
7.2.1.1	Validation sur le cercle	171
7.2.1.2	Validation des propriétés de Λ	173
7.2.2	Réalisation de la troncature	177
7.2.2.1	Approche naïve	180
7.2.2.2	Approche par intégration par parties	181

Ce chapitre sera consacré à la mise en œuvre pratique de la méthode de décomposition de domaine dans un code informatique d'éléments finis. Dans une première section, nous verrons quelques généralités concernant la simulation numérique par éléments finis du problème étudié, puis nous nous concentrerons plus particulièrement sur l'opérateur Λ et quelques unes de ses spécificités.

7.1 GÉNÉRALITÉS

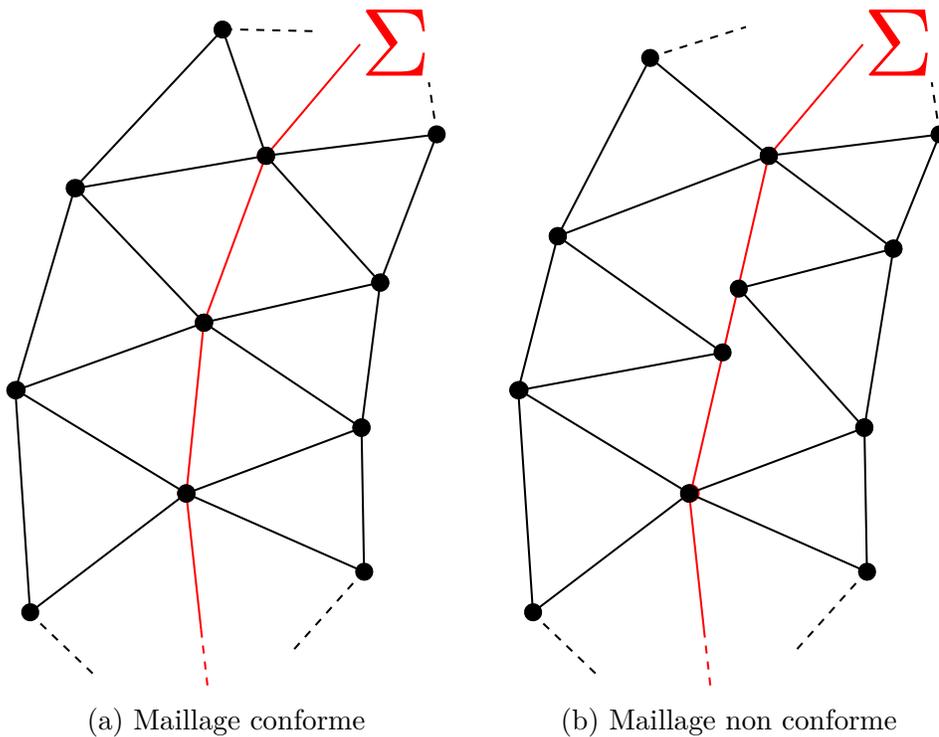
7.1.1 LES MAILLAGES

Toute simulation numérique d'un phénomène physique (et *a fortiori* de la propagation d'onde) repose sur un domaine physique. La première étape de la simulation consiste donc à approcher ce domaine physique par un maillage. Nous avons vu dans le chapitre précédent que le maillage de notre domaine Ω doit vérifier une propriété très importante : la conformité à l'interface Σ . La figure 7.1 montre deux exemples de maillages autour d'une interface Σ , l'un étant conforme, l'autre non. Pour réaliser des maillages conformes des différentes géométries étudiées, nous avons utilisé le mailleur `Gmsh` développé par Christophe Geuzaine [16]. En plus de la géométrie du problème, un paramètre h_0 est donné au mailleur comme objectif de longueur d'arête. Ce paramètre est important à deux niveaux. Tout d'abord, nous avons vu dans la partie concernant la discrétisation que la taille caractéristique du maillage doit être suffisamment petite pour assurer que les problèmes locaux, ainsi que le problème global, soient bien posés. En pratique, nous n'avons jamais été confronté à un cas pour lequel un des problèmes était mal posé. Le second aspect important concerne la précision de la solution. Il est clair que plus le maillage est fin (h_0 petit) et plus l'erreur due à la discrétisation sera faible. Pour les phénomènes de propagation d'onde, on parle de manière équivalente de *nombre de points par longueur d'onde*, c'est à dire combien de points de maillage sont utilisés pour représenter une onde de longueur λ . On a les relations suivantes

$$k = 2\pi \frac{f}{c}, \quad \lambda = \frac{c}{f}, \quad n_\lambda = \frac{\lambda}{h} = \frac{2\pi}{kh} \quad (7.1)$$

avec k le nombre d'onde (en m^{-1}), f la fréquence de l'onde (en Hz), c la vitesse de la lumière dans le vide¹, λ la longueur d'onde (en m), h la taille caractéristique d'un élément du maillage (en m) et n_λ le nombre de points par longueur d'onde. Plus le nombre de points par longueur d'onde est élevé, et mieux le maillage sera

1. $c = 299792458 \text{ m.s}^{-1}$

FIGURE 7.1 – Exemple de deux types de triangulation autour de l'interface Σ

capable de représenter des modes élevés. Dans la suite du document, nous utiliserons des maillages possédant de 10 (peu raffiné) à 270 (très raffiné) points par longueur d'onde.

Pour vérifier la qualité des maillages, plusieurs critères s'offrent à nous. Tout d'abord, il est intéressant de vérifier que l'objectif de pas de maillage h_0 est respecté. Pour l'exemple de la géométrie circulaire, nous avons représenté la distribution des tailles d'arêtes rapportées à la longueur objective h_0 et pour différentes valeurs de h_0 . Sur l'ensemble des maillages étudiés, on peut voir sur les figures 7.3 et 7.4 qu'environ 80% des arêtes sont comprises dans l'intervalle $[0.8 h_0, 1.2 h_0]$ et que plus de 95% des arêtes sont de taille inférieure à $1.2 h_0$. De plus, la moyenne des tailles des arêtes est inférieure à h_0 . Les tailles d'arêtes sont globalement bien contrôlées.

Un second critère important concernant la qualité des maillages correspond au

ratio suivant

$$\sigma_i = \frac{d_i}{r_i} \quad (7.2)$$

pour chaque élément i de la triangulation, où d_i est le diamètre de l'élément (pour un triangle, la plus longue arête) et r_i est le rayon de la plus grande boule inscrite dans l'élément (pour un triangle, le rayon de son cercle inscrit). Ce ratio permet de d'estimer la forme d'un triangle. Plus le ratio est faible, et plus le triangle associé est proche d'un triangle équilatéral. Au contraire un triangle avec un ratio très élevé (> 10) sera très étiré et de mauvaise qualité pour une méthode éléments finis. Nous avons représenté figure 7.2 des exemples de triangle pour quelques ratios caractéristiques. Pour plus de détails sur l'intérêt de ce ratio comme paramètre de contrôle de qualité de maillage, nous renvoyons le lecteur à l'annexe *Results of standard finite element approximation* pages 95-111 de [17]. Sur l'ensemble des maillages étudiés, ce ratio ne dépasse pas 7.5 et plus de 99% des triangles ont un ratio inférieur à 5. Ceci montre la bonne qualité des maillages générés par le logiciel Gmsh.

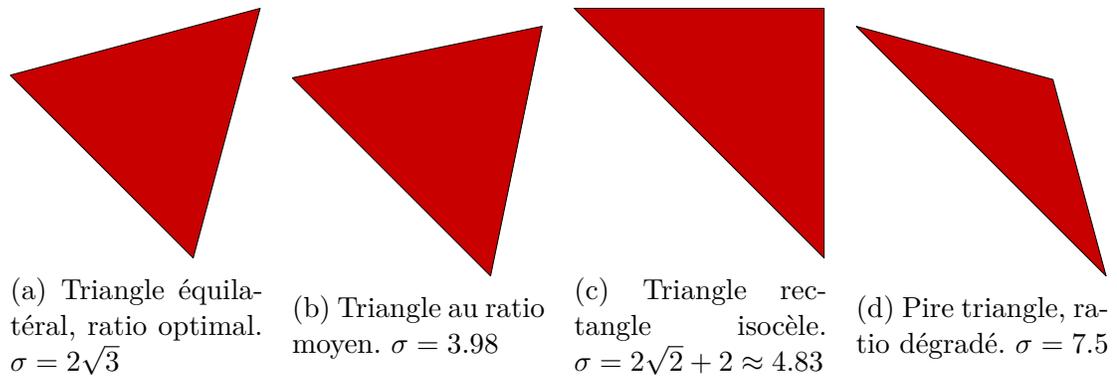
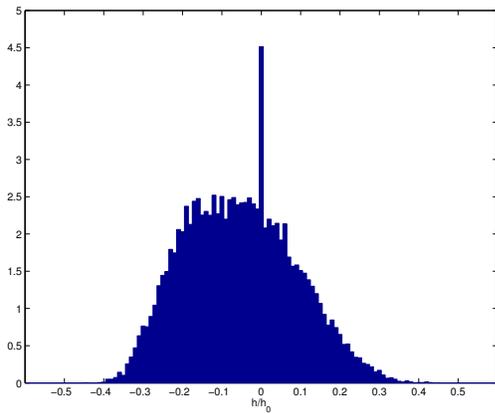
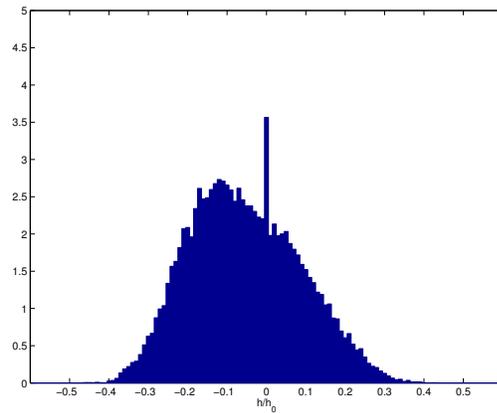


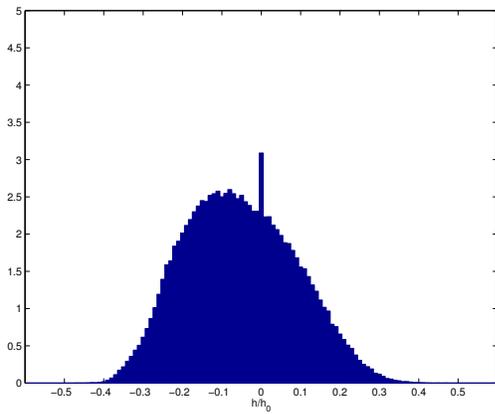
FIGURE 7.2 – Exemples de ratio σ pour différents triangles, allant du ratio optimal (triangle équilatéral) au pire ratio trouvé dans les maillages étudiés



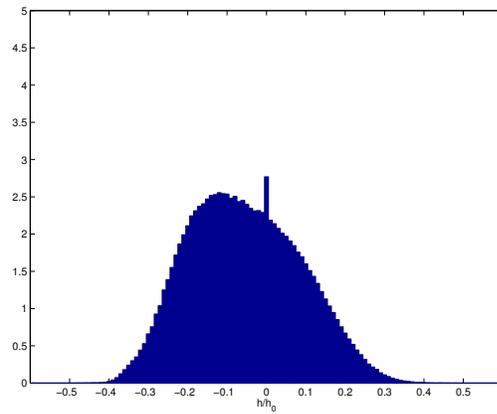
(a) $h_0 = 31.45e^{-3}$



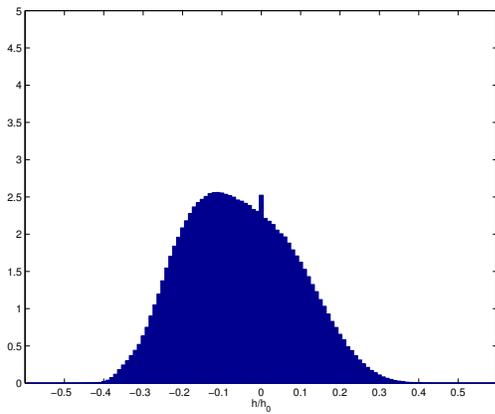
(b) $h_0 = 20.94e^{-3}$



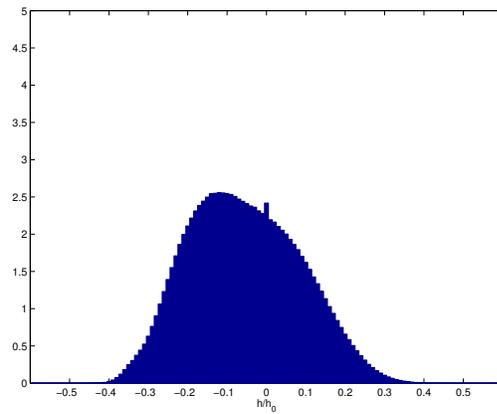
(c) $h_0 = 10.47e^{-3}$



(d) $h_0 = 6.98e^{-3}$



(e) $h_0 = 3.49e^{-3}$



(f) $h_0 = 2.33e^{-3}$

FIGURE 7.3 – Répartition de la taille des arêtes d'un maillage relative à la taille objective h_0 pour différents raffinements

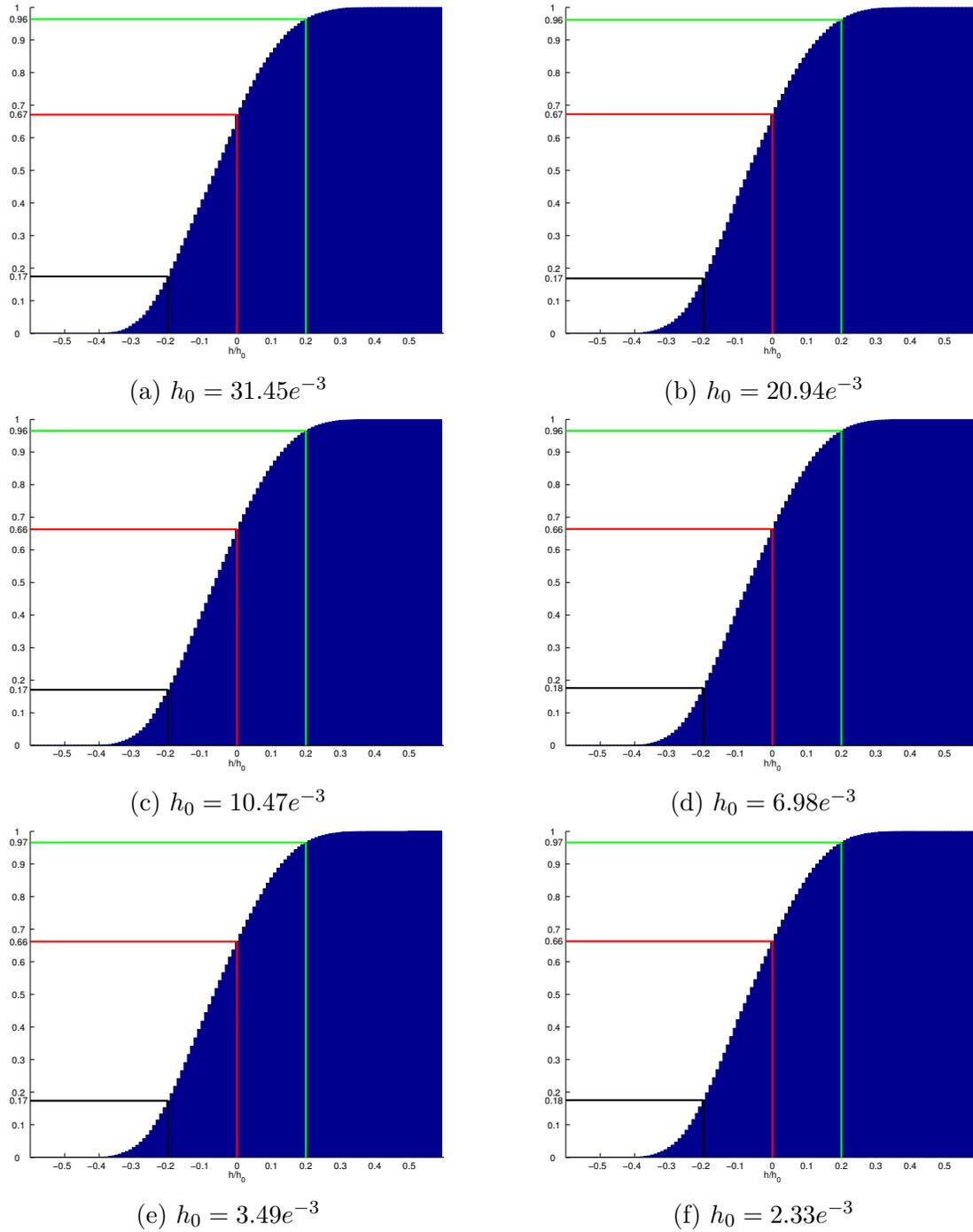


FIGURE 7.4 – Répartition cumulée de la taille des arêtes d'un maillage relative à la taille objective h_0 pour différents raffinements. Par exemple, dans (f), 97% des arêtes ont une taille inférieure à $1.2h_0$

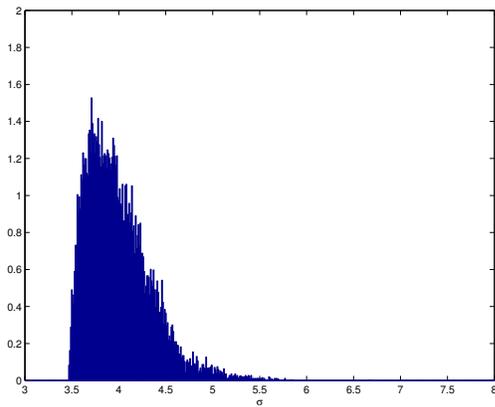
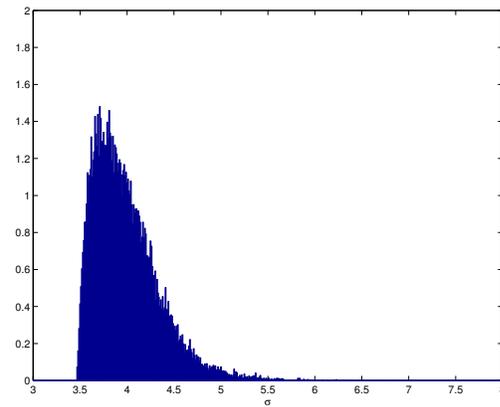
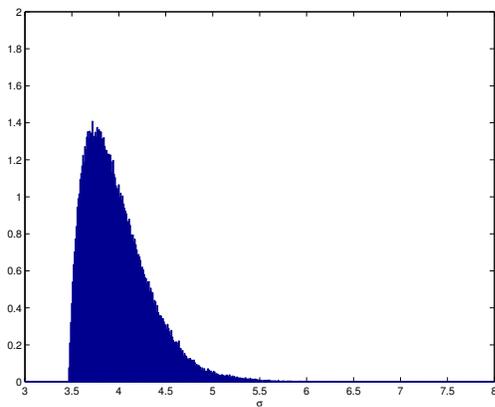
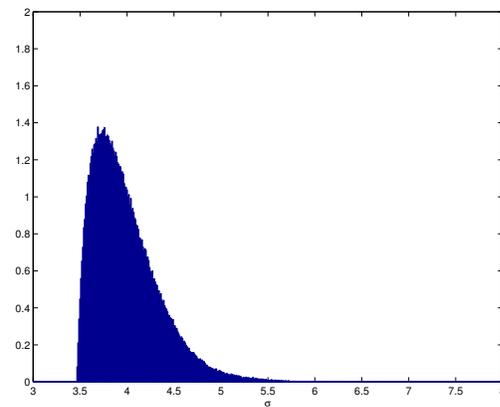
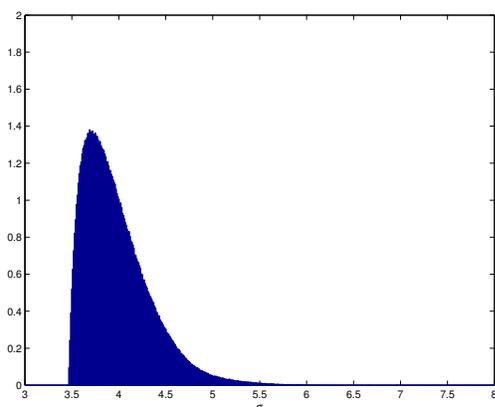
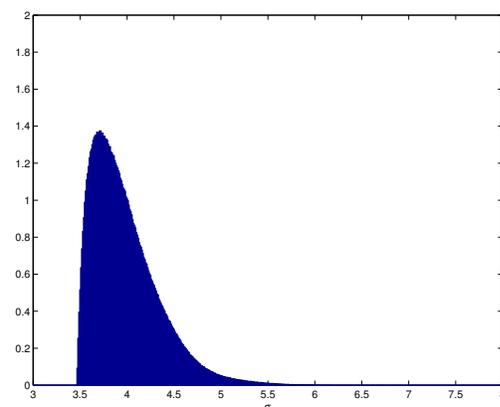
(a) $h_0 = 31.45e^{-3}$ (b) $h_0 = 20.94e^{-3}$ (c) $h_0 = 10.47e^{-3}$ (d) $h_0 = 6.98e^{-3}$ (e) $h_0 = 3.49e^{-3}$ (f) $h_0 = 2.33e^{-3}$

FIGURE 7.5 – Répartition des ratios σ d'un maillage pour différents raffinements. Tous les ratios sont inférieurs à 7.5.

7.1.2 CODES NUMÉRIQUES ET CHOIX DES ÉLÉMENTS

Pour implémenter les méthodes de décomposition de domaine que nous avons développées, nous avons commencé par travailler sur une maquette `Fortran` de décomposition de domaine développée par Francis Collino. Ce code utilise des éléments finis d'ordre 1 sur des maillages triangulaires (2D uniquement) pour la résolution de l'équation d'Helmholtz. Cette maquette nous a permis de vérifier rapidement la faisabilité de nos conditions de transmission et d'avoir nos premiers résultats sur des géométries simples. Cette maquette nous a également permis de mieux appréhender la construction de l'opérateur Λ et ses spécificités, que nous verrons plus loin dans ce chapitre.

Nous avons ensuite implémenté une méthode de décomposition de domaine, ainsi que nos conditions de transmission, dans le code `Montjoie` développé par Marc Duruflé à l'INRIA Bordeaux-Sud Ouest. Il s'agit d'un code éléments finis orienté pour l'ordre élevé, et écrit en C++. Il permet de résoudre un large spectre d'équations simulant la propagation d'ondes, en régime temporel et harmonique : équation de Helmholtz (et équation d'onde), équations de Maxwell, aéroacoustique et élastodynamique. Les maillages peuvent être 2D ou 3D, tétraédriques, quadrangulaires ou hexaédriques.

7.1.3 ALGORITHMES

Nous allons détailler ici l'implémentation de la méthode de décomposition de domaine dans sa variante basée sur l'algorithme de Jacobi, ainsi que pour GMRES. Les deux méthodes reposent sur une base commune, à savoir une phase d'initialisation et une fonction de *scattering*. Cette fonction consiste à calculer, à partir de traces entrantes, les traces sortantes. Ce calcul fait intervenir la résolution d'un problème local dans chacun des sous-domaines. N'importe quel solveur linéaire peut être utilisé pour cela. Il est même possible d'utiliser des solveurs différents dans chaque sous-domaine, ou encore d'utiliser un solveur itératif. Cependant, comme cette fonction a vocation à être appelée à chaque itération, il est intéressant de factoriser les matrices locales dans la phase d'initialisation, de telle sorte qu'ensuite la résolution d'un problème local soit moins coûteuse. Nous avons utilisé le solveur direct `PaStiX` développé à l'INRIA par Pierre Ramet [20],[23].

Le solveur `PaStiX` est un solveur parallèle direct pour matrice creuse. Il se base sur les méthodes de factorisation LL^T , LDL^T ou LU pour la résolution de systèmes linéaires. Massivement parallèle, l'intérêt de ce solveur dans notre cas est d'être *multi-threadé*, ajoutant ainsi un second niveau de parallélisme pour la résolution

des problèmes locaux au premier niveau de parallélisme induit par la décomposition de domaine.

Pour écrire les algorithmes, nous nous baserons sur les notations et les calculs effectués dans les sections 6.3.5 et 6.4.5 du chapitre précédent relatives aux formes matricielles des problèmes. La phase d'initialisation consiste principalement à construire les matrices des problèmes locaux, et plus généralement toutes les quantités qui restent invariantes au cours des itérations. Elle est décrite par l'algorithme 1. Pour la fonction de *scattering*, deux versions coexistent, selon que l'opérateur

Algorithme 1 Phase d'initialisation

```

1 : Lecture des maillages et paramètres
2 : si Opérateur donné =  $T$  alors
3 :   Construire la matrice  $\mathbb{T}$ 
4 : sinon
5 :   Construire la matrice  $\mathbb{L}$  et la matrice de masse  $\mathbb{M}_\Sigma$ 
6 : fin si
7 : pour  $j = 1, 2$  faire
8 :   Construire les matrices  $\tilde{K}_{h,j}$  (6.51)
9 :   Construire les matrices des problèmes locaux  $K_{h,j}$  (6.50 ou 6.99-6.100)
10 :   Factorisation LU des matrices locales  $K_{h,j}$  avec PaStiX
11 :   Construire les seconds membres  $\mathbf{F}_j$  (6.54 ou 6.104)
12 : fin pour
  
```

T soit donné directement par sa forme bilinéaire, ou par la forme bilinéaire de Λ . Ce sont les algorithmes 2 et 3. Il est important de noter que les algorithmes 2 et 3 ne font intervenir qu'un seul sous-domaine à la fois. Cette fonction peut donc être appelée en parallèle sur chacun des sous-domaines sans qu'il n'y ait de communication entre les sous-domaines. De même, la boucle de la phase d'initialisation (lignes 7 à 12) est complètement parallélisable puisqu'aucune interaction entre sous-domaines n'est nécessaire.

Nous utiliserons également une fonction qui, étant donné des traces entrantes \mathbf{g}_j , renvoie la solution volumique \mathbf{u}_j dans le sous-domaine Ω_j . Cette fonction **Solution** permet de récupérer la solution \mathbf{u} recherchée. Elle se décline, comme la fonction **Scattering**, en deux variantes, selon que l'opérateur T nous soit donné directement avec sa forme bilinéaire, ou avec la forme bilinéaire de Λ . Ce sont les algorithmes 4 et 5.

D'après la proposition 3 et l'équation (2.77), les algorithmes itératifs que nous

Algorithme 2 Fonction de *scattering* – Opérateur T

Requiert : Vecteur des traces entrantes $\mathbf{g}_{h,j}^{in}$ **Renvoie** : Vecteur \mathbf{S}_j tel que $\mathbf{S}_j = \mathbb{S}_{h,j} \mathbf{g}_{h,j}^{in}$

1 : Initialiser le second membre du problème local à partir des traces entrantes

$$\mathbf{g} \leftarrow \mathbb{B}_j^T \mathbf{g}_{h,j}^{in}$$

2 : Résoudre le problème local $\mathbb{K}_{h,j} \mathbf{x} = \mathbf{g}$ ($\mathbb{K}_{h,j}$ étant factorisée préalablement)

$$\mathbf{x} \leftarrow \mathbb{K}_{h,j}^{-1} \mathbf{g}$$

3 : Extraire la trace de la solution \mathbf{x} sur Σ

$$\boldsymbol{\varphi} \leftarrow \mathbb{B}_j \mathbf{x}$$

4 : Appliquer l'opérateur T en réalisant le produit matrice-vecteur

$$\mathbf{S}_j \leftarrow \mathbb{T}_r \boldsymbol{\varphi}$$

5 : Calculer les traces sortantes de Ω_j

$$\mathbf{S}_j \leftarrow -\mathbf{g}_{h,j}^{in} + 2ik \mathbf{S}_j$$

Algorithme 3 Fonction de *scattering* – Opérateur Λ

Requiert : Vecteur des traces entrantes $\mathbf{g}_{h,j}^{in}$ **Renvoie** : Vecteur \mathbf{S}_j tel que $\mathbf{S}_j = \mathbb{S}_{h,j} \mathbf{g}_{h,j}^{in}$

1 : Initialiser le second membre du problème local à partir des traces entrantes

$$\mathbf{g} \leftarrow \begin{bmatrix} \mathbb{B}_j^T \mathbf{g}_{h,j}^{in} \\ 0_{n_{h,\Sigma}} \end{bmatrix}$$

2 : Résoudre le problème local $\mathbb{K}_{h,j} \mathbf{x} = \mathbf{g}$ ($\mathbb{K}_{h,j}$ étant factorisée préalablement)

$$\mathbf{x} \leftarrow \mathbb{K}_{h,j}^{-1} \mathbf{g}$$

3 : Extraire les inconnues de surface $\boldsymbol{\varphi}$ de \mathbf{x}

$$\boldsymbol{\varphi} \leftarrow \mathbb{B}_{j,\Sigma} \mathbf{x}$$

4 : Appliquer l'opérateur Λ en réalisant le produit matrice-vecteur

$$\mathbf{S}_j \leftarrow \mathbb{L}^* \boldsymbol{\varphi}$$

5 : Calculer les traces sortantes de Ω_j

$$\mathbf{S}_j \leftarrow -\mathbf{g}_{h,j}^{in} + 2ik \mathcal{R}e(z) \mathbf{S}_j$$

Algorithme 4 Solution \mathbf{u} à partir des traces entrantes – Opérateur T

Requiert : Vecteur des traces entrantes $\mathbf{g}_{h,j}^{in}$ **Renvoie** : Vecteur \mathbf{u}_j solution du problème local dans Ω_j avec termes sources

1 : Initialiser le second membre du problème local à partir des traces entrantes

$$\mathbf{g} \leftarrow \mathbf{f}_j + \mathbb{B}_j^T \mathbf{g}_{h,j}^{in}$$

2 : Résoudre le problème local $\mathbb{K}_{h,j} \mathbf{u}_j = \mathbf{g}$ ($\mathbb{K}_{h,j}$ étant factorisée préalablement)

$$\mathbf{u}_j \leftarrow \mathbb{K}_{h,j}^{-1} \mathbf{g}$$

Algorithme 5 Solution \mathbf{u} à partir des traces entrantes – Opérateur Λ

Requiert : Vecteur des traces entrantes $\mathbf{g}_{h,j}^{in}$ **Renvoie** : Vecteur \mathbf{u}_j solution du problème local dans Ω_j avec termes sources

1 : Initialiser le second membre du problème local à partir des traces entrantes

$$\mathbf{g} \leftarrow \begin{bmatrix} \mathbf{f}_j + \mathbb{B}_j^T \mathbf{g}_{h,j}^{in} \\ 0_{n_h, \Sigma} \end{bmatrix}$$

2 : Résoudre le problème local $\mathbb{K}_{h,j} \mathbf{x}_j = \mathbf{g}$ ($\mathbb{K}_{h,j}$ étant factorisée préalablement)

$$\mathbf{x} \leftarrow \mathbb{K}_{h,j}^{-1} \mathbf{g}$$

3 : Extraire les inconnues volumique \mathbf{u}_j de \mathbf{x}

$$\mathbf{u}_j \leftarrow \mathbb{B}_{j, \Omega_j} \mathbf{x}$$

développons cherchent à résoudre le problème linéaire (écrit pour les traces entrantes)

$$(I - S_\pi) g^{in} = F. \quad (7.3)$$

La matrice (que l'on notera \mathbb{A}) représentant l'opérateur $I - S_\pi$ n'est jamais construite explicitement. Cependant, il nous sera utile d'avoir une fonction qui, étant donné un vecteur \mathbf{g} , renvoie l'action de cette matrice sur ce vecteur, c'est à dire une fonction calculant le produit matrice-vecteur $\mathbb{A} \mathbf{g}$ sans avoir la matrice \mathbb{A} sous un format explicite. Une telle fonction est facile à construire grâce à la fonction `Scattering` définie précédemment. L'algorithme 6 présente cette fonction.

