



**HAL**  
open science

# Méthodes d'ondelettes pour l'analyse numérique d'intégrales oscillantes

Daniel Goujot

► **To cite this version:**

Daniel Goujot. Méthodes d'ondelettes pour l'analyse numérique d'intégrales oscillantes. Mathématiques [math]. Université d'Evry-Val d'Essonne, 2004. Français. NNT : . tel-00008353

**HAL Id: tel-00008353**

**<https://theses.hal.science/tel-00008353>**

Submitted on 6 Apr 2005

**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 DE DOCTORAT

présentée à

**UNIVERSITÉ D'ÉVRY  
VAL-D'ESSONNE**

pour obtenir le titre de

**DOCTEUR DE L'UNIVERSITÉ  
D'ÉVRY VAL-D'ESSONNE**

Spécialité :

**Mathématiques appliquées**

soutenue par

**Daniel GOUJOT**

le 10 décembre 2004

titre

**Méthodes d'ondelettes pour  
l'analyse numérique  
d'intégrales oscillantes**

mention très honorable avec les félicitations du Jury :

M. Pierre-Gilles Lemarié-Rieusset	Directeur de thèse
M. Martin Costabel	Rapporteur
M. Jacques Laminie	Rapporteur
M. Yves Meyer	Président
M. Djalil Kateb	Examineur



## Résumé

Nous utilisons trois discrétisations connues pour leur localisation fréquentielle et spatiale: les bases d'ondelettes, les paquets d'ondelettes et les bases de cosinus locaux. Nous avons construit et programmé deux algorithmes:

— pour l'équation parabolique non-linéaire  $\Delta(u) + e^{c^*u} = f$  avec  $f$  présentant une singularité, notre algorithme calcule la compression optimale en dimension 1 et 2, avec résultats numériques pour la dimension 1.

— pour l'équation intégrale oscillante correspondant à la Combined Integral Field Equation qui est en rapport avec le problème de diffraction des ondes (Helmholtz) par un obstacle régulier 2D, lorsque la longueur d'onde diminue vers 0. Les trois discrétisations ci-dessus sont testées, et nous étudions sa bonne compressibilité dans une analyse précise des obstacles à la compression menée de manière asymptotique. Des résultats originaux, montrant que  $N$  degrés de liberté par longueur d'onde suffisent à hautes fréquences, ont été démontrés, et les matrices résultant de ce seuillage ont été étudiées, illustrations et preuves à l'appui.

## Abstract

We have worked with three families of functions known for their localisation in space and in frequency: the wavelet basis, the wavelet packet basis, and the Malvar local cosine basis. We have constructed and implemented two algorithms:

— for the equation  $\Delta(u) + e^{c^*u} = f$  with  $f$  having a tough singularity, we have proved sharp complexity estimates for the algorithm in 1D and 2D, in the wavelet basis, and have reached high precisions in 1D numerical results.

— for the oscillating boundary integral equation known as Combined Field Integral Equation, related to the Helmholtz diffraction problem around a 2D smooth obstacle, in the high frequency context. All above three function families have been tested, and we make a sharp analysis of the compressibility of the resulting system asymptotically for a decreasing wavelength. We have proved an original result, fully justifying the use of  $N$  degrees of freedom per wavelength. We have explained and proved the graphical repartitions in the thresholded matrix.





**Remerciements.** Cette thèse n'aurait pu aboutir sans la présence de plusieurs personnes, que je tiens à remercier ici.

Merci à Yves Meyer, qui a demandé plusieurs années à l'avance à faire partie de mon jury de thèse (!). C'est un geste qui m'a fait grand honneur, autant par son rôle fondateur dans le domaine des mathématiciens ondelettistes, que par sa gentillesse quand il a bien voulu, à plusieurs reprises, me recevoir dans son bureau, m'écouter, et me donner des directions de recherche. Je le remercie également pour avoir accepté le rôle de Président du Jury.

Je dois aussi remercier Martin Costabel, qui a accepté un rôle de rapporteur pour ma thèse, et qui a joué ce rôle avec un brio extraordinaire, aussi bien par la minutie de sa vérification allant jusqu'à refaire chaque calcul, par sa capacité à comprendre des implications de mes démonstrations que moi-même ignorais, et pour m'avoir consacré une journée toute entière pour me faire part de ses impressions de lecteur. Je le remercie aussi pour avoir su me révéler que certaines de mes propriétés méritaient d'être renommées en théorèmes, ce qui m'a donné une dose supplémentaire de courage face aux mathématiques.

Merci également à Jacques Laminie, qui a accepté lui aussi un rôle de rapporteur pour ma thèse, avec tout le travail de lecture (deux cent quarante-cinq pages!) et de bibliographie que cela implique. Je le remercie également car il a su me fournir une aide inestimable lors de tête à tête qui, tout au long de cette thèse, se sont déroulés à Orsay, en m'écoutant exposer longuement mes problématiques mathématiques avec leurs cadres, puis les structures des programmes informatiques (un programme de vingt mille lignes C++ pour chaque partie ...) utilisés pour mieux comprendre cette problématique, pour ensuite me donner un avis technique précieux clef en mains et me donner un laissez-passer pour utiliser un ordinateur monstrueusement doté en processeurs et en mémoire.

Je suis très sensible à la présence dans ce jury de Djalil Kateb, que j'ai eu le plaisir de rencontrer à Compiègne lors d'un exposé sur lequel se base partiellement mon exposé de thèse. Il a, chose très importante, pu m'indiquer des éléments de bibliographie que je ne connaissais pas, et il m'a initié au milieu des problèmes inverses, et m'a présenté à Mr Ha-Duong.

Je dois également remercier mon directeur de thèse, Pierre-Gilles Lemarié-Rieusset, qui, à la lecture d'une propriété d'un article qui devait devenir 5.2, a su détecter qu'il y avait de quoi creuser dans ce coin des mathématiques, et en effet il y avait beaucoup de choses à faire, l'analyse numérique de problèmes hautes fréquences étant encore dans ses débuts, avec certains des résultats importants connus seulement au stade unidimensionnel. J'ai ainsi eu le plaisir d'obtenir un certain nombre de résultats valorisants. Monsieur

Lemarié a aussi régulièrement lu des pages et des pages de démonstrations préliminaires de ces résultats, que je ne comprenais pas encore, et qui donc étaient plutôt difficiles à lire, comme toujours quand on écrit quelque chose qu'on ne comprend pas ; il a par exemple lu une démonstration de soixante-dix pages donnant l'équivalent de la propriété 4.3 en présence d'un coin, qui était une démonstration si indigeste qu'elle n'est pas passé dans ce manuscrit. Je dois également le remercier pour son rôle de directeur de thèse : il a su me guider avec psychologie, arrivant à comprendre comment j'étais fait, ce qui me faisait travailler, comment je pouvais réagir à telle ou telle situation ; en un mot, il a su me faire produire une thèse. Également, il est une personne très modeste, qui n'a jamais voulu qu'une démonstration dont il est l'inspirateur porte une référence à lui. Et pourtant, les démonstrations les plus claires de ce manuscrit ont le plus souvent été construites par lui-même ! Je resterais très longtemps endetté vis-à-vis de lui, non seulement pour ces démonstrations qu'il ne voulait pas revendiquer, mais bien plus important encore pour avoir structuré ma formation et plusieurs années de travail, ce en quoi il a pleinement réussi.

Ce travail de thèse a été réalisé au Département de mathématiques de l'université d'Évry, et a pu être mené à bien grâce au support financier de l'éducation nationale, et grâce au support financier, matériel, informatique et humain du département, qui est dirigé par Monique Jeanblanc.

Je remercie Valérie Picot pour les nombreux services qu'elle m'a rendus (doigts de fée), j'apprécie sa contribution majeure et éclatante à la vie du laboratoire.

J'adresse mes remerciements, pour la sympathie qu'ils m'ont témoignés, à Abdelmejid Bayad, à Fabrice Vandebrouck, à Francis Hirsch pour avoir créé ce laboratoire et donc rendu possible tout cela, à Gilles Hargé, à Gilles Lacombe, à Jamel Chikhi pour sa générosité, à Julia Matos, à Lucilla Corrias, à Pedro Ferreira pour ses précieux conseils de programmation et pour toutes les fois qu'il m'a présenté dans l'univers scientifique, à Stéphane Brouillaud et à Monique Jeanblanc pour m'avoir encadré ma pédagogie, à Stéphane Crepey, à Sylvie Mas-Gallic, à Thomas Simon pour son japonais et Victorita Dolean pour m'avoir plusieurs fois remplacé, et tous les autres membres du département ; j'ai une pensée émue pour toutes ces personnes pour avoir partagé une salle, un repas ou les ressources de leurs ordinateurs pendant ces années.

Je passe ensuite une dédicace spéciale à tous les jeunes gens que j'ai eu le plaisir de côtoyer durant ces quelques années à Évry, parmi lesquels Ali Zhioua, Binh Dao, Christophette Blanchet, Giulia Furioli, Ivan Marin et Rui Dos Santos.

Pour leurs encouragements et leur assistance aussi bien matérielle que morale qui m'ont permis de faire cette thèse dans de bonnes conditions, je remercie les ingénieurs systèmes Maouloud Ould Baba et Jean-Marc Dauget qui m'ont souvent fait des gestes spéciaux pour mes programmes.

Je remercie Albert Cohen pour m'avoir introduit au domaine et au sujet de la première partie, et pour m'avoir donné le virus des ondelettes avec ses cours de DEA et ses exposés, très, très minutieusement préparés et si pédagogiques.

Je remercie également Pascal Joly pour m'avoir présenté les méthodes matricielles multi-processeurs qui m'a fourni une intuition du calcul parallèle qui m'a été très utile dans cette thèse.

Il me faut aussi remercier Eric Darve qui m'a, lors d'un déménagement, permis de photocopier certaines références bibliographique de ma thèse.

Je remercie aussi Toufic Abboud qui m'a reçu plusieurs fois et aidé à mieux connaître la facette réalité dans la communauté plus appliquée, c'est à dire les libertés que la pratique valide.

Je remercie également Habib Ammari qui m'a, lui, aidé à mieux connaître la facette de la communauté moins appliquée, lors de nos rencontres.

Merci à Tahar Boulmezaoud, voisin occasionnel de bibliothèque qui a su me montrer des livres qui m'ont été très utiles. Merci aussi au personnel des différentes bibliothèques que j'ai été amené à visiter en cherchant ma bibliographie.

Je remercie aussi les personnes de la reprographie de l'université d'Évry qui se sont occupés de créer la version papier de cette thèse.

Je remercie de plus tous les auteurs des programmes du domaine public que j'ai utilisés intensément durant cette thèse, à savoir tous les contributeurs à linux, gcc, latex, vim et fwm. Sans eux, mes conditions de travail auraient sans doute été très différentes et beaucoup moins agréables. Je remercie également les concepteurs des librairies Arpack, Umfpack et Conmin.

Je remercie aussi ma famille, et mes amis qui se reconnaîtront ici, qui m'ont tant de fois encouragé à continuer la thèse, et à gérer les divers accros technique se révélant aux mauvais moment comme par exemple lors de la rédaction.

Je remercie de tout mon cœur, Ioana, mon épouse, qui m'a soutenu, tiré, poussé, et sans elle ce manuscrit de thèse n'existerait pas. Elle a été pour moi un la personne qui organisait mes idées, mes envies de travailler et mes envies de programmer. Je dois aussi la remercier pour avoir enduré toutes mes histoires sur ma vie de jeune chercheur perdu au milieu des objets mathématiques. Enfin, je la remercie pour son amour.





Il est possible, aussi bien dans la solution d'un problème que dans la matrice d'un système, de supprimer beaucoup d'informations en contrôlant la petitesse de l'erreur ainsi produite. Une fois que ce contrôle est théoriquement fondé, il devient important de prédire à l'avance quelles informations seront à supprimer, ce qui dans l'idéal permet d'économiser le calcul de ces informations.

Nous étudierons d'abord, dans la partie I (rédigée en anglais), cette prédiction dans un problème unidimensionnel et non-linéaire, cas dans lequel nous obtiendrons, pour la première fois, l'économie idéale, c'est à dire asymptotiquement la meilleure à une constante près.

Nous étudions également, dans la partie II (rédigée en français), cette prédiction dans un problème oscillatoire bidimensionnel dans lequel les bases d'ondelettes usuelles sont de peu d'utilité pour séparer l'information importante de l'information supprimable, lorsque la longueur d'onde est petite ; les bases de paquets d'ondelettes permettront d'obtenir cette supprimabilité.

Cette thèse est téléchargeable et consultable sur internet, avec les transparents de l'exposé de thèse, sous différents formats, à l'adresse suivante :

[http://www.maths.univ-evry.fr/pages\\_perso/goujot/these.html](http://www.maths.univ-evry.fr/pages_perso/goujot/these.html)

L'adresse devrait changer de loin en loin, toutefois vous le retrouverez avec un moteur de recherche utilisant Daniel Goujot Helmholtz Thèse.



# Table des matières

<b>I</b>	<b>Un problème non-linéaire et unidimensionnel.</b>	<b>6</b>
A	Introduction . . . . .	7
B	Wavelets and truncations (notation: $\lesssim \gtrsim \sim \Omega$ ) . . . . .	8
B-a	On the line . . . . .	9
B-b	On the segment (notations: $\tilde{\lambda}$ ) . . . . .	10
B-c	On other domains (notations: $d \langle   \rangle \Phi_j \Psi_j \Lambda j_\lambda V_j W_j$ $H \ \cdot\ _H$ ) . . . . .	11
B-d	Level-based truncation (notation: $j_\lambda$ ) . . . . .	12
B-e	Adaptive $H_0^1$ threshold-based N-term truncation (nota- tion: $\Psi$ ) . . . . .	13
B-f	Besov spaces (notation: $B_{p,p}^s$ ) . . . . .	13
B-g	Restricted truncation . . . . .	16
B-h	Interpolated cardinality (notations: $\Phi \#_{s,p}$ ) . . . . .	18
C	Algorithm (notations: $F u_\infty$ ) . . . . .	21
C-a	Assumptions . . . . .	21
C-b	Newton's algorithm . . . . .	22
C-c	Preconditioned gradient algorithm (notations: $P \rho \omega$ ) . . . . .	24
C-d	Adaptive decomposition and reconstruction (notations: $\succ \succcurlyeq$ ) . . . . .	25
C-e	Description of the Algorithm . . . . .	28
D	1D numerical example . . . . .	30
D-a	Definition . . . . .	30
D-b	Establishing assumption I.17 . . . . .	32
D-c	Establishing assumption I.18 . . . . .	33
D-d	Mother wavelet choice . . . . .	33
E	Extended a priori analysis of the algorithm . . . . .	34
E-a	Sufficient refinement for precision and truncation . . . . .	34
E-b	Convergence and cost: 1D and 2D . . . . .	38
E-c	Further dimensions . . . . .	41
E-d	Acknowledgments . . . . .	41
	Appendix . . . . .	41
<b>II</b>	<b>Un problème intégral de Helmholtz</b>	<b>47</b>
<b>1</b>	<b>Introduction à l'équation de Helmholtz</b>	<b>47</b>
<b>2</b>	<b>L'équivalence entre le problème de Helmholtz et le problème intégral de Helmholtz</b>	<b>53</b>
2.1	Le problème extérieur de Helmholtz . . . . .	53

2.1.1	L'opérateur de Helmholtz . . . . .	53
2.1.2	La condition de rayonnement de Sommerfeld : condition aux limites à l'infini . . . . .	59
2.1.3	Un domaine extérieur non borné . . . . .	62
2.1.4	Condition aux limites au bord du diffractant . . . . .	66
2.1.5	Unicité dans $H_{loc}^1({}^cS)$ , pour Dirichlet (et Neumann) . . . . .	71
2.2	Le problème interne au diffractant, et la nouvelle inconnue $J$ . . . . .	82
2.3	Le problème intégral de Helmholtz : la CFIE . . . . .	92
2.3.1	L'équation intégrale combinée . . . . .	92
2.3.2	Singularités de la CFIE dues aux coins . . . . .	96
2.3.3	Inégalité de Gårding pour la CFIE avec coins . . . . .	105
2.3.4	Existence et unicité dans la CFIE. . . . .	115
<b>3</b>	<b>La discrétisation de la CFIE.</b>	<b>117</b>
3.1	Besoin de l'analyse harmonique de la CFIE. . . . .	117
3.2	Discrétisations vérifiant une inégalité directe. . . . .	119
3.2.1	Les fonctions constantes par morceau. . . . .	120
3.2.2	Les paquets d'ondelettes . . . . .	122
3.2.3	Les familles de Malvar . . . . .	131
3.3	Ce qui a été fait sur cette équation . . . . .	132
3.3.1	La discrétisation . . . . .	133
3.3.2	Les accélérations avec multipôles . . . . .	133
3.3.3	Les accélérations avec la base de Malvar . . . . .	134
3.3.4	Les accélérations avec les ondelettes . . . . .	137
3.3.5	Prédiction et calculs dans la matrice de rigidité . . . . .	140
3.3.6	Préconditionnement : mélanges de ces méthodes . . . . .	141
3.3.7	Compatibilité de tout cela avec les coins : . . . . .	142
<b>4</b>	<b>L'intuition des numériques sur <math>\frac{\lambda_0}{h}</math>.</b>	<b>143</b>
4.1	Inversibilité de l'opérateur. . . . .	144
4.2	L'inégalité directe est inefficace . . . . .	146
4.3	L'opérateur focalise les hautes fréquences en l'absence de coins. . . . .	154
4.4	Les ondelettes discrétisent correctement. . . . .	174
4.5	Malvar . . . . .	184
<b>5</b>	<b>Étude de matrices issues des systèmes</b>	<b>185</b>
5.1	La justification de la discrétisation . . . . .	185
5.1.1	Dans la base canonique des fonctions d'échelle. . . . .	187
5.1.2	Les problèmes de résonances . . . . .	189
5.1.3	Hypothèses du théorème fondant l'application numérique. . . . .	193

5.2	Choix de la base de paquets d'ondelettes . . . . .	196
5.2.1	Bases d'ondelettes, Paquets d'ondelettes . . . . .	196
5.2.2	Stratégies possibles . . . . .	200
5.3	Position des coefficients importants de cette matrice . . . . .	208
5.3.1	Chiffres de compression . . . . .	208
5.3.2	Éviter le calcul des coefficients seuillés . . . . .	208
5.3.3	la localisation du Pic en Fourier des paquets d'ondelettes	223
5.3.4	Une approche numériquement économique et mathématiquement justifiée. . . . .	231
<b>6</b>	<b>Conclusion</b>	<b>232</b>
6.1	Annexe . . . . .	232
6.2	Bibliographie . . . . .	234

# I Un problème non-linéaire et unidimensionnel.

Blind Wavelet Compression of the Solution of a Nonlinear PDE with Singular Forcing Term within Optimal Order Cost: Stability of Restricted Approximation to Small Errors

Nous présentons, à part, un article détaillant une méthode de résolution en base adaptée d'ondelettes pour une équation différentielle elliptique non linéaire. C'est l'aspect prédiction de ce qui résulte du seuillage qui est commun avec la suite, plus homogène, de cette thèse. Dans [12], W. Hackbusch a utilisé l'exemple-test bidimensionnel  $\Delta u + e^{cu} = 0$  avec  $c > 0$  pour comparer les vitesses de convergence de trois méthodes multigrilles couplées à l'algorithme de Newton. Nous prenons le même exemple unidimensionnel sur  $[0,1]$  pour illustrer notre méthode de résolution adaptative, bien que cette méthode soit valable en dimension  $d \leq 2$ . Nous utilisons l'algorithme de Newton, une méthode de Galerkin, une reconstruction/décomposition adaptative introduite dans [10] par Wolfgang Dahmen, et une hybridation de résultats linéaires et non-linéaires de la théorie de l'approximation.

Pour mettre en valeur les capacités adaptatives de notre méthode, nous avons ajouté un second membre singulier dans l'exemple-illustration, élément de  $H^{-1} \cap B_{p,p}^{s-2}$ , avec  $p = \frac{21}{41}$  et  $s = 2 + \frac{58}{105}$  un peu plus grand que  $1 + \frac{1}{p} - \frac{1}{2}$ : ce second membre est le Laplacien de  $x \mapsto a|3x-1|^{2/3}$  avec  $a$  réel. L'algorithme décrit a été utilisé tel quel en programmation, et l'évaluation de sa convergence et l'optimalité de sa complexité (en  $\varepsilon^{-\frac{d}{s-1}} |\log \varepsilon|^a$  avec  $a > 0$  où  $\varepsilon$  est l'erreur en norme  $H^1$ ) est intégralement justifiée en dimension  $d \leq 2$ .

## Résumé

Level lifting of the wavelet expansion is related to an interpolation result for Sobolev  $H^s$  spaces; however, the nonlinear  $N$ -term approximation is linked with a non-linear interpolation result for Sobolev spaces  $W^{s,p}$  non-compactly included into  $L^2$ . [5] introduced the notion of restricted approximation, related to a restricted interpolation result for Lebesgue and Besov spaces, which has the advantage of controlling the number of nodes of the wavelet tree. Based on this, we construct an optimal order resolution algorithm extending beyond the linear elliptic case of [3], as we illustrate numerically. Recall that optimal order adaptivity implies blind compression of the unknown of the PDE. We illustrate on a univariate version of the bivariate PDE  $\Delta u + e^{cu} = 0$ ,  $c > 0$ , used to benchmark three nonadaptive multilevel methods [12].

The adaptiveness of our algorithm is highlighted by the addition in this illustration of a singular forcing term. This term is an element

of  $H^{-1}$  but does not belong to  $H^{-\frac{5}{6}}$ : more precisely, it is the second derivative of  $t \mapsto |3t-1|^{2/3}$ . This illustration passed the numerical implementation test ( $C\varepsilon^{-0.505}$  flops). The algorithm's convergence and cost ( $\varepsilon^{-\frac{d}{s-1}}$  where  $\varepsilon$  is the final error in  $H^1$  norm) in both univariate ( $d = 1$ ) and bivariate ( $d = 2$ ) general cases is shown to have optimal order, with  $s$  less than three.

## A Introduction

Optimal order cost is a difficult issue for an adaptive resolution of the steady Laplace equation on a unidimensionnal interval, and on a bidimensionnal square, with a nonlinear zero-order perturbation term. The wavelet basis [2] is a natural response to local lacks of smoothness, without nonlinear perturbation terms [3], and a wavelet framework for nonlinear elliptic equations was outlined in [10].

A sensitive issue is to localize (in the frequency-time and in the space-time) the biggest wavelet coefficients of the nonlinear part.

The wavelet framework gives a natural way to get the biggest wavelet coefficients of a function  $f$  once known that  $f$  has  $s$  derivatives in the Lebesgue space  $L^p(\mathbb{R}^d)$ ; this notion is extendable to the case  $s \in \mathbb{R}_+$  and  $p \in \mathbb{R}_+$ , with the Sobolev and Besov families of function spaces. When  $s = \frac{d}{p} - \frac{d}{2}$ , this natural way is to compare the modulus of the wavelet coefficient with  $\varepsilon^{-p/(2-p)}$ , where  $\varepsilon$  is the acceptable precision loss in  $L^2$  norm [2].

This threshold-based truncation method is not recommended in the presence of nonlinear terms: while it allows to control the number of active wavelets coefficients, i.e., the number of leaves in the wavelet tree, we need to control the number of nodes to maintain optimal order cost in use of collocations for the nonlinear contribution [7, 10].

A first choice to enforce this control has been that the  $s$  derivatives are in  $L^{p'}$  where  $p'$  is slightly greater than  $p$  [10]; however, the function resulting from the above threshold does not have as many derivatives in  $L^{p'}$  as  $f$ . Although this difference is very tiny, an algorithm using many times this threshold process cannot have optimal order cost given that the unknown has  $s$  derivatives in  $L^{p'}$ .

The raw enforcement of the needed estimates on the number of nodes [1] in the framework outlined in [10] was made in [4], which however does not address the drift provoked by unit cost collocations.

Collocation produces coefficients not of a wavelet, but of a scale function. To get the wavelet coefficients, we must apply repeatedly the wavelet decomposition filter on the result of the collocations.



This wavelet decomposition filter gradually propagates the drift provoked by the truncation through the scales. This propagation does explode when used with less Sobolev Regularity than  $H^\nu$ , where  $\nu$  is a small positive real number depending on the multiresolution, see property I.33. Hence, the error of collocation at scale  $2^{-j}$  must be smaller than  $\epsilon 2^{-\nu j}$  where  $\epsilon$  is the finally accepted precision on coarsest scale. To see this, consider the computation of a residual with no more than  $H^{-1}$  linear smoothness through iterated wavelet decomposition, with coefficients computed by collocations: the latter will either be too expensive, or the computation will numerically blow.

We present a new method to control this drift propagation. We replace the number of leaves of the wavelet tree by the restricted cardinality [5], and we use a threshold method without the smoothness gap described above. Our method compares  $\|\langle \tilde{\psi}_{jk} | f \rangle \psi_{jk}\|_{W^{1,p''}}$  with the threshold

$$C \epsilon^{\frac{2}{2-p'}} (\|f\|_{W^{s,p'}})^{-\frac{d}{2-p'}}$$

where  $\epsilon$  is the target accuracy in  $H^1$ -norm, and where  $p'' = \frac{2-p'}{p'} \frac{d}{s-1}$ , and is naturally related to the family of Besov spaces.

Section 1 presents the restricted truncation and the restricted cardinality. In section 2, we define some assumptions, then we construct gradually the algorithm, starting with a formal Newton algorithm and ending with the optimal order algorithm. In section 3, we present the univariate illustration and the 1D numerical results, and prove that it satisfies both assumptions. In section 4, we prove *a priori* the convergence with optimal order cost in the univariate and bivariate cases.

**Notations** *Unless otherwise specified, domain of function spaces is  $\Omega = (0,1)^d$  with  $d = 1$  or  $d = 2$  (except for the first section). All non-italic symbols have a meaning reached by consensus. When two expressions  $a$  and  $b$  depend on a parameter  $c$ , we will note by  $a \stackrel{c}{\lesssim} b$  the assertion  $\exists M, \forall c, a \leq Mb$ , and by  $a \stackrel{c}{\sim} b$  the assertion  $\exists m > 0, \exists M, \forall c, mb \leq a \leq Mb$ .*

## B Wavelets and truncations

Throughout this section, the domain of function spaces is a bounded open subset of  $\mathbb{R}^d$  with  $d > 0$ . In wavelet theory, the two parameters  $h$  and  $\tilde{h}$  stand for supported sequences indexed by  $\mathbb{Z}$ . For some well chosen values of

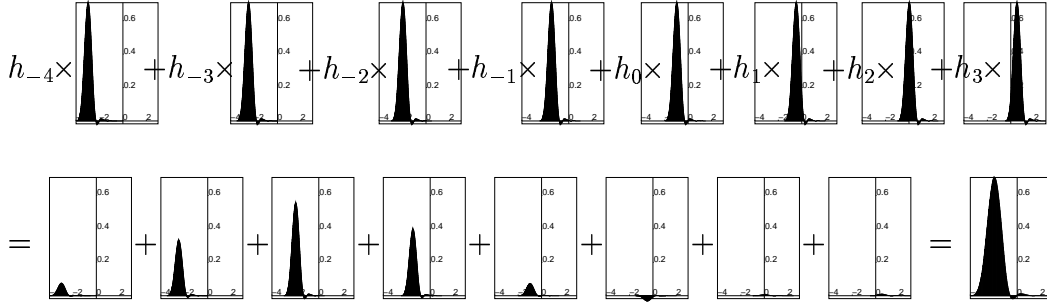
these parameters (satisfying the duality identity [2] and the Cohen criterion Th 17.2 from therein), there exist unique finite-supported functions so-called *primal* and *dual mother scale functions*, denoted respectively  $\phi$  and  $\tilde{\phi}$ , such that  $\int \phi = 1$ ,  $\int \tilde{\phi} = 1$ , and  $\int \phi(x)\tilde{\phi}(x-m)dx = \delta_m^0$  for all  $m \in \mathbb{Z}$ , and such that we have the following *refinement equations*:

$$\phi(x) = \sum_{m \in \mathbb{Z}} h_m \phi(2x-m) \quad \text{and} \quad \tilde{\phi}(x) = \sum_{m \in \mathbb{Z}} \tilde{h}_m \tilde{\phi}(2x-m) \quad (1)$$

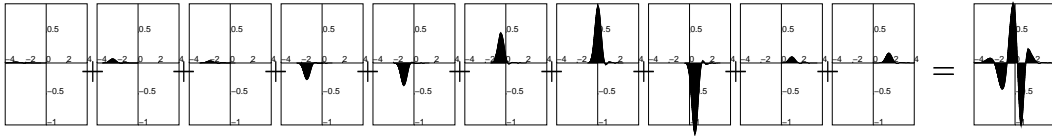
$h$  and  $\tilde{h}$  are called the *scaling filters*. The *primal* and *dual mother wavelets*  $\psi$  and  $\tilde{\psi}$  are defined as follows:

$$\psi(x) := \sum_{m \in \mathbb{Z}} (-1)^m \tilde{h}_{1-m} \phi(2x-m) \quad \text{and} \quad \tilde{\psi}(x) := \sum_{m \in \mathbb{Z}} (-1)^m h_{1-m} \tilde{\phi}(2x-m). \quad (2)$$

We will need further regularity assumptions; the first assumption is that  $\phi$  and  $\tilde{\phi}$  are both  $L^2(\mathbb{R})$ . We illustrate the left part of (1):



Next, we illustrate the definition of  $\psi$  in (2):



## B-a On the line

According to (1), the primal mother function  $\phi$  is an element of the space  $V_1^{\mathbb{R}}$  which we define as the closed subspace of  $L^2(\mathbb{R})$  spanned by the functions  $x \mapsto \phi(2x-m)$  indexed by  $m \in \mathbb{Z}$ . This space is translation invariant, more precisely,  $f \in V_1^{\mathbb{R}} \iff x \mapsto f(x+\frac{1}{2}) \in V_1^{\mathbb{R}}$ . Let us define likewise, for each  $j \in \mathbb{Z}$ , the *analysis space*  $V_j^{\mathbb{R}}$  as the closed subspace of  $L^2(\mathbb{R})$  spanned by the functions  $x \mapsto \phi(2^j x-m)$  indexed by  $m \in \mathbb{Z}$ . In particular,  $V_0^{\mathbb{R}} \subset$

$V_1^{\mathbb{R}}$  ( $\phi(\cdot - m)$  belongs to  $V_1^{\mathbb{R}}$  because  $\phi$  belongs to  $V_1^{\mathbb{R}}$ ). Therefore we have the following increasing sequence, which is an instance of **multiresolution analysis**:

$$\dots \subset V_{-1}^{\mathbb{R}} \subset V_0^{\mathbb{R}} \subset V_1^{\mathbb{R}} \subset V_2^{\mathbb{R}} \subset \dots \subset L^2(\mathbb{R})$$

Its intersection is  $\{0\}$ , and its union is dense in  $L^2(\mathbb{R})$ . For each  $j \in \mathbb{Z}$ , let us define the *detail space*  $W_j^{\mathbb{R}}$  as the closed subspace of  $L^2(\mathbb{R})$  generated by the functions  $x \mapsto \psi(2^j x - m)$  indexed by  $m \in \mathbb{Z}$ . This detail space  $W_j^{\mathbb{R}}$  is a linear complement of  $V_j^{\mathbb{R}}$  in  $V_{j+1}^{\mathbb{R}}$ . For each  $j \in \mathbb{Z}$ , the following so-called **wavelet basis** set is a Riesz basis of  $L^2(\mathbb{R})$ :

$$\{x \mapsto 2^{\frac{j}{2}} \phi(2^j x - m) : m \in \mathbb{Z}\} \cup \{x \mapsto 2^{\frac{l}{2}} \psi(2^l x - m) : l, m \in \mathbb{Z} \text{ and } l \geq j\} \quad (3a)$$

A **good point for wavelets** is that the corresponding dual Riesz basis is

$$\{x \mapsto 2^{\frac{j}{2}} \tilde{\phi}(2^j x - m) : m \in \mathbb{Z}\} \cup \{x \mapsto 2^{\frac{l}{2}} \tilde{\psi}(2^l x - m) : l, m \in \mathbb{Z} \text{ and } l \geq j\} \quad (3b)$$

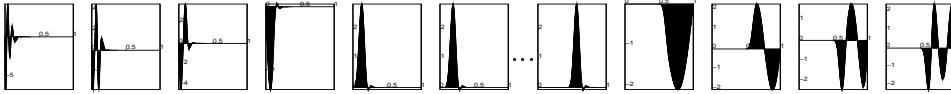
## B-b On the segment

Let us temporarily denote by  $S$  the set of the non-null crude restrictions of the elements of (3a) to  $[0,1]$ ;  $S$  doesn't give a Riesz basis of  $L^2([0,1])$ . We will hence use  $S$  with a set of modifications on its elements overlapping  $\{0\} \cup \{1\}$ . We will not detail these modifications introduced in [13], and used in [8]. Let us denote by respectively  $\Phi_j^{(0,1)}$  and  $\Psi_j^{(0,1)}$  the obtained sets of modified scale functions and modified wavelets of level  $j$ , where  $j \geq j_0$  (a minimal level  $j_0$  appears in this construction).  $\Phi_j^{(0,1)}$  spans the space

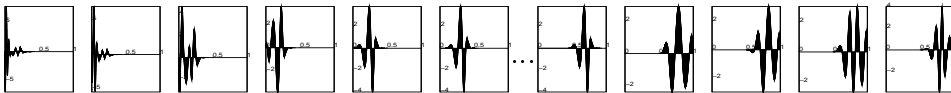
$$V_j^{(0,1)} := \{f_{|[0,1]} : f \in V_j^{\mathbb{R}}, f(0)=f(1)=0, f_{|_{2^{-j}[-1,a]}} \in \mathbb{R}[X], f_{|_{1+2^{-j}[-b,1]}} \in \mathbb{R}[X]\}$$

with  $a, b \in \mathbb{N}$ . In the following, we illustrate for  $j_0 = 4$ :

— scale function set  $\Phi_4^{(0,1)}$  with 16 elements and  $a = 0$  and  $b = 2$ :

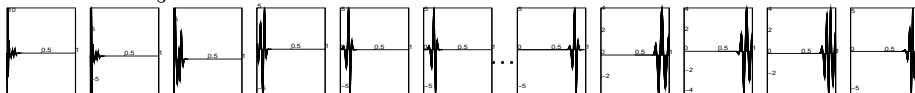


— wavelet set  $\Psi_4^{(0,1)}$  with 16 elements :



(4)

— wavelet set  $\Psi_5^{(0,1)}$  with 32 elements :



**Property I.1** *The set  $\Lambda^{(0,1)} := \Phi_{j_0}^{(0,1)} \cup \bigcup_{j \geq j_0} \Psi_j^{(0,1)}$  is a Riesz basis of  $L^2([0,1])$ , and is called the wavelet basis. This is theorem 33.4 of [2].*

**Notations I.2** *Thus  $\Lambda^{(0,1)}$  has a dual Riesz basis. When  $\lambda$  is an element of  $\Lambda^{(0,1)}$ , we write  $\tilde{\lambda}$  for its corresponding element in the dual basis. Replacing  $\Lambda^{(0,1)}$  with the Riesz basis  $\Phi_k^{(0,1)} \cup \bigcup_{j \geq k} \Psi_j^{(0,1)}$  allows us to define likewise  $\tilde{\lambda}$  for  $\lambda \in \Phi_k^{(0,1)}$  when  $k > j_0$ . Let  $\langle f|g \rangle$  be the scalar product  $\int_0^1 fg$ : we have  $\langle \lambda|\tilde{\mu} \rangle = \begin{cases} 1 & \text{if } \lambda = \mu \\ 0 & \text{otherwise} \end{cases}$  where  $\lambda, \mu \in \Phi_k^{(0,1)} \cup \bigcup_{j \geq k} \Psi_j^{(0,1)}$  and  $k \geq j_0$ .*

### B-c On other domains

The domain  $(0,1)^2$  or other bounded domains of  $\mathbb{R}^2$  may also be used instead of  $(0,1)$ , if a wavelet multiresolution with the same notations ( $j_0, \Phi_j^\Omega, \Psi_j^\Omega, \tilde{\lambda}, \dots$ ) is available, and satisfies I.1 where  $\Omega$  is replacing  $(0,1)$ .

**Property I.3** *The definitions  $\Phi_j^{(0,1)^2} := \Phi_j^{(0,1)} \otimes \Phi_j^{(0,1)}$  and  $\Psi_j^{(0,1)^2} := \Phi_j^{(0,1)} \otimes \Psi_j^{(0,1)} \cup \Psi_j^{(0,1)} \otimes \Phi_j^{(0,1)} \cup \Psi_j^{(0,1)} \otimes \Psi_j^{(0,1)}$  give a suitable multiresolution for  $\Omega = (0,1)^2$  the unit square (see theorem 33.4 of [2])*

**Notations I.4** *Take  $\Omega$  to be the domain of the problem and  $d$  to be its dimension. Take  $\langle f|g \rangle$  to be the scalar product  $\int_\Omega fg$ . From now, we will abbreviate  $\Phi_j^\Omega$  to  $\Phi_j$ ,  $\Psi_j^\Omega$  to  $\Psi_j$ ,  $V_j^\Omega = \text{span} \Phi_j$  to  $V_j$ ,  $W_j^\Omega = \text{span} \Psi_j$  to  $W_j$ , and  $\Lambda^\Omega = \Phi_{j_0} \cup \Psi_{j_0} \cup \Psi_{j_0+1} \cup \dots$  to  $\Lambda$ . For each  $\lambda \in \Phi_k \cup \Psi_k$ , let  $j_\lambda := k$  be the level of  $\lambda$ .*

**Property I.5** *Given  $s > 0$  and  $s \neq \frac{1}{2}$ , the wavelet basis  $\Lambda$  is (up to a renormalisation) a Riesz basis of the closure  $E$  in  $H^s$  of the set of continuous functions cancelling on  $\partial\Omega$ , provided that  $\psi \in H^{\varepsilon+s}(\mathbb{R})$  (c.f. theorem 33.6 of [2]):*

$$\left\| 2^{sj_\lambda} \langle \tilde{\lambda}|f \rangle \right\|_{\ell^2(\lambda \in \Lambda)} \stackrel{f \in E}{\sim} \left\| \|\lambda \langle \tilde{\lambda}|f \rangle\|_{H^s} \right\|_{\ell^2(\lambda \in \Lambda)} \stackrel{f \in E}{\sim} \|f\|_{H^s}$$

**Definition I.6** Let us assume that  $\psi \in H^{1+\varepsilon}(\mathbb{R})$ . For the space  $H := H_0^1 = H_0^1(\Omega)$ , choose the norm

$$\|f\|_H^2 := \left\| \sum_{\lambda \in \Phi_{j_0}} \langle f | \tilde{\lambda} \rangle \nabla \lambda \right\|_{L^2}^2 + \sum_{j=j_0}^{\infty} \sum_{\lambda \in \Psi_j} |\langle f | \tilde{\lambda} \rangle|^2 2^{2j\lambda}$$

To see that  $\|\cdot\|_H$  is a norm of  $H$ : consider that  $\|\|\nabla \lambda\|_{L^2} \langle f | \tilde{\lambda} \rangle\|_{\ell^2(\lambda \in \Phi_{j_0}^{(0,1)})} \stackrel{f \in V_{j_0}}{\sim} \|\nabla f\|_2$  ( $V_{j_0}$  of finite dimension), and apply property I.5 using  $s=1$ .

### B-d Level-based truncation

The coarsest levels of a wavelet basis are known to retain the smoothest part of a function. Thus, if a function  $f \in H$  (where  $H := H_0^1$ ) with  $H^s$  Sobolev regularity ( $s > 1$ ) has been interfered by some error  $\hat{f} - f$  which is known to be small in  $H$ -norm, then the first  $k$  coarsest levels of  $\hat{f}$  contain sufficient information to construct the following function  $P_k^{\text{lin}} \hat{f}$  which is both close to  $f$  and has fewer wavelet coefficients:

$$P_k^{\text{lin}} \hat{f} := \sum_{\lambda \in \Phi_{j_0}} \langle \tilde{\lambda} | \hat{f} \rangle \lambda + \sum_{j=j_0}^k \sum_{\lambda \in \Psi_j} \langle \tilde{\lambda} | \hat{f} \rangle \lambda$$

A Jackson equality [2] states about  $P_k^{\text{lin}}$  that  $f \in H \cap H^s$  if and only if

$$\left\| 2^{k(s-1)} \|f - P_k^{\text{lin}} f\|_H \right\|_{\ell^2(k \geq j_0)} < \infty. \quad (5)$$

**Theorem I.7** Let  $\kappa > 1$  be a freely customisable price to pay on precision. Let  $s > 1$  and assume that  $\forall g \in H^s(\Omega) \cap H \quad \|g\|_{H^s} \sim \|2^{sj\lambda} \langle \tilde{\lambda} | g \rangle\|_{\ell^2(\lambda \in A)}$ . For all  $\hat{f} \in H$  and  $f \in H^s(\Omega) \cap H$ , the least  $k$  such that  $\|f - P_k^{\text{lin}} \hat{f}\|_H \leq \kappa \|\hat{f} - f\|_H$  satisfies  $2^{k(s-1)} \leq 2^{j_0(s-1)} + C \|\hat{f} - f\|_H^{-1} \|f\|_{H^s}$  with  $C > 0$  independent of  $f, \hat{f}$ , see remark 34.9 of [2].

■ Let  $a := \|\hat{f} - f\|_H$  and  $A := \|f\|_{H^s}$ . Combining  $\|P_k^{\text{lin}}(\hat{f} - f)\|_H \leq a$  (continuity) and  $\|f - P_k^{\text{lin}} f\|_H \leq C' 2^{-(s-1)k} A$  (Jackson (5)), we have  $\|f - P_k^{\text{lin}} \hat{f}\|_H \leq \|f - P_k^{\text{lin}} f\|_H + \|P_k^{\text{lin}}(\hat{f} - f)\|_H \leq C' 2^{-(s-1)k} A + a$ : hence the least  $k \geq j_0$  such that  $\|f - P_k^{\text{lin}} \hat{f}\|_H \leq \kappa a$  satisfies either  $2^{-(s-1)k} \geq \frac{\kappa-1}{C'} \frac{a}{A}$  or  $k = j_0$ . ■

## B-e Adaptive threshold-based N-term truncation

In the level-based truncation above, we have selected each wavelet according to its level; all used levels (i.e. used sets  $\Psi_j$ ) are full.

Here, we will threshold each wavelet independently of its level, looking at its  $H_0^1$ -contribution in the wavelet expansion of  $f$ : instead of  $P^{\text{lin}}$ , we now use  $P_n^{\text{lin}}$  defined by

$$P_n^{\text{lin}}g := \sum_{\lambda \in \Phi_{j_0}} \langle \tilde{\lambda} | g \rangle \lambda + \sum_{\lambda \in \Psi_n(g)} \langle \tilde{\lambda} | g \rangle \lambda.$$

The Jackson inequality in (5) now can be formulated with the norm of  $f$  in the Sobolev space  $W^{s,p}$  [2]:

$$n^{\frac{s-1}{d}} \|f - P_n^{\text{lin}}f\|_H \stackrel{n,f}{\lesssim} \|f\|_{W^{s,p}}. \quad (6)$$

We have  $p \in (1,2)$  and

$$s = 1 + \frac{d}{p} - \frac{d}{2}. \quad (7)$$

$W^{s,p}$  contains more cusp functions than former space  $f \in H^s$ : here is the good side of adaptivity, **working sparsely on detail sets**  $\Psi_j$ .

Similarly to theorem (I.7),  $P_k^{\text{lin}}\hat{f}$  is a function both close to  $f$  and with controlled number of coefficients:

**Theorem I.8 (N-Term approximation)** *Let  $p \in (1,2)$  and  $s := 1 + \frac{d}{p} - \frac{d}{2}$ . Assume that  $s$  is small enough so that our wavelet base satisfies (6) for  $f \in W^{s,p} \cap H$  (again,  $H := H_0^1$ ). There exists  $\kappa > 1$  such that for all smooth function  $f \in W^{s,p} \cap H$  and for all approximation  $\hat{f} \in H$ , the least  $n$  such that  $\|f - P_n^{\text{lin}}\hat{f}\|_H \leq \kappa \|f - \hat{f}\|_H$  satisfies  $n^{(s-1)/d} \leq C \|f - \hat{f}\|_H^{-1} \|f\|_{W^{s,p}}$  with  $C > 0$  independent of  $f$  and  $\hat{f}$ . This is Theorem 5.2 of [3] and is proven there and in Appendix.*

The projector  $P_n^{\text{lin}}$  is no longer linear:  $P_n^{\text{lin}}(f + g)$  can differ from  $P_n^{\text{lin}}f + P_n^{\text{lin}}g$ , hence the wavelet subtree depends on the function: this is *adaptivity*.

## B-f Besov spaces

As (5) is a characterisation of  $H^s$ ,  $\|P_k^{\text{lin}}\hat{f}\|_{H^s} \lesssim \|f\|_{H^s} + \|f - \hat{f}\|_H$  is clearly a consequence of theorem I.7. But (6) isn't linked with a characterisation of  $W^{s,p}$  (see (14) of [6]): hence the additional smoothness condition  $P_n^{\text{lin}}\hat{f} \in B$

cannot result from theorem I.8. Even if in the threshold-based context the wavelet compression is known to recover from small errors on a smooth signal with localized singularities, this does not happen precisely enough for  $W^{s,p}$ -smoothness for optimal-order-cost programming: optimal-order-cost means that the amount of computation is directly proportional to the number of parameters required in a result achieving the target accuracy. The conservation of  $B$ -smoothness by truncation (with either  $B = H^s \cap H_0^1$  or  $B = W^{s,p} \cap H_0^1$ ) is **mandatory** (see (19c) or (19h) or C-c or [3]): we definitely need the truncation to be related to a characterisation of the used smoothness space, e.g. (5).

With a slight change on (6), we obtain  $\|2^{k(s-1)}\|f - P_{2^k}^{\text{lin}} f\|_H$ , which is the semi-norm of a function space (see equation (2.16) in [11]). However, this space isn't as common as Lebesgue spaces, it is the Besov space  $B_{p,p}^s$ . The **Besov choice is a trade-off** between simplicity of smoothness space function, adaptivity, and simplicity of wavelet characterisation. The first benefit of this choice is the following:

**Property I.9** *Given  $s \in \mathbb{R}_+$  and  $p \in \mathbb{R}_+$  with  $s \geq \frac{d}{p}$ , the wavelet basis  $\Lambda$  is (up to a renormalisation) a Riesz basis of the closure  $E$  in  $B_{p,p}^s$  of the set of continuous functions cancelling on  $\partial\Omega$ , provided that  $\psi \in B_{p,p}^{s+s}(\mathbb{R})$  and  $\tilde{\psi} \in B_{p,p}^{s-s}(\mathbb{R})$  (see [2], Remark 33.6 for the case  $p < 1$ , Theorem 33.4 otherwise):*

$$\left\| 2^{(s-\frac{d}{p}+\frac{d}{2})j\lambda} |\langle f|\tilde{\lambda}\rangle| \right\|_{\ell^p(\lambda \in \Lambda)} \stackrel{f \in E}{\sim} \left\| \|\lambda \langle f|\tilde{\lambda}\rangle\|_{B_{p,p}^s} \right\|_{\ell^p(\lambda \in \Lambda)} \stackrel{f \in E}{\sim} \|f\|_{B_{p,p}^s}$$

Besov spaces come with a bonus: the limitation  $\frac{s-1}{d} < \frac{1}{2}$  linked with  $1 < p$  in the exponent of theorem I.8 will no longer be needed. Just be careful that the triangular inequality of  $B_{p,p}^s$  is relaxed when  $p < 1$ .

Besov spaces  $B_{p,q}^s$  have three indices,  $s \in \mathbb{R}$  and  $p, q \in (0, \infty]$ , and we will stick to the case  $q = p$ . Even with this restriction, the obtained family isn't totally ordered for the inclusion (whereas  $H^s$ 's family is totally ordered), but is rather a bidimensional lattice (with min and max). Let us relate  $B_{p,p}^s$  with the complex  $\frac{1}{p} + is$ : the space related with  $z \in \mathbb{C}_+$  is included in the space related with  $z' \in \mathbb{C}_+$  if and only if

$$\arg(z-z') \in \begin{cases} [\text{atan}(d), \pi) & \text{if } \Omega \subset \mathbb{R}^d \text{ is bounded;} \\ [\text{atan}(d), \frac{\pi}{2}) & \text{else.} \end{cases}$$

Strip off  $\text{atan}(d)$  to obtain compact inclusion. Hence we have a continuous lattice. The compact inclusions remain true if we change the spaces related

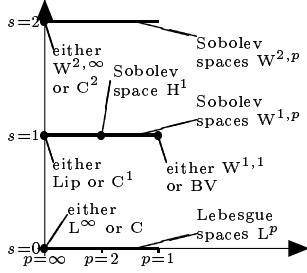


FIG. 1 – *Function spaces in a planar diagram.*

to some points, see Figure 1 (we legitimately associate many spaces to the same complex  $z$ , because the real interpolation of function spaces is somehow unaware of such alternatives). Thus, the Besov spaces are the continuation between well known Lebesgue, Sobolev, and continuity spaces. Here is the second benefit for Besov choice:

The theorem I.8 with  $W^{s,p}$  replaced by  $B_{p,p}^s$  has the additional result  $\|P_n^{\text{nl}} \hat{f}\|_{B_{p,p}^s} \lesssim \|f\|_{B_{p,p}^s} + \|f - \hat{f}\|_H$ . Therefore, due to the truncation,  $P_n^{\text{nl}} \hat{f}$  inherits  $B_{p,p}^s$  smoothness estimates from  $f$ . However:

**Remark I.10**  $P_n^{\text{nl}} \hat{f}$  does not inherit  $H^s$  smoothness even when  $f \in H^s$ .

■ Take  $f = \sum_{\lambda \in A} 2^{-(\frac{1}{2}+s)j\lambda} \frac{\lambda}{j_\lambda^2}$  and  $\hat{f}_{\varepsilon,J} = f + \varepsilon \alpha_J \sum_{\lambda \in \Psi_J} [1 + 2^J d_A(\lambda) \log(1 + 2^J d_A(\lambda))]^{-1}$  with  $d_A(\lambda)$  being the distance of  $A$  to the support of  $\lambda$ , and  $\alpha_J \stackrel{J}{\sim} 1$  being the positive real satisfying  $\|f - \hat{f}_{\varepsilon,J}\|_H = \varepsilon$ . We have  $f \in H^s \cap H \subset B_{p,p}^s$ , hence  $\|P_{n_{\varepsilon,J}}^{\text{nl}} \hat{f}_{\varepsilon,J}\|_{B_{p,p}^s} \lesssim \|f\|_{B_{p,p}^s} + \varepsilon$  with  $n_{\varepsilon,J}$  given by theorem I.8. However, for all  $C > 0$ , there exists  $\varepsilon > 0$  and  $K \in \mathbb{N}$  such that for all  $J > K$ , there is a term of scale  $J$  in the wavelet expansion  $P_{C\varepsilon^{-d/(s-1)}}^{\text{nl}} \hat{f}_{\varepsilon,J}$  making the  $H^s$ -norm of this expansion to be greater than  $\varepsilon \alpha_J 2^{Js}$ . As  $\varepsilon \alpha_J 2^{Js}$  cannot be bounded uniformly in  $J > K$ , the proof is complete (same explosion encountered when truncating the wavelet expansion of  $\hat{f}_{\varepsilon,J}$  by threshold  $C\varepsilon^{2/(2-p)}$ ). ■

Hence, the level-based truncation is specific to the conservation of  $H^s$ -type smoothness, the threshold-based truncation is specific to the conservation of non-linear smoothness, and we cannot replace one by another when using truncation to recover smoothness from small  $H_0^1$  errors.

Threshold-based truncation method has the drawback that, besides using partially full wavelet trees, the  $j_\lambda$  **isn't controlled** even if  $\|f\|_{W^{s,p}}$  stays



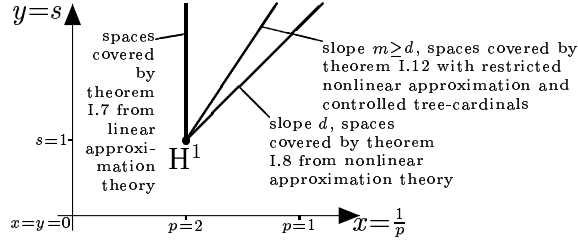


FIG. 2 – *Optimal Approximations in  $H^1$  of Besov Spaces  $B_{p,p}^s$ .*

bounded and  $\|e\|_H$  is small: take  $f = 2^{-j\lambda} \lambda$  (use of Besov space  $B_{p,p}^s$  does not help).

### B-g Restricted truncation

Numerically dealing with non-linear equations ends up in pointwise evaluation (for use in collocation). The cost of pointwise evaluation on an adapted grid depends on the number of nodes, and not on the number of leaves, in the wavelet binary tree. Hence the cardinality criterion is not enough to make optimal order resolution of nonlinear equations.

This is the cardinality of the wavelet tree, i.e., the subset of  $\Lambda$  containing  $\Psi_\eta(g) = \{\lambda \in \Psi : 2^{j\lambda} |\langle \tilde{\lambda} | g \rangle| \geq \eta\}$  and extended by cone conditions (see definition I.27). This cardinality is difficult to control.

In the context of PDE, more information on used functions is available. [10] added the realistic assumption  $f \in W^{s,p'}$  with  $\frac{1}{p'} < \frac{s-1}{d} + \frac{1}{2}$  instead of (7), and [6] added the assumption  $f \in H^{1+\varepsilon}$ . More recently, [4] dealt with N-sized tree approximation, an extensive analysis of specific corresponding smoothness being provided in [1]. In this paper, we will use the more handy  $B_{p',p'}^s$ -smoothness, which is naturally related to restricted truncation and restricted nonlinear approximation [5].

The truncation strategies (either level-based or threshold-based) are specifically related to the type of smoothness we want to recover (either linear  $H^s$  or nonlinear  $B_{p,p}^s$  smoothness). Similarly, there is a so-called restricted truncation strategy related to restricted nonlinear  $B_{p',p'}^s$  smoothness [5], which corresponds to the line with slope  $m \geq d$  in Figure 2.

The **original results of section 1** consists in showing that the truncation based on thresholding  $\|\langle f | \tilde{\lambda} \rangle \lambda\|_{W^{1,p''}}$  with some  $p'' < 2$  recovers restricted nonlinear smoothness from an  $H_0^1$ -small error, and that this truncation brings control on tree-cardinalities.

**Lemma I.11** *If  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a non-decreasing function, and if  $a > 0$  is such that  $I := \int_0^\infty ax^{-a-1}f(x)dx < \infty$ , then  $\forall x \in \mathbb{R} x^{-a}f(x) \leq I$ .*

$$\blacksquare I \geq \int_x^\infty ay^{-a-1}f(y)dy \geq \int_x^\infty ay^{-a-1}f(x) = x^{-a}f(x) \blacksquare$$

**Theorem I.12 (original result : restricted truncation)** *Let  $p \in (0,2)$  and  $s \geq 1 + \frac{d}{p} - \frac{d}{2}$  such that the Besov space  $B := B_{p,p}^s(\Omega) \cap H_0^1(\Omega)$  admits  $\|f\|_B := \|\langle \tilde{\lambda} | f \rangle 2^{(s-d/p+d/2)j\lambda}\|_{\ell^p(\lambda \in \Lambda)}$  as a pseudo-norm, and  $|f|_B := \|\langle \tilde{\lambda} | f \rangle 2^{(s-d/p+d/2)j\lambda}\|_{\ell^p(\lambda \in \Psi)}$  as a pseudo-seminorm (recall  $\Psi := \cup_{j \geq j_0} \Psi_j$ ). Choose a freely customisable value  $\kappa > 1$  for the admissible factor of precision loss.*

*For any smooth function  $f \in B$  and any approximation  $\hat{f} \in H := H_0^1(\Omega)$ , the greatest  $\eta$  such that  $\|f - P_\eta \hat{f}\|_H \leq \kappa \|f - \hat{f}\|_H$  satisfies  $\|P_\eta \hat{f}\|_B \leq C_\kappa (\|f\|_B + \|f - \hat{f}\|_H)$  and  $|P_\eta \hat{f}|_B \leq C_\kappa |f|_B$ , with  $C_\kappa$  independent of  $f$  and  $\hat{f}$ , and with notations*

$$P_\eta g = \sum_{\lambda \in \Phi_{j_0}} \langle \tilde{\lambda} | g \rangle \lambda + \sum_{\lambda \in \Psi_\eta(g)} \langle \tilde{\lambda} | g \rangle \lambda$$

and

$$\Psi_\eta(g) := \{\lambda \in \Psi : \|\langle \tilde{\lambda} | g \rangle \lambda\|_{W^{1,p'}} \geq \eta\}$$

with  $p' := \frac{2-p}{p} \frac{d}{s-1} < 2$ . An explicit value of  $\eta$  satisfying the three inequalities is  $\eta = C'_\kappa |f|_B^{-p/(2-p)} \|f - \hat{f}\|_H^{2/(2-p)}$  with  $C'_\kappa$  independent of  $f, \hat{f}$ .

In this theorem,  $\|f - \hat{f}\|_H$  can be replaced by a greater real number: see the proof. We postpone cardinality estimation to next theorem.

**■** For the proof, we switch to an equivalent definition with  $\Psi_\eta(g) := \{\lambda \in \Psi : 2^{rj\lambda} |\langle \tilde{\lambda} | g \rangle| \geq \eta\}$  with  $r := \frac{d}{2} + \frac{2-sp}{2-p} = 1 - \frac{p}{2-p} (s-1 - \frac{d}{p} + \frac{d}{2})$ .

We need two reformulations of the pseudo-seminorm  $| \cdot |_B$  based on Fubini:

$$|g|_B = \|p\eta^{p-1} \sum_{\lambda \in \Psi: 2^{rj\lambda} |\langle \tilde{\lambda} | g \rangle| \geq \eta} 2^{2(1-r)j\lambda} \|\cdot\|_{L^1(\eta>0)}^{\frac{1}{p}}\|; \quad (8a)$$

$$|g|_B = \|(2-p)\eta^{-(3-p)} \sum_{\lambda \in \Psi: 2^{rj\lambda} |\langle \tilde{\lambda} | g \rangle| < \eta} 2^{2j\lambda} \langle \tilde{\lambda} | g \rangle^2 \|\cdot\|_{L^1(\eta>0)}^{\frac{1}{p}}\|. \quad (8b)$$

Let  $a \in [ \|f - \hat{f}\|_H, \infty)$ . To obtain  $\|f - P_\eta \hat{f}\|_H^2 \leq \kappa a$ , we will use  $\beta := \frac{2}{\kappa-1}$ :

$$\|f - P_\eta \hat{f}\|_H^2$$

$$\begin{aligned}
&= \left[ \left\| \sum_{\lambda \in \Phi_{j_0}} \langle \tilde{\lambda} | f - \hat{f} \rangle \nabla \lambda \right\|_{L^2}^2 + \sum_{\lambda \in \Psi: 2^{r_{j\lambda}} |\langle \tilde{\lambda} | \hat{f} \rangle| \geq \eta} 2^{2j\lambda} \langle \tilde{\lambda} | f - \hat{f} \rangle^2 \right] + \sum_{\lambda \in \Psi: 2^{r_{j\lambda}} |\langle \tilde{\lambda} | \hat{f} \rangle| < \eta} 2^{2j\lambda} \langle \tilde{\lambda} | f \rangle^2 \\
&\leq \left[ a^2 - \sum_{\lambda \in \Psi: 2^{r_{j\lambda}} |\langle \tilde{\lambda} | \hat{f} \rangle| < \eta} 2^{2j\lambda} \langle \tilde{\lambda} | f - \hat{f} \rangle^2 \right] + \sum_{\substack{\lambda \in \Psi: 2^{r_{j\lambda}} |\langle \tilde{\lambda} | \hat{f} \rangle| < \eta \\ \text{and } 2^{r_{j\lambda}} |\langle \tilde{\lambda} | f \rangle| \geq (1+\beta)\eta}} 2^{2j\lambda} \langle \tilde{\lambda} | f \rangle^2 + \sum_{\substack{\lambda \in \Psi: 2^{r_{j\lambda}} |\langle \tilde{\lambda} | \hat{f} \rangle| < \eta \\ \text{and } 2^{r_{j\lambda}} |\langle \tilde{\lambda} | f \rangle| < (1+\beta)\eta}} 2^{2j\lambda} \langle \tilde{\lambda} | f \rangle^2 \\
&\leq (1 + \frac{1}{\beta})^2 a^2 + \sum_{\substack{\lambda \in \Psi: 2^{r_{j\lambda}} |\langle \tilde{\lambda} | \hat{f} \rangle| < \eta \\ \text{and } 2^{r_{j\lambda}} |\langle \tilde{\lambda} | f \rangle| \geq (1+\beta)\eta}} 2^{2j\lambda} (\langle \tilde{\lambda} | f \rangle^2 - (1 + \frac{1}{\beta})^2 \langle \tilde{\lambda} | f - \hat{f} \rangle^2) + \sum_{\lambda \in \Psi: 2^{r_{j\lambda}} |\langle \tilde{\lambda} | f \rangle| < (1+\beta)\eta} 2^{2j\lambda} \langle \tilde{\lambda} | f \rangle^2 \\
&\leq (1 + \frac{1}{\beta})^2 a^2 + 0 + ((1+\beta)\eta)^{2-p} |f|_B^p \quad (8c)
\end{aligned}$$

Set  $C'_\kappa$  to  $(\frac{2\beta+3}{\beta^2})^{\frac{1}{2-p}} \frac{1}{1+\beta}$  and  $\eta$  as given in theorem statement: we get  $\|f - P_\eta \hat{f}\|_H \leq (8c) = \kappa a$ : end of first step. Similarly but without using lemma I.11, we obtain that  $\|f - P_\theta \hat{f}\|_H^2 \leq (1 + \frac{1}{\beta})^2 a^2 + ((1+\beta)\eta)^{2-p} |f|_B^p a_\theta$  (where  $\int_0^\infty a_\theta \frac{d\theta}{\theta} = \frac{1}{2-p}$ ), which allow us to study the smoothness of  $P_\eta \hat{f}$ :

$$\begin{aligned}
|P_\eta \hat{f}|_B^p &= \int_\eta^\infty (2-p)\theta^{-(3-p)} \|P_\theta P_\eta \hat{f} - P_\eta \hat{f}\|_H^2 d\theta \\
&\leq \int_\eta^\infty (2-p)\theta^{-(3-p)} (2\|f - P_\theta P_\eta \hat{f}\|_H^2 + 2\|f - P_\eta \hat{f}\|_H^2) d\theta \\
&\leq 2((1 + \frac{1}{\beta})^2 a^2 + \kappa^2 a^2) \eta^{-(2-p)} + 2(1+\beta)^{2-p} |f|_B^p
\end{aligned}$$

Replacing  $\eta^{-(2-p)}$ , we get  $|P_\eta \hat{f}|_B^p \lesssim |f|_B^p$ . Final step:

$$\begin{aligned}
\|P_\eta \hat{f}\|_B^p &= |P_\eta \hat{f}|_B^p + \sum_{\lambda \in \Phi_{j_0}} 2^{p(s-d/p+d/2)j_0} |\langle \tilde{\lambda} | \hat{f} \rangle|^p \\
&\lesssim |P_\eta \hat{f}|_B^p + \sum_{\lambda \in \Phi_{j_0}} 2^{p(s-d/p+d/2)j_0} |\langle \tilde{\lambda} | f \rangle|^p + \left\| \sum_{\lambda \in \Phi_{j_0}} \langle \tilde{\lambda} | f - \hat{f} \rangle \nabla \lambda \right\|_{L^2}^p \\
&\leq \|f\|_B^p + a^p \quad \blacksquare
\end{aligned}$$

The theorem I.12, along with properties I.14 and I.15, is valid with no limitations on the dimension  $d$ . There is a straightforward generalisation to the case  $H = B_{v,v}^u$  ( $u \geq 0, v > 1$ ) instead of  $H = H^1 = B_{2,2}^1$ , and with some work we have the case  $H = W^{s,p}$  with ideas from [5].

## B-h Restricted cardinality

To complete similarity of above theorem I.12 with theorem I.8, we need an additional result on  $\#\{\lambda \in \Psi : \|\langle \tilde{\lambda} | g \rangle \lambda\|_{W^{s,p}} \geq \eta\}$ . This additional result should be even stronger, i.e., dealing with cardinals of wavelet trees. For clarification, we add an intermediate result on a **restricted cardinal**, which

is a particular power of  $\Phi$  from [5]. Recall that  $\Psi := \cup_{j \geq j_0} \Psi_j$ ,  $\Lambda := \Phi_{j_0} \cup \Psi$ , and  $\Phi := \cup_{j \geq j_0} \Phi_j$ .