Algorithme 6 Implémentation du produit matrice-vecteur **MatMult****Requiert** : Vecteur d'entrée \mathbf{g} **Renvoie** : Vecteur $\tilde{\mathbf{g}} = \mathbb{A}\mathbf{g}$

1 : Extraire les inconnues de chacun des sous-domaines

$$\mathbf{g}_1 \leftarrow \mathbf{g}(1 : n_{h,\Sigma}) \quad \text{et} \quad \mathbf{g}_2 \leftarrow \mathbf{g}(n_{h,\Sigma} + 1 : 2n_{h,\Sigma})$$

2 : **pour** $j = 1, 2$ **faire**3 : Calculer le *scattering* de \mathbf{g}_j

$$\mathbf{S}_j \leftarrow \text{Scattering}(\mathbf{g}_j)$$

4 : **fin pour**5 : Échanger les vecteurs de *scattering* et reformer un vecteur global

$$\tilde{\mathbf{g}} \leftarrow \begin{bmatrix} \mathbf{S}_2 \\ \mathbf{S}_1 \end{bmatrix}$$

6 : Former le vecteur voulu suivant (7.7)

$$\tilde{\mathbf{g}} \leftarrow \mathbf{g} - \tilde{\mathbf{g}}$$

7.1.3.1 ALGORITHME DE JACOBI

Nous pouvons maintenant écrire l'algorithme de Jacobi. Rappelons qu'il s'agit d'un algorithme de point fixe relaxé pour les traces entrantes qui s'écrit (6.55)

$$\begin{aligned} \mathbf{g}_{h,1}^{in,n+1} &= (1-r)\mathbf{g}_{h,1}^{in,n} + r\mathbf{g}_{h,2}^{out,n} \\ \mathbf{g}_{h,2}^{in,n+1} &= (1-r)\mathbf{g}_{h,2}^{in,n} + r\mathbf{g}_{h,1}^{out,n} \end{aligned} \quad (7.4)$$

On obtient l'algorithme 7. Comme pour tout processus itératif, nous avons besoin d'un critère d'arrêt pour stopper l'algorithme. Lorsque le problème est de taille raisonnable (en 2D), nous utiliserons comme critère de convergence l'erreur direct sur la solution \mathbf{u}_h . Dans la phase d'initialisation, nous résoudrons une première fois le système global (avec le solveur direct **PaStiX**) pour obtenir la solution \mathbf{u}_h^{ex} . Ensuite, à la fin de chaque itération de l'algorithme (ligne 10) nous mettons à jour le critère *res* selon la formule

$$res^2 = \frac{\|\mathbf{u}_{h,1}^{ex} - \mathbf{u}_{h,1}^n\|^2 + \|\mathbf{u}_{h,2}^{ex} - \mathbf{u}_{h,2}^n\|^2}{\|\mathbf{u}_h^{ex}\|^2} \quad (7.5)$$

où les solutions $\mathbf{u}_{h,j}^n$ sont calculées à partir de $\mathbf{g}_{h,j}^{in,n}$ par l'appel à la fonction **Solution**. De cette manière, *res* correspond à la norme L^2 de l'erreur relative sur

le domaine global Ω . Si le problème global est trop conséquent pour permettre le calcul de la solution exacte, alors nous utiliserons l'erreur inverse sur les traces entrantes $\mathbf{g}_{h,j}^{in,n}$ en calculant le résidu par

$$res = \frac{\|\mathbf{F} - \mathbb{A}\mathbf{g}\|}{\|\mathbf{F}\|} \quad (7.6)$$

Algorithme 7 Itérations - Algorithme de Jacobi

Requiert : Second membre \mathbf{F}_j , $j = 1, 2$, Tolérance ϵ .

Renvoie : Vecteurs des solutions $\mathbf{g}_{h,j}^i$, $i = 1, 2$.

1 : Initialiser $\mathbf{g}_{h,j}^{in,0} = 0$ pour $j = 1, 2$.

2 : $n \leftarrow 0$ et $res \leftarrow 1$

3 : **tant que** $res > \epsilon$ **faire**

4 : **pour** $j = 1, 2$ **faire**

5 : Calculer le *scattering* de $\mathbf{g}_{h,j}^{in,n}$

$$\mathbf{S}_j \leftarrow \text{Scattering}(\mathbf{g}_{h,j}^{in,n})$$

6 : Calculer la trace sortante $\mathbf{g}_{h,j}^{out,n}$

$$\mathbf{g}_{h,j}^{out,n} \leftarrow \mathbf{S}_j + \mathbf{F}_j$$

7 : **fin pour**

8 : Échanger les traces sortantes

$$\mathbf{g}_{h,1}^{in,n+1} \leftarrow \mathbf{g}_{h,2}^{out,n} \quad \text{et} \quad \mathbf{g}_{h,2}^{in,n+1} \leftarrow \mathbf{g}_{h,1}^{out,n}$$

9 : Appliquer la relaxation (7.4)

$$\mathbf{g}_{h,j}^{in,n+1} \leftarrow (1 - r) \mathbf{g}_{h,j}^{in,n} + r \mathbf{g}_{h,j}^{in,n+1}$$

10 : Calculer le résidu d'après (7.6) ou (7.5)

$$res = \sqrt{r_1^2 + r_2^2}, \quad \text{avec} \quad \begin{cases} r_1 = \|\mathbf{F}_1 - (\mathbf{g}_{h,1}^{in,n} + \mathbf{S}_2)\| / \|\mathbf{F}_1\| \\ r_2 = \|\mathbf{F}_2 - (\mathbf{g}_{h,2}^{in,n} + \mathbf{S}_1)\| / \|\mathbf{F}_2\| \end{cases}$$

11 : **fin tant que**

7.1.3.2 ALGORITHME GMRES

Alors que l'algorithme de Jacobi résout le problème comme un problème de point fixe (2.108), l'algorithme GMRES cherche à résoudre directement l'équation

vérifiée par la solution à la convergence de l'algorithme de Jacobi (2.109). Cette équation s'écrit, avec nos notations

$$\begin{aligned} \mathbf{g}_{h,1}^{in} - \text{Scattering}(\mathbf{g}_{h,2}^{in}) &= \mathbf{F}_1 \\ \mathbf{g}_{h,2}^{in} - \text{Scattering}(\mathbf{g}_{h,1}^{in}) &= \mathbf{F}_2 \end{aligned} \Leftrightarrow \mathbb{A} \mathbf{g}_h^{in} = \mathbf{F}. \quad (7.7)$$

Plus généralement, l'algorithme GMRES est un algorithme de résolution de système linéaire de type $Ax = b$. Ici, nos inconnues sont les $\mathbf{g}_{h,j}^{in}$, $j = 1, 2$, tandis que les seconds membres \mathbf{F}_j ont déjà été formés lors de la phase d'initialisation. Pour éviter de former explicitement la matrice du système linéaire (7.7), il suffit de fournir à l'algorithme GMRES une fonction `MatMult` (Algorithme 6) retournant l'action de la matrice sur un vecteur, c'est à dire `MatMult` implémente le produit matrice-vecteur de (7.7) sans construire explicitement la matrice \mathbb{A} .

L'objectif de l'algorithme GMRES est de minimiser à l'itération n la norme du résidu $r_n = b - Ax_n$ dans l'espace de Krylov $V_n = \text{span}(b, Ab, A^2b, \dots, A^nb)$. Pour cela, une base de cet espace est construite et agrandie au fur et à mesure des itérations. Si un grand nombre d'itérations est nécessaire, garder en mémoire cette base peut avoir un coût mémoire très important. Pour éviter ce problème, nous utilisons, de manière très classique, une variante de l'algorithme, dit algorithme GMRES avec *restart* m , dans laquelle l'algorithme est réinitialisé toutes les m itérations. Pour plus d'informations sur ces questions, nous renvoie le lecteur au livre de Yousef Saad [34] ainsi qu'aux articles [35] et [29]. Presque par construction, le critère d'arrêt de l'algorithme GMRES est basé sur l'erreur inverse (7.6). L'algorithme 8 décrit l'algorithme GMRES avec *restart*.

7.2 SPÉCIFICITÉS DE L'OPÉRATEUR Λ

Nous nous placerons ici dans le cas 2D où l'opérateur T est donné sous la forme $T = \Lambda^* \Lambda$, et Λ est construit à partir des potentiels de Riesz.

7.2.1 ASSEMBLAGE

Lors de l'assemblage de cet opérateur, nous sommes donc amenés à calculer des intégrales doubles du type

$$\int_{\Sigma} \int_{\Sigma} |x - y|^{\frac{1}{2}} \dot{\varphi}_i(x) \dot{\varphi}_j(y) dx dy, \quad (7.8)$$

Algorithme 8 Algorithme GMRES avec *restart* m . Problème de taille n

Requiert : Fonction **MatMult** calculant le produit matrice-vecteur, Vecteur solution initial \mathbf{x}_0 , second membre \mathbf{F} , Tolérance ϵ

Renvoie : Vecteur solution \mathbf{x}

- 1 : Calculer le résidu initial $\mathbf{r}_0 \leftarrow \mathbf{F} - \text{MatMult}(\mathbf{x}_0)$
- 2 : Calculer $\mathbf{v}_1 \leftarrow \mathbf{r}_0 / \|\mathbf{r}_0\|$
- 3 : **tant que** $\|\mathbf{r}_0\| > \epsilon$ **faire**
- 4 : **pour** $j = 1$ à m **faire**
- 5 : **pour** $i = 1$ à j **faire**
- 6 : Calculer le produit scalaire $h_{i,j} \leftarrow (\text{MatMult}(\mathbf{v}_j), \mathbf{v}_i)$
- 7 : **fin pour**
- 8 : Calculer un nouveau vecteur de la base de Krylov

$$\hat{\mathbf{v}}_{j+1} \leftarrow \text{MatMult}(\mathbf{v}_j) - \sum_{i=1}^j h_{i,j} \mathbf{v}_i$$

- 9 : Calculer sa norme $h_{j+1,j} \leftarrow \|\hat{\mathbf{v}}_{j+1}\|$
 - 10 : Le normer $\mathbf{v}_{j+1} \leftarrow \hat{\mathbf{v}}_{j+1} / h_{j+1,j}$
 - 11 : **fin pour**
 - 12 : Les $(h_{i,j})$ déterminent une matrice \mathbb{H} de taille $m + 1 \times m$ et les (\mathbf{v}_j) une matrice \mathbb{V} de taille $n \times m$.
 - 13 : Soit \mathbf{e}_1 la première colonne de la matrice identité de taille $m + 1 \times m + 1$
 - 14 : Soit \mathbf{y} le vecteur de taille m minimisant $\|\beta \mathbf{e}_1 - \mathbb{H} \mathbf{y}\|$. Cette résolution est faite par factorisation QR .
 - 15 : Calculer le nouveau vecteur solution approché $\mathbf{x}_0 \leftarrow \mathbf{x}_0 + \mathbb{V} \mathbf{y}$
 - 16 : Calculer le nouveau résidu $\mathbf{r}_0 \leftarrow \mathbf{F} - \text{MatMult}(\mathbf{x}_0)$
 - 17 : **fin tant que**
-

où les fonctions φ_i et φ_j sont les fonctions de base associées aux degrés de liberté i et j . Par simplification, nous supposons que l'espace d'approximation de $V_{h,\Sigma}$ est l'espace des fonctions P^1 . Dans ce cas, les dérivées $\dot{\varphi}_i$ et $\dot{\varphi}_j$ sont constantes, et l'intégrale se réduit à la somme de quatre intégrales du type

$$I = \int_K \int_L |x - y|^{\frac{1}{2}} dx dy. \quad (7.9)$$

où K et L sont des segments supports des fonctions φ_i et φ_j respectivement. Le calcul numérique précis de ces intégrales I n'est pas si évident, bien que l'intégrande soit continue. Pour effectuer ces intégrations, nous distinguerons plusieurs cas

- (i) Les segments K et L sont confondus.
- (ii) Les segments K et L ont un sommet S commun.
- (iii) Les segments K et L sont disjoints mais proches.
- (iv) Les segments K et L sont éloignés.

Si les segments sont disjoints (cas (iii) et (iv)), nous utiliserons pour l'intégration numérique la méthode de quadrature de Gauss-Legendre. Cette méthode consiste à remplacer l'intégrale par la somme

$$I \approx \sum_{p_i=1}^N \sum_{p_j=1}^N w_{p_i} w_{p_j} |x_{p_i} - y_{p_j}|^{\frac{1}{2}}. \quad (7.10)$$

où N est le nombre de points d'intégration utilisé, x_{p_i} et y_{p_j} sont les points d'intégrations (les zéros des polynômes de Legendre de degré N) et w_{p_i} et w_{p_j} sont les poids associés. Si les segments sont proches, nous utiliserons beaucoup de points d'intégration (par exemple 6 points) et si les segments sont éloignés, nous en utiliserons moins (3 points par exemple).

Si les segments sont confondus, on utilise la formule suivante

$$\int_0^h \int_0^h |x - y|^q dx dy = h^{q+2} \frac{2}{(q+2)(q+1)} \quad (7.11)$$

ce qui donne, avec $q = \frac{1}{2}$

$$I = \frac{8}{15} h^{\frac{5}{2}}. \quad (7.12)$$

Le cas le plus difficile à intégrer numériquement correspond au cas (ii). Dans ce cas, nous n'avons pas de formule exacte et le domaine d'intégration comprend le cas $x = y$. Nous allons utiliser les formules développées par Nicolas Salles et Marc Lenoir dans [36] qui permettent de réduire la dimension d'intégration tout en enlevant la singularité. Si le segment K (de longueur h_K) est donné par $[S, P_K]$ et L (de longueur h_L) par $[S, P_L]$, alors on a la formule suivante

$$\int_K \int_L |x - y|^{p-\frac{1}{2}} dx dy = \frac{2h_L}{2p+3} \int_K |x - P_L|^{p-\frac{1}{2}} dx + \frac{2h_K}{2p+3} \int_L |y - P_K|^{p-\frac{1}{2}} dy \quad (7.13)$$

ce qui donne, avec $p = 1$

$$I = \frac{2h_L}{5} \int_K |x - P_L|^{\frac{1}{2}} dx + \frac{2h_K}{5} \int_L |y - P_K|^{\frac{1}{2}} dy. \quad (7.14)$$

Comme $P_L \notin K$ et que $P_K \notin L$, les intégrandes sont maintenant parfaitement régulières et nous utiliserons une quadrature de type Gauss-Legendre pour finir l'intégration.

Ces techniques très puissantes d'intégration permettent d'obtenir des résultats précis très rapidement. Elles permettent également l'intégration de noyau plus singulier, et se généralisent aux dimensions supérieures. Néanmoins, ces techniques deviennent rapidement laborieuses à mettre en place lorsque pour des éléments d'ordre plus élevé.

7.2.1.1 VALIDATION SUR LE CERCLE

Pour valider l'implémentation de l'opérateur Λ , nous avons réalisé deux types d'expériences. Dans cette section, nous allons observer les valeurs propres de la matrice correspondant à la discrétisation de Λ sur le cercle. Ensuite, dans la section suivante, nous mettrons en évidence une des propriétés de l'opérateur Λ discret. Nous étudierons ici l'opérateur Λ_0 défini variationnellement par

$$\langle \Lambda_0 u, v \rangle = \sqrt{\frac{2}{\pi}} \int_{\Sigma} \int_{\Sigma} |x - y|^{\frac{1}{2}} \frac{\partial u(x)}{\partial \tau_x} \frac{\partial v(y)}{\partial \tau_y} dx dy \quad (7.15)$$

Rappelons que le symbole voulu pour l'opérateur Λ_0 doit se comporter comme $|s|^{\frac{1}{2}}$ (c'est un opérateur pseudo-différentiel d'ordre $\frac{1}{2}$). Comme précédemment, nous prendrons comme espace de discrétisation de $V_{h,\Sigma}$ l'espace des éléments finis P^1 . On notera φ_i les fonctions de bases (fonctions chapeaux) associées à cette discrétisation du cercle. Pour calculer les valeurs propres associées à l'opérateur discret Λ_h , nous commençons par construire les matrices

$$\begin{aligned} \mathbb{M}_{i,j} &= \int_{\Sigma_h} \varphi_i^h(x) \varphi_j^h(x) d\sigma(x) \\ \mathbb{L}_{i,j} &= \sqrt{\frac{2}{\pi}} \int_{\Sigma_h} \int_{\Sigma_h} \sqrt{|x - y|} \varphi_i^h(x) \varphi_j^h(y) d\sigma(y) d\sigma(x) \end{aligned} \quad (7.16)$$

\mathbb{M}_h correspond à la matrice de masse sur Σ et \mathbb{L}_h est la matrice associée à Λ_0 . Pour supprimer l'effet de la masse nous calculons la matrice $\tilde{\mathbb{L}}_h = \mathbb{M}_h^{-1} \mathbb{L}_h$. Finalement nous calculons les valeurs propres de la matrice $\tilde{\mathbb{L}}_h$. Cela revient à résoudre le problème aux valeurs propres généralisé $\mathbb{L}_h x = \lambda \mathbb{M}_h x$. Comme cette matrice est Toeplitz (de même que toutes les matrices de cette section), le calcul des valeurs propres s'effectue en calculant la transformée de Fourier d'une colonne de la matrice.

De manière assez classique, et afin d'accélérer les calculs, nous avons commencé par condenser la masse sur la diagonale. Dans notre cas (cercle discrétisé uniformément), la masse ainsi condensée devient la matrice $h Id$ où Id désigne la matrice identité. Les valeurs propres ainsi obtenues sont représentées figure 7.6.

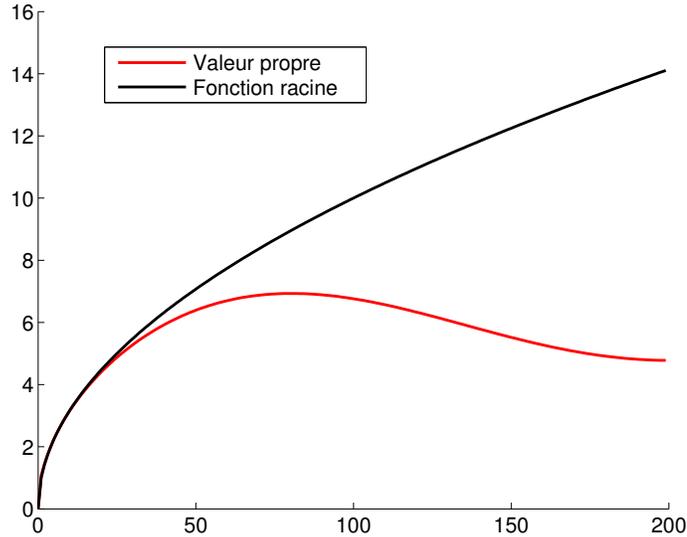


FIGURE 7.6 – Valeurs propres $\lambda_N(n)$ de l'opérateur discrétisé Λ_0 avec condensation de la masse versus le mode n . Nous avons également représenté la fonction racine carrée. On utilise $N = 401$ points de discrétisation sur le cercle.

De manière assez surprenante, les valeurs propres sont très éloignées de l'objectif d'approcher la fonction racine. Plus surprenant encore, la forme de cette courbe rouge dans la figure 7.6 semble assez “universelle”, au sens où elle est indépendante du nombre de points N utilisés pour discrétiser le cercle. Il semble exister une fonction F telle que les valeurs propres soient données par la fonction

$$\lambda_N(n) = \sqrt{N}F\left(\frac{n}{N}\right). \quad (7.17)$$

Après avoir recherché la cause du problème, nous nous sommes aperçu que l'écart important entre les valeurs propres et la fonction racine était dû à la condensation de la masse. En effet, si on prend en compte la masse exacte dans le calcul des valeurs propres, nous obtenons la figure 7.7. Encore une fois, cette courbe semble indépendante du nombre N de points utilisés. Cette expérience nous a permis de nous rendre compte de l'importance de garder la masse exacte dans la suite de nos calculs.

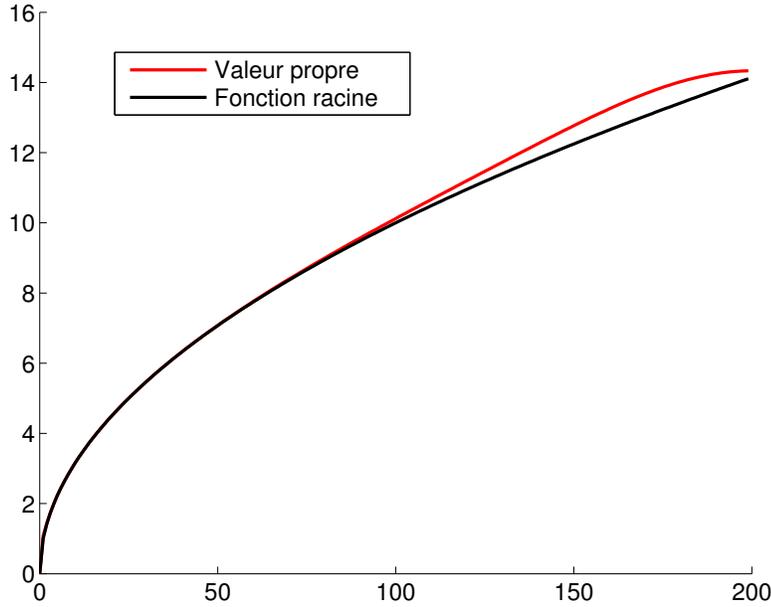


FIGURE 7.7 – Valeurs propres $\lambda_N(n)$ de l'opérateur discrétisé Λ_0 sans condensation de la masse versus le mode n . Nous avons également représenté la fonction racine carrée. On utilise $N = 401$ points de discrétisation sur le cercle.

7.2.1.2 VALIDATION DES PROPRIÉTÉS DE Λ

Dans ce paragraphe, nous nous proposons de retrouver expérimentalement les propriétés de l'opérateur Λ_0 (7.15)

$$\Lambda_0 \in \mathcal{I}som\left(H_{mo}^{\frac{1}{2}}(\Sigma), L_{mo}^2(\Sigma)\right) \quad (7.18)$$

ou encore comme Λ_0 est symétrique

$$\Lambda_0^* = \Lambda_0 \in \mathcal{I}som\left(L_{mo}^2(\Sigma), H_{mo}^{-\frac{1}{2}}(\Sigma)\right) \quad (7.19)$$

Pour cela, on va utiliser des espaces d'approximations et évaluer les constantes de continuité pour différentes géométries.

De par l'expression de Λ_0 , on voit directement que si $\mathbf{1}_\Sigma$ est la fonction constante sur Σ , on a $\Lambda_0 \mathbf{1}_\Sigma = 0$. Considérons le sous-espace (de codimension 1) de $H^{-\frac{1}{2}}(\Sigma)$

$$V = \left\{ \varphi \in H^{-\frac{1}{2}}(\Sigma), \langle \varphi, \mathbf{1}_\Sigma \rangle = 0 \right\} \quad (7.20)$$

et on cherche ici à vérifier numériquement l'éventualité de l'existence de deux constantes C_0 et C_1 telles que

$$C_0 \|u\|_{L^2}^2 \leq \|\Lambda_0 u\|_{H^{-\frac{1}{2}}}^2 \leq C_1 \|u\|_{L^2}^2, \quad \forall u \in V \quad (7.21)$$

en d'autres termes, on se demande si Λ_0 est un bien isomorphisme de V dans $L^2(\Sigma)$. Choisissons la norme suivante dans V

$$\|\Lambda_0 \varphi\|_{H^{-\frac{1}{2}}}^2 = \langle S\varphi, \varphi \rangle \quad (7.22)$$

où S est l'opérateur de simple couche ; si φ est assez régulière, on a

$$S\varphi(x) = \frac{1}{2\pi} \int_{\Sigma} \log|x-y| \varphi(y) d\sigma(y) \quad (7.23)$$

Il est connu (voir par exemple [30]) que cet opérateur S envoie $H^{-\frac{1}{2}}(\Sigma)$ dans $H^{\frac{1}{2}}(\Sigma)$, et donc le crochet de dualité dans la définition de la norme a bien un sens. Une démonstration de cette propriété montre tout d'abord que si φ est de moyenne nulle, le crochet est toujours positif. Ensuite, on montre que S est coercif sur $H^{-\frac{1}{2}}(\Sigma)$ si la longueur de la courbe est plus petite que 1, et reste coercif si on se restreint à V (c'est à dire aux fonctions de moyenne nulle). On peut alors réécrire les inégalités sous la forme

$$C_0 \|u\|_{L^2}^2 \leq \langle S \Lambda_0 u, \Lambda_0 u \rangle \leq C_1 \|u\|_{L^2}^2, \quad \forall u \in V \quad (7.24)$$

ou encore

$$C_0 \|u\|_{L^2}^2 \leq \langle \Lambda_0^* S \Lambda_0 u, u \rangle \leq C_1 \|u\|_{L^2}^2, \quad \forall u \in V \quad (7.25)$$

Pour tester numériquement ces estimations, on approche Λ_0 et S par une famille d'opérateurs Λ_h et S_h à l'aide d'un maillage de pas h de la courbe Σ , et un sous-espace de fonctions de dimension finie W_h de $H^{-\frac{1}{2}}(\Sigma)$. On suppose que la fonction constante est dans W_h et que $\Lambda_h \mathbf{1}_{\Sigma} = 0$. Ceci signifie en particulier que l'opérateur $K_h = \Lambda_h^* S_h \Lambda_h$ a pour valeur propre 0. Nous allons diagonaliser la famille d'opérateur K_h (on notera $\lambda_p(K_h)$ les valeurs propres de K_h). Comme l'opérateur Λ_0 que nous avons construit est bien un isomorphisme, les inégalités (7.21) doivent être vérifiées. On doit donc observer que toutes les valeurs propres non nulles sont bornées inférieurement et supérieurement, indépendamment de h

$$\begin{aligned} \lim_{h \downarrow 0} \min_{p, \lambda_p(K_h) \neq 0} \lambda_p(K_h) &> 0 \\ \lim_{h \downarrow 0} \max_{p, \lambda_p(K_h) \neq 0} \lambda_p(K_h) &< \infty \end{aligned} \quad (7.26)$$

En pratique, nous avons pris pour l'espace W_h l'espace des éléments finis P^1 sur une courbe Σ découpée en segments de longueur h . Nous avons ensuite formé les matrices

$$\begin{aligned}\mathbb{M}_{i,j} &= \int_{\Sigma_h} w_i^h(x) w_j^h(x) d\sigma(x) \\ \mathbb{L}_{i,j} &= -\sqrt{\frac{2}{\pi}} \int_{\Sigma_h} \int_{\Sigma_h} \sqrt{|x-y|} \dot{w}_i^h(x) \dot{w}_j^h(y) d\sigma(y) d\sigma(x) \\ \mathbb{S}_{i,j} &= -\frac{1}{2\pi} \int_{\Sigma_h} \int_{\Sigma_h} \log|x-y| \dot{w}_i^h(x) \dot{w}_j^h(y) d\sigma(y) d\sigma(x)\end{aligned}\tag{7.27}$$

où les indices i et j varient de 1 à N_h , le nombre de nœuds de Σ_h et w_h^i est la fonction de base (fonctions chapeaux) associée au nœud d'indice i . Enfin, nous avons formé la matrice

$$\mathbb{K}_h = (\mathbb{M}_h^{-1} \mathbb{L}_h)^T \mathbb{S}_h (\mathbb{M}_h^{-1} \mathbb{L}_h)\tag{7.28}$$

et résolu numériquement le problème de valeurs propres généralisé qui consiste à trouver $\lambda_p(K_h)$ et U_p tels que

$$\mathbb{K}_h U_p = \lambda_p(K_h) \mathbb{M} U_p.\tag{7.29}$$

Comme annoncé précédemment, nous trouverons toujours une valeur propre nulle (qui correspond à la fonction constante sur Σ).

Remarque 6. *On observera que par homogénéité des opérateurs, les valeurs propres seront toujours invariantes par changement d'échelles : $\lambda_p(\xi\Lambda_0) = \lambda_p(\Lambda_0)$. Ainsi, elles ne dépendent que de la forme de Σ et non de sa taille.*

Nous avons représenté sur la figure 7.8 pour différentes valeurs de h les valeurs propres trouvées lorsque la courbe Σ est un cercle. Notons que les courbes sont shiftées de telle sorte qu'elles ne se superposent pas. Dans chaque cas, on retrouve bien la valeur propre 0, puis toutes les autres valeurs propres se situent autour de 0.5 et sont bornées indépendamment du raffinement. Hormis la valeur propre nulle, aucune valeur propre ne tend vers 0 ou ne diverge avec h . Nous avons réalisé la même expérience pour d'autres courbes, dont une possède un coin. Les résultats sont présentés figures 7.9 et 7.10. Dans tous les cas, y compris lorsque Σ possède un coin, les valeurs propres sont quasiment indépendantes du raffinement, au sens où, hormis pour la valeur propre nulle (qui correspond aux fonctions constantes), les valeurs propres sont bornées indépendamment de h . Aucune valeur propre non

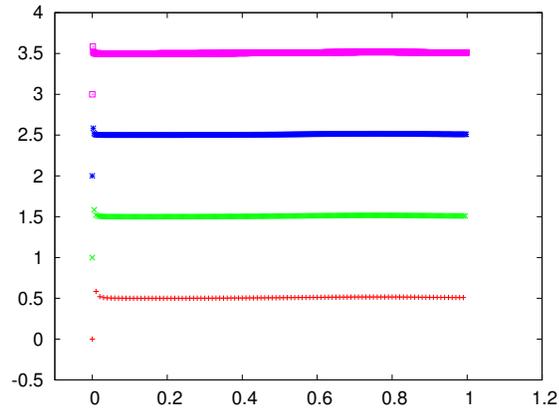


FIGURE 7.8 – Représentation des valeurs propres pour différents maillages du cercle. On représente $\frac{n}{N} \mapsto \lambda_n^N + a$. $a = 0, 1, 2, 3$ correspond à $N = 200, 400, 800, 1600$, N est le nombre de points du maillage. La valeur propre nulle est envoyée sur l'ordonnée a .

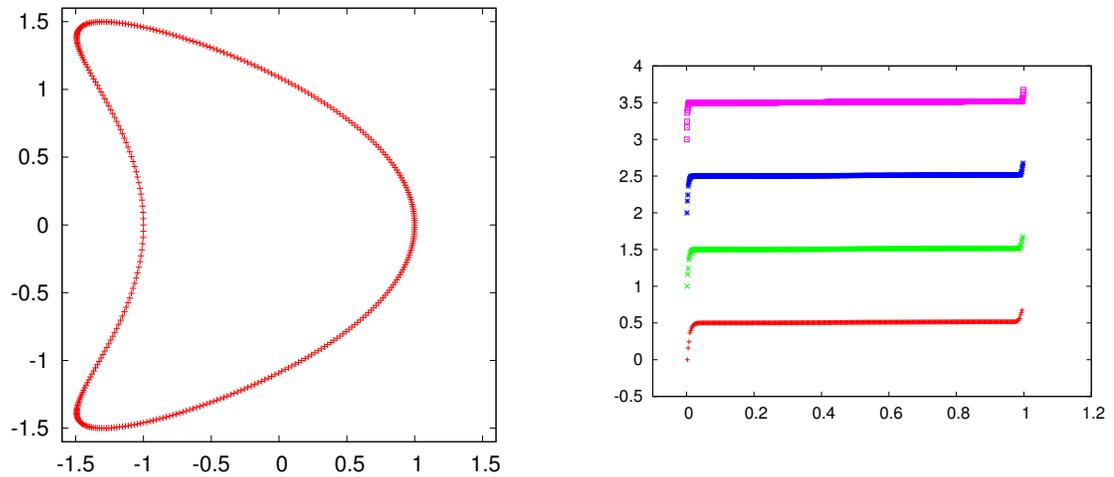


FIGURE 7.9 – Représentation des valeurs propres pour différents maillages du cerf-volant. On représente $\frac{n}{N} \mapsto \lambda_n^N + a$. $a = 0, 1, 2, 3$ correspond à $N = 200, 400, 800, 1600$, N est le nombre de points du maillage. La valeur propre nulle est envoyée sur l'ordonnée a .

nulle ne tend vers 0 ou ne diverge avec le raffinement.