**Definition I.13** For  $V$  a subset of  $\Lambda$  or of  $\Phi \cup \Psi$ , let its restricted cardinality be

$$\#_{s,p} V := \left\| 2^{(s-1-\frac{d}{p}+\frac{d}{2})j\lambda} \right\|_{\ell^{\frac{2p}{2-p}}(\lambda \in V \setminus \Phi_{j_0})}^{\frac{d}{s-1}}$$

We have  $\#_{s,2} V = \sup\{2^{dj\lambda} : \lambda \in V\}$  (related to  $H^s$ ) and  $\#_{1+\frac{d}{p}-\frac{d}{2},p} V = \#V$ . The restricted cardinality is related to the space  $B_{p,p}^s$ : (8a) gives

$$|g|_B^p = \|p\eta^{p-1} \cdot \#_{s,p} \{\lambda \in \Psi : 2^{rj\lambda} |\langle \tilde{\lambda} | g \rangle| \geq \eta\}\|_{L^1(\eta > 0)}^{\frac{s-1}{d} \frac{2p}{2-p}} \quad (8d)$$

**Lemma I.14** restricted cardinality controls size:  $\#(V \setminus \Phi_{j_0}) \stackrel{V \subset \Phi \cup \Psi}{\lesssim} \#_{s,p} V$ .

■ Let  $r := 1 - \frac{p}{2-p}(s-1-\frac{d}{p}+\frac{d}{2})$ . Set  $A_j := \begin{cases} \Phi_j \cup \Psi_j & \text{if } j > j_0 \\ \Psi_{j_0} & \text{if } j = j_0 \end{cases}$ . Let  $C := \sup_{j \geq j_0} 2^{-dj} \#A_j$  (depends on the multiresolution). We make a level-based cut around level  $J \geq j_0$  in  $V$ :  $\#(V \setminus \Phi_{j_0}) = a_J + b_J$  with  $a_J := \sum_{j=J}^{\infty} \#(V \cap A_j)$  and  $b_J := \sum_{j=j_0}^{J-1} \#(V \cap A_j) \leq C \frac{1}{2^{d-1}} 2^{dJ}$ . Take the biggest  $J$  such that  $a_J$  is greater than  $C \frac{1}{2^{d-1}} 2^{dJ}$ :

$$\begin{aligned} \#(V \setminus \Phi_{j_0}) &\leq 2a_J = 2 \sum_{j=J}^{\infty} \left( 2^{2(1-r)j} \#(V \cap A_j) \right)^{\frac{d}{d+2(1-r)}} \left( 2^{-dj} \#(V \cap A_j) \right)^{\frac{2(1-r)}{d+2(1-r)}} \\ &\leq 2 \left( \sum_{j=j_0}^{\infty} 2^{2(1-r)j} \#(V \cap A_j) \right)^{\frac{d}{d+2(1-r)}} \left( \sum_{j=J}^{\infty} 2^{-dj} \#(V \cap A_j) \right)^{\frac{2(1-r)}{d+2(1-r)}} \\ &= 2 \left( \#_{s,p} V \right) \left( 2^{-dJ} \#(V \cap A_J) + \sum_{j=J+1}^{\infty} 2^{-dj} \#(V \cap A_j) \right)^{\frac{2(1-r)}{d+2(1-r)}} \\ &\leq 2 \#_{s,p} V \left( C + 2^{-d(J+1)} a_{J+1} \right)^{\frac{2(1-r)}{d+2(1-r)}} \lesssim \#_{s,p} V \quad \blacksquare \end{aligned}$$

**Lemma I.15** If  $s > 1 - \frac{d}{p} + \frac{d}{2}$ , the restricted cardinality  $\#_{s,p}$  also controls the set generated by cone-conditions (i.e. the wavelet tree):  $\#(V^G \setminus \Phi_{j_0}) \stackrel{V \subset \Phi \cup \Psi}{\lesssim} \#_{s,p} V$  where  $V^G := \{\lambda \in \Phi \exists \mu \in V \text{ supp } \mu \cap \text{supp } \lambda \neq \emptyset \text{ and } j_\mu \geq j_\lambda\}$ .

■ We use the method and notations  $r$ ,  $\Lambda_j$ , and  $C$  of the above proof. Let  $M$  be  $\sup_{\lambda \in \Phi \cup \Psi} \sup_{j \geq j_0} \#(\{\lambda\}^G \cap \Phi_j)$  ( $M$  depends only on the multiresolution): We make a level-based cut around level  $J > j_0$  in  $V^G$ :  $\#(V^G \setminus \Phi_{j_0}) \leq a_J + b_J$  with  $a_J := \sum_{j=J-1}^{\infty} [j-J+2] M \#(V \cap \Lambda_j)$  and  $b_J := \sum_{j=j_0+1}^{J-1} \#(V^G \cap \Phi_j) \leq C \frac{1}{2^{d-1}} 2^{dJ}$ . Take the biggest  $J$  such that  $a_J$  is greater than  $C \frac{1}{2^{d-1}} 2^{dJ}$ :

$$\begin{aligned}
\#(V^G \setminus \Phi_{j_0}) &= a_J + b_J \leq 2a_J \\
&= 2M \sum_{j=J-1}^{\infty} \left[ 2^{2(1-r)j} \#(V \cap \Lambda_j) \right]^{\frac{d}{d+2(1-r)}} \left( \frac{[j-J+2]^{1+d/(2-2r)}}{2^{dj}} \#(V \cap \Lambda_j) \right)^{\frac{2(1-r)}{d+2(1-r)}} \\
&\leq 2 \left[ \sum_{j=j_0}^{\infty} 2^{2(1-r)j} \#(V \cap \Lambda_j) \right]^{\frac{d}{d+2(1-r)}} \left( \sum_{j=J-1}^{\infty} \frac{[j-J+2]^{1+d/(2-2r)}}{2^{dj}} \#(V \cap \Lambda_j) \right)^{\frac{2(1-r)}{d+2(1-r)}} \\
&= 2 [\#_{s,p} V] \left( 2^{-dJ} \#(V \cap \Lambda_J) + (2^{-dJ} \sum_{j=J}^{\infty} \frac{[j-J+2]^{1+d/(2-2r)}}{2^{dj-dJ}} \#(V \cap \Lambda_j)) \right)^{\frac{2(1-r)}{d+2(1-r)}} \\
&\leq 2 [\#_{s,p} V] \left( C + 2^{-dJ} a_{J+1} \sup_{j \geq J} \frac{[j-J+2]^{1+d/(2-2r)}}{[j-J+1] M 2^{d[j-J]}} \right)^{\frac{2(1-r)}{d+2(1-r)}} \lesssim \#_{s,p} V \quad \blacksquare
\end{aligned}$$

We have the following Jackson-type characterisation of  $B_{p,p}^s$ :

$$\left\| 2^{k(s-1)} \inf_{\#_{s,p} V \leq 2^{dk}} \left\| f - \sum_{\lambda \in V} \langle \tilde{\lambda} | f \rangle \lambda \right\|_{\mathbb{H}^1} \right\|_{\ell^p(k \geq j_0)} < \infty \quad (9)$$

Hence the restricted cardinality is far easier to deal with than the tree-cardinality in a PDE scheme. The small drawback than only a power of the restricted cardinality is subadditive didn't complicate the proof of, e.g., next theorem:

**Theorem I.16** *Theorem I.12 also implies that the cardinality, the wavelet tree cardinality (only if  $s > 1 + \frac{d}{p} - \frac{d}{2}$ ), and the restricted cardinality of  $\Psi_{\eta}(f+e)$  are  $\leq (C_{\kappa} \frac{1}{a} |f|_B)^{\frac{d}{s-1}}$ .*

■ The subadditivity property gives  $[\#_{s,p} \Psi_{\eta}(f+e)]^{\frac{s-1}{d} \frac{2p}{2-p}} \leq a_e + a_f$  with  $a_g := [\#_{s,p} \Psi_{\frac{1}{2}\eta}(g)]^{\frac{s-1}{d} \frac{2p}{2-p}} = \sum_{\lambda \in \Psi: 2^{rj\lambda} |\langle \tilde{\lambda} | f \rangle| \geq \frac{1}{2}\eta} 2^{2(1-r)j\lambda}$ . Using (8b) and I.11, we have  $a_f \leq |f|_B^p \frac{1}{1-e^{-p}} (\frac{1}{2}\eta)^{-p}$ . Besides,  $a_e \leq \sum_{\lambda \in \Psi: 2^{rj\lambda} |\langle \tilde{\lambda} | e \rangle| \geq \frac{1}{2}\eta} (\frac{1}{2}\eta)^{-2} 2^{s j \lambda} \langle \tilde{\lambda} | e \rangle^2 \leq \frac{4}{\eta^2} \|e\|_H^2 \leq \frac{4a^2}{\eta^2}$ . Replace now  $\eta$  and redefine once  $C_{\kappa}$ . ■

## C Algorithm

The framework of our adaptive algorithm works for the solution  $u_\infty$  of a general nonlinear equation  $F(u_\infty) = 0$ , with

$$F(u)(x) = f(x, u(x), \nabla u(x), \Delta u(x)). \quad (10)$$

First, we present some assumptions on this equation.

### C-a Assumptions

The solution  $u_\infty$  of support  $\Omega$  we are numerically computing should have a good Besov regularity:  $u_\infty \in B := B_{p,p}^s(\Omega) \cap H$  with  $H := H_0^1(\Omega)$ , and  $0 < p < 2$ , and  $s > 1 + \frac{d}{p} - \frac{d}{2}$ , where  $\Omega := (0,1)^d$  and  $d=1$  or  $d=2$ .

We need two assumptions.

The first assumption, I.17, enables Newton-type convergence.

**Assumption I.17** *There exists an  $H$ -neighbourhood  $U$  of  $u_\infty$  such that  $F$  is continuous  $U \rightarrow H^{-1}$ , and:*

- 1.3.1. *The differential application  $DF$  of  $F$  is bounded from  $U$  to the set of isomorphisms  $H \rightarrow H^{-1}$ ;*
- 1.3.2.  *$[DF]^{-1} \circ F$  is bounded  $B \rightarrow B$ :  $\|DF(u)^{-1}F(u)\|_B \stackrel{u \in U}{\lesssim} 1 + \|u\|_B$ ;*
- 1.3.3. *Moreover  $\forall u \in U$ ,  $DF(u)$  is a symmetric positive (possibly unbounded) endomorphism of  $L^2(\Omega)$ ;*
- 1.3.4.  *$DF$  is Lipschitz  $U \rightarrow L(H; H^{-1})$ .*

The assumption I.18, helping the truncation, deals with technicalities specific to our approach:  $DF(u) \cdot v - F(u)$  has to be decomposed in two parts denoted  $X_{u,v}$  and  $Y_{u,v}$ . The first part should be linear with classical operators, for which the existing work [3] provide an implemented optimal cost framework; the second part should be without derivative term at all.

**Assumption I.18** *The residue of equation  $[DF(u)] \cdot v = F(u)$  splits to the form  $X_{u,v} + Y_{u,v} = F(u) - DF(u) \cdot v$  (a second order part  $X_{u,v}$ , and a zero order nonlinear part  $Y_{u,v}$ ) where:*

- 1.3.1. *When  $u \in B$  and  $v \in B$ ,  $X_{u,v}$  is in  $B_{p,p}^{s-2}$ .*

- 1.3.2. There is an implemented optimal-order-cost method  $\mathbf{X}(\varepsilon, u, v)$  returning a finite wavelet expansion in the dual wavelet basis such that  $\|\mathbf{X}(\varepsilon, u, v) - X_{u,v}\|_{\mathbb{H}^{-1}} \stackrel{\varepsilon, u, v}{\lesssim} (1 + \|u\|_B + \|v\|_B)^a \varepsilon$  with  $a > 0$ ; its computational cost is  $\stackrel{\varepsilon, u, v}{\lesssim} (1 + \|u\|_B + \|v\|_B)^b \varepsilon^{-\frac{d}{s-1}} + \sharp_{s,p} u + \sharp_{s,p} v$  with  $b > 0$ .
- 1.3.3.  $Y_{u,v}(x)$  is of the form  $Y(u(x), v(x))$  with  $Y \in C^2(\mathbb{R}^2)$  independent of  $x \in \Omega$ .

## C-b Newton's algorithm

Assumption 1.3.1 allows us to use the Newton method which constructs a sequence  $u_0, u_1, u_2, \dots$  converging to  $u_\infty$ . Each step of Newton's algorithm (algorithm C-b) defines  $u_{i+1}$  as the solution of the linearised problem at  $u_i$ :

Let  $u_0$  be the initial guess.

For  $i = 0, 1, 2, \dots$ ; exit if  $u_i$  is close enough to  $u_\infty$ .

$$u_{i+1} := u_i - [DF(u_i)]^{-1} F(u_i). \quad (11a)$$

Algorithm 1 – Newton's algorithm.

**Property I.19** When the algorithm converges, the error is squared at each iteration:  $\forall i \geq i_0 \ C_F \|u_i - u_\infty\|_H \leq (C_F \|u_{i_0} - u_\infty\|_H)^{2^{i-i_0}}$ , with

$$C_F := \sup_{\substack{u, v \in U \\ u \neq v}} \frac{\|[DF(u)]^{-1}[DF(u) - DF(v)]\|_{\mathbb{L}(H, H)}}{2\|u - v\|_H}.$$

■  $C_F < \infty$  according to 1.3.1 and 1.3.4. We will replace  $F(u_i)$  by its Taylor expansion  $F(u_\infty) + \int_0^1 [DF(t \cdot u_i + (1-t) \cdot u_\infty)](u_i - u_\infty) dt$  in following formula:

$$u_{i+1} - u_\infty = [DF(u_i)]^{-1} [DF(u_i)](u_i - u_\infty) - [DF(u_i)]^{-1} F(u_i)$$

As  $F(u_\infty) = 0$ , we get  $u_{i+1} - u_\infty = \int_0^1 [DF(u_i)]^{-1} [DF(u_i) - DF(t \cdot u_i + (1-t) \cdot u_\infty)](u_i - u_\infty) dt$ . According to the definition of  $C_F$ , we have finally  $\|u_{i+1} - u_\infty\|_H \leq C_F \|u_i - u_\infty\|_H^2$ . ■

The constant  $C_F$  also allows us to quantify the convergence when a geometrically decreasing drift is made in step (11a), as shown in the following

property. For an optimal-cost algorithm, it is important for the resulting convergence rate to decrease with the *same rate*, hence we need a lemma:

**Lemma I.20** For  $\rho_1 \in (0,1]$ , let  $(\nu_i)_{i \in \mathbb{N}}$  be the sequence defined by  $\nu_0 = \frac{1}{2}$  and  $\nu_{i+1} = \frac{1}{4}\rho_1^i + \nu_i^2$ . Then,  $\nu_i \stackrel{i}{\sim} \rho_1^i$ . [proof in Appendix]

**Property I.21** Let  $\kappa_1 > 1$  and  $\rho_1 \leq 1$ . Suppose that  $\|u_0 - u_\infty\|_H \leq \frac{\nu_0}{\kappa_1 C_F}$  and that step (11a) is computed with an error geometrically decreasing so that  $\forall i \in \mathbb{N}$ :

$$\|u_{i+1} - u_\infty\|_H \leq \kappa_1 \|(u_i - DF(u_i)^{-1}F(u_i)) - u_\infty\|_H + \frac{1}{4C_F\kappa_1}\rho_1^i \quad (12)$$

Then the resulting algorithm converges, with the same rate:

$$\|u_i - u_\infty\|_H \leq \frac{\nu_i}{\kappa_1 C_F} \stackrel{i}{\sim} \rho_1^i$$

■ Proof of property I.19 gives  $\|(u_i - DF(u_i)^{-1}F(u_i)) - u_\infty\|_H \leq C_F \|u_i - u_\infty\|_H^2$ . Use lemma I.20 on  $[\kappa_1 C_F \|u_{i+1} - u_\infty\|_H] \leq [\kappa_1 C_F \|u_i - u_\infty\|_H]^2 + \frac{1}{4}\rho_1^i$ . ■

As  $\|(u_i - DF(u_i)^{-1}F(u_i)) - u_\infty\|_H$  is the error for exact Newton algorithm, (12) allows us to inject (see table C-b) a discretization drift of  $\frac{1}{4C_F\kappa_1^2}\rho_1^i$  and a restricted truncation which degrades precision by the factor  $\kappa_1$  (theorem I.12) in this Newton iteration (11a). That truncation insures that  $u_{i+1}$  has the optimal memory size given its  $H_0^1$ -distance to  $u_\infty$ , and that  $\|u_{i+1}\|_B$  is bounded (the next loop will need it!).

Let  $u_0$  be the initial guess.

For  $i = 0, 1, 2, \dots$ ; exit if  $u_i$  is close enough to  $u_\infty$ .

Compute  $v_i$  satisfying  $\|v_i - DF(u_i)^{-1} \cdot F(u_i)\|_H \leq \frac{1}{4C_F\kappa_1^2}\rho_1^i$ . (13)

Let  $u_{i+1}$  be the  $P_\eta(f+e)$  of restricted truncation theorem I.12  
with  $\kappa := \kappa_1$  and  $e := u_i - v_i - u_\infty$ , given  $f := u_\infty$  is in  $B_{p,p}^s$ .

Algorithm 2 – Newton’s geometrical-error-proof algorithm.



### C-c Preconditioned gradient algorithm

Step (13) will be implemented with a preconditioned gradient method. Recall that  $\Phi_{j_0}$  contains scale function of level  $j_0$ , that  $\Psi$  contains all wavelets of the wavelet basis, and that  $\tilde{\mu}$  is the dual element defined by notation I.2.

**Definition I.22** *Let the preconditioner  $P$  be the  $L^2$ -symmetric nondegenerate positive linear operator defined by the conditions  $\forall \mu \in \Phi_{j_0} P(\tilde{\mu}) \in V_{j_0}$ ,  $\forall \lambda, \mu \in \Phi_{j_0} \langle \lambda | -\Delta P(\tilde{\mu}) \rangle = \langle \lambda | \tilde{\mu} \rangle$ , and  $\forall \mu \in \Psi P(\tilde{\mu}) = \langle \mu | -\Delta \mu \rangle^{-1} \cdot \mu$ .*

The Laplacian preconditioner is fitted for  $\Phi_{j_0}$ , but not for boundaries of  $\Psi_j$ . (Instead we have fitted the boundaries of  $\Psi_j$  so that they are orthogonal and have orthogonal derivatives (using freedom degrees of modification B-b), so that, e.g., the first three elements of  $\Psi_j^{(0,1)}$  in (4) look like Laplacian eigenvectors.)

**Lemma I.23**  $\langle P^{-1}f | f \rangle = \|f\|_H^2$  and  $-P\Delta$  is an isomorphism of  $H$ . ■  
Easy. ■

**Property I.24** *Define the spectrum bounds  $\lambda_{\max} := \sup_{u \in U} \lambda_{\max}(PDF(u))$  and  $\lambda_{\min} := \inf_{u \in U} \lambda_{\min}(PDF(u))$ . As  $\lambda_{\max} < \infty$  and  $\lambda_{\min} > 0$ , the numbers  $\rho := \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}$  and  $\omega := \frac{2}{\lambda_{\min} + \lambda_{\max}}$  are real.*

■  $P^{1/2}DF(u)P^{1/2}$  is a symmetric positive  $U$ -uniform  $L^2$ -automorphism according to above lemma and to elements 1.31, 1.33 of assumption I.17.  $P^{1/2}$  is an isomorphism  $L^2 \rightarrow H$  and is  $L^2$ -symmetric. ■

In the step (13), we need to compute an approximation  $r_{i,k}$  of the residue  $F(u_i) - DF(u_i)v_{i,k}$  which, thank to assumption I.18, rewrites itself to  $X_{u_i, v_{i,k}} + Y_{u_i, v_{i,k}}$  where the second term contains all non-linear terms. We have to compute  $r_{i,k}$  as a **wavelet expansion in the dual wavelet basis  $\tilde{A}$** , so that we will be able to use the preconditioner  $P$  and to make the iteration  $v_{i,k+1} = v_{i,k} - \omega P r_{i,k}$  of a gradient method.

Assumption 1.32 provides an out-of-the-box method **X** approximating the linear part  $X_{u_i, v_{i,k}}$ : we have to construct a method approximating the zero-order nonlinear part  $Y_{u_i, v_{i,k}}$ ; that method contains a blind truncation, and an approximate computation of truncated coefficients.

The blind truncation of  $Y_{u_i, v_{i,k}}$  (i.e. localization of its main coefficients without actually computing any other coefficients) is based on localization heuristics freely adapted from [3] which add more wavelets around big wavelet coefficients of  $u_i$  and of  $v_{i,k}$  (see (17c) and (17d)).

The approximate computation of  $\langle Y_{u_i, v_{i,k}} | \lambda \rangle$  with a collocation is expensive to compute naively when the support of  $\lambda$  is large compared to the local refinement of  $u_i$  and  $v_{i,k}$ . As pointed out by [10], we have to factorise this computation according to (1) and (2), i.e., using the *adaptive decomposition* and the *adaptive reconstruction* (see the following).

### C-d Adaptive decomposition and reconstruction

Since the scale equations (1) and (2) above correspond to  $\Omega = \mathbb{R}$ , we have to reformulate them to use a wavelet multiresolution on general  $\Omega$ . Recall that  $\Phi$  contains the scale function of any scale, and that  $j_\lambda$  is the scale value of any wavelet  $\lambda$  or scale function  $\lambda$ .

**Definition I.25** Let  $\lambda \ll \mu$  (read "the scale equation of scale function  $\lambda$  contains  $\mu$ ") be the binary relation defined for  $\lambda$  and  $\mu$  elements of  $\Phi \cup \Psi$ , true if and only if  $\lambda \in \Phi$  and  $\mu \in \Phi_{j_\lambda+1}$  and  $\text{supp} \mu \subset \text{supp} \lambda$ . (1) becomes:  $\forall \lambda \in \Phi$

$$\lambda = \sum_{\mu \succ \lambda} \langle \lambda | \tilde{\mu} \rangle \mu \quad \text{and} \quad \tilde{\lambda} = \sum_{\tilde{\mu} \tilde{\succ} \tilde{\lambda}} \langle \tilde{\lambda} | \mu \rangle \tilde{\mu} \quad (14)$$

**Definition I.26** Let  $\lambda \lll \mu$  (read "the scale equation of wavelet  $\lambda$  contains  $\mu$ ") be the binary relation defined for  $\lambda$  and  $\mu$  elements of  $\Phi \cup \Psi$ , true if and only if  $\lambda \in \Psi$  and  $\mu \in \Phi_{j_\lambda+1}$  and  $\text{supp} \mu \subset \text{supp} \lambda$ . (2) becomes:  $\forall \lambda \in \Psi$

$$\lambda = \sum_{\mu \gg \lambda} \langle \lambda | \tilde{\mu} \rangle \mu \quad \text{and} \quad \tilde{\lambda} = \sum_{\tilde{\mu} \gg \tilde{\lambda}} \langle \tilde{\lambda} | \mu \rangle \tilde{\mu} \quad (15)$$

$\tilde{\succ}$  and  $\tilde{\gg}$  in (14) and (15) are the equivalent notations fitted for dual scale equations. To compute point-values of a wavelet expansion, the fastest way is to reconstruct the function as combination of scale functions belonging to the thinnest scale. In the context of adaptivity, this thinnest scale can vary between two points of  $\Omega$ , because we want to avoid making thinnest computations where thinnest wavelet coefficients are absent. Hence, we patch together different scales [10] to form the graded layer of thinnest scale functions, layer

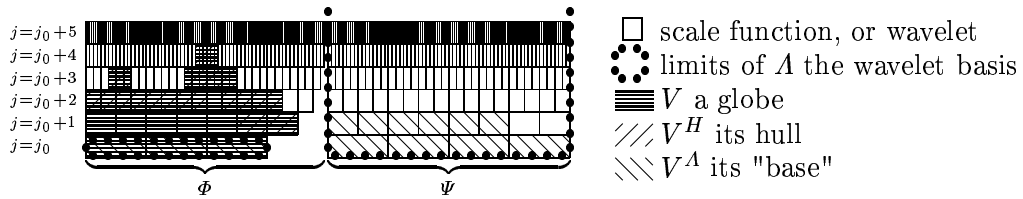
that we will call a **hull** (see  $V^H$  in next illustration). The expression of a function in the **hull** context does not allow to compute all of its wavelet coefficients, but only coefficients for the set  $V^A$  of wavelet that are locally coarser than the hull  $V^H$ . The transitions between  $V^A$  and  $V^H$ , called adaptive decomposition and adaptive reconstruction, use the notion of **globe**, which is a progressively refined set of scale functions, the globe  $V^{HG}$  containing all scale function locally coarser than the hull  $V^H$ .

**Definition I.27** A finite subset  $V$  of  $\Phi$  is said to be a **globe** if it contains  $\Phi_{j_0}$  and if it is  $\succ$ -stable<sup>1</sup> (cone condition, i.e.,  $\forall \lambda, \mu (\lambda \succ \mu \text{ and } \lambda \in V) \Rightarrow \mu \in V$ ).

**Definition I.28** ( $V^A := \Phi_{j_0} \cup V^\Psi$ ,  $V^\Psi :=$  wavelets contained in  $\text{span} V$ ) Let the application **basing** be defined on the set of globes by  $V \mapsto V^A$  where  $V^A := \Phi_{j_0} \cup V^\Psi$  with  $V^\Psi$  the set of wavelets  $\lambda \in \Psi$  such that  $\forall \mu \succ \lambda, \mu \in V$ .

**Definition I.29** ( $W^G$ , globe generated by  $W$ ) Let the application **globing** be defined on the set of finite subsets  $W$  of  $\Phi \cup \Psi$  by  $W \mapsto W^G$  where  $W^G$  is the smallest globe  $V$  such that  $W \cap \Phi \subset V$  and  $W \cap \Psi \subset V^\Psi$  (compatible with notation  $V^G$  in property I.15 controlling the size of wavelet tree).

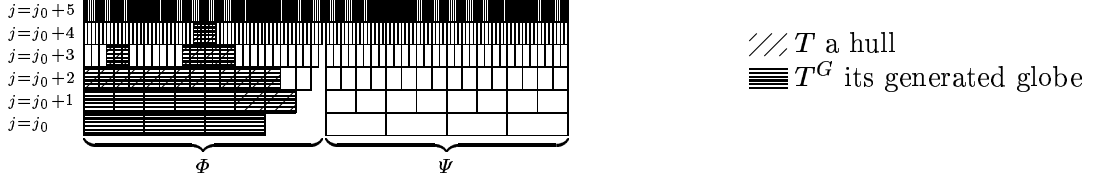
**Definition I.30** ( $V^H$ , thinnest part of  $V$ ) Let the application **hulling** be defined on the set of finite subsets of  $\Phi$  with  $V \mapsto V^H$  where  $V^H := \{\lambda \in V : \exists \mu \succ \lambda \ \mu \notin V\}$  contains the locally thinnest elements of  $V$ .



**Definition I.31** Define a **hull** as an element of the image of the application **hulling**. That application is injective, and its inverse is the application **globing** (restricted to the set of hulls).

1. If several functions have the same support, a globe must contain either all or none of them

The complement in  $\Phi$  of a hull  $T$  has two "connected components":  $T^G \setminus T$  with coarser scale functions, and  $\Phi \setminus T^G$  with arbitrary thinner scale functions:



The applications basing, bordering and globing have natural extension to coefficients, using the scale equations (15) and (14).

**Definition I.32** Let  $T$  be a hull, and  $(a_\lambda)_{\lambda \in T}$  a family. Inserting (14) and (15) anywhere leads to compute one real  $a_\lambda$  for each  $\lambda \in \Phi \cup \Psi$  which is coarser than  $T$ . Let the **dual adaptive decomposition** of this family be  $\sum_{\lambda \in T^{GA}} a\{\lambda\} \tilde{\lambda}$  which admits the following formal definition:

$$a\{\lambda\} := \begin{cases} a_\lambda & \text{if } \lambda \in T \\ \sum_{\mu \triangleright \lambda} \langle \lambda | \tilde{\mu} \rangle \cdot a\{\mu\} & \text{if } \lambda \in T^G \setminus T \text{ (then all these } \mu \text{ are in } T^G) \\ \sum_{\mu \triangleright \lambda} \langle \lambda | \tilde{\mu} \rangle \cdot a\{\mu\} & \text{if } \lambda \in T^{G\Psi} \text{ (then all these } \mu \text{ are in } T^G) \end{cases}$$

The dual adaptive decomposition uses (14) between each level, and each use will degrade the precision. Hence, we have a level-sensitive error growth.

**Property I.33** The dual adaptive decomposition I.32 (defining notation  $\{ \}$ ) is stable in "regularity"  $\nu \in (0,2)$ : if  $T$  is a hull and  $(e_\lambda)_{\lambda \in T}$  a family, then

$$\|2^{\nu j_\lambda} e\{\lambda\}\|_{\ell^2(\lambda \in T^{GA})} \lesssim \|2^{\nu j_\lambda} e_\lambda\|_{\ell^2(\lambda \in T)}.$$

The real  $\nu$  is the one such that there exists a customisable weight  $\chi : \Phi \rightarrow (m,1)$  with  $m$  some positive real, and a constant  $C_\chi \in (1,2^\nu)$  (measuring the quality of weighted stability of the scale equation (14)) such that  $\forall j \geq j_0 \forall (e_\lambda)_\lambda \in \mathbb{R}^{\Phi_{j+1}}$ :

$$\sum_{\mu \in \Phi_j} \left( \chi(\mu) \sum_{\lambda \in \Phi_{j+1}} \langle \mu | \tilde{\lambda} \rangle e_\lambda \right)^2 \leq C_\chi^2 \sum_{\lambda \in \Phi_{j+1}} \chi(\lambda)^2 e_\lambda^2 \quad (16)$$

In practical cases, there are values of  $\chi$  for which  $\nu$  and  $C_\chi - 1$  are small.

■ Definitions I.32 and I.33 give, for  $j \geq j_0$ :

$$\sum_{\lambda \in (T^G \setminus T) \cap \Phi_j} \chi(\lambda)^2 2^{2\nu j \lambda} e\{\lambda\}^2 \leq 2^{-2\nu} C_\chi^2 \sum_{\lambda \in T^G \cap \Phi_{j+1}} \chi(\lambda)^2 2^{2\nu j \lambda} e\{\lambda\}^2.$$

Using a backward induction on  $j$  such that  $j_0 \leq j \leq \sup_{\lambda \in T^G} j_\lambda$ , we have:

$$\sum_{\lambda \in T^G \cap \Phi_j} \chi(\lambda)^2 2^{2\nu j \lambda} e\{\lambda\}^2 \leq \sum_{l \geq j} (2^{-2\nu} C_\chi^2)^{l-j} \sum_{\lambda \in T \cap \Phi_l} \chi(\lambda)^2 2^{2\nu j \lambda} e_\lambda^2$$

We see the error growth due to repeating (14). The sum  $\sum_{j=j_0}^\infty$  gives:

$$\sum_{\lambda \in T^G} \chi(\lambda)^2 2^{2\nu j \lambda} e\{\lambda\}^2 \leq \frac{1}{1-2^{-2\nu} C_\chi^2} \sum_{\lambda \in T} \chi(\lambda)^2 2^{2\nu j \lambda} e_\lambda^2;$$

End of proof with  $\sum_{\lambda \in T^G \setminus \Phi_{j_0}} 2^{2\nu(j_\lambda+1)} e\{\lambda\}^2 \lesssim \sum_{\lambda \in T^G \setminus \Phi_{j_0}} 2^{2\nu j \lambda} e\{\lambda\}^2$  ■

**Definition I.34** *Let  $W$  be a finite subset of  $\Lambda$ , and  $(b_\lambda)_{\lambda \in W}$  a family. Let the **primal adaptive reconstruction** of this family be  $(\langle \sum_{\lambda \in W} b_\lambda \lambda | \tilde{\lambda} \rangle)_{\lambda \in W^{GH}}$ .*

The primal adaptive reconstruction and the dual adaptive decomposition are not converse methods. They were first defined in [10] respectively as *Transform A* and *Transform B*.

**Definition I.35** *We define the **overglobe**, on the hull of which two nonconsecutive scales  $\Phi_j$  and  $\Phi_k$  with  $|k-j| \geq 2$  do not interfere. Hence, we can almost use standard nonadaptive techniques. For  $V$  a globe, let its overglobe  $V^S$  be the globe defined by:*

$$V^S := \Phi_{j_0} \cup \left\{ \lambda \in \Phi : \exists \mu \in V \cap \Phi_{j_\lambda} \exists \lambda' \ll \lambda \exists \mu' \ll \mu \exists x \in \Omega \lambda'(x) \mu'(x) \neq 0 \right\}$$

**Property I.36 (clean overlapping)** *For  $V$  a globe, for any  $\lambda \in V^{SH}$  and  $\mu \in V^S \cap \Phi_{j_\lambda+2}$  we have  $\forall x \in \Omega \lambda(x) \mu(x) = 0$ . [proof in Appendix]*

### C-e Description of the Algorithm

The notations defined above allow us to describe the actual implementations of step (13), resulting in the algorithm given by table 3.

We compute the zero order nonlinear part  $Y_{u_i, v}$  of the Galerkin residue of linearised equation  $DF(u_i)v = F(u_i)$  against a scale function  $\lambda \in \Phi$  with a collocation denoted  $\Gamma_\lambda$  for the measure  $\lambda(x)dx$ . We assume on this collocation that  $\int \lambda f = \Gamma_\lambda(f)$  for  $f \in \Phi_{j_\lambda+1}$ , and that  $\lambda(x)\Gamma_\lambda(dx)$  is nonnegative.

This collocation should only be used on a refined enough hull (17f), otherwise its precision is lost. We compute the set  $H_{i,k}$  which is a hull (depending on  $u_i$  and on  $v_{i,k}$ ) containing elements  $\lambda$  of  $\Phi$  such that  $\langle Y(u_i, v_{i,k}) | \lambda \rangle$  is cheap to compute. We make a primal adaptive reconstruction on  $u_i$  and on  $v_{i,k}$  beforehand, and a dual adaptive decomposition subsequently (17g). Taking in account the relative stability of this decomposition, we need a real  $\nu \in (0, 2)$  such that  $2^\nu > C_\chi$ ,  $\frac{3d+4}{4} < s \leq 3 - \nu$  and  $s > \frac{\nu}{2} + \frac{d}{p} + \frac{d}{4}$ . In the wavelet basis chosen in our example, we have  $C_\chi = 1$  in (16), and we set  $s = \frac{299}{100}$ ,  $p = \frac{5}{12}$  and  $\nu = \frac{1}{100}$ . For the same 2D example (i.e. with  $d = 2$ ),  $s = \frac{299}{100}$ ,  $p = \frac{400}{497}$  and  $\nu = \frac{1}{100}$  fits.

Let  $(u_0, A_0)$  be the initial guess of a solution of  $F(u) = 0$  (see (10)).

For  $i = 0, 1, \dots$ ; exit when  $u_i$  is close enough to  $u_\infty$ . (Newton algorithm)

Let  $(v_{i,0}, A_{i,0}) := (0, \Phi_{j_0})$ .

For  $k$  in  $\{0, 1, \dots, N_3 - 1\}$  (gradient algorithm on  $DF(u_i)v = F(u_i)$ ):

Let  $x_{i,k} := \mathbf{x}(N_4^{-1} \rho_1^i \rho_2^k, u_i, v_{i,k})$  ( $\mathbf{x}$  is in assumption 1.3.2). (17a)

Let  $V_4^{i,k} := \{\lambda \in \Phi \cup \Psi : \exists \mu \in A_i \cup A_{i,k} \text{ supp } \mu \cap \text{supp } \lambda \neq \emptyset \text{ and } j_\lambda \leq j_\mu\}$ . (17b)

For  $\mu$  in  $A_i \cup A_{i,k}$  (refinement of  $V_4^{i,k}$  around  $\mu$ ):

Add  $\{\lambda \in \Phi \cup \Psi : \text{supp } \mu \cap \text{supp } \lambda \neq \emptyset \text{ and } j_\mu < j_\lambda < j_\mu + k_\mu\}$  to  $V_4^{i,k}$ ,  
where  $k_\mu \in \mathbb{R}$  satisfies  $2^{k_\mu} =$

$\max\left(\left(\frac{1}{N_4} \rho_1^i \rho_2^k\right)^{-\frac{1}{2}} \max(|\langle u_i | \tilde{\mu} \rangle|, |\langle v_{i,k} | \tilde{\mu} \rangle|)^{\frac{1}{d+2}} 2^{-\frac{1}{d+2} j_\mu}, 2\right)$  (17c)

$\left(\frac{1}{N_4} \rho_1^i \rho_2^k\right)^{-\frac{1}{2-\nu}} \max(|\langle u_i | \tilde{\mu} \rangle|, |\langle v_{i,k} | \tilde{\mu} \rangle|)^{\frac{2}{2d+2-\nu}} 2^{\frac{2\nu+d}{4d+4-2\nu} j_\mu}$ . (17d)

Let  $V_5^{i,k} := \{\mu \in \Phi : \exists \lambda \in V_4^{i,k} \lambda \ll \mu \text{ or } \lambda \ll \mu\}$ . (17e)

Let  $H_{i,k} := \text{refined hull} [(A_i \cup A_{i,k})^{GS} \cup (V_5^{i,k})^G]^H$ . (17f)

Let  $y_{i,k} := \text{adaptive decomposition of } (\Gamma_\lambda(Y(u_i, v_{i,k})))_{\lambda \in H_{i,k}}$ . (17g)

Let  $r_{i,k} := x_{i,k} + y_{i,k}$  ( $\approx$  **residue** of  $DF(u_i)v_{i,k} = F(u_i)$ ). (17h)

Let  $(v_{i,k+1}, A_{i,k+1})$  be the lifting  $(P_\eta(v_{i,k} + \omega Pr_{i,k}), A_\eta(v_{i,k} + \omega Pr_{i,k}))$   
with threshold  $\eta := N_2^{-1}(\rho_1^i \rho_2^k)^{\frac{2}{2-p}}$  (from theorem I.12). (17i)

Let  $(u_{i+1}, A_{i+1})$  be the lifting  $(P_\eta(u_i - v_{i,N_3}), A_\eta(u_i - v_{i,N_3}))$  with  
threshold  $N_1^{-1} \rho_1^{\frac{2i}{2-p}}$  (from theorem I.12). (17j)

Algorithm 3 – *Optimal cost algorithm.*

Let the parameter  $\rho_1 \in (0, 1)$  be the main geometrical ratio of convergence of Newton algorithm. Part C-c defined  $\rho$  as the natural convergence rate of the undiscretized version of a gradient algorithm with preconditioner  $\omega P$ .

Let  $\rho_2 \in (\rho, 1)$  be the minor geometrical ratio of inner gradient algorithm. We use the Galerkin residue in a gradient process of convergence rate  $\rho$  and preconditioned with  $\omega P$ . Let  $\kappa_1 \in (1, \infty)$ ,  $\kappa_2 \in (1, \rho^{-1}\rho_2)$  and  $N_2$  be three threshold parameters. There exists a constant  $\varepsilon = \varepsilon(F, U, u_\infty, \rho_1, \rho_2, \kappa_1, \kappa_2, N_2)$  such that if three parameters  $N_1$ ,  $N_3$  and  $N_4$  are greater than  $\frac{1}{\varepsilon}$ , and if the initial estimation  $u_0$  satisfies  $\|u_0 - u_\infty\|_{H_0^1} \leq \varepsilon$  and has a finite wavelet expansion, the algorithm converges at optimal order rate  $\rho_1^i$  (theorem I.45) in  $H_0^1$ -norm and has optimal order cost  $\rho_1^{-i/(s-1)}$  (theorem I.47) where  $s$  is the non-linear Besov regularity of the solution. The mother wavelet should be Lipschitz at least.

## D 1D numerical example

In this section, we present the univariate equation used for a numerical application of the above algorithm.  $d$  is 1 and  $\Omega$  is the interval  $(0, 1)$ , and we use the multiresolution analysis presented in section B-b.

### D-a Definition

[12] is a numerical-application-based comparison between three non-linear multi-grid algorithms; its numerical applications are made on  $-\Delta u - e^{cu} = 0$  (where  $c > 0$ ). In dimensions 1 and 2, the bigger the parameter  $c$  is, the more difficult the convergence becomes. In order to highlight the adaptivity of our numerical approximation, we used the same equation with a cusp-like second member shown in Figure 3.

$$\forall x \in [0, 1], \quad -\Delta u(x) - e^{cu(x)} = f_{-5/6} \quad (18)$$

In terms of (10), it amounts to defining  $F(u)(x)$  as  $-\Delta u(x) - e^{cu(x)} - f_{-5/6}$  where  $f_{-5/6}$  is the illustrated distribution-based Laplacian of function  $x \mapsto -|3x - 1|^{\frac{2}{3}}$ . Its difference with  $x \mapsto 2|3x - 1|^{-\frac{4}{3}}$  is a singularity around point  $\{\frac{1}{3}\}$  (which is not a distribution).

The assumption I.18 will be reached with the following definitions:

$$\begin{aligned} Y(x, y) &= e^{cx}(cy - 1) \\ X_{u, v} &:= \Delta(v - u) - f_{-5/6} \\ Y_{u, v} &:= e^{cu}(cv - 1) \end{aligned} \quad (19a)$$

---

2. so that (19d) and (19h) (resp. utility and cost) are proportional <sup>3</sup>same for (19c) and (19h)

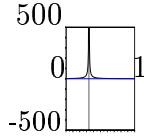


FIG. 3 – Second member  $f_{-5/6} = -\Delta(|3x - 1|^{2/3})$ .

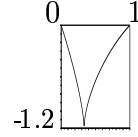


FIG. 4 – Solution  $u_\infty$  of equation (18) with  $c = 5$ .

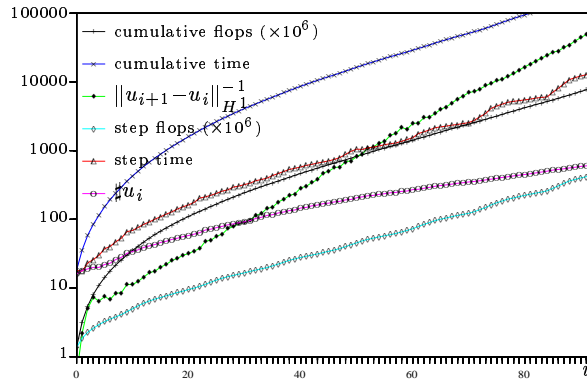


FIG. 5 – Error, flops and time measures when approximating  $u_\infty$  with  $c = 5$ .

In order to make the above algorithm clearer, it has been restricted to the case where the differential  $DF(u)$  is elliptic. There is an easy generalisation to the saddle-point case, with a work-around in step (17i). We set  $c = 5$  and we look for the solution  $u_\infty$  shown in Figure 4.

Algorithm C-e has been implemented, with  $\nu = 0.01$ ,  $s = 2.99$ ,  $p = \frac{5}{12}$ ,  $\rho_1 = 0.9$ ,  $\rho_2 = 0.95$ ,  $\omega = 0.15$ ,  $N_1 = 1$ ,  $N_2 = 3$ ,  $N_3 = 8$  and  $N_4 = 0.001$  in (17c) and  $N_4 = 5$  in (17d). According to the results in Figure 5, the number of floating point operation is slightly smaller than  $1750000 \times \varepsilon^{-0.505}$  for each iteration, so it has the right complexity rate,  $\varepsilon$  being the  $H^1$  error at the end of the iteration; a prototype is enough to validate this research implementation, hence even if its time complexity rate is greater than  $\varepsilon^{-0.505}$  (it is  $26\varepsilon^{-0.56}$  second per iteration), it is still admissible. The same example but in 2D and with optimal order and reasonable speed is to come.



## D-b Establishing assumption I.17

**Property I.37**  $u_\infty$  and  $F$  defined above satisfy the whole assumption I.17.

■ (Sketch) A result of spectral theory states that when  $E : v \mapsto -\Delta^{-1}(ce^{cu_\infty}v)$  is compact and symmetric  $H \rightarrow H$ , its spectrum is a point spectrum, and 0 is its only possible accumulation point. As  $T_u := \Delta^{-1} \circ DF(u)$  is of the form  $T_u = E + \text{Id} + C_u$  with  $C_u(v) = \Delta^{-1}(ce^{cu}v - ce^{cu_\infty}v)$ : we get that  $T_u$  is an  $H$ -isomorphism provided the spectral norm of  $C_u$  is smaller than the infimum  $r_E$  of modulus of eigenvalues of  $E + \text{Id} = \Delta^{-1} \circ DF(u_\infty)$ .

Now, let us detail one by one various points of this proof.

- **Proof that  $E$  is compact.**  $u_\infty$  is easily shown to be bounded, so that  $\Delta u_\infty = -e^{cu_\infty} - f_{5/6}$  is  $H^{-1}$ , thus  $u_\infty$  is in  $H := H_0^1([0,1])$ . Therefore, the operator  $v \mapsto ce^{cu_\infty}v$  is continuous  $H \rightarrow H$  [14];  $E$  is its composition with  $v \mapsto v$  (compact  $H \rightarrow H^{-1}([0,1])$ ) and with  $v \mapsto -\Delta^{-1}v$  (bounded  $H^{-1}([0,1]) \rightarrow H$ ), hence  $E$  is compact.

- **Proof that  $r_E > 0$ .** Assume for the sake of contradiction that  $r_E = 0$ . Then,  $E + \text{Id}$  isn't invertible;  $-1$  cannot be an accumulation point of spectrum of  $E$ , hence  $E + \text{Id} = \Delta^{-1} \circ DF(u_\infty)$  isn't injective, hence  $u_\infty$  is a double solution of  $F$ , which is clearly false.

- **Definition of  $U$**  such that for  $u \in U$ , the spectral norm of  $C_u$  is smaller than  $\frac{r_E}{2}$ . The spectral norm of  $C_u$  is smaller than its  $H \rightarrow H$  operator norm;

$$\|C_u v\|_H \lesssim \|c(e^{c(u-u_\infty)} - 1)e^{cu_\infty}v\|_{L^2} \lesssim ce^{c\|u_\infty\|_{L^\infty}} (e^{c\|u-u_\infty\|_{L^\infty}} - 1)\|v\|_H$$

Hence, there exists  $\varepsilon > 0$  such that  $\|u - u_\infty\|_H < \varepsilon$  implies that  $\|u - u_\infty\|_{L^\infty}$  is small enough for the spectral norm of  $C_u$  to be smaller than  $\frac{r_E}{2}$ . Let  $U := \{u \in H : \|u - u_\infty\|_H < \varepsilon\}$ .

Now that  $U$  is defined, we check assumption 21.

- **Proof that  $T_u = \Delta^{-1} \circ DF(u) = C_u + E + \text{Id}$  is an  $H$ -isomorphism uniformly with respect to  $u \in U$**  (i.e., 1.3.1): Spectrum of  $C_u$  is included in  $\{z \in \mathbb{C} : |z| \leq \frac{r_E}{2}\}$ , and spectrum of  $E$  in  $\{z \in \mathbb{C} : R \leq |z| \leq r_E\}$  for some  $R > 0$ . Hence spectrum of  $T_u$  is included in  $\{z \in \mathbb{C} : R+1+\frac{r_E}{2} \leq |z| \leq \frac{r_E}{2}\}$ , which is independent of  $u \in U$ .

- **Proof of  $\|[DF(u)]^{-1}F(u)\|_B \stackrel{u \in U \cap B}{\lesssim} 1 + \|u\|_B$**  (i.e., 1.3.2). As  $[DF(u)]^{-1}F(u) = T_u^{-1}\Delta^{-1}F(u) = (T_u^{-1} \circ \Delta^{-1}F)(u)$ , we split the proof as follows:

About  $T_u^{-1}$ : the operator  $-\text{Id} - T_u^{-1} = (-T_u - \text{Id}) \circ T_u^{-1} = (\Delta^{-1} \circ (ce^{cu}\text{Id})) \circ T_u^{-1} = \Delta^{-1} \circ (ce^{cu}T_u^{-1})$  is  $U$ -uniformly continuous  $B \rightarrow B$ , because this is the

composition of  $v \mapsto v$  (continuous  $B \rightarrow H$ ),  $v \mapsto T_u^{-1}v$  and  $v \mapsto ce^{cu}v$  (both  $U$ -uniformly continuous  $H \rightarrow H$ ),  $v \mapsto \Delta^{-1}v$  (continuous  $H \rightarrow H^3([0,1])$ ) and  $v \mapsto v$  (continuous  $H^3([0,1]) \rightarrow B$  as  $s \leq 3$ ).

About  $\Delta^{-1}F$ : the operator  $-\Delta^{-1}F(u) - u - \Delta^{-1}f_{-5/6} = \Delta^{-1}e^{cu}$  is  $U$ -uniformly  $B$ -bounded because  $e^{cu}$  is  $U$ -uniformly  $H$ -bounded as  $u$  is  $U$ -uniformly  $H$ -bounded [14], because  $\Delta^{-1}$  is continuous  $H \rightarrow H^3([0,1])$ , and because  $H^3([0,1])$  is continuously embedded in  $B$  (as  $s \leq 3$ ).

- **$DF(u)$  is symmetric**  $L^2 \rightarrow L^2$  (assumption 1.34): clear as  $DF(u) = -\Delta - ce^{cu}\text{Id}$ .

- **$DF$  is Lipschitz**  $U \rightarrow L(H; H^{-1})$ : as  $DF(u)v - DF(\tilde{u})v = ce^{cu}v - ce^{\tilde{c}\tilde{u}}v = ce^{cu}e^{c(u-u_\infty)}(1 - e^{c(\tilde{u}-u_\infty)})v$ , we have  $\|DF(u)v - DF(\tilde{u})v\|_{H^{-1}} \leq \|DF(u)v - DF(\tilde{u})v\|_{L^2} \leq ce^{c\|u_\infty\|_{L^\infty}}e^{c\|u-u_\infty\|_{L^\infty}}(e^{\|u-\tilde{u}\|_{L^\infty}} - 1)\|v\|_{L^2}$ ; conclusion by  $\|u - \tilde{u}\|_{L^\infty} \lesssim \|u - \tilde{u}\|_H$  and  $u \in U$ . ■

## D-c Establishing assumption I.18

**Property I.38** *Assume the following (strong) regularity property:  $\phi \in B_{\infty,\infty}^s$ ,  $\psi$  is orthogonal to  $\mathbb{R}_n[X]$  for  $n > s$  and  $\tilde{\psi} \in L^1$ .*

*Then the decomposition (19a) fulfils assumption I.18. The more general case  $F(u)(t) = -\Delta u(t) - G(u(t)) + f(t)$  satisfies assumption I.18 with  $Y(x,y) = G'(x)y - G(x)$  if  $\Delta^{-1}f \in B$  and  $G''' \in L^\infty(\mathbb{R})$ .*

■ Let us directly deal with this general case:  $X_{u,v} = \Delta(v-u) - f$  and  $Y_{u,v} = G'(x)y - G(u)$ . When  $d = 1$ , we only need the boundedness of  $G'''$  on compacts: indeed, we can bring this case to the case where  $G'''$  is bounded, by changing the definition of  $G$  out of the compact range of the solution  $u_\infty$ . 1.3.1 is straightforward. 1.3.2 is proved in corollary 3.10 of [3]. ■

## D-d Mother wavelet choice

We provide here the multiresolution satisfying  $2^\nu \geq C_\chi$  and the assumptions of theorem I.12 and property I.38. In order to restrict the support size of the used wavelet to 8 (while satisfying these criteria), we preferred the spectral factorization of the interpolating scale function (with 9 monomials) rather than the Coiflet or the Daubechies Wavelet (leading to wavelets of

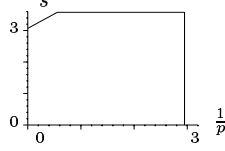


FIG. 6 – The border of a significant part of points corresponding to Besov spaces  $B_{p,p}^s(\mathbb{R})$  containing the chosen mother wavelet  $\phi$ .

respective support size 24 and 22 for similar  $C^3$  smoothness). This wavelet was used in construction (B) and in the illustrations there. We choose the following scaling filters  $h$  and  $\tilde{h}$ :

$m$	$h_m \perp_{\ell^2(\mathbb{Z})} (-1)^m \mathbb{R}_4[m]$	$m$	$\tilde{h}_m \perp_{\ell^2(\mathbb{Z})} (-1)^m \mathbb{R}_3[m]$
-4	0.10455117754191403821	-4	-0.22108234700026906300
-3	0.46828136624844826903	-3	0.135998416444055247947
-2	0.78556251203273865577	-2	1.61678343444011570751
-1	0.56288328040396058085	-1	1.30210406824836139263
0	0.10224411229042980891	0	-0.67481317893025672558
1	-0.043587990571616733725	1	-0.50474832650909841831
2	0.0076421981349174970968	2	0.37062811152289181576
3	0.012423343919207883828	3	0.046213451492551975908
		4	-0.091516020032481734685
		5	0.020432390327632570287

Figure 6 provides an insight on the regularity of  $\phi$ . In I.33, part B-b satisfies  $C_\chi = 1$  for  $m = 0.1$ .

## E Extended a priori analysis of the algorithm

### E-a Sufficient refinement for precision and truncation

**Property I.39** Steps (17b) to (17f) produce a hull  $H_{i,k} \subset \Phi$  sufficiently refined to ensure that, on the support of any  $\lambda \in H_{i,k}$ , functions  $u_i$  and  $v_{i,k}$  are equal to elements of  $V_{j_\lambda+1}$ .

■ As  $H_{i,k}$  is a hull, there exists a chain  $\nu_{j_0} \ll \nu_{j_0+1} \ll \dots \ll \nu_{j_\lambda} \ll \mu_{j_\lambda+1}$  with  $\nu_j \in \Phi_j$  and  $\nu_{j_\lambda} = \lambda$  and  $\nu_{j_\lambda+1} \notin H_{i,k}^G$ .  $A_i^{SG} \subset H_{i,k}^G$ , hence  $\nu_{j_\lambda+1} \notin A_i^{SG}$ .

Define  $k := \max\{k \geq j_0 : \nu_k \notin \Lambda_i^{SG}\}$ : as  $\Lambda_i^{SG}$  is a globe, the latter set is a connected subset of  $\mathbb{Z}$ . According to I.36, no element of  $\Lambda_i^G \cap \Phi_l$  is non-null on  $\text{supp}\nu_k$  with  $l \geq k+2$ . Hence,  $u_i$  and  $\sum_{\lambda \in \Lambda_i/j_\lambda \leq k} \langle \tilde{\lambda} | u_i \rangle \lambda$  are equal on  $\text{supp}\nu_k$ . As  $k \leq j_\lambda$ , and as  $\text{supp}\lambda \subset \text{supp}\nu_k$ , the proof for  $u_i$  is complete. A similar proof applies for  $v_{i,k}$ .  $\blacksquare$

**Definition I.40** For  $V$  a subset of  $\Phi \cup \Psi$  and  $f$  a function, let us define the seminorm  $M_V(f) := \left\| 2^{(\frac{d+2\nu}{4}-1)j_\lambda} \|\nabla f\|_{L^2(\text{supp}(\lambda))} \right\|_{\ell^4(\lambda \in V)}$  (see also [4] (2.21)) (as  $\frac{3d+4}{4} < 3 - \nu$ , we have  $\frac{d+2\nu}{4} < \frac{3d+4}{4} - \frac{1}{2} < 1$ ).

**Lemma I.41** The hull  $H_{i,k} \subset \Phi$  is sufficiently refined to ensure that the quantities  $M_{H_{i,k}}(u_i)^2$ ,  $M_{\Lambda \setminus H_{i,k}^{GA}}(u_i)^2$  and  $\|2^{-j_\lambda} \langle u_i | \lambda \rangle\|_{\ell^2(\lambda \in \Lambda \setminus H_{i,k}^{GA})}$  are  $\stackrel{N_{4,i,k}, u_i, v_{i,k}}{\lesssim} N_4^{-1} \|u_i\|_B^a \rho_1^i \rho_2^k$  for some power  $a > 0$ ; same inequalities for  $v_{i,k}$  as for  $u_i$ .

$\blacksquare$  According to steps (17b), (17d), (17e) and (17f), we have  $H_{i,k} \subset \Phi \setminus \Phi \cap V_4^{i,k} \subset \{\lambda \in \Phi \cup \Psi : j_\lambda \geq j_\mu + k_\mu\} \cup \{\lambda \in \Phi \cup \Psi : \text{supp}\lambda \cap \text{supp}\mu = \emptyset\}$ , and thus:

$$\begin{aligned} M_{H_{i,k}}(u_i) &\leq \sum_{\mu \in \Lambda_i} |\langle \tilde{\mu} | u_i \rangle| M_{H_{i,k}}(\mu) & (19b) \\ &\leq \sum_{\mu \in \Lambda_i} |\langle \tilde{\mu} | u_i \rangle| M_{\{\lambda \in \Phi : j_\lambda \geq j_\mu + k_\mu\}}(\mu) \\ &\lesssim \sum_{\mu \in \Lambda_i} |\langle \tilde{\mu} | u_i \rangle| 2^{\frac{2\nu+d}{4}j_\mu} 2^{-\frac{2-\nu}{2}k_\mu} \text{ (as } \nu < 2 \text{ and } \nabla\psi \in L^\infty) & (19c) \\ &\leq N_4^{-\frac{1}{2}} (\rho_1^i \rho_2^k)^{\frac{1}{2}} \|\langle \tilde{\mu} | u_i \rangle\|_{\ell^{2d/(2d+2-\nu)}(\mu \in \Lambda_i)} 2^{\frac{2\nu+d}{4}j_\mu} \end{aligned}$$

As  $s > \frac{3d+4}{4}$  and  $s > \frac{\nu}{2} + \frac{d}{p} + \frac{d}{4}$ , we can choose  $q > \max(\frac{2d}{2d+2-\nu}, p)$  such that  $\frac{1}{q} > \frac{2d+2-\nu}{2d} - \frac{4s-4d/p-2\nu+d}{4d} = \frac{1}{p} - \frac{4s-3d-4}{4d}$ . Therefore, Hölder gives:

$$\begin{aligned} M_{H_{i,k}}(u_i) &\lesssim N_4^{-\frac{1}{2}} (\rho_1^i \rho_2^k)^{\frac{1}{2}} \|\langle \tilde{\mu} | u_i \rangle\|_{\ell^q(\mu \in \Lambda_i)}^{2d/(2d+2-\nu)} \\ &\quad \times \|2^{-\frac{4s-4d/p-2\nu+d}{4}j_\mu}\|_{\ell^{1/[(2d+2-\nu)/2d-1/q]}(\mu \in \Lambda_i)}^{2d/(2d+2-\nu)} \\ &\lesssim N_4^{-\frac{1}{2}} (\rho_1^i \rho_2^k)^{\frac{1}{2}} \|u_i\|_B^{2d/(2d+2-\nu)} \end{aligned}$$

According to steps (17b), (17c), (17e) and (17f), we have  $\Lambda \setminus H_{i,k}^{GA} \subset \Psi \setminus \Psi \cap V_4^{i,k} \subset \{\lambda \in \Psi : j_\lambda \geq j_\mu + k_\mu\} \cup \{\lambda \in \Psi : \text{supp}\lambda \cap \text{supp}\mu = \emptyset\}$ .

As a first implication, the estimation for  $M_{H_{i,k}}(u_i)$  also applies for  $M_{\Lambda \setminus H_{i,k}^{GA}}(u_i)$ .

As a second implication:

$$\begin{aligned}
\|2^{-j\lambda}\langle\lambda|u_i\rangle\|_{\ell^2(\lambda\in\Lambda\setminus H_{i,k}^{GA})} &\leq \sum_{\mu\in\Lambda_i} |\langle\tilde{\mu}|u_i\rangle| \|2^{-j\lambda}\langle\mu|\lambda\rangle\|_{\ell^2(\lambda\in\Lambda\setminus H_{i,k}^{GA})} \\
&\leq \sum_{\mu\in\Lambda_i} |\langle\tilde{\mu}|u_i\rangle| \|2^{-j\lambda}\langle\mu|\lambda\rangle\|_{\ell^2(\lambda\in\Psi:j\lambda\geq j_\mu+k_\mu)} \\
&\lesssim \sum_{\mu\in\Lambda_i} |\langle\tilde{\mu}|u_i\rangle| 2^{-j_\mu} 2^{-2k_\mu} \text{ (as } \psi \in H^1(\mathbb{R})) \text{ (19d)} \\
&\leq C_\psi N_4^{-1} \rho_1^i \rho_2^k \|\langle\tilde{\mu}|u_i\rangle\|_{\ell^{d/(d+2)}(\mu\in\Lambda_i)} 2^{-j_\mu}
\end{aligned}$$

As  $s > \frac{3d+4}{4} > \frac{d+2}{2}$ , we can choose  $q > \max(\frac{d}{d+2}, p)$  such that  $\frac{1}{q} > \frac{d+2}{d} - \frac{1+s-d/p+d/2}{d} = \frac{1}{p} - \frac{s-(d+2)/2}{d}$ . Then:

$$\begin{aligned}
\|2^{-j\lambda}\langle\lambda|u_i\rangle\|_{\ell^2(\lambda\in\Lambda\setminus H_{i,k}^{GA})} &\lesssim C_\psi N_4^{-1} \rho_1^i \rho_2^k \|\langle\tilde{\mu}|u_i\rangle\|_{\ell^q(\mu\in\Lambda_i)} 2^{(s-d/p+d/2)j_\mu} \|\cdot\|_{\ell^{d/(d+2)}(\mu\in\Lambda_i)}^{d/(d+2)} \\
&\quad \times \|2^{-(1+s-d/p+d/2)j_\mu}\|_{\ell^{1/[1+(d+2)/d-1/q]}(\mu\in\Lambda)}^{d/(d+2)} \\
&\lesssim N_4^{-1} \rho_1^i \rho_2^k \|u_i\|_B^{d/(d+2)} \quad \blacksquare
\end{aligned}$$

**Property I.42** *The collocation error is sufficiently small so that distance between computed wavelet coefficients of  $y_{i,k}$  (at step (17g)) and the same coefficients of  $Y_{u_i, v_{i,k}}$  (the zero order nonlinear part of the formal residue of  $DF(u_i)v = F(u_i)$  at  $v = v_{i,k}$ ) is small:*

$$\|2^{-j\lambda}\langle y_{i,k} - Y_{u_i, v_{i,k}} | \lambda \rangle\|_{\ell^2(\lambda \in H_{i,k}^{GA})} \stackrel{N_{4,i,k}, u_i, v_{i,k}}{\lesssim} \frac{1}{N_4} (1 + \|u_i\|_B + \|v_{i,k}\|_B)^a \rho_1^i \rho_2^k$$

■ Let  $\lambda \in H_{i,k}$ . The coefficient of  $y_{i,k}$  corresponding to  $\lambda$  is harmed by the error  $\int Y(u(x), v(x)) \lambda(x) dx - \Gamma_\lambda(Y(u, v))$ , where  $u$  and  $v$  are the elements of  $V_{j_\lambda+1}$  equal to  $u_i$  and  $v_{i,k}$  on  $\text{supp } \lambda$  (see property I.39). Let us split  $Y(u(x), v(x))$  with this bidimensional Taylor formula of order 2:

$$\begin{aligned}
Y(u(x), v(x)) &= Y(u(y), v(y)) + DY(u(y), v(y)) \cdot \begin{pmatrix} u(x) - u(y) \\ v(x) - v(y) \end{pmatrix} + \\
&\quad \begin{pmatrix} u(x) - u(y) \\ v(x) - v(y) \end{pmatrix}^T \cdot \int_0^1 D^2 Y(tu(x) + (1-t)u(y), tv(x) + (1-t)v(y)) dt \cdot \begin{pmatrix} u(x) - u(y) \\ v(x) - v(y) \end{pmatrix}
\end{aligned}$$

The kernel  $\lambda - \Gamma_\lambda$  is orthogonal to the first line of the right hand of the equality, because it is orthogonal to  $V_{j_\lambda+1}$  (this is an assumption of the collocation  $\Gamma_\lambda$ ). Take the mean value on  $y \in \text{supp } \lambda$ , and use the Minkowski inequality

$$\left| \int Y(u(x), v(x)) \lambda(x) dx - \Gamma_\lambda(Y(u, v)) \right| \leq (N_\lambda u + N_\lambda v)^2 \|D^2 Y\|_{L^\infty} \quad (19e)$$

where  $N_\lambda f := \left[ \int_\Omega \int_\Omega |f(x) - f(y)|^2 \left( |\lambda(x)| dx + |\Gamma_\lambda|(dx) \right) \frac{\mathbb{1}_{\text{supp}\lambda}(y)}{\int_\Omega \mathbb{1}_{\text{supp}\lambda}} dy \right]^{1/2}$ .  $N_\lambda$  and  $M_{\{\lambda\}}$  restricted to the same finite dimensional quotient space  $V_{j+1}^{\text{supp}\lambda}/\mathbb{R}$  are equivalent norms. Moreover, scaling arguments show that  $2^{\frac{\nu}{2}j\lambda} N_\lambda(f) \stackrel{f,\lambda}{\sim} M_{\{\lambda\}}(f)$ . We use the Hölder inequality  $\|fg\|_{\ell^2} \leq \|f\|_{\ell^4} \|g\|_{\ell^4}$  (with a weighted measure naturally fitting this frame) to show:

$$\|2^{j\lambda} \left[ \langle Y_{u_i, v_{i,k}} | \lambda \rangle - \Gamma_\lambda(Y_{u_i, v_{i,k}}) \right] \|_{\ell^2(\lambda \in H_{i,k})} \stackrel{i,k, u_i, v_{i,k}}{\lesssim} (M_{H_{i,k}}(u_i) + M_{H_{i,k}}(v_{i,k}))^2.$$

Finish the proof with property I.33 and lemma I.41.  $\blacksquare$

**Property I.43** *The truncation to  $H_{i,k}^{GA}$  of the zero order nonlinear part  $Y_{u_i, v_{i,k}}$  of the formal residue of  $DF(u_i)v = F(u_i)$  at  $v = v_{i,k}$  results in a small  $H^{-1}$ -error:*

$$\|2^{-j\lambda} \langle Y_{u_i, v_{i,k}} | \lambda \rangle \|_{\ell^2(\lambda \in A \setminus H_{i,k}^{GA})} \stackrel{N_{4,i,k, u_i, v_{i,k}}}{\lesssim} \frac{1}{N_4} (1 + \|u_i\|_B + \|v_{i,k}\|_B)^a \rho_1^i \rho_2^k$$

$\blacksquare$  Proceed as for I.42, with  $\lambda - \Gamma_\lambda$  (where  $\lambda \in H_{i,k}$ ) replaced by  $\lambda \in A \setminus H_{i,k}^{GA}$ . Whereas  $\lambda - \Gamma_\lambda$  with  $\lambda \in H_{i,k}$  was orthogonal to  $V_{j\lambda+1}$ ,  $\lambda \in \Psi$  isn't orthogonal to  $V_{j\lambda+1}$ , so that the right hand side of (19e) contains this additional term:

$$\begin{aligned} & \left| \frac{1}{\int \mathbb{1}_{\text{supp}\lambda}} \int \int \mathbb{1}_{\text{supp}\lambda}(y) DY(u(y), v(y)) \cdot \left( \frac{u(x)-u(y)}{v(x)-v(y)} \right) \lambda(x) dx dy \right| \\ &= \left| \frac{1}{\int \mathbb{1}_{\text{supp}\lambda}} \int \int \mathbb{1}_{\text{supp}\lambda}(y) [DY(0,0) + \int_0^1 D^2 Y(tu(y), tv(y)) \cdot \left( \frac{u(y)}{v(y)} \right) dt] \cdot \left( \frac{u(x)}{v(x)} \right) \lambda(x) dx dy \right| \\ &\leq \max(DY(0,0), \|D^2 Y\|_{L^\infty(\mathbb{R}^2)}) \left[ 1 + \frac{\int \mathbb{1}_{\text{supp}\lambda} u}{\int \mathbb{1}_{\text{supp}\lambda}} + \frac{\int \mathbb{1}_{\text{supp}\lambda} v}{\int \mathbb{1}_{\text{supp}\lambda}} \right] (|\int u \lambda| + |\int v \lambda|) \end{aligned}$$

We thus obtain:

$$\begin{aligned} & \|2^{-j\lambda} \langle Y_{u_i, v_{i,k}} | \lambda \rangle \|_{\ell^2(\lambda \in A \setminus H_{i,k}^{GA})} \\ & \lesssim (M_{A \setminus H_{i,k}^{GA}}(u_i) + M_{A \setminus H_{i,k}^{GA}}(v_{i,k}))^2 + [1 + \|u_i\|_{L^\infty} + \|v_{i,k}\|_{L^\infty}] \\ & \quad \times \left( \|2^{-j\lambda} \int u_i \lambda\|_{\ell^2(v \in A \setminus H_{i,k}^{GA})} + \|2^{-j\lambda} \int v_{i,k} \lambda\|_{\ell^2(v \in A \setminus H_{i,k}^{GA})} \right) \end{aligned}$$

Finish the proof with  $\| \cdot \|_{L^\infty} \leq \| \cdot \|_B$  and lemma I.41.  $\blacksquare$

**Property I.44** *The overall error between the formal residue and the approximated residue  $r_{i,k}$  satisfies*

$$\|P(F(u_i) - DF(u_i)v_{i,k} - r_{i,k})\|_H \stackrel{i,k, u_i, v_{i,k}}{\lesssim} \frac{1}{N_4} (1 + \|u_i\|_B + \|v_{i,k}\|_B)^a \rho_1^i \rho_2^k \quad (19f)$$

■ We have  $P(F(u_i) - DF(u_i)v_{i,k} - r_{i,k}) = \omega P(X_{u_i, v_{i,k}} + Y_{u_i, v_{i,k}} - x_{i,k} - y_{i,k})$ . Use I.42 and I.43 and I.23, and assumption 1.3.2. ■

## E-b Convergence and cost : 1D and 2D

**Theorem I.45** *If  $N_1, N_3$  and  $N_4$  are greater than a constant  $C$ , and if  $\|u_0 - u_\infty\|_H \leq \frac{1}{C}$ , then  $\|u_i - u_\infty\|_H \lesssim \rho_1^i$ ,  $\|u_i\|_B \lesssim 1$  and  $(\#u_i)^{\frac{s-1}{d}} \lesssim \rho_1^{-i}$ .*

■ As our algorithm uses a Newton-like approach, the proof of convergence necessitates a careful induction, and a fine-tuning. We will construct four *independent* assumptions on  $N_1, N_3, N_4$  and  $\|u_0 - u_\infty\|_H$ , together with constants used in property  $(P_n)$  (defined in (19g)) so that  $(P_n)$  admits an induction proof. Seven lemmas will link together the six statements of  $(P_n)$ .

The assumptions on all other constants used in the algorithm ( $p, s, \rho_1, \rho_2, \kappa_1, \kappa_2, \omega, \nu$  and  $N_2$ ) are in part C-e, and we will use  $C_F$  as defined in property I.19. Assume that  $\|u_0 - u_\infty\|_H \leq \frac{1}{2\kappa_1 C_F}$  for property I.21. Let  $\kappa_2 \in (1; \rho^{-1}\rho_2)$ . Recall that  $\nu_i$  is defined by  $\nu_0 = \frac{1}{2}$  and  $\nu_{i+1} = \frac{1}{4}\rho_1^i + \nu_i^2$ , and that  $\nu_i \sim \rho_1^i$  (see I.20).

In order to maintain the objective  $\|u_i - u_\infty\|_H \leq \frac{\nu_i}{\kappa_1 C_F}$  of property I.21, assume on  $N_1$  that  $\forall i, \rho_1^i \kappa_1 (C'_{\kappa_1} N_1)^{-(2-p)/2} |u_\infty|_B^{p/2} \leq \frac{\nu_{i+1}}{\kappa_1 C_F}$ .

**Lemma 1** *There exists  $S_\nu$  such that  $\|u_i\|_B \leq S_\nu$  if  $\|u_i - u_\infty\|_H \leq \frac{\nu_i}{\kappa_1 C_F}$ .*

■ The constant  $C'_{\kappa_1}$  of theorem I.12 fits in step (17j), so that  $S_\nu := C_{\kappa_1} (\sup_i \frac{\nu_{i+1}}{\kappa_1 C_F} \rho_1^{-i} + \|u_\infty\|_B)$  fits. ■

**Lemma 2** *There exists  $S_F$  such that  $\|v_{i,\infty}\|_B \leq S_F$  if  $\|u_i\|_B \leq S_\nu$ .*

■ According to assumption 1.3.2,  $S_F := (1 + S_\nu) \sup_{u \in U} \frac{\|DF(u)^{-1}F(u)\|_B}{1 + \|u\|_B}$  fits because  $v_{i,\infty}$  is  $DF(u_i)^{-1}F(u_i)$ . ■

**Lemma 3** *There is  $S_{FF}$  fitting  $\|v_{i,0} - v_{i,\infty}\|_H \leq S_{FF} \rho_1^i$  if  $\|u_i - u_\infty\|_H \leq \frac{\nu_i}{\kappa_1 C_F}$ .*

■ Remarking that  $v_{i,0} = 0$  and  $u_0 - u_1$  was  $DF(u_0)^{-1}F(u_0)$  in the proof page 42,  $S_{FF} := \sup_i \frac{\nu_{i+1}}{\kappa_1 C_F} \rho_1^{-i} \sup_{u \in U} \|u - u_\infty\|_H^{-1} \|DF(u)^{-1}F(u)\|_H$  is finite. ■

**Lemma 4** Let  $w_{i,k} := v_{i,k} + \omega Pr_{i,k}$ . There exists  $S_\nu$  such that  $\|v_{i,k+1} - v_{i,\infty}\|_H \leq S_\nu \rho_1^i \rho_2^k$  if  $\|v_{i,\infty}\|_B \leq S_F$  and  $\|w_{i,k} - v_{i,\infty}\|_H \leq \frac{S_\nu}{\kappa_2} \rho_1^i \rho_2^k$ .

■ The constant  $C'_{\kappa_2}$  of theorem I.12 fits in step (17i), so that  $S_\nu := \max\left(\kappa_2 (C'_{\kappa_2} N_2)^{-(2-p)/2} S_F^{p/2}, S_{FF} \rho_2\right)$  fits. ■

**Lemma 5**  $\|v_{i,k+1}\|_B \leq C_{\kappa_2} (S_\nu \rho_1^i \rho_2^k + S_F)$  if  $\|v_{i,\infty}\|_B \leq S_F$  and  $\|w_{i,k} - v_{i,\infty}\|_H \leq \frac{S_\nu}{\kappa_2} \rho_1^i \rho_2^k$  (same proof).

Let  $a$  be the constant in property I.44 and let  $C_\omega$  be the constant of  $\lesssim$  in its equation (19f). Assume that  $N_4 \geq C_\omega \left(1 + S_\nu + C_{\kappa_2} (S_\nu + S_F)\right)^a \left(\frac{1}{\kappa_2} - \frac{\rho}{\rho_2}\right)^{-1} S_\nu^{-1}$ . We have:

**Lemma 6**  $\|w_{i,k} - v_{i,\infty}\|_H \leq \frac{S_\nu}{\kappa_2} \rho_1^i \rho_2^k$  if  $\|u_i\|_B \leq S_\nu$ ,  $\|v_{i,k}\|_B \leq C_{\kappa_2} (S_F + S_\nu \rho_1^i \rho_2^{k-1})$  and  $\|v_{i,k} - v_{i,\infty}\|_H \leq S_\nu \rho_1^i \rho_2^{k-1}$ .

■ I.44 and  $\|w_{i,k} - v_{i,\infty}\|_H \leq \|v_{i,k} - \omega P(F(u_i) - DF(u_i)v_{i,k}) - w_{i,k}\|_H + \|v_{i,k} - \omega P(F(u_i) - DF(u_i)v_{i,k}) - v_{i,\infty}\|_H$  ■

Finally, assume on  $N_3$  that  $S_\nu \rho_2^{N_3-1} \leq \frac{1}{4C_F \kappa_1^2}$ .

**Lemma 7**  $\|u_{i+1} - u_\infty\|_H \leq \frac{\nu_{i+1}}{\kappa_1 C_F}$  if  $\|u_i - u_\infty\|_H \leq \frac{\nu_i}{\kappa_1 C_F}$ ,  $\|v_{i,N_3} - v_{i,\infty}\|_H \leq S_\nu \rho_1^i \rho_2^{N_3-1}$  and  $N_3 \geq N_{13}$ . (proof: I.21)

The following assumption denoted  $(P_n)$  is proved by induction on  $n \in \mathbb{N}$ :

$$i \geq 0 \text{ and } (N_3+2)i \leq n \Rightarrow \begin{cases} 1 : \|u_i - u_\infty\|_H \leq \frac{\nu_i}{\kappa_1 C_F} \\ 2 : \|u_i\|_B \leq S_\nu \\ 3 : \|v_{i,\infty}\|_B \leq S_F \end{cases} \quad (19g)$$

$$i \geq 0, 0 \leq k \leq N_3 \text{ and } (N_3+2)i+k+1 \leq n \Rightarrow \begin{cases} 4 : \|v_{i,k} - v_{i,\infty}\|_H \leq S_\nu \rho_1^i \rho_2^{k-1} \\ 5 : \|v_{i,k}\|_B \leq C_{\kappa_2} (S_F + S_\nu \rho_1^i \rho_2^{k-1}) \\ 6 : \|w_{i,k} - v_{i,\infty}\|_H \leq \frac{S_\nu}{\kappa_2} \rho_1^i \rho_2^k \end{cases}$$

Case  $n=0$ :  $i=k=0$ ; 1 is given by definition of  $\nu_i$ , 2 by lemma 1, 3 by lemma 2.



Case  $n \in (N_3 + 2)\mathbb{Z}$ :  $k=0, 1$  is given by lemma 7, 2 by lemma 1, 3 by lemma 2.  
Case  $n-1 \in (N_3 + 2)\mathbb{Z}$ :  $k=0; 4$  is given by lemma 3, 5 by lemma 5, 6 by lemma 6.  
Other cases: 4 is given by lemma 4, 5 by lemma 5, 6 by lemma 6.

Now, let us prove that  $(\#u_i)^{\frac{s-1}{d}} \lesssim^i \rho_1^{-i}$ . Looking at formula  $\#_{s,p}A_\eta(f + e)^{\frac{s-1}{d}} \leq C_\kappa \|e\|_H^{-1} |f|_B$  of Theorem I.16, we have  $(\#_{s,p}u_i)^{\frac{s-1}{d}} \lesssim^i \rho_1^{-i}$ . The use of this theorem is legitimated by (P<sub>n</sub>.1), by step (17j) and because  $\|u_\infty\|_B$  is finite (we also have  $(\#_{s,p}v_{i,k})^{\frac{s-1}{d}} \lesssim^i \rho_1^{-i}$ , legitimated by (P<sub>n</sub>.4) and (P<sub>n</sub>.3), and by step (17i)). Lemma I.15 ends the proof.  $\blacksquare$

**Theorem I.46** *Computation cost and tree-cardinality of  $H_{i,k}$  are  $\lesssim^{i,k} (\rho_1^i \rho_2^k)^{\frac{-d}{s-1}}$ .*

$\blacksquare$  According to steps (17e) and (17f), we only need to prove the same for the cardinality of  $V_4^{i,k}$ . That set  $V_4^{i,k}$  is the union of  $\{\lambda \in \Phi \cup \Psi : \exists \mu \in A_i \cup A_{i,k} \text{ supp } \mu \cap \text{supp } \lambda \neq \emptyset \text{ and } j_\mu \geq j_\lambda\}$  whose size is  $\lesssim^{\#_{s,p}A_i \cup A_{i,k}}$  and of a refinement made step (17c) and (17d) whose cardinality is

$$\lesssim^{i,k} \sum_{\mu \in A_i \cup A_{i,k}} 2^{dk_\mu} \quad (19h)$$

$$\begin{aligned} &\lesssim (\rho_1^i \rho_2^k)^{-\frac{d}{2}} \|\langle u_i | \tilde{\mu} \rangle 2^{-j_\mu}\|_{\ell^{\frac{d}{d+2}}(\mu \in A_i \cup A_{i,k})}^{d/(d+2)} + (\rho_1^i \rho_2^k)^{\frac{-d}{2-\nu}} \|\langle u_i | \tilde{\mu} \rangle 2^{\frac{2\nu+d}{4}j_\mu}\|_{\ell^{\frac{2d}{2d+2-\nu}}(\mu \in A_i \cup A_{i,k})}^{2d/(2d+2-\nu)} \\ &+ (\rho_1^i \rho_2^k)^{-\frac{d}{2}} \|\langle v_{i,k} | \tilde{\mu} \rangle 2^{-j_\mu}\|_{\ell^{\frac{d}{d+2}}(\mu \in A_i \cup A_{i,k})}^{d/(d+2)} + (\rho_1^i \rho_2^k)^{\frac{-d}{2-\nu}} \|\langle v_{i,k} | \tilde{\mu} \rangle 2^{\frac{2\nu+d}{4}j_\mu}\|_{\ell^{\frac{2d}{2d+2-\nu}}(\mu \in A_i \cup A_{i,k})}^{2d/(2d+2-\nu)} \end{aligned}$$

For the  $\|\cdot\|$  norms, proceed as for (19d) and (19c). We conclude the study of cardinals with  $2 - \nu \geq s - 1$ . The cost of computation of the above set at step (17b) is  $\lesssim^{\#_{s,p}A_i \cup A_{i,k}}$ , and the cost of computation of the refinement is also  $\lesssim \sum_{\mu \in A_i \cup A_{i,k}} 2^{dk_\mu}$ .  $\blacksquare$

**Theorem I.47** *The cost to approximate  $u_\infty$  at a precision of  $N^{\frac{s-1}{d}}$  is proportional to the expected size  $N$  of a thresholded version of the exact solution  $u_\infty$ , at precision  $N^{\frac{s-1}{d}}$ , on the mere regularity assumption  $u_\infty \in B_{p,p}^s$ .*

$\blacksquare$  Thanks to last two theorems and to assumption 1.3.2, the cost of steps (17a) to (17d) is  $\lesssim^{i,k} (\rho_1^i \rho_2^k)^{\frac{-d}{s-1}}$ . Using the estimations on  $\#_{s,p}u_i$  and  $\#_{s,p}v_{i,k}$  made at end of proof of theorem I.45, and relying on the last theorem, the cost of all other steps of the algorithm presented subsection C-e is  $\lesssim^{i,k} (\rho_1^i \rho_2^k)^{\frac{-d}{s-1}}$ .  $\blacksquare$

### E-c Further dimensions

To reach the dimension 3, the keystone is the replacement of (19b) by:

$$M_a(b_{i,k}) \leq C \sum_{j \geq j_0} \|\langle b_{i,k} | \tilde{\mu} \rangle | M_a(\mu) \|_{\ell_{\tilde{\mu} \in A_i \cup A_{i,k} \cap A_j}^4}$$

We have restricted our algorithm to local elliptic cases; nevertheless, saddle-point problems are achievable, by inserting around (13) an Uzawa-type algorithm, or alternatively by inserting a negative-eigenvalues-eigenvector detector in that gradient algorithm. The use of a collocation (because of the nonlinear term) implies working in the  $L^\infty$  space: this is a limitation. To bypass the collocation for particular nonlinear terms, see [9].

### E-d Acknowledgments

The author is very grateful to his supervisor, Professor P.G. Lemarié-Rieusset for fruitful and stimulating discussions and continuous attention to this work, and for pointing essential landmarks for the proof of convergence. The author would like to acknowledge Professor A. Cohen for his enthralling discussions and lectures, and for giving all the guidelines to construct a framework of optimal rate resolution of nonlinear PDEs when they were out of public knowledge.

## Appendix

### Proof of theorem I.8, page 13

■ Define the truncation projector  $P_V g := \sum_{\lambda \in \Phi_{j_0}} \langle \tilde{\lambda} | g \rangle \lambda + \sum_{\lambda \in V} \langle \tilde{\lambda} | g \rangle \lambda$ .

**Lemma 1** ( $\Psi_\eta(g)$ 's size) : we have  $n_\eta(g) := \#\Psi_\eta(g) \leq C_1 \eta^{-p} \|g\|_{\mathbb{W}^{s,p}}^p$ .

$$\begin{aligned} \blacksquare \quad n_\eta(g) &= 2([\tfrac{1}{2}n_\eta(g)]^{\frac{1}{2}}[\tfrac{1}{2}n_\eta(g)]^{\frac{s-1}{d}})^p \leq 2(\tfrac{1}{\eta} \|P_{n_\eta(g)}^{\text{nl}} g - P_{\frac{1}{2}n_\eta(g)}^{\text{nl}} g\|_H [\tfrac{1}{2}n_\eta(g)]^{\frac{s-1}{d}})^p \\ &\leq 2(\tfrac{1}{\eta} \|g - P_{\frac{1}{2}n_\eta(g)}^{\text{nl}} g\|_H [\tfrac{1}{2}n_\eta(g)]^{\frac{s-1}{d}})^p \\ &\lesssim \eta^{-p} \|g\|_{\mathbb{W}^{s,p}}^p \quad (\text{with (6)}) \quad \blacksquare \end{aligned}$$

**Lemma 2** (truncation error on  $g$ ) :  $\|P_{\Psi_\eta(g)}(g) - g\|_H \leq C_2 \eta^{\frac{2-p}{2}} \|g\|_{\mathbb{W}^{s,p}}^{\frac{p}{2}}$ .

$$\begin{aligned}
\blacksquare \quad \|P_{\Psi_\eta(g)}(g) - g\|_H^2 &= \sum_{\lambda \in \Psi: 2^{j\lambda} |\langle \tilde{\lambda} | g \rangle| < \eta} \int_0^{2^{j\lambda} |\langle \tilde{\lambda} | g \rangle|} 2\theta d\theta = \int_0^\eta \theta (n_\theta(g) - n_\eta(g)) d\theta \\
&\leq \int_0^\eta \theta n_\theta(g) d\theta \\
&\leq C_1 \|g\|_{W^{s,p}}^p \int_0^\eta \theta^{1-p} d\theta \quad \blacksquare
\end{aligned}$$

Let  $a := \|\hat{f} - f\|_H$  and  $A := \|f\|_B$ . Decompose, then use the lemmas:

$$\begin{aligned}
\|f - P_{\Psi_\eta \hat{f}} \hat{f}\|_H &\leq \|P_{\Psi_\eta \hat{f}}(\hat{f} - f)\|_H + \|f - P_{\Psi_\eta \hat{f}} f\|_H \\
&\leq \|P_{\Psi_\eta \hat{f}}(\hat{f} - f)\|_H + \|f - P_{\Psi_{2\eta} f \cup (\Psi_\eta \hat{f} \cup \Psi \setminus \Psi_{2\eta} f)} f\|_H \\
&\leq a + [\|f - P_{\Psi_{2\eta} f} f\|_H^2 + 4\|(f - \hat{f}) - P_{\Psi_\eta \hat{f} \cup \Psi \setminus \Psi_{2\eta} f}(f - \hat{f})\|_H^2]^{\frac{1}{2}} \\
&\leq a + [C_2(2\eta)^{2-p} A^p + 4a^2]^{\frac{1}{2}}
\end{aligned}$$

Hence  $n := n_\eta \hat{f}$  with  $\eta := \frac{1}{2} (\frac{5a^2}{C_2} A^{-p})^{\frac{1}{2-p}}$  fits  $\|f - P_n^{\text{lin}} \hat{f}\|_H \leq \kappa \|\hat{f} - f\|_H$  with  $\kappa := 4$ . The estimation on  $n$  is easy:  $n_\eta(\hat{f}) \leq n_{\frac{1}{2}\eta}(f) + n_{\frac{1}{2}\eta}(\hat{f} - f) \leq C_1 \eta^{-p} A^p + a^2 (\frac{1}{2}\eta)^{-2} \leq [2^{-p} C_1 (\frac{C_2}{5})^{\frac{p}{2-p}} + 16 (\frac{C_2}{5})^{\frac{2}{2-p}}] (\frac{A}{a})^{\frac{2p}{2-p}} \sim (\frac{A}{a})^{\frac{d}{s-1}}$ .  $\blacksquare$

### Proof of lemma I.20, page 23

We first need this definition:

**Definition I.48** Let  $(\mu_{i,j})_{i,j \in \mathbb{N}}$  be the family recursively defined as follows [it is linked to I.20 with  $\nu_i = \sum_{j=0}^{2^i-2} \mu_{i,j} \rho_1^j$ ]:

$$\mu_{0,0} = \frac{1}{2}, \quad \mu_{i,j} := 0 \text{ if } j \geq 2^{i-2} \quad \text{and} \quad \sum_{j=0}^{2^i-1} \mu_{i+1,j} X^j = \frac{1}{4} X^i + \left( \sum_{j=0}^{2^i-2} \mu_{i,j} X^j \right)^2$$

**Lemma** All  $\mu_{i,j}$  are nonnegative reals,  $\mu_{i,0} = 2^{-2^{i-1}}$ ,  $\sum_{j=1}^{2^i-2} \mu_{i,j} = \frac{1}{2}$ , and  $\forall i, j \geq 0, 1 \leq j \leq i-2 \Rightarrow \mu_{i,j} \leq \pi_{i,j}$  where  $\pi_{i,j} = 2^{2^{j-1} - 2^{i-2} - 1}$ .

$\blacksquare$  A simple induction on  $i$  proves that all  $\mu_{i,j}$  are nonnegative, and that  $\mu_{i,0} = 2^{-2^{i-1}}$  for  $i \geq 0$ . Replacing  $X$  by 1 in the inductive definition of the  $\mu_{i,j}$ , we inductively get  $\sum_{j=0}^{2^i-2} \mu_{i,j} = \frac{1}{2}$ .

Now, let us prove  $1 \leq j \leq i-2 \Rightarrow \mu_{i,j} \leq \pi_{i,j}$ , with an induction on  $i \geq 3$ . For the initialisation, we prove that it is true for  $i = 3$  by direct computation:

	$i=0$	$i=1$	$i=2$	$i=3$
$j$	$\mu_{0,j}$	$\mu_{1,j}$	$\mu_{2,j}$	$\mu_{3,j}$
$j=0$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{16}$
$j=1$	0	0	$\frac{1}{4}$	$\frac{1}{8}$

	$i=0$	$i=1$	$i=2$	$i=3$
$j$	$\pi_{0,j}$	$\pi_{1,j}$	$\pi_{2,j}$	$\pi_{3,j}$
$j=0$	unused	unused	unused	unused
$j=1$	unused	unused	unused	$\frac{1}{4}$

Let  $i \geq 3$ , we will prove  $\forall j \in [1, i-1]$ ,  $\mu_{i+1, j} \leq \pi_{i+1, j}$  using  $\forall j \in [1, i-2]$ ,  $\mu_{i, j} \leq \pi_{i, j}$ .

First case, for  $j \in [2, i-2]$ :

$$\begin{aligned}
\mu_{i+1, j} &:= 2 \cdot \mu_{i, 0} \cdot \mu_{i, j} + \sum_{k=1}^{k=j-1} \mu_{i, k} \cdot \mu_{i, j-k} \\
&\leq 2^1 \cdot 2^{-2^{i-1}} \cdot 2^{2^{j-1}-2^{i-2}-1} + \sum_{k=1}^{k=j-1} 2^{2^{k-1}-2^{i-2}-1} \cdot 2^{2^{j-k-1}-2^{i-2}-1} \\
&= 2^{2^{j-1}-2^{i-1}-1} \cdot \left[ \underbrace{2^{1-2^{i-2}}}_{\leq \frac{1}{2} \text{ because } i \geq 3} + \sum_{k=1}^{k=j-1} 2^{1-2^{j-2}-1} \cdot \underbrace{2^{(2^{k-1}+2^{j-k-1})-(2^{j-2}+1)}}_{\leq 1 \text{ because } k-1, j-k-1 \in [0, j-2]} \right] \\
&\leq \pi_{i+1, j} \cdot \left[ \frac{1}{2} + \underbrace{(j-1) \cdot 2^{-2^{j-2}}}_{\leq \frac{1}{2} \text{ because } j \geq 2} \right] \\
&\leq \pi_{i+1, j} \cdot 1
\end{aligned}$$

Second case, with  $j = 1$ : same proof, the above sum being null.

Third case, with  $j = i-1$ : use the same estimations for the sum, along with  $\mu_{i, i-1} \leq \frac{1}{2}$ :

$$\begin{aligned}
\mu_{i+1, j} = \mu_{i+1, i-1} &:= 2 \cdot \mu_{i, 0} \cdot \mu_{i, i-1} + \sum_{k=1}^{k=j-1} \mu_{i, k} \cdot \mu_{i, j-k} \\
&\leq 2 \cdot 2^{-2^{i-1}} \cdot \frac{1}{2} + 2^{2^{j-1}-2^{i-1}-2} = 2^{2^{j-1}-2^{i-1}-1} \left[ \underbrace{2^{1-2^{j-1}} + \frac{1}{2}}_{\leq 1 \text{ because } j \geq 2} \right]
\end{aligned}$$

Then comes the proof of lemma I.20:

■ The above lemma gives

$$\begin{aligned}
\nu_i &= \mu_{i, 0} + \sum_{j=1}^{i-2} \mu_{i, j} \rho_1^j + \sum_{j=i-1}^{2^{i-2}} \mu_{i, j} \rho_1^j \\
&\leq 2^{-2^{i-1}} + \sum_{j=1}^{i-2} 2^{2^{j-1}-2^{i-2}-1} \rho_1 + \sum_{j=i-1}^{2^{i-2}} \mu_{i, j} \rho_1^{i-1} \\
&\leq 2^{-2^{i-1}} + 2^{-2^{i-3}-1} \rho_1 \sum_{j=1}^{i-2} 2^{2^{j-1}-2^{i-3}} + \rho_1^{i-1} \sum_{j=1}^{2^{i-2}} \mu_{i, j}
\end{aligned}$$

The right sum is equal to  $\frac{1}{2}$ , and the left sum becomes:

$$\sum_{j=1}^{i-2} 2^{2^{j-1}-2^{i-3}} = 1 + \sum_{j=1}^{i-3} 2^{2^{j-1}-2^{i-3}} \leq 1 + \sum_{j=1}^{i-3} 2^{-2^{i-4}} = 1 + (i-3)2^{-2^{i-4}} \leq \frac{3}{2}$$

### Proof of lemma I.36, page 28

We need the following property, stating that two scale functions of same level whose pointwise product isn't null are the "daughters" of a common scale function:

**Property** *Let  $j > j_0$  and  $\lambda, \mu \in \Phi_j$ . If  $\exists x \in \Omega$   $\lambda(x)\mu(x) \neq 0$  then  $\exists \nu \in \Phi_{j-1}$ ,  $\nu \prec \lambda$  and  $\nu \prec \mu$ . This property is true for classical wavelet basis.*

■ Let  $V$  be a globe. Let  $\lambda \in V^{SH} \cap \Phi_j$ ,  $\mu'' \in V^S \cap \Phi_{j+2}$  and  $x \in \Omega$ . Assume for the sake of contradiction that  $\lambda(x)\mu''(x) \neq 0$ . As  $\lambda$  belongs to the hull  $V^{SH}$ , let  $\lambda'$  belong to  $\Phi \setminus V^S$  with  $\lambda \prec \lambda'$ . As  $\mu''$  belongs to the overglobe  $V^S \cap$ , let  $\nu'' \in V \cap \Phi_{j+2}$ ,  $\mu' \prec \mu''$ ,  $\nu' \prec \nu''$  and  $x' \in \Omega$  such that  $\mu'(x')\nu'(x') \neq 0$ . According to the upper property, let  $\mu$  be such that  $\mu \prec \mu'$  and  $\mu \prec \nu'$ .  $\mu \prec \nu' \prec \nu'' \in V \cap \Phi_{j+2}$ , hence  $\mu' \in V$  and  $\mu \in V$ .  $x$  is inside interior of  $\text{supp}\mu''$ , and  $\mu'' \prec \mu' \prec \mu$ , hence  $x$  is inside interior of  $\text{supp}\mu$ .  $\lambda' \succ \lambda \in \Phi_j$ , hence  $\lambda' \in \Phi_{j+1}$ .  $\lambda \prec \lambda'$ ,  $\mu \prec \mu'$ ,  $\lambda(x) \neq 0$ ,  $\mu(x) \neq 0$  and  $\mu \in V$ , hence  $\lambda' \in V^S$ : this is absurd. ■

## Références

- [1] R. G. Baraniuk, R. DeVore, G. Kyriazis, and X. M. Yu. Near best tree approximation. *Advances in Computational Mathematics*, 1(1):357–373, 2002.
- [2] A. Cohen. Wavelet methods in numerical analysis. In *Handbook of numerical analysis*, volume 7 (part 3), pages 417–711. Elsevier, Amsterdam, 2000.
- [3] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods for elliptic operator equations: Convergence rates. *Math. Comput.*, 70(233):27–75, 2001.
- [4] A. Cohen, W. Dahmen, and R. DeVore. Sparse multiscale evaluation of nonlinear composition of functions. *SIAM J. Math. Anal.*, 35:279–303, 2003.
- [5] A. Cohen, R. DeVore, and R. Hochmuth. Restricted nonlinear approximation. *Constructive Approximation*, 16(1):85–113, 2000.
- [6] A. Cohen, R. DeVore, G. Kerkycharian, and D. Picard. Maximal spaces with given rate of convergence for thresholding algorithms. *Appl. Comput. Harmon. Anal.*, 11(2):167–191, 2001.
- [7] A. Cohen, S. M. Kaber, S. Müller, and M. Postel. Fully adaptive multi-resolution finite volume schemes for conservation laws. *Math. Comput.*, 72(241):183–225, 2003.
- [8] A. Cohen and R. Masson. Wavelet methods for second-order elliptic problems, preconditioning, and adaptivity. *SIAM J. Sci. Comput.*, 21(3):1006–1026, 1999.
- [9] W. Dahmen and C. A. Micchelli. Using the refinement equation for evaluating integrals of wavelets. *SIAM J. Numer. Anal.*, 30(2):507–537, 1993.
- [10] W. Dahmen, R. Schneider, and Y. Xu. Nonlinear functionals of wavelet expansions – adaptive reconstruction and fast evaluation. *Numer. Math.*, 86(1):49–101, 2000.
- [11] R. A. DeVore. Nonlinear approximation. *Acta Num.*, 7:51–150, 1998.
- [12] W. Hackbusch. Comparison of different multi-grid variants for nonlinear equations. *Z. Angew. Math. Mech.*, 72(2):148–151, 1992.
- [13] A. Jouini and P.-G. Lemarié-Rieusset. Analyse multi-résolution bi-orthogonale sur l’intervalle et applications. (Biorthogonal multiresolution analysis on the interval and applications). *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, 10(4):453–476, 1993.

- [14] T. Runst and W. Sickel. Sobolev spaces of fractional order, Nemytskij operators and nonlinear partial differential equations. In *de Gruyter Series in Nonlinear Analysis and Applications*, volume 3. de Gruyter, Berlin, 1996.

## II Un problème intégral de Helmholtz

### 1 Introduction à l'équation de Helmholtz

Le choix du problème oscillatoire bidimensionnel de cette thèse a été motivé par l'aspect curieux que prennent certaines matrices de systèmes numériques, lors de la résolution numérique du problème intégral oscillatoire unidimensionnel qui lui correspond. Ici nous choisissons le problème intégral de Helmholtz, en dimension deux.

Bradie-Coifman-Grossman [BCG93] a proposé une ébauche de la propriété 5.2. C'était un indice que certaines bases à bonne localisation spatiale et fréquentielle, à savoir les bases de Malvar et les bases de Fourier, pouvaient fournir un espace de discrétisation où la matrice, normalement pleine pour un problème intégral, était assez bien approximable par une matrice creuse. La quantification de cet adjectif « approximable » dans la phrase précédente nous a amené, via le théorème 5.1, à l'analyse de l'opérateur à discrétiser, à l'analyse de l'efficacité de la discrétisation, et de son inversibilité, toutes ces analyses devant être faites à haute fréquence, c'est à dire uniformément en une longueur d'onde  $\lambda_0$  tendant vers 0. Ainsi, des résultats nouveaux, propriétés 2.27 et 2.28 ont été démontrées dans un domaine d'analyse dure.

Un avantage de la formulation intégrale, sur la formulation conventionnelle du problème intégral de Helmholtz, est que la matrice de rigidité des systèmes rencontrés dans la résolution numérique de la formulation intégrale est de dimension beaucoup plus petite que celle de la formulation conventionnelle. En dimension deux, c'est une matrice de taille en  $O(\frac{1}{\lambda_0} \times \frac{1}{\lambda_0})$  au lieu d'une matrice en  $O(\frac{1}{\lambda_0^2} \times \frac{1}{\lambda_0^2})$ , où  $\lambda_0 > 0$  est la longueur d'onde du problème de Helmholtz. Lorsque la longueur d'onde  $\lambda_0$  est petite, son inverse  $\frac{1}{\lambda_0}$  est un bon étalon de la complexité, car le pas de discrétisation  $h$  est le plus souvent pris égal à  $\frac{\lambda_0}{10}$ .

En revanche, ces matrices sont pleines au lieu d'être creuses.

L'aspect curieux des matrices de rigidité des systèmes rencontrés dans la résolution numérique du problème intégral de Helmholtz consiste en des segments plus ou moins courbes, dans la matrice, le long desquels se regroupent les coefficients les plus gros ; les autres coefficients sont plus petits de plusieurs ordres de grandeur.

Lorsque une matrice d'un système présente des coefficients plus gros que



d'autres de plusieurs ordres de grandeur, on a avantage à utiliser un produit matrice-vecteur le plus efficace possible. Un tel produit peut être plus efficace dans la mesure où il ne prend pas en compte les petits coefficients de la matrice : même si une erreur en résulte, le produit matrice-vecteur est grandement accéléré (de plusieurs ordres de grandeur).

Afin de donner du sens à une étude mathématique sur ces regroupements de gros coefficients, il a fallu faire une étude asymptotique, en  $\lambda_0$  tendant vers 0 : en effet, la meilleure façon d'écrire des assertions empiriques comme

- « Lorsque la longueur d'onde vaut  $\lambda_0 = \frac{1}{10}$ , 5% des coefficients de la matrice suffisent pour faire un produit matrice-vecteur avec une erreur relative de 0.2. »
- « Lorsque la longueur d'onde vaut  $\lambda_0 = \frac{1}{10}$ , les coefficients hors des segments d'épaisseur 40 sont plus petit d'un facteur au moins 50. »

sous forme mathématiquement rigoureuse sont les assertions asymptotiques suivantes :

- « Lorsque  $\lambda_0 \rightarrow 0$ , il suffit d'avoir  $C \frac{1}{\varepsilon} \lambda_0^{1.6}$  coefficients pour obtenir un produit matrice-vecteur avec une erreur relative de  $\varepsilon$  ; des expérimentations numériques ont montré que  $C \approx 9$  sur un exemple. »
- « Les coefficients dont le couple de fréquences est à une distance d'au moins  $e \lambda_0^{1/2}$  sont plus petits que  $C e^{-2}$  ; des expérimentations numériques ont montré que  $C \approx 50$  sur un exemple. »

Dans ces deux assertions,  $C$  est une constante indépendante de  $\lambda_0$  et de  $e$  ; ces assertions sont intéressantes parce que  $C$  est indépendant de  $\lambda_0$ . Sans cette indépendance, ces assertions ne seraient pas utilisables numériquement.

Un ingrédient original (achevé) nécessaire pour de telles assertions (inachevée) a été l'étude du nombre de conditionnement de cette matrice, et plus encore de la norme de son inverse, et également l'erreur d'approximation faite.

Le problème de Helmholtz consiste à résoudre, à l'extérieur d'un ensemble au bord connexe  $S$  à bord régulier par morceaux, l'équation de Helmholtz, c'est à dire qu'on recherche une fonction  $E : {}^c S \rightarrow \mathbb{C}$  à valeurs complexe, définie sur le complémentaire  ${}^c S$  et vérifiant, sur  ${}^c S$ , l'équation

$$\Delta E + \frac{4\pi^2}{\lambda_0^2} E = 0 \quad \text{sur } {}^c S, \quad (20)$$

dite *équation de Helmholtz*.  $S$  est un sous-ensemble de  $\mathbb{R}^2$  qu'on appelle *diffRACTANT*, en référence aux diffractions qu'il occasionne dans l'analogie de l'équation des ondes. La longueur d'onde  $\lambda_0$  est un paramètre du problème.

Nous ajoutons deux conditions aux limites :

– La condition à l’infini de *rayonnement de Sommerfeld*

$$\lim_{\|x\| \rightarrow \infty} r^{\frac{1}{2}} \left( \frac{\partial E}{\partial r}(x) - \frac{2i\pi}{\lambda_0} E(x) \right) = 0, \quad (21a)$$

qui bloque, en quelle que sorte, la création d’énergie à l’infini. Ici  $r = \|x\|$  est le rayon, et la limite doit être uniforme en  $\frac{x}{\|x\|}$ .

– La condition au bord de Dirichlet non homogène

$$E(x) = -E^{inc}(x), \quad \forall x \in \partial S \quad (21b)$$

avec  $E^{inc}$  le second membre, et  $\partial S$  le bord de  $S$ .

Le mot-clef *inc* en exposant du second membre est une référence à l’équation des ondes, où  $E^{inc}$  est l’onde incidente, et où (21b) est la nullité du champ total sur le bord du diffractant  $S$  qui dans ce contexte diffracte le champ incident en créant le champ supplémentaire  $E$ .

Résumons le problème de Helmholtz :

$$\begin{cases} \Delta E(x) + \frac{4\pi^2}{\lambda_0^2} E(x) = 0 & x \in {}^c S \\ E(x) = -E^{inc}(x) & x \in \partial S \\ r^{\frac{1}{2}} \left( \frac{\partial E}{\partial r}(x) - \frac{2i\pi}{\lambda_0} E(x) \right) \rightarrow 0 & \|x\| \rightarrow \infty \end{cases} \quad (21c)$$

L’inconnue,  $E$ , est une fonction  ${}^c S \rightarrow \mathbb{C}$ . Le théorème 2.14 montrera l’unicité ; l’existence est aussi vraie, elle est citée sans démonstration par la remarque 2.18 et nous ne l’utiliserons pas ; cette existence est habituellement prouvée par l’existence pour le problème intégral simple couche de première espèce, comme Gauss l’a fait le premier.

Afin d’obtenir une équation équivalente mais avec une autre fonction  $J$  qui ait un ensemble de définition,  $\partial S$ , qui est plus compact que  ${}^c S$ , nous avons besoin d’un opérateur linéaire de prolongement,  $P_S$  :

$$P_S : E \mapsto P_S(E)$$

où  $P_S(E)$  est une fonction  $\mathbb{R}^2 \rightarrow \mathbb{C}$  qui vérifie  $P_S E(x) = E(x)$  pour  $x \notin S$ , qui vérifie l’équation de Helmholtz

$$\Delta E + \frac{4\pi^2}{\lambda_0^2} E = 0 \quad \text{sur } \overset{\circ}{S}, \quad (21d)$$

dans l’intérieur de  $S$ , et qui est continue (il y a recollement des valeurs au bord).

Attention, pour certains couples  $(S, \lambda_0)$ , il se peut qu'un phénomène connu sous le nom de *mode propre* fasse que  $P_S$  n'est pas uniquement défini (voir pas du tout défini) par les contraintes que nous avons donné ci-dessus. C'est pourquoi nous ne parlerons de cet opérateur  $P_S$  que dans les introductions, jamais pour énoncer un théorème.

Les deux théorèmes 2.13 et 2.17 de représentations interne et externe montrent que l'image de  $P_S$  est incluse dans l'ensemble des fonctions  $\mathbb{R}^2 \rightarrow \mathbb{C}$  pouvant se factoriser sous la forme

$$(G \star J)(x) = \int_0^L G(x - s(t))J(t)dt, \quad x \in \mathbb{R}^2 \quad (21e)$$

où  $s(t)$  est pour  $t \in [0, L]$  une paramétrisation par abscisse curviligne du bord  $\partial S$ , où  $G$  (définie en (28) en utilisant la fonction de Hankel  $H_0^{ank}$ ) est la solution fondamentale de l'équation de Helmholtz

$$\Delta E + \frac{4\pi^2}{\lambda_0^2} E = 0 \quad \text{sur } \mathbb{R}^2, \quad (21f)$$

vérifiant la condition à l'infini (21a), et où  $J$  est une fonction  $\partial S \rightarrow \mathbb{C}$ . Afin d'alléger les formules, la fonction  $J \circ s : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{C}$  sera aussi notée  $J$  sans que cela pose de problème d'ambiguïté.

Le remplacement, naturel, de  $E$  par  $G \star J$  dans la condition de Dirichlet (21b) mène à une équation dont  $J$  est l'inconnue; malheureusement, le même phénomène dit de *mode propre* cause parfois la non-injectivité de cette équation.

Nous arrivons donc au problème qui sera numériquement discrétisé. Au lieu de la condition de Dirichlet (21b), nous allons utiliser une condition dite de Robin (ou Leontovitch) qui est une condition d'impédance :

$$\frac{i\alpha}{\lambda_0} P_S(E)(x) + 2\nu_1^i P_S(E)(x) = L_\alpha(x) \quad x \in \partial S \quad (21g)$$

où  $i$  est le nombre complexe habituel  $\sqrt{-1}$ , où  $\alpha$  est un paramètre réel non nul, et où  $\nu_1^i P_S(E)$  est la dérivée normale intérieure de  $P_S(E)$ . Notons que pour que la solution  $E$  de (21g) respecte la condition (21b), il faut que

$$L_\alpha = -\frac{i\alpha}{\lambda_0} E^{inc} - 2\nu_1^i E^{inc}.$$

Ainsi, nous discrétiserons numériquement le problème suivant :

$$\frac{i\alpha}{\lambda_0} (G \star J)(x) + 2\nu_1^i (G \star J)(x) = L_\alpha(x) \quad x \in \partial S \quad (21h)$$

où  $J$  est une fonction  $\partial S \rightarrow \mathbb{C}$ ; d'après le théorème 2.30, ce problème admettra une unique solution  $J \in L^2(\partial S)$ , et de plus  $E = J \star G$  sera l'unique solution du problème non intégral (21c). Notez que nous avons besoin de la dérivée normale  $\nu_1^i E^{inc}$  de  $E^{inc}$  pour définir le second membre ci-dessus, donc le problème intégral (21h) ne permet de résoudre le problème non-intégral (21c) qu'après avoir calculé  $E^{inc}$  qui satisfasse (20), et ensuite en avoir calculé  $-\nu_1 E^{inc}$ ; cette étape intermédiaire aurait pu être évitée par le remplacement du problème (21h) par son adjoint.

D'après la propriété 2.23, ce problème (21h) admet cette forme développée :

$$J(t) + \int_0^L K(t,t')J(t')dt' = L_\alpha(t) \quad t \in [0,L] \quad (22)$$

où le noyau  $K$  de cet opérateur (qui est un *opérateur de Fredholm de seconde espèce*) vaut

$$K(t,t') := \begin{cases} 0 & \text{si } t = t' \\ \frac{i\alpha}{\lambda_0} G(s(t) - s(t')) - 2\langle s'(t)^\perp | \nabla G \rangle(x - s(t'))|_{x=s(t)} & \text{sinon} \end{cases} \quad (23)$$

Une fois la fonction  $G$  définie à l'aide de la fonction de Hankel, le lecteur trouvera une autre expression de  $K$  plus connue en (51).

Un exemple de discrétisation numérique dans un espace vectoriel de dimension finie  $V_h$  du problème (21h) consiste à chercher  $J_h \in V_h$  tel que

$$\int_0^L I_h(t)J_h(t)dt + \int_0^L \int_0^L K(t,t')I_h(t')J_h(t)dt'dt = \int_0^L L_\alpha(t)I_h(t)dt \quad \forall I_h \in V_h \quad (24)$$

Notons que cette dernière équation contient un opérateur de Fredholm de seconde espèce, et que  $G$  et  $L_\alpha$  dépendent de  $\lambda_0$ .

Par exemple, on peut prendre pour  $V_h$  l'ensemble des fonctions continues et affines par morceaux correspondant à une partition régulière de  $\partial S$ .

La difficulté principale réside dans le fait qu'une fonction de cette base, représentant un courant « local », rayonne dans tout l'espace environnant et interagit donc avec tous les autres éléments de la base. Cette interaction très forte conduit à des matrices complexes pleines. Nous verrons cependant que cette interaction est limitée à certaines bandes précises de fréquences et angulaires, ce qui permettra avec une base fréquentielle telle que les paquets d'ondelettes d'obtenir des matrices de rigidité dont beaucoup de termes seront négligeables.

L'objectif numérique habituel pour le problème (21c) est le calcul de la section radar, qui est une fonction dépendant de deux variables  $\beta$  et  $\gamma$  qui vaut

$$\int_0^L \overline{e_\gamma(s(t))} (J_\beta \star G)(s(t)) dt \quad (25)$$

avec  $J_\beta$  la solution du problème (21h) lorsque  $E^{inc}$  est l'onde plane

$$x \mapsto e^{\frac{2i\pi}{\lambda_0} x_1 \cos \beta + x_2 \sin \beta}$$

arrivant avec l'angle  $\beta$ , et avec  $e_\gamma$  qui vaut l'onde plane arrivant avec l'angle  $\gamma$ .

Avant de faire varier  $\lambda_0$ , nous avons réuni dans la partie 2 de multiples résultats sur le problème de Helmholtz (21c) en deux dimensions : ce problème a fait l'objet de moins d'intérêt que son homologue en trois dimensions, et donc la partie 2 qui établit l'équivalence entre le problème (21c) et le problème intégral (21h) rassemble des résultats pour la plupart publiés par différentes personnes ; il a paru utile de les rassembler ici. Nous avons dû adapter certains résultats sur l'opérateur de Laplace au problème de Helmholtz.

Ensuite, nous avons réuni dans la partie 3 les résultats existants essentiellement utilisés pour valider la solution du problème discrétisé (24). ainsi qu'un état de l'art sur l'utilisation du problème numérique (24).

La partie 4 regroupe des estimations sur l'inversibilité du problème, l'inversibilité de son équivalent numérique, ainsi qu'un argument mathématique qui justifie, pour la première fois en dimension deux, la phrase « nous discrétisons le problème en mettant dix degrés de liberté par longueur d'onde » si courante dans le contexte des équations oscillatoires comme l'équation de Helmholtz. Notre preuve est faite pour le problème intégral (21h), sujet de cette thèse, dans le cas où le bord du diffractant  $S$  est régulier ; elle suppose que la plus petite valeur singulière de l'opérateur (21h) ne tend pas vers 0 lorsque  $\lambda_0$  tend vers 0.

La partie 5 présente les résultats numériques en ondelettes, avec calcul des grandeurs ayant fait l'objet d'estimations dans la troisième partie. L'application numérique est faite pour un diffractant  $S$  sans coin, représenté figure 23. Nous montrons également que lorsque le diffractant présente des coins et une cavité, les estimations de la troisième partie ne sont pas numériquement possibles. Nous montrons également dans cette partie 5 qu'il est possible de seuller la matrice à l'aveugle.

Les résultats de la troisième et de la cinquième partie sont originaux.

## 2 L'équivalence entre le problème de Helmholtz et le problème intégral de Helmholtz

Nous allons légitimer l'équivalence des problèmes (21c) et (21h). Lorsque le diffractant ne présente pas de coin, il suffit presque de regrouper les résultats de Nédélec [N01] et Dautray-Lions [DL85]. Par complétude nous fournissons ici cette équivalence lorsque le diffractant présente éventuellement des coins : en effet, dans la sous-section 3.2, nous verrons que le problème numérique (24) est justifié pour  $h$  suffisamment petit, même si ce  $h$  doit être en  $O(\lambda_0^{3/2})$ .

Pour expliquer par la table page 195 que ces coins sont gênants numériquement, nous devons montrer que cette équivalence tient pour un diffractant avec coins. Reprenons, dans ce chapitre, l'analyse du problème numérique. Nous montrons par exemple qu'il admet une solution unique, et que cette solution a une interprétation solide. Pour cela, nous utilisons beaucoup de résultats publiés, et aussi quelques démonstrations originales comme celles de la propriété 2.19, de la remarque 2.22, de la propriété 2.23 et du lemme 2.24. Grosso modo, nous travaillerons avec la même démarche que Costabel et Stephan [CS85], que nous devons adapter de l'opérateur de Laplace à l'opérateur de Helmholtz, et à un diffractant pouvant présenter des parties courbes (comme fait par Costabel-Dauge [CD93]).

### 2.1 Le problème extérieur de Helmholtz

Nous allons construire le problème (21c) en fournissant toutes les justifications nécessaires.

#### 2.1.1 L'opérateur de Helmholtz

Dans cette section, nous présentons l'opérateur de Helmholtz et l'équation de Helmholtz que nous utiliserons par la suite. Nous présentons également plusieurs solutions de l'équation de Helmholtz.

L'ingrédient principal pour définir l'opérateur de Helmholtz est l'opérateur de Laplace  $\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$  en deux dimensions. L'opérateur de Helmholtz est

$$\Delta + \frac{4\pi^2}{\lambda_0^2} \text{Id.} \quad (26)$$

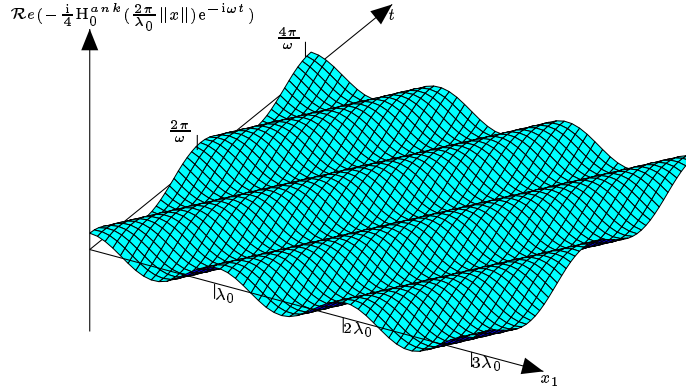


FIG. 7 – Onde plane : graphe de sa partie réelle, en fonction de  $x_1$  et de  $t$ .

La longueur d'onde  $\lambda_0$  est la grandeur caractéristique qui se retrouve dans les solutions de l'équation de Helmholtz (21f). Par commodité, nous supposons une fois pour toutes que

$$\lambda_0 < \frac{1}{2} \quad (27a)$$

afin de pouvoir, par exemple, écrire  $1 = O(|\log \lambda_0|)$  ce qui simplifiera un certain nombre d'expressions.

La longueur d'onde  $\lambda_0$  se retrouve par exemple dans la solution

$$e^{\frac{2i\pi}{\lambda_0}(\vec{e}_1|x) - i\omega t} \quad (27b)$$

qu'on appelle *onde plane* et qui est paramétrée par le temps  $t$ . Cette onde se propage dans la direction du vecteur unitaire  $\vec{e}_1$  à la vitesse  $\frac{\lambda_0 \omega}{2\pi}$ , et  $i$  est le nombre complexe  $\sqrt{-1}$ . Dans la direction  $\vec{e}_1$ , cette onde plane est périodique de période  $\lambda_0$  en  $x$ . Cette onde plane, à  $t$  fixé, est solution de l'équation de Helmholtz (21f). La figure 7 présente le graphe de la partie réelle de l'onde plane, en fonction de  $x_1$  et de  $t$  (pas de dépendance en  $x_2$ ). En suivant les lignes de niveaux (non tracées) de ce graphe, nous percevons que cette onde se propage dans la direction  $x_1$  croissant lorsque le temps  $t$  progresse.

L'onde plane est aussi une solution de l'équation des ondes

$$\Delta E - \frac{4\pi^2}{\lambda_0^2 \omega^2} \frac{\partial^2}{\partial t^2} E = 0 \quad \text{sur } \mathbb{R}^2.$$

Comme nous nous attachons aux difficultés de l'opérateur de Helmholtz liées à un domaine (non borné) avec cavité, la dépendance au temps n'est pas significative. Avec une hypothèse (classique) de dépendance en temps  $e^{-i\omega t}$ , l'équation des ondes devient l'équation de Helmholtz (21f).

L'équation de Helmholtz admet d'autres solutions, sous forme radiale. Pour les expliciter, nous utiliserons la fonction de Hankel  $H_0^{ank}$ ; la définition de cette fonction se trouve dans la démonstration qui suit. Nous recherchons d'autres solutions qui vérifient l'équation de Helmholtz (21f) partout dans  $\mathbb{R}^2$  sauf en l'origine. Plus précisément, nous remplaçons le second membre de (21f) par  $\delta$ , un Dirac placé en l'origine. Les solutions ainsi obtenues sont usuellement appelées *solutions fondamentales* de l'opérateur de Helmholtz; la convolution est leur principal usage, que nous verrons par exemple dans le théorème de représentation 2.13.

**Propriété 2.1** *Les candidats invariants par rotation pour être solution fondamentale de l'opérateur de Helmholtz (21f) sont soit*

$$G(x) = -\frac{i}{4} H_0^{ank} \left( \frac{2\pi}{\lambda_0} \|x\| \right) \quad x \in \mathbb{R}^2, \quad (28)$$

soit

$$\overline{G(x)} = \frac{i}{4} \overline{H_0^{ank} \left( \frac{2\pi}{\lambda_0} \|x\| \right)} \quad x \in \mathbb{R}^2, \quad (29a)$$

soit une combinaison affine des deux;  $H_0^{ank}$  est la fonction de Hankel de premier type et de paramètre nul.

Pour la démonstration de cette propriété, nous aurons besoin du lemme 2.2 qui suit, qui présente la solution fondamentale de l'opérateur de Laplace.

■ Recherchons une solution fondamentale de l'équation de Helmholtz (21f)  $E = E(x,t) = f(\|x\|)$  sous la forme radiale, et telle que

$$\Delta E + \frac{4\pi^2}{\lambda_0^2} E = \delta(x)$$

Pour un rayon  $r$  strictement positif, la formule du Laplacien en coordonnées cylindriques donne

$$f''(r) + \frac{1}{r} f'(r) + \frac{4\pi^2}{\lambda_0^2} f(r) = 0. \quad (29b)$$



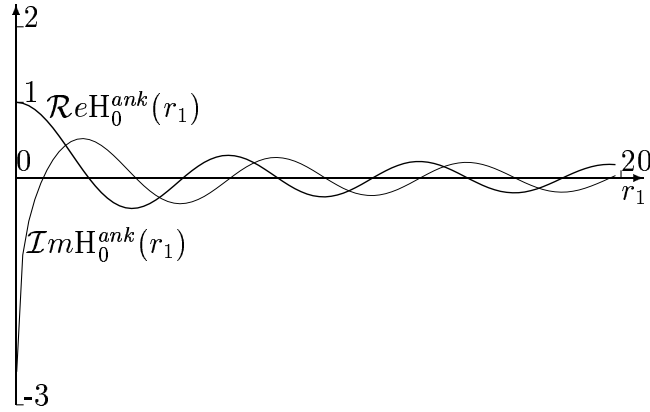
Avec le changement de variable  $r_1 = \frac{\lambda_0}{2\pi} r$  (et le changement d'inconnue  $f_1(r_1) = f(\frac{\lambda_0}{2\pi} r_1) = f(r)$ ), nous avons alors

$$\frac{4\pi^2}{\lambda_0^2} f_1''(r_1) + \frac{2\pi}{\lambda_0} \frac{1}{r_1} \frac{2\pi}{\lambda_0} f_1'(r_1) + \frac{4\pi^2}{\lambda_0^2} f_1(r_1) = 0$$

ce qui, en simplifiant, donne une équation de Bessel

$$f_1''(r_1) + \frac{1}{r_1} f_1'(r_1) + f_1(r_1) = 0 \quad (29c)$$

et les solutions de cette équation sont les combinaisons linéaires de la partie réelle et de la partie imaginaire de la fonction de Hankel  $H_0^{ank}(r)$ , dont voici les représentations graphiques (celle de la partie réelle est en gras) :



La partie réelle de  $H_0^{ank}$ , souvent notée  $J_0$ , est une fonction cylindrique de Bessel ; la partie imaginaire, souvent notée  $Y_0$  (ou encore  $N_0$ ) est une fonction cylindrique de Neumann. La fonction  $H_0^{ank}$  et la fonction conjuguée  $\overline{H_0^{ank}}$  sont usuellement notés  $H_0^1$  et  $H_0^2$  respectivement : ce sont les fonctions de Hankel de premier et second type. Nous n'utiliserons pas toutes ces notations, car  $H^1$  et  $H_0^1$  sont aussi deux notations usuelles pour des espaces fonctionnels de Sobolev que nous allons beaucoup utiliser.

La seule notation que nous conserverons de cette démonstration sera  $H_0^{ank}$ .

Kress [K89], p180, et Watson [W22], p73 (1) p64 (2) p62 (3) p15 (1) p60 note \*, donnent un développement limité au voisinage de 0 de cette fonction :

$$H_0^{ank}(t) = \frac{2i}{\pi} \ln t + 1 - 2i \frac{\ln 2 - C}{\pi} + O(t^2 \ln t) \quad (29d)$$

Ici,  $C \approx 0.57721$  est la constante d'Euler. De plus, Watson [W22], p74 (6), donne le développement limité suivant, avec  $C$  une (autre) constante complexe :

$$H_0^{ank'}(t) = \frac{2i}{\pi} \frac{1}{t} - \frac{i}{\pi} t \ln t - Ct + O(t^3 \ln t), \quad (29e)$$

Nous avons donc montré comme condition nécessaire que la solution fondamentale est une combinaison linéaire de (28) et de (29a). Cette condition nécessaire assure l'annulation de  $\Delta E + \frac{4\pi^2}{\lambda_0^2} E$  sur  $\mathbb{R}^2 \setminus \{0\}$ . Le développement limité (29d) nous permet de regarder ce qui se passe en 0. Comme, d'après le lemme suivant,  $\frac{1}{2\pi} \log$  est la solution fondamentale de Laplace, nous voyons que, de ces combinaisons linéaires de (28) et de (29a), seule les combinaisons affines (c'est à dire barycentriques) subsistent.

Il faut maintenant montrer qu'une combinaison affine  $E$  de (28) et de (29a) est effectivement solution de l'équation

$$\Delta E + \frac{4\pi^2}{\lambda_0^2} E = \delta \quad \text{sur } \mathbb{R}^2.$$

Nous avons vu ci-dessus que le membre gauche de cette équation a pour support l'origine 0 de  $\mathbb{R}^2$ . Soit  $w$  une fonction  $C^\infty(\mathbb{R}^2)$  à support compact qui vaut 1 au voisinage de l'origine. Il suffit de prouver, pour  $\varepsilon > 0$ , que

$$w\left(\frac{x}{\varepsilon}\right)\Delta E(x) + w\left(\frac{x}{\varepsilon}\right)\frac{4\pi^2}{\lambda_0^2}E(x) = \delta(x) \quad x \in \mathbb{R}^2. \quad (29f)$$

Or le membre gauche de cette équation ne dépend en fait pas de  $\varepsilon$  : donc il suffit de montrer que

$$\lim_{\varepsilon \rightarrow 0} \left( w\left(\frac{x}{\varepsilon}\right)\Delta E(x) + w\left(\frac{x}{\varepsilon}\right)\frac{4\pi^2}{\lambda_0^2}E(x) \right) = \delta(x) \quad x \in \mathbb{R}^2. \quad (29g)$$

avec  $\lim$  étant ici la convergence dans l'espace des distributions. Les développements limités ci-dessus décomposent  $E$  de la façon suivante :

$$E(x) = \frac{1}{2\pi} \log \|x\| + A(x)$$

où  $A$  est une fonction qui est  $C^1$ . Vu le lemme (2.2), il suffit alors de montrer que

$$\lim_{\varepsilon \rightarrow 0} \left( w\left(\frac{x}{\varepsilon}\right)\Delta A(x) \right) = 0 \quad x \in \mathbb{R}^2. \quad (29h)$$

et que

$$\lim_{\varepsilon \rightarrow 0} \left( w\left(\frac{x}{\varepsilon}\right)\frac{4\pi^2}{\lambda_0^2}E(x) \right) = 0 \quad x \in \mathbb{R}^2$$

Comme  $E$  est une fonction  $L^1_{loc}$ , la seconde limite ci-dessus est clairement nulle. Dans la première limite (29h), nous utilisons l'identité

$$w\left(\frac{x}{\varepsilon}\right)\Delta A(x) = \Delta \left( w\left(\frac{x}{\varepsilon}\right)A(x) \right) - \frac{1}{\varepsilon^2} \Delta w\left(\frac{x}{\varepsilon}\right)A(x) - \frac{2}{\varepsilon} \langle \nabla w\left(\frac{x}{\varepsilon}\right) | \nabla A \rangle(x)$$

ce qui développe cette limite (29h) comme la somme de ces trois termes :

$$\begin{aligned}\lim_{\varepsilon \rightarrow 0} \Delta \left( w \left( \frac{x}{\varepsilon} \right) A(x) \right) &= 0 \quad x \in \mathbb{R}^2; \\ \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^2} \Delta w \left( \frac{x}{\varepsilon} \right) A(x) &= 0 \quad x \in \mathbb{R}^2. \\ \lim_{\varepsilon \rightarrow 0} \left\langle \frac{1}{\varepsilon} \nabla w \left( \frac{x}{\varepsilon} \right) \middle| \nabla A(x) \right\rangle &= 0 \quad x \in \mathbb{R}^2;\end{aligned}$$

$A$  est  $L^1_{loc}$ , donc le premier terme est vrai sans  $\Delta$  : comme on est dans l'espace des distributions, la limite avec  $\Delta$  est vraie aussi.  $A$  est continu en 0, donc le second terme vaut

$$\left( \int \Delta w \right) A(0) \delta = \left( - \int \langle \nabla w | \nabla 1 \rangle \right) A(0) \delta = 0 \times A(0) \delta = 0$$

où nous avons utilisé une intégration par parties.  $\nabla A$  est  $L^1_{loc}$  et est borné, donc le troisième terme est pour  $x$  au voisinage de 0 en

$$O(\varepsilon \int |\nabla w|)$$

qui tend donc vers 0 dans l'espace des distributions (nous utilisons la compacité du support de  $w$ ) lorsque  $\varepsilon$  tend vers 0. Ainsi, (29h) est vrai, donc de même (29g) et (29f) aussi : la preuve est terminée. ■

**Lemme 2.2** *La solution fondamentale de l'opérateur de Laplace est  $\frac{1}{2\pi} \log \|x\|$ .*

■ Soit  $f$  une fonction  $C^2$  à support compact. Il faut prouver que  $\int_{x \in \text{supp} f} \Delta f(x) \log \|y - x\| dx = 2\pi f(y)$ . Les personnes familières avec la théorie des distributions trouveront ce lemme facile, car le Laplacien peut être déplacé entre les opérandes d'une convolution, et comme le Laplacien de ce log vaut  $2\pi\delta$  : voir par exemple Schwartz [S61] p 314. Une démonstration élémentaire est aussi possible, et la voici : la fonction

$$\log_\varepsilon(x) = \frac{1}{4} \max\left(0, 1 - \frac{\|x\|^4}{\varepsilon^4}\right) + \min\left(0, \frac{\|x\|^2}{\varepsilon^2} - 1\right) + \log \max(\|x\|, \varepsilon)$$

converge vers  $\log \|x\|$  dans  $L^1_{loc}$ , d'après le théorème de convergence dominée (fonction dominatrice  $\frac{5}{4} + |\log \|x\||$ ). Cette convergence a aussi lieu dans l'espace des distributions, ce qui donne :

$$\lim_{\varepsilon \rightarrow 0} \int_{x \in \text{supp} f} \Delta f(x) \log_\varepsilon(y - x) dx = \int_{x \in \text{supp} f} \Delta f(x) \log \|y - x\| dx$$

Ici, la limite  $\lim_{\varepsilon \rightarrow 0}$  est localement uniforme en  $y$ . Cette fonction  $\log_\varepsilon$  vaut  $\log \|x\|$  là où  $\|x\| \geq \varepsilon$ ; elle vaut  $\frac{\|x\|^2}{\varepsilon^2} - \frac{\|x\|^4}{4\varepsilon^4} - \frac{3}{4} + \log \varepsilon$  pour  $\|x\| \leq \varepsilon$ . Le raccord est correct pour toutes les dérivées d'ordre 1 et 2 car :

$$\frac{\partial \log_\varepsilon}{\partial x_i}(x) = \begin{cases} -\frac{x_i \|x\|^2}{\varepsilon^4} + \frac{2x_i}{\varepsilon^2} & \text{si } \|x\| < \varepsilon \\ \frac{x_i}{\|x\|^2} & \text{sinon} \end{cases}$$

$$\frac{\partial^2 \log_\varepsilon}{\partial x_i \partial x_j}(x) = \begin{cases} -2\frac{x_i x_j}{\varepsilon^4} - \delta_i^j \frac{\|x\|^2}{\varepsilon^4} + \delta_i^j \frac{2}{\varepsilon^2} & \text{si } \|x\| < \varepsilon \\ -\frac{2x_i x_j}{\|x\|^4} + \delta_i^j \frac{1}{\|x\|^2} & \text{sinon} \end{cases}$$

Ainsi,  $\log_\varepsilon$  est  $C^2$ , et donc l'intégration par parties suivante est indolore :

$$\int_{x \in \text{supp} f} \Delta f(x) \log_\varepsilon(y - x) dx = (-1)^2 (-1)^2 \int_{x \in \text{supp} f} f(x) (\Delta \log_\varepsilon)(y - x) dx.$$

La fonction  $f$  est  $C^2$ , donc un développement de Taylor à reste intégral de  $f$  en  $y$  montre que

$$\lim_{\varepsilon \rightarrow 0} \int_{x \in \text{supp} f} f(x) \Delta_y \log_\varepsilon \|y - x\| dx = 2\pi f(y).$$

car on a

$$\Delta \log_\varepsilon = \begin{cases} -4\frac{\|x\|^2}{\varepsilon^4} + \frac{4}{\varepsilon^2} & \text{si } \|x\| < \varepsilon \\ 0 & \text{sinon} \end{cases}.$$

■

La solution fondamentale du Laplacien  $\frac{1}{2\pi} \log$  peut être devinée en prenant le cas  $\lambda_0 = \infty$  dans l'équation différentielle (29b) vérifiée par la fonction de Hankel  $H_0^{ank}$  :  $\log$  est une solution de l'équation différentielle résultante  $f'' + \frac{1}{r} f' = 0$ ; la solution constante est une autre solution, mais elle n'a pas d'intérêt.