Ceci valide notre implémentation de l'opérateur Λ_0 en tant qu'isomorphisme de $H^{\frac{1}{2}}(\Sigma)$ dans $L^2(\Sigma)$ (ou de $L^2(\Sigma)$ dans $H^{-\frac{1}{2}}(\Sigma)$).

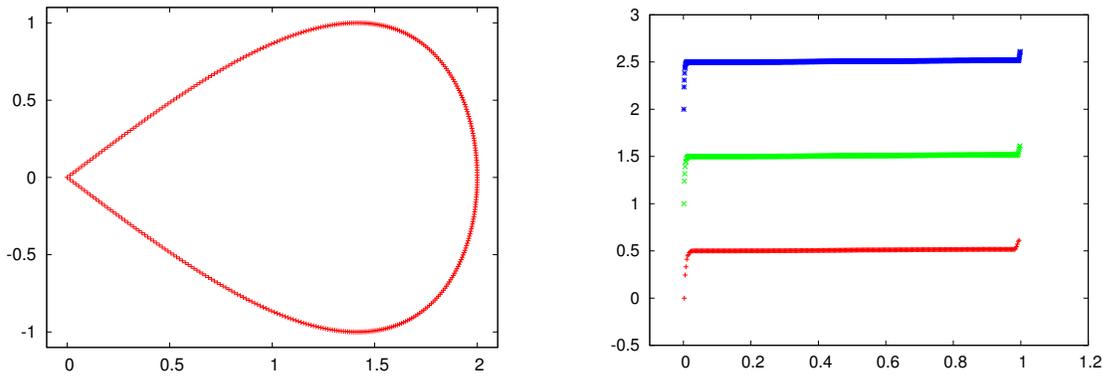


FIGURE 7.10 – Représentation des valeurs propres pour différents maillages de la goutte d'eau. On représente $\frac{n}{N} \mapsto \lambda_n^N + a$. $a = 0, 1, 2$ correspond à $N = 400, 800, 1600$, N est le nombre de points du maillage. La valeur propre nulle est envoyée sur l'ordonnée a .

7.2.2 RÉALISATION DE LA TRONCATURE

Dans cette section, nous allons étudier comment réaliser la troncature de l'opérateur défini à l'aide des potentiels de Riesz (4.9). La troncature pour cet opérateur est plus subtile que pour les autres opérateurs à notre disposition car il fait intervenir des dérivées tangentielles. Nous étudierons donc l'opérateur Λ_0 défini sur la courbe Σ d'abscisse curviligne x (supposée régulière) et dont la forme bilinéaire est

$$b(u, v) = \int_{\Sigma} \int_{\Sigma} |x(s) - x(t)|^{\frac{1}{2}} \partial_s u(x(s)) \partial_t v(x(t)) dx(s) dy(t). \quad (7.30)$$

Une première idée est d'appliquer directement la troncature χ sur le noyau intégral de telle sorte que la forme bilinéaire de l'opérateur tronqué $\Lambda_{0,\chi}$ vérifie

$$b_{\chi}(u, v) = \int_{\Sigma} \int_{\Sigma} \chi(|x(s) - x(t)|) |x(s) - x(t)|^{\frac{1}{2}} \partial_s u(x(s)) \partial_t v(x(t)) dx(s) dy(t), \quad (7.31)$$

avec $\chi(r) = \chi_0(r/\delta)$ où le paramètre δ est un paramètre de scaling gouvernant la largeur de la troncature. La fonction de troncature $\chi_0(r)$ définie pour $r \geq 0$ doit être choisie telle que

- elle n'influence pas le noyau si les points $x(s)$ et $x(t)$ sont proches : $\chi_0(r) = 1$ pour $r \leq r_0$
- elle annule le noyau si les points sont éloignés : $\chi_0(r) = 0$ pour $r \geq r_1$

- elle est suffisamment régulière. Pour $r_0 \leq r \leq r_1$, nous prendrons le polynôme d'ordre 5 tel que χ_0 soit c^2 . Ce polynôme (unique) s'écrit

$$\forall r \in [r_0, r_1], \quad \chi_0(r) = \frac{(r_1 - r)^3 [(r_1 - r)^2 + 5(r - r_0)(r + r_1 - 2r_0)]}{(r_1 - r_0)^5} \quad (7.32)$$

Dans la suite, nous prendrons $r_0 = \frac{1}{4}$ et $r_1 = 1$. La figure 7.11 représente la fonction χ_0 pour ces paramètres. Nous avons vu lors de la construction de cet

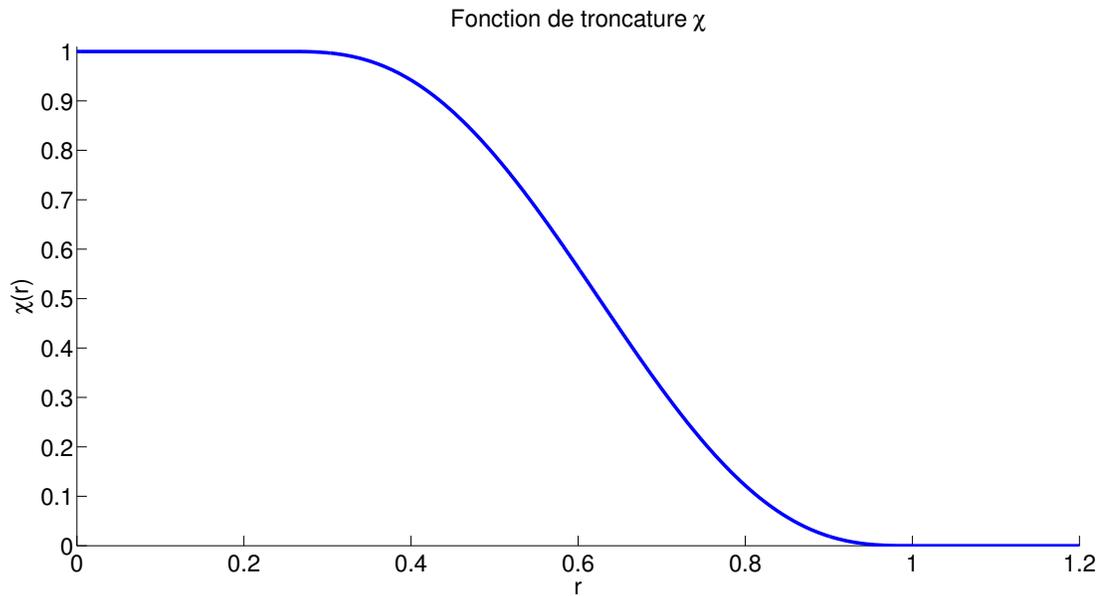


FIGURE 7.11 – Représentation de la fonction de troncature utilisée. La fonction est de classe c^2 .

opérateur, que la troncature ne modifie pas les propriétés théoriques de l'opérateur nécessaires aux démonstrations de la convergence exponentielle. Tous les résultats de la partie théorique et de la section concernant la discrétisation restent valables avec l'opérateur tronqué de cette manière. Cependant, nous avons observé numériquement que les convergences des algorithmes n'étaient pas à la hauteur de notre attente.

Nous avons donc tenté de comprendre pourquoi la convergence était dégradée avec cette manière de troncature. Pour analyser numériquement l'effet de la troncature sur l'opérateur Λ_0 , nous avons pris le cas d'une interface Σ circulaire, de rayon 1 discrétisé avec $N = 401$ points équidistants et nous avons construit la matrice

$\mathbb{L}_{0,\chi}$ associé à la discrétisation de $\Lambda_{0,\chi}$ sur cette interface pour des éléments finis d'ordre 1, avec comme largeur de troncature $\delta = 0.2$. Figure 7.12, nous avons représenté la colonne centrale des matrices en 7.12a (les matrices sont Toeplitz grâce à la géométrie de Σ) ainsi que leurs valeurs propres en 7.12b.

On peut voir que la troncature a une influence non négligeable sur la colonne de la matrice. La singularité (située sur la diagonale de la matrice) est bien préservée, et les éléments sont bien nuls suffisamment loin de la diagonale. Cependant, dans la zone intermédiaire, correspondant à la zone où la fonction de troncature χ n'est pas constante, les valeurs de la matrice sont fortement modifiées. Même si peu de ces valeurs sont fortement modifiées, les conséquences sur les valeurs propres de la matrice sont importantes et quasiment toutes les valeurs propres de $\mathbb{L}_{0,\chi}$ sont très différentes de \mathbb{L}_0 .

Le problème semble venir de la zone de transition de la fonction de troncature.

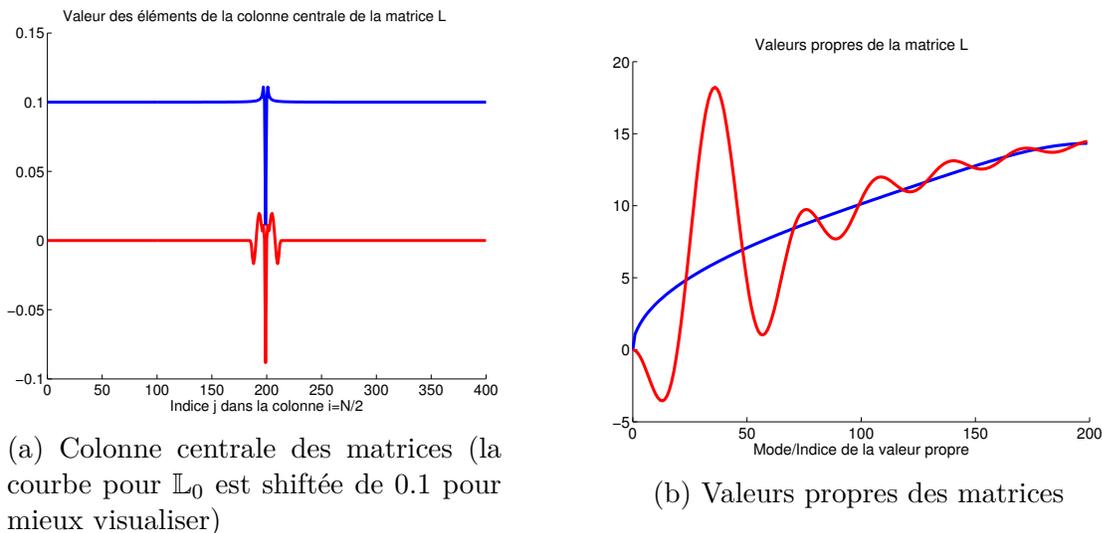


FIGURE 7.12 – Caractéristiques des matrices des opérateurs Λ_0 et $\Lambda_{0,\chi}$. En bleu, pour la matrice \mathbb{L}_0 non tronquée. En rouge, pour la matrice $\mathbb{L}_{0,\chi}$ tronquée.

Intuitivement, et à la place du comportement observé ici, on aurait plutôt envie d'appliquer la fonction de troncature directement sur la colonne de la matrice, pour éviter de trop la perturber tout en s'assurant que tout élément éloigné de la diagonal soit nul. Pour cela nous proposons deux autres approches pour tronquer cet opérateur.

7.2.2.1 APPROCHE NAÏVE

Introduisons la forme bilinéaire

$$\tilde{b}_\chi(u, v) = \int_\Sigma \int_\Sigma |x(s) - x(t)|^{\frac{1}{2}} \partial_s(\sqrt{\chi(s, t)} u(s)) \partial_t(\sqrt{\chi(s, t)} v(t)) dx(s) dy(t), \quad (7.33)$$

Dans ce cas de figure, la troncature n'est plus supportée par le noyau intégral, mais par les fonctions de base. Dans le cas d'une discrétisation par des éléments finis d'ordre 1, et en notant S_i^\pm et S_j^\pm les paires de segments supportant les fonctions de bases $\Phi_{\Sigma, i}$ et $\Phi_{\Sigma, j}$, alors l'élément de la matrice $\tilde{\mathbb{L}}_{0, \chi}$ est donné par la somme des quatre intégrales

$$\int_{S_i^{\epsilon_i}} \int_{S_j^{\epsilon_j}} |x(s) - x(t)|^{\frac{1}{2}} \partial_s(\sqrt{\chi(s, t)} \Phi_{\Sigma, i}(s)) \partial_t(\sqrt{\chi(s, t)} \Phi_{\Sigma, j}(t)) dx(s) dy(t), \quad (7.34)$$

pour $\epsilon_i = \pm$ et $\epsilon_j = \pm$. Une approche naïve consiste alors à négliger les variations de χ par rapport à celle des fonctions de base. Comme les variations de ces fonctions sont de l'ordre de h , où h est la distance entre deux points successifs de Σ , cette approche consiste à supposer que

$$h \frac{d}{dR} \sqrt{\chi(R)} \ll 1. \quad (7.35)$$

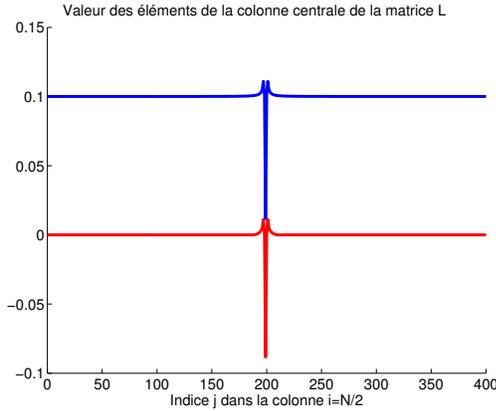
En utilisant cette approximation et en notant x_l le point associé au degré de liberté l , les éléments de la matrice deviennent

$$\tilde{b}_\chi(\Phi_{\Sigma, i}, \Phi_{\Sigma, j}) \approx \chi(|x_i - x_j|) \sum_{\epsilon_i = \pm} \sum_{\epsilon_j = \pm} \left(\int_{S_i^{\epsilon_i}} \int_{S_j^{\epsilon_j}} |x(s) - x(t)|^{\frac{1}{2}} \partial_s(\Phi_{\Sigma, i}(s)) \partial_t(\Phi_{\Sigma, j}(t)) dx(s) dy(t) \right) \quad (7.36)$$

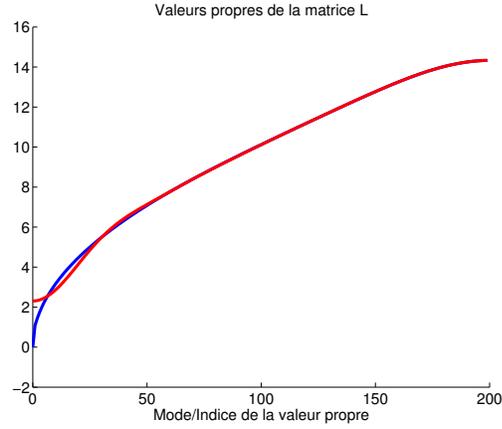
ou encore

$$\tilde{b}_\chi(\Phi_{\Sigma, i}, \Phi_{\Sigma, j}) \approx \chi(|x_i - x_j|) b(\Phi_{\Sigma, i}, \Phi_{\Sigma, j}). \quad (7.37)$$

Cette approche revient donc à tronquer directement la colonne de la matrice \mathbb{L}_0 sans troncature et l'implémentation est très simplifiée. C'est l'approche que nous avons retenu pour la suite des expériences numériques. Figure 7.13, nous avons une nouvelle fois représenté une colonne et les valeurs propres de la matrice $\tilde{\mathbb{L}}_{0, \chi}$.



(a) Colonne centrale des matrices (la courbe pour \mathbb{L}_0 est shiftée de 0.1 pour mieux visualiser)



(b) Valeurs propres des matrices

FIGURE 7.13 – Caractéristiques des matrices des opérateurs Λ_0 et $\tilde{\Lambda}_{0,\chi}$. En bleu, pour la matrice \mathbb{L}_0 non tronquée. En rouge, pour la matrice $\tilde{\mathbb{L}}_{0,\chi}$ tronquée.

7.2.2.2 APPROCHE PAR INTÉGRATION PAR PARTIES

La seconde approche consiste à appliquer la troncature après intégration par parties de la forme bilinéaire. Prenons K un noyau intégral et b_K la forme bilinéaire

$$b_K(u, v) = \int_{\Sigma} \int_{\Sigma} K(|x(s) - x(t)|) \partial_s u(x) \partial_t v(y) dx(s) dx(t). \quad (7.38)$$

Après intégration par parties, et en supposant K suffisamment régulier, on a

$$b_K(u, v) = \int_{\Sigma} \int_{\Sigma} \partial_s \partial_t K(|x(s) - x(t)|) u(x) v(y) dx(s) dx(t). \quad (7.39)$$

L'idée est d'appliquer la troncature à ce stade du calcul. On aurait alors

$$\hat{b}_{\chi}(u, v) = \int_{\Sigma} \int_{\Sigma} \chi(|x(s) - x(t)|) \partial_s \partial_t K(|x(s) - x(t)|) u(x) v(y) dx(s) dx(t). \quad (7.40)$$

Cependant, dans notre cas, le noyau est $K(R) = R^{\frac{1}{2}}$, $R = |x(s) - x(t)|$, et n'est pas suffisamment régulier pour écrire \hat{b}_{χ} sous cette forme. Nous allons donc chercher une forme bilinéaire correctrice $\chi(u, v)$ telle que

$$\hat{b}_{\chi}(u, v) = b_{\chi}(u, v) + c_{\chi}(u, v). \quad (7.41)$$

En notant $R = |x(s) - x(t)|$ et $\tilde{R} = \frac{x(s)-x(t)}{|x(s)-x(t)|}$, calculons $\partial_s \partial_t K(|x(s) - x(t)|)$

$$\partial_s K(R) = \left(\tilde{R} \cdot \dot{x}(s) \right) K'(R) \quad (7.42)$$

puis

$$\partial_s \partial_t K(R) = -\dot{x}(s) \cdot \dot{x}(t) \frac{K'(R)}{R} + \left(\tilde{R} \cdot \dot{x}(s) \right) \left(\tilde{R} \cdot \dot{x}(t) \right) R \frac{d}{dR} \frac{K'(R)}{R} \quad (7.43)$$

On cherche alors une fonction Υ telle que

$$\chi(R) \partial_s \partial_t K(R) = -\dot{x}(s) \cdot \dot{x}(t) \frac{(\chi K)'(R)}{R} - \left(\tilde{R} \cdot \dot{x}(s) \right) \left(\tilde{R} \cdot \dot{x}(t) \right) R \frac{d}{dR} \frac{(\chi K)'(R)}{R} + \Upsilon \quad (7.44)$$

et Υ serait le noyau intégral de la forme bilinéaire correctrice voulue. On a d'une part

$$\chi(R) \frac{K'(R)}{R} = \frac{(\chi K)'(R)}{R} - \frac{\chi'(R) K(R)}{R} \quad (7.45)$$

et en dérivant

$$\frac{d}{dR} \frac{\chi(R) K'(R)}{R} = \frac{d}{dR} \frac{(\chi K)'(R)}{R} - \frac{d}{dR} \frac{\chi'(R) K(R)}{R}. \quad (7.46)$$

D'autre part, on a

$$\chi(R) \frac{d}{dR} \frac{K'(R)}{R} = \frac{d}{dR} \frac{\chi(R) K'(R)}{R} - \chi'(R) \frac{K'(R)}{R}. \quad (7.47)$$

En combinant ces deux dernières équations, on trouve

$$\chi(R) \frac{d}{dR} \frac{K'(R)}{R} = \frac{d}{dR} \frac{(\chi K)'(R)}{R} - \frac{d}{dR} \frac{\chi'(R) K(R)}{R} - \chi'(R) \frac{K'(R)}{R}. \quad (7.48)$$

Alors, Υ vaut

$$\begin{aligned} \Upsilon(s, t) &= \dot{x}(s) \cdot \dot{x}(t) \frac{\chi'(R) K(R)}{R} \\ &\quad + \left(\tilde{R} \cdot \dot{x}(s) \right) \left(\tilde{R} \cdot \dot{x}(t) \right) R \left(\frac{d}{dR} \frac{\chi'(R) K(R)}{R} - \chi'(R) \frac{K'(R)}{R} \right) \end{aligned} \quad (7.49)$$

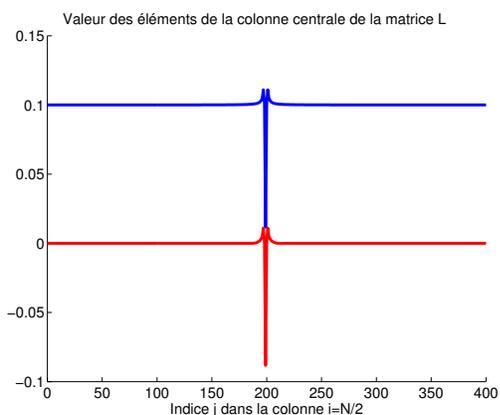
et

$$\hat{b}_\chi(u, v) = b_\chi(u, v) + \int_\Sigma \int_\Sigma \Upsilon(s, t) u(x(s)) v(x(t)) dx(s) dx(t). \quad (7.50)$$

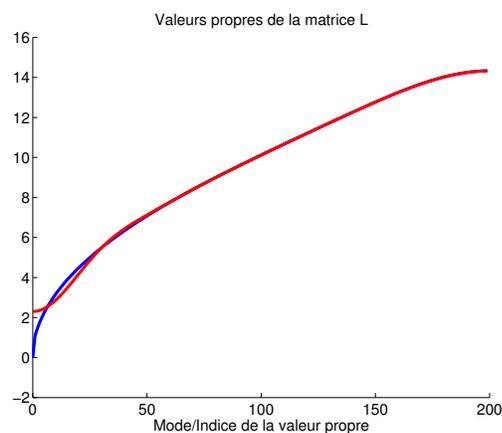
On peut voir d'après (7.49) que Υ ne fait intervenir que les dérivées de la fonction de troncature. Cette approche (7.50) pour tronquer l'opérateur Λ_0 correspond donc à l'approche directe (7.31) à laquelle on ajoute une correction basée sur les dérivées de la fonction de troncature. Écrite sous cette expression, la forme bilinéaire \widehat{b}_χ est bien définie car Υ est intégrable. En effet, Υ est nulle en tout point où la fonction de troncature χ est constante, et en particulier lorsque $x(s)$ et $x(t)$ sont proches ou confondus (c'est à dire là où le noyau K est singulier).

Nous avons implémenté cette approche et représenté figure 7.14 une colonne de la matrice corrigée $\widehat{\mathbb{L}}_0$ ainsi que ses valeurs propres, comparées à celles de la matrice originale \mathbb{L}_0 . Comme on peut le voir, cette approche par intégration par parties est très fortement similaire à l'approche naïve. Les oscillations que l'on observait sur les valeurs propres dans l'approche directe (voir figure 7.12b) sont parfaitement corrigées et les valeurs propres sont maintenant très semblables aux valeurs propres sans troncature.

Parmi les méthodes de troncature à notre disposition, nous préférons l'approche naïve qui, même si elle se base sur une approximation, permet de très peu modifier l'opérateur non tronqué, sans induire de surcoût car il suffit d'appliquer la troncature directement sur les colonnes de la matrice.



(a) Colonne centrale des matrices (la courbe pour \mathbb{L}_0 est shiftée de 0.1 pour mieux visualiser)



(b) Valeurs propres des matrices

FIGURE 7.14 – Caractéristiques des matrices des opérateurs Λ_0 et $\widehat{\Lambda}_{0,\chi}$. En bleu, pour la matrice \mathbb{L}_0 non tronquée. En rouge, pour la matrice $\widehat{\mathbb{L}}_{0,\chi}$ tronquée.

8

Résultats numériques

Sommaire

8.1	Cas de la géométrie circulaire	186
8.1.1	Étude de l'évolution du résidu pour plusieurs raffinements	187
8.1.1.1	$kR_1 = 10$	187
8.1.1.2	$kR_1 = 30$	195
8.1.2	Comparaison avec les valeurs propres analytiques . . .	201
8.1.2.1	Calculs des valeurs propres numériques . . .	201
8.1.2.2	Comparaison	201
8.1.3	Influence de la troncature	208
8.2	Cas du carré	210
8.3	Diverses géométries	216
8.3.1	Domaine étoilé	216
8.3.2	Cavité résonnante	220
8.4	Multi sous-domaines	222

Ce chapitre est consacré à l'analyse des résultats numériques. Dans une première section, nous ferons une analyse détaillée du cas de la géométrie circulaire simple. Nous étudierons ensuite le comportement des algorithmes que nous avons présentés sur d'autres géométries moins régulières.

8.1 CAS DE LA GÉOMÉTRIE CIRCULAIRE

Dans cette section, nous allons mettre en évidence quelques caractéristiques intéressantes des algorithmes itératifs proposés dans ce document. La géométrie correspond à celle de la géométrie circulaire simple décrite figure 3.1, avec $R_1 = 1$, $R_2 = 2$ et $R_\Sigma = 1.5$. Comme précédemment, en $r = R_1$, une condition de Dirichlet est imposée, et en $r = R_2$, nous imposons une condition absorbante d'ordre 1. Nous étudierons ce problème pour deux valeurs de k , telles que $kR_1 = 10$ et $kR_1 = 30$. Comme $R_1 = 1$, ceci correspond à $k = 10$ (soit une fréquence $f \approx 477 MHz$) et à $k = 30$ (soit $f \approx 1.43 GHz$). Nous avons regroupé dans les tables 8.1 et 8.2 les paramètres optimisés que nous avons utilisés pour l'opérateur et pour les conditions de type Després. Les optimisations ont été réalisées telles que décrites en section 3.2.3. Nous comparerons ces conditions de transmission optimisées avec

TABLE 8.1 – Paramètres pour le cas $kR_1 = 10$

	Jacobi		Gauss-Seidel		GMRES	
	Després	Opérateur	Després	Opérateur	Després	Opérateur
z	$1.53 + 0.02i$	0.217	$1.39 + 0.038i$	$0.26 - 0.01i$	$0.97 + 0.18i$	$0.36 - 0.017i$
β		2.86		2.04		0.898
r	0.54	0.56	0.59	0.61		

les conditions basées sur l'approximation de Padé complexe (1.18) développées par Yassine Boubendir, Xavier Antoine et Christophe Geuzaine dans [5]. Pour cette méthode, les paramètres utilisés sont les suivants

- Nombre de fractions $N_p = 8$
- Angle de la branche de la racine carré $\alpha = \frac{\pi}{4}$
- Paramètre d'amortissement $\varepsilon \approx 0.62$.

TABLE 8.2 – Paramètres pour le cas $kR_1 = 30$

	Jacobi		Gauss-Seidel		GMRES	
	Després	Opérateur	Després	Opérateur	Després	Opérateur
z	$1.32 + 0.03i$	$0.22 + 0.01i$	$1.29 - 0.04i$	$0.27 - 0.02i$	$0.685 + 0.05i$	$0.37 - 0.01i$
β		2.48		1.93		0.565
r	0.52	0.57	0.645	0.72		

Dans la suite, nous ferons référence aux conditions de transmission de Després optimisées par Després, celles utilisant un opérateur complètement non local $T = \Lambda^* \Lambda$ par *Opérateur*. Quand l'opérateur est localisé par la fonction de troncature, nous y ferons référence par *Opérateur tronqué*. Sauf mention contraire (notamment section 8.1.3, la longueur de la troncature sera prise égale à $\delta = 0.1$). Enfin, les conditions de [5] basées sur l'approximation de Padé complexe par *Padé*. Sur la figure 8.1, sont représentées les solutions aux problèmes considérés pour les deux valeurs de k .

Ces cas test vont nous permettre de mettre en évidence trois caractéristiques intéressantes de notre méthode de décomposition de domaine. Dans un premier temps nous allons étudier l'évolution de l'erreur (norme du résidu) au cours des itérations pour différents raffinements. Nous observerons ensuite les valeurs propres du problème itératif, que nous comparerons aux valeurs propres analytiques calculées dans le chapitre 3. Enfin, nous étudierons l'influence de la taille de la troncature sur les différents algorithmes.

8.1.1 ÉTUDE DE L'ÉVOLUTION DU RÉSIDU POUR PLUSIEURS RAFFINEMENTS

8.1.1.1 $kR_1 = 10$

Pour chacun des trois algorithmes itératifs suivants : Jacobi, Gauss-Seidel et GMRES, nous avons comparé les différentes conditions de transmission de Després, avec opérateur (tronqué ou non) et Padé. La figure 8.2 représente l'évolution des erreurs pour la méthode de Jacobi. Chacune des sous-figures 8.2a à 8.2d correspond à un type de conditions de transmission. De manière générale, et dans de nombreux cas, la convergence semble s'effectuer en deux étapes. Une première étape de convergence rapide, puis une seconde étape (fin de convergence) plus lente. Lors de la première étape de convergence, ce sont les modes bas de l'er-

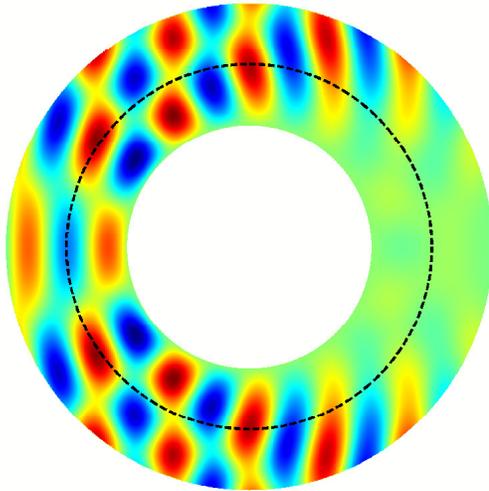
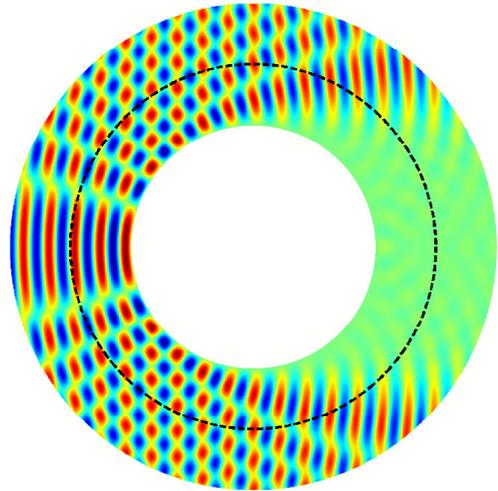
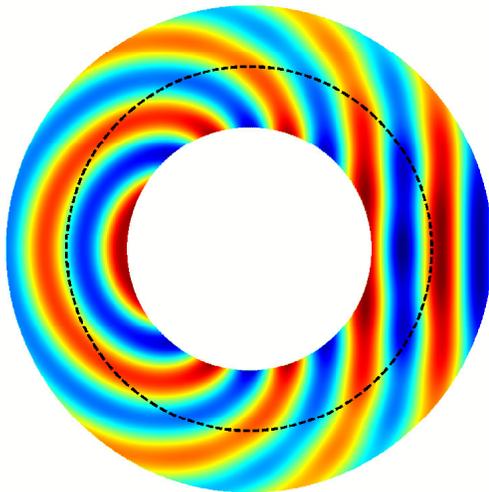
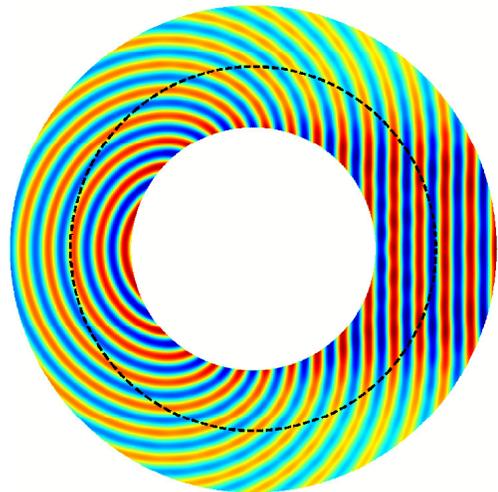
(a) Solution (en champ total) pour $kR_1 = 10$ ($f = 477 \text{ MHz}$)(b) Solution (en champ total) pour $kR_1 = 30$ ($f = 1.43 \text{ GHz}$)(c) Solution (en champ diffracté) pour $kR_1 = 10$ ($f = 477 \text{ MHz}$)(d) Solution (en champ diffracté) pour $kR_1 = 30$ ($f = 1.43 \text{ GHz}$)

FIGURE 8.1 – Représentation de la solution aux problèmes de diffraction du cercle pour différentes valeurs de k . En pointillé, est représentée l'interface de transmission Σ .

leur, correspondant aux ondes propagatives et pour lesquelles la solution possède beaucoup d'énergie, qui sont réduits. Dans la seconde étape de la convergence, ce sont les modes les plus élevés, plus difficiles à faire converger, et dont le taux de

convergence modal est plus élevé.