### 2.1.2 La condition de rayonnement de Sommerfeld : condition aux limites à l'infini

À la différence du cas du Laplacien, l'opérateur de Helmholtz n'a pas qu'une seule solution fondamentale canonique (voir la proposition 2.1). Une condition aux limites à l'infini dite *condition de rayonnement de Sommerfeld* (vue en (21a)) permettra de discriminer entre ces solutions.

Les figures 8 et 9 présentent les graphes des parties réelles des fonctions  $G$  (28) et  $\overline{G}$  (29a) multipliées par  $e^{-i\omega t}$ . Attention, l'axe des abscisses est  $\|x\|$  et non plus  $x_1$  comme dans la figure 7. Voici la différence entre ces deux

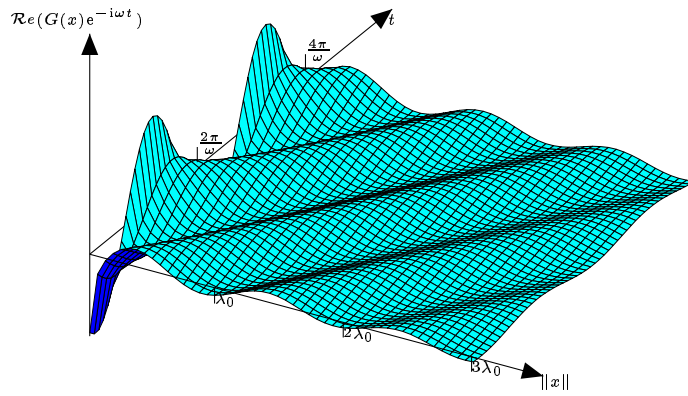


FIG. 8 – La solution fondamentale  $G$  (28) rayonnante vers l'infini avec dépendance au temps en  $e^{-i\omega t}$ .

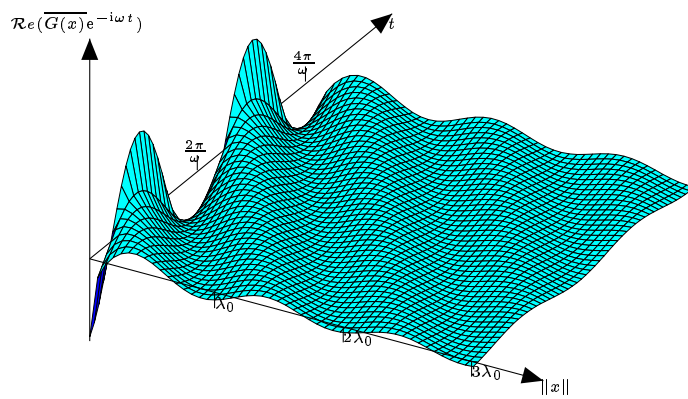


FIG. 9 – La solution fondamentale  $\overline{G}$  (29a) se concentrant avec dépendance au temps en  $e^{-i\omega t}$ .

solutions, qu'on comprend en suivant les lignes de niveaux de ces surfaces : la solution fondamentale (28) correspond à des ondes émises, l'onde s'éloigne de l'origine, tandis que la solution fondamentale (29a) correspond à des ondes reçues, l'onde se concentre vers l'origine, lorsqu'on prend en compte cette dépendance au temps en  $e^{-i\omega t}$ .

Pour distinguer entre ces deux solutions  $G$  et  $\bar{G}$  leurs combinaisons affines, et les ondes planes (27b), nous avons besoin d'une condition aux limites à l'infini. Comme l'opérateur de Helmholtz est oscillatoire, nous avons besoin d'une condition aux limites oscillatoire.

Sommerfeld [S12] a introduit la condition aux limites (21a), qu'on appelle communément la *condition de rayonnement de Sommerfeld* que nous rappelons :

$$\lim_{\|x\| \rightarrow \infty} r^{\frac{1}{2}} \left( \frac{\partial E}{\partial r}(x) - \frac{2i\pi}{\lambda_0} E(x) \right) = 0$$

où  $r = \|x\|$  est le rayon. L'exposant  $\frac{1}{2}$  est spécifique au cas bidimensionnel (en dimension  $d$ , il faut prendre  $\frac{d-1}{2}$ ).

Nous nous contenterons d'une condition plus faible, utilisée par exemple par Wilcox [W56] :

$$0 = \lim_{r \rightarrow \infty} \int_{\|x\|=r} \left| \frac{\partial E}{\partial r}(x) - \frac{2i\pi}{\lambda_0} E(x) \right|^2 dx \quad (30)$$

Cette condition de rayonnement permet de discriminer entre ces deux solutions fondamentales  $G$  (28) et  $\bar{G}$  (29a) lorsqu'on y remplace  $E$  par ces solutions fondamentales ; de plus, elle élimine leurs combinaisons affines, et les ondes planes (27b) :

**Propriété 2.3** *De toutes les solutions vues ci-dessus en (27b) et dans la propriété 2.1, seule la solution radiale émise  $G$  (28) vérifie la condition de rayonnement de Sommerfeld sous sa forme forte (21a) ou sous sa forme faible (30).*

■ Pour les ondes planes, on a

$$r^{1/2} \left( \frac{\partial}{\partial r} E(x) - \frac{2i\pi}{\lambda_0} E(x) \right) = r^{1/2} (\langle x | \vec{e}_1 \rangle - 1) \frac{2i\pi}{\lambda_0} E(x)$$

qui explose dans toutes les directions, sauf à proximité immédiate de la demi-droite  $x \in \mathbb{R}_+ \vec{e}_1$ . Donc l'onde plane ne vérifie pas la condition de rayonnement de Sommerfeld, ni (21a) ni (30).

On voit bien que  $\vec{e}_1$ , qui est la seule direction dans laquelle la condition de rayonnement de Sommerfeld (21a) est satisfaite, correspond à une direction dans laquelle l'onde se propage : cette condition modélise assez bien le fait que l'onde (et donc l'énergie) ne peut venir de l'infini.

Pour qu'une fonction radiale  $E(x) = f(\frac{2\pi}{\lambda_0}\|x\|)$  vérifie cette équation, il faut que

$$f'(r) - if(r) = o(\frac{1}{r}). \quad (31)$$

Watson [W22] montre, p196 partie 7.2, que  $H_0^{ank}$  et sa dérivée  $H_0^{ank'}$  ont des équivalents en l'infini tels que

$$H_0^{ank}(r) - (1 - i)e^{ir}/\sqrt{\pi r} = O(r^{-3/2}) \quad (32a)$$

et

$$H_0^{ank'}(r) - (1 + i)e^{ir}/\sqrt{\pi r} = O(r^{-3/2}). \quad (32b)$$

On en déduit que  $G + \eta(\bar{G} - G)$  ne vérifie la condition de rayonnement de Sommerfeld que lorsque  $\eta$  est nul. ■

### 2.1.3 Un domaine extérieur non borné

Nous résolvons l'équation de Helmholtz (20) dans le complémentaire d'un fermé  $S$  appelé le *diffractant*, qui a par exemple les formes suivantes :



(33)



(34)

Il faudra ajouter deux conditions aux limites : une condition à l'infini (30) ; et une condition sur la valeur de la solution  $E$  au bord de  $S$ , détaillée dans la partie suivante en (35c).

Nous utiliserons, sur le diffractant  $S$ , les hypothèses suivantes :

1. compact, connexe, et de complémentaire connexe ;
2. Lipschitzien (donc sans épines) ;
3. avec une continuité par morceaux de type  $C^p$  par morceaux pour un certain entier  $p \geq 2$  ;
4. sans raccord des demi-tangentes entre les morceaux  $C^p$ .

Un ensemble  $S$  est dit Lipschitzien et  $C^p$  par morceaux s'il est délimité par une courbe qui est localement et à une rotation près le graphe d'une fonction Lipschitzienne et  $C^p$  par morceaux. Pour la partie 4 et la justification des résultats numériques présentés, une régularité  $C^5$  a été nécessaire.

Comme l'opérateur de Helmholtz (26) est une perturbation d'ordre inférieur du Laplacien, et que l'espace fonctionnel  $H_{loc}^1({}^cS)$  est le bon espace pour rechercher une solution pour Laplacien d'après la propriété 2.7, alors cet espace sera aussi le bon pour rechercher la solution du problème de Helmholtz extérieur 21c.

Nous définissons ci-dessous sur l'ouvert  ${}^cS$  les deux espaces  $\mathcal{D}_0({}^cS)$  et  $\mathcal{D}(\overline{{}^cS})$  contenant les fonctions régulières respectivement sans et avec valeur de bord. Ces notations sont classiques, par exemple utilisées par Temam [T79].

**Définition 2.4** *L'espace  $\mathcal{D}_0({}^cS)$  est l'ensemble des fonctions  $\mathbb{R}^2 \rightarrow \mathbb{C}$  indéfiniment dérivables à support compact inclus dans l'ouvert  ${}^cS$ . L'espace  $\mathcal{D}(\overline{{}^cS})$  est l'ensemble des fonctions  ${}^cS \rightarrow \mathbb{C}$  indéfiniment dérivables à support précompact inclus dans le fermé  $\overline{{}^cS}$ , telles que toutes les fonctions dérivées sont prolongeables en des fonctions  $\overline{{}^cS} \rightarrow \mathbb{C}$ .*

Nous définissons sur l'ouvert non borné  ${}^cS$  la norme  $\|\cdot\|_{H^1({}^cS)}$  et les espaces de Sobolev  $H^1({}^cS)$ ,  $H_0^1({}^cS)$  et  $H^{-1}({}^cS)$  :

**Définition 2.5** *La norme de l'espace  $H^1({}^cS)$  est définie par*

$$\|E\|_{H^1({}^cS)} = \sqrt{\int_{{}^cS} |E(x)|^2 + |\nabla E(x)|^2 dx} \quad (35a)$$

*pour une fonction régulière  $E \in \mathcal{D}(\overline{{}^cS})$ , où  $\nabla E(x)$  est le vecteur  $(\frac{\partial E}{\partial x_1}(x), \frac{\partial E}{\partial x_2}(x))$ .*

*L'espace  $H^1({}^cS)$  est la complétion de  $\mathcal{D}(\overline{{}^cS})$  pour cette norme. L'espace  $H_0^1({}^cS)$  est la complétion de  $\mathcal{D}_0({}^cS)$  pour cette norme. Nous prolongeons l'application norme (35a) par densité de  $E \in \mathcal{D}(\overline{{}^cS})$  aux espaces  $H^1({}^cS)$  et  $H_0^1({}^cS)$ .*

*L'espace  $H^{-1}({}^cS)$  est l'identification du dual de  $H_0^1({}^cS)$  à travers le crochet de dualité de  $L^2({}^cS)$  qui est l'espace Hilbertien des fonctions  ${}^cS \rightarrow \mathbb{C}$  de carré intégrable.*

Malheureusement, la propriété 2.16 montrera que l'espace  $H^1({}^cS)$  n'est pas le bon pour une solution de l'équation de Helmholtz (20) car la seule solution



possible  $H^1({}^cS)$  sera la solution nulle. Nous devons nous contenter, pour la suite, d'un espace  $H_{loc}^1({}^cS)$  complet mais non normé :

**Définition 2.6** *L'espace  $H_{loc}^1(\mathbb{R}^2)$  est l'ensemble des distributions  $\mathbb{R}^2 \rightarrow \mathbb{C}$  telles que pour tout sous-espace  $C$  ouvert borné de  $\mathbb{R}^2$ , la restriction à  $C$  soit  $H^1(C)$ . L'espace  $H_{loc}^1({}^cS)$  est la trace de  $H_{loc}^1(\mathbb{R}^2)$  sur  ${}^cS$ .*

Le théorème 2.14 montrera que le problème (20) admet dans  $H_{loc}^1({}^cS)$  au plus une solution, avec des conditions aux limites à l'infini.

Les fonctions qui appartiennent à ces espaces fonctionnels  $L^2({}^cS)$ ,  $\mathcal{D}_0({}^cS)$ ,  $\mathcal{D}(\overline{{}^cS})$ ,  $H^1({}^cS)$ ,  $H_{loc}^1({}^cS)$ ,  $H_0^1({}^cS)$  ont pour support le complémentaire  ${}^cS$  du diffractant  $S$ .

Les espaces  $H^1({}^cS)$  et  $H^{-1}({}^cS)$  sont les bons espaces, aux poids près, pour bien poser un problème avec le Laplacien. La propriété 2.16 montrera que ce n'est pas le cas pour un problème avec l'opérateur de Helmholtz.

**Propriété 2.7** *Le Laplacien  $\Delta$  est continu de  $H^1({}^cS)$  dans  $H^{-1}({}^cS)$ .*

■ Prenons deux fonctions  $\phi \in \mathcal{D}(\overline{{}^cS})$  et  $\psi \in \mathcal{D}_0({}^cS)$ . On a

$$\int_{{}^cS} \Delta\phi(x)\psi(x)dx = \int_{\partial S} \partial_n\phi(x)\psi(x)dx - \int_{{}^cS} \langle \nabla\phi(x) | \nabla\psi(x) \rangle dx.$$

Une inégalité de Cauchy-Schwarz montre que ce membre droit est plus petit en module que  $\|\phi\|_{H^1({}^cS)}\|\psi\|_{H^1({}^cS)}$  lorsque  $\psi \in H_0^1({}^cS)$ , car  $\psi$  est nul sur  $\partial S$ . ■

Nous aurons besoin de propriétés plus précises sur ce Laplacien, qui nécessitent la définition d'un espace fonctionnel sur  $\partial S$  pour définir des conditions aux bords.

Une interprétation physique de l'équation de Helmholtz (20) est l'étude d'un mode propre de l'équation des ondes (avec vitesse  $c$ ) à la fréquence  $\omega = \frac{2\pi c}{\lambda_0}$ . Par exemple, il peut s'agir de l'équation de Maxwell en trois dimensions,

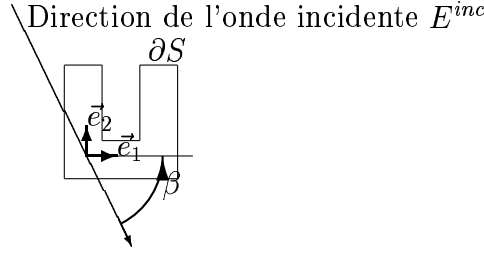
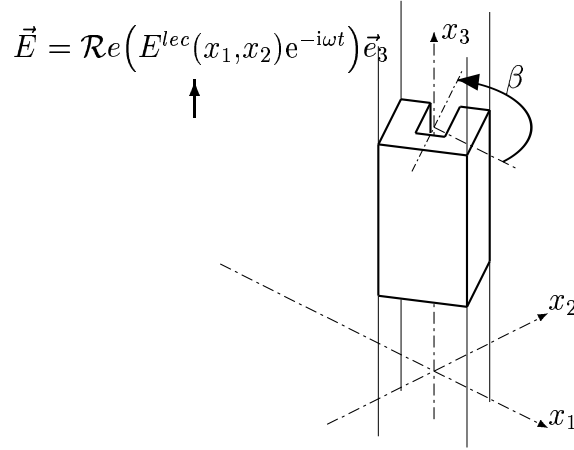


FIG. 10 – L'angle  $\beta$ , après le changement de référentiel.

autour d'un profilé métallique dont la section est  $S$  :



L'angle  $\beta$  est un paramètre du second membre  $E^{inc}$  et donc du problème (21c); ce paramètre est indépendant du temps. Plus particulièrement, si nous supposons que ce mode propre a son champ électrique colinéaire à une génératrice de ce profilé métallique (polarisation dite transverse magnétique et notée TM, voir Pan-Tretiakov-Gilbert [PTG03]), alors la composante  $E^{lec} \in H_{loc}^1(cS)$  dans cette direction  $\vec{e}_3$  du champ électrique vérifie l'équation de Helmholtz (20). Il en résulte que le diffractant est tourné d'un angle  $\beta$  par rapport à la base  $(\vec{e}_1, \vec{e}_2)$  dans laquelle l'onde plane a pour vecteur directeur le vecteur de coordonnées  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ . Par invariance, nous pouvons de façon équivalente supposer que  $\beta$  est l'angle avec lequel l'onde plane arrive par rapport à un référentiel  $(\vec{e}_1, \vec{e}_2)$  dans lequel le diffractant est fixe, voir la figure 10.

Les équations de Maxwell prédisent qu'au bord du diffractant, le champ électrique  $E^{lec}$  s'annule.

Dans cette interprétation, le champ électrique  $E^{lec} \in H_{loc}^1(cS)$  ne vérifie plus la condition de rayonnement de Sommerfeld : en effet, on a

$$E^{lec} = E^{inc} + E^{diff}$$

où le champ incident  $E^{inc} \in H_{loc}^1({}^cS)$  est donné par l'onde plane

$$E^{inc}(x) = e^{\frac{2i\pi}{\lambda_0} \langle \vec{e}_1 \cos \beta - \vec{e}_2 \sin \beta | x \rangle} \quad (35b)$$

qui est l'excitation, et où  $E^{diff} \in H_{loc}^1({}^cS)$  vérifie la condition de rayonnement de Sommerfeld car c'est la conséquence de l'excitation  $E^{inc}$  qui n'a lieu que sur  $\partial S$  qui est borné.

L'onde plane incidente  $E^{inc}$  est  $H_{loc}^1(\mathbb{R}^2)$  ce qui physiquement s'interprète par le fait que l'énergie est localement finie.

Comme le champ électrique émis  $E^{diff}$  est une conséquence de la diffraction par le diffractant  $S$  de la portion de  $E^{inc}$  qui touche  $S$ , le champ électrique émis  $E^{diff}$  est produit par une source d'énergie finie par unité de temps, c'est à dire que  $E^{diff}$  est dans  $H_{loc}^1({}^cS)$  et tend même vers 0 en l'infini, car la propagation à vitesse finie de l'énergie est une des caractéristiques essentielles des équations hyperboliques Nédélec [N01]. Ainsi,  $E^{diff}$  est le bon candidat pour être la fonction  $E$  inconnue que nous étudierons (unicité, existence).

Dautray-Lions [DL85], t2 XI p623, note que l'analogue pour le Laplacien du problème extérieur bidimensionnel (21c) ne peut pas être considéré comme le cas limite du cylindre de longueur tendant vers l'infini dans l'analogue pour le Laplacien du problème extérieur tridimensionnel. L'intérêt du problème (21c) est donc mathématique, il vise à comprendre ce qui se passe dans le cas bidimensionnel, pour tirer des méthodes qui seront peut-être utiles pour le cas tridimensionnel (non abordé dans cette thèse).

Nous avons donc vu une interprétation physique de l'équation de Helmholtz, permettant d'analyser les modes propres de l'équation des ondes dans le cas transverse magnétique (TM). Il existe d'autres applications de l'équation de Helmholtz: il y a le cas linéarisé de l'équation des ondes acoustiques; Wendland [W90] mentionne un écoulement bidimensionnel stationnaire à rotationnel et divergence nuls, et Colton-Kress [CK92] détaille page 13 la version tridimensionnelle de cette application. Dans cette application, le diffractant  $S$  est un objet au bord duquel la pression est constante.

#### 2.1.4 Condition aux limites au bord du diffractant

L'équation de Helmholtz dans  $\partial S$  (20) a beaucoup de solutions, par exemple les restrictions à ce domaine  ${}^cS$  de l'onde plane et de la solution fondamentale  $G$  (28). Même si nous imposons la condition de rayonnement de Sommerfeld, il reste une infinité de solutions, qu'on obtient en translatant

la solution fondamentale  $G$  (tout en prenant bien garde à ce que la singularité de cette solution soit dans  $S$ ).

Nous avons besoin d'une autre condition aux limites pour ne plus rendre possible qu'une seule solution. Pour cela, nous pouvons traduire le fait que, dans l'interprétation de Maxwell TM ci-dessus,  $E^{lec}$  s'annule sur  $S$ . Comme  $H_{loc}^1({}^cS)$  contient des fonctions singulières (par exemple  $x \mapsto \log(1 + |\log \|x - x_0\||)$ ), nous avons besoin de définir plus précisément la *valeur au bord* d'une fonction  $H_{loc}^1({}^cS)$ :

**Définition 2.8** *L'application trace  $\nu_0$  est la complétion à  $\phi \in H_{loc}^1({}^cS)$  de l'application qui à  $\phi \in \mathcal{D}(\overline{{}^cS})$  associe  $\phi|_{\partial S}$ .*

Dans l'interprétation ci-dessus,  $E^{lec} = E + E^{inc}$  s'annule au bord du diffractant  $S$ . Ceci donne la condition à la limite (21b) qui plus formellement s'écrit

$$\nu_0 E = -\nu_0 E^{inc} \quad \text{sur } \partial S. \quad (35c)$$

où  $E^{inc}$  est une fonction régulière bornée définie sur  $\mathbb{R}^2$ . Nous pouvons même conserver la définition de  $E^{inc}$  (35b) pour le problème (21c) étudié par cette thèse.

La rédaction de la thèse sera facilitée par la notation suivante, qui est un paramétrage  $s$  de  $\partial S$ .

**Définition 2.9** *Soit  $L$  le périmètre de  $S$ , et soit  $s$  une application  $\mathbb{R}/L\mathbb{Z} \rightarrow \partial S$  qui soit une paramétrisation par abscisse curviligne de  $\partial S$ , dans le sens direct. Nous utiliserons souvent  $s'(t)^\perp$  pour noter la normale extérieure au diffractant  $S$ , avec  $^\perp$  la rotation indirecte d'angle  $\frac{\pi}{2}$  d'un vecteur, et où  $s'(t)$  est la dérivée (vectorielle) de  $s$  en  $t$ .*

Une conséquence immédiate de l'hypothèse (2) que  $S$  est Lipschitzien est que

$$c \inf_{m \in \mathbb{Z}} |t - t' - Lm| \leq \|s(t) - s(t')\| \leq \inf_{m \in \mathbb{Z}} |t - t' - Lm|$$

où  $c$  est une constante, et où l'inf est un ingrédient pour définir la distance de  $\mathbb{R}/L\mathbb{Z}$ .

La trace  $\nu_0$  a de bonnes propriétés, données par Grisvard [G85]:

**Propriété 2.10** *Lorsque le diffractant  $S$  vérifie les hypothèses 1 à 4 de la page 62, cette application  $\nu_0$  est prolongeable continûment à  $H_{loc}^1({}^cS)$ , et son image est incluse dans  $L^2(\partial S)$ .*

*Son image est l'espace de Sobolev  $H^{\frac{1}{2}}(\partial S)$ , qui est homéomorphe à  $H^{\frac{1}{2}}(\mathbb{R}/L\mathbb{Z})$  via la paramétrisation  $s$ .*

*Les espaces  $H^m(\mathbb{R}/L\mathbb{Z})$  avec  $m \in \mathbb{R}$  sont les espaces de Sobolev dont la norme est*

$$f \mapsto \sqrt{\sum_{z \in \mathbb{Z}} (1 + z^2)^m \left| \int_0^L f(t) e^{i \frac{2\pi}{L} z t} dt \right|^2}.$$

*et qui contient les distributions  $f : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{C}$  telles que cette norme est finie.*

*Nous définissons les espaces de Sobolev  $H^m(\partial S)$ , comme étant homéomorphe à  $H^m(\mathbb{R}/L\mathbb{Z})$  via la paramétrisation  $s$ , pour  $m$  valant  $-\frac{1}{2}$ ,  $0$ ,  $\frac{1}{2}$  et  $1$ .*

*De plus, il existe un relèvement  $\nu_0^{-1}$  de cette application  $\nu_0$ , c'est à dire qu'il existe une application  $\nu_0^{-1}$  continue de  $H^{\frac{1}{2}}(\partial S)$  dans  $H^1({}^cS)$  telle que  $\nu_0 \circ \nu_0^{-1} = Id_{H^{\frac{1}{2}}(\partial S)}$ .*

Cette propriété a été prouvée par Lions-Magenes [LM68]. Nous n'en reproduisons pas la démonstration ici.

Voici tout d'abord une observation sur la régularité des solutions de l'équation (20) dans l'ouvert  ${}^cS$ :

**Remarque 2.11** *Une solution  $E \in H_{loc}^1({}^cS)$  de l'équation (20) est analytique et  $H^m(C)$  pour tout compact  $C$  inclus dans  ${}^cS$ .*

La démonstration de cette remarque se fait par récurrence sur  $m$ , et se base d'une part sur l'ellipticité de  $\Delta$  qui est facile à démontrer directement sur la définition de l'espace  $H_{compact}^m({}^cS)$ , et d'autre part sur l'identité

$$\phi \Delta^{-1}(\phi \Delta(\phi E)) = \phi^2 E - \phi \Delta^{-1}(E \Delta \phi) - 2\phi \Delta^{-1}\langle \nabla \phi | \nabla E \rangle$$

où  $\phi$  est une fonction  $\mathcal{D}_0({}^cS)$  qui vaut 1 sur  $C$ . Nous n'utiliserons pas cette remarque 2.11 dans les démonstrations, mais seulement pour éclaircir certaines démonstrations, car la démonstration de cette remarque (sans l'analyticité) fait intervenir des notions (Fourier pour  $\mathcal{D}_0({}^cS)$ ) qui ne sont pas plus simples que les résultats de cette partie 2. De plus, cette remarque 2.11 résultera du théorème de représentation 2.13 (comme dans le cas tridimensionnel, voir Colton-Kress [CK92]).

Lions-Magenes [LM68] ont aussi montré que, dans certains cas, par exemple lorsque  $E$  vérifie l'équation (20), la dérivée normale de  $E$  a aussi une limite sur  $\partial S$ :

**Propriété 2.12** *Une solution  $E \in H_{loc}^1({}^cS)$  de l'équation de Helmholtz (20) a une dérivée normale à  $S$  qui est dans  $H^{-\frac{1}{2}}(\partial S)$ , qui est aussi l'identification du dual de  $H^{\frac{1}{2}}(\partial S)$  à travers le crochet de dualité de  $L^2(\partial S)$ . Nous notons  $\nu_1 E$  cette dérivée normale (extérieure).*

■ (30) Soit  $H \in H_{loc}^1({}^cS)$ . Montrons que

$$\int_{\|x\|=r} \frac{2\pi}{\lambda_0} \frac{\partial E}{\partial r} H \quad (35d)$$

est convergent pour presque tout  $r$  assez grand (c'est immédiat en admettant la continuité de  $\frac{\partial E}{\partial r}$  donnée par la remarque 2.11).

L'inégalité de Cauchy-Schwarz montre que

$$\left| \frac{\partial E}{\partial r}(x) H(x) \right| \leq \left| \frac{\partial E}{\partial r}(x) \right|^2 + |H(x)|^2 \leq |\nabla E(x)|^2 + |H(x)|^2$$

Les fonctions  $E$  et  $H$  sont  $H_{loc}^1({}^cS)$ , donc la fonction

$$r \rightarrow \int_{\|x\|=r} |E|^2 + |\nabla E|^2 + |H|^2 + |\nabla H|^2$$

est localement intégrable: donc cette fonction est finie presque partout, et donc (35d) est finie pour presque tout  $r$  tel que  $r > r_0$ , où  $r_0$  est le plus petit rayon tel que  $S$  est inclus dans la boule de centre 0 et de rayon  $r_0$ .

### Une formule de Stokes.

Soit  $r \in (r_0, r_0 + 1]$  tel que l'intégrale (35d) est finie. Utilisons une formule de Stokes dans l'espace  $S_r$  délimité par  $S$  et  $\{x : \|x\| = r\}$ , visibles sur la figure (11):

$$S_r := \{x \in {}^cS : \|x\| < r\}$$

$E$  est  $H^1({}^cS)$  localement,  $\Delta E$  est dans  $H^1({}^cS)$  localement comme  $\Delta E = -\frac{4\pi^2}{\lambda_0^2} E$ . Ainsi, la formule de Stokes donne ceci, en supposant  $E$  régulière:

$$\int_0^L \nu_1 E(s(t)) \nu_0 H(s(t)) dt = \int_{\|x\|=r} \frac{\partial E}{\partial r} H - \int_{S_r} H \Delta E - \int_{S_r} \langle \nabla H | \nabla E \rangle. \quad (36)$$

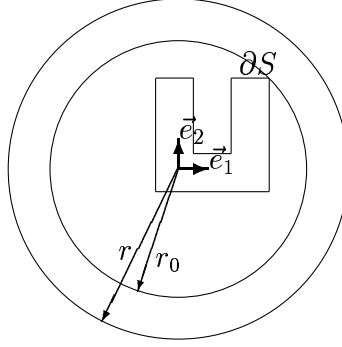


FIG. 11 – *Le grand cercle et le diffractant délimitent  $S_r$ .*

avec

$$\nu_1 E(s(t)) = \langle s'(t)^\perp | \nabla E(s(t)) \rangle$$

où la définition 2.9 a noté  $^\perp$  la rotation indirecte : ainsi, ceci est une dérivée normale au sens habituel.

Nous allons montrer maintenant que le membre droit de (36) dépend continûment de  $E, H \in H_{loc}^1(cS)$ , ce qui nous fournira la définition de  $\nu_1 E$  comme un élément de  $H^{-1/2}(\partial S)$  dépendant continûment de  $E \in H_{loc}^1(cS)$ . Dans cette démarche, la surjectivité de  $\nu_0 : H_{loc}^1(cS) \rightarrow H^{1/2}(\partial S)$ , fournie par la proposition (2.10), est très importante.

Pour montrer la continuité du membre droit de (36), nous devons montrer que

$$\int_0^L \nu_1 E(s(t)) \nu_0 H(s(t)) dt \leq C \|E\|_{H^1(S_r)} \|H\|_{H^1(S_r)}. \quad (37a)$$

Or l'intégration de (36) pour  $r$  entre  $r_0$  et  $r_0 + 1$  donne

$$\begin{aligned} & \int_0^L \nu_1 E(s(t)) \nu_0 H(s(t)) dt \\ &= \int_{\partial S} \nu_1 E \nu_0 H \\ &= \int_{r_0}^{r_0+1} \int_{\partial S} \nu_1 E \nu_0 H dr \\ &= \int_{r_0}^{r_0+1} \int_{\|x\|=r} \frac{\partial E}{\partial r} H dr - \int_{r_0}^{r_0+1} \int_{S_r} (H \Delta E + \langle \nabla H | \nabla E \rangle) dr \\ &= \int_{r_0-1 \leq \|x\| \leq r_0} \frac{\partial E}{\partial r} H + \int_{S_r} \left( \frac{4\pi^2}{\lambda_0^2} H E - \langle \nabla H | \nabla E \rangle \right) \min(1, r_0 + 1 - r) dr \end{aligned}$$

ce qui est visiblement plus petit que  $C \|E\|_{H^1(S_{r_0})} \|H\|_{H^1(S_{r_0})}$ , et avec une constante  $C$  qui ne dépend que de  $r_0$ , de  $\lambda_0$  et de  $S$ . La dépendance en

$r_0$  n'est pas importante : il suffit de prendre  $r_0$  minimal tel que  $S$  soit inclus dans la boule de centre 0 et de rayon  $r_0 - 1$ .

Ainsi, nous avons montré (37a), et donc  $\nu_1 E$  est bien dans  $H^{-\frac{1}{2}}(\partial S)$  lorsque  $E$  est une solution  $H_{loc}^1({}^c S)$  de l'équation de Helmholtz (20). ■

### 2.1.5 Unicité dans $H_{loc}^1({}^c S)$ , pour Dirichlet (et Neumann)

Le théorème de représentation suivant va permettre de caractériser une solution  $E$  en ne connaissant que  $\nu_0 E$  et  $\nu_1 E$ . Il s'en suivra un résultat immédiat d'unicité. C'est la première étape, la suivante étant un schéma numérique qui n'utilise de  $E$  que ses valeurs au bord  $\nu_0 E$  et  $\nu_1 E$ .

**Théorème 2.13 (représentation)** *Lorsque  $E$  est une solution  $H_{loc}^1({}^c S)$  de (20) vérifiant la condition de rayonnement de Sommerfeld (30),  $E$  vérifie presque partout l'égalité suivante :*

$$\nu_1 E \star G + (s'^{\perp}(\cdot)\nu_0 E) \star \nabla G = \mathbb{1}_{cS} E \quad (38)$$

où  $G$  est la solution fondamentale (28), et où  $\star$  est la convolution au sens où cette égalité s'écrit aussi sous la forme

$$\begin{aligned} \int_0^L \nu_1 E(s(t))G(x - s(t))dt + \int_0^L \nu_0 E(s(t))\langle s'(t)^{\perp} | \nabla \rangle G(x - s(t))dt \\ = \begin{cases} E(x) & \text{si } x \in {}^c S \\ 0 & \text{sinon} \end{cases} \quad x \notin \partial S \end{aligned}$$

Ainsi, dans la convolution  $f \star g$ , où  $f$  a pour support  $\partial S$  et  $g$  a son support inclus dans  $\mathbb{R}^2$ , la notation  $\star$  joue un rôle un peu dissymétrique ;  $f \star g$  est dans cette thèse une notation abrégée pour la notation consensuelle  $(fd\sigma) \star g$  où  $d\sigma$  est la mesure de Lebesgue unidimensionnelle de l'ensemble  $\partial S$  vu comme sous-ensemble de  $\mathbb{R}^2$ .

L'analyticité de  $E$  sur  ${}^c S$  est une conséquence de ce théorème.

Ce théorème est montré par Wilcox [W56] et Colton-Kress [CK92] dans le cas tridimensionnel et avec une hypothèse de régularité  $C^2$  sur  $E$  (Colton-Kress, sur  $\partial S$ , ne demande que la continuité et l'existence des dérivées de Fréchet normales, mais demande une continuité  $C^2$  sur  $S$  et aussi la condition plus forte (21a)). Ce théorème est aussi mentionné par Costabel [C88], en remplaçant (20) par une hypothèse de compacité sur le support de  $E$  (le second membre de (38) est à adapter en conséquence).



■ Soit  $x_0 \in \mathbb{R}^2 \setminus \partial S$ . Nous allons montrer (38) en  $x = x_0$ . Nous allons raisonner uniquement pour les  $x_0$  où les fonctions maximales  $M(E)$  et  $M(\nabla E)$  sont finies (ces fonctions maximales sont explicitées en (39g) et (39h), et leur finitude serait immédiate en admettant la remarque 2.11), et qui sont des points de Lebesgue de  $E$ . Ces conditions sont vraies presque partout en  $x_0$  pour  $L^1_{loc}$  et donc pour  $L^2_{loc}(^cS)$  d'après Rudin [R87] théorèmes 7.4 et 7.7.

### La condition de Sommerfeld limite l'explosion à l'infini sur $E$ .

La condition de rayonnement de Sommerfeld (30) contient un module que nous développons ainsi :

$$\left| \frac{\partial E}{\partial r}(x) - \frac{2i\pi}{\lambda_0} E(x) \right|^2 = \left| \frac{\partial E}{\partial r}(x) \right|^2 + \frac{4\pi^2}{\lambda_0^2} |E(x)|^2 - 2\mathcal{I}m\left(\frac{2\pi}{\lambda_0} \frac{\partial E}{\partial r}(x) \overline{E(x)}\right) \quad (39a)$$

Ceci est intégrable par  $\int_{\|x\|=r}$  pour presque tout  $r$  (pour le dernier terme, ceci a déjà été montré, voir (35d)). L'idée suivante est de décomposer l'intégrale  $\int_{\|x\|=r}$  du dernier terme : nous montrerons que cette intégrale est indépendante de  $r$ .

Soit  $r \in R$  tel que la boule centrée en 0 et de rayon  $r$  contienne  $S$  : Nous pouvons intégrer (39a) et ensuite utiliser la formule de Stokes (36) avec  $H = \overline{E}$  et en ignorant la partie réelle, ce qui donne :

$$\int_{\|x\|=r} (39a) = \int_{\|x\|=r} \left| \frac{\partial E}{\partial r} \right|^2 + \int_{\|x\|=r} \frac{4\pi^2}{\lambda_0^2} |E|^2 - \frac{4\pi}{\lambda_0} \mathcal{I}m\left(\int_{\partial S} \nu_1 E \overline{\nu_0 E}\right) \quad (39b)$$

D'après la condition de rayonnement de Sommerfeld (30), ceci a pour limite 0 lorsque  $r$  tend vers l'infini ; comme d'autre part l'expression

$$\mathcal{I}m\left(\int_{\partial S} \nu_1 E \overline{\nu_0 E}\right)$$

est indépendante de  $r$ , on a montré que

$$\int_{\|x\|=r} \left| \frac{\partial E}{\partial r} \right|^2 + \int_{\|x\|=r} \frac{4\pi^2}{\lambda_0^2} |E|^2 = O(1 + \|\nu_0 E\|_{\mathbb{H}^{\frac{1}{2}}(\partial S)} \|\nu_1 E\|_{\mathbb{H}^{-\frac{1}{2}}(\partial S)}) \quad (39c)$$

lorsque  $r \rightarrow \infty$ . Cette implication a été trouvée par Wilcox [W56] ; Sommerfeld [S12] en faisait une hypothèse, dite condition de finitude.

### Cas $x_0$ dans $S$ .

Supposons que  $x_0$  soit dans  $S \setminus \partial S$ . Soit  $r > 0$  suffisamment grand tel que  $S$  soit inclus dans la boule ouverte de centre 0 et de rayon  $r$  (voir la figure 12).

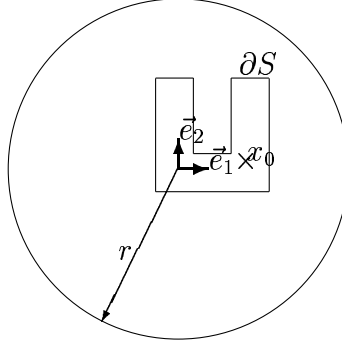


FIG. 12 – Cas  $x_0$  dans  $S$ .

Considérons la solution fondamentale translaturée  $G(x) = -\frac{i}{4}H_0^{ank}(\frac{2\pi}{\lambda_0}\|x - x_0\|)$  vue en (28). La fonction  $G$  est  $H_{loc}^1(cS)$ , donc nous pouvons la substituer à  $H$  dans la formule de Stokes (36).

$G$  est même une solution de l'équation de Helmholtz (20), donc nous pouvons échanger  $G$  et  $E$  dans la formule de Stokes (36).

Nous obtenons ainsi deux formules de Stokes. Si nous les soustrayons, nous avons formule de Green :

$$\begin{aligned} & \int_{\partial S} \nu_1 E \nu_0 G - \int_{\partial S} \nu_0 E \nu_1 G \\ &= \int_{\|x\|=r} \left( \frac{\partial E}{\partial r} G - \frac{\partial G}{\partial r} E \right) + \int_{S_r} E \Delta G - \int_{S_r} G \Delta E. \end{aligned} \quad (39d)$$

Comme  $E$  et  $G$  sont toutes deux solutions de (20), nous avons  $E \Delta G = G \Delta E$ , et donc les deux dernières intégrales se compensent pour presque tout  $r$ .

D'après les équivalents à l'infini (32a) et (32b) de  $H_0^{ank}$  et  $H_0^{ank'}$ , ce qu'il reste du second membre de (39d) vérifie :

$$\begin{aligned} & \int_{\|x\|=r} \left( \frac{\partial E}{\partial r} G - \frac{\partial G}{\partial r} E \right) \\ &= \frac{1}{4i} \int_{\|x\|=r} \left( \frac{\partial E}{\partial r}(x) H_0^{ank} \left( \frac{2\pi}{\lambda_0} \|x - x_0\| \right) - \frac{2\pi}{\lambda_0} H_0^{ank'} \left( \frac{2\pi}{\lambda_0} \|x - x_0\| \right) E(x) \right) dx \\ &= \frac{1-i}{4i} \sqrt{\frac{\lambda_0}{2\pi^2 r}} e^{\frac{2i\pi}{\lambda_0} r} \int_{\|x\|=r} \left( \frac{\partial E}{\partial r}(x) - \frac{2i\pi}{\lambda_0} E(x) \right) dx \\ & \quad + O(r^{-3/2}) \int_{\|x\|=r} \frac{\partial E}{\partial r} + O(r^{-3/2}) \int_{\|x\|=r} \frac{\partial E}{\partial r} \end{aligned}$$

Grâce à l'inégalité de Cauchy-Schwarz, la condition de rayonnement de Som-

merfeld (30) montre que

$$\int_{\|x\|=r} \left( \frac{\partial E}{\partial r}(x) - \frac{2i\pi}{\lambda_0} E(x) \right) dx = o(\sqrt{r}) \quad (39e)$$

et la propriété de finitude (39c) montre que  $\int_{\|x\|=r} \frac{\partial E}{\partial r}$  et  $+\int_{\|x\|=r} \frac{\partial E}{\partial r}$  sont des  $O(\sqrt{r})$ .

Il apparaît finalement que le membre droit de (39d) tend vers 0 lorsque  $r$  tend vers l'infini. Son membre gauche, qui est indépendant de  $r$ , est alors nul.

Il reste à reformuler  $\nu_1 E \nu_0 G - \nu_0 E \nu_1 G$  en le membre gauche de (38) dans l'énoncé. En utilisant la définition de  $G$ , on a :

$$\begin{aligned} & \int_{\partial S} \nu_1 E \nu_0 G - \int_{\partial S} \nu_0 E \nu_1 G \\ &= \int_0^L \left( \nu_1 E(s(t)) G(s(t) - x_0) - \nu_0 E(s(t)) \langle s'(t)^\perp | \nabla G(s(t) - x_0) \rangle \right) dt \\ &= \int_0^L \left( \nu_1 E(s(t)) G(x_0 - s(t)) + \langle \nu_0 E(s(t)) s'(t)^\perp | \nabla G(x_0 - s(t)) \rangle \right) dt \\ &= (\nu_1 E(s(t)) \star G)(x_0) + (\nu_0 E s'(\cdot)^\perp \star \nabla G)(x_0) \end{aligned}$$

Dans le cas  $x_0 \in S$ , le membre droit de (38) est nul.

### Cas $x_0$ hors de $S$ .

Soit  $x_0 \notin S$ . Soit  $r > 0$  suffisamment grand tel que  $B(x_0, \varepsilon)$  et  $S$  soient tous deux inclus dans la boule ouverte de centre 0 et de rayon  $r$ . Soit  $\varepsilon > 0$  tel que la boule de centre  $x_0$  et de rayon  $\varepsilon$  n'intersecte ni  $S$  ni  $\{x : \|x\| = r\}$ . Soit  $\alpha > 0$  tel que  $\alpha < \varepsilon$ .

Pour obtenir une identité de Green similaire à (39d), reprenons la formule de Stokes (36), avec les deux mêmes remplacements que pour (39d), sauf que la  $B(x_0, \alpha)$  est maintenant exclue de ce domaine  $S_r$  (figure 13) :

$$S_r := \{x \in {}^c S : \|x\| < r, \|x - x_0\| \geq \alpha\}$$

Cette exclusion a pour effet d'ajouter ceci dans le membre droite de l'identité de Green (39d) :

$$\int_{\|x-x_0\|=\alpha} \left( \left\langle \frac{x-x_0}{\|x-x_0\|} | \nabla G \right\rangle E - \left\langle \frac{x-x_0}{\|x-x_0\|} | \nabla E \right\rangle G \right) \quad (39f)$$

Comme dans le cas précédent, nous avons encore

$$\int_{\|x\|=r} \left( \frac{\partial E}{\partial r} G - \frac{\partial G}{\partial r} E \right) + \int_{S_r} E \Delta G - \int_{S_r} G \Delta E = 0$$

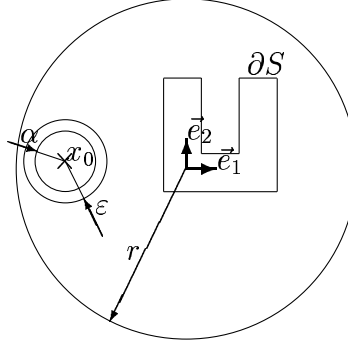


FIG. 13 – Cas  $x_0$  hors de  $S$  :  $S_r$  est délimité par le petit cercle, le grand cercle et  $\partial S$ .

Observons ensuite que, comme le reste de l'égalité (39d) est fini et indépendant de  $\alpha$ , alors le morceau (39f) est aussi fini et indépendant de  $\alpha$ . Nous pouvons donc le multiplier par  $\alpha$ , puis l'intégrer pour  $\alpha$  entre 0 et  $\varepsilon$  ce qui donne

$$(39f) = \frac{2}{\varepsilon^2} \int_{\|x-x_0\|<\varepsilon} \left( \langle x-x_0 | \nabla G(x) \rangle E(x) - \left\langle \frac{x-x_0}{\|x-x_0\|} | \nabla E(x) \right\rangle \|x-x_0\| G(x) \right) dx$$

Or on a

$$\int_{\|x-x_0\|<\varepsilon} \langle x-x_0 | \nabla G(x) \rangle E(x) dx = \int_{\|x-x_0\|<\varepsilon} \frac{2\pi}{\lambda_0} \|x-x_0\| \frac{1}{4i} H_0^{ankt} \left( \frac{2\pi}{\lambda_0} \|x-x_0\| \right) E(x) dx$$

Dès le début de la démonstration, nous supposons que la fonction maximale  $M(E)$  est finie en  $x_0$  (ceci résulterait de la remarque 2.11), ce qui s'écrit ainsi :

$$M(E)(x_0) = \limsup_{\varepsilon \rightarrow 0} \frac{1}{\pi \varepsilon^2} \int_{\|x-x_0\| \leq \varepsilon} |E(x)| dx < \infty. \quad (39g)$$

Alors le développement limité (29e) de  $H_0^{ankt}$  donne.

$$\begin{aligned} & \frac{2}{\varepsilon^2} \int_{\|x-x_0\|<\varepsilon} \langle x-x_0 | \nabla G(x) \rangle E(x) dx \\ &= \frac{1}{\pi \varepsilon^2} \int_{\|x-x_0\|<\varepsilon} E(x) dx + \left( M(E)(x_0) + O(1) \right) O\left( \varepsilon(1 + |\log \varepsilon|) \right) \end{aligned}$$

L'intégrale de droite converge vers  $E(x_0)$  lorsque  $\varepsilon \rightarrow 0$  par définition de l'hypothèse que  $x_0$  est un point de Lebesgue de  $E$ .

Nous avons aussi supposé que la fonction maximale  $M(\nabla E)$  était finie en  $x_0$ , ce qui s'écrit ainsi :

$$M(E)(x_0) = \limsup_{\varepsilon \rightarrow 0} \frac{1}{\pi \varepsilon^2} \int_{\|x-x_0\| \leq \varepsilon} |\nabla E(x)| dx < \infty. \quad (39h)$$

Le développement limité (29d) permet de conclure cette preuve :

$$\begin{aligned}
& \left| \frac{2}{\varepsilon^2} \int_{\|x-x_0\|<\varepsilon} \left\langle \frac{x-x_0}{\|x-x_0\|} \mid \nabla E(x) \right\rangle \|x-x_0\| G(x) dx \right| \\
& \leq (M(\nabla E)(x_0) + O(1)) \sup_{\|x-x_0\|\leq\varepsilon} \|x-x_0\| \frac{1}{4i} H_0^{ank} \left( \frac{2\pi}{\lambda_0} \|x-x_0\| \right) \\
& \leq (M(\nabla E)(x_0) + O(1)) O\left(\varepsilon(1 + |\log \varepsilon|)\right)
\end{aligned}$$

qui tend donc vers 0 lorsque  $\varepsilon \rightarrow 0$ .

Dans la démonstration, la preuve sur le membre droit de (39d) dépend fortement du choix de la solution fondamentale choisie. Une autre solution fondamentale radiale aurait changé le facteur  $\frac{2i\pi}{\lambda_0}$  dans (39e), et alors (30) aurait été inutilisable dans cette preuve.  $\blacksquare$

**Théorème 2.14** *Il y a unicité dans le problème de Helmholtz extérieur non intégral (21c). L'unicité reste valide lorsqu'on remplace, dans ce problème, la condition à la limite  $E(x) = E^{inc}(x)$  sur  $\partial S$  (21b) par la nullité de  $\text{Im} \int_{\partial S} \nu_1 E \overline{\nu_0 E}$  (ou même par sa négativité).*

Si on ajoutait comme condition que la dérivée normale  $\nu_1 E$  au bord de  $S$  est aussi imposée dans  $H^{-\frac{1}{2}}(\partial S)$ , alors l'unicité proviendrait immédiatement du théorème 2.13 précédent. Sans cette condition supplémentaire, nous avons besoin d'un lemme qui démontre, outre ce résultat dans le cas où  $S$  est une boule, l'impossibilité d'une solution  $H^1({}^c S)$  non nulle ( $H^1({}^c S)$  est plus fort que  $H_{loc}^1({}^c S)$ ).

**Lemme 2.15** *Dans le cas où le diffractant  $S$  est la boule  $B_R$  de rayon  $R$  centrée à l'origine, le théorème ci-dessus est vrai. La seule solution  $E$  possible, si elle existe, est donnée en coordonnées polaires  $(r, \theta)$  par*

$$E(r, \theta) = \sum_{m \in \mathbb{Z}} a_m H_{|m|}^{ank} \left( \frac{2\pi}{\lambda_0} r \right) e^{mi\theta} \quad r > R, \theta \in [0, 2\pi]. \quad (39i)$$

où les  $a_m$  sont déterminés de façon unique par les coefficients de la transformée de Fourier  $L$ -périodique de  $\nu_0 E^{inc}$  (ces derniers valent  $a_m H_{|m|}^{ank} \left( \frac{2\pi}{\lambda_0} R \right)$ ).

La décomposition ci-dessus est une décomposition en harmoniques ; c'est un peu l'analogue de la décomposition de Laurent (avec les fonctions de Hankel

joutant le rôle des puissances négatives de  $r$ ), voir Rellich [R43], d'où la démonstration suivante vient.

■ Cette démonstration est tirée de Rellich [R43], qui toutefois y faisait des hypothèses fortes sur  $E$  : il y est supposé deux fois dérivable. Ici, nous n'avons que  $E \in H_{loc}^1({}^cS)$ , même si l'analyticité dans  ${}^cS$  (mais pas au bord) est donnée par le théorème de représentation 2.13. Rellich travaille en décomposant  $E$  dans les harmoniques du plan, et montre que si  $E$  est  $L^2({}^cS)$ , alors  $E$  est nul.

Commençons par prouver l'unicité dans le cas  $S = B_R$ , premier point du lemme. Pour tout rayon  $r > R$ ,  $E$  restreint à  $\partial B_r$  (boule centrée à l'origine) est dans  $H^{\frac{1}{2}}(\partial B_r)$  qui admet une décomposition de Fourier : donc il existe des fonctions qu'on note  $H_m$  telles que

$$E(r, \theta) = \sum_{m \in \mathbb{Z}} H_m\left(\frac{2\pi}{\lambda_0} r\right) e^{mi\theta} \quad r > R, \theta \in [0, 2\pi] \quad (39j)$$

et ces fonctions sont  $L_{loc}^2({}^cS)$  d'après une inégalité de Bessel utilisant  $E \in L_{loc}^2({}^cS)$ .

Supposons que  $E$  n'est pas nul, nous allons montrer que  $\mathcal{I}m \int_{\partial S} \nu_1 E \overline{\nu_0 E}$  n'est pas nul. Pour un certain rayon  $r_0$ , la restriction de  $E$  à  $\partial B_{r_0}$  est non nulle ; et donc pour un certain  $m$ , on a  $H_m(r_0) \neq 0$ .

### La fonction $H_m$ vérifie une équation différentielle.

Soit  $g$  une fonction  $(R, \infty) \rightarrow \mathbb{C}$  régulière et à support compact. Comme  $H_m$  est un coefficient de Fourier et comme  $\Delta E + \frac{4\pi^2}{\lambda_0^2} E = 0$ , une formule de Green donne

$$\begin{aligned} & \frac{8\pi^3}{\lambda_0^2} \int_R^\infty H_m\left(\frac{2\pi}{\lambda_0} r\right) g(r) r dr \\ &= \frac{4\pi^2}{\lambda_0^2} \int_{{}^cS} E e^{-mi\theta} g(r) dx \\ &= - \int_{{}^cS} \Delta E e^{-mi\theta} g(r) dx \\ &= - \int_{{}^cS} E \Delta(e^{-mi\theta} g(r)) dx \quad (39k) \\ &= - \int_{{}^cS} E \left( e^{-mi\theta} g''(r) + \frac{1}{r} e^{-mi\theta} g'(r) - \frac{m^2}{r^2} e^{-mi\theta} g(r) \right) \\ &= -2\pi \int_R^\infty H_m\left(\frac{2\pi}{\lambda_0} r\right) \left( g''(r) + \frac{1}{r} g'(r) - \frac{m^2}{r^2} g(r) \right) r dr \\ &= - \int_R^\infty \left( \frac{8\pi^3}{\lambda_0^2} H_m''\left(\frac{2\pi}{\lambda_0} r\right) + \frac{4\pi^2}{\lambda_0} \frac{1}{r} H_m'\left(\frac{2\pi}{\lambda_0} r\right) - 2\pi \frac{m^2}{r^2} H_m\left(\frac{2\pi}{\lambda_0} r\right) \right) g(r) r dr \quad (39\ell) \end{aligned}$$

ce qui montre que  $H_m$  vérifie dans l'espace des distributions l'équation de Bessel (cas général) suivante :

$$f''(r) + \frac{1}{r}f'(r) + \left(1 - \frac{m^2}{r^2}\right)f(r) = 0 \quad (40)$$

Comme  $H_m \in L^2_{loc}([R, \infty))$ , alors  $H'_m \in H^{-1}_{loc}((R, \infty))$ , donc (40) donne  $H''_m \in H^{-1}((R, \infty))$ , donc  $H'_m \in L^2_{loc}((R, \infty))$ , donc (40) donne  $H''_m \in H^1((R, \infty))$ , et ainsi de suite ... Ainsi cette équation différentielle montre que  $H_m$  est  $H^3_{loc}((R, \infty))$ , et donc c'est une fonction  $C^2_{loc}((R, \infty))$ , qui vérifie l'équation de Bessel (40) au sens classique.

Les solutions de cette équation différentielle sont des fonctions usuelles :  $H_m$  est une combinaison linéaire de la fonction de Hankel  $H_{|m|}^{ank}$  et de la fonction conjuguée.

**La définition de  $H_m(r)$  est dérivable en  $r = R$ .**

En (39j), nous avons défini  $H_m$  par

$$H_m\left(\frac{2\pi}{\lambda_0}r\right) = \frac{1}{2\pi} \int_0^{2\pi} E(r, \theta) e^{-mi\theta} \quad r \geq R$$

et nous allons dériver cette égalité, de façon indirecte car nous ne connaissons pas assez d'hypothèse de régularité sur  $\frac{\partial}{\partial r} E|_{r=R}$ . Reprenons la longue manipulation ci-dessus, dans le cas où la fonction  $g$  contient  $R$  dans son support. À (39k), il faut ajouter

$$\int_{\partial S} \nu_1 E e^{-mi\theta} g(r) - \int_{\partial S} \nu_0 E e^{-mi\theta} g'(r) = \int_{\partial S} \nu_1 E e^{-mi\theta} g(R) - 2\pi H_m\left(\frac{2\pi}{\lambda_0}R\right) g'(R) R$$

et à (39l) il faut de plus ajouter  $2\pi$  fois

$$H_m\left(\frac{2\pi}{\lambda_0}R\right) g'(R) R - \frac{2\pi}{\lambda_0} H'_m\left(\frac{2\pi}{\lambda_0}R\right) g(R) R - H_m\left(\frac{2\pi}{\lambda_0}R\right) g(R) + H_m\left(\frac{2\pi}{\lambda_0}R\right) g(R)$$

ce qui fait qu'au membre gauche de (40) il faut ajouter

$$\frac{\lambda_0^2}{4\pi^2} \delta\left(r - \frac{\lambda_0}{2\pi}R\right) \left(\frac{4\pi^2}{\lambda_0} f'(r) r - \int_0^{2\pi} \nu_1 E(R, \theta) e^{-mi\theta} d\theta\right).$$

où  $\delta$  est la masse de Dirac. Nous avons donc montré que

$$\frac{2\pi}{\lambda_0} H'_m\left(\frac{2\pi}{\lambda_0}R\right) = \frac{1}{2\pi} \int_0^{2\pi} \nu_1 E(R, \theta) e^{-mi\theta} d\theta \quad (41a)$$

Bien sûr, tout ceci marche aussi si on restreint  $E$  à  ${}^c B_r$  au lieu de le restreindre à  ${}^c B_R$ , ce qui donne (41a) mais avec  $r$  au lieu de  $R$  et  $\frac{\partial}{\partial r} E(r, \theta)$  au lieu de  $\nu_1 E(R, \theta)$ .

**La fonction  $H_m$  vérifie une condition de rayonnement normalisée.**

La condition de Sommerfeld (30) de  $E$  permet de montrer que  $H_m$  vérifie la condition radiale de rayonnement de Sommerfeld (31), car on a :

$$\left| H'_m\left(\frac{2\pi}{\lambda_0}r\right) - iH_m\left(\frac{2\pi}{\lambda_0}r\right) \right| = \left| \int_0^{2\pi} \left( \frac{\lambda_0}{2\pi} \frac{\partial}{\partial r} E(r,\theta) - iE(r,\theta) \right) e^{-mi\theta} \right|$$

et ensuite une inégalité de Cauchy-Schwarz permet d'utiliser (30), ce qui montre la condition de rayonnement radiale normalisée (31).

**La fonction  $H_m$  est une fonction de Hankel.**

La fonction de Hankel  $H_{|m|}^{ank}$  admet le comportement asymptotique suivant à l'infini, donné par Watson [W22] p197 :

$$H_m^{ank}(r) = \frac{1-i}{i^m} e^{ir} / \sqrt{\pi r} + O(r^{-\frac{3}{2}}) \quad (42)$$

$$H_m^{ank'}(r) = \frac{1+i}{i^m} e^{ir} / \sqrt{\pi r} + O(r^{-\frac{3}{2}}) \quad (43)$$

Comme  $H_m$  vérifie la condition radiale normalisée de Sommerfeld (31), on voit que  $H_m$  est proportionnel à  $H_{|m|}^1$ . Posons  $a_m = \frac{H_m(R)}{H_{|m|}^1(R)}$ . Comme  $H_m(r_0) \neq 0$  ( $m$  et  $r_0$  ont été choisis spécifiquement pour assurer cette non-nullité), on en déduit que  $H_m(R) \neq 0$ , et que donc  $a_m \neq 0$ .

**Une base de fonctions harmoniques.**

Posons

$$a_m = \int_0^{2\pi} \nu_0 E(\theta) e^{mi\theta} d\theta H_{|m|}^{ank}\left(\frac{2\pi}{\lambda_0}R\right)^{-1} < \infty$$

Observons que  $a_m = \frac{H_m(R)}{H_{|m|}^1(R)}$ . Ces nombres  $a_m$  satisfont (39i) d'après la définition de  $H_m$ . Nous avons vu au paragraphe précédent que l'un des  $a_m$  était non nul. L'égalité de Parseval nous montre que

$$\sum_{m \in \mathbb{Z}} (m+1) |a_m H_{|m|}^{ank}\left(\frac{2\pi}{\lambda_0}R\right)|^2 = \|\nu_0 E\|_{H^{\frac{1}{2}}(\partial S)} < \infty$$

Lorsqu'on multiplie (41a) par  $\overline{a_m H_{|m|}^{ank}\left(\frac{2\pi}{\lambda_0}R\right)}$ , on obtient

$$\frac{1}{2\pi} \int_0^{2\pi} \nu_1 E \overline{a_m H_{|m|}^{ank}\left(\frac{2\pi}{\lambda_0}R\right)} e^{mi\theta} d\theta = \frac{2\pi}{\lambda_0} |a_m|^2 H_{|m|}^{ank'}\left(\frac{2\pi}{\lambda_0}R\right) \overline{H_{|m|}^{ank}\left(\frac{2\pi}{\lambda_0}R\right)} \quad (44a)$$



Grâce à l'égalité de Parseval, la somme sur  $m$  dans l'argument de la conjugaison à gauche du signe égal converge dans  $H^{\frac{1}{2}}(\partial S)$  vers  $\nu_0 E$ , donc la somme sur  $m$  de l'intégrale de droite converge absolument vers  $\frac{1}{2\pi R} \int_{\partial S} \nu_1 E \overline{\nu_0 E}$ . On en déduit que la somme sur  $m \in \mathbb{Z}$  de (44a) converge, et donne

$$\frac{1}{2\pi} \int_0^{2\pi} \nu_1 E \overline{\nu_0 E} d\theta = \sum_{m \in \mathbb{Z}} \frac{2\pi}{\lambda_0} |a_m|^2 H_{|m|}^{ank'} \left( \frac{2\pi}{\lambda_0} R \right) \overline{H_{|m|}^{ank} \left( \frac{2\pi}{\lambda_0} R \right)}.$$

**Stricte positivité de  $\mathcal{I}m \int_{\partial S} \nu_1 E \overline{\nu_0 E}$ .**

Comme l'un des  $a_m$  est non nul, il suffirait de montrer que

$$r \mathcal{I}m \left( H_{|m|}^{ank'}(r) \overline{H_{|m|}^{ank}(r)} \right) = \frac{2}{\pi}. \quad (44b)$$

Or  $H_{|m|}^{ank}$  vérifie (40), ce qui montre que la dérivée du membre gauche de (44b) vaut

$$\mathcal{I}m \left( H_{|m|}^{ank'}(r) \overline{H_{|m|}^{ank}(r)} \right) - \frac{r}{r} \mathcal{I}m \left( H_{|m|}^{ank'}(r) \overline{H_{|m|}^{ank}(r)} \right) - r \left( 1 - \frac{m^2}{r^2} \right) \mathcal{I}m \left( |H_{|m|}^{ank}(r)|^2 \right)$$

ce qui est nul. Les équivalents (42) et (43) montrent alors (44b). Alors, le membre de droite de (44a) est une somme de termes positifs, et l'un de ces termes est même strictement positif: ainsi  $\mathcal{I}m \int_{\partial S} \nu_1 E \overline{\nu_0 E}$  est strictement positif.

Cette unicité nous a permis, au passage, de construire  $E$  d'après la transformée de Fourier de  $\nu_0 E$ ; les conditions nécessaires montrées ci-dessus nous permettent de montrer que la solution est nécessairement de la forme (39i). Pour montrer l'existence pour ce cas où  $S$  est un disque, il suffirait maintenant de vérifier que cette forme est une fonction  $H_{loc}^1({}^c S)$  vérifiant l'équation de Helmholtz (20), et la condition de Sommerfeld (30). ■

■ Maintenant, nous montrons le théorème 2.14, c'est à dire que nous travaillons pour  $S$  différent d'une boule. Pour prouver l'unicité, il suffit de raisonner pour  $\nu_0 E$  imposé nul, qui est à son tour une hypothèse plus forte que l'hypothèse alternative  $\mathcal{I}m \int_{\partial S} \nu_1 E \overline{\nu_0 E} = 0$ . Il suffit donc de raisonner pour la seule hypothèse  $\mathcal{I}m \int_{\partial S} \nu_1 E \overline{\nu_0 E} \leq 0$ , couplée à la condition de Sommerfeld (30).

Soit  $R > 0$  tel que le diffractant  $S$  soit inclus dans la boule ouverte  $B_R$  de rayon  $R$  centrée à l'origine (voir la figure 12, sauf que  $r$  est maintenant noté  $R$ ). Comme  $\overline{E}$  est aussi une solution  $H_{loc}^1({}^c S)$  en dehors de  $S$  qui est solution de l'équation de Helmholtz (20), nous pouvons la substituer à  $H$  dans la

formule de Stokes (36). En prenant la partie imaginaire de cette formule de Stokes, on obtient la négativité de la partie imaginaire

$$\mathcal{I}m \int_{\|x\|=r} \frac{\partial E}{\partial r} \overline{E}.$$

Alors  $E$  restreint à  ${}^cB_R$  satisfait toutes les hypothèses du lemme 2.15 ci-dessus, donc nous avons la nullité de  $E$  sur  ${}^cB_R$ . D'après le théorème de représentation 2.13,  $E$  hérite de  $G$  l'analyticité sur l'ouvert  ${}^cS$ :  $E$  est donc nulle sur  ${}^cS$  tout entier, il y a contradiction. ■

Nous avons ce corollaire immédiat :

**Propriété 2.16 (Rellich)** *L'équation (20) n'a qu'une seule solution  $H^1({}^cS)$  qui vérifie la condition de rayonnement de Sommerfeld (30) : la solution nulle.*

Notons que ce résultat reste vrai sans condition de Sommerfeld, c'est le théorème de Rellich.

■ Procédons à un raisonnement par l'absurde, en supposant  $E$  non nul. L'expression (39b) a pour limite 0 selon la condition de Sommerfeld ; le théorème d'unicité 2.14 ci-dessus montre que le dernier terme de (39b) ne peut pas être nul. Ainsi, ceci montre que

$$\int_{\|x\|=r} \left| \frac{\partial E}{\partial r} \right|^2 + \int_{\|x\|=r} \frac{4\pi^2}{\lambda_0^2} |E|^2$$

converge vers une valeur non nulle lorsque  $r$  tend vers l'infini : ceci interdit à  $E$  d'être  $H^1({}^cS)$  globalement. ■

Nous sommes maintenant prêts pour changer d'inconnue : au lieu de travailler sur  $E$ , nous allons travailler sur  $\nu_0 E$  et  $\nu_1 E$ . Ainsi, comme (38) sera automatiquement une solution de l'équation de Helmholtz (20) quelles que soient les valeurs de  $\nu_0 E$  et  $\nu_1 E$ , il n'y aura plus qu'à s'assurer que les conditions aux limites (à l'infini (30) et au bord de l'objet) sont bonnes.

Pour un diffractant sans coin, Bamberger et Ha-Duong [BH87] ont construit une application sesquilinéaire basée sur les quatre opérateurs

$$\begin{aligned} J &\mapsto \nu_0(J \star G), \\ J &\mapsto \nu_0\left((Js'(\cdot)^\perp) \star \nabla G\right), \\ J &\mapsto \nu_1(J \star G), \end{aligned}$$

$$J \mapsto \nu_1 \left( (J s'(\cdot)^\perp) \star \nabla G \right), \quad (44c)$$

(qu'on appelle souvent les opérateurs simple couche, double couche, double couche adjoint, et hypersingulier), qui est coercive, et qui donne le couple  $(\nu_0 E, \nu_1 E)$  via une formule variationnelle.

Nous préférons changer encore d'inconnue, pour ne plus avoir qu'une fonction, au lieu de deux, et pour pouvoir être compatible avec des coins dans le diffractant. Nous n'obtiendrons pas de coercivité, mais néanmoins nous obtiendrons une inégalité de Gårding qui la remplacera efficacement, voir le théorème 3.1.

## 2.2 Le problème interne au diffractant, et la nouvelle inconnue $J$ .

Le théorème de représentation 2.13 était valide pour une fonction sur  ${}^c S$ . Néanmoins, il est tentant de rechercher la solution  $E$  sous la forme  $E = J \star G$  où  $J$  est une fonction à support  $\partial S$ : en effet, la solution fondamentale  $G$  (28) garantit que  $J \star G$  vérifie l'équation de Helmholtz aussi bien dans  ${}^c S$  (20) que dans  $S$  (21d).

Le théorème d'unicité 2.14 et la condition au bord (35c) suggèreraient naturellement la recherche d'une fonction  $J$  à support  $\partial S$  dont la valeur au bord de  $S$  vérifie

$$\nu_0 (J \star E) = -\nu_0 E^{inc} \quad \text{sur } \partial S.$$

Seulement, le théorème d'unicité 2.14 concerne la restriction de  $E = J \star G$  à  ${}^c S$  et non la fonction  $J$ : et en effet, si la longueur d'onde  $\lambda_0$  correspond à un des modes propres du problème de Helmholtz intérieur (21d) dans  $S$  (et il existe de tels modes propres), alors il existe des  $J$  non nuls tels que  $J \star G$  est nul sur  $\partial S$ , et même sur  ${}^c S$ ; de plus, lorsque la longueur d'onde  $\lambda_0$  est proche d'une longueur d'onde  $\lambda$  qui correspond à un des modes propres, Amini-Kirkup [AK95] montrent que le nombre de conditionnement souffre d'un facteur  $|\lambda_0 - \lambda|^{-1}$ .

Pour éviter cet écueil, nous devons mieux étudier le problème de Helmholtz intérieur (21d) dans  $S$ , ce qui nous permettra d'introduire une seconde condition au bord, (48i), du côté interne de  $\partial S$ .

Le théorème de représentation 2.13 était valide pour une fonction définie sur  ${}^c S$ . Le théorème ci-dessous est son analogue pour  $S$  au lieu de  ${}^c S$ . Ainsi, nous verrons que si  $E = J \star G$ , alors

$$J = \nu_1 E + \nu_1^i E$$

où  $\nu_1^i E$  est la dérivée normale sur  $\partial S$  de  $E$  au bord intérieur de  $S$ .

**Théorème 2.17 (représentation)** *Lorsque  $E \in H^1(S)$  vérifie l'équation de Helmholtz (21d),  $E$  admet une dérivée normale sur  $\partial S$  notée  $\nu_1^i E$  qui est  $H^{-\frac{1}{2}}(\partial S)$  (et dépend continûment de  $E \in H^1(S)$ ), et  $E$  vérifie presque partout l'égalité suivante :*

$$\nu_1^i E \star G - (s'^{\perp}(\cdot)\nu_0^i E) \star \nabla G = \mathbb{1}_S E \quad (45)$$

où  $G$  est la solution fondamentale (28).

La version tridimensionnelle de théorème a par exemple été montrée par Colton-Kress [CK92], avec les mêmes restrictions que pour leur version du théorème 2.13.

■ Pour l'existence de  $\nu_1^i$ , c'est exactement la même démonstration que pour 2.12, sauf que la démonstration est beaucoup plus simple dans un domaine  $S$  qui est compact, contrairement au domaine  ${}^c S$ . La formule de Stokes (36) devient alors

$$- \int_{\partial S} \nu_0^i E \nu_1^i G = \int_S E \Delta G + \int_S \langle \nabla E | \nabla S \rangle \quad (46a)$$

qui est vraie pour  $G$  une fonction  $H^1(S)$ .

Soit  $x_0 \in \mathbb{R}^2 \setminus \partial S$ . Nous allons montrer (45) en  $x = x_0$ . Nous allons raisonner uniquement pour les  $x_0$  où les fonctions maximales  $M(E)$  et  $M(\nabla E)$  sont finies, et qui sont des points de Lebesgue de  $E$ . Notons que ces finitudes seraient immédiates d'après la remarque 2.11 et son analogue pour  $S$  au lieu de  ${}^c S$ .

**Cas  $x_0$  hors de  $S$ .**

Considérons la solution fondamentale  $G(x) = -\frac{i}{4} H_0^{ank}(\frac{2\pi}{\lambda_0} \|x - x_0\|)$  vue en (28). La fonction  $G$  est  $H^1(S)$ . Alors la formule de Stokes (46a) est valide. De plus,  $E$  et  $G$  y ont un rôle symétrique car  $G$  est une solution de l'équation de Helmholtz (21d) sur  $S$ , donc on a aussi la formule de Stokes (46a) mais avec  $E$  et  $G$  échangés. En soustrayant les deux formules de Stokes, nous obtenons

$$- \int_{\partial S} (\nu_0^i E \nu_1^i G - \nu_0^i G \nu_1^i E) = \int_S (E \Delta G - G \Delta E) \quad (46b)$$

Or  $E$  et  $G$  vérifient tous deux l'équation de Helmholtz (21d) dans  $S$ : le membre droit ci-dessus est donc nul, et c'est tout ce qu'il nous fallait pour montrer (45) en  $x_0$ .

**Cas  $x_0$  dans  $S$ .**

Pour obtenir une identité de Green similaire à (46b), reprenons la formule de Stokes (46a), sauf que nous excluons du domaine d'intégration  $S$  une boule  $B(x_0, \varepsilon)$  pour un certain  $\varepsilon$  tel que cette boule est incluse dans l'intérieur de  $S$ . Cette formule de Stokes devient alors

$$-\int_{\partial S} \nu_0^i E \nu_1^i G - \int_{\|x-x_0\|=\varepsilon} E(x) \left\langle \frac{x-x_0}{\|x-x_0\|} \middle| \nabla G(x) \right\rangle dx = \int_{S \setminus B(x_0, \varepsilon)} (E \Delta G + \langle \nabla E | \nabla S \rangle)$$

Par symétrie, nous avons aussi la même formule sauf que  $E$  et  $G$  sont échangés. Lorsqu'on soustrait ces deux formules, nous obtenons (46b) sauf qu'à son membre droit il faut ajouter (39f).

La compensation faite dans le membre droit de (46b) marche encore ici, et l'estimation de (39f) marche elle aussi ici, ce qui donne (45) en  $x_0$ . ■

La remarque ci-dessous est basée sur la somme de cette seconde formule de représentation (45) avec la première (38) afin d'éliminer le terme  $\nu_0 E$  : ceci aura le grand avantage de diminuer le nombre d'inconnue, afin d'atteindre plus facilement un schéma numérique avec existence et unicité de la fonction  $J$ . Cette remarque est citée, sans sa démonstration, pour illustration néanmoins nous ne l'utiliserons pas :

**Remarque 2.18** *Il existe au moins une solution  $E \in H_{loc}^1({}^c S)$  au problème de Helmholtz extérieur non intégral (21c) où  $E^{inc} \in H_{loc}^1(\mathbb{R}^2)$ . La solution  $E \in H_{loc}^1({}^c S)$  dépend continûment et linéairement du paramètre  $E^{inc} \in H_{loc}^1(\mathbb{R}^2)$ .*

*La solution  $E$  est déconvolvable par  $G$ , et on a précisément*

$$(\nu_1 E - \nu_1^i E^{inc}) \star G = \mathbb{1}_{{}^c S} E - \mathbb{1}_S E^{inc} \quad (47)$$

où  $\nu_1$  est la dérivée externe et  $\nu_1^i$  la dérivée normale intérieure.

La démonstration de cette remarque se trouve par exemple dans Nédélec [N01] (mais il faut passer de la dimension 3 à la dimension 2, passage dont la validité a été vérifiée, même si elle est hors du sujet de cette thèse). Elle commence par le cas où  $S$  est la boule, où le raisonnement se fait dans les harmoniques cylindriques, comme dans le lemme 2.15.

La remarque 2.18 ci-dessus suggère, pour trouver  $E$ , de chercher d'abord  $\nu_1 E + \nu_1^i E^{inc}$  comme une fonction  $H^{-\frac{1}{2}}(\partial S)$  qu'on appellera  $J$ . Numériquement,  $J$  est une inconnue intermédiaire, qui a par exemple le gros mérite sur  $E$  d'être à support compact, et donc numériquement directement utilisable.

Les méthodes résolvant directement le problème de Helmholtz extérieur (21c) se heurtent à une difficulté théorique très sérieuse : la résolution numérique ne peut s'effectuer que dans un domaine borné. Si on résout le problème de Helmholtz extérieur, le domaine de résolution est non-borné, donc il doit être tronqué avant d'être traité numériquement. Cette troncature est réalisée en imposant des conditions sur la frontière extérieure du domaine de résolution qui permettent de « simuler » un domaine infini. Néanmoins, ces conditions restent approchées et des réflexions purement numériques sur la frontière extérieure apparaissent. De plus, la complexité est en  $\lambda_0^{-2}$  où  $\lambda_0$  est la longueur d'onde.

Les méthodes intégrales consistent à calculer (21h) les courants électriques  $J$  sur le bord  $\partial S$  de l'objet diffractant  $S$ . Nous montrerons au théorème 2.30 que c'est un problème équivalent au calcul du champ électrique  $E$  dans le domaine extérieur infini. À partir du courant  $J$ , le champ électrique  $E$  peut être recalculé dans toute partie du domaine extérieur avec  $E = J \star G$ , et l'équation de Helmholtz et les conditions de rayonnement à l'infini sont automatiquement satisfaites par  $E$ .

Comme  $\nu_1 E$  et  $\nu_1^i E$  sont  $H^{-\frac{1}{2}}(\partial S)$  d'après le théorème de représentation 2.17 et la propriété 2.12, nous voyons que  $J \in H^{-\frac{1}{2}}(\partial S)$  est un choix logique ; la propriété ci-dessous renforce ce choix :

**Propriété 2.19** *Si  $J \in H^{-\frac{1}{2}}(\partial S)$ , alors  $J \star G$  est une solution  $H_{loc}^1({}^c S)$  (dépendant continûment de  $J$ ) vérifiant (sur son domaine  ${}^c S$ ) l'équation de Helmholtz (20) et la condition de rayonnement de Sommerfeld (30).  $J \star G$  est aussi  $H^1(S)$  et vérifie sur  $S$  l'équation de Helmholtz (21d).*

*De plus, si  $J \in L^2(\partial S)$ , alors  $J \star G$  est continu.*

■ Nous allons utiliser, comme intermédiaire, la fonction

$$E : x \mapsto \int_0^L J(y) \frac{1}{2\pi} \log(\|s(y) - x\|) dy. \quad (48a)$$

**Montrons que  $E = J \star \frac{1}{2\pi} \log \|\cdot\|$  est dans  $H_{loc}^1(\mathbb{R}^2)$  et que  $\Delta E = 0$  sur  ${}^c \partial S$ .**

Supposons d'abord que  $J$  est dans  $L^2(\partial S)$ . On voit que cette fonction  $E$  est dans  $L_{loc}^\infty({}^c S)$ , car  $J \in L^2(\partial S)$  et le reste de l'intégrale est aussi  $L^2(\partial S)$ .

Soit  $f \in \mathcal{D}(\mathbb{R}^2)$ , montrons que  $|\int_{\mathbb{R}^2} E(x)\Delta f(x)dx| \leq C\|f\|_{\mathbf{H}^1(\mathbb{R}^2)}$  :

$$\begin{aligned}
& \int E(x)\Delta f(x)dx \\
&= \int_{x \in \text{supp}f} \Delta f(x) \int_0^L J(t) \frac{1}{2\pi} \log \|s(t) - x\| dt dx \\
&= \int_0^L J(t) \int_{x \in \text{supp}f} \Delta f(x) \frac{1}{2\pi} \log \|s(t) - x\| dx dt \\
&= \int_0^L J(t) f(s(t)) dt
\end{aligned} \tag{48b}$$

où le lemme 2.2 sur la solution fondamentale est la raison de la dernière inégalité. La définition dans la propriété 2.12 de la norme  $\mathbf{H}^{-\frac{1}{2}}(\partial S)$  donne alors

$$\left| \int E(x)\Delta f(x)dx \right| \leq \|J\|_{\mathbf{H}^{-\frac{1}{2}}(\partial S)} \|\nu_0 f\|_{\mathbf{H}^{\frac{1}{2}}(\partial S)} \leq C \|J\|_{\mathbf{H}^{-\frac{1}{2}}(\partial S)} \|f\|_{\mathbf{H}^1(\mathbb{R}^2)}$$

où la constante  $C$  vient de la propriété 2.10, et ne dépend que de  $S$ . Ainsi,  $E$  est dans  $\mathbf{H}^{-1}(\mathbb{R}^2)$  lorsque  $J \in \mathbf{L}^2(\partial S)$ .

La nullité de  $\Delta E$  est claire d'après le résultat de Fubini ci-dessus.

Nous devons prolonger ce résultat à  $J \in \mathbf{H}^{-\frac{1}{2}}(\partial S)$  au lieu de  $J \in \mathbf{L}^2(\partial S)$ . Pour cela, nous remarquons que l'application  $J \mapsto E$ , telle que définie par (48a), est continue de  $\mathbf{H}^{-\frac{1}{2}}(\partial S)$  dans l'espace des distributions à support  $K$ , où  $K$  est un compact de  $\mathbb{R}^2$  contenant  $S$ . Comme  $\mathbf{L}^2(\partial S)$  est dense dans  $\mathbf{H}^{-\frac{1}{2}}(\partial S)$ , l'estimation ci-dessus se généralise lorsque  $\text{supp}f \subset K$  par continuité à tout  $J \in \mathbf{H}^{-\frac{1}{2}}(\partial S)$ , ce qui montre que l'application  $J \mapsto E$  est continue de  $\mathbf{H}^{-\frac{1}{2}}(\partial S)$  dans  $\mathbf{H}^1(K)$ . Ainsi, l'application  $J \mapsto E$  est continue  $\mathbf{H}^{-\frac{1}{2}}(\partial S) \rightarrow \mathbf{H}_{loc}^1(\mathbb{R}^2)$ , et est à valeur dans l'ensemble des fonctions à Laplacien nul sur  ${}^c\partial S$ .

### Formule de représentation intérieure pour $E$ , lorsque $\Delta E = 0$ .

Grâce à la nullité de ce Laplacien, nous avons la formule de représentation (cas du Laplacien) suivante :

$$\nu_1^i E \star \left( \frac{1}{2\pi} \log \|\cdot\| \right) - (s'^{\perp}(\cdot) \nu_0^i E) \star \nabla \left( \frac{1}{2\pi} \log \|\cdot\| \right) = \mathbb{1}_S E. \tag{48c}$$

Il suffit de reprendre mot pour mot la démonstration du théorème 2.17 analogue mais pour l'opérateur de Laplace au lieu de l'opérateur de Helmholtz : en effet, l'identité de Green fournit facilement ce résultat.

**Formule de représentation extérieure pour  $E$ , avec  $\Delta E = 0$ .**

Nous allons montrer la formule de représentation analogue, qui est

$$\nu_1 E \star \left( \frac{1}{2\pi} \log \|\cdot\| \right) + (s'^{\perp}(\cdot) \nu_0 E) \star \nabla \left( \frac{1}{2\pi} \log \|\cdot\| \right) = \mathbb{1}_{c_S} E. \quad (48d)$$

Soit  $x_0$  un point qui n'est pas sur  $S$ . La démonstration du théorème 2.13 marche bien ici, excepté pour le premier terme de (39d) qui est ici

$$\int_{\|x\|=r} \left( \frac{1}{2\pi} \log \|x - x_0\| \frac{\partial E}{\partial r} - E \frac{1}{2\pi} \frac{\partial \log \|x - x_0\|}{\partial r} \right) \quad (48e)$$

et pour lequel il faut trouver une nouvelle majoration, en l'absence de condition de rayonnement à l'infini. Pour prouver notre formule de représentation extérieure, il faut montrer que (48e) tend vers 0 lorsque  $r$  tend vers l'infini. Nous y arriverons ci-dessus en faisant un développement limité avec reste intégral en  $s(t)$  voisin de  $x_0$ , puis nous montrerons que (48e) contient l'intégrale de la fonction  $J \in H^{-\frac{1}{2}}$  contre une fonction Lipschitzienne dont la norme tend vers 0 lorsque  $r$  tend vers l'infini.

**Un développement de Taylor à reste intégral de (48e).**

En remplaçant la valeur de  $E$  par sa définition (48a), un Fubini facile ( $x_0$  n'est pas sur  $\partial S$ , donc le log et ses dérivées sont uniformément Lipschitziens) montre que (48e) vaut :

$$\int \frac{J(t)}{4\pi^2} \int_{\|x\|=r} \log \|x - x_0\| \frac{\partial \log \|x - s(t)\|}{\partial r} - \log \|x - s(t)\| \frac{\partial \log \|x - x_0\|}{\partial r} dt$$

Or l'intégrande interne vaut en fait :

$$\int_0^1 \frac{\partial}{\partial t'} \left( \log \|x - x_0 + (1-t')(x_0 - s(t))\| \frac{\partial \log \|x - x_0 + t'(x_0 - s(t))\|}{\partial r} \right) dt'$$

Comme  $s$  est une paramétrisation par abscisse curviligne, la norme Lipschitzienne de l'intégrande de cette intégrale en  $t'$  est plus petite que

$$C \left( 1 + \left| \log \sup_t \|s(t) - x\| \right| + \left| \log \|x_0 - x\| \right| \right) \frac{\left( \|x - x_0\| + \sup_t \|s(t) - x_0\| \right)^4}{\left( \|x - x_0\| - \sup_t \|s(t) - x_0\| \right)^6}$$

avec  $C$  indépendant de  $t$  et de  $t'$  (notons que  $\nabla^3(\log \|z\|) \sim (\mathcal{R}e \log z)''' = \mathcal{R}e \frac{1}{6} z^{-3}$ ). La puissance 4 provient du fait que la dérivation selon  $t$  peut ne dériver qu'un terme en  $x_0 - s(t)$  ce qui ne fait pas apparaître de puissance supplémentaire en  $\frac{1}{\|x-x_0\|}$ .



On en déduit que (48e) est l'intégrale de  $J$  contre un élément dont la norme Lipschitzienne est en  $O(\frac{r}{r^2})$ . Comme l'ensemble des fonctions Lipschitziennes  $C^{0,1}(\partial S)$  est continûment inclus dans  $H^{\frac{1}{2}}(\partial S)$ , on en déduit que (48e) est en  $\frac{1}{r}$ . Donc (48e) tend vers 0 lorsque  $r \rightarrow \infty$ .

### Égalité des valeurs au bord de $E$ .

Nous allons établir que les valeurs au bord interne  $\nu_0^i E$  et externe  $\nu_0 E$  sont égales. Comme ces deux valeurs dépendent continûment de  $E \in H_{loc}^1({}^c S)$  et donc de  $J \in H^{-\frac{1}{2}}(\partial S)$ , il suffit de prouver l'égalité

$$\nu_0^i E = \nu_0 E \quad (48f)$$

c'est à dire

$$\nu_0^i(J \star G) = \nu_0(J \star G).$$

pour  $J \in L^\infty(\partial S)$  seulement, par densité de  $L^\infty(\partial S)$  dans  $H^{-\frac{1}{2}}(\partial S)$  (on aurait même pu remplacer  $L^\infty$  par  $L^p$  avec  $p > 1$ , voir Gaier [G76]; pour  $p = 2$ , en admettant l'ellipticité du Laplacien il y a une méthode simple basée sur  $[J \star G] = [J \star \frac{1}{2\pi} \log \|\cdot\|] = [\Delta^{-1}(Jd\sigma)]$  qui est nul car  $Jd\sigma \in H^{-\frac{1}{2}-\varepsilon}$  et  $H^{-\frac{1}{2}-\varepsilon} \subset C^{\frac{1}{2}-\varepsilon}$ ). Or, dans ce cas, on peut montrer la continuité au voisinage d'un point  $x_0 \in \partial S$ . Il existe  $\varepsilon > 0$  tel que la boule de centre  $x_0$  et de rayon  $\varepsilon$  ne contient qu'un coin au plus du diffractant, et que le morceau de diffractant inclus dans cette boule soit connexe.

Alors, pour tout point  $x$  de la boule  $B(x_0, \frac{1}{2}\varepsilon)$  et pour tout  $r < \frac{1}{2}\varepsilon$ , la portion  $S_{x,r}$  de  $\partial S$  intersectant  $B(x,r)$  (représentée en gras figure 14) est plus petite que  $Cr$ , avec  $C$  une constante. Quitte à augmenter cette constante, on peut supposer cette estimation vraie même lorsque  $r$  est plus grand.

Prenons deux points  $x$  et  $y$  dans la boule  $B(x_0, \frac{1}{8}\varepsilon)$ . Notons  $r$  le double de la distance entre ces deux points. Soit  $R$  la distance maximale entre  $x$ ,  $y$  et  $\partial S$  (voir figure 14).

Il suffit de montrer que ceci est petit :

$$|E(x) - E(y)| \leq \int |J(t)| \left| \log \|s(t) - x\| - \log \|s(t) - y\| \right| dt$$

Le morceau correspondant à  $\partial S \setminus (S_{x,r} \cup S_{y,r})$  est petit car

$$\begin{aligned} & \int_{\partial S \setminus (S_{x,r} \cup S_{y,r})} \left| \log \|s(t) - x\| - \log \|s(t) - y\| \right| dt \\ &= \sum_{0 \leq k \leq \frac{\log R - \log r}{\log 2}} \int_{t: 2^k r \leq \|s(t) - x\| \leq 2^{k+1} r} \left| \log \|s(t) - x\| - \log \|s(t) - y\| \right| dt \end{aligned}$$

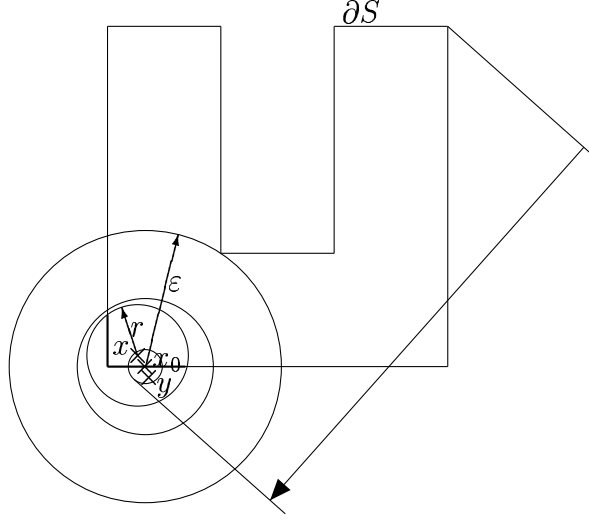


FIG. 14 – *En gras*,  $S_{x,r}$ .

$$\leq \sum_{0 \leq k \leq \frac{\log R - \log r}{\log 2}} \frac{C2^{k+1}r\|x - y\|}{2^k r} \leq 2Cr \frac{\log R - \log r}{\log 2}$$

et le morceau restant est plus petit que

$$2 \int_{S_{x,r}} |\log \|s(t) - x\|| dt + 2 \int_{S_{y,r}} |\log \|s(t) - y\|| dt$$

qui sont tous deux petits pour la raison similaire que

$$\begin{aligned} & \sum_{-\infty \leq k < 0} C2^{k+1}r(|\log(2^k r)| + |\log(2^{k+1} r)|) \\ & \leq C(1 + 2 \log 2)2r \log 2 + 4Cr(|\log r| + |\log 2r|) \end{aligned}$$

Ainsi,  $|E(x) - E(y)|$  tend vers 0 lorsque  $\|x - y\|$  tend vers 0 : nous avons prouvé la continuité au voisinage de  $x_0$ , ce pour  $x_0$  un point quelconque de  $\partial S$ . Ainsi, (48f) est vrai.

Kress [K89] théorème 6.13 montre une continuité plus forte du type  $C^\alpha$  lorsque  $J \in L^\infty$  (et dans la parenthèse ci-dessus nous l'avons montré avec  $J \in L^2$ ). Mais la seule continuité qui passe par densité à  $J \in H^{-\frac{1}{2}}(\partial S)$  est l'égalité (48f) dans  $H^{\frac{1}{2}}(\partial S)$ .

**Rapport entre  $J$  et les dérivées au bord de  $E = J \star \frac{1}{2\pi} \log \|\cdot\|$ .**

Nous avons maintenant les deux formules de représentations (48c) et (48d). En les additionnant, et comme on a  $\nu_0^i E = \nu_0 E$ , il vient

$$(\nu_1 E + \nu_1^i E) \star \left(\frac{1}{2\pi} \log \|\cdot\|\right) = E.$$

Or on a  $E = J \star \left(\frac{1}{2\pi} \log \|\cdot\|\right)$  par définition : ainsi, on a

$$J = \nu_1 E + \nu_1^i E.$$

avec  $E$  une fonction  $H^1(\mathbb{R}^2)$  de Laplacien nul. Ce résultat est une pseudo-surjectivité de l'application  $\nu_1$  (c.f. propriété 2.12). Pour la vraie surjectivité, nous aurions besoin de l'existence du problème de Helmholtz intérieur (21d) dans  $S$ , ce qui implique que  $\frac{4\pi^2}{\lambda_0}$  ne doit pas être valeur propre du Laplacien sur  $S$ .

Cette pseudo-surjectivité nous permet tout de même de mieux convoler  $H^{-\frac{1}{2}}(\partial S)$  avec une fonction régulière :

**Montrons que  $J \star G - E$  est dans  $H_{loc}^1(\mathbb{R}^2)$ .**

Soit  $F = G - \frac{1}{2\pi} \log \|\cdot\|$ . Soit  $K$  un compact.

Soit  $H$  une fonction  $H^1(\mathbb{R}^2)$  à support compact. On voit d'après les développements limités (29d) et (29e) que  $F \in C^1$ . Il est alors clair que  $H \star F \in H^1(K)$ . Fubini suivi des deux formules de Stokes (36) et (46a) donne :

$$\begin{aligned} \int_{\mathbb{R}^2} (J \star F) H &= \int_{\mathbb{R}^2} \int_0^L H(x) J(s(t)) F(s(t) - x) dt dx \\ &= \int_{\partial S} (H \star F) J \\ &= \int_{\partial S} (H \star F) \nu_1 E + \int_{\partial S} (H \star F) \nu_1^i E \\ &= \int_{\|x\|=r} (H \star F) \frac{\partial E}{\partial r} - \int_{\mathbb{R}^2} \langle \nabla E | \nabla (H \star F) \rangle - \int_{\mathbb{R}^2} (H \star F) \Delta E \end{aligned}$$

Comme  $\Delta E = 0$ , le dernier terme disparaît. En intégrant pour  $r$  entre deux valeurs comme pour la fin de la démonstration de la propriété 2.12, on montre que

$$\left| \int_{\mathbb{R}^2} H(J \star F) \right| \leq C \|E\|_{H^1(cS)} \|H \star F\|_{H^1(cS)} \leq C \|E\|_{H^1(K)} \|H\|_{H^1(\mathbb{R}^2)}$$

avec  $E \in H^1(K)$  dépendant continûment de  $J \in H^{-\frac{1}{2}}(\partial S)$ . Ceci montre que  $J \star F \in H_{loc}^1(cS)$  est vrai continûment en  $J \in H^{-\frac{1}{2}}(\partial S)$ .

### Condition de rayonnement de Sommerfeld.

Dautray-Lions [DL85] t2 p755 montre, lorsque  $J$  est une distribution à support compact, que  $J \star G$  vérifie la condition de rayonnement de Sommerfeld forte (21a) et donc la condition faible (30) aussi.

Exceptionnellement, nous nous abstenons de reprendre ici cette démonstration, car elle n'a pas besoin d'être adaptée et n'apporte rien de particulier.

### Vérification de l'équation de Helmholtz.

D'après la propriété 2.7, l'opérateur de Helmholtz (26) est continu de  $H_{loc}^1({}^cS)$  dans  $H_{loc}^{-1}$ . Alors l'ensemble des fonctions  $H_{loc}^1({}^cS)$  qui sont solution de l'équation de Helmholtz (20) est un ensemble fermé. Comme  $J \star G$  converge dans  $H_{loc}^1({}^cS)$ , comme  $G \in H_{loc}^1(\mathbb{R}^2 \setminus \{0\})$  et vérifie l'équation de Helmholtz (20), et comme  $J$  a pour support  $\partial S$ , alors  $J \star G$  est dans cet ensemble fermé.

C'est à dire que  $J \star G$  est solution de l'équation de Helmholtz (20) dans  ${}^cS$ .

Le même raisonnement montre que  $J \star G$  est aussi solution de l'équation de Helmholtz (21d) dans  $S$ . ■

Nous pouvons donc utiliser la fonction

$$J = \nu_1 E + \nu_1^i E$$

comme l'inconnue à rechercher, et une fois cette fonction calculée, le champ électrique  $E$  pourra se déduire de  $J$  par simple convolution avec la solution fondamentale de l'opérateur de Helmholtz (28) :

$$E = J \star G.$$

Ce remplacement de  $E$  par  $J$ , comme nous l'avons dit dans l'introduction, permet d'avoir une fonction  $J$  définie seulement sur un sous-ensemble compact de dimension 1, à savoir le bord  $\partial S$  du diffractant

$$J : \partial S \rightarrow \mathbb{R}$$

plutôt qu'une fonction

$$E : {}^cS \rightarrow \mathbb{R}$$

définie sur le complémentaire non borné du diffractant  $S$ .

## 2.3 Le problème intégral de Helmholtz : la CFIE

### 2.3.1 L'équation intégrale combinée

Maintenant que nous avons introduit la fonction inconnue  $J$ , précisons les équations que nous imposons à  $J$ . Nous cherchons à construire (21h). D'après le théorème d'unicité ci-dessus, il suffirait par exemple d'imposer la valeur de  $\nu_0(J \star G)$  pour choisir  $J$ . L'équation résultante

$$\nu_0(J \star G) = g \quad (48g)$$

est la version bidimensionnelle scalaire de l'Electric Field Integral Equation (EFIE), basée sur la condition extérieure (35c) et la condition intérieure  $\nu_0^i E = g$ . Seulement, l'opérateur correspondant (qui est symétrique)

$$\begin{aligned} J &\mapsto \nu_0(J \star G) \\ \mathbb{H}^{-\frac{1}{2}}(\partial S) &\rightarrow \mathbb{H}^{\frac{1}{2}}(\partial S) \end{aligned} \quad (48h)$$

a des problèmes lorsque le problème interne a des modes propres pour certaines longueurs d'ondes  $\lambda_0$  (précisément, lorsque  $-\frac{4\pi^2}{\lambda_0^2}$  est une des valeurs propres de l'opérateur non borné de Laplace  $\mathbb{H}_0^1(S) \rightarrow \mathbb{H}_0^1(S)$ ).

Pour une application numérique, nous avons besoin d'un autre opérateur dont l'inversibilité ne dépend pas si critiqueusement de la longueur d'onde  $\lambda_0$ , même si nous perdons la symétrie de l'opérateur. C'est indispensable pour une étude numérique asymptotique en  $\lambda_0$  petit.

La proposition suivante introduit l'opérateur suivant, qui contient une condition aux limites de type Leontovitch ou encore Impedance Boundary Condition (IBC).

$$J \mapsto \frac{i\alpha}{\lambda_0} \nu_0(J \star G) + 2\nu_1^i(J \star G) \quad (48i)$$

avec  $\alpha$  un réel non nul (physiquement mesurable en mètre par Ohms) ; elle démontre son injectivité, et la surjectivité est démontrée par le théorème 2.30. Nous obtenons ainsi l'équation du problème (21h) qui s'appelle la *Combined Field Integral Equation* (CFIE) ; elle est basée sur la condition au bord extérieure (35c) et la condition au bord intérieur basée sur la fonction

$$\frac{i\alpha}{\lambda_0} \nu_0 E + 2\nu_1^i E. \quad (49)$$

**Propriété 2.20** *Soit  $\alpha \neq 0$  un nombre réel. Il y a unicité dans le problème intégral de Helmholtz tel que posé en (21h), lorsque la fonction  $E^{inc}$  est*

$H_{loc}^1(\mathbb{R}^2)$  et vérifie l'équation de Helmholtz (21f) sur  $\mathbb{R}^2$ . L'éventuelle solution  $J$  vérifie aussi  $\nu_0(J \star G) = -\nu_0 E^{inc}$ , et l'égalité  $J \star G = -E^{inc}$  dans  $H^1(S)$ , et alors  $J \star G$  est solution du problème non-intégral de Helmholtz extérieur (21c).

La suggestion de prendre une combinaison linéaire à coefficients complexes de  $\nu_0$  et  $\nu_1^i$  a été faite par Brakhage-Werner [BW65].

Notons que pour résoudre le problème de Helmholtz extérieur par un problème intégral, il aurait fallu changer d'opérateur, en remplaçant (21h) par son adjoint. En effet, le problème (21h) suppose  $\nu_1 E^{inc}$  connu.

L'ajout de  $\lambda_0$  dans (49) est suggéré par Kress [K85], d'après l'étude numérique de la minimisation du nombre de conditionnement de l'équation numérique résultant de (21h) pour le cas où  $\partial S$  est un cercle (expérience répétée par Amini [A93]). Si de plus  $|\alpha| \leq 4\pi$ , Canning [C94] explique (pour l'opérateur adjoint) que le courant résultant ne rayonne pas (au sens physique) vers l'infini dans la direction formant un angle de

$$\arccos -\frac{\alpha}{4\pi} \quad (50)$$

(indépendant de  $\lambda_0$ ) avec la normale extérieure du diffractant. Voici son argument : dans l'expression, lorsqu'on remplace  $G$  dans  $K$  (23) par sa définition (28) contenant la fonction de Hankel  $H_0^{ank}$ , on obtient cette fonction

$$K(t, t') := \begin{cases} 0 & \text{si } t = t' \\ \frac{\alpha}{4\lambda_0} H_0^{ank}\left(\frac{2\pi}{\lambda_0}|s(t) - s(t')|\right) + \frac{i\pi}{\lambda_0} \langle s'(t)^\perp | \frac{s(t) - s(t')}{|s(t) - s(t')|} \rangle H_0^{ank'}\left(\frac{2\pi}{\lambda_0}|s(t) - s(t')|\right) & \text{sinon} \end{cases} \quad (51)$$

qui est le noyau de l'équation de Fredholm de seconde espèce (21h). Lorsque  $t$  est fixé, nous voyons que les

deux facteurs dans ce noyau se compensent lorsque  $\frac{1}{\lambda_0}(s(t') - s(t))$  est grand tout en formant l'angle (50) avec le vecteur  $s'(t)^\perp$ , vu les équivalents de  $H_0^{ank}$  donnés par (32a) et (32b). Nous vérifierons numériquement cette prédiction dans la section 5.1.1.

Ainsi, le facteur  $\lambda_0$  dans (21h) rend cette dernière équation plus homogène en  $\lambda_0$  au sens physique.

■ Par linéarité, il suffit de montrer que si  $J \in H^{-\frac{1}{2}}(\partial S)$  vérifie (21h) avec  $E^{inc} = 0$ , alors  $J = 0$ . D'après la propriété 2.19 précédente, on voit que  $E := J \star G$  est  $H^1(S)$ , vérifie  $\nu_0 E = \nu_0^i E$  et vérifie l'équation de Helmholtz sur  $S$  (20) et sur  ${}^c S$  (21d) avec la condition de Sommerfeld (30).

L'égalité de Stokes (46a) pour  $S$  avec  $G$  remplacé par  $\overline{E}$  donne :

$$-\int_{\partial S} \nu_0^i E \overline{\nu_1^i E} = \int_S \overline{E} \Delta E + \int_S |\nabla E|^2 = -\frac{4\pi^2}{\lambda_0^2} \int_S |E|^2 + \int_S |\nabla E|^2 \quad (52a)$$

Or on a

$$\nu_1^i E = -\frac{i\alpha}{2\lambda_0} \nu_0^i E \quad :$$

la partie imaginaire de cette équation (52a) donne alors  $\int_{\partial S} |\nu_0^i E|^2 = 0$ . Par cascade, on a aussi  $\nu_1^i E = 0$  et  $\nu_0 E = 0$ . La formule de représentation (45) dans  $S$  montre alors que  $E = 0$  dans  $S$ . Le théorème d'unicité 2.14 montre que, si  $\nu_0 E = 0$ , alors  $E = 0$  dans  ${}^c S$ .

Nous ne supposons plus que  $E^{inc} = 0$ . Nous allons maintenant montrer que  $J \star G = -E^{inc}$  dans  $S$ . Observons que  $E^{inc} \in H^1(S)$ , ce qui fait qu'on peut reprendre tout le début de cette démonstration avec  $E = J \star G - E^{inc}$  au lieu de  $E = J \star G$ ; comme  $E^{inc}$  est  $H_{loc}^1(\mathbb{R}^2)$  et vérifie l'équation de Helmholtz sur  $\mathbb{R}^2$ , le théorème de représentation interne 2.17 appliqué avec  $S$  remplacé par un voisinage borné de  $S$  montre que  $E^{inc}$  est analytique dans un voisinage de  $S$ ; donc on a  $\nu_1^i E^{inc} = -\nu_1 E^{inc}$ . et  $\nu_0^i E^{inc} = \nu_0 E^{inc}$ .

Il s'en suit que  $E = J \star G - E^{inc}$  est tel que (49) est nul, ce qui en reprenant toute cette démonstration montre que  $E$  est nul sur  $S$  (on ne peut pas reprendre la fin de cette démonstration qui utilise le théorème d'unicité 2.14, car  $E$  ne vérifie pas la condition de rayonnement de Sommerfeld (30)). De la nullité de  $E$  sur  $S$ , on en tire aussi  $\nu_0^i (J \star G) = \nu_0^i E^{inc}$  qui permet de clore cette démonstration.

Remarquons que cette démonstration a utilisé sur  $E^{inc}$  que  $E^{inc} \in H^1(V)$  avec  $V$  un voisinage de  $S$ : l'hypothèse  $E^{inc} \in H_{loc}^1(\mathbb{R}^2)$  est en fait un peu forte. ■

**Remarque 2.21** *Wagner-Chew [WC95], Deng-Ling [DL99b] et Golik [G98] [G00] travaillent avec  $\alpha \nu_0^i E + 2\nu_1^i E$  au lieu de (49) : c'est une erreur regrettable, comme nous allons le montrer ci-dessous.*

Les démonstrations de l'utilité de  $\alpha$  non nul dans (49) faites par Brakhage-Werner [BW65] et Colton-Kress [CK83] supposent toutes les deux que  $\alpha$  n'est pas imaginaire pur.

Nous montrons ci-dessous, en utilisant les harmoniques cylindriques, qu'un tel remplacement est même fatal pour la propriété ci-dessus dans le cas où  $S$  est une boule :

**Remarque 2.22** Cette propriété est fausse si le nombre complexe  $i$  est oublié dans (49) lorsque  $S$  est une boule, c'est à dire que certaines valeurs de  $\lambda_0$  rendent cette propriété fausse.

■ La démonstration du lemme 2.15 contient déjà des idées pour cette démonstration : la fonction  $(r, \theta) \mapsto e^{im\theta} H_m^{ank}(\frac{2\pi r}{\lambda_0})$  résout l'équation de Helmholtz extérieure (20) dans  ${}^cS$ , et sa partie réelle résout l'équation de Helmholtz intérieure (21d) dans  $S$ ; nous utilisons des coordonnées cylindriques  $(r, \theta)$  centrées sur la boule  $S$ .

Soit  $m \geq 0$  un entier quelconque. Soit  $\lambda_0$  tel que  $\frac{2\pi R}{\lambda_0}$  est un zéro de la fonction

$$x \mapsto \mathcal{R}e(\alpha H_m^{ank}(x) - \frac{4\pi}{\lambda_0} H_m^{ank'}(x)). \quad (52b)$$

L'existence de  $\lambda_0$  est claire d'après les développements limités (42) et (43); comme  $\mathcal{R}e H_m^{ank}$  est solution de l'équation différentielle de Bessel (40), on a même  $\mathcal{R}e H_m^{ank}(\frac{2\pi R}{\lambda_0}) \neq 0$ .

Soit la fonction  $E$  définie par

$$E(r, \theta) = \begin{cases} e^{im\theta} \frac{H_m^{ank}(\frac{2\pi r}{\lambda_0})}{H_m^{ank}(\frac{2\pi R}{\lambda_0})} & \text{si } r > R \\ e^{im\theta} \frac{\mathcal{R}e H_m^{ank}(\frac{2\pi r}{\lambda_0})}{\mathcal{R}e H_m^{ank}(\frac{2\pi R}{\lambda_0})} & \text{sinon} \end{cases}$$

Cette fonction vérifie l'équation de Helmholtz dans  ${}^cS$  (20) et dans  $S$  (21d). La somme des deux formules de représentations (38) et (45) donne alors

$$(\nu_1 E + \nu_1^i E) \star G + (s^{\perp}(\cdot)(\nu_0 E - \nu_0^i E)) \star \nabla G = E$$

qui montre que  $E$  est la convolée de  $G$  par la fonction

$$J(r, \theta) = \frac{2\pi}{\lambda_0} e^{im\theta} \left( \frac{H_m^{ank'}(\frac{2\pi r}{\lambda_0})}{H_m^{ank}(\frac{2\pi R}{\lambda_0})} - \frac{\mathcal{R}e H_m^{ank'}(\frac{2\pi r}{\lambda_0})}{\mathcal{R}e H_m^{ank}(\frac{2\pi R}{\lambda_0})} \right) \delta(r=R) = \frac{2i}{\pi R} \frac{\delta(r=R) e^{im\theta}}{H_m^{ank}(\frac{2\pi R}{\lambda_0}) \mathcal{R}e H_m^{ank}(\frac{2\pi R}{\lambda_0})}$$

(où on a utilisé l'identité 6.1) et pourtant on a  $\alpha \nu_0^i E + 2\nu_1^i E = 0$ . ■

Travailler avec  $\frac{\alpha}{\lambda_0} \nu_0^i E + 2\nu_1^i E$  n'offre pas plus de garantie : la même démonstration, mais avec (52b) privé de son premier  $x$ , montre l'existence de résonances pour certains  $\lambda_0$ .

Donc le nombre complexe  $i$  dans (48i) est *indispensable* dans la propriété ci-dessus. L'oubli de ce  $i$ , fait par Wagner-Chew [WC95], Deng-Ling [DL99b] et Golik [G98] [G00], est à l'origine de problèmes latents de résonance pour



certaines valeurs de  $\lambda_0$ , et donc remet en question tous les résultats numériques disponibles pour la CFIE lorsqu'ils incluent une dépendance en  $\lambda_0$ .

Bamberger-Ha-Duong [BH87] a des résultats numériques basés sur (48i), mais pour un problème où la condition au bord extérieur (35c) a été remplacée par un analogue externe de (49). Ils ont dû remplacer l'inconnue  $J$  par un couple d'inconnues, et ils utilisent l'opérateur hypersingulier (44c) qui est plus difficile à traiter dans les coins du diffractant.

### 2.3.2 Singularités de la CFIE dues aux coins

En présence de coins, l'espace  $H^{-\frac{1}{2}}(\partial S)$  choisi pour  $J$  n'est pas commode à discrétiser ; il est possible d'utiliser à la place l'espace  $L^2(\partial S)$ . En bonus, ce choix permettra aussi de décomposer l'opérateur (55) en l'identité additionné à un opérateur contractant et à un opérateur compact, et donc d'étudier la dépendance en  $\lambda_0$  de l'efficacité d'une discrétisation.

L'espace  $L^2(\partial S)$  est le plus petit des espaces Hermitiens de Sobolev pour lequel Costabel [C88] montre la continuité de l'opérateur  $J \mapsto (49)$ , et c'est le plus gros pour lequel la contractance est facile à montrer dans la décomposition annoncée ci-dessus.

Lorsque le diffractant  $S$  n'a pas de coins, Kress [K89], théorème 6.13, Nédélec [N01], théorème 3.1.2, et Dautray-Lions [DL85], proposition III.3.12, montrent que  $J \mapsto 2\nu_1^i(J \star G)$  est l'opérateur identité ajouté à l'opérateur de noyau

$$(t, t') \mapsto -\frac{i}{4} \frac{2\pi}{\lambda_0} \langle s'(t)^\perp | \frac{s(t) - s(t')}{|s(t) - s(t')|^2} \rangle |s(t) - s(t')| H_0^{ankl} \left( \frac{2\pi}{\lambda_0} |s(t) - s(t')| \right), \quad (52c)$$

qui est un noyau régulier à prolonger par continuité au cas où  $t = t'$ . Cette régularité résultera de la propriété 2.24 ci-dessous ; Mikhlin [M59] est l'un de ceux qui ont montré l'analogie de cette régularité pour l'opérateur de Laplace.

Montrons que, dans le cas où  $S$  peut avoir des coins, cette formule (52c) reste valable (Chandler [C84b] fait l'analogie pour le Laplacien), même si la régularité du noyau est perdue aux coins éventuels :

**Propriété 2.23** *Soit  $J \in L^2(\partial S)$ . Alors, pour  $t \in [0, L]$  :*

$$-\nu_1^i(J \star G)(t) = -\frac{\theta(t)}{2\pi} J(t) + \int_0^L \langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle J(t') dt'$$

où  $\theta(t)$  vaut l'angle interne dans  $S$  vu par  $s(t)$ , et où l'intégrale est définie au sens de Lebesgue. Comme l'égalité ci-dessus n'est valide que presque sûrement,  $\theta(t)$  est remplaçable par  $\pi$ .

Dans cette propriété,  $\theta(t)$  vaut  $\pi$  sauf aux coins ( $\theta(t_0) = \frac{\pi}{2}$  pour  $s(t_0)$  un coin normal à angle droit,  $\theta(t_0) = \frac{3\pi}{2}$  pour  $s(t_0)$  un angle droit rentrant) qui sont de mesure négligeable.

Notons que l'intégrale apparaissant dans cette propriété n'a pas le même opérateur que dans le second terme du membre gauche de l'équation de représentation (45) : c'est l'opérateur adjoint. De plus, ce n'est pas non plus l'intégrale obtenue en développant  $\nu_1^i(J \star G)$  : en effet, l'intégrale et la dérivée ont été *intervertis*.

■  $E := J \star G$  est  $H_{loc}^1({}^cS)$  et  $H^1(S)$  d'après la propriété 2.19. Ignorons, le temps de cette démonstration, le facteur  $-\frac{i}{4}$  présent dans la définition (28) de  $G$ .

Calculons, sur l'intérieur de  $\partial S$ , le gradient de  $E$ , qui ne pourrait avoir de sens que contre une fonction  $\phi$  régulière à support borné sur  $\mathbb{R}^2$ . Comme  $\Delta E - \frac{4\pi^2}{\lambda_0^2}E = 0$ , avec  $\varepsilon > 0$ , la formule de Green  $\int_{\partial S} v \partial_n u = \int_{\Omega} v \Delta E + \langle \nabla E | \nabla v \rangle$  (avec produit scalaire non Hermitien) donne

$$\begin{aligned} & \int_0^L \phi(s(t) + \varepsilon s'(t)^\perp) \langle s'(t)^\perp | \nabla_t \rangle E(s(t) + \varepsilon s'(t)^\perp) dt \\ &= \int_{D_\varepsilon} \frac{4\pi^2}{\lambda_0^2} E(x) \phi(x) + \langle \nabla E(x) | \nabla \phi(x) \rangle dx \end{aligned}$$

où  $D_\varepsilon$  est l'ensemble des points  $S$  à une distance au moins  $\varepsilon$  de  ${}^cS$ . Grâce aux régularités connues sur  $E$  dans le membre droit, toute l'égalité ci-dessus a une limite lorsque  $\varepsilon$  tend vers 0 par valeur positive dès que  $\phi$  est  $H^1({}^cS)$ , malgré la présence de la dérivée intérieure normale de  $E$ . Remplaçons  $E$ , ce qui donne :

$$\begin{aligned} \int_{D_0} \frac{4\pi^2}{\lambda_0^2} E(x) \phi(x) + \langle \nabla E(x) | \nabla \phi(x) \rangle dx &= \int_{D_0} \int_0^L J(t') \frac{4\pi^2}{\lambda_0^2} H_0^{ank} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) \phi(x) \\ &\quad + J(t') \langle \nabla \phi(x) | \nabla_x \rangle H_0^{ank} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) dt' dx \\ &= \int_0^L J(t') \int_{D_0} \frac{4\pi^2}{\lambda_0^2} H_0^{ank} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) \phi(x) + \langle \nabla \phi(x) | \nabla_x \rangle H_0^{ank} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) dx dt' \end{aligned}$$

Pour justifier l'interversion d'intégrales (Fubini) ci-dessus, montrons que la dernière intégrale double converge absolument. Comme  $J \in L^2(\partial S)$ , il suffit par exemple de montrer que l'intégrale sur  $D_0 = S$  est bornée indépendamment de  $t'$ . Or c'est le cas, comme  $D_0$  est un compact de  $\mathbb{R}^2$ , et comme on

a, vu que  $\phi$  et  $\nabla\phi$  sont bornés :

$$\begin{aligned} & \left| \frac{4\pi^2}{\lambda_0^2} \text{H}_0^{\text{ank}} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) \phi(x) + \langle \nabla\phi(x) | \nabla_x \rangle \text{H}_0^{\text{ank}} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) \right| \\ & \leq C(|\log(\|s(t') - x\||) + \frac{1}{\|s(t') - x\|}) \end{aligned}$$

où  $C = C(\phi, \lambda_0)$  est indépendant de  $t'$  et de  $x$  d'après (29d) et (29e).

Ensuite, on remplace  $D_0$  par  $D_0 \setminus B(s(t'), \varepsilon) \cap D_0$  en enlevant la boule de centre  $s(t')$  et de rayon  $\varepsilon$  : lorsque  $\varepsilon$  tend vers 0, il y a convergence vers la intégrale ci-dessus. Comme  $\frac{4\pi^2}{\lambda_0^2} \text{H}_0^{\text{ank}} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) = \Delta_x (\text{H}_0^{\text{ank}} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right))$ , la même formule de Green donne

$$\begin{aligned} & \int_{D_0 \setminus B(s(t'), \varepsilon) \cap D_0} \frac{4\pi^2}{\lambda_0^2} \text{H}_0^{\text{ank}} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) \phi(x) + \langle \nabla\phi(x) | \nabla_x \rangle \text{H}_0^{\text{ank}} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) dx \\ & = \int_{(\partial S) \setminus (\partial S) \cap B(s(t'), \varepsilon)} \phi(x) \langle s'(x)^\perp | \nabla_x \rangle \text{H}_0^{\text{ank}} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) dx \\ & + \int_{x \in D_0, \|x - s(t')\| = \varepsilon} \phi(x) \left\langle \frac{s(t') - x}{\|s(t') - x\|} | \nabla_x \right\rangle \text{H}_0^{\text{ank}} \left( \frac{2\pi}{\lambda_0} \|s(t') - x\| \right) dx \end{aligned}$$

La dernière intégrale tend vers  $-2i \frac{\theta(t)}{\pi} \phi(s(t'))$ , et dans la précédente, l'intégrande est intégrable, donc au bilan on a bien trouvé ce qu'il fallait montrer, mais intégré contre la fonction  $\phi$ .  $\blacksquare$

Donnons un équivalent du noyau apparaissant dans la propriété 2.23. Cet équivalent est le premier résultat original de la partie II de cette thèse.

**Propriété 2.24** *Lorsque  $t$  et  $t'$  sont à distance inférieure à  $\lambda_0$  et de part et d'autre d'un coin  $t_0$ , on a*

$$\begin{aligned} \langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle & = \frac{\frac{1}{2\pi} |t' - t_0| \sin \theta(t_0)}{(t - t_0)^2 + (t' - t_0)^2 + 2(t - t_0)(t' - t_0) \cos \theta(t_0)} \\ & + O\left(\frac{1}{\lambda_0}\right) \end{aligned} \quad (52d)$$

*et pour  $t$  et  $t'$  à distance entre eux inférieure à  $\lambda_0$  dans une partie connexe régulière de  $\partial S$ , on a*

$$\langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle = O\left(1 + \frac{(t - t')^2}{\lambda_0} \left| \log \frac{|t - t'|}{\lambda_0} \right| \right).$$

Ces développements limités sont uniformes en  $\lambda_0 < 1$ ; on peut les dériver en  $t$  une fois membre à membre, les deux  $O()$  devenant respectivement  $O(\frac{1}{|t-t'|})$  et  $O(\frac{1}{\lambda_0})$ .

Lorsque  $|t - t'| > \lambda_0$  (modulo  $L$ ), il faut remplacer les quatre restes par respectivement  $O(\frac{1}{\sqrt{\lambda_0|t-t'|}})$ ,  $O(\sqrt{\frac{|t-t'|}{\lambda_0}})$ ,  $O(\frac{1}{\sqrt{\lambda_0^3|t-t'|}})$  et  $O(\sqrt{\frac{|t-t'|}{\lambda_0^3}})$ .

Le second développement limité est une preuve de la régularité du noyau (52c) dans le cas où  $\partial S$  est régulier.

■ Pour simplifier, posons  $t_0 = 0$ .

Le développement limité (29e) donne :

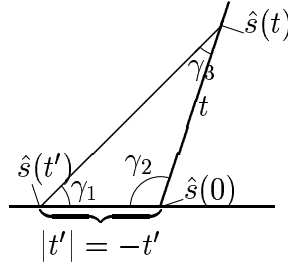
$$\begin{aligned} & \langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle \\ &= -\frac{i}{4} \frac{2\pi}{\lambda_0} \text{H}_0^{\text{ank}'} \left( \frac{2\pi}{\lambda_0} \|s(t) - s(t')\| \right) \frac{\langle s'(t)^\perp | s(t) - s(t') \rangle}{\|s(t) - s(t')\|} \end{aligned} \quad (52e)$$

$$= \frac{1}{2\pi} \frac{\langle s'(t)^\perp | s(t) - s(t') \rangle}{\|s(t) - s(t')\|^2} + O\left(\frac{1}{\lambda_0}\right) \quad (52f)$$

#### Quatrième quadrant.

Supposons que  $t > 0$  et que  $t' < 0$ .

Au voisinage de  $t_0 = 0$ ,  $s$  est équivalent à  $\hat{s}$  (version affine de  $s$ ),  $\hat{s}$  étant schématiquement :



avec  $\hat{s}$  caractérisée par  $\hat{s}(0) = s(0)$ ,  $\hat{s}'(0_+) = s'(0_+)$ ,  $\hat{s}'(0_-) = s'(0_-)$  et  $\hat{s}'' = 0$ .

On a

$$\|\hat{s}(t) - s(t)\| = \|\hat{s}(0_+) + t\hat{s}'(0_+) - s(0_+) - ts'(0_+) - \int_0^t (t-r)s''(r)dr\| \leq \|s''\|_\infty \frac{t^2}{2}$$

et

$$\|\hat{s}(t) - \hat{s}(t')\|^2 = t^2 + t'^2 - 2ett' \geq \min(1 + e, 1)(t^2 + t'^2) \quad (52g)$$

ce qui fait qu'on a, quitte à supposer  $\|s''\|_\infty \sqrt{t^2 + t'^2} \leq \min(1 + e, 1)$  :

$$\|s(t) - s(t')\| \geq \min\left(\sqrt{\frac{1+e}{2}}, \frac{1}{2}\right) \sqrt{t^2 + t'^2}$$

Lorsqu'on mélange (52f) avec les développements limités

$$s'(t) = \hat{s}'(t) - \langle s''(0_+)^{\perp} | s'(0_+) \rangle s'(0_+)^{\perp} t + O(t^2) \quad (52h)$$

et

$$s(t) = \hat{s}(t) - \frac{1}{2} \langle s''(0_+)^{\perp} | s'(0_+) \rangle s'(0_+)^{\perp} t^2 + O(t^3)$$

et

$$s(t') = \hat{s}(t') - \frac{1}{2} \langle s''(0_-)^{\perp} | s'(0_-) \rangle s'(0_-)^{\perp} t'^2 + O(t'^3), \quad (52i)$$

on a, en développant jusqu'à  $O(\frac{1}{\lambda_0})$  pour  $t > 0$  et  $-t' > 0$  :

$$(52f) = \frac{1}{2\pi} \frac{\langle \hat{s}'(t)^{\perp} | \hat{s}(t) - \hat{s}(t') \rangle}{\|\hat{s}(t) - \hat{s}(t')\|^2} + O\left(\frac{1}{\lambda_0}\right) = \frac{1}{2\pi} \frac{-t' \sin \theta(0)}{t^2 + t'^2 - 2ett'} + O\left(\frac{1}{\lambda_0}\right)$$

où on a utilisé

$$\hat{s}(t) = \hat{s}(0_+) + \hat{s}'(0_+)t, \quad (52j)$$

$$\hat{s}(t') = \hat{s}(0_-) + \hat{s}'(0_-)t',$$

et

$$\langle s'(0_+)^{\perp} | s'(0_-) \rangle = -\langle s'(0_-)^{\perp} | s'(0_+) \rangle = \cos(\theta(0) - \frac{\pi}{2}) = \sin(\theta(0)). \quad (52k)$$

Les développements limités (29e) et (32b) permettent de prouver  $O\left(\frac{1}{\sqrt{\lambda_0|t-t'|}}\right)$  à partir de (52e) lorsque  $|t - t'| \geq 0$ .

### Second quadrant.

Supposons que  $t < 0$  et que  $t' > 0$ . La fonction  $\hat{s}$  convient encore parfaitement ici, et elle satisfait (52g). Alors (52f) donne le résultat demandé aussi, l'égalité (52k) fournissant un changement de signe qui donne le facteur  $|t'|$  indiqué.