Pour la figure 8.2a, nous avons utilisé des conditions de type Després. On peut voir que l'algorithme stagne pour presque tous les raffinements de maillages. Le taux de convergence est très proche de 1 et les modes les plus élevés ne convergent pas. Seul le maillage le plus grossier (10 points par longueur d'onde) laisse espérer une convergence en un nombre raisonnable d'itérations. Ceci s'explique par le fait qu'un maillage grossier ne peut pas représenter les modes les plus élevés et pour lesquels des conditions de Després ne permettent pas une convergence.

L'utilisation de l'opérateur complètement non local en 8.2b permet d'éviter cette stagnation. Cette figure montre deux phénomènes différents. Tout d'abord, lorsque l'on raffine le maillage (le nombre de points par longueur d'onde augmente), le taux de convergence final (en fin de convergence, pour les modes élevés) tend vers une constante différente de 1. En effet les courbes des erreurs sont toutes parallèles. Ceci laisse penser que, quelque soit le raffinement, tous les taux de convergence modaux sont strictement inférieur à 1. Le second phénomène mis en évidence par 8.2b est que le nombre d'itérations nécessaire pour atteindre la convergence tend à diminuer avec le raffinement. Nous expliquerons plus en détail ce phénomène dans la section suivante relative aux comparaisons avec les taux modaux analytiques.

L'ajout de la troncature sur l'opérateur non local n'a que très peu d'influence sur la convergence (voir figure 8.2c). Pour les cas les plus raffinés (nombre de points par longueur d'onde supérieur à 30), l'influence n'est quasiment pas perceptible. Pour les cas moins raffinés, et pour cette situation, l'influence de la troncature semble même positive, le nombre d'itérations pour atteindre la convergence à 10^{-11} est plus faible avec troncature que sans troncature. Ce résultat, qui peut paraître contre-intuitif n'est en réalité pas en contradiction avec la théorie développée. En effet, si l'opérateur non local que l'on utilise vérifie un certain nombre de propriétés théoriques (isomorphisme, etc.) et que l'optimisation est réalisée pour cet opérateur non tronqué, rien n'indique que cet opérateur non tronqué soit l'opérateur optimal à utiliser. Il n'est donc pas impossible que la troncature, en plus de localiser l'opérateur et donc de réduire le temps de calcul, réduise aussi le nombre d'itérations à réaliser en améliorant le taux de convergence. Enfin, la figure 8.2d montre les courbes de convergence quand l'opérateur d'approximation par Padé de [5] est utilisé. Sur ce cas circulaire très basique, cet algorithme est très performant et surpasse nos conditions de transmission pour tous les raffinements. Ceci est d'autant plus spectaculaire pour les maillages peu raffinés. Il est néanmoins intéressant de voir que la dépendance du nombre d'itérations requis pour

atteindre la convergence est plutôt défavorable. Quand le nombre de point par longueur d'onde augmente, le nombre d'itérations augmente sensiblement aussi.

Ces résultats sont résumés dans la table 8.3, qui présentent le nombre d'itéra-

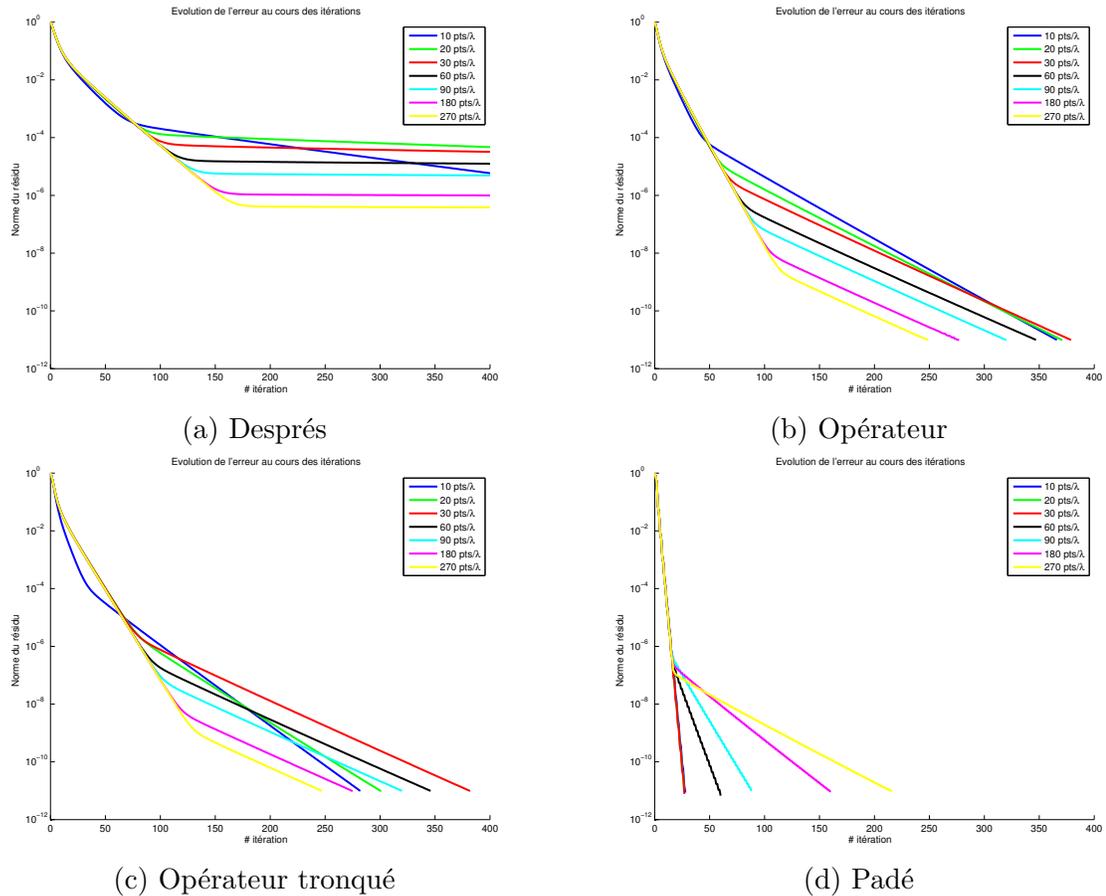


FIGURE 8.2 – Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme de Jacobi

tions requis par l'algorithme de Jacobi pour atteindre la convergence voulue pour chacune des situations décrites précédemment. Cette table montre bien le comportement des conditions de transmission vis à vis du pas du maillage (ou du raffinement). Alors que pour des conditions de type Padé, le nombre d'itérations augmente avec le raffinement, les conditions basées sur un opérateur non local (ou localisé par troncature) sont très peu sensible au raffinement. Le nombre d'itérations a même tendance à baisser avec le raffinement.

TABLE 8.3 – Nombre d’itérations de Jacobi requises pour atteindre la convergence (10^{-11}) pour $kR_1 = 10$

	Després	Opérateur	Opérateur tronqué	Padé
10 pts/ λ	>1500	366	282	28
20 pts/ λ	>1500	371	301	27
30 pts/ λ	>1500	379	382	27
60 pts/ λ	>1500	347	346	60
90 pts/ λ	>1500	320	320	88
180 pts/ λ	>1500	277	275	160
270 pts/ λ	>1500	249	247	216

Nous nous sommes ensuite intéressé à l’algorithme de Gauss-Seidel. Rappelons que cet algorithme est très similaire à l’algorithme de Jacobi, la seule différence étant que lors d’une itération, les sous-domaines ne sont plus résolus indépendamment mais l’un après l’autre, en utilisant les données aux interfaces qui ont été calculées le plus récemment. Séquentiel, cet algorithme a l’avantage de tirer parti des informations plus rapidement (dès qu’elles ont été calculées). On peut donc s’attendre à une convergence plus rapide. La figure 8.3 représente comme précédemment les courbes de convergence lors de l’utilisation de différentes conditions de transmission, et pour différents raffinements. On peut voir qu’effectivement, de manière générale, les algorithmes convergent plus vite que pour l’algorithme de Jacobi. Le nombre d’itérations est plus ou moins divisé par 2 par rapport au cas Jacobi.

L’algorithme de Gauss-Seidel étant basé sur le même principe que l’algorithme de Jacobi (à savoir un algorithme de point fixe), il est normal de retrouver toutes les conclusions que nous avons pu faire précédemment. Ainsi, on voit que l’algorithme de Gauss-Seidel utilisé avec des conditions de transmission de type Després (8.3a) stagne pour des raffinements élevés. Avec des conditions de transmissions basées sur nos opérateurs intégraux (tronqués ou non) (8.3b-8.3c), on voit clairement une amélioration de la convergence et le phénomène de stagnation disparaît. De plus, le taux de convergence final semble indépendant du raffinement (ou au moins borné strictement par 1) car les courbes sont parallèles en fin de convergence. En-

fin, les conditions de Padé restent très performantes puisqu'elles permettent une convergence plus rapide dans tous les cas de figure. Néanmoins, la même remarque s'applique ici. Le taux de convergence (et donc le nombre d'itérations requis pour obtenir la convergence) dépend du raffinement de maillage. Plus le maillage est fin (et donc le pas de maillage h faible), et plus la convergence requiert d'itérations. Nous avons une nouvelle fois résumé ces résultats dans la table 8.4.

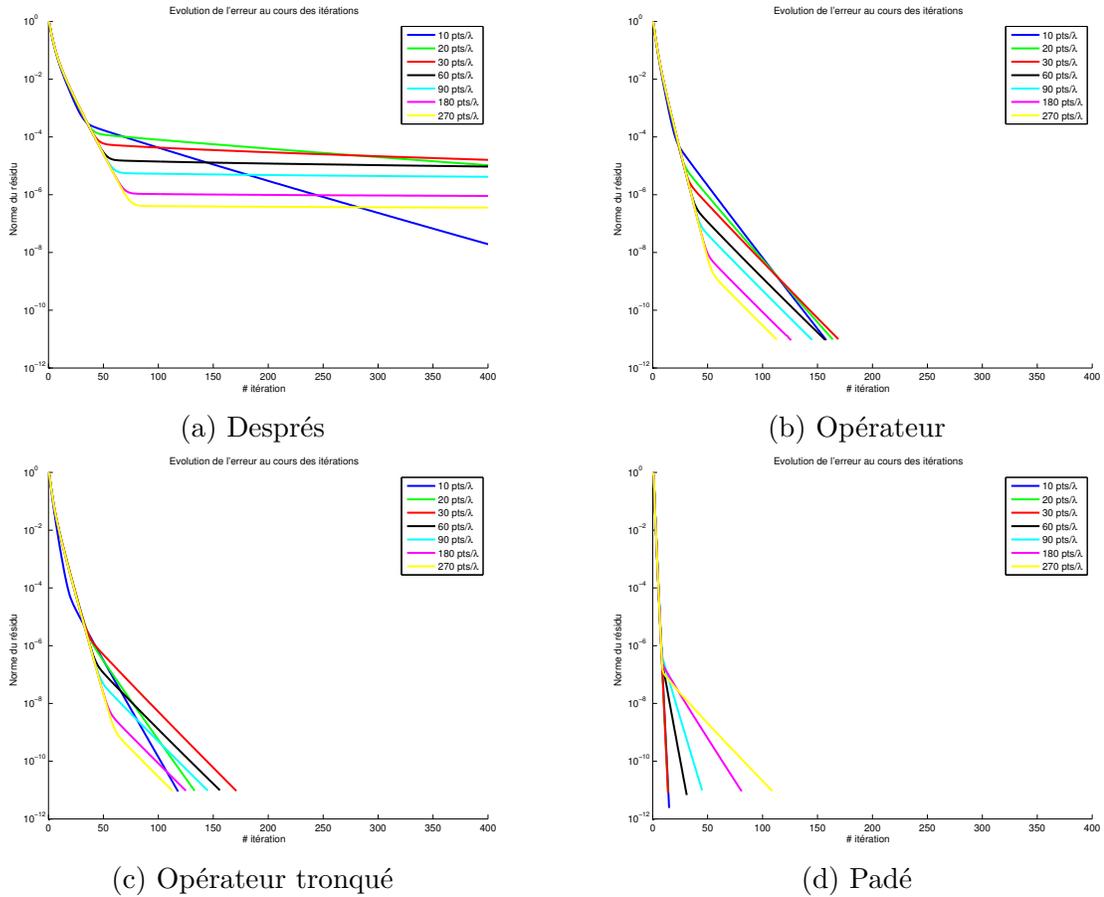


FIGURE 8.3 – Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme de Gauss-Seidel

TABLE 8.4 – Nombre d’itérations de Gauss-Seidel requises pour atteindre la convergence (10^{-11}) pour $kR_1 = 10$

	Després	Opérateur	Opérateur tronqué	Padé
10 pts/ λ	714	158	118	15
20 pts/ λ	>1500	164	133	14
30 pts/ λ	>1500	169	171	14
60 pts/ λ	>1500	157	156	31
90 pts/ λ	>1500	145	145	45
180 pts/ λ	>1500	126	125	81
270 pts/ λ	>1500	113	113	109

Finalement, nous avons réalisé les mêmes expériences pour l’algorithme GMRES. Cet algorithme a la particularité d’être très différent des algorithmes précédents (Jacobi et Gauss-Seidel) car il ne se base plus sur une méthode de point fixe pour la résolution du problème, mais sur les méthodes de Krylov. Dans ces méthodes, l’estimation du taux de convergence des algorithmes est, dans le cas général, plus complexe. Il est donc intéressant de regarder comment l’algorithme GMRES se comporte avec différentes conditions de transmission. Les courbes de convergence sont représentées figure 8.4.

Une première constatation, assez flagrante, est que cet algorithme est nettement plus performant que les deux précédents (notons que l’intervalle de l’axe des abscisse a été divisé par 2). Ceci n’est pas vraiment surprenant car l’algorithme de Jacobi est un algorithme relativement basique tandis que l’algorithme GMRES réalise, à chaque itération, une minimisation du résidu sur l’espace de Krylov. Une seconde observation serait que l’utilisation de conditions de transmission de Després n’entraîne plus la stagnation de l’algorithme. Même avec ces conditions de transmission, l’algorithme converge. Cependant, la convergence dépend clairement du raffinement du maillage. Encore une fois, plus le maillage est fin, est plus la convergence est lente. Le même comportement se retrouve lorsque les conditions de type Padé sont utilisées. Bien que moins flagrant, le nombre d’itérations augmente avec le raffinement.

Dans le cas de l'opérateur non local (possiblement tronqué), la convergence de l'algorithme de GMRES est quasiment indépendante du raffinement du maillage. De plus, la troncature n'a presque aucun effet (ou alors un effet positif) sur la convergence, par rapport à l'opérateur non tronqué. Ces résultats sont très intéressants car ils montrent d'une part que les résultats théoriques que nous avons démontrés sur le problème continu se retrouvent après discrétisation et implémentation. D'autre part, ces résultats pour le GMRES montrent que les conditions de transmission utilisées influencent aussi les méthodes de Krylov, et notamment l'utilisation des opérateurs intégraux que nous avons développés permet d'obtenir une convergence de ces méthodes quasi indépendante du pas du maillage. Encore

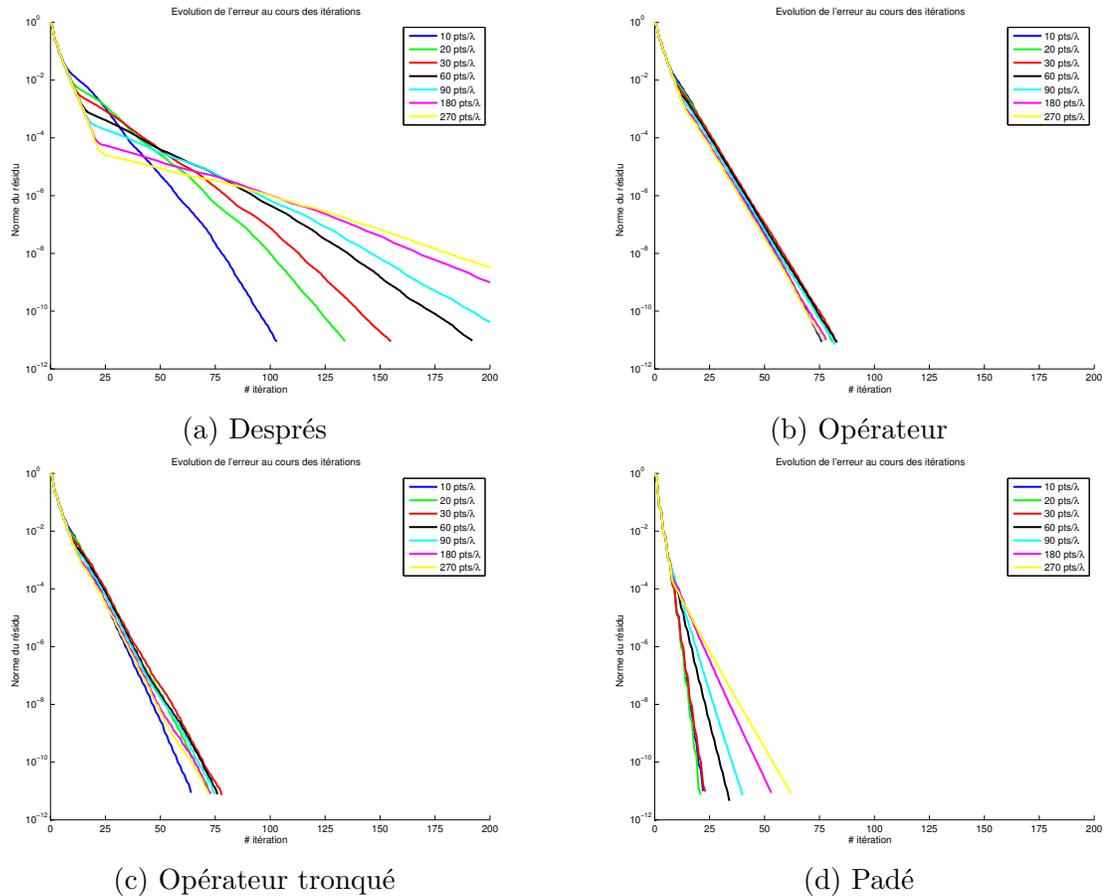


FIGURE 8.4 – Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme GMRES

une fois, la table 8.5 synthétise le nombre d'itérations nécessaire à l'algorithme GMRES pour obtenir la convergence et montre bien les résultats décrits ci-dessus.

TABLE 8.5 – Nombre d'itérations de GMRES requises pour atteindre la convergence (10^{-11}) pour $kR_1 = 10$

	Després	Opérateur	Opérateur tronqué	Padé
10 pts/ λ	103	76	64	22
20 pts/ λ	134	81	73	21
30 pts/ λ	155	83	78	23
60 pts/ λ	192	83	76	34
90 pts/ λ	214	82	75	40
180 pts/ λ	258	78	73	53
270 pts/ λ	285	77	72	62

8.1.1.2 $kR_1 = 30$

Nous reprenons ici les mêmes expériences que dans la section précédente, mais pour une valeur de k différente. Ici $k = 30$, ce qui correspond à une fréquence $f = 1.47 GHz$. Nous allons montrer que les résultats obtenus précédemment restent valables. Nous avons utilisé les mêmes maillages que dans le cas $kR_1 = 10$, donc d'après l'équation (7.1), le nombre de points par longueur d'onde est divisé par 3. C'est pourquoi ici le nombre de points par longueur d'onde maximal est de 90. Commençons par l'algorithme de Jacobi. La figure 8.5 présente les courbes de convergence pour cet algorithme et pour chacune des conditions de transmission testées. Les courbes sont très similaires au cas précédent avec fréquence plus faible. En effet, les mêmes remarques s'appliquent ici. Ainsi, l'algorithme de Jacobi avec conditions de transmission de type Després peine à converger pour des raffinements supérieur à 10 points par longueur d'onde. Au contraire, lors de l'utilisation de conditions de transmission avec opérateur non local, l'algorithme converge et on n'a plus de phénomène de stagnation. De plus, le taux de convergence (en fin de convergence) est quasi indépendant du raffinement du maillage (les courbes deviennent parallèles). Comme précédemment, les conditions de transmission basées

sur l'approximation de Padé sont particulièrement efficaces pour la résolution de ce problème. Cependant, même si elles permettent une convergence très rapide, leur taux de convergence se dégrade avec le raffinement. L'effet est moins visible que précédemment car nous n'avons pas pu raffiner davantage le maillage. Nous

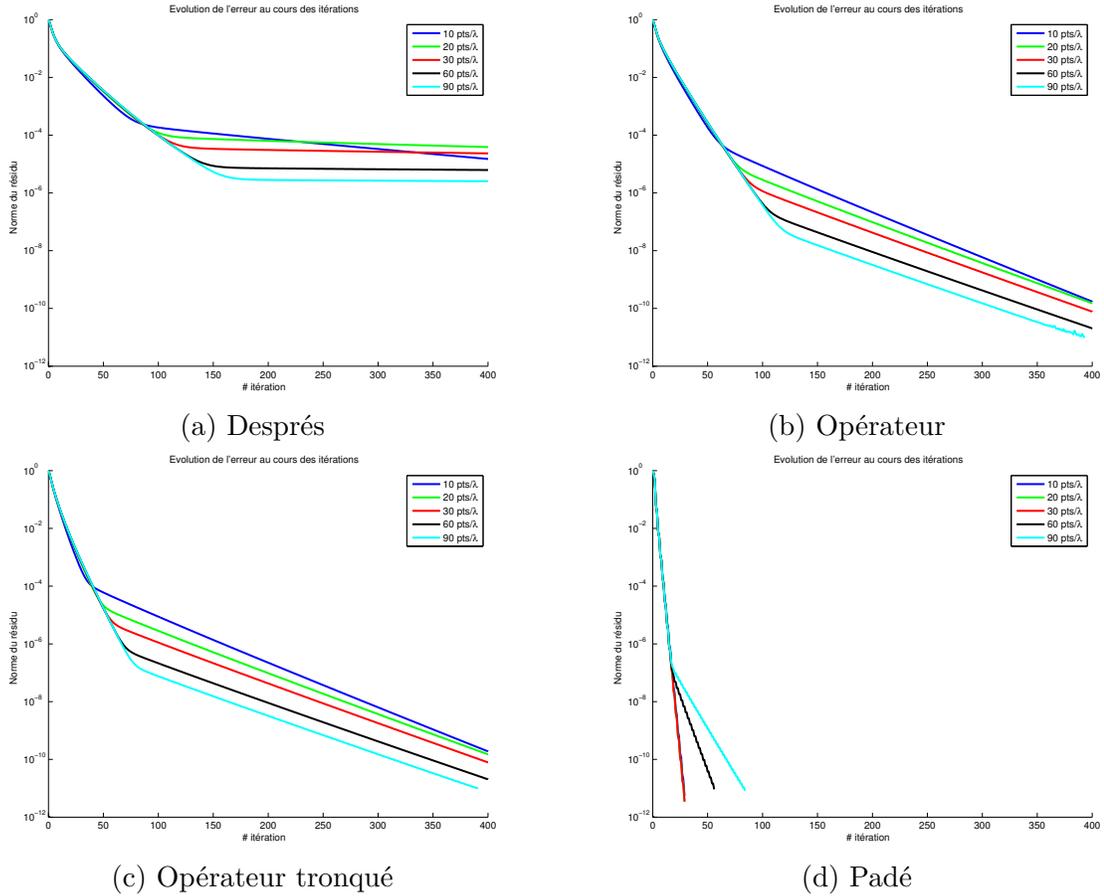


FIGURE 8.5 – Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme de Jacobi

résumons les performances des différentes situations pour l'algorithme de Jacobi dans la table 8.6.

TABLE 8.6 – Nombre d’itérations de Jacobi requises pour atteindre la convergence (10^{-11}) pour $kR_1 = 30$

	Després	Opérateur	Opérateur tronqué	Padé
10 pts/ λ	>1500	482	485	29
20 pts/ λ	>1500	487	486	29
30 pts/ λ	>1500	467	467	29
60 pts/ λ	>1500	425	424	56
90 pts/ λ	>1500	393	391	84

Pour l’algorithme de Gauss-Seidel, qui n’est qu’une variante de l’algorithme de Jacobi, les conclusions sont bien évidemment les mêmes. Cet algorithme se comporte comme dans le cas précédent, et comme l’algorithme de Jacobi. En effet, l’algorithme de Gauss-Seidel converge environ deux fois plus rapidement que l’algorithme de Jacobi. De plus, alors que l’utilisation de conditions de transmission de type Després font stagner l’algorithme, les conditions basées sur les opérateurs complètement non locaux ou localisés grâce à la troncature rendent l’algorithme de Gauss-Seidel quasiment indépendant du raffinement du maillage. Une nouvelle fois, l’algorithme basé sur les conditions de type Padé surpasse toutes les autres conditions de transmission, même si la même remarque peut être faite concernant sa dépendance au raffinement du maillage.

TABLE 8.7 – Nombre d’itérations de Gauss-Seidel requises pour atteindre la convergence (10^{-11}) pour $kR_1 = 30$

	Després	Opérateur	Opérateur tronqué	Padé
10 pts/ λ	814	159	160	15
20 pts/ λ	>1500	164	164	15
30 pts/ λ	>1500	159	159	15
60 pts/ λ	>1500	145	146	29
90 pts/ λ	>1500	136	134	43

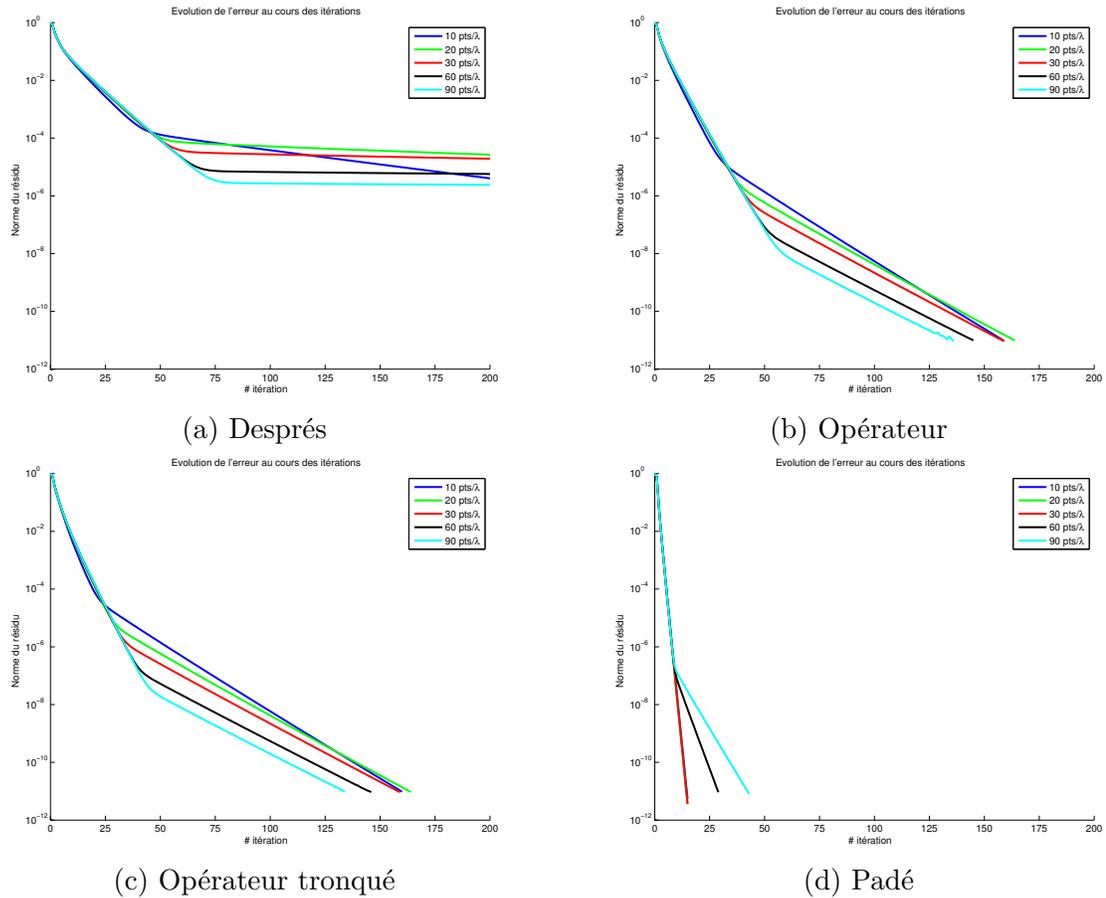


FIGURE 8.6 – Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme de Gauss-Seidel

Finalement, pour l'algorithme GMRES, et alors qu'il ne constitue pas une variante de l'algorithme de Jacobi, son comportement est là encore assez similaire à l'algorithme de Jacobi (voir le récapitulatif pour GMRES table 8.8).

De manière générale, l'algorithme de GMRES constitue un excellent solveur itératif pour les problèmes étudiés, convergeant nettement plus rapidement que les méthodes de point fixe comme Jacobi ou Gauss-Seidel. Nous avons aussi vu, sur les exemples proposés, que les conditions de transmission que nous avons développées dans ce document permettent de rendre les algorithmes itératifs (de type point fixe, mais aussi basés sur les méthodes de Krylov) quasiment indépendants du raffinement du maillage, ce que ne permettent pas les conditions de transmis-

sion locales comme celle de type Després ou Padé. Néanmoins, sur les cas basiques de cette section et les raffinements proposés, les conditions de type Padé restent plus performantes.

Nous retrouvons en tous cas les propriétés théoriques très intéressantes de nos conditions de transmission au niveau numérique.