**Premier et troisième quadrants, partie régulière.**

En insérant les développements limités

$$s(t') = s(t) - s'(t)(t - t') + \frac{1}{2}s''(t)(t - t')^2 + O((t - t')^3) \quad (52\ell)$$

et (29e) et (32b) dans (52e), le résultat sort directement. De plus, comme  $\|s(t') - s(t)\|^2 = (t - t')^2 + O((t - t')^4)$ , nous trouvons même l'équivalent

$$\langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle = -\frac{\langle s'(t)^\perp | s''(t) \rangle}{4\pi} + O\left(\frac{(t - t')^2}{\lambda_0} \left| \log \frac{|t - t'|}{\lambda_0} \right| \right).$$

**Dérivation des développements limités.**

D'après Watson [W22], p74 (5), la dérivée seconde de la fonction  $H_0^{ank}$  admet le développement limité  $-(29d) - \frac{1}{t}(29e)$ , ce qui donne

$$H_0^{ank''}(t) = -\frac{2i}{\pi} \frac{1}{t^2} - \frac{i}{\pi} \ln t + C + O(t^2 \ln t)$$

En combinant ce développement limité avec le développement limité de  $H_0^{ank'}$  vu en (29e), on a :

$$\begin{aligned} & \frac{\partial}{\partial t} \langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle \\ &= -\frac{i}{4} \left(\frac{2\pi}{\lambda_0}\right)^2 H_0^{ank''} \left(\frac{2\pi}{\lambda_0} \|s(t) - s(t')\|\right) \frac{\langle s'(t)^\perp | s(t) - s(t') \rangle}{\|s(t) - s(t')\|} \frac{\langle s'(t) | s(t) - s(t') \rangle}{\|s(t) - s(t')\|} \\ & \quad - \frac{i}{4} \frac{2\pi}{\lambda_0} H_0^{ank'} \left(\frac{2\pi}{\lambda_0} \|s(t) - s(t')\|\right) \frac{\langle s''(t)^\perp | s(t) - s(t') \rangle}{\|s(t) - s(t')\|} \\ & \quad + \frac{i}{4} \frac{2\pi}{\lambda_0} H_0^{ank'} \left(\frac{2\pi}{\lambda_0} \|s(t) - s(t')\|\right) \frac{\langle s'(t)^\perp | s(t) - s(t') \rangle \langle s'(t) | s(t) - s(t') \rangle}{\|s(t) - s(t')\|^3} \quad (52m) \\ &= -\frac{1}{\pi} \frac{\langle s'(t)^\perp | s(t) - s(t') \rangle \langle s'(t) | s(t) - s(t') \rangle}{\|s(t) - s(t')\|^4} + \frac{1}{2\pi} \frac{\langle s''(t)^\perp | s(t) - s(t') \rangle}{\|s(t) - s(t')\|^2} \\ & \quad + O\left(\frac{1}{\lambda_0}\right) \quad (52n) \end{aligned}$$

**Quatrième quadrant.**

Une nouvelle fois, supposons  $t > 0$  et  $t' < 0$ . En utilisant les développements limités (52h) à (52i), puis les égalités (52j) à (52k), on a

$$\frac{\partial}{\partial t} \langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle$$

$$\begin{aligned}
&= -\frac{1}{\pi} \frac{\langle \hat{s}'(t)^\perp | \hat{s}(t) - \hat{s}(t') \rangle \langle \hat{s}'(t) | \hat{s}(t) - \hat{s}(t') \rangle}{\|\hat{s}(t) - \hat{s}(t')\|^4} + \frac{1}{2\pi} \frac{\langle s''(t)^\perp | s(t) - s(t') \rangle}{\|s(t) - s(t')\|^2} \\
&\quad - \frac{1}{\pi} \frac{\langle s''(0_+)^\perp | s'(0_+) \rangle \langle \hat{s}'(0_+) | \hat{s}(t) - \hat{s}(t') \rangle \langle \hat{s}'(t) | \hat{s}(t) - \hat{s}(t') \rangle}{\|\hat{s}(t) - \hat{s}(t')\|^4} t \\
&\quad + \frac{1}{\pi} \frac{\langle \hat{s}'(t)^\perp | \hat{s}(t) - \hat{s}(t') \rangle \langle s''(0_+)^\perp | s'(0_+) \rangle \langle \hat{s}'(0_+)^\perp | \hat{s}(t) - \hat{s}(t') \rangle}{\|\hat{s}(t) - \hat{s}(t')\|^4} t + O\left(\frac{1}{\lambda_0}\right) \\
&= -\frac{1}{\pi} \frac{(-t' \sin \theta(0))(t - et')}{(t^2 + t'^2 - 2ett')^2} + O\left(\frac{1}{\lambda_0} + \frac{1}{|t - t'|}\right)
\end{aligned}$$

### Second quadrant.

Supposons que  $t < 0$  et que  $t' > 0$ . Tout marche de la même manière, l'égalité (52k) fournissant ici aussi un changement de signe qui donne le facteur  $|t'|$  indiqué.

### Premier et troisième quadrants, partie régulière.

En insérant (52ℓ) dans (52n), le résultat  $O(\frac{1}{\lambda_0})$  sort directement. Watson [W22] montre, p196 partie 7.2, que  $H_0^{ank''}$  a l'équivalent en l'infini suivant :

$$H_0^{ank''}(r) + (1 - i)e^{ir}/\sqrt{\pi r} = O(r^{-3/2}) \quad (53)$$

ceci avec (32b) permet de montrer que, lorsque  $|t - t'| > \lambda_0$ , on a une majoration en  $O(\sqrt{\frac{|t-t'|}{\lambda_0^3}})$ . ■

La seule singularité de ce noyau (52d) est donc, aux coins, un terme en  $1/\|(t, t') - (t_0, t_0)\|$  avec  $t_0$  l'abscisse curviligne du coin.

Nous pouvons vérifier à la main la continuité  $L^2(\partial S) \rightarrow L^2(\partial S)$  de cette singularité :

**Propriété 2.25** *Cette singularité détaillée par le lemme 2.24 est un opérateur continu, et contractant  $L^2(\partial S) \rightarrow L^2(\partial S)$  avec un ratio strictement plus petit que  $\frac{1}{2}$ .*

■ Soit  $f$  et  $g$  deux fonctions  $L^2(\mathbb{R})$ . Pour la démonstration, prenons  $t_0 = 0$ . Notons

$$Mf(t_0) = \limsup_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_{\|t-t_0\| \leq \varepsilon} |E(t)| dt$$

la fonction maximale de  $f$ , version unidimensionnelle de (39g). On a :

$$\begin{aligned}
& \left| \iint \frac{f(x)g(y)}{\|(x,y) - (t_0,t_0)\|} dx dy \right| \\
& \leq \iint_{|x-t_0| \leq |y-t_0|} \frac{|f(x)g(y)|}{|y-t_0|} dx dy + \iint_{|y-t_0| \leq |x-t_0|} \frac{|f(x)g(y)|}{|x-t_0|} dx dy \\
& \leq \|f\|_{L^2} \|Mg\|_{L^2} + \|g\|_{L^2} \|Mf\|_{L^2} \\
& \leq C \|f\|_{L^2} \|g\|_{L^2}
\end{aligned} \tag{54a}$$

ce qui montre la continuité. Chandler [C84b] suggère, pour la contractivité, d'utiliser l'isométrie

$$\begin{aligned}
I : L^2((0,1)) &\rightarrow L^2((0,\infty)) \\
f \mapsto t &\mapsto If(t) = f(e^{-t})e^{-t/2}
\end{aligned}$$

qui est reliée à la transformation de Mellin d'après Costabel-Stephan [CS85]. Un changement de variable montre que c'est bien une isométrie. Pour simplifier la suite, posons  $\theta = \theta(0)$ . Or on a

$$\begin{aligned}
I(t \mapsto \int_{-\infty}^0 \frac{1}{2\pi} \frac{t' \sin \theta}{t^2 + t'^2 + 2tt' \cos \theta} f(t') dt')(t) \\
&= \int_0^{\infty} \frac{1}{2\pi} \frac{t' \sin \theta}{e^{-2t} + t'^2 - 2e^{-t}t' \cos \theta} f(-t') dt' e^{-t/2} \\
&= \int_0^{\infty} \frac{1}{2\pi} f(-t') e^{-t/2} \mathcal{I}m \frac{1}{e^{-t} - t' e^{i\theta}} dt' \\
&= \int_0^{\infty} \frac{1}{2\pi} f(e^{-t'}) e^{-t'} e^{-t/2} \mathcal{I}m \frac{1}{e^{-t} - e^{-t'} e^{i\theta}} dt' \\
&= \int_{\mathbb{R}} \frac{1}{2\pi} If(t') \mathcal{I}m \frac{e^{\frac{t'-t}{2}}}{e^{t'-t} - e^{i\theta}} dt'
\end{aligned} \tag{54b}$$

Nous reconnaissons ici une convolution, et le lemme 2.26 calculera une transformée de Fourier qui donne, pour  $\xi \in \mathbb{R}$  :

$$-i\pi \frac{e^{(\xi - \frac{1}{2})(\theta - \pi)}}{\sinh \pi(\xi - \frac{1}{2})} = \int_{\mathbb{R}} e^{-i(\xi - \frac{1}{2})t} \frac{e^t}{e^t - e^{i\theta}} dt = \int_{\mathbb{R}} e^{-i\xi t} \frac{e^{\frac{t}{2}}}{e^t - e^{i\theta}} dt$$

Comme  $If$  est  $L^2(\mathbb{R})$ , sa transformée de Fourier existe, donc (54b) vaut

$$\begin{aligned}
(54b) &= -\frac{1}{4\pi} \int_{\mathbb{R}} \widehat{If}(\xi) \left( i\pi \frac{e^{(\xi - \frac{1}{2})(\theta - \pi)}}{\sinh \pi(\xi - \frac{1}{2})} + i\pi \frac{e^{(-\xi + \frac{1}{2})(\theta - \pi)}}{\sinh \pi(-\xi + \frac{1}{2})} \right) d\xi \\
&= \frac{i}{2} \int_{\mathbb{R}} \widehat{If}(\xi) \frac{\sinh(\xi - \frac{1}{2})(\theta - \pi)}{\sinh(\xi - \frac{1}{2})\pi} d\xi
\end{aligned}$$



Comme  $I$  et la transformée de Fourier sont des isométries, il reste donc à montrer que la fraction  $\frac{\sinh(\xi-i/2)(\theta-\pi)}{\sinh(\xi-i/2)\pi}$  est bornée, en norme, par plus petit que 1. Or l'identité

$$4|\sinh(t+it')|^2 = |e^t e^{it'} - e^{-t} e^{-it'}|^2 = e^{2t} + e^{-2t} + 2\mathcal{R}e e^{it'} \overline{e^{-it'}} = 2 \cosh 2t - 2 \cos 2t'$$

permet de montrer que  $\frac{|\sinh(\xi-i/2)(\theta-\pi)|}{|\sinh(\xi-i/2)\pi|}$  atteint son maximum, qui est plus petit que 1 strictement (car  $0 < \theta < 2\pi$ ). D'après Costabel-Stephan [CS85] et Chandler [C84b], le maximum est atteint pour  $\xi = 0$ , mais nous n'avons pas besoin de cette remarque.

L'isométrie  $I$ , la transformée de Fourier et la translation de vecteur  $\frac{i}{2}$  sont les ingrédients de la transformation de Mellin, explicitement exhibée par Costabel-Stephan [CS85] et Fabes-Jodeit-Lewis [FJL77] pour démontrer cette propriété ; Chandler [C84b] reprend aussi cette démonstration. ■

Le lemme suivant est attribué par Costabel-Stephan [CS85] à Eskin [E73] et à Shamir [S64].

**Lemme 2.26** *La transformée de Fourier de  $\frac{e^t}{e^t - e^{i\theta}}$  vaut  $-i\pi \frac{e^{\xi(\theta-\pi)}}{\sinh \pi \xi}$  à la fréquence  $\xi$  telle que  $\mathcal{I}m \xi = -\frac{i}{2}$ , si  $\theta$  est un élément non nul de  $(0, 2\pi)$ .*

■ Remplaçons  $\xi$  par  $\xi - \frac{i}{2}$  : nous devons montrer que l'intégrale sur  $\mathbb{R}$  de la fonction

$$f(t) = \frac{e^{(\frac{1}{2}-i\xi)t}}{e^t - e^{i\theta}}$$

vaut  $\pi \frac{e^{(\xi-\frac{i}{2})(\theta-\pi)}}{\cosh \pi \xi}$ , ce que nous allons montrer avec la méthode des résidus. Sur le domaine  $\{z \in \mathbb{C} : \mathcal{I}m(z) \in [0, 2\pi]\}$ , cette fonction  $f$  est holomorphe, sauf en le pôle  $i\theta$  où son résidu vaut

$$\frac{e^{(\frac{1}{2}-i\xi)t}}{(e^t - e^{i\theta})e^{i\theta}} = \frac{e^{(\frac{1}{2}-i\xi)i\theta}}{e^{i\theta}} = e^{(\xi-\frac{i}{2})\theta}.$$

Donc l'intégrale de Cauchy de  $f$  en sens direct le long du bord de  $\{z \in \mathbb{C} : \mathcal{I}m(z) \in [0, 2\pi], |\mathcal{R}e(z)| < A\}$  vaut  $2i\pi e^{(\xi-\frac{i}{2})\theta}$ . Or, cette intégrale vaut

$$\int_{-A}^A (f(t) - f(t + 2i\pi)) dt + i \int_0^{2\pi} (f(it + A) - f(it - A))$$

où, lorsque  $A$  tend vers l'infini, la seconde intégrale est un  $O(e^{-A/2})$  qui tend vers 0, et la première, qui vaut

$$\int_{-A}^A (1 + e^{2\pi\xi}) f(t) dt,$$

tend vers  $\int_{\mathbb{R}} f dt$  multiplié par  $1 + e^{2\pi\xi}$  : comme  $\frac{2i}{2e^{\pi\xi}} = e^{-\pi(\xi - \frac{i}{2})}$ , la démonstration est terminée. ■

### 2.3.3 Inégalité de Gårding pour la CFIE avec coins

L'inégalité de Gårding est un homologue affaibli de la coercivité, suffisant pour fonder une discrétisation numérique, et qui est plus fréquent que la coercivité lorsqu'on travaille avec des opérateurs non autoadjoints. Pour un opérateur  $L^2(\partial S) \rightarrow L^2(\partial S)$ , une inégalité de Gårding consiste en l'existence d'un opérateur compact  $L^2(\partial S) \rightarrow H^1(\partial S)$  tel que la somme deux opérateurs est coercive. Cette partie montre l'inégalité de Gårding (56).

La CFIE utilise l'opérateur double couche  $\nu_1^i(\cdot \star G)$ , pour lequel l'inégalité de Gårding est un résultat peu connu, son analogue pour le Laplacien étant fait par Costabel-Stephan [CS85] ou Chandler [C84b]. Pour l'opérateur  $\nu_0^i(\cdot \star G)$  comme pour l'opérateur  $\nu_1^i(\cdot s'^{\perp} \star \nabla G)$ , des inégalités de Gårding dans les espaces respectivement  $H^{-\frac{1}{2}}(\partial S)$  et  $H^{\frac{1}{2}}(\partial S)$  sont par exemple citées par Costabel [C88].

**Théorème 2.27** *L'application  $J \mapsto \nu_0^i(J \star G)$  est continue de  $L^2(\partial S)$  dans  $H^1(\partial S)$ , avec une norme d'opérateur en  $O(\sqrt{\frac{|\log \lambda_0|}{\lambda_0}})$ . Cette application est aussi continue de  $L^2(\partial S)$  dans  $L^2(\partial S)$ , avec une norme d'opérateur en  $O(|\log \lambda_0|)$ .*

Nous arrivons à prouver cette propriété grâce à des développements limités. Ce ne serait pas aussi facile si  $S$  n'était pas  $C^2$  par morceaux mais simplement Lipschitzien : Coifman-McIntosh-Meyer [CMM82] ont montré cette propriété dans le cas où le diffractant  $S$  est simplement supposé Lipschitzien avec des résultats de Calderon.

La continuité de  $H^{-\frac{1}{2}}(\partial S)$  dans  $H^{\frac{1}{2}}(\partial S)$  de l'application  $J \mapsto \nu_0^i(J \star G)$  résulte simplement du fait que  $J \star G$  est  $H^1(S)$  comme montré par la propriété 2.19. Mais pour avoir une inégalité de Gårding sur l'opérateur double couche, nous avons dû utiliser des espaces plus réguliers, par exemple en remplaçant  $H^{-\frac{1}{2}}(\partial S)$  par  $L^2(\partial S)$ . Cette inégalité utilisera fondamentalement la propriété 2.25 de contractance.

Steinbach-Wendland [SW01] a montré une inégalité de Gårding similaire pour un autre opérateur  $H^{\frac{1}{2}}(\partial S) \rightarrow H^{\frac{1}{2}}(\partial S)$ , dans une norme relative à l'inverse de l'opérateur simple couche.

Il est possible de montrer une norme d'opérateur  $L^2(\partial S) \rightarrow L^2(\partial S)$  en  $O(|\log \lambda_0| \sqrt{\lambda_0})$  pour l'application  $J \mapsto \nu_0^i(J \star G)$ , en utilisant le cas  $p = 0$  de la démonstration de la future propriété 4.3 page 155.

■ Soit  $J \in L^2(\partial S)$ . D'après la propriété 2.19,  $J \star G$  est une fonction continue. La définition de  $G$  par (28) et le développement limité (29d) donnent :

$$\begin{aligned} \nu_0^i(J \star G)(t) &= \frac{i}{4} \int_0^L \left( \frac{2i}{\pi} \log \frac{\|s(t) - s(t')\|}{L/\pi} - H_0^{ank} \left( \frac{2\pi}{\lambda_0} \|s(t) - s(t')\| \right) \right) J(t') dt' \\ &+ \frac{1}{2\pi} \int_0^L \log \frac{\|s(t) - s(t')\|}{L/\pi} J(t') dt' \end{aligned} \quad (54c)$$

**Un opérateur à noyau  $k(t, t')$  qui est  $C^1$ .**

D'après les développements limités (29d) et (29e), la première intégrale de (54c) a un opérateur dont le noyau  $k(t, t')$  est  $C^1$ , qui est donc en particulier borné. Nous avons besoin d'une estimation sur la dépendance en  $\lambda_0$  des normes  $C^1$  et  $L^\infty$  de ce noyau.

Watson [W22] montre, p196 partie 7.2, que les fonctions

$$H_0^{ank} \left( \frac{1}{t} \right) e^{-i\frac{1}{t}} \frac{1}{\sqrt{t}} \quad \text{et} \quad H_0^{ank'} \left( \frac{1}{t} \right) e^{-i\frac{1}{t}} \frac{1}{\sqrt{t}}$$

admettent en  $t \rightarrow 0^+$  des développements limités à tout ordre.

Une composition d'un développement limité avec ces développements limités montre que

$$t \mapsto H_c^{ank} \left( \frac{\sqrt{1-t^2}}{t} \right) e^{-i\frac{\sqrt{1-t^2}}{t}} \left( \frac{\sqrt{1-t^2}}{t} \right)^{-1/2}$$

admet en  $t \rightarrow 0^+$  des développements limités à tout ordre, pour  $c = 0$  ou pour  $c = 1$ . D'après les développements limités (29d) et (29e), les fonctions

$$x \mapsto H_0^{ank}(x) - \frac{2i}{\pi} \log \frac{x^2}{x^2 + 1} \quad \text{et} \quad x \mapsto H_1^{ank}(x) + \frac{2i}{\pi x}$$

sont continues sur  $\mathbb{R}_+$ . De plus, les fonctions

$$t \mapsto \frac{2i}{\pi} \log \left[ \left( \frac{\sqrt{1-t^2}}{t} \right)^2 / \left( \left( \frac{\sqrt{1-t^2}}{t} \right)^2 + 1 \right) \right] \quad \text{et} \quad t \mapsto \frac{2i}{\pi} / \frac{\sqrt{1-t^2}}{t}$$

admettent des développements limités en zéro à tout ordre. En ajoutant ces développements limités, on voit qu'il existe des coefficients complexes  $a_{c,i}$  tels que  $\forall n > 0, \exists M_n > 0, \forall x \in \mathbb{R}^+$  :

$$\left| H_0^{ank}(x) - \frac{2i}{\pi} \log \left( \frac{x^2}{1+x^2} \right) - \sum_{m=1}^n \left( \frac{a_{0,m} e^{ix}}{(x^2+1)^{\frac{m}{2}-\frac{1}{4}}} + \frac{a_{1,m}}{(x^2+1)^{\frac{m}{2}}} \right) \right| < \frac{M_n}{(1+x^2)^{\frac{n}{2}+\frac{1}{4}}} \quad (54d)$$

$$\left| H_1^{ank}(x) + \frac{2i}{\pi x} - \sum_{m=1}^n \left( \frac{a_{2,m} e^{ix}}{(x^2 + 1)^{\frac{m}{2} - \frac{1}{4}}} + \frac{a_{3,m}}{(x^2 + 1)^{\frac{m}{2}}} \right) \right| < \frac{M_n}{(1 + x^2)^{\frac{n}{2} + \frac{1}{4}}} \quad (54e)$$

L'intérêt de ces deux estimations poussées est qu'elles permettent de remplacer  $x$  par une fonction  $u(t, t')$  lorsque nous disposons d'estimations  $C^p$  connues sur  $u^2(t, t')$ , voir (89g).

Comme on a  $\lambda_0 < \frac{1}{2}$  (27a), ces estimations permettent de montrer que

$$|k(t, t')| = O(|\log \lambda_0|) \quad \text{et} \quad \left| \frac{\partial k}{\partial t}(t, t') \right| = O\left( \min\left( \frac{1}{\lambda_0}, \frac{1}{\sqrt{\lambda_0 |t - t'|}} \right) \right)$$

Montrons qu'un opérateur de la forme  $J \mapsto \int_0^L k(\cdot, t) J(t) dt$  est continu  $L^2(\partial S) \rightarrow L^2(\partial S)$  lorsque  $k$  est une fonction  $L^2([0, L]^2)$ . Une inégalité de Hölder donne en effet, pour  $f \in L^2(\partial S)$  :

$$\begin{aligned} \left( \int_0^L \int_0^L k(t', t) J(t) f(t') dt dt' \right)^2 &\leq \|k\|_2^2 \int_0^L \int_0^L |J(t)|^2 |f(t')|^2 dt dt' \\ &\leq \|k\|_2^2 \|J\|_{L^2(\partial S)}^2 \|f\|_{L^2(\partial S)}^2 \end{aligned} \quad (54f)$$

ce qui montre aisément la continuité désirée.

La première intégrale de (54c) a donc une norme  $L^2(\partial S)$  en  $O(|\log \lambda_0| \|J\|_{L^2(\partial S)})$ . Pour montrer que cette intégrale est  $H^1(\partial S)$ , nous devons dériver sous l'intégrale. L'opérateur

$$(t, t') \mapsto \frac{1}{\varepsilon} \left| k(t + \varepsilon, t') - k(t, t') - \frac{\partial k}{\partial t}(t, t') \varepsilon \right| \quad (54g)$$

est borné uniformément en  $\varepsilon$  petit, donc le théorème de convergence dominée montre que la première intégrale de (54c) est faiblement dérivable dans  $L^2(\partial S)$  par rapport à  $t$ , et la dérivée a une norme  $L^2(\partial S)$  en  $O\left(1 + \frac{|\log \lambda_0|}{\lambda_0}\right)^{\frac{1}{2}} \|J\|_{L^2(\partial S)}$ .

### Cas où $S$ est une boule.

Dans ce cas, la seconde intégrale de (54c) est étudiée dans [C84a] par exemple, et est la valeur en  $J$  d'un opérateur qui est diagonal en Fourier, car cet opérateur est une convolution. Calculons explicitement les coefficients de Fourier de cet opérateur :

Préparons la fonction noyau de convolution avec un changement de variables en prenant  $t, t' \in [0, 2\pi]$  :

$$\log \frac{\left| s\left(\frac{L}{2\pi}t\right) - s\left(\frac{L}{2\pi}t'\right) \right|}{L/\pi} = \log \frac{|e^{it} - e^{it'}|}{2} = \log \sin \frac{|t - t'|}{2} = \frac{1}{2} \log \frac{1 - \cos(t - t')}{2} \quad (54h)$$

Or  $\frac{1}{2} \log \frac{1-\cos t}{2}$  est la valeur en  $r = 1$  d'une fonction dont la dérivée en  $t$  vaut

$$\frac{1}{2} \left( \log(1 + r^2 - 2r \cos t) \right)' = \frac{2r \sin t}{1 + r^2 - 2r \cos t} = 2\mathcal{I}m \frac{1}{1 - re^{it}}$$

Lorsque  $r < 1$ , cette dérivée vaut  $2\mathcal{I}m \sum_{n=1}^{\infty} r^n e^{int} = 2 \sum_{n=1}^{\infty} r^n \sin(nt)$  et donc on a

$$\frac{1}{2} \log(1 + r^2 - 2r \cos t) = C - 2 \sum_{n=1}^{\infty} \frac{r^n}{n} \cos(nt)$$

avec  $C$  une constante. Le produit scalaire de cette égalité avec  $e^{in't}$  converge lorsque  $r \rightarrow 1$ , ce qui montre que l'opérateur de la seconde intégrale de (54c) a des coefficients de Fourier en  $O(\frac{1}{n})$  : alors cette seconde intégrale est  $H^1(\partial S)$  uniformément en  $J \in L^2(\partial S)$  d'après la propriété 2.10.

### Cas plus général où $S$ n'a pas de coin.

Dans ce cas, nous déformons  $\partial S$ , ce qui fait qu'au membre droit de (54h) il faut ajouter ceci :

$$\begin{aligned} & \frac{1}{2} \log \left( 1 + \left( \sin \frac{|t-t'|}{2} \right)^{-2} \left( \frac{\pi^2}{L^2} \left\| s\left(\frac{L}{2\pi}t\right) - s\left(\frac{L}{2\pi}t'\right) \right\|^2 - \sin^2 \frac{|t-t'|}{2} \right) \right) \quad (54i) \\ & = \frac{1}{2} \log \left( 1 + \left( \sin \frac{|t-t'|}{2} \right)^{-2} \left( \int_t^{t'} \int_t^{t'} \left( \frac{\langle s'(\frac{L}{2\pi}u) | s'(\frac{L}{2\pi}v) \rangle}{4} - \frac{\cos(u-v)}{4} \right) dv du \right) \right) \end{aligned}$$

L'objectif est de montrer que ceci est  $C^1$  en  $t$  et  $t'$ . Comme  $\langle \frac{L}{2\pi} s'(u) | \frac{L}{2\pi} s'(u) \rangle = 1$  puis comme  $\langle s''(\frac{L}{2\pi}w) | s'(\frac{L}{2\pi}w) \rangle = 0$ , on a

$$\begin{aligned} & \frac{\langle s'(\frac{L}{2\pi}u) | s'(\frac{L}{2\pi}v) \rangle}{4} - \frac{\cos(u-v)}{4} \\ & = \int_u^v \frac{\langle \frac{L}{2\pi} s''(\frac{L}{2\pi}w) | s'(\frac{L}{2\pi}v) \rangle + 1 - 1 + 2 \sin^2(v-w)}{4} dw \\ & = \int_u^v \int_w^v \frac{-\langle \frac{L}{2\pi} s''(\frac{L}{2\pi}w) | \frac{L}{2\pi} s''(\frac{L}{2\pi}w') \rangle + 4 \sin(w'-w) \cos(w'-w)}{4} dw' dw \end{aligned}$$

La dernière double-intégrale ci-dessus admet un développement limité du type

$$a(v-u)^2 + o((v-u)^2)$$

où  $a$  est nombre réel dépendant continûment de  $u$ , car  $s$  est  $C^2$ . Donc la dérivée en  $t$  de l'argument du logarithme dans (54i) admet un équivalent du type

$$\sum_{n=0}^2 a_n (t-t')^n + o((t-t')^2)$$

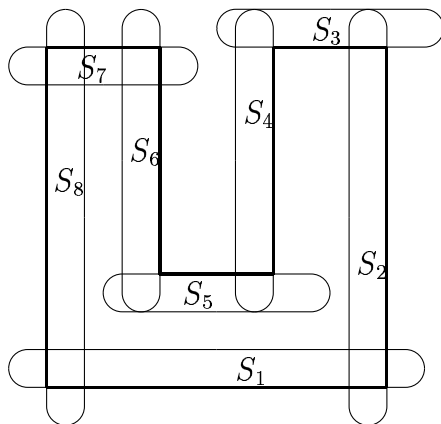


FIG. 15 – Une partition chevauchante de  $\partial S$ .

où  $a_n$  est un nombre réel dépendant continûment de  $t'$ . L'argument du logarithme dans (54i), qui est

$$\frac{\frac{\pi^2}{L^2} \left\| s\left(\frac{L}{2\pi}t\right) - s\left(\frac{L}{2\pi}t'\right) \right\|^2}{\sin^2 \frac{|t-t'|}{2}},$$

est donc  $C^1$  en  $t$  uniformément en  $t'$ . Comme cet argument est aussi dans  $[c, C]$  pour  $c > 0$  et  $C \geq c$  deux constantes réelles, (54i) est  $C^1$  en  $t$  uniformément en  $t'$ .

Le même raisonnement ci-dessus autour de (54g) montre alors que  $\nu_0^i(J \star G)$  est  $H^1(\partial S)$  uniformément en  $J \in L^2(\partial S)$  dans ce cas aussi.

### Cas général où $S$ peut avoir des coins.

Nous allons prolonger les morceaux  $C^p$  de  $\partial S$  de façon à nous ramener au cas précédent. Soit  $S_1, S_2, \dots, S_n$  des fermés  $C^p$  tels que

$$\{\partial S_m \cap \partial S, m = 1..n\} \tag{54j}$$

soit une partition dans un sens affaibli de  $\partial S$  qui soit adaptée à la continuité  $C^p$  par morceau de  $\partial S$ . Ce n'est pas une partition au sens strict, car nous devons tolérer le chevauchement pour un nombre fini de points de  $\partial S$ , comme sur la figure 15. Comme la propriété 2.19 montre que  $J \star G$  est continu, il suffit de montrer que la seconde intégrale de (54c) est  $H^1(\partial S_m \cup \partial S)$  en  $t$  pour un  $m$  donné : le recollement des valeurs de  $J \star G$  aux coins suffit alors à montrer

que les dérivées (au sens faible) sur chaque morceau forment, ensemble, la dérivée (au sens faible) de  $J \star G$  sur  $\partial S$ .

La restriction de la seconde intégrale de (54c) à  $t' \in s^{-1}(\partial S_m \cup \partial S)$  est une fonction  $H^1(cS)$  en  $t$  comme montré par le cas précédent, où  $S$  n'avait pas de coins.

La dérivée en  $t$  de la restriction de la seconde intégrale de (54c) à  $t' \notin s^{-1}(\partial S_m \cup \partial S)$  est un opérateur à noyau en

$$O\left(\frac{1}{\|s(t) - s(t')\|}\right) = O\left(\frac{1}{|t - t_0| + |t' - t_0|} + \frac{1}{|t - t_1| + |t' - t_1|}\right)$$

où  $t_0$  et  $t_1$  sont les extrémités de  $s^{-1}(\partial S_m \cup \partial S)$ . Nous pouvons reprendre le début de la démonstration de la propriété 2.25 tel quel, ce qui montre que la dérivée en  $t$  de la restriction de la seconde intégrale de (54c) à  $t' \notin s^{-1}(\partial S_m \cup \partial S)$  est  $L^2(\partial S)$ : la restriction de la seconde intégrale de (54c) à  $t' \notin s^{-1}(\partial S_m \cup \partial S)$  est donc  $L^2(\partial S)$  avec la même démonstration, et donc est aussi  $H^1(\partial S)$ .

Ainsi, la seconde intégrale de (54c) est  $H^1(\partial S)$  en  $t$  continûment par rapport à  $J \in L^2(\partial S)$ .

La démonstration est terminée. Signalons une idée pour une autre démonstration, qui est toutefois fautive: (48b) donne la première inégalité ci-dessous

$$\left| \int E(x) \Delta f(x) dx \right| \leq \|J\|_{L^2(\partial S)} \|\nu_0 f\|_{L^2(\partial S)} \leq ??? C \|J\|_{L^2(\partial S)} \|f\|_{H^{\frac{1}{2}}(\mathbb{R}^2)}$$

mais la seconde inégalité est malheureusement fautive: ainsi,  $J \star G$  a peu de chances d'être  $H^{\frac{3}{2}}(\mathbb{R}^2)$ . Notons que  $J \star G$  est  $H^{\frac{3}{2}}(S)$  et  $H^{\frac{3}{2}}(\partial S)$  (résultat difficile, voir Costabel-Dauge [CD93]). ■

**Théorème 2.28** *L'application  $J \mapsto \nu_1^i(J \star G)$  est continue de  $L^2(\partial S)$  dans  $L^2(\partial S)$ ; c'est la somme d'une application continue coercive de norme en  $O(1)$  et dont l'inverse a une norme en  $O(1)$  lorsque  $\lambda_0 \rightarrow 0$ , et d'une application continue  $L^2(\partial S) \rightarrow H^1(\partial S)$  de norme en  $O(\sqrt{\frac{|\log \lambda_0|}{\lambda_0^3}})$  qui est aussi une application continue  $L^2(\partial S) \rightarrow L^2(\partial S)$  de norme en  $O(\sqrt{\frac{|\log \lambda_0|}{\lambda_0}})$ .*

Cette décomposition permet d'isoler les singularités relatives aux coins; l'idée vient de Elschner [E93], sauf en ce qui concerne la dépendance en  $\lambda_0$ .

■ Soit  $J \in L^2(\partial S)$ . Reprenons la « partition chevauchante » définie par  $S_1, S_2, \dots, S_n$  en (54j). Nous utiliserons  $S_0$  comme un synonyme de  $S_n$ , et  $S_{n+1}$  comme un synonyme de  $S_1$ .

Notons  $t_m$  l'abscisse du coin à l'intersection entre  $\partial S_{m-1}$  et  $\partial S_m$ . On peut définir rigoureusement  $t_m$  en disant qu'il est l'unique point d'accumulation de  $\partial S_{m-1} \cap \partial S$  et  $\partial S_m \cap \partial S$  à la fois. Le terme « point d'accumulation » a ici été introduit afin d'éviter les points (en nombre fini) à cause desquels  $(\partial S_m \cap \partial S)_{1 \leq m \leq n}$  ne forme pas une partition de  $\partial S$ , mais seulement une « partition chevauchante ».

Pour chaque  $m$  tel que  $1 \leq m \leq n$ , et pour  $c = 0$  puis  $c = 1$ , soit une fonction  $\phi_{m,c} : \partial S_m \cap \partial S \rightarrow [0,1]$  qui vaille  $1 - c$  au voisinage de  $t_m$  et qui vaille  $c$  au voisinage de  $t_{m+1}$ . Prolongerons  $\phi_{m,c}$  par zéro sur le reste de  $\partial S$ ; nous pouvons imposer l'égalité presque sûre suivante :

$$\phi_{m,0} + \phi_{m,1} = \mathbb{1}_{\partial S_m \cap \partial S} \quad \text{en } \partial S$$

Nous pouvons en déduire que

$$\sum_{m=1}^n \sum_{c=0}^1 \phi_{m,c}(t) = 1 \quad \text{et} \quad \sum_{m=1}^n \sum_{c=0}^1 \phi_{m,c}(t)^2 \leq 1 \quad t \in [0,L]. \quad (54k)$$

D'après la propriété 2.23, le nombre  $-\nu_1^i(J \star G)(s(t))$  est la somme des termes suivants :

$$-\frac{1}{2}J(t) \quad (54\ell)$$

$$\int_0^L \sum_{m=1}^n (\phi_{m,0}(t) + \phi_{m,1}(t)) \langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle J(t') (\phi_{m,0}(t') + \phi_{m,1}(t')) dt' \quad (54m)$$

$$\int_0^L \sum_{m=1}^n \phi_{m,0}(t) \langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle J(t') \phi_{m-1,1}(t') dt' \quad (54n)$$

$$\int_0^L \sum_{m=1}^n \phi_{m,1}(t) \langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle J(t') \phi_{m+1,0}(t') dt' \quad (54o)$$

$$\int_0^L \sum_{m=1}^n \phi_{m,0}(t) \langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle J(t') (1 - \phi_{m,0}(t') - \phi_{m,1}(t') - \phi_{m-1,1}(t')) dt' \quad (54p)$$

$$\int_0^L \sum_{m=1}^n \phi_{m,1}(t) \langle s'(t)^\perp | \nabla G(s(t) - s(t')) \rangle J(t') (1 - \phi_{m,0}(t') - \phi_{m,1}(t') - \phi_{m-1,1}(t')) dt' \quad (54q)$$

### Un opérateur coercif.

À l'opérateur coercif  $J \mapsto (54\ell)$ , nous pouvons ajouter un opérateur de norme  $L^2$  strictement plus petite que  $\frac{1}{2}$ , ce qui est le cas des parties principales

$$-\int_0^L \sum_{m=1}^n \frac{1}{2\pi} \frac{\phi_{m,0}(s(t))(t' - t_m) \phi_{m-1,1}(s(t')) J(t') dt' \sin \theta(t_m)}{(t - t_m)^2 + (t' - t_m)^2 + 2(t - t_m)(t' - t_m) \cos \theta(t_m)} \quad (54r)$$



et

$$\int_0^L \sum_{m=1}^n \frac{1}{2\pi} \frac{\phi_{m,1}(s(t))(t' - t_{m+1})\phi_{m+1,0}(s(t'))J(t')dt' \sin \theta(t_{m+1})}{(t - t_{m+1})^2 + (t' - t_{m+1})^2 + 2(t - t_{m+1})(t' - t_{m+1}) \cos \theta(t_{m+1})} \quad (54s)$$

des termes (54n) et (54o) selon la propriété 2.25.

Sur chacun des intervalles  $\partial S_m$ , chacune des parties principales (54r) et (54s) est  $C$ -contractante avec un facteur  $C$  strictement plus petit que  $\frac{1}{2}$ . Soit  $C$  la plus grande de ces constantes: on a  $C < \frac{1}{2}$ .

L'intégrale en  $t$  contre une fonction  $f \in L^2(\partial S)$  de (54r) et (54s) vérifie respectivement :

$$\begin{aligned} \left| \int_0^L (54r) f(t) dt \right| &\leq \sum_{m=1}^n C \|f \phi_{m,0}\|_{L^2} \|J \phi_{m-1,1}\|_{L^2} \\ \left| \int_0^L (54s) f(t) dt \right| &\leq \sum_{m=1}^n C \|f \phi_{m,1}\|_{L^2} \|J \phi_{m+1,0}\|_{L^2} \end{aligned}$$

Une inégalité de Cauchy-Schwarz montre que la somme  $|\int_0^L ((54r) + (54s)) f(t) dt|$  est plus petite que

$$C \sqrt{\sum_{m=1}^n (\|f \phi_{m,0}\|_{L^2}^2 + \|f \phi_{m,1}\|_{L^2}^2)} \sqrt{\sum_{m=1}^n (\|J \phi_{m-1,1}\|_{L^2}^2 + \|J \phi_{m+1,0}\|_{L^2}^2)}$$

qui est lui-même plus petit que  $C \|f\|_{L^2} \|J\|_{L^2}$  grâce à (54k).

Ainsi, nous avons

$$\left(\frac{1}{2} - C\right) \|J\|_{L^2} \leq \|(54\ell) + (54r) + (54s)\|_{L^2} \leq \left(\frac{1}{2} + C\right) \|J\|_{L^2}$$

donc nous avons bien là un opérateur coercif.

### Recollement des valeurs pour la continuité du reste.

Nous pouvons encore ajouter un opérateur  $L^2 \rightarrow L^2$  à l'opérateur coercif ci-dessus sans perdre la coercivité, pourvu que sa norme d'opérateur  $L^2 \rightarrow L^2$  soit plus petite que  $\frac{1}{2} - C$ . Cette possibilité est très importante pour passer de l'espace  $\prod_{m=1}^n H^1(\partial S_m \cap \partial S)$  (obtenable par exemple en utilisant la propriété 2.24 pour le terme (54m)) à l'espace  $H^1(\partial S)$  (demandé dans l'énoncé).

Soit  $\phi_\varepsilon(t)$  la fonction  $\max(0, 1 - \inf_{m=1}^n |\frac{t-t_m}{\varepsilon}|)$ . On suppose  $\varepsilon < \lambda_0$ . Le terme

$$\phi_\varepsilon(t)(-\nu_1^i(J \star G)(t) - (54\ell) - (54r) - (54s)) \quad (54t)$$

est l'image de  $J$  par un opérateur dont le noyau, d'après la décomposition (54ℓ) à (54q), a un module en

$$\phi_\varepsilon(t) \left( 0 + \int_{t,t'} \begin{cases} O(1) & \text{si } t, t' \text{ dans le même } \partial S_m \\ 2O(\frac{1}{\lambda_0}) + 2O(\frac{1}{\sqrt{\lambda_0}}) & \text{sinon} \end{cases} \right) \quad (54u)$$

lorsque  $|t - t'| < \lambda_0$ , et en

$$\phi_\varepsilon(t) \left( 0 + \int_{t,t'} \begin{cases} O(\sqrt{\frac{|t-t'|}{\lambda_0}}) & \text{si } t, t' \text{ dans le même } \partial S_m \\ 2O(\frac{1}{\sqrt{\lambda_0|t-t'|}}) + 2O(\frac{1}{\sqrt{\lambda_0}}) & \text{sinon} \end{cases} \right) \quad (54v)$$

lorsque  $|t - t'| \geq \lambda_0$ , d'après la propriété 2.24 et, pour les termes (54p) et (54q), d'après le développement limité (32b) utilisé sur (52e). L'intégrale du carré de (54u) en  $t$  et  $t'$  est en

$$O(\varepsilon\lambda_0 + \varepsilon\lambda_0 \frac{1}{\lambda_0^2} + \varepsilon\lambda_0 \frac{1}{\lambda_0})$$

et l'intégrale du carré de (54v) en  $t$  et  $t'$  est en

$$O(\varepsilon \frac{1}{\lambda_0} + \varepsilon \frac{|\log \lambda_0|}{\lambda_0} + \varepsilon \frac{1}{\lambda_0}).$$

Donc l'opérateur  $J \mapsto (54t)$  a d'après le raisonnement en (54f), une norme  $L^2 \rightarrow L^2$  en  $O(\sqrt{\frac{\varepsilon}{\lambda_0}} |\log \lambda_0|)$ . Il existe donc une constante  $c$ , en rapport avec  $C$ , telle que l'opérateur  $J \mapsto (54t)$  avec  $\varepsilon = c\lambda_0/|\log \lambda_0|$  ait une norme plus petite que  $\frac{1}{2} - C$  indépendamment de  $\lambda_0$ .

### **Le reste est un opérateur compact.**

Nous avons donc montré que  $J \mapsto \nu_1^i(J \star G)$  était la somme d'un opérateur coercif  $L^2(\partial S) \rightarrow L^2(\partial S)$ , et de l'opérateur qui à  $J$  associe

$$(1 - \phi_\varepsilon(t))(-\nu_1^i(J \star G)(t) - (54\ell) - (54r) - (54s)). \quad (54w)$$

### **Le reste est un opérateur continu, cas $L^2 \rightarrow L^2$ .**

Nous allons montrer que cet opérateur est à valeur dans  $L^2(\partial S)$ , ensuite nous évaluerons la dérivée en  $t$ . Ce terme (54w) est l'image de  $J$  par un opérateur dont le noyau, d'après la décomposition (54ℓ) à (54q), a un module borné par (54u) avec  $\phi_\varepsilon = 1$  ou (54v) avec  $\phi_\varepsilon = 1$  selon que  $|t - t'| < \lambda_0$

ou non. L'intégrale du carré de (54u) avec  $\phi_\varepsilon = 1$  en  $t$  et  $t'$  sur le domaine  $|t - t'| < \lambda_0$  est en

$$O(\lambda_0 + \lambda_0^2 \frac{1}{\lambda_0^2} + \lambda_0^2 \frac{1}{\lambda_0})$$

L'intégrale du carré de (54v) avec  $\phi_\varepsilon = 1$  en  $t$  et  $t'$  sur le domaine  $|t - t'| \geq \lambda_0$  est en

$$O(\frac{1}{\lambda_0} + \frac{|\log \lambda_0|}{\lambda_0} + \frac{1}{\lambda_0})$$

et donc le raisonnement en (54f) montre que la norme  $L^2$  de (54w) est en  $O(\sqrt{\frac{|\log \lambda_0|}{\lambda_0}})$ .

**Le reste est un opérateur compact, cas  $L^2 \rightarrow H^1$ .**

La dérivée en  $t$  de ce terme (54w) est l'image de  $J$  par un opérateur dont le noyau, d'après la décomposition (54ℓ) à (54q), a un module en

$$0 + \int_{t,t'} \begin{cases} O(\frac{1}{\lambda_0}) & \text{si } t, t' \text{ dans le même } \partial S_m \\ 2O(\frac{1}{|t-t'|}) + 2O(\lambda_0^{-\frac{3}{2}}) & \text{sinon} \end{cases} + \frac{|\phi'_\varepsilon(t)|}{\phi_\varepsilon(t)}(54u) \quad (54x)$$

lorsque  $|t - t'| < \lambda_0$ , et en

$$0 + \int_{t,t'} \begin{cases} O(\sqrt{\frac{|t-t'|}{\lambda_0^3}}) & \text{si } t, t' \text{ dans le même } \partial S_m \\ 2O(\frac{1}{\sqrt{\lambda_0^3|t-t'|}}) + 2O(\lambda_0^{-\frac{3}{2}}) & \text{sinon} \end{cases} + \frac{|\phi'_\varepsilon(t)|}{\phi_\varepsilon(t)}(54v) \quad (54y)$$

lorsque  $|t - t'| \geq \lambda_0$ , d'après la propriété 2.24 et, pour les termes (54p) et (54q), d'après les développements limités (32b) et (53) utilisés sur (52e).

Le terme en  $O(\frac{1}{|t-t'|})$  présent dans (54x) mérite un traitement spécial : le raisonnement (54f) ne marche pas, mais heureusement il suffit de reprendre le raisonnement (54a). Ainsi, nous pouvons éliminer ce terme de (54x) pour la suite de la démonstration, il suffira à la fin d'ajouter  $O(1)$  à l'estimation de la norme d'opérateur  $L^2(\partial S) \rightarrow H^1(\partial S)$ .

L'intégrale du carré de (54x) en  $t$  et  $t'$  sur le domaine  $|t - t'| < \lambda_0$  est en

$$O(\frac{1}{\lambda_0} + 0 + \lambda_0^2 \frac{1}{\lambda_0^3} + \frac{1}{\varepsilon}(\varepsilon \lambda_0 + \varepsilon \lambda_0 \frac{1}{\lambda_0^2} + \varepsilon \lambda_0 \frac{1}{\lambda_0}))$$

L'intégrale du carré de (54y) en  $t$  et  $t'$  sur le domaine  $|t - t'| \geq \lambda_0$  est en

$$O(\frac{1}{\lambda_0^3} + \frac{|\log \lambda_0|}{\lambda_0^3} + \frac{1}{\lambda_0^3} + \frac{1}{\varepsilon}(\varepsilon \frac{1}{\lambda_0} + \varepsilon \frac{|\log \lambda_0|}{\lambda_0} + \varepsilon \frac{1}{\lambda_0})).$$

La somme de ces deux valeurs et le remplaçant de  $\varepsilon$  donne l'estimation  $O(\sqrt{\frac{|\log \lambda_0|}{\lambda_0^3}})$  de la norme d'opérateur  $L^2(\partial S) \rightarrow H^1(\partial S)$  de l'opérateur  $J \mapsto (54w)$ . ■

En rassemblant les résultats des deux lemmes ci-dessus, nous avons une inégalité de Gårding sur l'opérateur  $A$

$$A : J \mapsto \frac{i\alpha}{\lambda_0} \nu_0^i(J \star G) + 2\nu_1^i(J \star G) \quad (55)$$

qui est présent dans l'équation (21h) que nous voulons résoudre :

**Théorème 2.29 (inégalité de Gårding)** *Il existe un opérateur compact  $T : L^2(\partial S) \rightarrow H^1(\partial S)$  et une constante positive  $c_A < 1$  indépendante de  $\lambda_0$  tels que pour tout  $J \in L^2(\partial S)$ ,*

$$\langle AJ - TJ - J|J \rangle \geq c_A \|J\|_{L^2(\partial S)}^2. \quad (56)$$

La constante  $c_A$  est indépendante de  $\lambda_0$  ; malheureusement, l'opérateur  $T$ , tout comme  $A$ , n'est pas indépendant de  $\lambda_0$  : par exemple, on a les estimation

$$\begin{aligned} \|TJ\|_{L^2(\partial S)} &= O(\lambda_0^{-1} \|J\|_{L^2(\partial S)} |\log \lambda_0|) \\ \|TJ\|_{H^1(\partial S)} &= O(\lambda_0^{-3} \|J\|_{L^2(\partial S)} |\log \lambda_0|) \end{aligned}$$

Lorsqu'un sous-ensemble du bord diffractant est droit, on peut montrer que la dernière estimation est optimale au facteur  $|\log \lambda_0|$  près.

Cette inégalité de Gårding fait par exemple que les petites et les grandes valeurs singulières de  $A$  n'ont pas de point d'accumulation, ce qui rend très rapide le calcul numérique des normes d'opérateur  $L^2(\partial S) \rightarrow L^2(\partial S)$  de  $A$  et de  $A^{-1}$ . Malheureusement, nous n'avons pas réussi à utiliser pleinement cette inégalité de Gårding pour par exemple prouver que  $\|A^{-1}\|$  est borné indépendamment de  $\lambda_0$  bien que les expérimentations numériques pour le cas du diffractant sans coin, et une analogie par Peterson-Mittra-Ray [PMR99] avec l'optique géométrique puisse le suggérer.

#### 2.3.4 Existence et unicité dans la CFIE.

**Théorème 2.30** *Pour toute fonction  $E^{inc}$  qui est régulière au voisinage de  $S$  et qui vérifie l'équation de Helmholtz (21d) sur  $S$ , le problème (21h) qui*

consiste à trouver une fonction  $J \in L^2(\partial S)$  telle que

$$\frac{i\alpha}{\lambda_0} \nu_0^i(J \star G) + 2\nu_1^i(J \star G) = -\frac{i\alpha}{\lambda_0} \nu_0 E^{inc} - 2\nu_1^i E^{inc} \quad (57)$$

admet une unique solution, c'est à dire que le problème dont  $J$  est solution est bien posé au sens de Hadamard. De plus, la fonction  $E = J \star G$  est solution du problème (21c).

■ L'unicité dans  $L^2(\partial S)$  découle simplement de celle dans  $H^{-\frac{1}{2}}(\partial S)$  vue à la propriété 2.20. L'inégalité de Gårding (56) va permettre de montrer l'existence.

Nous savons que  $(A - T - \text{Id})$  est un opérateur  $c_A$ -contractant. Comme  $T$  est compact, la décomposition

$$A = \text{Id} + (A - T - \text{Id}) + T$$

montre que le spectre essentiel (constitué des points d'accumulation du spectre complet qui est l'ensemble des nombres complexes  $\lambda$  tels que  $A - \lambda \text{Id}$  n'est pas bijectif) de  $A$  est le spectre essentiel de  $J + (A - T - \text{Id})$  qui est inclus dans l'ensemble des nombres complexes de rayon  $c_A$  et de centre 1. Comme  $c_A < 1$ , on voit que 0 est dans un voisinage (inclus dans  $\mathbb{C}$ ) où le spectre de  $A$  n'a pas de points d'accumulation: les conditions sont réunies pour la fin de la démonstration, basé sur le principe de l'alternative de Fredholm. L'unicité ci-dessus montre alors que 0 n'est pas valeur propre de  $A$ . Alors, le spectre complet de  $A$  n'intersecte pas un voisinage de 0, et donc  $A$  est inversible  $L^2(\partial S) \rightarrow L^2(\partial S)$ .

Pour prouver que  $E = J \star G$  vérifie (21c), il suffit alors d'utiliser le théorème 2.14. ■

### 3 La discrétisation de la CFIE.

#### 3.1 Besoin de l'analyse harmonique de la CFIE.

Supposons qu'on a un sous-espace vectoriel  $V_h$  de  $L^2(\partial S)$  qui contient « à peu près » tous les vecteurs propres de  $A - \text{Id}$  ayant les valeurs propres de partie imaginaire plus grande que  $\frac{1-cA}{2}$  dans le sens où la norme

$$\|(\text{Id} - P_h)AP_h\|_{L^2(\partial S) \rightarrow L^2(\partial S)} = \|(\text{Id} - P_h)(A - \text{Id})P_h\|_{L^2(\partial S) \rightarrow L^2(\partial S)} \quad (58)$$

est strictement plus petit que  $\frac{1-cA}{2}$ , avec  $P_h$  la projection orthogonale sur  $V_h$ . L'utilisation d'un tel espace  $V_h$  se ferait alors via l'équation variationnelle suivante

$$P_hAJ_h = P_hL_\alpha. \quad (59a)$$

avec  $L_\alpha = AJ$  et  $J$  la solution exacte de l'équation (57), et où  $J_h$  est l'approximation numérique de  $J$ , recherchée dans l'espace de discrétisation  $V_h$ . Comme  $J$  est la solution,  $AJ$  vaut le second membre de l'équation non discrétisée (57), et donc le terme  $P_hAJ$  est la discrétisation du second membre. Lorsque  $V_h$  est de dimension finie, l'équation variationnelle (59a) est une équation en dimension finie, c'est même un système carré, qu'on appelle *système discrétisé de Galerkin*.

Intuitivement, lorsque (58) est suffisamment petit,  $P_hAP_h$  modélise « suffisamment bien »  $A$  sur  $V_h$ , car les vecteurs propres gênants (qui sont en nombre fini comme  $T$  est compact) sont neutralisés : alors nous pouvons nous attendre à ce que

$$\|J - J_h\|_{L^2(\partial S)} \leq c\|J - P_hJ\|_{L^2(\partial S)} \quad (59b)$$

avec  $c$  une constante indépendante de  $J$ . Ceci peut même se montrer :

**Théorème 3.1** *Lorsque le projecteur orthogonal  $P_h$  sur l'espace de discrétisation  $V_h$  vérifie*

$$\|(Id - P_h)AP_h\|_{L^2(\partial S) \rightarrow L^2(\partial S)} < \frac{1}{2}\|A^{-1}\|_{L^2(\partial S) \rightarrow L^2(\partial S)}^{-1} \quad (60)$$

alors on a (59b) avec  $c$  dépendant de  $\lambda_0$ . De plus, lorsque

$$\|P_hA(Id - P_h)\|_{L^2(\partial S) \rightarrow L^2(\partial S)} < \frac{1}{2}\|A^{-1}\|_{L^2(\partial S) \rightarrow L^2(\partial S)}^{-1}, \quad (61)$$

qui est l'analogie de (60) mais avec  $A^*$  à la place de  $A$ , on a

$$\|J_h - J\|_{L^2(\partial S)} \leq 2\|J - P_hJ\|_{L^2(\partial S)} \quad (62)$$

La partie 4 de cette thèse se base sur l'estimation (62). L'estimation plus classique (58) est moins utile pour une étude asymptotique en  $\lambda_0$ , car le nombre  $\|A\|\|A^{-1}\|$  de conditionnement de  $A$  grandit lorsque  $\lambda_0$  tend vers 0.

Lorsque le diffractant  $S$  présente des coins, il est assez difficile d'obtenir (60), comme les expérimentations numériques de la figure 5 page 195 le montreront, avec un espace  $V_h$  qui reste invariant par translation. La section 3.3.7 contient des références à ce sujet.

■ Utilisons abusivement la notation  $\|\cdot\|$  aussi bien pour la norme de  $L^2(\partial S)$  que pour la norme d'opérateur  $L^2(\partial S) \rightarrow L^2(\partial S)$ . Posons

$$c_h = \|(\text{Id} - P_h)AP_h\| \quad \text{et} \quad c_h^* = \|P_hA(\text{Id} - P_h)\|$$

Soit  $u_h \in V_h$ , on a :

$$\begin{aligned} & \|P_hAu_h\| \|(A^*)^{-1}\| \|u_h\| \\ & \geq \langle P_hAu_h | (A^*)^{-1}u_h \rangle \\ & = \langle u_h | u_h \rangle - \langle (\text{Id} - P_h)AP_hu_h | (A^*)^{-1}u_h \rangle \geq (1 - c_h \|(A^*)^{-1}\|) \|u_h\|^2 \end{aligned}$$

ce qui montre que  $\|P_hAu_h\| \geq (\|(A^*)^{-1}\|^{-1} - c_h) \|u_h\|$  c'est à dire que  $P_hAP_h$  est aussi « bien » inversible que  $A$  : donc la discrétisation de  $A$  dans  $V_h$  est fidèle. Par exemple, en prenant  $u_h = J_h - P_hJ$  et en utilisant  $\|(A^*)^{-1}\| = \|A^{-1}\|$  et  $\|P_h\| = 1$ , on a :

$$\left(\frac{1}{\|A^{-1}\|} - c_h\right) \|J_h - P_hJ\| \leq \|P_hAP_h(J_h - P_hJ)\| = \|P_hA(\text{Id} - P_h)^2J\|$$

et nous rappelons que (59a) définit  $J_h$ . Pour prouver (62) on utilise que  $\text{Id} - P_h$  est un projecteur. Finalement, on a

$$\|J_h - P_hJ\| \leq \frac{c_h^*}{\|A^{-1}\|^{-1} - c_h} \|J - P_hJ\| \quad (63a)$$

■

L'énoncé de ce théorème 3.1 fait apparaître le terme  $J - P_hJ$  qui est l'erreur minimale sur la solution due à l'espace de discrétisation choisi. Pour estimer cette erreur à partir du terme  $L_\alpha - P_hL_\alpha$  qui est l'erreur due à la discrétisation du second membre  $L_\alpha$ , nous devons supposer que

$$\|(\text{Id} - P_h)(A - \text{Id})(\text{Id} - P_h)\| \quad (64)$$

est plus petit que  $\frac{1}{2}$  :

**Propriété 3.2** *Si la norme d'opérateur (64) est plus petite que  $\frac{1}{2}$ , alors*

$$\|J - P_hJ\|_{L^2(\partial S)} \leq 2\|(\text{Id} - P_h)AP_h\|_{L^2(\partial S) \rightarrow L^2(\partial S)} \|J\|_{L^2(\partial S)} + 2\|P_hL_\alpha - L_\alpha\|_{L^2(\partial S)}.$$

■ On a

$$\begin{aligned}
& \frac{1}{2} \|(\text{Id} - P_h)J\| \\
& \geq \|(\text{Id} - P_h)(A - \text{Id})(\text{Id} - P_h)\| \|(\text{Id} - P_h)J\| \\
& \geq \|(\text{Id} - P_h)\text{Id}(\text{Id} - P_h)^2 J\| - \|(\text{Id} - P_h)AP_h J\| - \|(\text{Id} - P_h)AJ\| \\
& = \|(\text{Id} - P_h)J\| - \|(\text{Id} - P_h)AP_h J\| - \|(\text{Id} - P_h)L_\alpha\|
\end{aligned}$$

ce qui donne le résultat demandé. Une étude plus poussée donne

$$(1 - \|(\text{Id} - P_h)(A - \text{Id})(\text{Id} - P_h)\|) \|J - P_h J\| \leq \|(\text{Id} - P_h)AP_h J\| + \|(\text{Id} - P_h)L_\alpha\|$$

■

### 3.2 Discrétisations vérifiant une inégalité directe.

D'après la propriété précédente,  $V_h$  doit être choisi de sorte que  $A$  vérifie (60) et (61) (qui sont deux hypothèses indispensables pour valider le résultat (62) du théorème 3.1).

Dans le cas sans coin, nous savons que l'opérateur  $A - \text{Id}$  impliqué dans (58) est compact, et même à valeur dans  $H^1(\partial S)$ , d'après la sous-section 2.3.3. Si par exemple  $V_h$  est fidèle à  $H^1(\partial S)$  au sens où on a une inégalité directe du type Jackson

$$\|f - P_h f\|_{L^2(\partial S)} \leq Ch \|f\|_{H^1(\partial S)}, \quad (65a)$$

alors lorsque  $h$  diminue, (58) décroît en

$$O(h \|A\|_{L^2(\partial S) \rightarrow H^1(\partial S)}),$$

donc (58) et le théorème 3.1 montrent que la discrétisation est stable pour  $h$  assez petit à  $\lambda_0$  fixé, c'est à dire qu'il existe  $h_0 = h_0(\lambda_0)$  tel que pour tout  $h < h_0$ , on a (59b).

Nous citerons trois possibilités usuelles pour les espaces  $V_h$  parmi les sous-espaces de  $L^2(\partial S)$ , en choisissant des espaces de dimension  $\frac{L}{h}$  où  $L$  est la longueur du diffractant  $S$ , ayant fait l'objet des définitions 2.9 et (33) :

- l'espace des fonctions (à valeur complexe) constantes par morceaux sur une partition du bord du diffractant  $\partial S$  en morceaux de taille  $h$  ;
- Les fonctions d'échelles, où l'espace  $V_h = V_{2^{-j}}$  est engendré par  $\frac{L}{h}$  translations bien réparties d'une fonction  $\psi_{j,0}^{(0)}$  auto-similaire au sens fractal ;



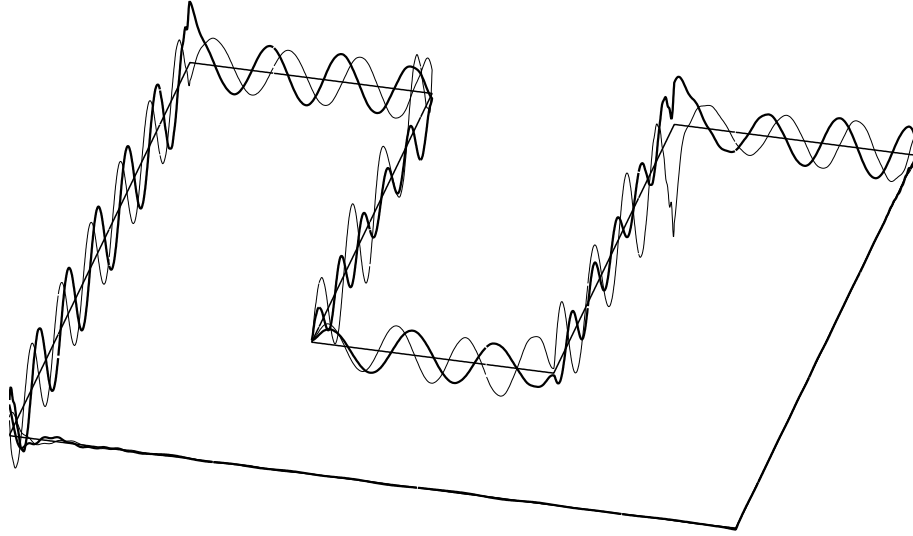


FIG. 16 – *Partie réelle (en gras) et imaginaire de la solution  $J$  pour  $\lambda_0 = \frac{1}{4}$ ,  $\beta = 31^\circ$ .*

- Les cosinus locaux, où l'espace  $V_h$  est engendré par  $\sqrt{L/h}$  translatées d'une fonction « bosse », multipliée par  $\sqrt{L/h}$  fonctions en cosinus de fréquences inférieures à  $1/2h$ .

Les deux dernières propositions pour  $V_h$  admettent des bases canoniques (de dimension  $L/h$ ).

### 3.2.1 Les fonctions constantes par morceau.

Lorsque  $S$  est donné par (33) (qui a pour périmètre  $L = 16$ ),  $\lambda_0 = \frac{1}{4}$  et  $\beta = 31^\circ$ , la solution  $J$  a l'allure indiquée sur la figure 16. Une première idée pour la discrétiser est de découper le bord du diffractant  $\partial S$  en petits segments droits de longueur  $h$  où on appelle  $h$  le *pas de discrétisation*, et de décider que  $V_h$  ne contiendra que des fonctions constantes sur chacun des  $\frac{L}{h}$  segments, ce qui donne par exemple la fonction de la figure 17. Ceci est un cas particulier du cas suivant des paquets d'ondelettes, et ainsi le théorème 4.8 montrera que pour  $h$  assez petit ceci mène à un problème bien posé. Le

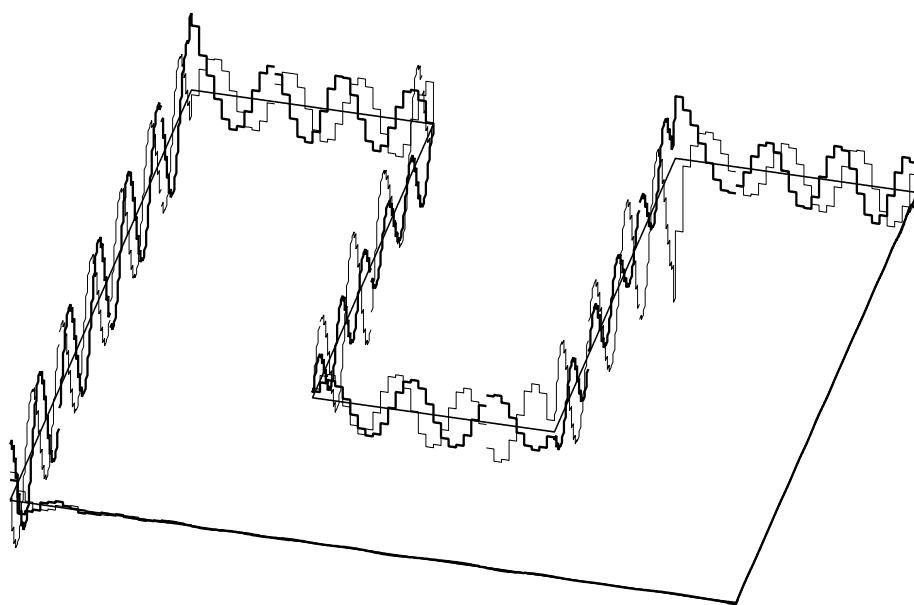


FIG. 17 – Une fonction  $P_h J$  avec un nombre fini de degrés de liberté, pour calculer la fonction à valeurs complexes  $J$  de la figure 16, avec le pas  $h = \frac{1}{32}$ .

calcul de la matrice du système résultant de cette discrétisation par (59a) mène au calcul d'intégrales doubles qui sont nombreuses (il y en a  $\frac{L^2}{h^2}$ ).

Il est tentant d'approcher ces intégrales, par exemple en isolant la singularité du noyau de l'opérateur de  $A$  qu'on discrétise dans une famille de Fourier  $L$ -périodique, et en discrétisant le reste (de noyau  $C^p$ ) avec des méthodes de qualocation de Chandler-Sloan [CS90], Sloan [S91] [S00] ou Kress-Sloan [KS93], ou des méthodes de Nyström présentée par Rokhlin [R83] et Alpert-Beylkin-Coifman-Rokhlin [ABC93]. Notons qu'il se trouve beaucoup d'autres façons de discrétiser avec en quelque sorte des masses de Dirac, qu'on retrouvera dans la bibliographie ; néanmoins nous n'avons cité ci-dessus que les approches qui ont leur justification mathématique.

Dans cette famille de discrétisation, nous avons une localisation en espace, mais pas de localisation en fréquence.

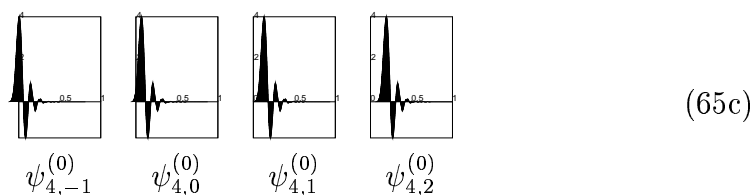
### 3.2.2 Les paquets d'ondelettes

L'espace des fonctions d'échelles admet une base orthonormale invariante par translation, qu'on appellera base canonique. Cette base cohabite avec d'autres bases orthonormales connues sous le nom de paquets d'ondelettes.

Nous supposons que  $\frac{1}{h}$  est une puissance de deux, que nous notons  $2^j$  avec  $j \in \mathbb{N}$  fixé. L'espace de discrétisation  $V_{2^{-j}}$  est défini comme l'espace engendré par les fonctions

$$\{\psi_{j,m}^{(0)}, m \in \mathbb{Z}/L2^j\mathbb{Z}\} \quad (65b)$$

dont nous dessinons quatre éléments :



$V_{2^{-j}}$  est un sous-espace de  $L^2(\partial S)$  qui est, via le paramétrage  $s$  de la définition 2.9, isométrique à  $L^2(\mathbb{R}/L\mathbb{Z})$ , avec  $L$  la longueur du bord du diffractant  $S$  (33). Par conséquent, les fonctions dessinées ci-dessus sont en fait  $L$ -périodiques. Pour plus de commodité, nous supposons que  $L$  est une puissance de deux, quitte à modifier  $S$  et donc  $\lambda_0$ .

Nous supposons que

$$\psi_{j,m}^{(0)}(x) = \psi_{j,0}^{(0)}(x + m2^{-j}), \quad (65d)$$

c'est à dire que  $V_{2^{-j}}$  est engendré par les translatées d'une même fonction  $\psi_{j,0}^{(0)}$ , et que donc les fonctions dessinées ci-dessus sont les translatées l'une de l'autre. Le choix de la fonction  $\psi_{j,0}^{(0)}$  n'est pas quelconque, il doit satisfaire deux propriétés remarquables :

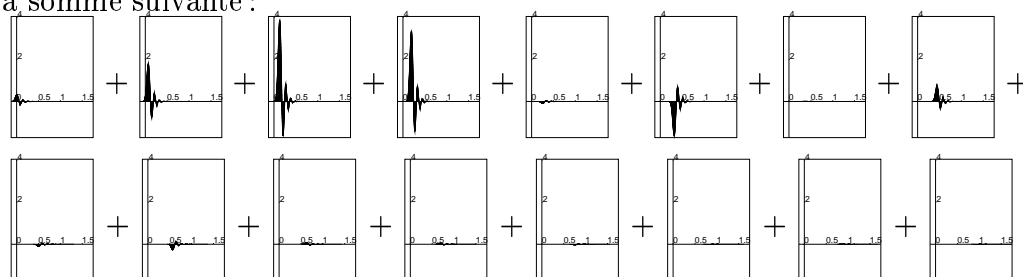
- les fonctions de (65b) forment une famille orthonormale.
- chaque fonction est *auto-similaire*, c'est à dire semblable dans un certain sens à ses dilatées.

Ainsi, les quatre fonctions dessinées ci-dessus sont en fait orthonormales, au sens où

$$\int_0^L \psi_{j,m}^{(0)}(x) \psi_{j,p}^{(0)}(x) dx = 0 \quad (65e)$$

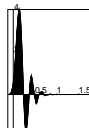
dès que  $m$  et  $p$  sont différents (dans  $\mathbb{Z}/L2^j\mathbb{Z}$ ).

Voyons maintenant l'allure d'un élément de  $V_{2^{-j}}$ . Comme  $V_{2^{-j}}$  est engendré par la famille de fonctions (65b), un élément quelconque de  $V_{2^{-j}}$  est la combinaison linéaire de fonctions en  $\psi_{j,*}^{(0)}$ . Par exemple, un élément de  $V_h$  est la somme suivante :



Les termes de cette somme sont des éléments de la base (65b) de  $V_{2^{-j}}$  multipliés par des coefficients. En les additionnant, nous trouvons une fonction de  $V_{2^{-j}}$ , car  $V_{2^{-j}}$  est un espace vectoriel.

Nous allons maintenant définir ce qu'est l'auto-similarité de  $\psi_{4,0}^{(0)}$ . Pour l'illustration ci-dessus, nous n'avons pas choisi au hasard les coefficients (ni le nombre d'élément de la base utilisée), au contraire nous avons pris un cas très particulier pour expliquer l'*auto-similarité*. Cette auto-similarité consiste à ce que le résultat de cette somme soit



qui est à dilatation près (et à une repériodisation près) la fonction  $\psi_{4,0}^{(0)}$  illustrée en (65c). Précisément, nous avons, avec  $m_0, m_1 \in \mathbb{N}$  et  $g_*^{(0)} \in \mathbb{R}$ , pour

tout  $x \in \mathbb{R}/L\mathbb{Z}$  :

$$\sum_{m=-m_0}^{m=m_1} g_m^{(0)} \psi_{j,m}^{(0)}(x) = \psi_{j,0}^{(0)}\left(\frac{x}{2}\right) + \psi_{j,L2^{j-1}}^{(0)}\left(\frac{x}{2}\right) \quad (65f)$$

qui est l'équation d'échelle d'auto-similarité fractale de la fonction  $\psi_{j,0}^{(0)}$  (qui est pour cette raison appelée une *fonction d'échelle*) servant à définir  $V_{2^{-j}}$ . L'existence d'une telle identité (65f) justifie l'appellation *fonction d'échelle* pour la fonction  $\psi_{j,0}^{(0)}$ . Toutes les fonctions de (65f) sont en fait des translatées ou dilatées de la fonction  $\psi_{j,0}^{(0)}$  : en effet (65f) se reformule en

$$\sum_{m=-m_0}^{m=m_1} g_m^{(0)} \psi_{j,0}^{(0)}(x - 2^{-j}m) = \psi_{j,0}^{(0)}\left(\frac{x}{2}\right) + \psi_{j,0}^{(0)}\left(\frac{x}{2} - \frac{L}{2}\right)$$

qui n'utilise qu'une fonction, la fonction  $\psi_{j,0}^{(0)}$ .

Nous avons prévenu que la base (65b) cohabitait avec d'autres bases orthonormales connues sous le nom de paquets d'ondelettes, et nous allons construire une deuxième base. L'identité (65f) nous incite à définir la fonction

$$\psi_{j-1,m}^{(0)}(x) = [\psi_{j,2m}^{(0)}\left(\frac{x}{2}\right) + \psi_{j,2m+L2^{j-1}}^{(0)}\left(\frac{x}{2}\right)]/\sqrt{2}$$

qui est, pour  $m \in \frac{1}{2}\mathbb{Z}/L2^{j-1}\mathbb{Z}$ , un élément remarquable de  $V_{2^{-j}}$ . Comme  $m$  a en fait  $L2^j$  valeurs distinctes ci-dessus avec  $0 \leq m < L2^{j-1}$ , l'ensemble des fonctions  $\{\psi_{j-1,m}^{(0)} : m \in \frac{1}{2}\mathbb{Z}/L2^{j-1}\mathbb{Z}\}$  est une famille de fonctions de même cardinal que la famille (65b) qui serait un candidat méritant pour être une autre base de  $V_{2^{-j}}$ . Seulement, cette famille n'est pas génératrice (ne génère pas l'élément  $\sum_{m=1}^{L2^j} (-1)^m \psi_{j,m}^{(0)}$  car  $\psi_{j,0}^{(0)} \in L^2(\partial S)$ ) donc cette famille-candidat n'est pas éligible à être une base de  $V_{2^{-j}}$ . En ne gardant de cette famille que les indices entiers (c'est à dire  $m \in \mathbb{Z}/L2^{j-1}\mathbb{Z}$ ), nous obtenons une famille libre, qui engendre un sous-espace de  $V_{2^{-j}}$  noté  $V_{2^{1-j}}$ . Cohen et Daubechies [CD93] montre à partir de (65e) et (65f) qu'un autre élément particulier de  $V_{2^{-j}}$ , à savoir la fonction

$$\psi_{j-1,0}^{(1)}(x) = \left[ \sum_{m=-m_0}^{m=m_1} g_m^{(1)} (-1)^m g_{1-m}^{(0)} \psi_{j,m}^{(0)}(x) \right] / \sqrt{2} = [\psi_{j,0}^{(0)}\left(\frac{x}{2}\right) + \psi_{j,L2^{j-1}}^{(0)}\left(\frac{x}{2}\right)] / \sqrt{2}$$

avec le choix particulier  $g_m^{(1)} = (-1)^m g_{1-m}^{(0)}$ , est orthogonal à ce sous-espace. Cette fonction s'appelle une *ondelette*. De plus, la famille orthonormale formée par cette fonction et ses translatées définies similairement à (65d) par

$$\psi_{j-1,m}^{(1)}(x) = \psi_{j-1,0}^{(1)}(x - 2^{1-j}m)$$

engendrent le complémentaire orthogonal de  $V_{2^{1-j}}$  dans  $V_{2^{-j}}$ . Ainsi, la famille

$$\{\psi_{j-1,m}^{(n)}, m \in \mathbb{Z}/L^{2^j-1}\mathbb{Z}, n \in \{0,1\}\} \quad (65g)$$

avec la convention  $\psi_{j-1,m}^{(0)} = \psi_{j-1,m}^{(0)}$  est une seconde base orthonormale de  $V_{2^{-j}}$ .

Maintenant que nous avons deux bases orthonormales, nous allons regarder la structuration par blocs de la matrice de passage correspondante. La matrice de passage notée  $P_{L^{2^j}}$  entre cette base et la base (65b), carrée de taille  $L^{2^j}$ , admet la structure par blocs suivante

$$P_{L^{2^j}} = \begin{bmatrix} P_{L^{2^j}}^{(0)} \\ P_{L^{2^j}}^{(1)} \end{bmatrix} \quad (65h)$$

où  $P_{L^{2^j}}^{(0)}$  et  $P_{L^{2^j}}^{(1)}$  sont des matrices « rectangulaires circulantes » à  $L^{2^j-1}$  lignes et  $L^{2^j}$  colonnes, de pente non pas 1 mais  $\frac{1}{2}$ . Par exemple,  $\sqrt{2}P_{L^{2^j}}^{(0)}$  est :

$$\begin{array}{c} \psi_{j-1,0}^{(0)} \\ \psi_{j-1,1}^{(0)} \\ \psi_{j-1,2}^{(0)} \\ \vdots \\ \psi_{j-1,t}^{(0)} \\ \vdots \\ \psi_{j-1,L^{2^j-1}}^{(0)} \end{array} \begin{bmatrix} \psi_{j,0}^{(0)} & \psi_{j,1}^{(0)} & \dots & & \dots & \psi_{j,e}^{(0)} & \dots & & \dots & \psi_{j,L^{2^j-1}}^{(0)} \\ g_0^{(0)} & g_1^{(0)} & \dots & g_{m_1}^{(0)} & 0 & \dots & \dots & 0 & g_{-m_0}^{(0)} & g_{1-m_0}^{(0)} & \dots & \dots & g_{-1}^{(0)} \\ g_{-2}^{(0)} & g_{-1}^{(0)} & \dots & g_{m_1-2}^{(0)} & g_{m_1-1}^{(0)} & g_{m_1}^{(0)} & 0 & \dots & \dots & 0 & g_{-m_0}^{(0)} & \dots & g_{-3}^{(0)} \\ g_{-4}^{(0)} & g_{-3}^{(0)} & & & & & & & & & & & g_{-5}^{(0)} \\ \ddots & \ddots & \ddots & & \ddots & \ddots & \ddots & & & & & & \ddots \\ \ddots & \ddots & & \ddots & \ddots & \ddots & \ddots & \ddots & & & & & \ddots \\ \ddots & \ddots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & & & \ddots \\ g_2^{(0)} & \dots & g_{m_1}^{(0)} & 0 & \dots & 0 & g_{-m_0}^{(0)} & g_{1-m_0}^{(0)} & g_{2-m_0}^{(0)} & g_{3-m_0}^{(0)} & \dots & g_0^{(0)} & g_1^{(0)} \end{bmatrix}$$

où nous avons annoté chaque colonne par la fonction correspondant au coefficient, et chaque ligne par la fonction résultat de la combinaison linéaire correspondante. Pour simplifier les notations, prenons la convention que  $g_m^{(0)} = 0$  lorsque  $m \in \mathbb{Z} \setminus [-m_0, m_1]$  : alors, le terme général ligne  $l$  colonne  $c$  de la matrice  $P_{L^{2^j}}^{(0)}$  est

$$\frac{1}{\sqrt{2}} \sum_{m \in \mathbb{Z}} g_{c-2l+L^{2^j}m}^{(0)}$$

où  $c \in \mathbb{Z}/L^{2^j}\mathbb{Z}$  et  $l \in \mathbb{Z}/L^{2^j-1}\mathbb{Z}$ , et où le signe somme ne sert qu'à remplir les coins en bas à gauche et en haut à droite de la matrice  $P_{L^{2^j}}^{(0)}$ . La matrice  $P_{L^{2^j}}^{(1)}$  est définie de façon analogue à  $P_{L^{2^j}}^{(0)}$ , en remplaçant  $g_*^{(0)}$  par  $g_*^{(1)}$ .

Nous allons maintenant construire une troisième base de  $V_{2^{-j}}$ , connue sous le nom de base d'ondelettes. Nous avons fait cette construction à  $j$

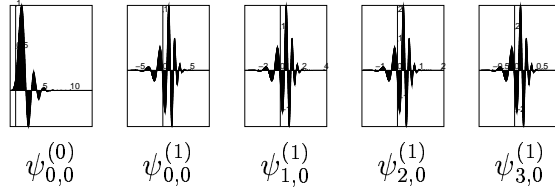
donné tel que  $2^{-j}$  égale le pas de discrétisation  $h$ , et nous avons donné une matrice de passage entre deux bases orthonormales. Cette matrice  $P_{L2^j}$ , qui est orthogonale, peut être définie sur le même principe pour  $j$  ayant d'autres valeurs, ce qui nous permet de concevoir  $P_{L2^{j'}}$ ,  $P_{L2^{j'}}^{(0)}$  et  $P_{L2^{j'}}^{(1)}$  pour  $j'$  entier. Nous allons utiliser des valeurs de  $j'$  inférieures à  $j$  pour construire d'autres bases orthogonales : en notant  $\times$  le produit matriciel, nous voyons que

$$\left[ \begin{array}{c} \left[ \begin{array}{c} \left[ \begin{array}{c} P_{L2^{j-3}}^{(0)} \\ P_{L2^{j-3}}^{(1)} \end{array} \right] \times P_{L2^{j-2}}^{(0)} \\ P_{L2^{j-2}}^{(1)} \end{array} \right] \times P_{L2^{j-1}}^{(0)} \\ P_{L2^{j-1}}^{(1)} \\ P_{L2^j}^{(1)} \end{array} \right] \times P_{L2^j}^{(0)} \quad (65i)$$

est une définition par blocs d'une matrice orthogonale, qui est la matrice de passage entre la base canonique (65b) et une troisième base qui est une généralisation de la base (65g). En réutilisant les mêmes notations (notamment  $\psi_{*,*}^{(*)}$ ), cette troisième base s'écrit (dans l'ordre conféré par (65i)) comme :

$$\{\psi_{j-1,m}^{(0)}, m \in \mathbb{Z}/L2^{j-4}\mathbb{Z}\} \cup \{\psi_{j-1,m}^{(1)}, m \in \mathbb{Z}/L2^{j'}\mathbb{Z}, j' \in \{j-4, j-3, j-2, j-1\}\} \quad (65j)$$

Cette troisième famille, qui est orthogonale, est connue sous le nom de *base d'ondelettes*. Cette troisième famille (65j) se ramène, à translation près et dans l'ordre, aux cinq fonctions suivantes (car  $j - (j - 4) + 1 = 5$ ) que nous représentons dans le cas  $j = 4$  :



Bien sûr, nous pouvons remplacer  $j - 4$  dans (65j) par d'autres entiers par exemple  $j - j'$  ; le nombre  $j'$ , qui dans (65j) valait 4, est le nombre d'échelles représentées, c'est à dire le nombre de valeurs différentes en premier indice de  $\psi_{*,*}^{(1)}$ . Une base d'ondelettes contenant  $j'$  échelle contient  $j' + 1$  groupes de fonctions, chacun des groupes étant constitué de translatées d'une même fonction.

Dans la matrice de passage (65i), nous n'avons autorisé de produit que pour le sous-bloc supérieur de chaque groupe de deux sous-blocs. Pour passer de la base d'ondelette au concept plus général de *base de paquets d'ondelettes*, il suffit de construire une matrice différente de (65i) en s'autorisant des produits sur tous les sous-blocs, au lieu de se restreindre aux produits sur les

$$\left[ \begin{array}{c} \left[ \begin{array}{c} \left[ \begin{array}{c} P_{L2j-3}^{(0)} \\ P_{L2j-3}^{(1)} \end{array} \right] \times P_{L2j-2}^{(0)} \\ \left[ \begin{array}{c} P_{L2j-3}^{(1)} \\ P_{L2j-3}^{(0)} \end{array} \right] \times P_{L2j-2}^{(1)} \end{array} \right] \times P_{L2j-1}^{(0)} \\ \left[ \begin{array}{c} \left[ \begin{array}{c} P_{L2j-3}^{(0)} \\ P_{L2j-3}^{(1)} \end{array} \right] \times P_{L2j-2}^{(1)} \\ \left[ \begin{array}{c} P_{L2j-3}^{(1)} \\ P_{L2j-3}^{(0)} \end{array} \right] \times P_{L2j-2}^{(0)} \end{array} \right] \times P_{L2j-1}^{(1)} \\ \left[ \begin{array}{c} \left[ \begin{array}{c} P_{L2j-3}^{(0)} \\ P_{L2j-3}^{(1)} \end{array} \right] \times P_{L2j-2}^{(0)} \\ \left[ \begin{array}{c} P_{L2j-3}^{(1)} \\ P_{L2j-3}^{(0)} \end{array} \right] \times P_{L2j-2}^{(1)} \end{array} \right] \times P_{L2j-1}^{(0)} \\ \left[ \begin{array}{c} \left[ \begin{array}{c} P_{L2j-3}^{(0)} \\ P_{L2j-3}^{(1)} \end{array} \right] \times P_{L2j-2}^{(1)} \\ \left[ \begin{array}{c} P_{L2j-3}^{(1)} \\ P_{L2j-3}^{(0)} \end{array} \right] \times P_{L2j-2}^{(0)} \end{array} \right] \times P_{L2j-1}^{(1)} \end{array} \right] \times P_{L2j}^{(0)} \\ \left[ \begin{array}{c} \left[ \begin{array}{c} P_{L2j-3}^{(0)} \\ P_{L2j-3}^{(1)} \end{array} \right] \times P_{L2j-2}^{(1)} \\ \left[ \begin{array}{c} P_{L2j-3}^{(1)} \\ P_{L2j-3}^{(0)} \end{array} \right] \times P_{L2j-2}^{(0)} \end{array} \right] \times P_{L2j}^{(1)} \end{array} \right] \times P_{L2j}^{(1)}
\end{array} \right]$$

FIG. 18 – Une matrice de passage équilibrée.

sous-blocs supérieurs. Pour être précis, nous n'allons pas faire de produit des sous-blocs inférieurs avec

$$\begin{bmatrix} P_{L2j'-1}^{(0)} \\ P_{L2j'-1}^{(1)} \end{bmatrix}, \text{ mais plutôt avec } \begin{bmatrix} P_{L2j'-1}^{(1)} \\ P_{L2j'-1}^{(0)} \end{bmatrix} \quad (66)$$

afin que la numérotation des nouvelles fonctions  $\psi_{*,*}^{(2)}$ ,  $\psi_{*,*}^{(3)}$ ,  $\psi_{*,*}^{(4)}$  ... apparaissant dans les paquets d'ondelettes soient mieux en rapport avec la fréquence optimale de leur transformée de Fourier (voir la propriété 5.7 et la remarque 5.8). Oublier de faire le remplacement (66) reviendra, dans la formule de récurrence (89x) définissant les paquets d'ondelettes, à échanger  $4n + 2$  et  $4n + 3$ .

Nous pouvons construire, par exemple, en partant de (65j), la matrice plus équilibrée de la figure 18. C'est, au même titre que la matrice (65i), une généralisation possible de la première matrice de passage  $P_{L2j}$  (65h).

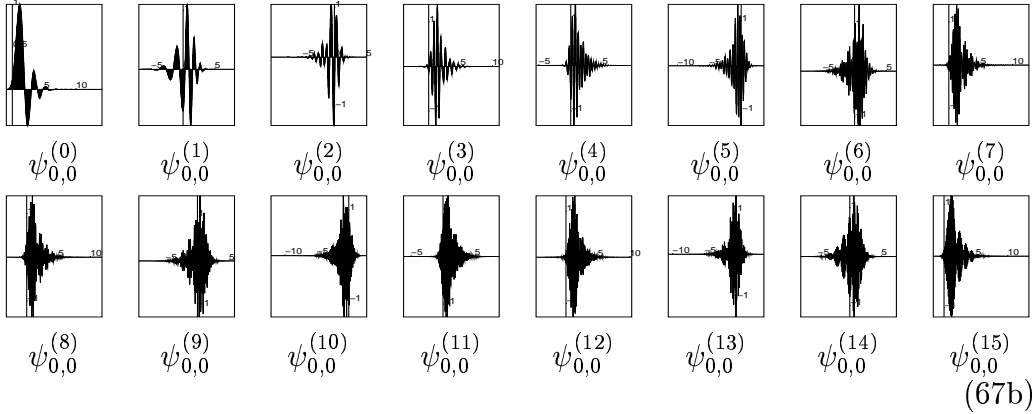
Le remplacement (66) se voit nettement dans cette matrice figure 18, il brise un peu la répétition. La famille orthonormale résultante admet une notation analogue à (65g), respectant l'ordre conféré par la matrice de passage



de la figure 18 :

$$\{\psi_{j-4,m}^{(n)}, m \in \mathbb{Z}/L2^{j-4}\mathbb{Z}, n \in \{0,1,\dots,2^4-1\}\} \quad (67a)$$

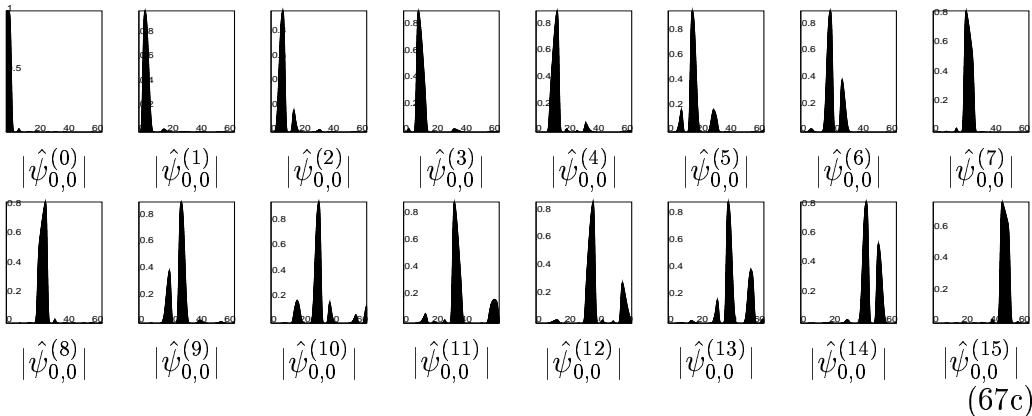
Cette nouvelle notation contient de nouvelles fonctions  $\psi_{*,*}^{(n)}$  pour  $2 \leq n \leq 15$ , qui ne sont pas des dilatations des deux fonctions  $\psi^{(0)}$  ou  $\psi^{(1)}$  rencontrées précédemment. Voici la représentation graphique, à translations près, de toutes les fonctions présentes dans la famille (67a), dans l'ordre :



Le remplacement (66) se justifie facilement à la lumière de ces représentations graphiques: sans lui, les fonctions ci-dessus apparaîtraient dans le désordre par rapport à leur fréquence, vu qu'on obtiendrait, au lieu de la famille (67a), la famille

$$\{\psi_{j-4,m}^{(n)}, m \in \mathbb{Z}/L2^{j-4}\mathbb{Z}, n \in \{0,1,3,2,7,6,4,5,15,14,12,13,8,9,11,10\}\}$$

qui ne serait pas dans l'ordre fréquentiel tel que visible sur les graphes ci-dessus, et aussi ci-dessous sur les modules de leurs transformées de Fourier :



où la convention pour la transformation de Fourier d'une fonction  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  est

$$\hat{\phi}(\xi) = \int_{\mathbb{R}} \phi(t) e^{-i\xi t} dt. \quad (67d)$$

et où la fonction  $\phi$  doit être déperiodisée proprement (de sorte à donner un support plus petit que  $L$ ) si c'est une fonction  $\mathbb{R}/L\mathbb{Z}$ .

Nous constatons dans les représentations ci-dessus que les éléments de cette base de paquet d'ondelettes sont rangés par fréquence croissante. Wickerhauser [W94] démontre même cette constatation, pour deux filtres  $g_*^{(0)}$  particuliers (le filtre de Haar, et le filtre de Shannon). La numérotation des fonction  $\psi_{0,0}^{(n)}$  illustrées ci-dessus tient compte de façon essentielle de ce remplacement (66), sinon il faudrait permuter comme indiqué ci-dessus les indices, en insérant une permutation connue sous le nom de *code de Gray*. Malheureusement, Wickerhauser [W94] est le seul à ne pas avoir oublié de faire ce remplacement (66). Deng-Ling [DL99a][DL99b][DL02] maintient que, sans faire ce remplacement, les fonctions  $\psi_{0,0}^{(n)}$  ont pour fréquence principale  $(n + \frac{1}{2})\pi$ , ce qui est visiblement faux d'après les représentations ci-dessus. Coifman-Meyer-Wickerhauser [CMW92] et Golik [G98][G00] ne font pas non plus ce remplacement, mais se sont bien gardés de donner la fréquence principale de  $\psi_{0,0}^{(n)}$ .

Nous avons donc en tout quatre bases pour l'espace  $V_{2^{-j}}$ . Présentons une caractérisation de toutes les bases de paquets d'ondelettes. Selon le même principe que pour les deux dernières bases, nous pouvons encore modifier par simplification la matrice de passage de la figure 18, pour obtenir, par exemple, cette matrice quelconque :

$$\left[ \begin{array}{c} \left[ \begin{array}{c} \left[ \begin{array}{c} P_{L2^{j-3}}^{(0)} \\ P_{L2^{j-3}}^{(1)} \end{array} \right] \times P_{L2^{j-2}}^{(0)} \\ \left[ \begin{array}{c} P_{L2^{j-3}}^{(1)} \\ P_{L2^{j-3}}^{(0)} \end{array} \right] \times P_{L2^{j-2}}^{(1)} \end{array} \right] \times P_{L2^{j-1}}^{(0)} \times P_{L2^j}^{(0)} \\ P_{L2^{j-1}}^{(1)} \\ \left[ \begin{array}{c} \left[ \begin{array}{c} P_{L2^{j-3}}^{(0)} \\ P_{L2^{j-3}}^{(1)} \end{array} \right] \times P_{L2^{j-3}}^{(1)} \\ P_{L2^{j-3}}^{(0)} \end{array} \right] \times P_{L2^j}^{(1)} \end{array} \right] \end{array} \quad (67e)$$

La base de  $V_{2^{-j}}$  correspondant à cette matrice de passage s'écrit, en respectant l'ordre prescrit par cette matrice de passage :

$$\begin{aligned} & \{\psi_{j-4,m}^{(n)}, m \in \mathbb{Z}/L2^{j-4}\mathbb{Z}, n \in \{0,1,2,3\}\} \cup \{\psi_{j-2,m}^{(1)}, m \in \mathbb{Z}/L2^{j-2}\mathbb{Z}\} \cup \\ & \{\psi_{j-3,m}^{(n)}, m \in \mathbb{Z}/L2^{j-3}\mathbb{Z}, n \in \{4,5\}\} \cup \{\psi_{j-2,m}^{(3)}, m \in \mathbb{Z}/L2^{j-2}\mathbb{Z}\} \end{aligned} \quad (67f)$$

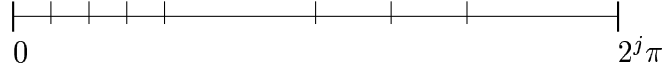


FIG. 19 – La partition du segment  $[0, 2^j \pi)$  correspondant à la partition fréquentielle de la base (67f) et à la matrice de passage (67e).

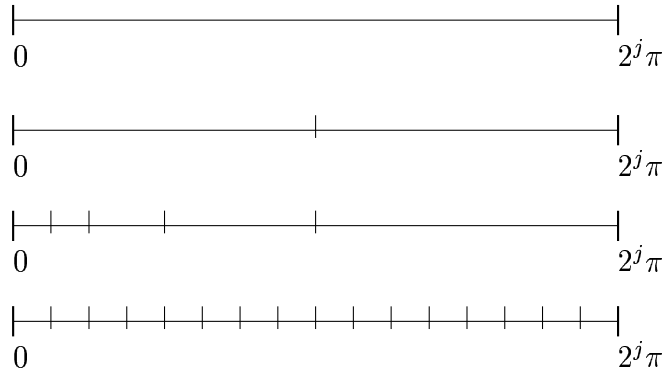


FIG. 20 – Les quatre partitions du segment  $[0, 2^j \pi)$  correspondant respectivement aux bases (65b), (65g), (65j) et (67a).

**Remarque 3.3** Nous remarquons que le segment  $[0, 2^j \pi)$  admet comme partition (illustrée figure 19) les segments demi-ouverts  $[n2^{j'} \pi, (n+1)2^{j'} \pi)$  indexés par  $n$  et  $j'$  tels que  $\psi_{j',*}^{(n)}$  apparaît dans la base ci-dessus, ce qui donne :

$$[0, 2^j \pi) = 2^{j-4}[0, \pi) \cup 2^{j-4}[\pi, 2\pi) \cup 2^{j-4}[2\pi, 3\pi) \cup 2^{j-4}[3\pi, 4\pi) \cup 2^{j-2}[\pi, 2\pi) \cup 2^{j-3}[4\pi, 5\pi) \cup 2^{j-3}[5\pi, 6\pi) \cup 2^{j-2}[3\pi, 4\pi)$$

Cette remarque est vérifiée par les quatre autres bases vues ci-dessus (voir la figure 20, et par tous les paquets d'ondelettes d'après Kahane-Lemarié [KL95]. Nous avons vu ci-dessus que, de plus, Wickerhauser [W94] montre qu'il est représentatif des fréquences inhérentes aux fonctions  $\psi_{j',*}^{(n)}$ .

Pour pouvoir illustrer plus facilement une base de paquets d'ondelettes, nous choisissons la représentation sous une forme particulière d'arbre binaire, qui est un arbre dont les noeuds peuvent avoir soit zéro enfant, soit deux enfants (distincts). Chaque noeud est une notation condensée des deux symboles

$$\left. \right] \quad \text{et} \quad \left. \right] \times P_*^{(*)}$$

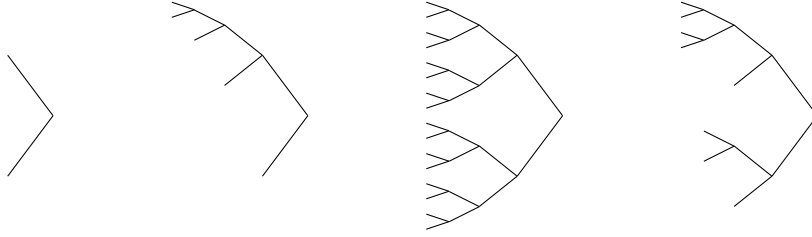


FIG. 21 – Notations compactes par arbres des bases (65g), (65j), (67a) et (67f).

Ainsi, les bases (65g), (65j), (67a) et (67f) sont notées, via leur matrice de passage respective, par des arbres représentés dans la figure 21.

Dans la discrétisation en ondelettes ou en paquets d'ondelettes, les éléments  $\psi_{j,m}^{(n)}$  de la base d'ondelettes ou de la base de paquets d'ondelettes choisie sont localisés en espace grâce à leur indice  $m$  si  $j$  est grand, et aussi localisé en fréquence grâce à leur exposant  $n$  si  $j$  est petit.

### 3.2.3 Les familles de Malvar

Coifman-Meyer [CM91] et Matviyenko [M96] ont proposé un espace de discrétisation formé de fonctions en

$$t \mapsto \cos\left(\left(n\pi + \frac{\pi}{2}\right) \frac{t - t_i}{t_i - t_{i+1}}\right) f\left(\frac{t - t_i}{t_i - t_{i+1}}\right) \quad (67g)$$

où  $t_i$ , pour  $i = 0, 1, \dots, I$  est, par exemple, une suite de progression arithmétique avec  $t_0 = 0$  et  $t_I = L$ , et où le terme

$$t \mapsto f\left(\frac{t - t_i}{t_i - t_{i+1}}\right)$$

est une fonction à support  $[\frac{t_{i-1}+t_i}{2}, \frac{t_{i+1}+t_{i+2}}{2}]$  qu'on appelle une *bosse*, construite avec une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  continue qui vaut plutôt 1 au voisinage de  $\frac{1}{2}$ . Coifman-Meyer [CM91] propose de prendre

$$f(t) = \sin\left(\frac{\pi}{2} \cos^2 \pi \frac{2t - 1}{4}\right) \quad t \in \left[-\frac{1}{2}, \frac{3}{2}\right]$$

Matviyenko [M96] propose de prendre plutôt une fonction trigonométrique paire vérifiant  $f(t) + f(-t) = 1$  et ayant en  $\pm\frac{1}{2}$  le plus possible de dérivées s'annulant. Ainsi, il obtient un espace de discrétisation et un projecteur non-orthogonal associé  $P_h$  tel que  $\|P_h\|_{L^2 \rightarrow L^2}$  et  $\|\text{Id} - P_h\|_{L^2 \rightarrow L^2}$  sont tous deux plus petits que  $\frac{1+\sqrt{2}}{2}$ .

Il existe néanmoins un certain nombre de contraintes sur  $f$ , que nous n'avons pas rappelées. Par exemple, si  $f(1-x) = f(x) \neq 0$  pour  $|x| < \frac{1}{2}$ , la famille de fonctions ainsi formée est une base dite base de Malvar, et est semi-orthogonale, au sens où si deux fonctions n'ont pas la même valeur pour  $i$  dans (67g), alors elles sont orthogonales entre elles. Si de plus on a  $f(x)^2 + f(-x)^2 = 1$  comme proposé par Coifman-Meyer [CM91], alors la base de Malvar correspondante est orthogonale. Averbuch-Braverman-Coifman-Israeli-Sidi [ABC00] précise qu'il vaut mieux choisir la fenêtre  $f$  la plus régulière possible.

L'entier  $n$  ci-dessus, pris entre 0 et  $N-1$ , est une fréquence. L'espace de discrétisation  $V_h$  engendré ces  $N \times I$  fonctions est de dimension  $NI$ . Il n'admet qu'une seule base canonique qui est constitué de ces  $N \times I$  fonctions, ce qui implique par exemple que lorsqu'on change  $I$  et/ou  $N$ , l'espace engendré par cette nouvelle famille de fonctions ne peut pas coïncider avec l'espace  $V_h$ .

Lorsque  $N$  vaut l'infini, l'espace  $V_h$  est une base de Riesz de  $L^2$ .

Dans la discrétisation de Malvar, les éléments de la base de  $V_h$  sont localisés en espace grâce à la fonction  $f$ , entre  $I$  segments différents; ils sont aussi localisés en fréquence, grâce au terme  $\cos((n\pi + \frac{\pi}{2})\frac{t-t_i}{t_i-t_{i+1}})$ , en  $N$  segments différents.

### 3.3 Ce qui a été fait sur cette équation

Dans l'étude de l'équation intégrale de Helmholtz, c'est le traitement de l'opérateur de Helmholtz intégral qui contient le plus de complexité latente. Cet opérateur présente deux difficultés:

- la singularité diagonale qui est un opérateur pseudo-différentiel (faiblement) singulier;
- l'aspect ondulatoire du noyau, qui pose des problèmes de type phase instationnaire.

La singularité diagonale est logarithmique, avec quelques difficultés supplémentaires aux coins du diffractant précisées par la propriété 2.24; c'est un opérateur de  $L^2$  dans  $L^2$ , qui est singulier.

Le noyau, en dehors de la singularité, présente une nature ondulatoire, qui est en

$$e^{i\frac{2\pi}{\lambda_0}r} r^{-1/2} \lambda_0^{-1/2} \quad (67h)$$

d'après les développements limités (32a) et (32b); Cette nature ondulatoire du noyau entraîne un phénomène de compensation locale, qui complique les calculs et les estimations.

### 3.3.1 La discrétisation

Ce processus présente le passage de l'équation analytique à sa mise sous forme d'un système. Il existe deux méthodes :

- la méthode de Nyström présentée par Rokhlin [R83] et Alpert-Beylkin-Coifman-Rokhlin [ABC93]. La solution est calculée comme une somme de masses de Dirac coefficientées, convolées par la solution fondamentale  $G$  (28) de l'équation de Helmholtz.
- La *méthode des moments* est la méthode de Petrov-Galerkin (Nečas [N66]) appliquée à une équation intégrale. Pour ce problème de Helmholtz intégral, elle a été introduite dans [H68] en 1968 (voir [H90]). On prend l'inconnue comme une somme coefficientée de fonctions élémentaires (*expansion functions* ou *base functions*, nous en avons cité des exemples à la section 3.2 ; si c'est des fonctions constantes par morceaux, on parle de *pulse basis*) ; on ne vérifie l'équation qu'en un certain nombre de points (collocation, ou *point matching*), ou alors on vérifie l'égalité contre un certain nombre de poids (les *test functions*). Lorsque ces *test functions* sont égales aux *expansion functions*, on obtient la méthode de Galerkin.

Il existe d'autres méthodes pour équation de Fredholm de première ou seconde espèce dont le noyau est continu (c'est à dire, d'après la propriété 2.24, lorsque le diffractant  $S$  n'a pas de coin), nous pouvons citer la collocation, par exemple. Un très bon bilan de ces méthodes a été fait par Kress [K89] (chapitres 12 à 17), et Wendland [W90].

La taille du système résultant de ces familles *expansion functions* et *test functions* pose des problèmes pratiques, pour la résolution de problèmes où les diffractants atteignant une certaine taille.

D'où l'utilité de l'accélération de ces résolutions, sous la forme de l'accélération du produit matrice-vecteur.

Il existe trois outils rapides de résolutions de cette équation :

- les multipôles,
- les bases de Malvar,
- les bases de paquet d'ondelettes.

### 3.3.2 Les accélérations avec multipôles

Historiquement, la méthode des multipôles a été introduite par Hockney-Eastwood [HE81] pour résoudre le problème à  $N$ -corps. Greengard-Rokhlin [GR87] a ensuite donné la définition moderne des multipôles, avec précision

réglable. Rokhlin [R83] a ensuite employé cette méthode pour l'équation de Helmholtz, dans le cas où la longueur d'onde n'est pas petite par rapport au diffractant. Puis Rokhlin [R85] a expliqué comment elle pouvait succéder aux méthodes de raytracing pour les cas où la longueur d'onde est très grande.

Coifman-Rokhlin-Wandzura [CRW93] remarque que le noyau est un noyau de convolution analytique, et propose cette méthode multipôles qui regroupe temporairement les inconnues par paquets locaux ; dans chaque paquet, le précalcul de la position moyenne et de différents moments permet d'accélérer le calcul du résidu.

Cette astuce locale de factorisation (qui s'écrit bien en Fourier) permet de passer de la complexité  $(\frac{L}{\lambda_0})^2$  à la complexité  $(\frac{L}{\lambda_0})^{3/2}$  ; l'ensemble de ces paquets peut, à son tour, être regroupé par paquet avec la même idée, et ainsi de suite, ce qui permet d'atteindre la complexité  $\frac{L}{\lambda_0} \log \frac{L}{\lambda_0}$  même si la constante est bien grande. Song-Lu-Chew [SLC97] en présente les caractéristiques suivantes :

- Il n'y a pas besoin de régularité sur le diffractant.
- Cette méthode ne marche qu'avec un opérateur à noyau analytique, voir Bradie-Coifman-Grossman [BCG93].
- Les coefficients de la matrice de rigidité sont calculés avec une complexité en  $O(1)$ .
- l'erreur de méthode est difficilement calculable, voir Amini-Profit [AP00].
- Il y a des problèmes de conditionnement lorsqu'il se présente des phénomènes de réverbérations.

Les problèmes de conditionnements ont été résolus plus tardivement, avec les bonnes propriétés de conditionnement de deux autres méthodes, les cosinus locaux de Malvar et les paquets d'ondelettes ; ces deux méthodes sont venues faire de la concurrence à la méthode multipôle.

### 3.3.3 Les accélérations avec la base de Malvar

La base de Malvar a été parachevée par Coifman-Meyer [CM91] et Matvienko [M96]. Elle a l'avantage de prendre un problème avec des sources isotropes, et à les recombinaison de sorte à former des sources ne rayonnant que dans deux (ou quatre) directions très sélectives Canning [C91].

Canning [C90] considère un diffractant sans épaisseur (donc avec l'EFIE à la place de la CFIE) de taille  $100\lambda_0$ . La matrice de rigidité de la méthode des moments est compressée avec une transformée de Fourier à fenêtres non glissantes. Ainsi, seuls les 7% plus gros coefficients de cette matrice suffisent

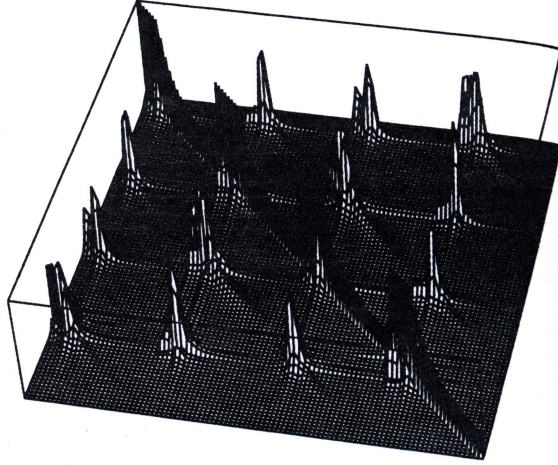


FIG. 22 – *Canning [C95] constate dans cette illustration que toutes les fréquences ne sont pas représentées partout dans la discrétisation d'opérateur de Helmholtz intégral dans une famille de Malvar.*

à résoudre le problème: c'est un gain en mémoire et en complexité, même en tenant compte du préconditionneur. Ainsi, dans un grand problème de taille  $1500\lambda_0$  étudié par Canning [C91], la compression est à 0.26% (c'est indispensable pour rendre possible la résolution d'un si grand problème). Canning [C91] présente aussi un problème à très forte résonance, en utilisant le préconditionneur LU Incomplet (ILU) qu'il a présenté dans [C90], même si toutefois le résidu reste à 1% ce qui d'après Golik [G00] n'est pas suffisant.

Ces bases sont plus précises en Fourier, néanmoins elles ne permettent pas un découpage aussi souple en espace.

Il est toujours aussi coûteux de calculer un terme, tout seul, de la matrice de rigidité.

De manière plus détaillée, les résultats obtenus avec la base de Malvar sont les suivants :

- Grâce aux propriétés fréquentielles des bases de Malvar, il explique que les hautes fréquences (longueur d'ondes plus petites que  $\lambda_0$ ) sont naturellement coupées par cet opérateur (ce sont des modes évanescents).
- Pour les autres fréquences, il montre, constatations numériques à l'appui, qu'elles ne sont pas toutes représentées partout dans l'opérateur dans une illustration reproduite figure 22. Avec un raisonnement en termes de multipôles, il explique que la réflexion de ce qui se passe à



une échelle spatiale de  $\lambda_1 > \lambda_0$  est naturellement reflété du matériau avec un angle de sortie valant

$$\arccos \frac{\lambda_0}{\lambda_1} \quad (67i)$$

et que donc la longueur d'onde  $\lambda_1$  n'est pas représentée partout dans l'opérateur (ce sont des modes de propagation).

- Ainsi, la matrice de rigidité du problème discrétisé peut être seuillée pour en fait ne contenir que  $\frac{L}{\lambda_0} \log \frac{L}{\lambda_0}$  termes.

La singularité diagonale de l'opérateur n'est pas détaillée, il se contente de prendre des fenêtres de Fourier suffisamment fines ( $\sqrt{L/\lambda_0}$  fenêtres de Fourier) et une estimation numérique sur les coefficients apparaissant dans la diagonale.

F.X. Canning suggère de résoudre le système obtenu avec une factorisation Doolittle-LU par blocs. L'analogie de la décomposition temps-fréquence avec l'estimation (67i) permet de déduire que la factorisation n'est pas plus complexe que  $\frac{L}{\lambda_0} \log \frac{L}{\lambda_0}$ .

Bradie-Coifman-Grossman [BCG93] étudie cet opérateur dans sa version intégrale 2D (en faisant abstraction de la singularité diagonale et du facteur  $\lambda_0^{-1/2}$  dans (67h)), ainsi que dans sa version intégrale 3D. Il étudie avec attention l'effet d'un seuillage sur la matrice. En fixant le seuil à  $\varepsilon$ , il montre qu'il ne reste que  $\frac{L}{\lambda_0} \log \frac{L}{\lambda_0}$  termes dans la matrice, mais sans véritablement expliquer où trouver ces termes importants lorsque le diffractant a des coins. [BCG93] s'aide de la nature oscillatoire du noyau en dehors de la diagonale. Toutefois, [BCG93] ne fait pas l'estimation de la norme spectrale de la matrice seuillée lorsque  $\frac{L}{\lambda_0}$  (et donc la longueur d'onde) augmente. Pour un diffractant sans coin, on peut estimer d'après [BCG93] que cette norme spectrale est en  $(\frac{L}{\lambda_0})^{\frac{5j-1}{4j-1}}$  où  $j$  est le nombre d'intégration par parties effectuées. Toutefois, ces résultats sont à relativiser car le facteur  $\lambda_0^{-1/2}$  a été négligé.

Nous devons parler de bases de Malvar au pluriel, car il y a de multiples façons de partitionner l'espace temps-fréquence; cependant, la modification au vol d'une partition est impossible car il faut recalculer les intégrales depuis zéro (l'espace  $V_h$  engendré par une base de Malvar dépend fortement de la base de Malvar). On peut se contenter de projeter l'intégrande sur l'espace des fonctions constantes par morceaux comme Averbuch-Braverman-Coifman-Israeli-Sidi [ABC00], cependant ceci détruit la précision acquise dans ce même article avec le calcul des intégrales avec une méthode de Richardson accélérée.

- Le rôle de la singularité n'est pas explicité séparément.

- L’erreur résultante sur la solution d’un seuil indépendant de  $\lambda_0$  n’est pas étudiée ; on peut estimer qu’elle se dégrade, c’est à dire qu’elle est en  $\lambda_0^{-5/4}$ .

Dans Malvar, on ne peut pas moduler la taille d’une fonction par rapport à la fréquence.

Voir aussi [BBJ94] qui aborde une méthode similaire.

Afin de remédier à cette limitation, une série de travaux s’est penchée sur les méthodes alternatives des ondelettes.

### 3.3.4 Les accélérations avec les ondelettes

[BCR91] puis [DPS93] pensent à utiliser les ondelettes pour un problème avec un opérateur pseudo-différentiel générique (non oscillatoire). Ainsi, [DHS02] montre qu’on peut atteindre moins que  $\frac{L}{\lambda_0}$  en complexité, lorsque la solution ayant  $\frac{L}{\lambda_0}$  degrés de liberté est elle-même compressible en ondelettes. Cet article fait une application sur l’opérateur de Laplace. [ABC93] sort du cadre des ondelettes, et montre qu’en sacrifiant la continuité, mais en gardant le maximum de moments nuls, il obtient une matrice plus creuse, et aussi un coût global plus petit sur la solution. Tous ces articles travaillent sur des noyaux non oscillatoires.

Steinberg et Leviathan travaillent dans [SL93] avec un noyau oscillatoire ; il constate sur la matrice l’effet des extrémités du diffractant, et aussi celui des ondes évanescentes vu ci-dessus (voir aussi Wang-Welland [WW92]). Toutefois, [SL93] n’a pas à prendre en compte la propagation des ondes émises par les longueurs d’ondes plus grandes que  $\lambda_0$ , comme ces ondes ne rencontrent jamais son diffractant (plat) en dehors du domaine d’émission. Il constate un taux de compression sensible de la matrice. [SL93] discrétise en utilisant l’ondelette infinie symétrique orthogonale de Battle-Lemarié, qu’il restreint au diffractant par simple troncation. [CS96] utilise l’ondelette de Meyer sur un diffractant non plat, et montre que les ondelettes facilitent l’utilisation numérique des espaces de Sobolev ( $H^{1/2}$ ,  $H^{-1/2}$ ) dans lesquels l’équation EFIE (48g) est bien posée. Dans le cas où le diffractant a une zone à forte courbure, [SL96] montre qu’il suffit d’avoir des ondelettes de support de la taille du plus petit rayon de courbure pour stabiliser la méthode des moments ; de plus, ces ondelettes n’ont l’air nécessaires qu’aux zones de fortes courbures. Dans tous ces articles, tous les diffractants considérés avec la méthode d’ondelettes ont un périmètre d’au plus  $4\lambda_0$  : il n’y a pas d’étude asymptotique avec  $\lambda_0$  décroissant.

Wagner et Chew font dans [WC95] l’étude asymptotique du cas où  $\lambda_0$  tend vers 0 : ainsi, le diffractant a une taille qui augmente par rapport à la

longueur d'onde. Il observe que la matrice du système a, après le seuillage,  $c(\frac{L}{\lambda_0})^2$  coefficients, pour un diffractant de périmètre  $L$ ; la constante  $c$  donne aux ondelettes l'avantage sur les multipôles (et sur la méthode des moments) lorsque le périmètre du diffractant est inférieur à  $500\lambda_0$ , surtout en l'absence de coins dans le diffractant. Le choix du nombre de moments nuls de l'ondelette de Daubechies a été fixé à 8, car c'est celui pour lequel le remplissage de la matrice des moments et sa transformation en base d'ondelettes prennent le même temps.

Pour améliorer la seuillabilité de la matrice, [NG97] propose par exemple de ne prendre que des fonctions symétriques compactes dans la base d'ondelette retenue, ce qui oblige à remplacer l'orthogonalité par de la bi-orthogonalité ou de la semi-orthogonalité (définie dans [C92b]). Cependant, selon [GYO00], ceci se paye par un nombre de conditionnement qui est considérablement plus mauvais, aussi bien pour la bi- que la semi-orthogonalité. Pour une erreur de la solution en 5% [G00], la bi-orthogonalité est une catastrophe au point de vue du nombre de coefficients seuillables, et la semi-orthogonalité est meilleure que l'orthogonalité. Toutefois, le très mauvais conditionnement conduit [G00] à un plus grand nombre d'itération dans les méthodes itératives, ce qui disqualifie la semi-orthogonalité aussi.

Toutefois, l'approche de compression par ondelettes est citée comme décevante par [G98] et [DL97] pour les problèmes oscillatoires: l'analyse fréquentielle des ondelettes est insuffisante pour capturer de façon suffisamment fine le côté oscillatoire de l'opérateur de Helmholtz.

Canning, dans [C96], utilise les ondelettes pour couper les hautes fréquences, pour ensuite insérer une méthode de Malvar pour les basses fréquences, dans le cas où il y a beaucoup plus que 10 degrés de liberté par longueur d'onde. La partition de ce segment correspond grosso modo à un découpage de ces fréquences par petits paquets. La discrétisation en paquets d'ondelettes facilite le découplage de ces segments de fréquences. Une inégalité indique que plus un segment de fréquence est petit, plus les éléments de la base de paquets d'ondelettes résultante sont de support large.

Pour un diffractant cylindrique à section circulaire (sans coins), Baharav et Leviatan [BL96] partitionnent  $[0, \frac{10\pi}{\lambda_0})$  (voir la remarque 3.3) de façon à minimiser la compressibilité (la norme  $\log^2$ , c.f. [CW92]) du second membre. La base résultante de paquets d'ondelettes (avec Haar) fait que la suppression de 88% des coefficients de la matrice de rigidité n'entraîne qu'une erreur de 1% sur le résidu. Le même pourcentage n'est que de 40% pour un Fourier à fenêtre, et de 65% pour les ondelettes (de Haar aussi).

Ensuite, Golik [G98] travaille avec le filtre de Daubechies à huit moments nuls, et fait le choix du partitionnement de façon à minimiser la somme

des normes  $l^1$  des seconds membres impliqués dans les résolutions demandées. Le découpage obtenu découpe surtout les fréquences entre 0 et  $\frac{2\pi}{\lambda_0}$ . Une étude asymptotique donne un produit matrice-vecteur en  $O((\frac{L}{\lambda_0})^{\frac{4}{3}})$  (norme  $l^2$  du résidu en 1%), et remarque que la présence de coins dans le diffractant ne devraient pas avoir d'incidence sur ce taux de complexité, mais seulement augmenter la constante (tout en restant moins cher que les multipôles pour les diffractants plus petits que  $500\lambda_0$ ). Puis, [DL99a] travaille avec le filtre de Daubechies à 7 moments nuls, et, en optimisant le découpage suivant la somme des modules des coefficients de la matrice de rigidité, obtient  $O((\frac{L}{\lambda_0})^{1.4})$ , avec un seuil proportionnel à la moyenne de la norme  $l^1$  d'une colonne. Le découpage obtenu n'est globalement optimal pour aucun critère [DL99b], par exemple aucun test n'est fait avec les nœuds plus profonds que la solution définitive; C'est le prix à payer lorsqu'on essaye d'optimiser le découpage selon la matrice de rigidité résultante (au lieu de se contenter du second membre comme l'a fait [G98]).

Puis Deng-Ling [DL99a] observe que le découpage se fait essentiellement autour de la fréquence  $\frac{2\pi}{\lambda_0}$  (c'est à dire la longueur d'onde  $\lambda_0$ ); ceci est exploité dans [DL99b] pour définir un découpage fréquentiel qui se raffine le plus autour de cette fréquence, comme c'est la fréquence majeure de l'opérateur: il n'y a plus de coût de recherche du meilleur découpage. Ainsi, cette fréquence de l'opérateur est utilisée avec des éléments de la base qui sont à la bonne fréquence et peu nombreux même s'ils sont de support bien plus longs. [DL99b] arrive aussi à ne plus stocker de matrice pleine: le coût mémoire du calcul de la matrice de rigidité est même estimé à  $O((\frac{L}{\lambda_0})^{1.3})$ , et même  $O((\frac{L}{\lambda_0})^{1.25})$  lorsque le diffractant n'a pas de coin, ce qui lui permet de considérer un problème de taille  $819.2\lambda_0$ ; le coût de calcul reste, lui, en  $O((\frac{L}{\lambda_0})^2 \log^2 \frac{L}{\lambda_0})$  [DL02].

[G00] illustre que pour une même tolérance relative sur la solution, il n'est pas utile d'utiliser les ondelettes au lieu des paquets d'ondelettes, pour des problèmes de taille  $12.8\lambda_0$  à  $102.4\lambda_0$  et même plus. [G00] illustre aussi que la semi-orthogonalité est peu utile, et que la bi-orthogonalité est une catastrophe en tout point pour une même tolérance relative sur la solution. Ces résultats supposent qu'il n'y a pas de bon préconditionneur.

[DL02] observe que l'essentiel de l'opérateur, au sens de la construction d'un préconditionneur, est sur la diagonale, et pas loin de la longueur d'onde  $\lambda_0$  ce qui représente les interactions entre segments parallèles du bord du diffractant (voir diffractants en formes de peignes) et les interactions internes à chaque segment. De plus, ils construisent un préconditionneur (avec illustration numérique) en  $O((\frac{L}{\lambda_0}) \log(\frac{L}{\lambda_0}))$ . La construction prend  $O((\frac{L}{\lambda_0})^{\frac{4}{3}})$ .

Toutefois, ces deux articles font la même erreur, en oubliant la permuta-

tion de Gray (échange des résultats du filtre pour les paquets de fréquence correspondant à  $n$  impair [W94]), ce qui met en cause leur raisonnement fréquentiel. Ainsi, [DL99b] n’obtient-il que  $O((\frac{L}{\lambda_0})^{1.25})$ ; avec deux fois moins de coefficients que [G98], [DL99b] obtient tout de même une erreur sur la solution de 3%.

Nous avons ainsi vu trois stratégies pour trouver ce découpage

1. de façon à maximiser la compressibilité du (des) second(s) membre(s) [BL96] [G98] [G00], avec la stratégie Bottom-Up.
2. de façon à maximiser la compressibilité de la matrice de rigidité, avec la stratégie Top-Down [DL99a].
3. indépendamment du problème, de façon à zoomer autour de la fréquence de la longueur d’onde  $\lambda_0$  [DL99b] [DL02].

Voir le résultat de nos essais dans la partie 5.2.2.

Outre ce découpage, pour configurer les paquets d’ondelettes, il faut aussi choisir une ondelette. Nous avons jusque là cité des ondelettes orthogonale (Haar, Daubechies). Selon Cohen et Daubechies [CD93], la bi-orthogonalité et la semi-orthogonalité sont à éviter car elles rendent instable la méthode des paquets d’ondelettes.

Jusqu’ici, tous les articles font la matrice de rigidité des (paquets d’)ondelettes à partir de la matrice de rigidité de la méthode des moments (Petrov-Galerkin, avec discrétisation en fonction constantes par morceaux, et Diracs comme fonctions tests). Ce choix est par exemple motivé par [SL93] car la valeur des ondelettes est souvent difficile à calculer ([SL93] a préféré travailler avec des ondelettes spline orthogonales, donc avec un support infini à tronquer). Peut-être que cette approximation n’est pas importante comme il s’agit d’une erreur en hautes fréquences; toutefois, il résulte que  $N$  degrés de liberté par longueur d’onde donne une précision qui se dégrade lorsque  $\lambda_0$  tend vers 0, selon le théorème 4.8.

### 3.3.5 Prédiction et calculs dans la matrice de rigidité

[WC95], [DL99b] et [DL02] maintiennent aussi que la position des gros coefficients dans la matrice de rigidité est difficile à prédire, et serait très utile pour passer en-dessous de  $O((\frac{L}{\lambda_0})^2 \log^2(\frac{L}{\lambda_0}))$  pour la construction de cette matrice: or l’étude temps-fréquence présente dans l’article multipôles [R93] a été généralisée au cas des paquets d’ondelettes dans la partie 5 de cette thèse.

[ABC00] propose de calculer les termes de la matrice de rigidité selon la version asymptotique de l’opérateur de Helmholtz intégral; cependant ils

ne parviennent pas à avoir une intégration moins coûteuse que la longueur du domaine d'intégration, comptée en longueur d'ondes. [ABC00] mentionne la méthode de phase stationnaire inventée dans [S56] (en 1856) et justifiée dans [W18]. Toutefois, cette méthode, asymptotique par essence, est citée par [ABC00] comme convergeant trop lentement pour qu'on assimile sa limite à la valeur d'une intégrale oscillatoire.

[C96] et [ABC00] suggèrent de calculer les coefficients de la matrice de rigidité en ne gardant que les produits d'exponentielles qui font moins de  $6\pi$  de variation de phase globale. Mais cela impose de décomposer en exponentielles complexes approximantes les fonctions de la base, et Coifman-Meyer-Wickerhauser [CMW92] explique qu'une telle décomposition d'un paquet d'ondelette n'est possible que si la fonction enveloppe (la fonction bosse) dépend un minimum du numéro de paquet d'ondelette.

[DL99b] a réussi à calculer la matrice de rigidité en moins de  $O((\frac{L}{\lambda_0})^2)$  pour l'occupation mémoire (mais pas en complexité) en n'utilisant pas la transformation rapide de paquets d'ondelettes.

[G98] suggère de calculer la matrice de rigidité initiale (en fonctions d'échelle) sous une forme non-standard compressante (avec, pourquoi pas, la méthode des multipôles), factorisée partiellement via la méthode des moments, et d'ensuite calculer la décomposition en paquets d'ondelettes, par-dessus la décomposition non-standard, ce qui pourrait faire moins que  $O((\frac{L}{\lambda_0})^2)$ .

### 3.3.6 Préconditionnement : mélanges de ces méthodes

Dans sa thèse [D99], E. Darve utilise la discrétisation dans la base de Malvar pour faire un preconditionneur, et combine ce preconditionneur avec la méthode des multipôles : ainsi, il a pu atteindre des nombres de conditionnement tels que le nombre d'itération de la résolution itérative est 5 pour une précision de  $10^{-6}$ . Il remarque aussi que la parité aux jonctions, dans la base de Malvar, est très importante pour atteindre de tels nombres de conditionnements.

L'article [DL02] suggère l'utilisation de la discrétisation avec les paquets d'ondelettes décrits dans [DL99b] pour faire un preconditionneur, et de combiner ce preconditionneur avec la méthode des multipôles : c'est là une solution astucieuse, même si elle souffre de la simplification abusive de la permutation du code de Gray et plus généralement de justifications mathématiques. Le nombre d'itération est plus petit que celui avec les preconditionneurs identité et simplement bloc-diagonaux de la méthode des moments ; le temps consacré au preconditionneur dans toute la résolution est en  $O((\frac{L}{\lambda_0})^{\frac{4}{3}})$ .

[CR03] propose de construire automatiquement une base, qui n'est ni celle des ondelettes, des paquets d'ondelettes, de Malvar ou de la méthode multipôles, en utilisant un processus de diagonalisation d'endomorphisme en bases orthonormées : ainsi, il discrimine très clairement entre les longueurs d'onde évanescentes (plus petites que  $\lambda_0$ ) et les autres. Avec un processus de diagonalisation d'un endomorphisme plus sensible à la direction d'émission que celui considéré dans [CR03], il est peut-être possible de discriminer parfaitement les autres longueurs d'onde selon leur direction d'émission, comme essayent de le faire les bases de Malvar et celles des paquets d'ondelettes ; de plus, les coins et les courbures n'auraient plus d'effet négatif sur le temps de résolution.

Notons que la conclusion de [DL99a] dit du bien d'une compression en paquets d'ondelettes non-standard de la matrice, avec une complexité tombant à  $O(\frac{L}{\lambda_0})$ . Mais Deng et Ling semblent parfois approximatifs dans leurs estimations (voir [DL02] ou [DL99b]).

### 3.3.7 Compatibilité de tout cela avec les coins :

De manière générale, toutes les approches ayant fait l'objet de justifications mathématiques augmentent notablement le nombre de degrés de libertés au voisinage des coins, comme présenté par Kress [K89]. Ceci a pour conséquence notable que l'espace  $V_h$  n'est plus invariant par translation. Nous n'étudierons pas ces espaces de discrétisation  $V_h$  plus adaptés comme proposés par Djaoua [D81], Costabel [C83], Costabel-Stephan [CS85] [CS87] [CS93], Stephan-Suri [SS89], Elschner [E93], Kress-Sloan [KS93], Schneider [S93], Schneider [S03], ou Petersdorff-Schwab [PS97].

Au lieu de cela, nous allons, grande nouveauté, faire des estimations asymptotiques lorsque  $\lambda_0$  tend vers 0, mais dans le cas où le diffractant ne présente pas de coin.

## 4 L'intuition des numériciens sur $\frac{\lambda_0}{h}$ .

L'intuition dit que pour discrétiser un signal oscillant  $J$ , il faut mettre au moins un degré de liberté par *ventre*, c'est à dire au moins deux degrés de liberté par *longueur d'onde* pour toute longueur d'onde susceptible d'être représentée dans la fonction  $J$ .

Cette intuition a été utilisée dans tous les articles faisant varier la longueur d'onde  $\lambda_0$  du problème de Helmholtz (24), qui demandent dix degrés de liberté (au lieu de deux) par longueur d'onde. Averbuch-Braverman-Coifman-Israeli-Sidi [ABC00] montrent qu'ainsi, les formules d'intégration existantes sont très satisfaisantes. Pour l'équation des ondes (non intégrale), Bamberger [B81] explique que la discrétisation en différences finies a une vitesse de propagation qui n'est pas la même que celle de l'équation des ondes lorsqu'il y a moins que 10 degrés de liberté par longueur d'onde.

Cependant, il faut encore montrer que le problème discrétisé avec  $N$  degrés de liberté par longueur d'onde reste valide lorsque  $\lambda_0$  tend vers 0, avec  $N$  indépendant de  $\lambda_0$ .