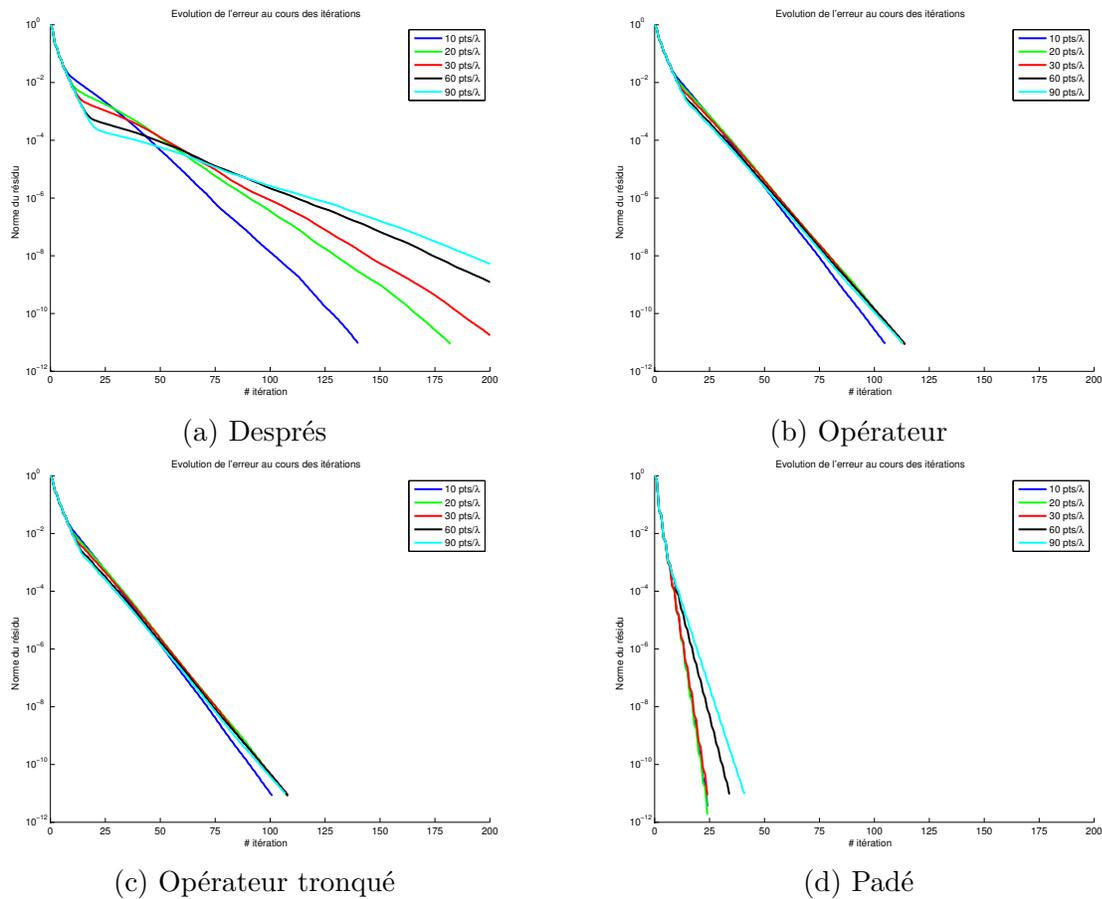


FIGURE 8.7 – Évolution des résidus pour différents raffinements et différents types de condition de transmission. Algorithme GMRES

TABLE 8.8 – Nombre d’itérations de GMRES requises pour atteindre la convergence (10^{-11}) pour $kR_1 = 30$

	Després	Opérateur	Opérateur tronqué	Padé
10 pts/ λ	140	105	101	24
20 pts/ λ	182	113	108	24
30 pts/ λ	205	114	108	24
60 pts/ λ	254	114	108	34
90 pts/ λ	286	113	107	41

Il est important de noter que nous avons utilisé, pour les simulations utilisant les conditions de type Padé, un nombre constant de fractions rationnelles $N = 8$. Ce nombre est déjà relativement élevé et permet à ces conditions de transmission de ne pas trop ressentir la dépendance en h . En effet, comme expliqué section 1.2.2.3, l’approximation de Padé consiste à approcher un opérateur non local du type $\sqrt{1 + \partial_t^2}$ par des fractions rationnelles. Plus le nombre de fractions rationnelles est élevé, et mieux l’opérateur *racine carrée* est approché, particulièrement pour les modes très évanescents. Ainsi, réduire N dans nos expériences conduirait à une plus grande sensibilité de la méthode au raffinement. A l’inverse, nous intuitons qu’il est possible de trouver $N = N(h)$ de telle sorte que la sensibilité des conditions de transmission de type Padé au raffinement ne soit plus visible.

Néanmoins, agir ainsi revient à adapter (et donc changer) l’opérateur à chacun des maillages. De plus, l’ajout de fractions rationnelles a un coût. En effet, chaque fraction induit une augmentation de la taille de la matrice locale puisque le dénominateur induit un problème linéaire à résoudre. Ainsi, chaque matrice ajoute $n_{h,\Sigma}$ lignes et colonnes dans la matrice locale. Sur certains cas où le nombre de points sur Σ est important, mettre trop de fractions rationnelles conduirait à des problèmes locaux trop importants (voire même importants que le problème global initial). Pour toutes ces raisons, nous avons décidé de conserver un nombre de fractions rationnelles constant dans toutes nos expériences numériques.

8.1.2 COMPARAISON AVEC LES VALEURS PROPRES ANALYTIQUES

8.1.2.1 CALCULS DES VALEURS PROPRES NUMÉRIQUES

Pour mieux comprendre les résultats numériques concernant les courbes de convergence que nous avons présentées précédemment, nous allons dans cette section étudier les valeurs propres numériques des matrices d'itération et les comparer aux valeurs propres analytiques dont les expressions sont données au chapitre 3 d'Analyses Modales. La comparaison est effectuée dans le cas de l'algorithme de Jacobi, et pour des conditions de transmission de type Després ou opérateur complètement non local uniquement. Il s'agit d'étudier les valeurs propres de la matrice correspondant à l'opérateur $(1 - r)I + r S_{h,\pi}$ (voir notamment les équations (2.71), (3.22) ou encore (6.25) pour les définitions sur le problème continu, modal et discret, respectivement). Dans le chapitre 7 concernant la mise en œuvre numérique, nous avons mis en place un algorithme permettant de renvoyer, étant donné un vecteur d'entrée \mathbf{g} , le vecteur correspondant au produit $(I - S_{h,\pi}) \mathbf{g}$. C'est l'algorithme 6 implémentant la fonction `MatMult`. En réutilisant cet algorithme, il est aisé d'obtenir le produit $S_{h,\pi} \mathbf{g}$. Dès lors, pour obtenir la colonne j de $S_{h,\pi}$, il suffit d'appeler cette fonction avec comme vecteur d'entrée le vecteur \mathbf{g}_j tel que $(\mathbf{g}_j)_i = \delta_{i,j}$, où $\delta_{i,j}$ est le symbole de Kronecker. Pour construire explicitement la matrice $S_{h,\pi}$, on itère sur tous les vecteurs \mathbf{g}_j , $j = 1, \dots, 2n_{h,\Sigma}$. Une fois la matrice construite explicitement, nous faisons appel à la bibliothèque `LAPACK` [1] pour le calcul des valeurs propres.

Une fois les valeurs propres $\lambda_{S,m}$ de $S_{h,\pi}$ obtenues, il est facile de calculer les taux modaux associés, définis par $|1 - r + r \lambda_{S,m}|$ et de les comparer aux taux modaux analytiques, que nous avons calculés dans le chapitre d'analyse modale. Ici, nous n'étudierons que le cas $k = 10$, les résultats étant identiques pour $k = 30$.

8.1.2.2 COMPARAISON

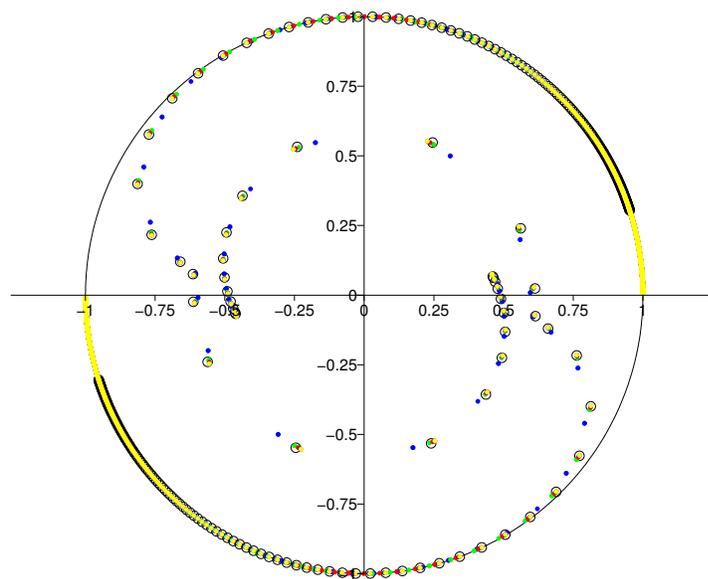
Représentons tout d'abord les valeurs propres dans le plan complexe. La figure 8.8 représente les valeurs propres analytiques (rond) et les valeurs propres numériques (points) pour les différents raffinements. La figure 8.8a correspond aux valeurs propres dans le cas où les conditions de transmission de type Després sont utilisées, tandis que pour 8.8b, ce sont les valeurs propres dans le cas de l'opérateur non local. Il est clair que les valeurs propres numériques sont très proches, voir confondues avec les valeurs propres analytiques. On peut voir également que plus on raffine le maillage, et plus les valeurs propres numériques sont proches des valeurs propres analytiques. Les valeurs propres numériques convergent vers

les valeurs propres analytiques. Évidemment, ce comportement était attendu. En effet, plus h est petit et plus on s'approche du problème continu. Notons que dans le cas des conditions de Després, les valeurs propres tendent vers 1. En fait, elles s'approchent de plus en plus vers 1 avec le raffinement. À l'inverse, avec les conditions de transmission basées sur les opérateurs non locaux, les valeurs propres restent éloignées de 1. On retrouve les propriétés des lemmes 4 et 5 sur les propriétés des opérateurs de scattering modaux et de l'analyse modale de convergence. Ces propriétés restent vraies numériquement après discrétisation.

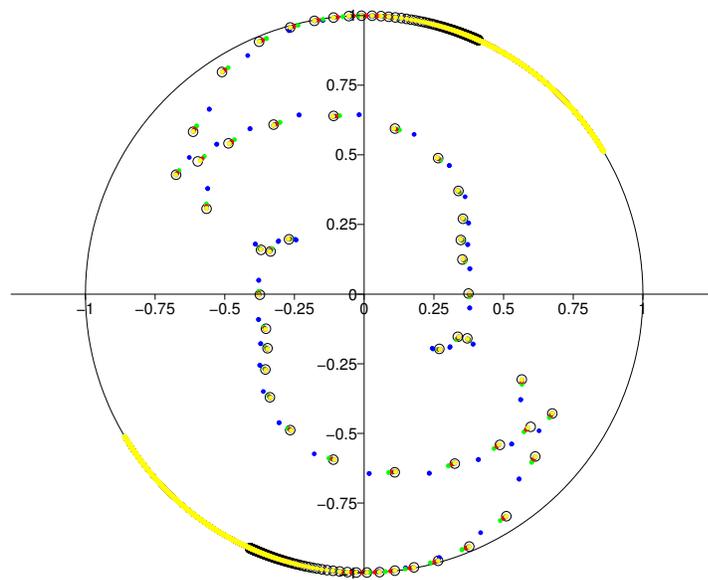
Sur les figures suivantes (figures 8.9 à 8.12), nous nous intéressons aux taux de convergence modaux, c'est à dire à $\tau_{r,m} = |1 - r + r\lambda_m|$ où r est le paramètre de relaxation. Le nombre de modes pour un maillage fixé correspond au nombre de points sur l'interface Σ . Donc plus le maillage est fin, et plus il y aura de modes représentés dans la solution associée à ce maillage. Comme nous analysons des raffinements très différents, et pour ne pas écraser les courbes, les figures 8.9 et 8.10 correspondent à la restriction des figures 8.11 et 8.12 sur la bande de modes $[0, 150]$. Ces deux figures permettent de voir que plus le maillage est fin, et plus les taux de convergence modaux sont proches des taux analytiques. De plus, la figure 8.11, représentant les taux pour les conditions de transmission de Després, permet d'expliquer facilement pourquoi la méthode de Jacobi stagne pour ces conditions de transmission. En effet, les taux modaux tendent très rapidement vers 1 (tout en restant strictement inférieur à 1), et il existe de nombreux modes pour lesquels l'algorithme ne converge pas.

À l'inverse, lorsqu'on utilise les conditions avec opérateurs non locaux, la figure 8.12 montre qu'aucun des taux modaux ne s'approche de 1. On retrouve aussi le fait que plus on raffine le maillage, et plus les taux numériques s'approchent des taux analytiques. Néanmoins, on observe ici un comportement particulier. En effet, si les taux modaux tendent, mode par mode, vers les taux analytiques, ce n'est pas le cas de leur maximum. On a seulement une convergence faible des taux numérique vers les taux analytiques quand le raffinement augmente (ou que h tend vers 0). Néanmoins, une caractéristique très intéressante des courbes présentées en 8.12 est que pour chaque raffinement, le maximum des taux de convergence modaux semble à peu près constant et donc indépendant du pas du maillage h .

Dans le chapitre 6 concernant la discrétisation de la méthode, nous ne sommes pas parvenu à montrer l'indépendance en h du taux de convergence qui intervient dans le théorème 6. Les expériences numériques semblent pourtant suggérer une telle indépendance.



(a) Conditions de Després



(b) Opérateur non local

FIGURE 8.8 – Comparaison des valeurs propres numériques et analytiques pour le cas de la géométrie circulaire ($kR_1 = 10$). Les valeurs propres analytiques sont représentées par des cercles noirs, les valeurs propres numériques par des points de couleurs : En bleu 10 points par longueur d'onde, en vert 20 points, en rouge 30 points, en noir 60 points, en cyan 90 points, en magenta 180 points et en jaune 270 points.

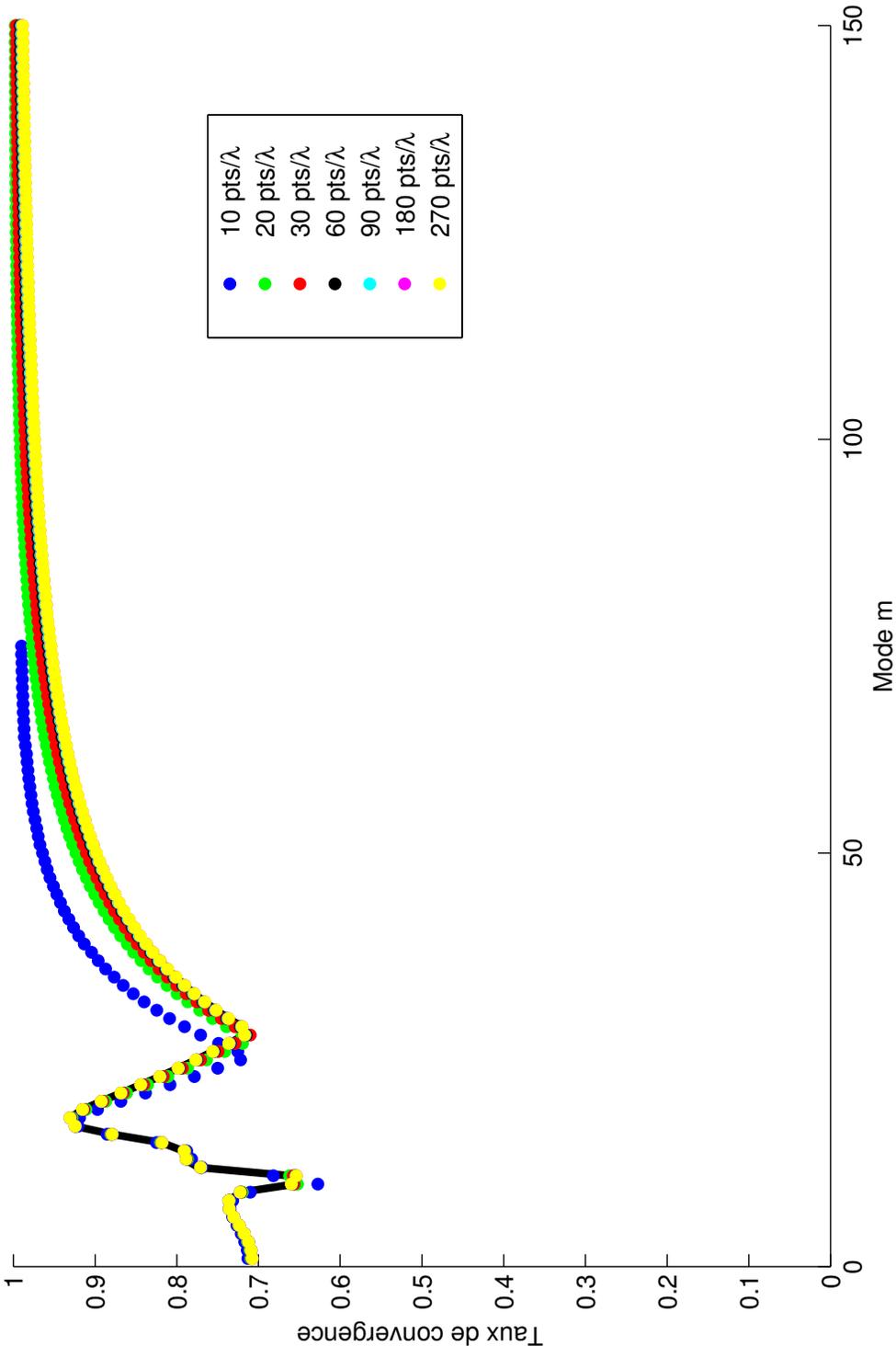


FIGURE 8.9 – Représentation des taux de convergence modaux analytiques et numériques pour différents raffinements, avec conditions de transmission de type Després. Le trait plein noir représente les taux modaux analytiques

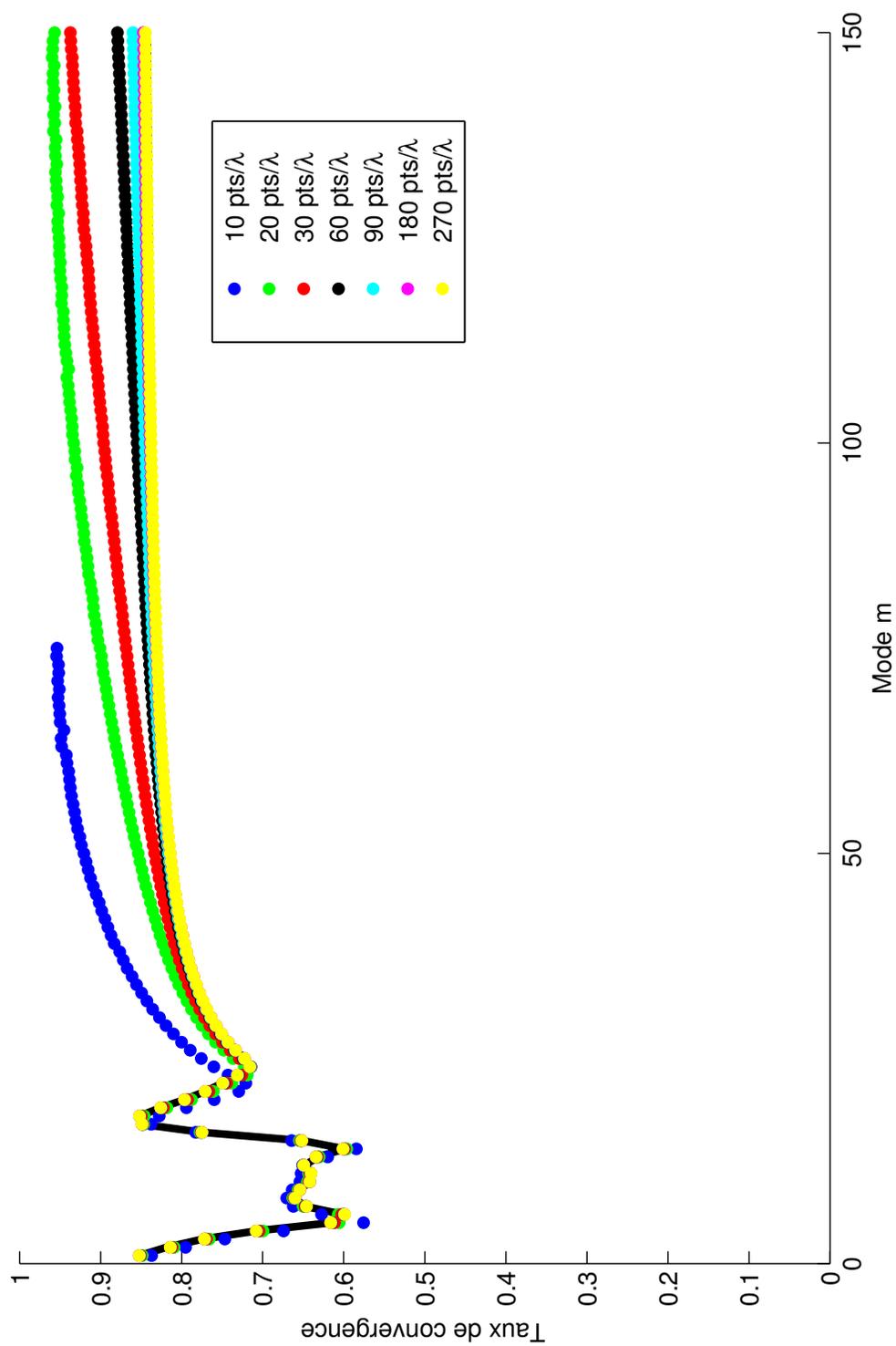


FIGURE 8.10 – Représentation des taux de convergence modaux analytiques et numériques pour différents raffinements, avec conditions de transmission avec opérateur non local. Le trait plein noir représente les taux modaux analytiques

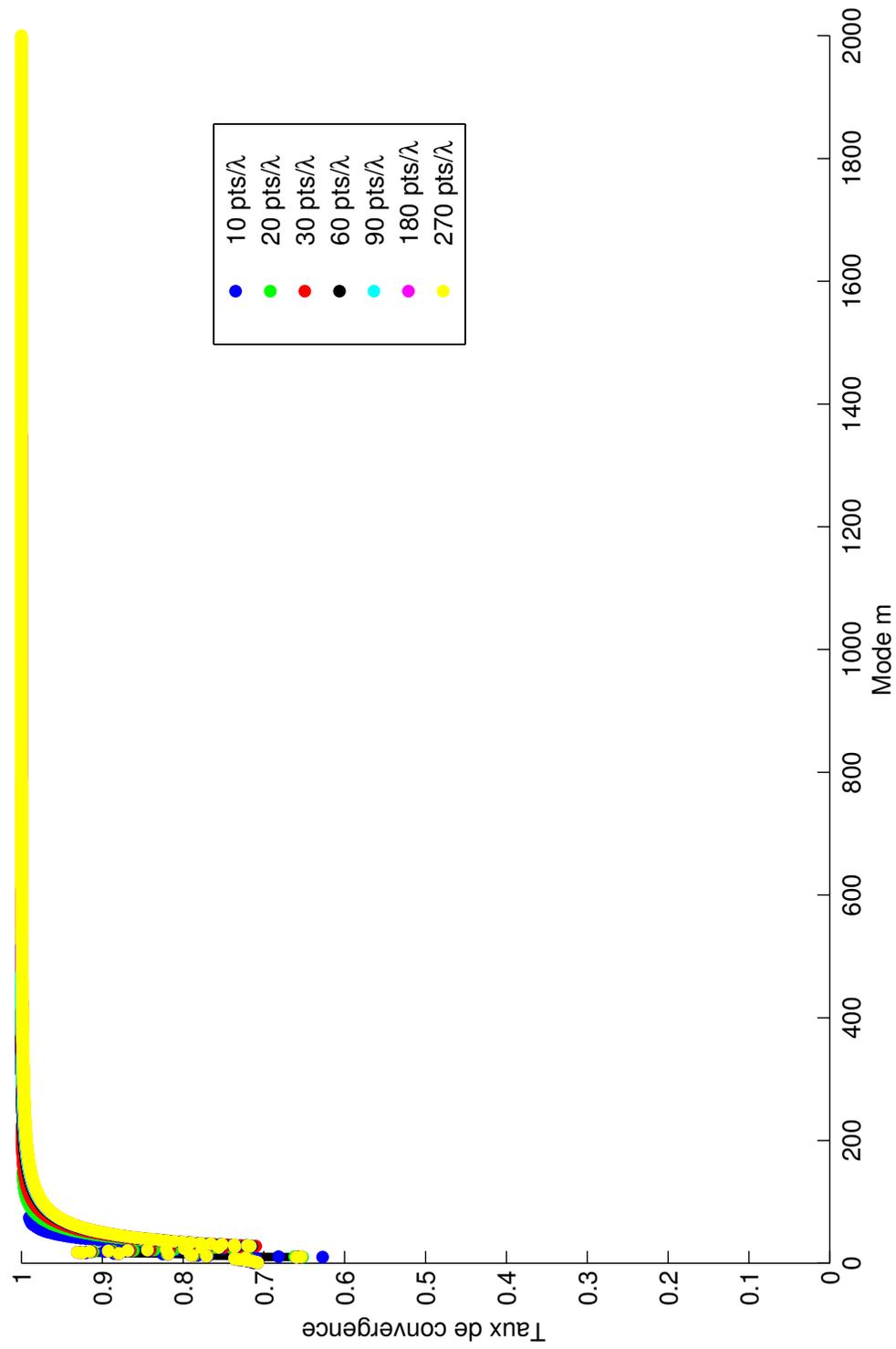


FIGURE 8.11 – Représentation des taux de convergence modaux analytiques et numériques pour différents raffinements, avec conditions de transmission de type Després. Le trait plein noir représente les taux modaux analytiques

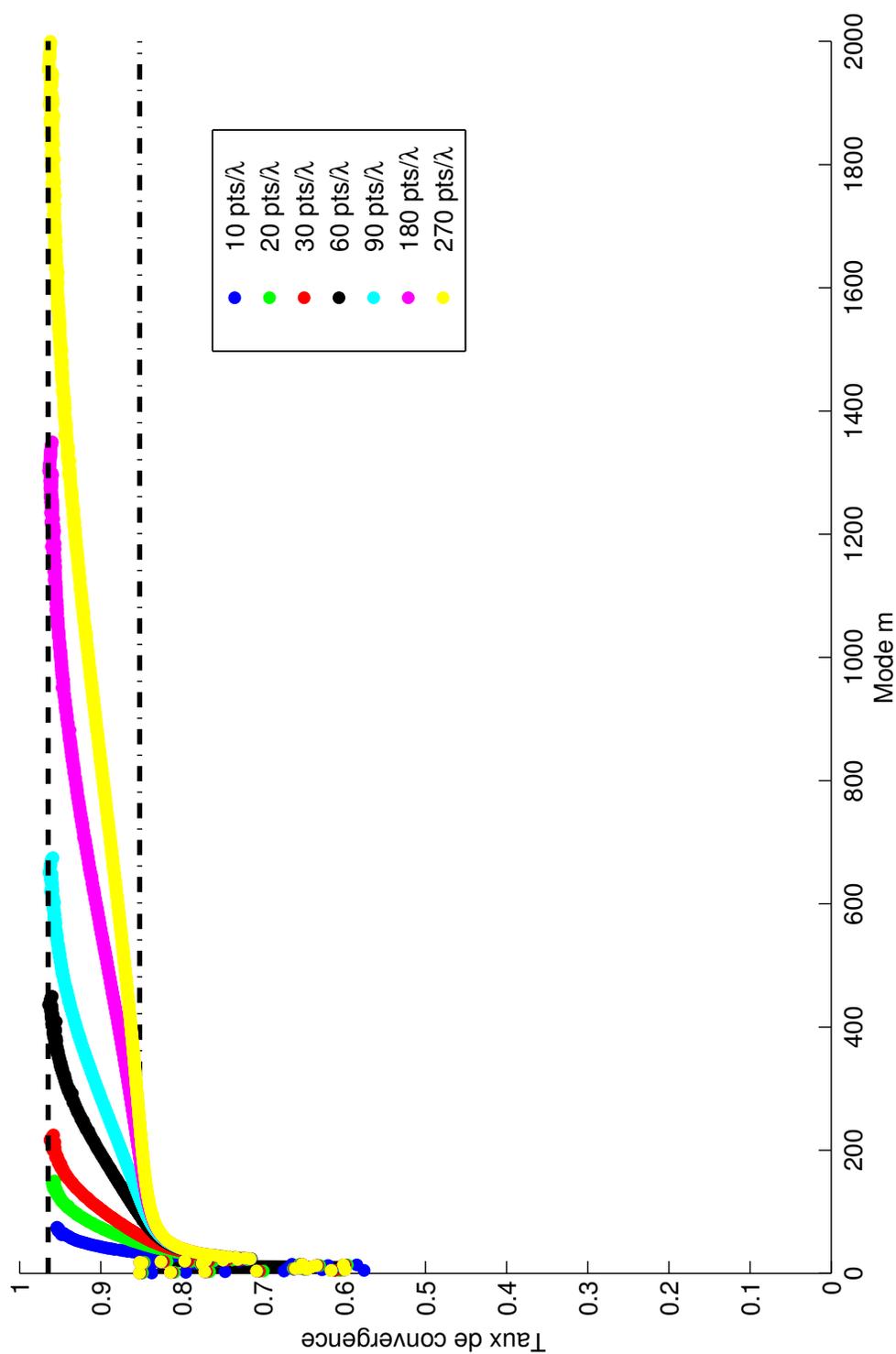


FIGURE 8.12 – Représentation des taux de convergence modaux analytiques et numériques pour différents raffinement, avec conditions de transmission avec opérateur non local. Le trait plein noir représente les taux modaux analytiques. Le trait mixte point-pointillé symbolise le taux modal analytique limite (≈ 0.86), tandis que le trait pointillé correspond au maximum de tous les taux numériques (≈ 0.96)

8.1.3 INFLUENCE DE LA TRONCATURE

Intéressons-nous désormais à l'influence de la troncature sur la convergence des algorithmes. Nous avons vu dans la section 7.2.2 que la troncature telle que nous l'avons implémentée (c'est à dire directement sur les colonnes de la matrice de l'opérateur intégral) n'influait que très peu l'opérateur et ses valeurs propres. Regardons comment la taille de la troncature influence les algorithmes de Jacobi ou GMRES. Rappelons que dans (7.31), la fonction de troncature est définie comme

$$\chi(r) = \chi_0\left(\frac{r}{\delta}\right) \quad (8.1)$$

et que pour tout $x > 1$, $\chi_0(x) = 0$. La taille de la troncature est donc gouvernée par ce paramètre δ . Dans quelques uns des cas présentés précédemment, nous allons faire évoluer ce paramètre δ et observer le nombre d'itérations que l'algorithme (de Jacobi ou GMRES) requiert pour atteindre la convergence (résidu à 10^{-11}). La figure 8.13 présente les résultats de ces expériences. Elle met en évidence l'existence d'une taille de troncature optimale. Avoir une troncature plus large n'aide plus à la convergence (et peut même la dégrader). Cette propriété est vraiment très important car la troncature permet de réduire fortement le coût associé à l'utilisation d'opérateurs intégraux. Sans troncature, les opérateurs que nous avons développés seraient trop coûteux à calculer et apporteraient une contribution dense non négligeable à la matrice des problèmes locaux (voir les équations matricielles (6.99) et (6.100)). Avec la troncature, le surcoût est négligeable et on montre ici qu'en plus, la troncature n'induit pas non plus de dégradation de la convergence de l'algorithme.

Une autre propriété qui semble se dégager des figures de 8.13 est que la largeur de troncature optimale paraît inversement proportionnelle au nombre d'onde k . Dans le cas de l'algorithme de Jacobi, l'optimum semble se trouver pour $k\delta \approx \frac{1}{2}$, tandis pour l'algorithme de GMRES, la taille de troncature optimale semble un peu plus grande, pour $k\delta \approx 1.5$. Ces longueurs de troncature optimales correspondent, dans notre cas, à de faibles longueurs (seulement quelques fractions de la longueur totale de l'interface). Ainsi, $k\delta \approx 1.5$ correspond, si $k = 10$ à un paramètre δ valant environ 1.6% de la longueur de l'interface Σ . Tous les autres cas donne des longueurs inférieures. Après troncature, la matrice de l'opérateur est donc très creusée.

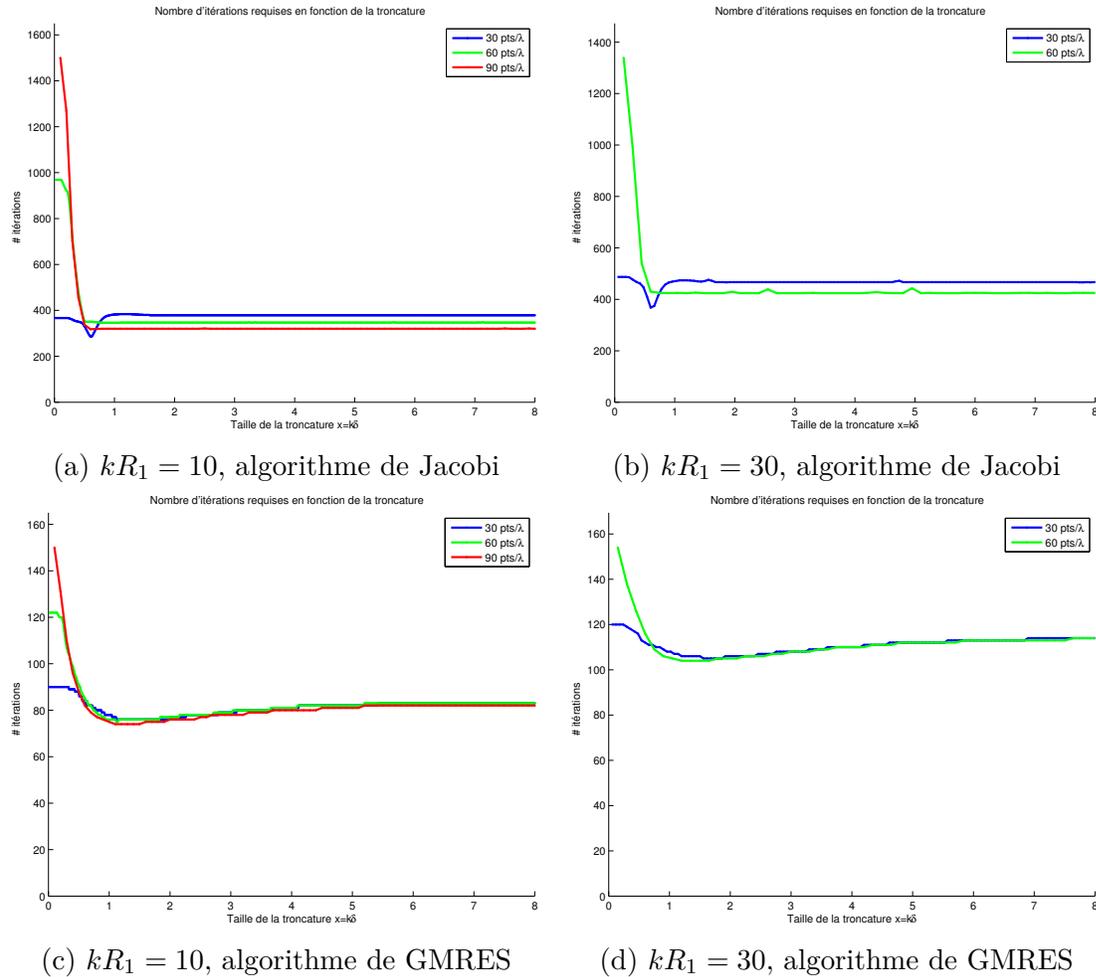


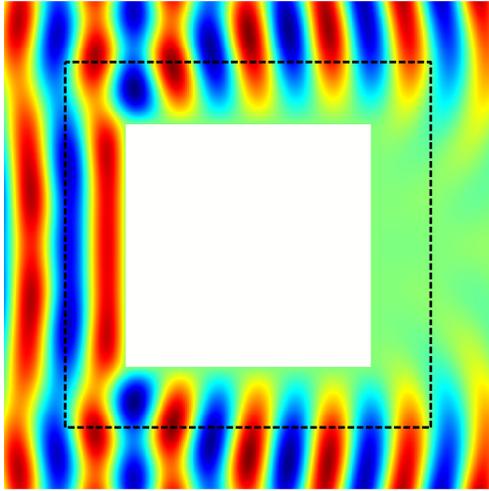
FIGURE 8.13 – Évolution du nombre d’itérations nécessaire à la convergence de l’algorithme pour différents raffinements. À gauche, le nombre d’onde vaut $k = 10$, à droite $k = 30$. En haut, on utilise l’algorithme de Jacobi, et en bas l’algorithme GMRES. L’abscisse est normalisée et vaut $x = k\delta$ où δ est la longueur de la troncature.

8.2 CAS DU CARRÉ

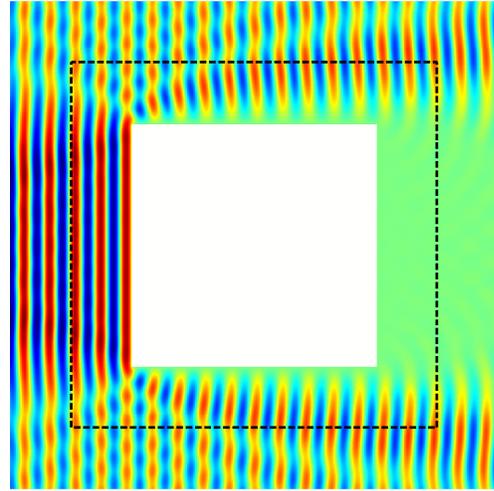
Presque toutes les démonstrations du chapitre de théorie générale reposent sur une hypothèse de régularité de la surface Σ . Cependant, il est très courant dans les applications numériques d'avoir des interfaces avec des coins. Même pour les interfaces régulières du problème continu, après discrétisation les interfaces sont de toutes façons polygonales. C'est pourquoi nous allons étudier numériquement dans cette section le cas de la diffraction d'un carré. L'interface de transmission sera carré également. Le carré diffractant est de côté $c_1 = 2.0$ et une condition absorbante est placée sur un carré de côté $c_2 = 4.0$. L'interface est placée au milieu, et est de côté $c_\Sigma = 3.0$. Nous prendrons encore une fois $k = 10$ et $k = 30$. Comme les épaisseurs entre les différentes interfaces sont identiques au cas du cercle, nous utiliserons les mêmes paramètres d'opérateurs et de relaxation (voir tables 8.1 et 8.2). Les figures 8.14 représentent les solutions des problèmes en champs total et diffracté.

Nous n'étudierons ici que l'opérateur tronqué ($\delta = 0.1$ sauf mention contraire) et nous le comparerons aux conditions de transmission de type Padé. La figure 8.15 présente les résultats de convergence pour le nombre d'onde $k = 10$. Les figures de gauche (8.15a, 8.15c et 8.15e) montrent la convergence pour les conditions avec opérateur tronqué, celles de droite (8.15b, 8.15d et 8.15f) pour les conditions basées sur l'approximation de Padé pour chacun des trois algorithmes itératifs (respectivement Jacobi, Gauss-Seidel et GMRES). Les résultats sont très similaires au cas du cercle. Ainsi, l'algorithme GMRES reste le plus performants dans presque toutes les situations, tandis que l'algorithme de Gauss-Seidel réduit environ par 2 le nombre d'itérations par rapport à l'algorithme de Jacobi. Il semblerait que la rupture de régularité de l'interface n'influence que peu la convergence des algorithmes.

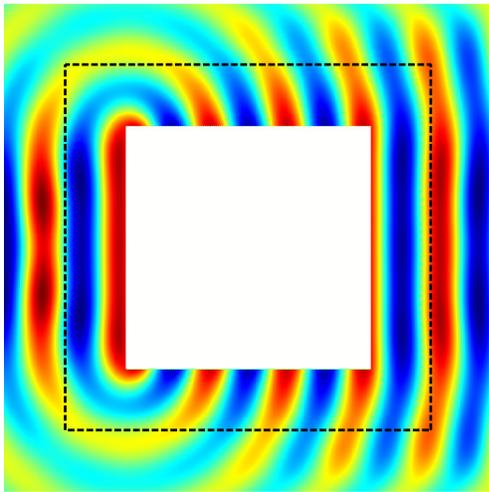
Concernant les comportements des conditions de transmission face au raffinement de maillage, les conclusions sont là encore identiques. Si l'on ne tient pas compte du maillage le plus grossier, qui représente probablement trop mal les hauts modes, les taux de convergence en fin de convergence pour les algorithmes avec opérateur tronqué sont quasiment indépendants du raffinement. Le nombre d'itérations nécessaire pour obtenir la convergence avec ces conditions de transmission est soit stable, soit a tendance à diminuer quand on raffine le maillage. À l'inverse, avec des conditions de type Padé, il est clair que la convergence se dégrade avec le raffinement, y compris pour l'algorithme GMRES.



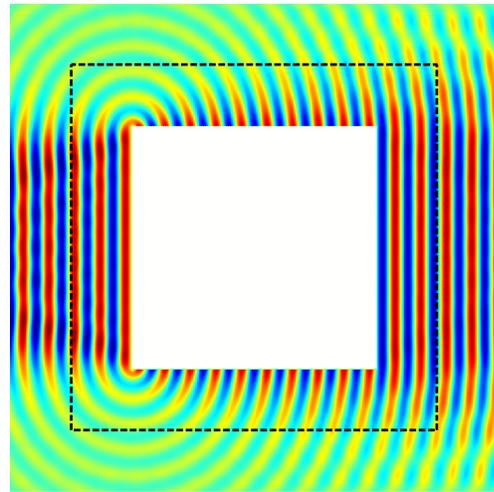
(a) Solution (en champ total) pour $k = 10$ ($f = 477 \text{ MHz}$)



(b) Solution (en champ total) pour $k = 30$ ($f = 1.43 \text{ GHz}$)

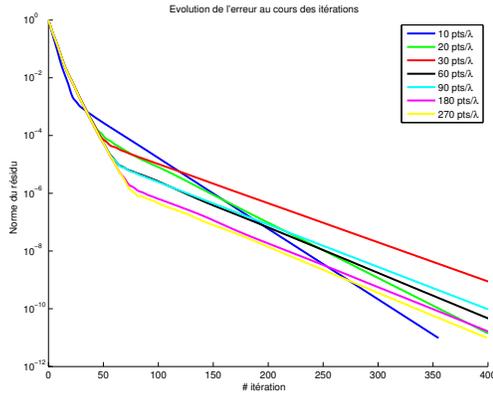


(c) Solution (en champ diffracté) pour $k = 10$ ($f = 477 \text{ MHz}$)

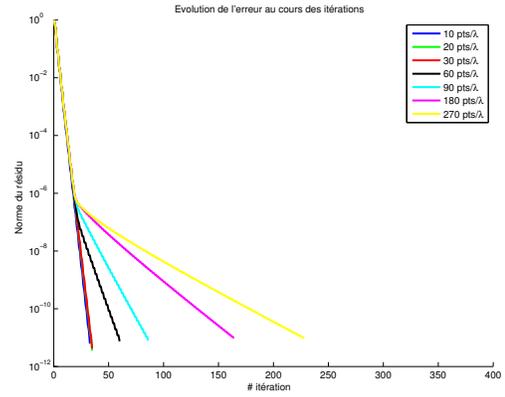


(d) Solution (en champ diffracté) pour $k = 30$ ($f = 1.43 \text{ GHz}$)

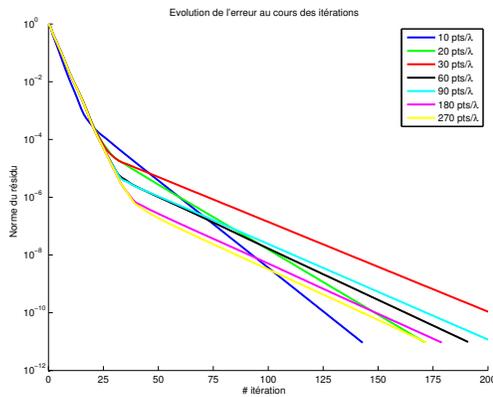
FIGURE 8.14 – Représentation de la solution aux problèmes de diffraction du carré pour différentes valeurs de k . En pointillé, est représentée l'interface de transmission Σ .



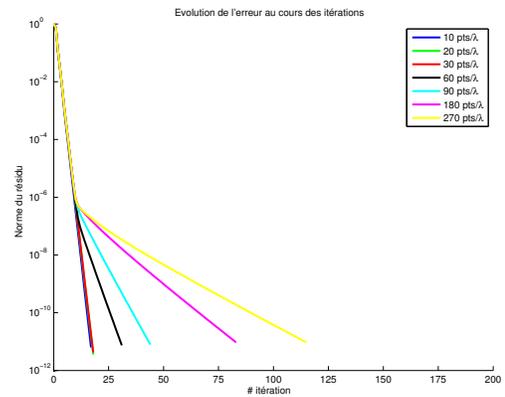
(a) Opérateur – Jacobi



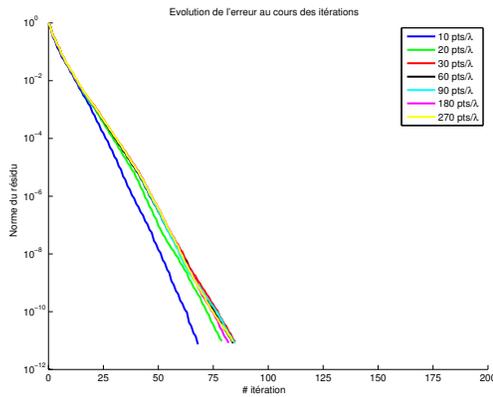
(b) Padé – Jacobi



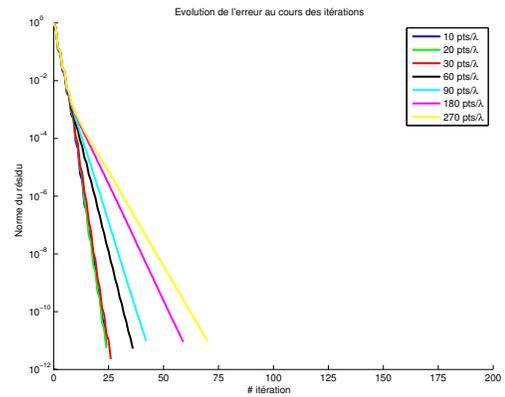
(c) Opérateur – Gauss-Seidel



(d) Padé – Gauss-Seidel



(e) Opérateur – GMRES



(f) Padé – GMRES

FIGURE 8.15 – Évolution des résidus pour différents raffinements et différents types de condition de transmission. Cas du nombre d'onde $k = 10$

La table 8.9 donne les nombres d'itérations requis par chacun des algorithmes pour les deux types de conditions de transmission étudiés. Même si les conditions de type Padé requièrent moins d'itérations, leur comportement face au raffinement (dégradation de la convergence quand on raffine le maillage) est symptomatique des conditions utilisant des opérateurs locaux. En raffinant le maillage, ce dernier permet de représenter de plus en plus de modes élevés, pour lesquels les taux de convergence modaux sont proches de 1 dans le cas d'opérateurs locaux. La convergence se dégrade donc. Dans le cas d'opérateurs non locaux, tels que ceux que nous avons développés, et puisque les taux de convergence modaux sont tous inférieurs à un taux $\tau^* < 1$, les hauts modes ne posent pas de problème et la convergence est identique quelque soit le raffinement.

TABLE 8.9 – Nombre d'itérations nécessaire pour atteindre une erreur de 10^{-11} pour chacune des méthodes. Cas $k = 10$.

	Opérateur tronqué			Padé		
	Jacobi	Gauss-Seidel	GMRES	Jacobi	Gauss-Seidel	GMRES
10 pts/ λ	355	143	68	33	17	24
20 pts/ λ	408	171	79	35	18	24
30 pts/ λ	544	234	85	35	18	26
60 pts/ λ	443	191	84	60	31	36
90 pts/ λ	468	203	85	86	44	42
180 pts/ λ	415	179	82	164	83	59
270 pts/ λ	399	172	84	228	115	70

Comme pour le cas circulaire, nous avons représenté les taux modaux numériques pour chacun des raffinements et pour l'algorithme de Jacobi (cette fois, nous ne connaissons pas les taux analytiques). La figure 8.16 montre ces taux. Comme précédemment, et malgré une interface Σ non régulière, les expériences numériques laissent penser qu'il existe un taux de convergence modal $\tau^* < 1$ tel que pour tous les raffinements de maillage, le maximum des taux modaux soit borné par τ^* . Encore une fois, cette propriété n'est réalisable qu'avec des opérateurs non locaux.

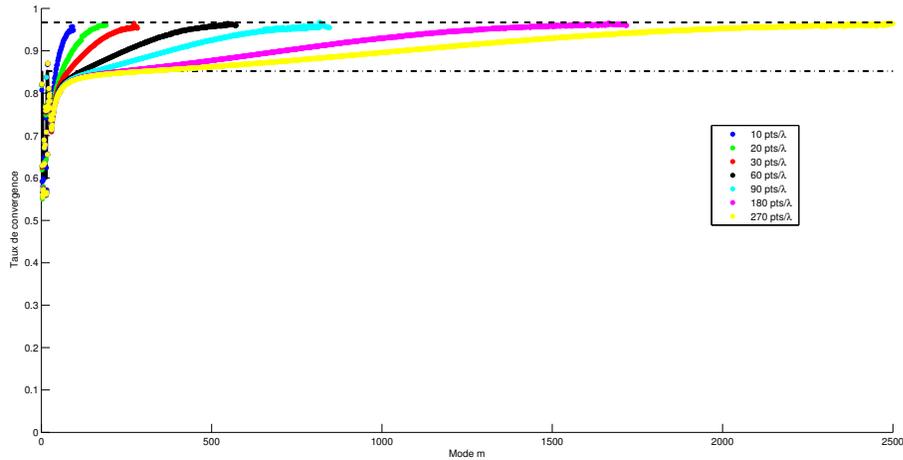
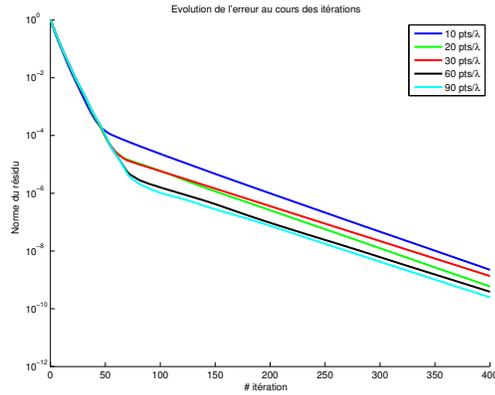


FIGURE 8.16 – Représentation des taux de convergence modaux pour plusieurs raffinements. Cas du carré avec $k = 10$. Le trait mixte point-pointillé correspond à la limite théorique des taux modaux analytiques (sur le cercle). Le trait pointillé simple correspond au maximum de tous les taux modaux numériques.

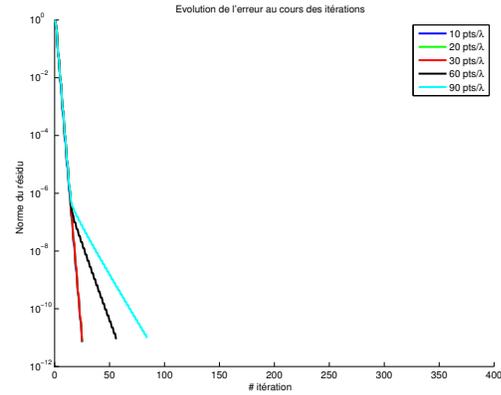
Les résultats sont très similaires dans le cas où $k = 30$, comme nous pouvons le voir sur la figure 8.17. Table 8.10, on retrouve le nombre d’itérations pour atteindre la convergence dans chacune des situations. Elle montre encore une fois que le raffinement ne dégrade pas la convergence des méthodes itératives lorsqu’on utilise des opérateurs non locaux tronqués, contrairement aux opérateurs locaux.

TABLE 8.10 – Nombre d’itérations nécessaire pour atteindre une erreur de 10^{-11} pour chacune des méthodes. Cas $k = 30$.

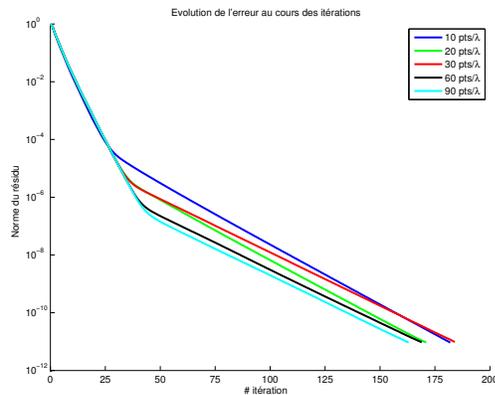
	Opérateur tronqué			Padé		
	Jacobi	Gauss-Seidel	GMRES	Jacobi	Gauss-Seidel	GMRES
10 pts/λ	579	182	105	25	13	24
20 pts/λ	535	171	114	25	13	24
30 pts/λ	577	184	115	27	13	24
60 pts/λ	534	169	118	56	30	36
90 pts/λ	517	163	119	84	43	43



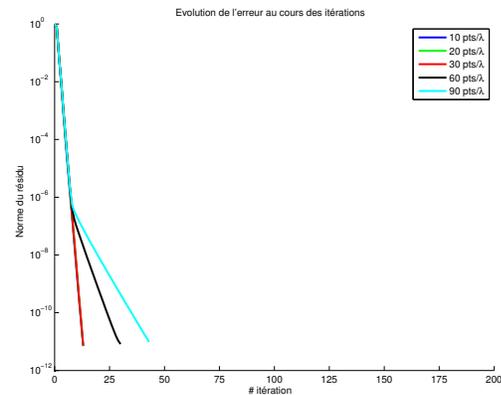
(a) Opérateur – Jacobi



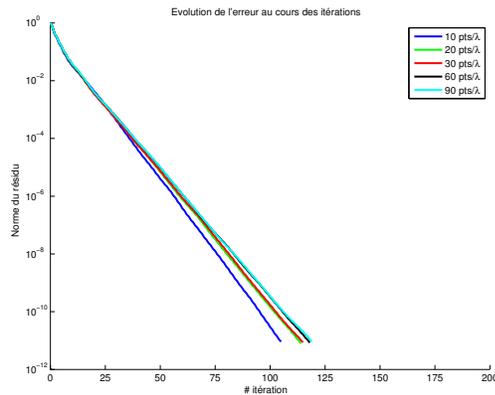
(b) Padé – Jacobi



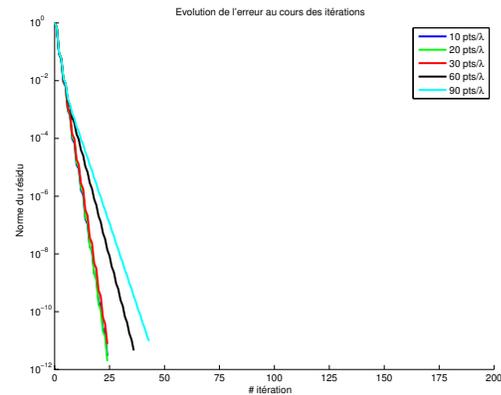
(c) Opérateur – Gauss-Seidel



(d) Padé – Gauss-Seidel



(e) Opérateur – GMRES



(f) Padé – GMRES

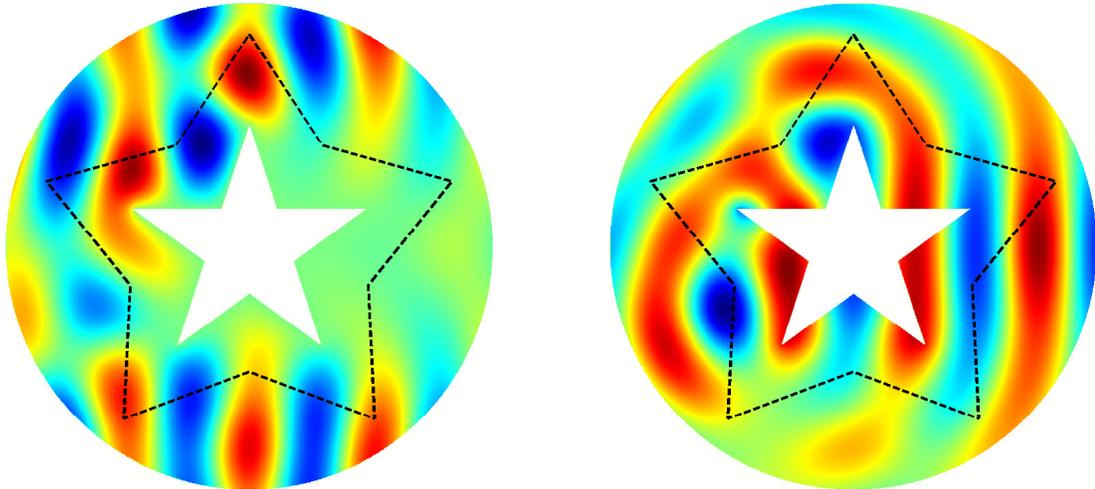
FIGURE 8.17 – Évolution des résidus pour différents raffinements et différents types de condition de transmission. Cas du nombre d'onde $k = 30$

8.3 DIVERSES GÉOMÉTRIES

Nous allons ici présenter deux géométries différentes un peu moins académiques que celles étudiées jusqu'alors. La première géométrie est un domaine étoilé, l'interface de transmission étant également étoilée. La seconde géométrie est un peu plus complexe et consiste en une cavité résonnante.

8.3.1 DOMAINE ÉTOILÉ

L'objet diffractant dans cette section est une étoile régulière à 5 branches (voir figure 8.18 présentant la solution du problème). Les points extérieurs de l'étoile diffractante sont sur un cercle de rayon $R_0 = 1$, tandis que les points intérieurs sont sur un cercle de rayon $R_1 = \frac{\cos 2\pi/5}{\cos \pi/5} R_0 \approx 0.38$. Nous n'avons pas pris comme interface de transmission une étoile régulière afin d'équilibrer la taille des deux sous-domaines. Ainsi, les points extérieurs de l'interface de transmission sont sur un cercle de rayon $R_2 = \frac{7}{4}R_0 = 1.75$, tandis que les points intérieurs sont un cercle de rayon $R_3 = \frac{1}{2} \left(1 + \frac{\cos 2\pi/5}{\cos \pi/5}\right) R_2 \approx 1.21$. Enfin, le domaine est fermé par une interface circulaire de rayon 2 sur laquelle on place une condition absorbante d'ordre 1.



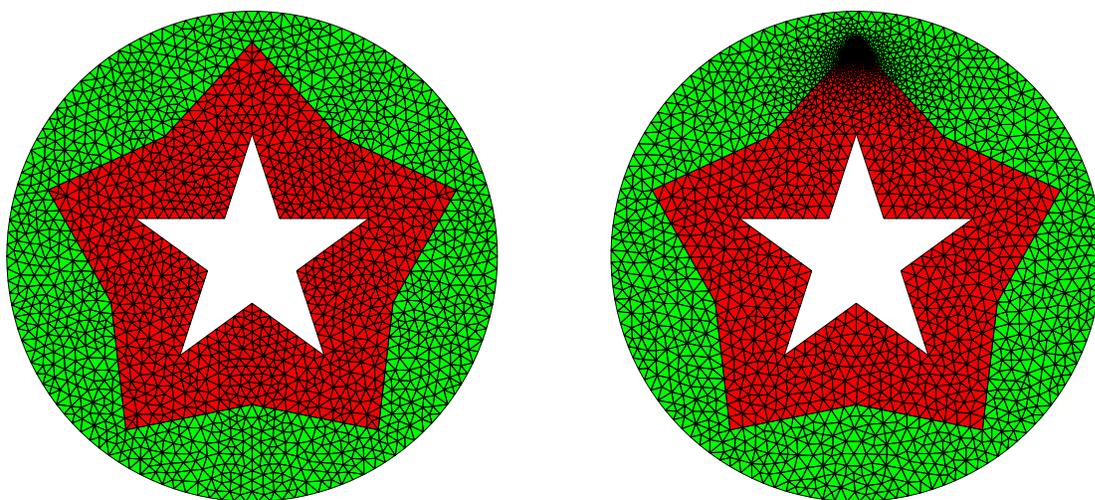
(a) Solution du problème en champ total

(b) Solution du problème en champ diffracté

FIGURE 8.18 – Solution au problème de diffraction d'une étoile à 5 branches. Les pointillés représentent l'interface de transmission

La figure 8.18 présente la solution du problème pour un nombre d'onde valant $k = 2\pi$, en champ total et en champ diffracté et la position de l'interface de transmission. Le domaine est discrétisé avec le paramètre de maillage $h_0 = 0.1$, ce qui correspond à 10 points par longueur d'onde. Le maillage associé est représenté figure 8.19a. Au lieu de raffiner uniformément le maillage comme nous l'avons fait lors des précédentes expériences numériques (cercle et carré), nous étudierons le comportement des algorithmes lorsqu'on raffine localement autour d'un point du maillage situé sur l'interface de transmission. La figure 8.19b présente un exemple de raffinement local et la table 8.11 présente les caractéristiques des maillages pour les différents raffinements h_{raff} que nous avons appliqués. On voit sur les graphiques de répartition des tailles d'arêtes (dont l'abscisse est en échelle logarithmique) que cette méthode de raffinement local ne fait qu'ajouter des arêtes dont les tailles sont comprises entre h_{raff} et h_0 . Logiquement, plus on raffine et plus la taille minimale des arêtes diminue. Cependant, le nombre d'arêtes ajoutées est relativement faible.

Nous étudierons les algorithmes de Jacobi et GMRES pour les conditions de trans-



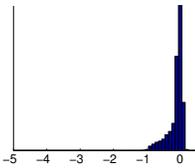
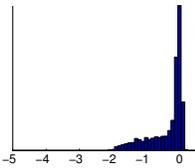
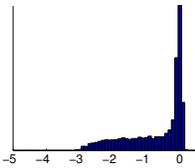
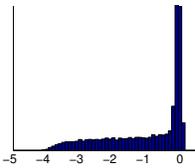
(a) Maillage uniforme, $h_0 = 0.1$, 10 points par longueur d'onde

(b) Maillage non uniforme, un point de l'interface est fortement raffiné ($\frac{h_0}{100}$ ici)

FIGURE 8.19 – Exemples de maillages utilisés pour le problème du domaine étoilé. Le sous-domaine Ω_1 est en rouge, et Ω_2 en vert.

mission avec opérateur non local tronqué et de type Padé. Même si la géométrie est très différente du cas du cercle, nous utiliserons pour l'opérateur tronqué les

TABLE 8.11 – Caractéristiques des maillages raffinés. Les figures correspondent à la répartition statistique des valeurs de $\log \frac{h}{h_0}$ où h est la taille de chacune des arêtes des maillages.

Raffinement	$\frac{h_0}{10}$	$\frac{h_0}{100}$	$\frac{h_0}{1000}$	$\frac{h_0}{10000}$
$h_{raff} =$	$\frac{h_0}{10}$	$\frac{h_0}{100}$	$\frac{h_0}{1000}$	$\frac{h_0}{10000}$
Nombre de nœuds	1896	2386	2970	3697
Nombre d'arêtes	5493	6963	8715	10885
Répartition des arêtes				

paramètres donnés dans la table 8.1.

TABLE 8.12 – Nombre d'itérations nécessaire pour atteindre une erreur de 10^{-11} pour chacune des méthodes. Cas étoilé

	Opérateur tronqué		Padé	
	Jacobi	GMRES	Jacobi	GMRES
$h_{raff} = \frac{h_0}{10}$	553	81	89	42
$h_{raff} = \frac{h_0}{100}$	618	82	777	98
$h_{raff} = \frac{h_0}{1000}$	559	82	6270	200
$h_{raff} = \frac{h_0}{10000}$	590	83	>15 000	341

Les historiques de convergence montrant l'évolution du résidu au cours des itérations pour chacune des méthodes se trouvent figure 8.20. Pour ces expériences, la dépendance des conditions locales (comme Padé) au raffinement est particulièrement flagrante (voir également table 8.12). Pour la méthode de Jacobi avec

conditions de transmission de type Padé, le nombre d'itérations semble directement proportionnelle à $\frac{1}{h_{raff}}$. Si l'augmentation du nombre d'itérations est plus limité pour GMRES, il passe néanmoins de 42 itérations pour le maillage le moins raffiné à 341 pour le maillage le plus raffiné. En utilisant les conditions de transmission non locales tronquées, l'algorithme GMRES converge en 81 à 83 itérations. Il est quasiment insensible au raffinement. C'est exactement le comportement désiré lorsque nous avons développé les conditions de transmission non locales.

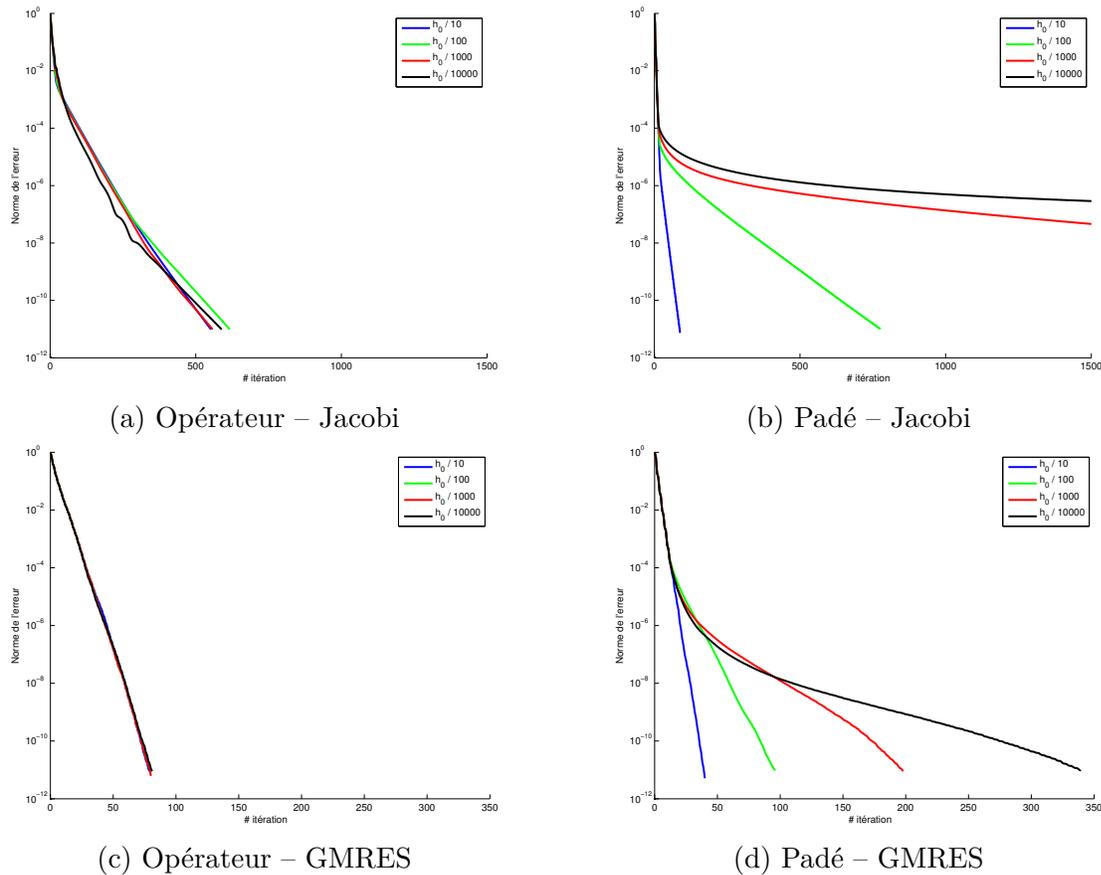


FIGURE 8.20 – Historique de convergence des algorithmes de Jacobi et GMRES utilisant les conditions avec opérateur non local (8.20a et 8.20c) ou les conditions de type Padé (8.20b et 8.20d).

8.3.2 CAVITÉ RÉSONNANTE

Le problème de la cavité résonnante correspond à la diffraction d'un disque parfaitement conducteur (sur lequel on impose une condition de Dirichlet) dans lequel on applique une ouverture étroite qui conduit à une cavité. Le rayon du disque diffractant est de $R_1 = 0.8$, et le rayon de la cavité est $R_0 = 0.5$. La condition absorbante est placée en $R_2 = 1.25$. L'ouverture est de taille $e = 2R_0 \sin \frac{19\pi}{360}$. La figure 8.21a présente la géométrie du problème. L'interface de transmission est construite telle qu'elle se situe à une distance 0.225 de l'objet, sauf bien entendu dans l'ouverture où l'interface est placée à $\frac{e}{4}$ de l'objet (voir figure 8.21b).

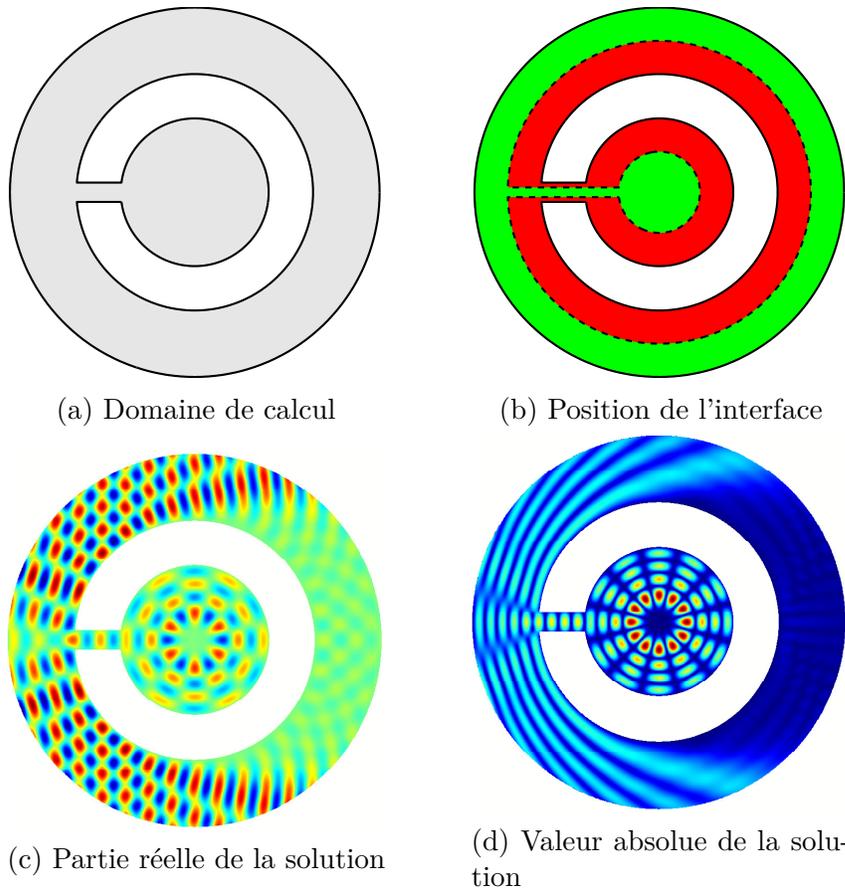


FIGURE 8.21 – Présentation du problème de cavité résonnante. En haut, les domaines de calcul Ω (gris), Ω_1 (rouge) et Ω_2 (vert). En bas, la partie réelle et la valeur absolue de la solution (champ total).

Il est évident que cette manière de décomposer le domaine n'est pas la plus naturelle pour ce type de problème, mais l'une des limitations des conditions de transmission basées sur des opérateurs intégraux est que l'interface doit être une courbe fermée. Cet exemple est construit pour étudier le comportement des algorithmes itératifs dans le cas de problèmes résonnants. Le nombre d'onde k est donc choisi tel que la fréquence associée soit une fréquence de résonance de la cavité. Plus précisément, on a

$$\exists n \in \mathbb{N} \text{ tel que, } J_n(kR_0) = 0. \quad (8.2)$$

où J_n est la fonction de Bessel d'ordre n de première espèce. On peut voir, d'après la représentation de la solution, que $n = 6$ car la dépendance en θ de la solution dans la cavité correspond à $e^{6i\theta}$. De plus, en excluant le premier 0 en $r = 0$, on voit que nous avons choisi k tel que kR_0 soit le quatrième zéros de la fonction de Bessel J_6 . Ceci correspond à $k \approx 40.64$, ou une fréquence $f \approx 1.94 \text{ GHz}$. Notons également que l'ouverture agit comme un guide d'onde. Si la fréquence était inférieure à la fréquence de coupure du guide d'onde, la solution serait quasi nulle dans la cavité. Ici, le nombre d'onde de coupure vaut $k_c = \frac{\pi}{e} \approx 9.52$, soit $f_c \approx 454 \text{ MHz}$.

Ici, nous allons comparer les algorithmes de Jacobi et GMRES utilisant des conditions de type Després, Padé ou des conditions avec un opérateur non local tronqué. Nous utiliserons les paramètres de la table 8.13, obtenus par optimisation sur le cercle sans cavité résonnante.

TABLE 8.13 – Paramètres pour le cas de la cavité résonnante

	Jacobi		GMRES	
	Després	Opérateur	Després	Opérateur
z	$1.30 + 0.03i$	$0.18 - 0.004i$	$0.668 - 0.04i$	$0.35 - 0.10i$
β		2.95		0.67
r	0.53	0.64		

La figure 8.22 présente les historiques de convergence au cours des itérations des deux algorithmes de Jacobi et GMRES pour chacune des conditions de transmission. Si les résultats du GMRES sont assez classiques, avec la convergence la plus rapide pour les conditions de type Padé, les courbes concernant l'algorithme de Jacobi montrent un phénomène intéressant. Avec les conditions de type Després ou

basées sur l'opérateur intégral tronqué, l'algorithme converge vers la solution relativement lentement. Par contre, l'utilisation des conditions de type Padé conduit à la divergence de l'algorithme. Après calcul des valeurs propres pour ce problème, nous nous sommes effectivement aperçu de l'existence de quelques valeurs propres de module supérieur à 1, entraînant la divergence des méthodes itératives de type point-fixe (Jacobi et Gauss-Seidel). Ce phénomène, assez rarement observé lors de nos tests numériques, montre que les conditions de transmission basées sur des fractions rationnelles n'assurent pas nécessairement la convergence de tous les algorithmes itératifs. Ces conditions sont moins robustes.

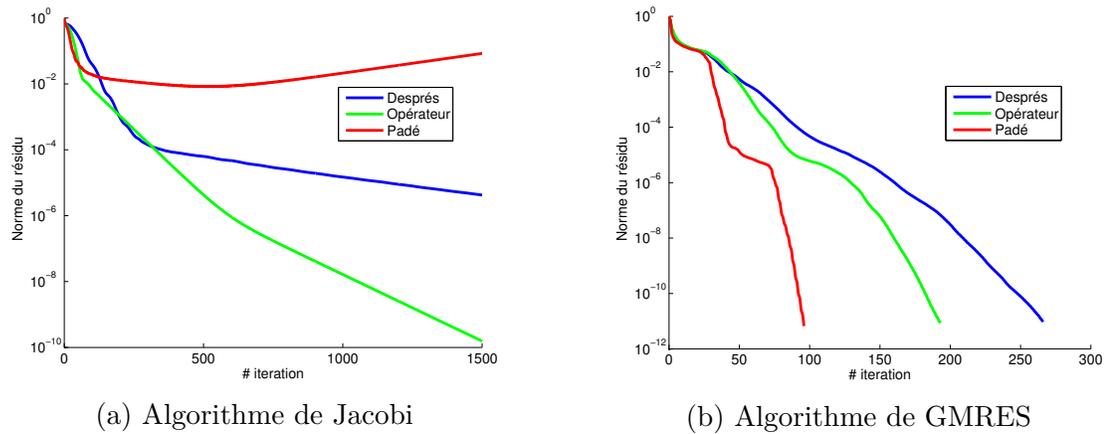


FIGURE 8.22 – Historique de convergence des algorithmes de Jacobi et GMRES pour différentes conditions de transmission, sur le cas de la cavité résonnante.

8.4 MULTI SOUS-DOMAINES

Dans l'ensemble du document, nous nous sommes focalisés sur une décomposition de deux sous-domaines. Cependant, nous l'avons indiqué au début du chapitre 2 de théorie générale, la méthode de nous avons développée s'adapte à un nombre quelconque de sous-domaines, à la condition que les interfaces entre les domaines ne s'intersectent pas, c'est à dire que le domaine est découpé en pelures d'oignon. Avec cette décomposition, chaque domaine ne communique qu'à un seul autre domaine à travers une interface donnée. Nous avons donc implémenté la décomposition de domaine en multi sous-domaines.

Nous présenterons ici l'exemple de la géométrie circulaire décomposée en 2, 4, 8 ou 16 sous-domaines. À chaque fois, les sous-domaines sont d'épaisseur identiques,

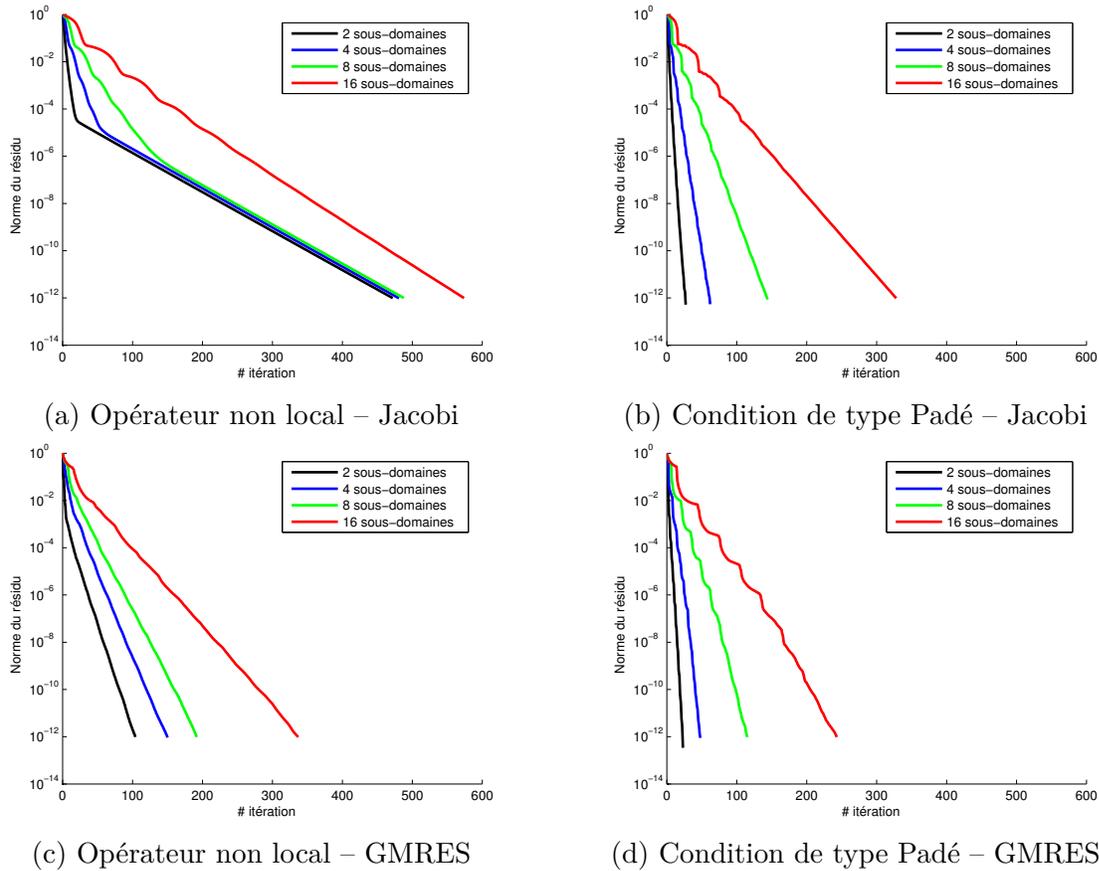


FIGURE 8.23 – Évolution des résidus au cours des itérations des algorithmes de Jacobi et GMRES. À gauche, les conditions de transmissions non locales sont utilisées. À droite, nous utilisons les conditions de type Padé.

et un même maillage est utilisé pour toutes les simulations. Ainsi, par exemple, le sous-domaine Ω_1 du problème décomposé en 2 sous-domaines correspond à l'agglomération des sous-domaines Ω_1 à Ω_8 du problème décomposé en 16 sous-domaines. De plus, nous utiliserons les mêmes conditions de transmission sur toutes les interfaces : soit les conditions de type Padé, soit les conditions non locales. Enfin, nous testerons le comportement des algorithmes de Jacobi et GMRES.

La figure 8.23 présente l'évolution des résidus pour chacune des situations. Excepté l'algorithme de Jacobi utilisé avec des conditions non locales, le comportement des algorithmes est relativement classique et attendu. En effet, on observe une dégradation de la convergence avec le nombre de sous-domaines. Ce résultat est connu

et semblable à de nombreuses méthodes itératives. L'algorithme de Jacobi avec des conditions de transmission non locales (figure 8.23a) montre cependant un comportement particulier. En effet, le nombre d'itérations nécessaire pour obtenir la convergence est presque identique pour les trois premières décompositions (2, 4 et 8 sous-domaines). Il semblerait donc que, sur ce cas, la convergence soit relativement dégradée sur ces décompositions, et que cette dégradation ne soit pas due au nombre de sous-domaines. Si nous n'avons pas réussi à identifier le problème, nous pouvons suspecter un effet de maillage (résonance locale par exemple).

9

Conclusions et perspectives

Sommaire

9.1	Conclusion	225
9.2	Sur les jonctions	227
9.3	Sur les conditions de transmission	228
9.4	Retour sur l'analyse modale	230

Ce chapitre est une conclusion rapide sur le travail réalisé et détaillé dans ce document. Nous présenterons également plusieurs perspectives concernant les conditions de transmission non locales telles que nous les avons développées. Il s'agit principalement d'idées que nous pensons d'intérêt mais que nous n'avons pas eu le temps de détailler. Nous espérons que ce chapitre permettra qu'elles ne tombent pas dans l'oubli.

9.1 CONCLUSION

Dans ce document, nous avons tout d'abord introduit un cadre général pour les méthodes de décomposition de domaine. Ces méthodes reposent sur des conditions de transmission de type impédantes. Nous avons pu exhiber, pour une classe

d'opérateurs très généraux, des conditions suffisantes pour garantir d'une part le caractère bien posé des problèmes locaux et d'autre part la convergence de la méthode. De plus, en ajoutant une propriété d'isomorphisme à ces opérateurs, nous pouvons garantir une convergence exponentiellement rapide des méthodes de décomposition de domaine de type Jacobi.

Ensuite, nous avons développé une analyse modale de plusieurs cas simples. Cette analyse a un double objectif. Tout d'abord, elle permet de mieux appréhender la théorie générale développée dans les chapitres précédents, en mettant en évidence plusieurs caractéristiques fondamentales pour obtenir la convergence géométrique de la méthode. Mais l'intérêt principale de notre analyse modale est de fournir une estimation analytique précise du taux de convergence de la méthode. Ceci nous permet de mettre en place une procédure d'optimisation pour réduire ce taux de convergence et ainsi, accélérer la résolution de problèmes dans un contexte numérique.

La seconde partie du document est axée sur la construction d'opérateurs vérifiant les conditions garantissant la convergence exponentielle de la méthode de décomposition de domaine. Une première méthode de construction se base sur la théorie des potentiels de Riesz. Ces potentiels correspondent aux puissances fractionnaires du Laplacien. Sous certaines conditions de régularité de la courbe, nous montrons que l'opérateur défini par les potentiels de Riesz vérifie la propriété d'isomorphisme voulue. La deuxième méthode de construction d'opérateurs fait intervenir les semi-normes de Gagliardo. Elles permettent de définir des normes des espaces de Sobolev H^s et ces normes nous donnent facilement les propriétés voulues. Tous ces opérateurs sont des opérateurs intégraux. Ils peuvent donc être relativement coûteux à implémenter et pour remédier à ce problème, nous avons introduit une méthode de troncature qui permet à la fois de localiser les opérateurs (les intégrales sont réduites à un domaine beaucoup plus petit) et de préserver les propriétés requises pour garantir la convergence exponentielle.

Enfin, la troisième partie du manuscrit s'intéresse aux aspects numériques de notre méthode de décomposition de domaine. Nous avons mis en place une méthode de discrétisation du problème pour laquelle nous montrons que les problèmes locaux restent bien posés, et que l'algorithme de décomposition de domaine converge même après discrétisation. De plus, notre discrétisation possède la propriété très intéressante de conserver la solution, c'est à dire qu'à la convergence, la solution discrète du problème décomposé correspond exactement à la solution discrète du problème initial non décomposé.

Le dernier chapitre de la thèse correspond aux résultats numériques. Nous avons présenté et analysé plusieurs cas test. Les deux premiers cas test sont très académiques et correspondent à un objet circulaire et un objet carré. L'analyse des résultats numériques confirme les propriétés attendues. On trouve ainsi que l'utilisation des opérateurs non locaux permet d'obtenir une convergence quasi indépendante du raffinement du maillage. Au contraire, les autres méthodes testées (conditions de transmission de Després ou fractions rationnelles de type Padé) sont négativement sensibles au raffinement : plus le maillage est fin et plus la convergence se dégrade. Néanmoins, il est à noter que sur ces cas très simple, les conditions de transmission de type Padé permettent une convergence plus rapide. L'intérêt des conditions de transmission que nous avons développées est particulièrement mis en évidence sur les deux cas tests suivants, à savoir le cas du raffinement local et le cas de la cavité résonnante. Le cas du raffinement local montre de manière flagrante l'intérêt de la convergence exponentielle. Même pour des raffinements extrêmement forts, la vitesse de convergence est insensible au raffinement. Enfin, nous avons présenté un cas mettant en évidence la robustesse de notre méthode de décomposition de domaine. Il s'agit du cas de la cavité résonnante. Sur ce cas, certaines méthodes, pour lesquelles aucune preuve de convergence n'est établie, divergent et ne permettent pas d'obtenir la solution. L'utilisation de conditions de transmission qui garantissent la convergence des méthodes de décomposition de domaine permet d'assurer la robustesse des codes numériques.

Pour conclure succinctement, si nos conditions de transmission ne permettent pas toujours d'avoir la convergence la plus rapide (notamment pour les problèmes simples), elles présentent le double avantage d'être très robustes (la convergence de la méthode est prouvée, même après discrétisation) et d'être insensible au raffinement du maillage, même s'il est localement fort.

9.2 SUR LES JONCTIONS

L'une des conditions les plus restrictives pour l'utilisation de conditions de transmission non locales telles que celles développées ici restent la nécessité d'avoir un découpage du domaine en pelures d'oignon. Si un tel découpage est naturel pour certaines applications (objet recouvert de couches successives de matériaux par exemple), on peut avoir parfois besoin d'effectuer d'autres types de découpages. Nous avons donc tenté d'appliquer nos conditions de transmission à d'autres découpages. La figure 9.1 présente les deux découpages étudiés. Dans le premier, un simple carré est coupé en deux dans la longueur (figure 9.1a). Dans le second cas, le carré est percé en son milieu par un trou de taille r , r étant petit devant le

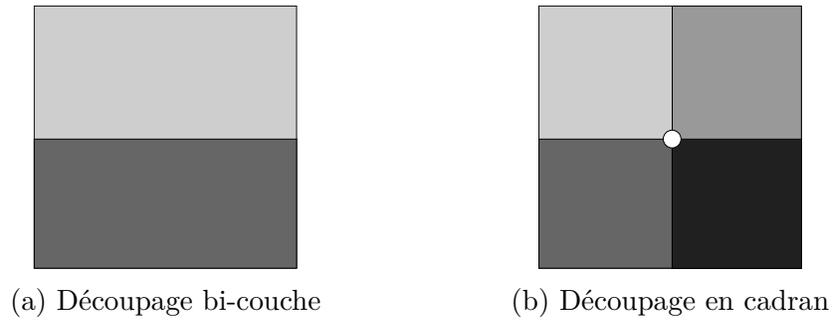


FIGURE 9.1 – Présentation des découpages étudiés. Dans le cas de quatre sous-domaines, le carré est troué en son centre : les interfaces de transmission ne se touchent pas.

côté du carré. Le domaine est ensuite découpé en quatre (figure 9.1b). Sur chaque bord du domaine, on place une condition absorbante d'ordre 1. Le côté du domaine carré vaut 8 et on prend $k = 2\pi$. Il est évident que ces types de découpage ne rentrent pas dans la théorie développée tout au long du document.

Les expériences numériques tendent à montrer qu'effectivement, la convergence est fortement dégradée et n'est plus exponentielle : on retrouve la dépendance en fonction du raffinement. Les figures 9.2 représentent l'erreur $e_h^n = |u_h^{exact} - u_h^n|$ de l'algorithme de Jacobi pour $n = 1000$. Après 1000 itérations, la convergence n'est toujours pas atteinte, et on s'aperçoit que l'erreur se concentre effectivement aux points de jonction.

Il est clair que la limitation d'interfaces de transmission fermées et non intersectantes reste une contrainte forte, que nous n'avons pas encore réussi à dépasser. Trouver une manière de lever cette limitation ouvrirait beaucoup de perspectives quant aux domaines d'application de notre méthodes de décomposition de domaine. Pour le moment, notre méthode est en effet peu adaptée aux problèmes de propagation d'ondes autres que la diffraction d'objets recouverts par des matériaux (notamment propagation d'ondes sismiques par exemple).

9.3 SUR LES CONDITIONS DE TRANSMISSION

Nous l'avons vu dans la section concernant l'état de l'art, beaucoup d'améliorations ont été apportées depuis le développement des conditions de transmission de Després. Des opérateurs tangentiels ont tout d'abord été ajoutés (conditions de type $\partial_n u + ik u + \beta \partial_{tt} u$) et enfin l'utilisation de fractions rationnelles de type

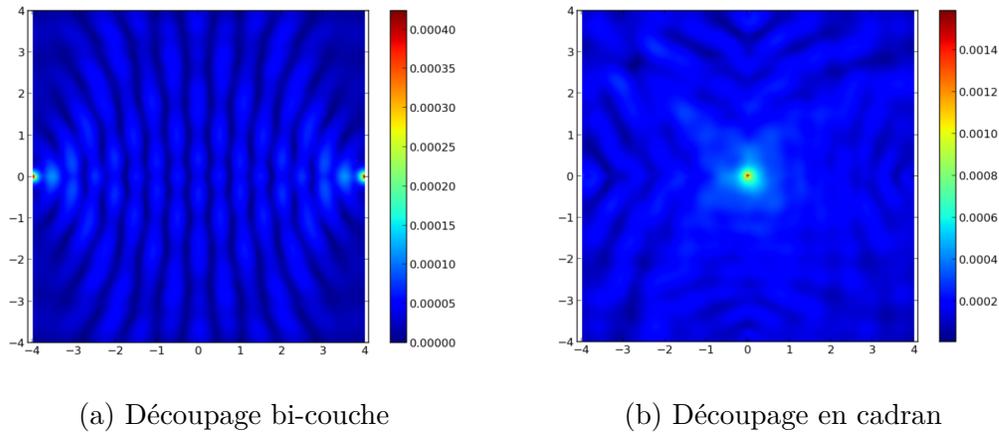


FIGURE 9.2 – Erreur par rapport à la solution après 1000 itérations de l’algorithme de Jacobi (la convergence n’est pas atteinte). Le rayon du trou dans 9.2b est de $r = 0.001$.

Padé a permis d’améliorer significativement la convergence des méthodes de décomposition de domaine.

L’ajout d’opérateurs impédants ayant des propriétés d’isomorphisme particulières ouvre de nouvelles perspectives dans la construction de conditions de transmission assurant une convergence rapide de méthodes de décomposition de domaine. L’objectif serait de combiner ces conditions de transmission pour obtenir une convergence indépendante du raffinement tout en profitant de la rapidité de convergence des opérateurs de type fractions rationnelles lorsque le maillage est assez grossier. La manière de combiner ces différentes conditions de transmission reste pour l’instant inconnue, mais requerrait de conserver la propriété d’isomorphisme de l’opérateur impédant.

De nombreuses questions relatives aux conditions de transmission pour les méthodes de décomposition de domaine restent ouvertes. Par exemple, nous pouvons légitimement nous demander ce que pourrait donner des conditions de transmission évoluant avec les itérations

- Opérateurs alternant avec les itérations : nous pourrions utiliser par exemple des opérateurs locaux pour les itérations paires, et des opérateurs non locaux pour les itérations impaires. La convergence peut-elle être prouvée ? Qu’en est-il de la convergence exponentielle ?

- Opérateurs différents à partir d'une certaine itération : l'objectif serait d'utiliser des conditions de transmission locales (de type Padé par exemple) en début de convergence pour profiter de leur performance sur les modes propagatifs et faiblement évanescents, puis utiliser ensuite des conditions non locales comme celles que nous avons développées pour traiter les ondes très évanescentes.

Bien évidemment, ces méthodes ne rentrent pas dans la théorie développée dans cette thèse, mais ces questions nous semblent suffisamment intéressantes pour être mentionnées ici.

9.4 RETOUR SUR L'ANALYSE MODALE

Pour finir, nous mentionnons ici un fait qui nous a surpris et qui ouvre aussi de nouvelles perspectives. Rappelons brièvement quelques éléments de l'analyse modale (voir section 1 du chapitre 3).

Dans le cas de la géométrie circulaire simple (dans le vide), nous avons pu donner une estimation du taux de convergence modal. Celui-ci est directement lié aux normes de valeurs propres et s'écrit (voir équation (3.23))

$$\tau_{r,m} = |1 - r + r\sqrt{s_{1,m}s_{2,m}}| \quad (9.1)$$

où les coefficients de *scattering* sont donnés par (3.10)

$$s_{1,m} = -\frac{Q_{1,m} - i\bar{t}_m}{Q_{1,m} + it_m}, \quad s_{2,m} = -\frac{Q_{2,m} - it_m}{Q_{2,m} + i\bar{t}_m}. \quad (9.2)$$

Nous en déduisons l'asymptotique suivante (3.17)

$$s_{1,m}s_{2,m} \sim \frac{\frac{m}{kR_\Sigma} - i\bar{t}_m}{\frac{m}{kR_\Sigma} + it_m} \frac{\frac{m}{kR_\Sigma} - it_m}{\frac{m}{kR_\Sigma} + i\bar{t}_m} \quad (9.3)$$

et en concluons que l'unique manière d'obtenir la convergence exponentielle (liée à la condition $\lim s_{1,m}s_{2,m} \neq 1$) est d'avoir un symbole t_m équivalent à m quand $m \rightarrow \infty$. En effet, toute autre asymptotique pour t_m conduit à une limite $\lim s_{1,m}s_{2,m} = 1$ et donc à une perte de la convergence exponentielle.

Supposons maintenant que nous nous libérons de la contrainte $T_1 = T$ et $T_2 = T^*$,

c'est à dire que nous nous autorisons à prendre deux opérateurs différents de part et d'autre de l'interface. Dans ces conditions, et en notant $t_{1,m}$ et $t_{2,m}$ les symboles respectifs de T_1 et T_2 , il est aisé d'obtenir les nouveaux coefficients de *scattering*

$$s_{1,m} = -\frac{Q_{1,m} - it_{2,m}}{Q_{1,m} + it_{1,m}}, \quad s_{2,m} = -\frac{Q_{2,m} - it_{1,m}}{Q_{2,m} + it_{2,m}} \quad (9.4)$$

et l'asymptotique, pour $m \rightarrow \infty$

$$s_{1,m}s_{2,m} \sim \frac{\frac{m}{kR_\Sigma} - it_{2,m}}{\frac{m}{kR_\Sigma} + it_{1,m}} \frac{\frac{m}{kR_\Sigma} - it_{1,m}}{\frac{m}{kR_\Sigma} + it_{2,m}} = \frac{\frac{m}{kR_\Sigma} - it_{2,m}}{\frac{m}{kR_\Sigma} + it_{2,m}} \frac{\frac{m}{kR_\Sigma} - it_{1,m}}{\frac{m}{kR_\Sigma} + it_{1,m}}. \quad (9.5)$$

Dans ces conditions, il n'est plus nécessaire d'utiliser des symboles $t_{1,m}$ et $t_{2,m}$ équivalent à m quand $m \rightarrow \infty$ pour obtenir une limite $\lim s_{1,m}s_{2,m}$ différente de 1. Par exemple, en prenant $t_{1,m} = 1$ et $t_{2,m} \sim m^2$ (ce qui correspond à prendre des conditions de Després dans un sous-domaine et des conditions d'ordre 2 dans l'autre sous-domaine), on obtient

$$\lim_{m \rightarrow \infty} s_{1,m}s_{2,m} = -1 \quad (9.6)$$

et

$$\lim_{m \rightarrow \infty} \tau_{r,m} = |1 - r + ri| = \sqrt{1 - 2r + 2r^2} \quad (9.7)$$

et pour $0 < r < 1$, on a $\lim \tau_{r,m} < 1$. Cet unique argument ne permet pas de conclure quant à la convergence de la méthode, et encore moins de la convergence exponentielle. Cependant, lorsque $T_1 = T_2$ ou $T_1 = T_2^*$ (ce qui couvre la quasi totalité des conditions de transmission de la littérature), nous avons montré au moins dans le cas de la géométrie circulaire que l'utilisation d'opérateurs locaux (de type Després, d'ordre 2 ou de type fractions rationnelles de Padé) empêche la convergence exponentielle car $\lim \tau_{r,m} = 1$. En s'autorisant des opérateurs T_1 et T_2 différents, cette conclusion n'est plus vraie et le fait d'avoir un taux de convergence strictement inférieur à 1 pour les modes très évanescents (toujours pour ce cas particulier) semble suggérer qu'il serait possible d'avoir une convergence exponentielle tout en utilisant des opérateurs locaux.

Cette observation ouvre ainsi de nombreuses perspectives mais une nouvelle théorie doit être développée pour répondre aux questions que nous nous étions posées au début de ce document, à savoir

- Les problèmes locaux sont-ils bien posés ?
- Sommes-nous assurés de la convergence ?
- La convergence est-elle effectivement exponentielle ?

A

Annexes

Sommaire

A.1	Développement limité du noyau $B(s, t)$ en $s = t$	233
A.2	Calcul des coefficients	236
A.2.1	Un argument d'analycité	236
A.2.2	Calcul par déformation de contour	238

A.1 DÉVELOPPEMENT LIMITÉ DU NOYAU $B(s, t)$ EN $s = t$

L'objet de cette section est de montrer l'intégrabilité du noyau

$$B(s, t) = \frac{\partial^2(N_\Sigma - N_C)}{\partial t \partial s}(s, t), \quad (\text{A.1})$$

où N_Σ et N_C sont défini en (4.23). Plus précisément, nous allons montrer que

$$B(s, t) = o\left(\frac{1}{|s - t|^{\frac{1}{2}}}\right). \quad (\text{A.2})$$

En $s = t$, et pourvu que l'abscisse curviligne x soit c^2 , les développements limités de x et \dot{x} donnent

$$\begin{aligned} x(s) &= x(t) + (s-t)\dot{x}(t) + \frac{(s-t)^2}{2}\ddot{x}(t) + o((s-t)^2) \\ \dot{x}(s) &= \dot{x}(t) + (s-t)\ddot{x}(t) + o(s-t) \\ x(s) - x(t) &= (s-t) \left[\dot{x}(t) + \frac{s-t}{2}\ddot{x}(t) + o(s-t) \right] \\ |x(s) - x(t)|^2 &= |s-t|^2 \left[|\dot{x}(t)|^2 + \dot{x}(t) \cdot \ddot{x}(t)|s-t| + o(|s-t|) \right] \end{aligned} \quad (\text{A.3})$$

Puisque x est une abscisse curviligne, il vient $|\dot{x}(t)|^2 = 1$, et par dérivation $\dot{x}(t) \cdot \ddot{x}(t) = 0$. D'où

$$|x(s) - x(t)|^2 = |s-t|^2 [1 + o(|s-t|)] \quad (\text{A.4})$$

et finalement, on obtient le développement limité de R

$$R = |x(s) - x(t)| = |s-t| [1 + o(|s-t|)]. \quad (\text{A.5})$$

En utilisant les équations (A.3) et (A.5), on peut calculer les développements limités des produits de l'équation (4.27) :

$$\begin{aligned} K_1 &= \frac{x(s) - x(t)}{R} \cdot \dot{x}(s) = \frac{(s-t) \left[\dot{x}(t) + \frac{s-t}{2}\ddot{x}(t) + o(s-t) \right]}{|s-t| [1 + o(|s-t|)]} \cdot \dot{x}(s) \\ &= \frac{s-t}{|s-t|} \left[\dot{x}(t) + \frac{s-t}{2}\ddot{x}(t) + o(s-t) \right] \cdot \dot{x}(s) \end{aligned} \quad (\text{A.6})$$

De plus, par le développement limité de $\dot{x}(s)$ (A.3), on a les produits scalaires suivant :

$$\begin{aligned} \dot{x}(t) \cdot \dot{x}(s) &= 1 + o(s-t) \\ \ddot{x}(t) \cdot \dot{x}(s) &= o(1) \end{aligned} \quad (\text{A.7})$$

En injectant ces deux dernières équations dans l'expression de K_1 , on obtient :

$$K_1 = \frac{s-t}{|s-t|} [1 + o(|s-t|)] \quad (\text{A.8})$$

De même,

$$\begin{aligned}
K_2 &= \frac{x(s) - x(t)}{R} \cdot \dot{x}(t) = \frac{(s-t) \left[\dot{x}(t) + \frac{s-t}{2} \ddot{x}(t) + o(s-t) \right]}{|s-t| [1 + o(|s-t|)]} \cdot \dot{x}(t) \\
&= \frac{s-t}{|s-t|} \left[\dot{x}(t) + \frac{s-t}{2} \ddot{x}(t) + o(s-t) \right] \cdot \dot{x}(t) \\
&= \frac{s-t}{|s-t|} [1 + o(s-t)]
\end{aligned} \tag{A.9}$$

et

$$K_1 K_2 = 1 + o(|s-t|) \tag{A.10}$$

De l'équation (4.27), il vient alors :

$$\begin{aligned}
\frac{\partial^2 N_\Sigma}{\partial s \partial t} &= \frac{1}{4R^{\frac{3}{2}}} [-2 + 3 + o(|s-t|)] = \frac{1}{4R^{\frac{3}{2}}} [1 + o(s-t)] \\
&= \frac{1}{4|s-t|^{\frac{3}{2}}} [1 + o(s-t)]
\end{aligned} \tag{A.11}$$

Calculons maintenant le développement limité de $\frac{\partial^2 N_C}{\partial s \partial t}$

$$\begin{aligned}
\frac{\partial^2 N_C}{\partial t \partial s} &= \frac{2 - \cos^2 \frac{|s-t|}{a}}{4a^{\frac{3}{2}} \sin^{\frac{3}{2}} \frac{|s-t|}{a}} = \frac{1}{4a^{\frac{3}{2}}} \frac{2 - (1 + o(|s-t|))^2}{\left(\frac{|s-t|}{a} + o(|s-t|^2) \right)^{\frac{3}{2}}} \\
&= \frac{1}{4|s-t|^{\frac{3}{2}}} [1 + o(|s-t|)]
\end{aligned} \tag{A.12}$$

Puis en injectant ces résultats (A.11 et A.12) dans l'équation (A.1), on obtient le développement limité voulu :

$$\begin{aligned}
B(s, t) &= \frac{\partial^2 N_\Sigma}{\partial s \partial t} - \frac{\partial^2 N_C}{\partial s \partial t} = \frac{1}{4|s-t|^{\frac{3}{2}}} [1 - 1 + o(s-t)] \\
&= o\left(\frac{1}{|s-t|^{\frac{1}{2}}} \right)
\end{aligned} \tag{A.13}$$

A.2 CALCULS DES COEFFICIENTS DE FOURIER DES OPÉRATEURS DE GAGLIARDO

Soit α un nombre plus petit que 1 (éventuellement négatif) ; le but de cet annexe est de calculer explicitement

$$\hat{\Lambda}_p^{(\alpha)} = \int_{-\pi}^{\pi} \frac{1 - \cos p\theta}{(2 - 2\cos\theta)^{\frac{1}{2}+\alpha}} d\theta. \quad (\text{A.14})$$

pour tout $p > 0$ et d'établir la formule

$$\hat{\Lambda}_p^{(\alpha)} = -2 \cos(\alpha\pi) \Gamma(-2\alpha) \left(\frac{\Gamma(p + \frac{1}{2} + \alpha)}{\Gamma(p + \frac{1}{2} - \alpha)} - \frac{\Gamma(\frac{1}{2} + \alpha)}{\Gamma(\frac{1}{2} - \alpha)} \right) := L_p(\alpha) \quad (\text{A.15})$$

A.2.1 UN ARGUMENT D'ANALYCITÉ

Fixons $p > 0$, l'idée du calcul est d'établir une formule valable pour α assez petit négatif, puis de l'étendre par exemple à des $\alpha > 0$ en utilisant des arguments d'analycité. Considérons donc

$$\alpha \in U \mapsto \hat{\Lambda}_p^{(\alpha)}, \quad (\text{A.16})$$

où U est l'ouvert du plan complexe

$$U = \{\alpha \in \mathbb{C}, \operatorname{Re} \alpha > -1\}. \quad (\text{A.17})$$

Pour θ non nul dans $[-\pi, \pi]$, l'application

$$\alpha \in U \mapsto f(\theta, \alpha) = \frac{1 - \cos p\theta}{(2 - 2\cos\theta)^{\frac{1}{2}+\alpha}} \quad (\text{A.18})$$

est analytique sur U ; de plus, en utilisant

$$|1 - \cos p\theta| \leq \frac{p^2 \theta^2}{2} \quad \text{et} \quad \frac{1}{\sqrt{2 - 2\cos\theta}} \leq \frac{\pi}{2|\theta|} \quad (\text{A.19})$$

on a

$$|f(\theta, \alpha)| \leq \frac{p^2}{2} \left(\frac{\pi}{2}\right)^{|1+2\alpha|} |\theta|^{1-2\operatorname{Re} \alpha} \quad (\text{A.20})$$

Si K est un compact de U , on a $|\alpha| \leq C_K$ et $\operatorname{Re} \alpha \geq \alpha_K \geq -1$ et

$$|f(\theta, \alpha)| \leq \frac{p^2}{2} \left(\frac{\pi}{2}\right)^{1+2C_K} |\theta|^{1-2\alpha_K} = f_K(\theta) \quad (\text{A.21})$$

et comme $1 - 2\alpha_K > -1$, $f_K(\theta) \in L^1([-\pi, \pi])$. On a donc une famille de fonctions $f(\theta, \alpha)$, analytique presque partout en $\alpha \in U$ et dominée, uniformément sur tout compact de U , par une fonction intégrable. La théorie nous indique alors que $\alpha \mapsto \int_{-\pi}^{\pi} f(\theta, \alpha) d\theta$ est analytique sur U . Soit β un nombre dans $]0, 1[$, on pose

$$\hat{F}_p^{(\beta)} = \int_{-\pi}^{\pi} \frac{e^{ip\theta}}{(2 - 2\cos\theta)^{\frac{1-\beta}{2}}} d\theta. \quad (\text{A.22})$$

Dans le paragraphe suivant, nous établirons la formule

$$\hat{F}_p^{(\beta)} = 2 \cos\left(\beta \frac{\pi}{2}\right) \Gamma(\beta) \frac{\Gamma\left(p + \frac{1-\beta}{2}\right)}{\Gamma\left(p + \frac{1+\beta}{2}\right)}, \quad \beta \in]0, 1[\quad (\text{A.23})$$

où $\Gamma(z)$ est la fonction Gamma. On admettra cette formule pour le moment. Si α est dans l'intervalle $] -\frac{1}{2}, 0[$, alors -2α est bien dans $]0, 1[$. Or on a,

$$\hat{\Lambda}_p^\alpha = \hat{F}_0^{-2\alpha} - \hat{F}_p^{-2\alpha}, \quad \alpha \in] -\frac{1}{2}, 0[\quad (\text{A.24})$$

et donc, toujours si α est dans $] -\frac{1}{2}, 0[$,

$$\hat{\Lambda}_p^{(\alpha)} = -2 \cos(\alpha\pi) \Gamma(-2\alpha) \left(\frac{\Gamma\left(p + \frac{1}{2} + \alpha\right)}{\Gamma\left(p + \frac{1}{2} - \alpha\right)} - \frac{\Gamma\left(\frac{1}{2} + \alpha\right)}{\Gamma\left(\frac{1}{2} - \alpha\right)} \right) := L_p(\alpha) \quad (\text{A.25})$$

Maintenant, si l'on regarde l'expression de $\hat{\Lambda}_p^{(\alpha)}$, puisque la fonction $z \mapsto \Gamma(z)$ est analytique sur \mathbb{C} privé des entiers négatifs ou nuls, $L_p(\alpha)$ est analytique sur U sauf, éventuellement, en $\alpha = 0$ ou $\alpha = \frac{1}{2}$. Considérons l'ouvert U_ϵ constitué de U privé de deux petites boules centrées autour de $\alpha = 0$ ou $\alpha = \frac{1}{2}$, on dispose alors de deux formules : la formule intégrale $\int_{-\pi}^{\pi} f(\theta, \alpha) d\theta$ et la formule $L_p(\alpha)$, qui sont toutes deux analytiques sur U_ϵ et qui coïncident sur un petit intervalle. On en déduit qu'elles coïncident sur U_ϵ . Enfin, comme la forme intégrale est continue en $\alpha = \frac{1}{2}$ ou $\alpha = 0$, la formule impliquant les fonctions Gamma se prolongera forcément par continuité. Par exemple, en $\alpha = \frac{1}{2}$, on aura

$$\lim_{\alpha \rightarrow \frac{1}{2}} L_p(\alpha) = p\pi = \hat{\Lambda}_p^{(\frac{1}{2})} \quad (\text{A.26})$$

En $\alpha = 0$ la formule est plus complexe, comme $x \Gamma(x) \rightarrow 1$ quand $x \rightarrow 0$, il vient

$$\lim_{\alpha \rightarrow 0} L_p(\alpha) = 2 \left(\psi' \left(p + \frac{1}{2} \right) - \psi' \left(\frac{1}{2} \right) \right) = \hat{\Lambda}_p^{(0)}, \quad (\text{A.27})$$

avec

$$\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)} \quad (\text{A.28})$$

cette fonction croît avec x comme le logarithme de x .

A.2.2 CALCUL PAR DÉFORMATION DE CONTOUR

Soit β un nombre dans $]0, 1[$, on veut montrer que

$$\hat{F}_p^{(\beta)} = \int_{-\pi}^{\pi} \frac{e^{ip\theta}}{(2 - 2\cos\theta)^{\frac{1-\beta}{2}}} d\theta. \quad (\text{A.29})$$

est donné par la formule

$$\hat{F}_p^{(\beta)} = 2 \cos\left(\beta \frac{\pi}{2}\right) \Gamma(\beta) \frac{\Gamma\left(p + \frac{1-\beta}{2}\right)}{\Gamma\left(p + \frac{1+\beta}{2}\right)}, \quad \beta \in]0, 1[\quad (\text{A.30})$$

où $\Gamma(z)$ est la fonction Gamma. Pour ramener l'intégrale à une intégrale connue, on l'interprète comme une intégrale de contour que l'on va déformer.

$$\hat{F}_p^{(\beta)} = \int_{[-\pi, \pi]} g(z) dz. \quad (\text{A.31})$$

Pour cela, on définit la fonction dans le plan complexe

$$g : z \mapsto 2^{\frac{\beta-1}{2}} e^{\frac{\beta-1}{2} \log(1-\cos z)} e^{ipz} \quad (\text{A.32})$$

où le log est défini *via* la détermination principale

$$\log(1 - \cos z) = \log |1 - \cos z| + i \arg(1 - \cos z), \quad \arg t \in]-\pi, \pi[\quad (\text{A.33})$$

La fonction n'est donc pas définie (elle présente un saut) sur les lignes pour lesquelles $1 - \cos z$ est réel négatif, soit $z = i\xi + 2k\pi\mathbb{Z}$, $\xi \in \mathbb{R}$: la droite des imaginaires purs est une ligne de singularité dans le domaine $\Omega = \{z \in \mathbb{C}, \operatorname{Re} z \in]-\pi, \pi[\}$.

Si $z = \epsilon + i\xi = i(\xi - i\epsilon)$, on a

$$1 - \cos z = -(\cosh \xi - 1) + i\epsilon \sinh \xi + O(\epsilon^2) \quad (\text{A.34})$$

et l'argument du log va tendre vers π si z tend vers l'axe imaginaire avec $\epsilon \operatorname{Im} z > 0$ et vers $-\pi$ pour $\epsilon \operatorname{Im} z < 0$. Nous avons représenté figure A.1 le contour sur lequel nous effectuons l'intégration. Maintenant, on a par application des formules de Cauchy,

$$\int_{-\pi}^{0-} g(z) dz = \int_{\gamma_{-\pi}^{\uparrow}} g(z) dz + \int_{\gamma_{-\pi+i\infty}^{\rightarrow}} g(z) dz + \int_{\gamma_{0-+i\infty}^{\downarrow}} g(z) dz \quad (\text{A.35})$$

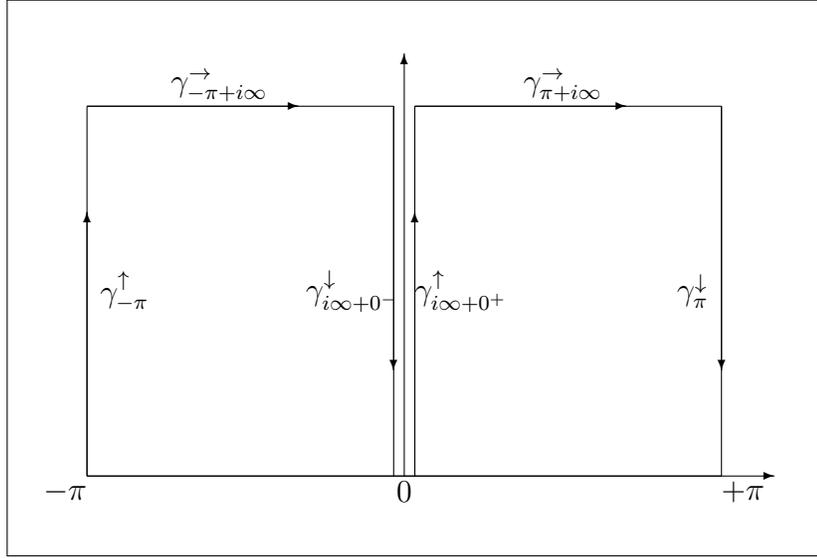


FIGURE A.1 – Schéma représentant le contour d'intégration.

$$\int_{0+}^{\pi} g(z)dz = \int_{\gamma_{\pi}^{\downarrow}} g(z)dz + \int_{\gamma_{\pi+i\infty}^{\leftarrow}} g(z)dz + \int_{\gamma_{i\infty+0+}^{\uparrow}} g(z)dz \quad (\text{A.36})$$

Si l'on somme les deux résultats, les intégrales sur $\gamma_{-\pi}^{\uparrow}$ et $\gamma_{\pi}^{\downarrow}$ s'éliminent par périodicité, celles sur $\gamma_{-\pi+i\infty}^{\rightarrow}$ et sur $\gamma_{\pi+i\infty}^{\leftarrow}$ sont nulles car g est exponentiellement décroissante avec la partie imaginaire de z ¹, il reste les deux branches qui donnent

$$\int_{-\pi}^{\pi} g(z)dz = \int_0^{\infty} [g(i\xi + 0^-) - g(i\xi + 0^+)] id\xi \quad (\text{A.37})$$

On calcule le saut, on a

$$[ig(i\xi + 0^-) - ig(i\xi + 0^+)] = e^{-p\xi} (2 \cosh \xi - 2)^{\frac{-1+\beta}{2}} i \left(e^{-i\frac{1-\beta}{2}\pi} - ie^{i\frac{1-\beta}{2}\pi} \right) \quad (\text{A.38})$$

soit

$$i[g(i\xi + 0^-) - g(i\xi + 0^+)] = 2e^{-p\xi} (2 \cosh \xi - 2)^{\frac{-1+\beta}{2}} \sin \left(\frac{1-\beta}{2}\pi \right) \quad (\text{A.39})$$

et ainsi

$$\hat{F}_p^{(\beta)} = 2 \cos\left(\frac{\beta\pi}{2}\right) \int_0^{\infty} e^{-p\xi} (2 \cosh \xi - 2)^{\frac{-1+\beta}{2}} d\xi \quad (\text{A.40})$$

1. la propriété $\beta < p + 1$ est **essentielle** ici

On pose $x = e^{-\xi}$, et on utilise $2 \cosh \xi - 2 = x^{-1}(1-x)^2$, il reste

$$\hat{F}_p^{(\beta)} = 2 \cos\left(\frac{\beta\pi}{2}\right) \int_0^1 x^{p-\frac{\beta-1}{2}-1} (1-x)^{\beta-1} dx \quad (\text{A.41})$$

soit

$$\hat{F}_p^{(\beta)} = 2 \cos\left(\frac{\beta\pi}{2}\right) \frac{\Gamma(\beta)\Gamma\left(p + \frac{1-\beta}{2}\right)}{\Gamma\left(p + \frac{\beta+1}{2}\right)} = 2 \cos\left(\frac{\beta\pi}{2}\right) \Gamma(\beta) \frac{\Gamma\left(p + \frac{1}{2} - \frac{\beta}{2}\right)}{\Gamma\left(p + \frac{1}{2} + \frac{\beta}{2}\right)} \quad (\text{A.42})$$

ce qui est la formule annoncée. De fait, on a utilisé

$$\int_0^1 x^{a-1} (1-x)^{b-1} dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad (\text{A.43})$$

cette relation est facile à montrer car

$$\Gamma(a)\Gamma(b) = \int_0^\infty \int_0^\infty e^{-\xi_1 - \xi_2} \xi_1^{a-1} \xi_2^{b-1} d\xi_1 d\xi_2 \quad (\text{A.44})$$

ou ($t = \xi_1 + \xi_2$)

$$\Gamma(a)\Gamma(b) = \int_0^\infty \int_0^t e^{-t} \xi_1^{a-1} (t - \xi_1)^{b-1} d\xi_1 dt \quad (\text{A.45})$$

ou ($\xi_1 = tx$)

$$\Gamma(a)\Gamma(b) = \int_0^\infty \int_0^1 e^{-t} t^{a+b-1} x^{a-1} (1-x)^{b-1} dx dt \quad (\text{A.46})$$

et

$$\Gamma(a)\Gamma(b) = \Gamma(a+b) \int_0^1 x^{a-1} (1-x)^{b-1} dx. \quad (\text{A.47})$$

Bibliographie

- [1] E. Anderson, Z. Bai, C. Bischof, L. S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. Lapack users' guide, 1999.
- [2] John Macleod Ball, Yves Capdeboscq, and Basang Tsering Xiao. On uniqueness for time harmonic anisotropic Maxwell's equations with piecewise regular coefficients. *Mathematical Models and Methods in Applied Sciences*, 22(11) :1–9, 2012.
- [3] Mario Bebendorf. *Hierarchical Matrices : A Means to Efficiently Solve Elliptic Boundary Value Problems*. Springer Berlin Heidelberg, 2008.
- [4] Yassine Boubendir. An analysis of the BEM-FEM non-overlapping domain decomposition method for a scattering problem. *Journal of Computational Applied Mathematics*, 204 :282–291, 2007.
- [5] Yassine Boubendir, Xavier Antoine, and Christophe Geuzaine. A quasi-optimal non-overlapping domain decomposition algorithm for the Helmholtz equation. *Journal of Computational Physics*, 231 :262–280, 2012.
- [6] Robert Brayton, Stephen Director, Gary Hachtel, and Luis Vidigal. A new algorithm for statistical circuit design based on quasi-newton methods and function splitting. *IEEE Transactions on Circuits and Systems*, 26(9) :784–794, 1979.
- [7] Eliane Bécache, Ciarlet Patrick, Hazard Christophe, and Lunéville Eric. *La Méthode des Eléments Finis. De la Théorie à la Pratique. II. Compléments*. Les Presses de l'ENSTA, 2010.
- [8] Xiao-Chuan Cai, Mario Casarin, Franck Elliott, and Olof Widlund. Overlapping schwarz algorithms for solving Helmholtz's equation. *Contemporary Mathematics*, 218 :391–399, 1998.

- [9] Francis Collino, Souad Ghanemi, and Patrick Joly. Domain decomposition method for harmonic wave propagation : a general presentation. *Computer Methods in Applied Mathematics*, 184 :171–211, 2000.
- [10] Bruno Després. *Methodes de decomposition de domaine pour les problemes de propagation d'ondes en regime harmonique : le theoreme de Borg pour l'equation de Hill vectorielle P*. PhD thesis, Paris 9, 1991.
- [11] Charbel Farhat, Michel Lesoinne, Patrick LeTallec, Kendall Pierson, and Daniel Rixen. Feti-dp : a dual–primal unified feti method–part i : A faster alternative to the two-level feti method. *International Journal for Numerical Methods in Engineering*, 50 :1523–1544, 2001.
- [12] Charbel Farhat and François-Xavier Roux. A method of finite element tearing and interconnecting and its parallel solution algorithm. *International Journal for Numerical Methods in Engineering*, 32 :1205–1227, 1991.
- [13] Martin Gander. Schwarz methods in the course of time. *Electronic Transactions on Numerical Analysis*, 31 :228–255, 2008.
- [14] Martin Gander, Laurence Halpern, and Frédéric Magoulès. An optimized Schwarz method with two-sided Robin transmission conditions for the Helmholtz equation. *International Journal for Numerical Methods in Fluids*, 55 :163–175, 2007.
- [15] Martin Gander, Frédéric Magoulès, and Frédéric Nataf. Optimized Schwarz methods without overlap for the Helmholtz equation. *SIAM Journal on Scientific Computing*, 24 :38–60, 2002.
- [16] Christophe Geuzaine and Jean-François Remacle. Gmsh : A 3-d finite element mesh generator with built-in pre-and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11) :1309–1331, 2009.
- [17] Vivette Girault and Pierre-Arnaud Raviart. *Finite Element Methods for Navier-Stokes Equations*. Springer-Verlag Berlin Heidelberg, 1986.
- [18] Dan Givoli. Computational absorbing boundaries. In *Computational Acoustics of Noise Propagation in Fluids - Finite and Boundary Element Methods*. Marburg, Steffen and Nolte, Bodo, 2008.

- [19] S.P. Han. A globally convergent method for nonlinear programming. *Journal of Optimization Theory and Applications*, 22(3) :297–309, 1977.
- [20] Pascal Hénon, Pierre Ramet, and Jean Roman. PaStiX : A High-Performance Parallel Direct Solver for Sparse Symmetric Definite Systems. *Parallel Computing*, 28(2) :301–321, 2002.
- [21] Riyad Kechroud, Xavier Antoine, and Abderrahmane Soulaïmani. Numerical accuracy of a padé-type non-reflecting boundary condition for the finite element solution of acoustic scattering problems at high-frequency. *International Journal for Numerical Methods in Engineering*, 00 :1–6, 2000.
- [22] Jin-Au Kong. *Electromagnetic wave theory*. John Wiley & Sons, Inc, 1986.
- [23] Xavier Lacoste, Mathieu Faverge, Pierre Ramet, Samuel Thibault, and George Bosilca. Taking advantage of hybrid systems for sparse direct solvers via task-based runtimes. In *Proceedings of HCW'2014 workshop of IPDPS*, 2014.
- [24] Seung-Cheol Lee, Marinos Vouvakis, and Jin-Fa Lee. A non-overlapping domain decomposition method with non-matching grids for modeling large finite antenna arrays. *Journal of Computational Physics*, 203 :1–21, 2005.
- [25] Pierre-Louis Lions. On the Schwarz alternating method. I. In *First international symposium on domain decomposition methods for partial differential equations*, pages 1–42. Paris, France, 1988.
- [26] Yvon Maday and Frédéric Magoulès. Optimized Schwarz methods without overlap for highly heterogeneous media. *Computer Methods in Applied Mathematics and Engineering*, 196 :1541–1553, 2007.
- [27] Frédéric Magoulès, Peter Iványi, and Barry H.V. Topping. Non-overlapping Schwarz methods with optimized transmission conditions for the Helmholtz equation. *Computer Methods in Applied Mathematics and Engineering*, 193 :4794–4818, 2004.
- [28] Frédéric Magoulès, François-Xavier Roux, and Stéphanie Salmon. Optimal discrete transmission conditions for a non-overlapping domain decomposition method for the Helmholtz equation. *SIAM Journal on Scientific Computing*, 25 :1497–1515, 2004.

- [29] Ronald B. Morgan. A restarted gmres method augmented with eigenvectors. *SIAM Journal on Matrix Analysis and Applications*, 16(4) :1154–1171, 1995.
- [30] Jean-Claude Nedelec. *Acoustic and Electromagnetic Equations : Integral Representations for Harmonic Problems*. Springer-Verlag, 2001.
- [31] Zhen Peng and Jin-Fa Lee. Non-conformal domain decomposition method with second order transmission conditions for time-harmonic electromagnetics. *Journal of Computational Physics*, 229 :5615–5629, 2010.
- [32] Zhen Peng, Vineet Rawat, and Jin-Fa Lee. One way domain decomposition method with second order transmission conditions for solving electromagnetic wave problems. *Journal of Computational Physics*, 229 :1181–1197, 2010.
- [33] M.J.D. Powell. A fast algorithm for nonlinearly constrained optimization calculations. In G.A. Watson, editor, *Numerical Analysis*, volume 630 of *Lecture Notes in Mathematics*, pages 144–157. Springer Berlin Heidelberg, 1978.
- [34] Youcef Saad. *Iterative methods for sparse linear systems : Second edition*. Society for Industrial and Applied Mathematics, 2003.
- [35] Youcef Saad and Martin H. Schultz. Gmres : a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3) :856–869, 1986.
- [36] Nicolas Salles. *Calculation of singularities in variational integral equations methods*. PhD thesis, Université Paris Sud - Paris XI, Sept 2013.
- [37] Hermann Amandus Schwarz. Über einen grenzübergang durch alternierendes verfahren. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, 15 :272–286, 1870.
- [38] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton mathematical series. Princeton University Press, 1970.
- [39] Christiaan Stolk. A rapidly converging domain decomposition method for the Helmholtz equation. *Journal of Computational Physics*, 241 :240–252, 2013.

-
- [40] Bruno Stupfel. Absorbing boundary conditions on arbitrary boundaries for the scalar and vector wave equations. *IEEE Transactions on Antennas and Propagation*, 42 :773–780, 1994.
- [41] Bruno Stupfel. Improved transmission conditions for a one-dimensional domain decomposition method applied to the solution of the Helmholtz equation. *Journal of Computational Physics*, 229 :851–874, 2010.
- [42] Marinos Vouvakis, Zoltan Cendes, and Jin-Fa Lee. A FEM domain decomposition method for phonic and electromagnetic band gap structure. *IEEE Transactions on Antennas and Propagation*, 54 :721–733, 2006.
- [43] Marinos Vouvakis, Kezhong Zhao, Seung-Mo Seo, and Jin-Fa Lee. A domain decomposition approach for non-conformal couplings between finite and boundary elements for unbounded electromagnetic problems in R^3 . *Journal of Computational Physics*, 225 :975–994, 2007.