Kress et Sloan [KS93] font par exemple la démonstration de l'existence de cet entier  $N$ , mais uniquement à  $\lambda_0$  fixé: Ils commencent par isoler la singularité logarithmique de la fonction de Hankel sous la forme de la fonction de noyau

$$(t, t') \mapsto \frac{i\alpha}{2\pi\lambda_0} \log \frac{\sin(\frac{\pi|t-t'|}{L})}{\lambda_0} \quad (67j)$$

telle que visible dans le développement limité (29d). Ensuite, le reste est considéré comme étant un opérateur compact même si oscillatoire: alors en effet, lorsqu'on diminue  $h$  à  $\lambda_0$  constant, on finit par stabiliser les valeurs propres de l'équation discrétisée assez loin de 0, ce qui permet alors de justifier les résultats numériques. Amini [A99] a la même approche.

Kress [K85] a calculé les valeurs propres signalées en (71) lorsque le diffractant est le disque, puis a fait des graphes pour étudier de visu leur comportement et ainsi corroborer l'ajout de  $\lambda_0$  dans (49). Amini [A93] a ensuite fait des graphes similaires pour le cas où le diffractant n'est plus un disque.

Amini-Profit [AP00] a repris ces graphes, et les a adaptés pour corroborer la notion de «  $N$  degrés de liberté par longueur d'onde » pour la méthode multipôles pour un diffractant quelconque.

Afin d'avoir  $N$  indépendant de  $\lambda_0$ , il faudrait étudier cet opérateur compact pour  $\lambda_0$  petit. Mais cet opérateur n'est pas étudiable de cette façon: c'est une fonction oscillatoire de longueur d'onde  $\lambda_0$  qui diminue; la valeur moyenne autour de laquelle se font ces oscillations est une fonction qui tend



à se comporter aussi mal que (67j) lorsque  $\lambda_0$  tend vers 0, ce qui fait que l'opérateur défini par (67j) a une norme  $L^2 \rightarrow L^2$  encore plus mauvaise que  $A$  lorsque  $\lambda_0$  tend vers 0.

Au lieu de (67j), nous avons choisi de considérer, pour séparer la singularité, la fonction de noyau

$$(t, t') \mapsto \frac{1}{\lambda_0} \text{H}_0^1\left(\frac{2L \sin\left(\frac{\pi|t-t'|}{L}\right)}{\lambda_0}\right)$$

ou plus précisément (afin d'intégrer  $\alpha$ ) le noyau  $K$  lui-même, mais dans le cas du diffractant circulaire ( $S$  n'est plus (33) mais c'est un disque de circonférence  $L$ ), auquel cas il a l'expression

$$\hat{K}(t, t') = \frac{\alpha}{4\lambda_0} \text{H}_0^{\text{ank}}\left(\frac{2L \sin\left(\frac{\pi|t-t'|}{L}\right)}{\lambda_0}\right) - \frac{i\pi}{\lambda_0} \sin\left(\frac{\pi|t-t'|}{L}\right) \text{H}_0^{\text{ank}'}\left(\frac{2L \sin\left(\frac{\pi|t-t'|}{L}\right)}{\lambda_0}\right) \quad (68)$$

où  $\hat{K}$  est la notation sous laquelle il sera introduit plus bas.

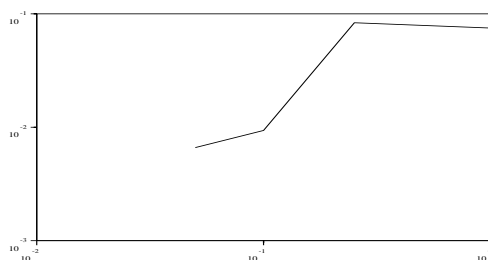
Tout comme Amini [A99] l'a fait pour le cas de l'équation de Laplace, nous étudions la perturbation  $K - \hat{K}$  venant du fait que  $S$  n'est pas circulaire. Cette perturbation se comporte bien mieux pour  $t'$  voisin de  $t$ ; et pour les autres  $t'$ , il oscille autour d'une valeur moyenne nulle, de sorte que sa norme  $L^2$  n'est qu'en  $\lambda_0^{-3/2}$ , soit pas plus mauvais que l'opérateur  $A$  lui-même.

## 4.1 Inversibilité de l'opérateur.

D'après le théorème 2.29, l'opérateur  $A$  (55) est la somme de l'identité, d'un opérateur contractant, et d'un opérateur compact: ainsi,  $\|A^{-1}\|^{-1}$  est asymptotiquement un  $O(1)$  lorsque  $\lambda_0$  tend vers 0.

Les expérimentations numériques ci-dessous montrent que  $\|A^{-1}\|$  n'est hélas pas en  $O(1)$ , même si  $S$  n'a *pas* de coin.

Voici un graphique de  $\|A^{-1}\|^{-1}$  en fonction de  $\lambda_0$ , lorsque  $S$  est le diffractant (34), et lorsque le paramètre  $\alpha$  de  $A$  visible en (55) vaut 4 :



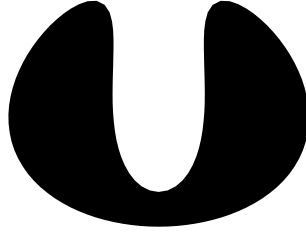


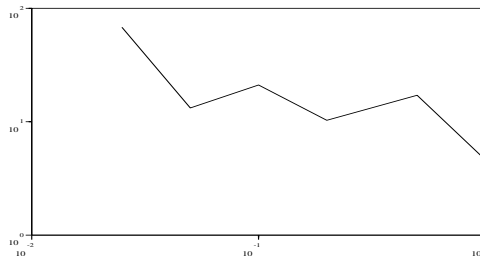
FIG. 23 – *Un choix sans coin pour le diffractant  $S$ .*

Il apparaît donc que lorsque  $S$  présente des coins, la plus petite valeur singulière de  $A$  peut être très petite. Pour le calculer, nous observons que l'inégalité  $\|P_h A u_h\| \geq (\|(A^*)^{-1}\|^{-1} - c_h) \|u_h\|$  page 118 peut être retrouvée avec les hypothèses (60) et (61), ce qui ramène le calcul au calcul de  $\|(P_h A|_{V_h})^{-1}\|$ . Il s'agit donc de calculer la plus petite valeur singulière d'une matrice finie, ce que nous faisons avec ARPACK++, qui est lui-même une interface en C++ au module ARPACK96 en Fortran (<http://www.caam.rice.edu/software/ARPACK/>).

Nous voyons que  $\|A^{-1}\|$  n'est pas borné non plus, lorsque le diffractant n'est plus (33) mais (34) (sans coins) qui correspond à la figure 23 qui a pour équation paramétrique

$$t' \mapsto c \begin{bmatrix} 6 \sin(t') + 3 \sin(2t') + 2 \sin(3t') \\ 6 - \cos(t') - 6 \cos(2t') \end{bmatrix} \quad (69a)$$

où  $t'$  n'est pas l'abscisse curviligne, et où  $c$  est tel que le périmètre  $L$  de  $S$  vaille 16. Voici un graphique de  $\|A^{-1}\|^{-1}$  en fonction de  $\lambda_0$ , lorsque  $S$  est le diffractant (34), et lorsque le paramètre  $\alpha$  de  $A$  visible en (55) vaut 4 :



Dans le cas où le diffractant  $S$  est une boule, le calcul des valeurs propres de  $A$  est possible, et nous constatons que  $\|A^{-1}\|$  est borné indépendamment de  $\lambda_0$  :

**Propriété 4.1** Lorsque le diffractant  $S$  est un disque, la norme de l'inverse de l'opérateur  $A$  est en  $O(1)$  indépendamment du rayon  $R$  et de la longueur d'onde  $\lambda_0$  tels que  $\lambda_0 < R$ . De plus, cet opérateur  $A$  est diagonal dans la base orthogonale de Fourier  $(e_\omega)_{\omega \in \mathbb{Z}}$  de  $L^2(\partial S)$  ainsi définie :

$$e_\omega(t) = e^{i\frac{2\pi}{L}\omega t} \quad (70)$$

■ Dans le cas de la boule, nous avons déjà exhibé, dans la démonstration de la remarque 2.22, pour  $m \in \mathbb{Z}$ , que le  $m$ -ième élément  $e_m$  de la base était un vecteur propre de  $J \mapsto \nu_0(J \star G)$  et de  $J \mapsto \nu_1(J \star G)$  de valeur propre valant respectivement

$$\frac{\pi R}{2i} \mathrm{H}_m^{\mathrm{ank}}\left(\frac{2\pi R}{\lambda_0}\right) \mathcal{R}e \mathrm{H}_m^{\mathrm{ank}}\left(\frac{2\pi R}{\lambda_0}\right) \text{ et } \frac{\pi^2 R}{i\lambda_0} \mathrm{H}_m^{\mathrm{ank}}\left(\frac{2\pi R}{\lambda_0}\right) \mathcal{R}e \mathrm{H}_m^{\mathrm{ank}'}\left(\frac{2\pi R}{\lambda_0}\right),$$

donc (70) est vecteur propre de  $A$  défini par (55) de valeur propre associée

$$\frac{2\pi R}{\lambda_0} \mathrm{H}_{|m|}^{\mathrm{ank}}\left(\frac{2\pi R}{\lambda_0}\right) \left( \frac{\alpha}{4} \mathcal{R}e \mathrm{H}_{|m|}^{\mathrm{ank}}\left(\frac{2\pi R}{\lambda_0}\right) + i\pi \mathcal{R}e \mathrm{H}_{|m|}^{\mathrm{ank}'}\left(\frac{2\pi R}{\lambda_0}\right) \right) \quad (71)$$

dont l'inverse est de module plus petit que  $\max(1, \frac{2\pi}{|\alpha|})$  dès que  $\frac{2\pi R}{\lambda_0} > \frac{1}{\alpha}$ , selon les graphiques des figures 26, 24 et 28 faits avec le logiciel de calcul Maple. Notez que ce n'est pas une démonstration. ■

Cette propriété est peut-être généralisable à tous les diffractant convexes. Toutefois, toutes les estimations démontrées dans la suite s'appliquent au cas d'un diffractant non convexe, donc nous ne supposons plus que  $\|A^{-1}\|$  est borné indépendamment de  $\lambda_0$  par une constante. D'après la condition (60) du théorème 3.1, un espace de discrétisation  $V_h$  devient légitime dès que son projecteur orthogonal  $P_h$  vérifie les estimations

$$\|(\mathrm{Id} - P_h)AP_h\| \leq \frac{1}{2}\|A^{-1}\| \text{ et } \|P_hA(\mathrm{Id} - P_h)\| \leq \frac{1}{2}\|A^{-1}\| \quad (72)$$

qui proviennent de (60) et (61).

## 4.2 L'inégalité directe est inefficace

D'après les théorèmes 2.27 et 2.28, l'opérateur  $J \mapsto \frac{i\alpha}{\lambda_0} \nu_0(J \star G) - 2\nu_1(J \star G) - J$  a une norme d'opérateur  $L^2(\partial S) \rightarrow H^1(\partial S)$  en  $O(\lambda_0^{-3/2} |\log \lambda_0|)$ ,

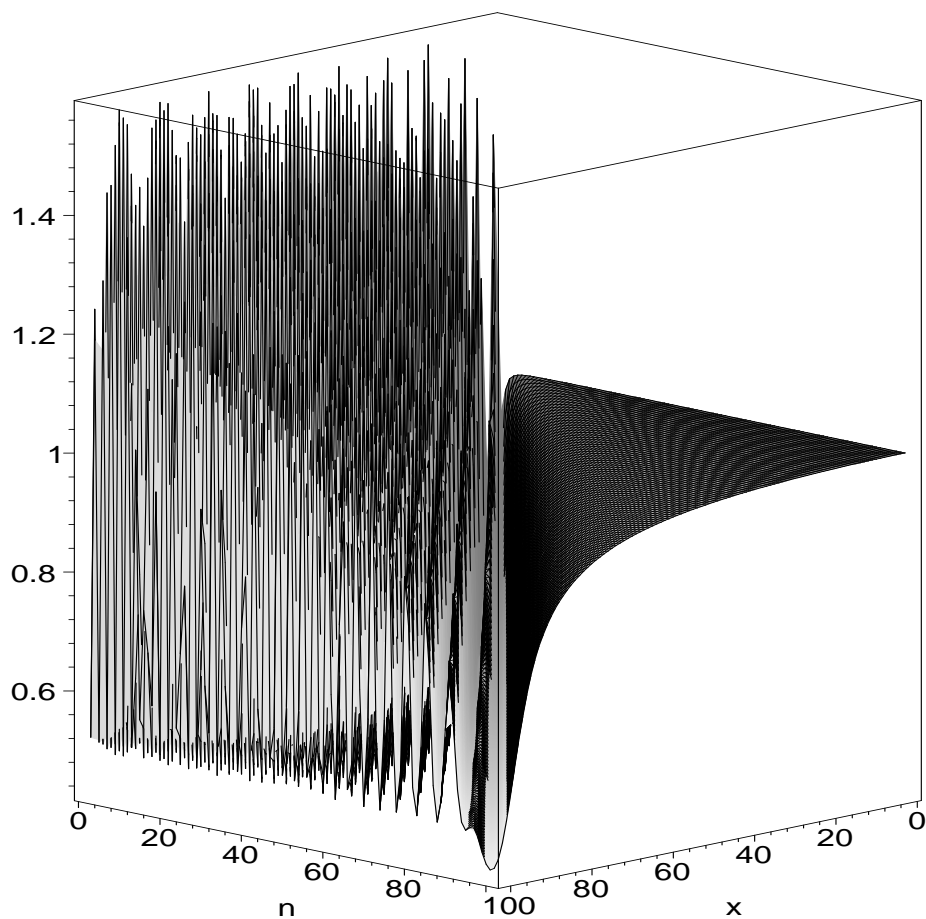


FIG. 24 – Graphe Maple permettant de supposer que la plus petite valeur singulière de  $A$  n'a pas 0 comme point d'accumulation lorsque  $\lambda_0$  tend vers 0 : cas  $\alpha = -4$  et  $S = \text{disque}$ . On a  $n = m$ ,  $x = \frac{2\pi R}{\lambda_0}$ , et en ordonnée on a l'inverse du module de (71).

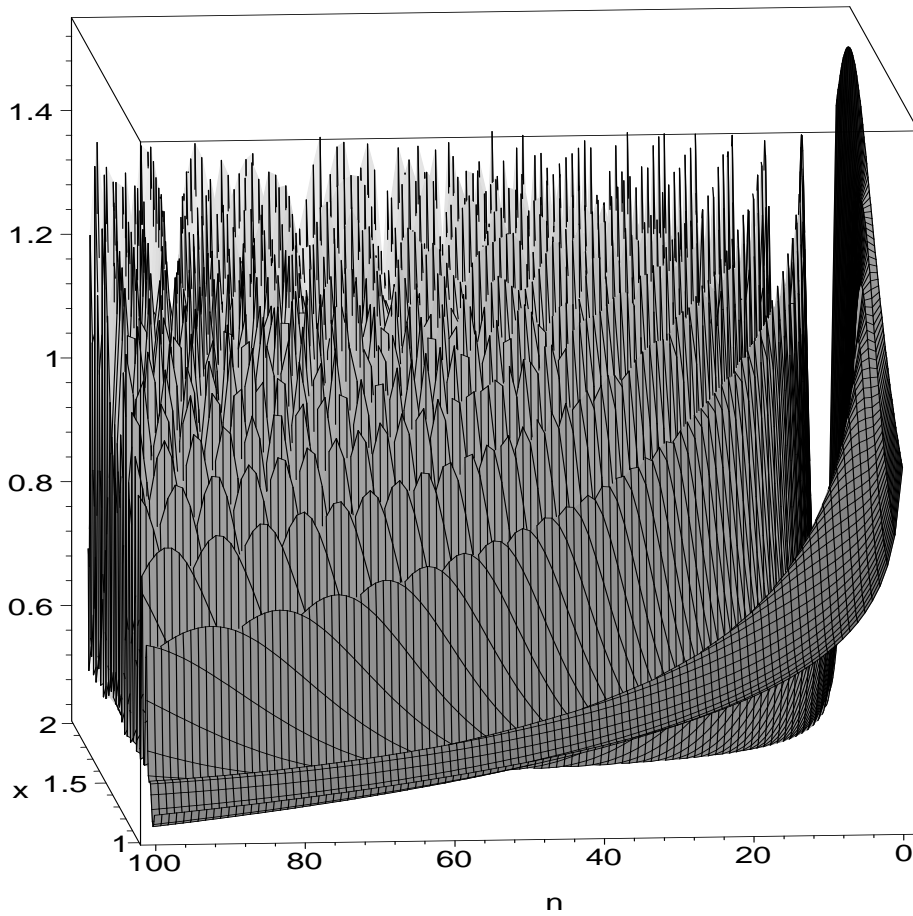


FIG. 25 – Graphe Maple agrandissant la zone  $m \leq \frac{2\pi R}{\lambda_0} \leq 2m$  de la figure 24 du cas  $\alpha = -4$ . On a  $n = m$ ,  $x = \frac{1}{m} \frac{2\pi R}{\lambda_0}$ , et en ordonnée on a l'inverse du module de (71).

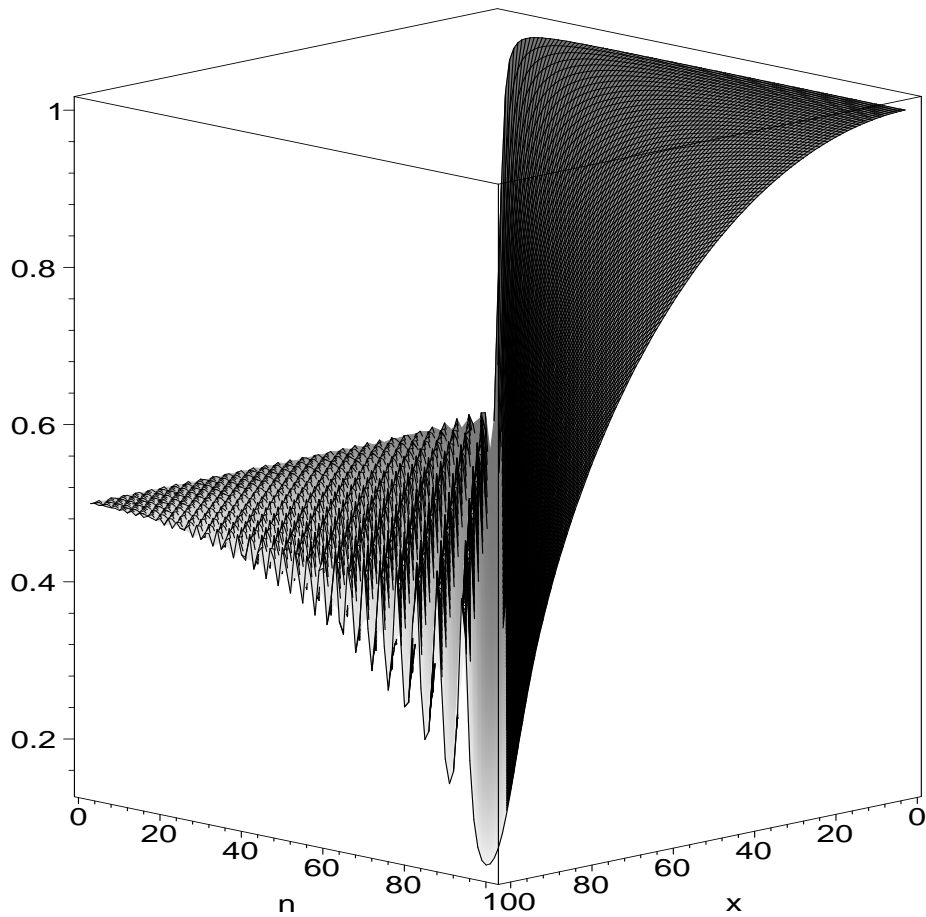


FIG. 26 – Graphe Maple permettant de supposer que la plus petite valeur singulière de  $A$  n'a pas 0 comme point d'accumulation lorsque  $\lambda_0$  tend vers 0 : cas  $\alpha = 4$  et  $S = \text{disque}$ . On a  $n = m$ ,  $x = \frac{2\pi R}{\lambda_0}$ , et en ordonnée on a l'inverse du module de (71).

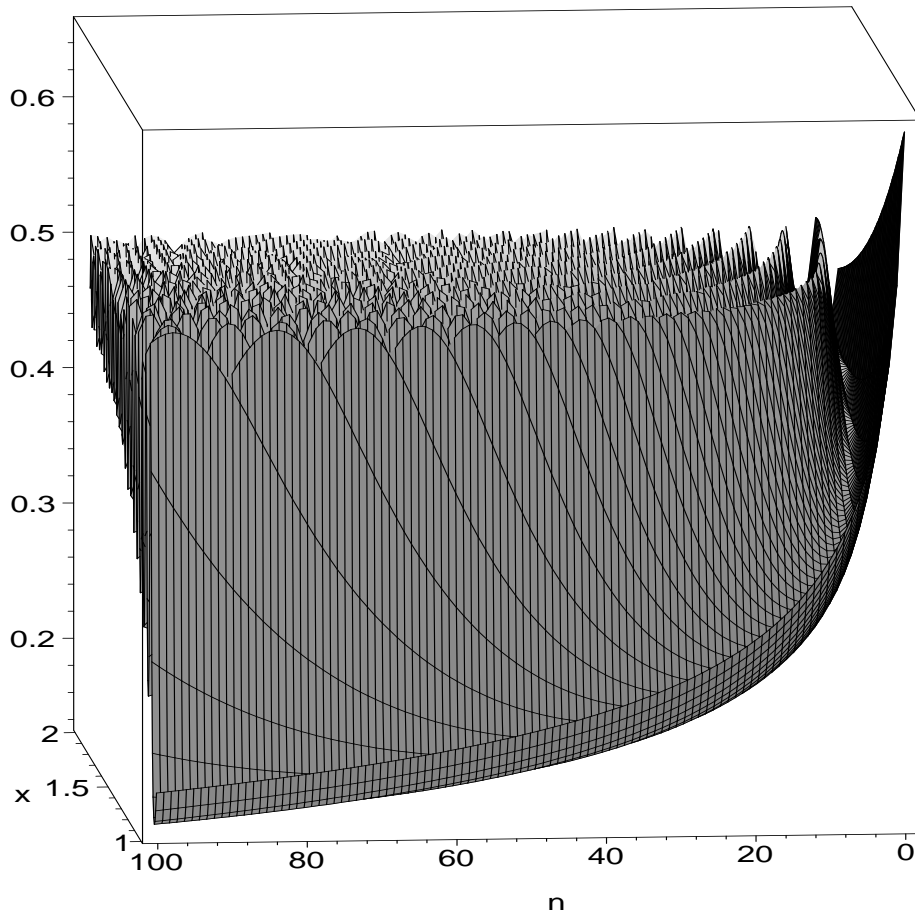


FIG. 27 – Graphe Maple agrandissant la zone  $m \leq \frac{2\pi R}{\lambda_0} \leq 2m$  de la figure 26 du cas  $\alpha = 4$ . On a  $n = m$ ,  $x = \frac{1}{m} \frac{2\pi R}{\lambda_0}$ , et en ordonnée on a l'inverse du module de (71).

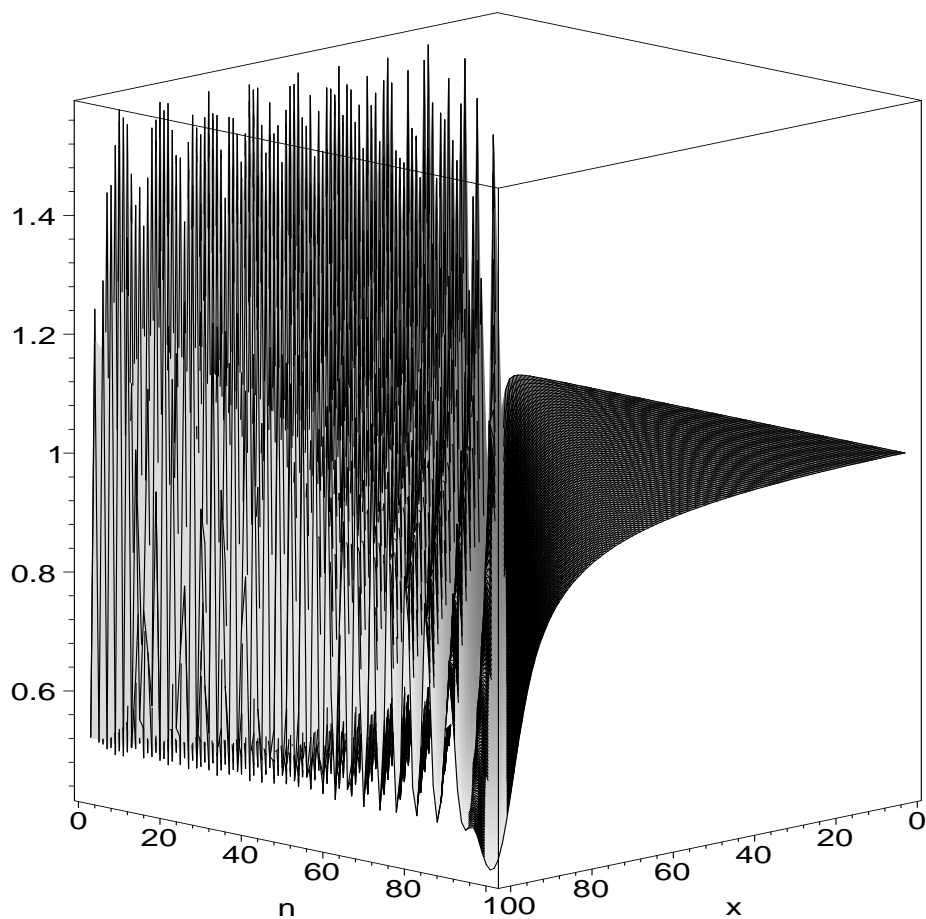


FIG. 28 – Graphe Maple permettant de supposer que la plus petite valeur singulière de  $A$  n'a pas 0 comme point d'accumulation lorsque  $\lambda_0$  tend vers 0 : cas  $\alpha = -4\pi$  et  $S = \text{disque}$ . On a  $n = m$ ,  $x = \frac{2\pi R}{\lambda_0}$ , et en ordonnée on a l'inverse du module de (71).



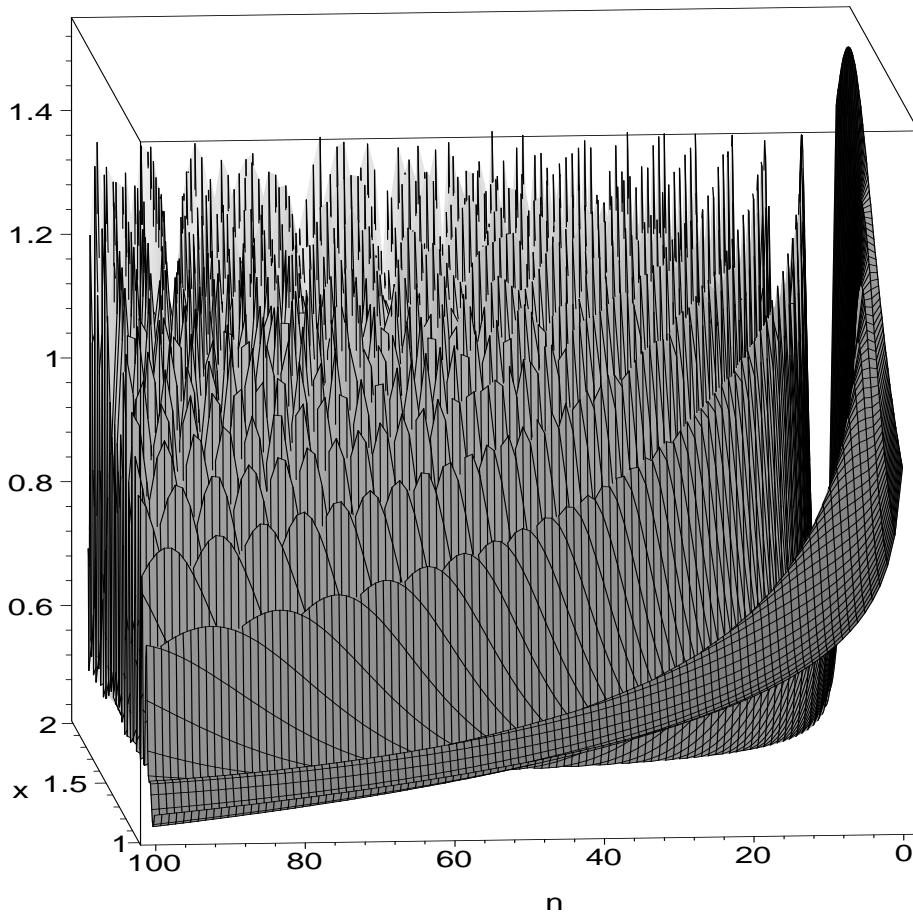


FIG. 29 – Graphe Maple agrandissant la zone  $m \leq \frac{2\pi R}{\lambda_0} \leq 2m$  de la figure 28 du cas  $\alpha = -4\pi$ . On a  $n = m$ ,  $x = \frac{1}{m} \frac{2\pi R}{\lambda_0}$ , et en ordonnée on a l'inverse du module de (71).

pour pouvoir déduire l'estimation (72) à partir d'une inégalité directe du type (65a) sur  $V_h$ , il faut que  $h$  soit en  $\lambda_0^{3/2}$ , c'est à dire que  $h$  doit décroître beaucoup plus vite que  $\lambda_0$  : il est impossible de se tenir à dix degrés de liberté par longueur d'onde.

Il n'existe pas d'inégalité directe du type (65a) portant sur  $H^1({}^cS)$  avec un facteur plus petit que  $h$  lorsqu'on utilise l'espace  $H^1({}^cS)$ .

Comme expliqué dans l'introduction (22) et justifié par la propriété 2.23, L'opérateur  $A$ , défini en (55), est un opérateur intégral de Fredholm de seconde espèce de la forme

$$Af(t) = f(t) + \int_0^L K(t,t')f(t')dt'$$

où  $K$  a été défini dans (23), et reformulé en (51) en remplaçant  $G$  par sa définition (28) avec la fonction de Hankel  $H_0^{ank}$ .

La norme d'opérateur  $L^2(\partial S) \rightarrow H^1(\partial S)$  de  $A$  n'est intuitivement pas plus petite que  $O(\lambda_0^{-3/2})$ , car son noyau  $K$  est un équivalent en  $\lambda_0^{-3/2}$ , à cause des équivalents (32b) et (53) sur les dérivées de la fonction de Hankel  $H_0^{ank}$ .

Le remplacement de  $H^1(\partial S)$  par des espaces de Hilbert  $H^m(\partial S)$  avec un exposant  $m > 1$  n'est pas possible à cause de la singularité logarithmique de  $H_0^{ank}$  visible dans (29d).

Le remplacement de  $H^1(\partial S)$  par des espaces de Besov (comme  $B_{1/2}^{1/2,1}$ ) issus de l'approximation non-linéaire est ici inefficace car l'oscillation de  $H_0^{ank}$  qui pose problème est présente dans un grand domaine de définition du noyau de  $A$ .

Ainsi, l'utilisation d'inégalités directes du type (65a) ne permet pas de corroborer l'intuition des numériciens.

La solution pour justifier l'usage de dix degrés de liberté par longueur d'onde est d'abandonner l'usage d'un espace tel que  $H^1(\partial S)$ , qui tient compte de toutes les fréquences, y incluant celle  $\frac{2\pi}{\lambda_0}$  de longueur d'onde  $\lambda_0$  caractéristique de l'opérateur  $A$ . Au lieu de cela, nous n'allons pas mesurer toutes les fréquences, mais seulement les fréquences de longueur d'onde bien plus petites que  $\lambda_0$ , en vérifiant directement (60) et (61) sur le projecteur  $P_h$  sur  $V_h$ .

### 4.3 L'opérateur focalise les hautes fréquences en l'absence de coins.

Nous allons étudier la plus petitesse de

$$\int_0^L \int_0^L K(t,t') b_1(t) e^{2i\pi \frac{k_1}{\lambda_0} t} b_2(t') e^{2i\pi \frac{k_2}{\lambda_0} t'} dt dt' \quad (73)$$

avec  $|k_1|$  et  $|k_2|$  sensiblement plus grands que 1 et multiples de  $\frac{\lambda_0}{L}$ , et  $b_1$  et  $b_2$  deux fonctions  $C^p$  indépendantes de  $k_1$ ,  $k_2$  et  $\lambda_0$ . Les démonstrations ne prendront pas en compte que  $|k_1|$  et  $|k_2|$  sont sensiblement plus grands que 1, ce qui permettra de démontrer, dans la partie 5, que la matrice de Galerkin est creuse (pour les paquets d'ondelettes comme pour Malvar). Par contre, nous nous limitons au cas  $\lambda_0 < \frac{1}{2}$ , par commodité, principalement pour être sûr que  $|\log \lambda_0|$  ne peut pas être petit, ce qui simplifiera les expressions. Cette restriction de la longueur d'onde  $\lambda_0$  n'est pas importante, car le principal objectif de cette partie 4 est d'obtenir (72) pour  $\lambda_0$  petit.

Pour montrer que (73) est petit, nous allons utiliser des intégrations par parties dans le principe de la phase stationnaire, car la fonction de Hankel  $H_0^{ank}$  et ses dérivées sont oscillantes. Posons

$$u(t,t') = 2\pi \|s(t) - s(t')\| \quad (74)$$

qui est une fonction indépendante de  $\lambda_0$ ; cette fonction commande ce qui dépend de  $\lambda_0$  dans le noyau (23). Plus précisément, c'est le gradient de cette fonction qui commande la vitesse des oscillations; ces oscillations ont moralement une longueur d'onde toujours supérieure à  $\lambda_0$ , en vertu des développements limités (32a) et (32b), et de la propriété suivante :

**Propriété 4.2** *Les dérivées selon  $t$  et  $t'$  (définies presque partout) de la fonction  $u$  sont bornées presque sûrement par  $2\pi$ , même si  $s$  a des coins :*

$$\left| \frac{\partial u}{\partial t}(t,t') \right| \leq 2\pi \quad \text{et} \quad \left| \frac{\partial u}{\partial t'}(t,t') \right| \leq 2\pi$$

■ Par symétrie seule la première inégalité est à montrer. La fonction  $s$  qui décrit le contour de  $S$  (définition 2.9) est dérivable par morceaux, et est continue : donc  $u$  est Lipschitzien ; trouvons une borne presque sûre sur  $\frac{\partial u}{\partial t}$ . Lorsque  $s$  est dérivable en  $t$ , on a :

$$\frac{\partial u}{\partial t}(t,t') = 2\pi \frac{\langle s'(t) | s(t) - s(t') \rangle}{\|s(t) - s(t')\|} \quad (75a)$$

et la preuve est terminée comme  $s$  est paramétré par abscisse curviligne :  $\|s'(t)\| = 1$ . ■

Ce qu'il faut repérer dans (73) et (51) est que  $\nabla u$  a une certaine image, et que lorsque le vecteur  $2\pi \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$  est loin de cette image, alors une intégration par parties se passe bien et nous pouvons utiliser la méthode de phase stationnaire. Une première manière de s'assurer de ceci est que  $2\pi|k_1 + k_2|$ , que nous noterons  $|k_+|$ , soit plus grand que  $6\pi$ , ce que fait la propriété 5.5 ci-dessous. Une seconde manière est de s'en assurer que  $2\pi|k_1 - k_2|$ , que nous noterons  $|k_-|$ , soit plus grand que  $6\pi$ , ce que fera la propriété 4.4.

Notons que  $u = O(|t - t'|)$ , ce qui fait que  $O(u)$  est une abréviation pratique pour  $O(|t - t'|)$ , abréviation que nous utiliserons sans modération.

**Propriété 4.3** *En l'absence de coin dans le diffractant, et si  $s$  est  $C^{p+3}$ , l'intégrale (73) est en*

$$O\left(|\log \lambda_0| \lambda_0^{p-\frac{1}{2}} (\|b_1\|_{\mathbb{H}^p} + \|b_2\|_{\mathbb{H}^p}) (1 + |k_+|)^p d_+^{-p}\right)$$

où

$$d_+ = \min_{1 \leq q \leq 2} \min_{t, t'} |k_+^2 - \left(\frac{\partial u}{\partial t} + \frac{\partial u}{\partial t'}\right)^2|^q$$

est en quelque sorte le produit de  $\lambda_0$  par la distance (ou son carré) entre le vecteur  $2\pi \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$ , et entre la diagonale de la matrice de Galerkin dans la base de Fourier, épaissie de l'amplitude de variation de  $\nabla u$  ; rappelons que  $k_+ = 2\pi(k_1 + k_2)$ .

■ Les fonctions  $b_1$  et  $b_2$  sont des fonctions  $L$ -périodiques : c'est difficilement manipulable. Supposons donc que  $b_1^2 + b_2^2$  est nul sur au moins un intervalle du diffractant, quitte à découper l'intégrale (73) en quatre morceaux. Soit  $z$  un point dans l'intérieur de cet intervalle. Dépériodisons les fonctions  $b_1$  et  $b_2$ , en rendant nul  $b_1(t)$  et  $b_2(t)$  lorsque  $t \notin [z, z + L]$ . Soit  $I_1 \subset \mathbb{R}$  et  $I_2 \subset \mathbb{R}$  les supports de  $b_1$  et  $b_2$ . Cette dépériodisation permet de signer sainement  $u$  : Soit

$$\bar{u}(t, t') := \frac{t' - t}{|t' - t|} u(t, t') = 2\pi \frac{t' - t}{|t' - t|} \|s(t) - s(t')\|.$$

**Lemme 4.3.1** *Comme  $s$  est  $C^3$ , ce signage est sain :  $\bar{u}$  est  $C^3$  (sur  $I_1 \times I_2$ ), et  $(t, t') \mapsto \frac{\bar{u}(t, t')}{t' - t}$  et  $(t, t') \mapsto \frac{t' - t}{\bar{u}(t, t')}$  sont  $C^2$  ;  $\frac{\partial^q \bar{u}}{\bar{u}^3}$  est continu pour  $0 < q \leq p$ , comme  $s$  est  $C^{p+3}$  ; on a aussi la continuité de  $\frac{\partial^{p+1} \bar{u}}{\bar{u}^2}$  (avec une puissance différente au dénominateur).*

■ Dans la démonstration de ce lemme, on fait abstraction du facteur  $2\pi$  dans la définition de  $u$  et de  $\bar{u}$ . La fonction  $u$  admet les dérivées suivantes :

$$\begin{aligned}\frac{\partial}{\partial t}u(t,t') &= \frac{\langle s'(t)|s(t) - s(t')\rangle}{\|s(t) - s(t')\|} \\ \frac{\partial^2}{\partial t^2}u(t,t') &= \frac{\langle s''(t)|s(t) - s(t')\rangle}{\|s(t) - s(t')\|} + \frac{\langle s'(t)^\perp|s(t) - s(t')\rangle^2}{\|s(t) - s(t')\|^3} \\ \frac{\partial}{\partial t} \frac{\partial}{\partial t'}u(t,t') &= -\frac{\langle s'(t)^\perp|s(t) - s(t')\rangle \langle s'(t')^\perp|s(t) - s(t')\rangle}{\|s(t) - s(t')\|^3} \\ \frac{\partial^3}{\partial t^3}u(t,t') &= \frac{\langle s'''(t)|s(t) - s(t')\rangle}{\|s(t) - s(t')\|} + 3\frac{\langle s''(t)^\perp|s(t) - s(t')\rangle \langle s'(t)^\perp|s(t) - s(t')\rangle}{\|s(t) - s(t')\|^3} \\ &\quad - 3\frac{\langle s'(t)|s(t) - s(t')\rangle \langle s'(t)^\perp|s(t) - s(t')\rangle^2}{\|s(t) - s(t')\|^5} \\ \frac{\partial^2}{\partial t^2} \frac{\partial}{\partial t'}u(t,t') &= -\frac{\langle s''(t)^\perp|s(t) - s(t')\rangle \langle s'(t')^\perp|s(t) - s(t')\rangle}{\|s(t) - s(t')\|^3} + \frac{\langle s'(t')|s(t) - s(t')\rangle \langle s'(t)^\perp|s(t) - s(t')\rangle^2}{\|s(t) - s(t')\|^5} \\ &\quad + 2\frac{\langle s'(t)|s(t) - s(t')\rangle \langle s'(t)^\perp|s(t) - s(t')\rangle \langle s'(t')^\perp|s(t) - s(t')\rangle}{\|s(t) - s(t')\|^5}\end{aligned}$$

Nous savons que  $\frac{t'-t}{|t'-t|} \frac{s(t)-s(t')}{\|s(t)-s(t')\|}$  tend vers  $s'(t) = \frac{s''(t)^\perp}{\|s''(t)\|}$  lorsque  $t'$  tend vers  $t$ . Ainsi,  $\frac{\langle s'(t)^\perp|s(t)-s(t')\rangle}{\|s(t)-s(t')\|^2}$  et  $\frac{\langle s''(t)|s(t)-s(t')\rangle}{\|s(t)-s(t')\|^2}$  sont continues en  $(t,t') \in I_1 \times I_2$ . Comme d'autre part  $s'$ ,  $s''$ ,  $s'''$  et  $\frac{t'-t}{|t'-t|} \frac{s(t)-s(t')}{\|s(t)-s(t')\|}$  sont continues aussi, on a montré que toutes les dérivées ci-dessus multipliées par  $\frac{t'-t}{|t'-t|}$  convergent dans  $\mathbb{R}$  lorsque  $t-t'$  diminue, et que donc  $\bar{u}$ ,  $D\bar{u}$ ,  $D^2\bar{u}$  et  $D^3\bar{u}$  sont quatre fonctions continues sur  $I_1 \times I_2$ , pourvu que  $s$  soit  $C^3$ . Un développement limité à reste intégral montre alors que  $\bar{u}(t,t')$ ,  $\frac{\bar{u}(t,t')}{t-t'}$  et  $\frac{t-t'}{\bar{u}(t,t')}$  sont  $C^3$  en  $(t,t')$  si  $s$  est  $C^4$ .

Posons  $\partial_+ := \frac{\partial}{\partial t} + \frac{\partial}{\partial t'}$ . Comme  $s$  est  $C^{q+3}$ , on a

$$\begin{aligned}\partial_+^q (\bar{u}(t,t')^2) &= \sum_{r=0}^q C_q^r \langle \partial_+^r (s(t) - s(t')) | \partial_+^{q-r} (s(t) - s(t')) \rangle \\ &= \sum_{r=0}^q C_q^r \langle (s^{(r)}(t) - s^{(r)}(t')) | (s^{(q-r)}(t) - s^{(q-r)}(t')) \rangle \\ &= \sum_{r=0}^q C_q^r \langle (t-t') s^{(r+1)} \left(\frac{t+t'}{2}\right) + O((t-t')^3) | (t-t') s^{(q-r+1)} \left(\frac{t+t'}{2}\right) + O((t-t')^3) \rangle \\ &= (t-t')^2 \sum_{r=0}^q C_q^r \langle s^{(r+1)} \left(\frac{t+t'}{2}\right) | s^{(q-r+1)} \left(\frac{t+t'}{2}\right) \rangle + O((t-t')^4) \\ &= (t-t')^2 \frac{\partial^q}{\partial z^q} \langle s'(z) | s'(z) \rangle + O((t-t')^4)\end{aligned}$$

où  $z = \frac{t+t'}{2}$ . Lorsque  $q$  est non nul, on se met à dériver  $\|s'\|^2$  qui est constant égal à 1 comme  $s$  est une courbe unicursale, donc la dernière somme ci-dessus est nulle. Un développement limité à reste intégral montre alors que  $\frac{\partial_+^q(\bar{u}^2)}{\bar{u}^4}$  est continu si  $0 < q \leq p$ . Comme  $\partial_+(\bar{u}^2) = 2\bar{u}\partial_+\bar{u}$ ,  $\frac{\partial_+\bar{u}}{\bar{u}^3}$  est  $C^0$  si  $s$  est  $C^4$ . Une récurrence facile basée sur le développement de  $\partial_+^q(\bar{u}^2)$  montre que, pour  $0 < q \leq p$ ,  $\frac{\partial_+^q\bar{u}}{\bar{u}^3}$  est  $C^0$ .

Pour  $q = p + 1$ , le raisonnement ci-dessus ne donne pas  $O((t - t')^4)$  mais seulement  $O((t - t')^3)$ ; alors,  $\frac{\partial_+^{p+1}(\bar{u}^2)}{\bar{u}^3}$  est continu, et  $\frac{\partial_+^{p+1}\bar{u}}{\bar{u}^2}$  est continu. ■

En dehors de la fonction  $u$ , nous devons aussi étudier la fonction apparaissant dans le noyau (23), à savoir

$$v(t, t') = \begin{cases} \lim_{\varepsilon \rightarrow 0^+} \langle s'(t)^\perp | \frac{s(t) - s(t+\varepsilon)}{\|s(t) - s(t+\varepsilon)\|^2} \rangle & \text{si } t = t' \\ \langle s'(t)^\perp | \frac{s(t) - s(t')}{\|s(t) - s(t')\|^2} \rangle & \text{sinon} \end{cases} \quad (75b)$$

**Lemme 4.3.2** *Comme  $s$  est  $C^{p+3}$ ,  $v$  est  $C^p$ .*

■ Pour montrer que  $v$  est  $C^p$ , c'est juste une affaire de développement limité à reste intégral. ■

Comme  $H_0^{ank}$  est solution de l'équation différentielle de Bessel (29c), on a :

$$\begin{bmatrix} H_0^{ank'}(r) \\ H_0^{ank''}(r) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -\frac{1}{r} \end{bmatrix} \begin{bmatrix} H_0^{ank}(r) \\ H_0^{ank'}(r) \end{bmatrix}, r > 0. \quad (75c)$$

À la lumière de (73) et de (23), posons

$$Y(t, t') = H_0^{ank} \left( \frac{1}{\lambda_0} u(t, t') \right) e^{2i\pi \frac{1}{\lambda_0} (k_1 t + t_2 t')} \quad (75d)$$

et

$$Z(t, t') = -H_0^{ank'} \left( \frac{1}{\lambda_0} u(t, t') \right) e^{2i\pi \frac{1}{\lambda_0} (k_1 t + t_2 t')}. \quad (75e)$$

L'identité (75c) se généralise ainsi :

$$\begin{bmatrix} \nabla Y(t, t') \\ \nabla(u(t, t') Z(t, t')) \end{bmatrix} = \frac{1}{\lambda_0} \begin{bmatrix} 2i\pi \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} & -\frac{1}{u(t, t')} \nabla u(t, t') \\ u(t, t') \nabla u(t, t') & 2i\pi \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \end{bmatrix} \begin{bmatrix} Y(t, t') \\ u(t, t') Z(t, t') \end{bmatrix} \quad (75f)$$

où la matrice centrale, définie par blocs, a quatre lignes et deux colonnes. Prenons le produit de  $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}^T$  par (75f), ce qui avec les notations  $\partial_+ := \frac{\partial}{\partial t} + \frac{\partial}{\partial t'}$  et  $k_+ = 2\pi(k_1 + k_2)$ , donne :

$$\partial_+ \begin{bmatrix} Y \\ uZ \end{bmatrix} = \frac{1}{\lambda_0} \begin{bmatrix} ik_+ & -\frac{\partial_+ u}{u} \\ u\partial_+ u & ik_+ \end{bmatrix} \begin{bmatrix} Y \\ uZ \end{bmatrix} \quad (75g)$$

Le déterminant de cette matrice vaut  $(\partial_+ u)^2 - k_+^2$ . Nous allons intégrer par parties avec la dérivation vectorielle suivante :

$$D : \begin{bmatrix} f \\ g \end{bmatrix} \mapsto \left( (t, t') \mapsto \partial_+ \left( \frac{1}{(\partial_+ \bar{u})^2 - k_+^2} \begin{bmatrix} ik_+ & -\bar{u}\partial_+ \bar{u} \\ \frac{\partial_+ \bar{u}}{\bar{u}} & ik_+ \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} \right) \right) \quad (75h)$$

Tout va bien, car le premier lemme montre que  $\frac{\partial_+ \bar{u}}{\bar{u}}$  continu. Nous allons intégrer parties  $p$  fois. On a  $D^T \begin{bmatrix} Y \\ uZ \end{bmatrix} = \frac{1}{\lambda_0} \begin{bmatrix} Y \\ uZ \end{bmatrix}$ , donc (73) vaut

$$\frac{1}{\lambda_0} \iint \begin{bmatrix} f \\ g \end{bmatrix}^T (\lambda_0 D^T)^p \begin{bmatrix} Y \\ uZ \end{bmatrix} dt dt' = \frac{1}{\lambda_0} (-\lambda_0)^p \iint \begin{bmatrix} Y \\ uZ \end{bmatrix}^T D^p \begin{bmatrix} f \\ g \end{bmatrix} dt dt' \quad (75i)$$

avec  $f(t, t') = \frac{\alpha}{4} b_1(t) b_2(t')$  et  $g(t, t') = i\pi v(t, t') b_1(t) b_2(t')$  qui sont deux fonctions indépendantes de  $\lambda_0$ . Notons que  $Y$  et  $Z$  dépendent de  $\lambda_0$ . Chacune des  $p$  dérivations  $\partial_+$  présente dans  $D$  va dériver tour à tour chaque fonction présente subséquentement, que ce soit  $f$  ou  $g$ , ou une fonction de la dérivation à coefficients non constants  $D$ . Dériver  $f(t, t')$  ou  $g(t, t')$  jusqu'à  $p$  fois est possible de façon contrôlée vu les estimations de régularités disponibles sur  $v$ ,  $b_1$  et  $b_2$ ; dériver  $v(t, t')$  jusqu'à  $p$  fois n'est pas un problème grâce au second lemme. Dériver chacun des termes de la matrice

$$M = \begin{bmatrix} ik_+ & -\bar{u}\partial_+ \bar{u} \\ \frac{\partial_+ \bar{u}}{\bar{u}} & ik_+ \end{bmatrix}$$

présente dans la définition de  $D$  ne grossit pas leur importance, grâce au premier lemme :  $\partial_+^q \left( \frac{\partial_+ \bar{u}}{\bar{u}} \right)$  est un produit de facteurs du genre  $\frac{\partial_+^q \bar{u}}{\bar{u}}$ , donc  $\partial_+^q \left( \frac{\partial_+ \bar{u}}{\bar{u}} \right)$  est en  $\bar{u}^2$  si  $q < p$ , et c'est en  $u$  si  $q = p$ : ainsi, dans tous les termes du développement des dérivées de (75i), les matrices  $M$  aux coefficients plus ou moins dérivés restent toutes de taille en  $1 + |k_+|$ .

Dériver le dénominateur de  $D$  dans (75i) fait apparaître un facteur multiplicatif de la forme suivante, où  $n \leq p$ :

$$\frac{\partial_+^n (\partial_+ \bar{u})^2}{(\partial_+ \bar{u})^2 - k_+^2} \quad (75j)$$

qui a un module contrôlé par  $\frac{\bar{u}^5}{(\partial_+ \bar{u})^2 - k_+^2}$  car  $s \in \mathbb{C}^{p+3}$  et  $n \leq p$ .

Comme au plus un seul facteur multiplicatif du type (75j) peut apparaître pour chacune des  $p$  dérivations de (75i), et comme (75i) ne contient lui-même que  $p$  représentants (avant ou après dérivation) du dénominateur de  $D$ , tous les termes du développement des dérivées de (75i) sont plus petits qu'une constante fois la racine de

$$\iint \left( \frac{|k_+| + 1}{(\partial_+ \bar{u})^2 - k_+^2} \right)^{2p} \max(|Y|, |uZ|)^2 \max\left(1, \left( \frac{\bar{u}^5}{(\partial_+ \bar{u})^2 - k_+^2} \right)^{2p}\right) dt dt' \\ \times (\|f\|_{\mathbb{H}^p}^2 + \|g\|_{\mathbb{H}^p}^2) \lambda_0^{2p-2}.$$

D'après (75d), (75e), (29d) et (29e), on a

$$\lambda_0^{-1/2} \max(|Y|, |uZ|) = O\left(\begin{cases} \lambda_0^{-1/2} |\log \frac{u}{\lambda_0}| & \text{si } |u| < \lambda_0 \\ u^{-1/2} & \text{sinon} \end{cases}\right)$$

qui a une norme  $L^2$  en  $O(|\log \lambda_0|)$ , ce qui donne le résultat.  $\blacksquare$

**Propriété 4.4** *L'intégrale (73) est en*

$$O\left(\frac{1}{|k_-|} \frac{1}{1 + \min\left(\frac{\|b_1\|_{L^2}}{\|b_1\|_{H^p}}, \frac{\|b_2\|_{L^2}}{\|b_2\|_{H^p}}\right) \min\left(\left(\frac{|k_+|}{\lambda_0}\right)^p, \frac{1}{|\log \lambda_0|} \left(\frac{|k_-|}{\lambda_0}\right)^{p-1}\right)}\right) \|b_1\|_{L^2} \|b_2\|_{L^2} \quad (76)$$

$$+ \frac{\lambda_0}{|k_-|} (\|b_1\|_{H^1(\partial S)} + \|b_2\|_{H^1(\partial S)}) + \frac{\lambda_0^{3/2}}{|k_-|^2} (\|b'_1\|_{W^{1,1}} + \|b'_2\|_{W^{1,1}})$$

avec les suppositions que les éléments  $b_1$  et  $b_2$  de la base sont  $H^p$  avec  $p > 0$ , que  $s$  est  $C^{p+3}$ , et que  $|k_-|$  est plus grand qu'une constante  $K_0$  indépendante de  $p$ , de  $b_1$ , de  $b_2$ , de  $k_1$ , de  $k_2$ , de  $S$  et de  $\lambda_0$ . Nous utiliserons la notation  $W^{m,p}$  pour l'espace de Sobolev habituel. La constante du  $O()$  ne dépend que de  $S$  et de  $p$ .

$\blacksquare$  Nous supposons que  $K_0 \geq 6\pi$ . La preuve est compliquée, nous allons d'abord la faire au cas où le diffractant  $S$  est un disque de rayon  $R = \frac{L}{2\pi}$ , puis pour le cas général. Rappelons que  $k_1 \frac{L}{\lambda_0}$  et  $k_2 \frac{L}{\lambda_0}$  sont entiers.

**Cas du disque, avec  $b_1(t)$  et  $b_2(t)$  valant 1 pour tout  $t$ .**

Dans ce cas, nous avons vu dans la démonstration de la propriété 4.1 sur la norme de  $A^{-1}$  que l'opérateur  $A$  était diagonal dans la base  $(e_m)_{m \in \mathbb{Z}}$  définie en (70), et que la valeur propre associée à  $e_m$  était (71), qu'on note temporairement  $v_m$ . Watson [W22] donne le développement en série suivant :

$$\begin{aligned} H_m^{ank}(2r) &= \sum_{n=0}^{\infty} \frac{(-1)^n r^{m+2n}}{n!(m+n)!} - i \sum_{n=0}^{m-1} \frac{(m-n-1)! r^{-m+2n}}{n! \pi} \\ &+ i \sum_{n=0}^{\infty} \frac{(-1)^n r^{m+2n}}{n!(m+n)! \pi} \left(2C - \sum_{p=0}^n \frac{1}{p} - \sum_{p=0}^{m+n} \frac{1}{p} + 2 \log r\right). \end{aligned}$$



En utilisant la formule de Stirling  $n! \sim \sqrt{2n\pi + 1} \left(\frac{n}{e}\right)^n$  nous trouvons deux constantes  $m_0$  et  $r_0$  telles que pour tout  $m \geq m_0$  et  $r < r_0 m$ ,

$$\frac{r^m}{(m-1)!} H_m^{ank}(2r) = \frac{i}{\pi} \sum_{n=0}^{m/2} \frac{(m-n-1)!}{(m-1)!n!} r^{2n} + O\left(\left(\frac{r}{m}\right)^m \log \frac{r}{m}\right)$$

et

$$\frac{m!}{r^m} \mathcal{R}e H_m^{ank}(2r) = \sum_{n=0}^{m/2} \frac{(-1)^n m! r^{2n}}{(m+n)!n!} + O\left(\left(\frac{r}{m}\right)^m\right)$$

uniformément en  $m$  et  $r$ . En utilisant ces deux développements limités avec la formule de dérivation  $H_m^{ank'}(2r) = \frac{1}{2} H_{m-1}^{ank}(2r) - \frac{1}{2} H_{m+1}^{ank}(2r)$  dans (71) avec  $r = \frac{\pi R}{\lambda_0}$ , on a

$$(71) = \left( \sum_{n=0}^{m/2} \frac{(m-n-1)!}{(m-1)!n!} r^{2n} \right) \left( -\frac{i\alpha r}{2\pi m} \sum_{n=0}^{m/2} \frac{(-1)^n m! r^{2n}}{(m+n)!n!} + \sum_{n=0}^{m/2} \frac{(-1)^n (m-1)! r^{2n}}{(m+n-1)!n!} - \frac{r^2}{m^2+m} \sum_{n=0}^{m/2} \frac{(-1)^n (m+1)! r^{2n}}{(m+n+1)!n!} \right) + O\left(\left(\frac{r}{m}\right)^m \log \frac{r}{m}\right) O\left(\left(\frac{r}{m-1}\right)^{m-1}\right) + O\left(\frac{r}{m}\right)$$

où le reste  $O\left(\frac{r}{m}\right)$  contient les produits entre un des restes et une des sommes provenant des deux développements limités ci-dessus, c'est à dire

$$\begin{aligned} & \frac{\alpha r}{2m} \frac{i}{\pi} \sum_{n=0}^{m/2} \frac{(m-n-1)!}{(m-1)!n!} r^{2n} O\left(\left(\frac{r}{m}\right)^m\right) \\ & \frac{\alpha r}{2m} \sum_{n=0}^{m/2} \frac{(-1)^n m! r^{2n}}{(m+n)!n!} O\left(\left(\frac{r}{m}\right)^m \log \frac{r}{m}\right) \\ & -i\pi \frac{i}{\pi} \sum_{n=0}^{m/2} \frac{(m-n-1)!}{(m-1)!n!} r^{2n} O\left(\left(\frac{r}{m-1}\right)^{m-1}\right) \\ & -i\pi \sum_{n=0}^{(m-1)/2} \frac{(-1)^n (m-1)! r^{2n}}{(m+n-1)!n!} O\left(\left(\frac{r}{m}\right)^m \log \frac{r}{m}\right) \\ & \frac{i\pi r^2}{m(m+1)} \frac{i}{\pi} \sum_{n=0}^{m/2} \frac{(m-n-1)!}{(m-1)!n!} r^{2n} O\left(\left(\frac{r}{m+1}\right)^{m+1}\right) \end{aligned}$$

et

$$\frac{i\pi r^2}{m(m+1)} \sum_{n=0}^{(m+1)/2} \frac{(-1)^n (m+1)! r^{2n}}{(m+n+1)!n!} O\left(\left(\frac{r}{m}\right)^m \log \frac{r}{m}\right)$$

dont les termes sont en  $O\left(\frac{r}{m}\right)$  si on suppose que  $2 \log m_0 < -\sqrt{m_0} \log r_0$  (faire un raisonnement différent selon que  $n < \sqrt{m}$ , auquel cas on a  $n \log m <$

$m \log r_0$  et donc  $m^n < (\frac{r}{m})^m$  qui donne assez de marge de manœuvre, ou non, auquel cas le facteur  $(n!)^2$  contient  $n^{2n}$  qui vaut  $m^n$  et facilite les choses). De plus, nous avons suffisamment de marge dans les estimations de ces termes pour que la somme soit en  $O(\frac{r}{m})$  elle aussi.

En recombinant ces sommes, nous voyons que (71) vaut

$$\begin{aligned} \frac{i\alpha}{2\pi r} \sum_{n=1}^{m/2} \frac{r^{2n}}{n!} \sum_{p=1}^n (-1)^p \binom{n-1}{p} f_n(m+p) + \sum_{n=0}^{m/2} \frac{r^{2n}}{n!} \sum_{p=0}^n (-1)^p \binom{n}{p} f_n(m+p) + \\ r^{-2} \sum_{n=2}^{m/2} \frac{r^{2n}}{n!} \sum_{p=2}^n (-1)^p \binom{n-2}{p} f_n(m+p) + O\left(\frac{r}{m}\right) \end{aligned}$$

avec  $f_n(x) = \prod_{i=1}^n \frac{1}{x-i}$ . En appliquant le théorème de Rolle et avec un raisonnement par récurrence sur  $n$ , nous voyons qu'il existe  $c \in [0, n]$  tel que  $\sum_{p=0}^n \binom{n}{p} (-1)^p f(p) = f^{(n)}(c)$  (appliquer le théorème de Rolle à  $t \mapsto \sum_{p=0}^n \binom{n}{p} (-1)^p (f(tp) - f(p)t^n)$  puis enchaîner l'hypothèse de récurrence). Alors les sommes sur  $p$  ci-dessus valent respectivement  $f_n^{(n-1)}(m+1+c)$ ,  $f_n^{(n)}(m+c')$  et  $f_n^{(n-2)}(m+2+c'')$ , avec  $c, c', c'' \geq 0$ . Comme  $n \leq \frac{m}{2}$  implique  $|f_n^{(p)}(m+c)| \leq \frac{(n+p)!}{n!} (\frac{m}{2})^{-n-p}$ , nous voyons donc que si  $r_0$  est suffisamment petit ( $r_0 \leq \frac{1}{4}$ ), on a (71) =  $1 + O(\frac{r}{m})$ .

La preuve est terminée. Nous corroborons le résultat de la propriété 4.4 par une étude numérique: les figures 30 et 31 montrent que cette valeur propre  $v_m - 1$  de l'opérateur  $A - \text{Id}$  est en  $O(\frac{1}{\lambda_0 m})$  uniformément en  $\lambda_0$  et en  $m > \frac{1}{\lambda_0}$ .

### Cas du disque, avec $b_1$ et $b_2$ quelconques.

Maintenant, nous avons une estimation pour (73), mais uniquement lorsque  $b_1$  et  $b_2$  sont constants égaux à 1, grâce au cas précédent. Pour passer au cas général, nous allons nous baser sur le fait que l'ajout de  $b_1$  et  $b_2$  reviennent à des produits de fonctions, qui dans la base de Fourier (70) s'interprètent par la convolution par les transformées de Fourier  $L$ -périodiques de  $b_1$  et de  $b_2$ .

Soit  $A = \min(\frac{\|b_1\|_{L^2}}{\|b_1\|_{\mathbb{H}^p}}, \frac{\|b_2\|_{L^2}}{\|b_2\|_{\mathbb{H}^p}})$  l'expression intervenant dans l'estimation de l'énoncé. Utilisons-là. Nous avons, uniformément en  $b_1$  et en  $\omega \in \mathbb{Z}$ :

$$\hat{b}_1\left(\frac{2\pi}{L}\omega\right) = O\left(\min(\|b_1\|_{L^2} e_1(\omega), \frac{\|b_1\|_{\mathbb{H}^p} e_2(\omega)}{(1+|\omega|)^p})\right) = O\left(\frac{\max(e_1(\omega), e_2(\omega))}{\frac{1}{\|b_1\|_{L^2}} + \frac{1+|\omega|^p}{\|b_1\|_{\mathbb{H}^p}}}\right)$$

avec  $e_1$  et  $e_2$  deux fonctions  $\mathbb{Z} \rightarrow \mathbb{R}_+$  de norme  $\ell^2$  unitaire. Ainsi nous voyons que  $\left((1 + A|\omega|^p)b_1\left(\frac{2\pi}{L}\omega\right) / \|b_1\|_{L^2}\right)_{\omega \in \mathbb{Z}}$  a une norme  $\ell^2$  bornée uniformément en  $b_1$ . Il en est de même pour  $b_2$ .

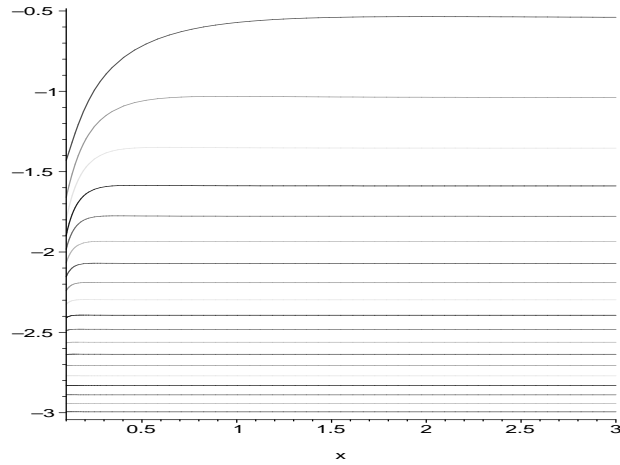


FIG. 30 – Graphes du logarithme  $\log |v_m - 1|$  du module de la valeur propre de  $A - Id$  en fonction de  $\frac{2\pi R}{\lambda_0}$ , à  $\lambda_0 m$  fixé. Les 19 courbes correspondent à  $\lambda_0 m$  fixé respectivement à  $4\pi R$ ,  $6\pi R$ , ...,  $40\pi R$ .

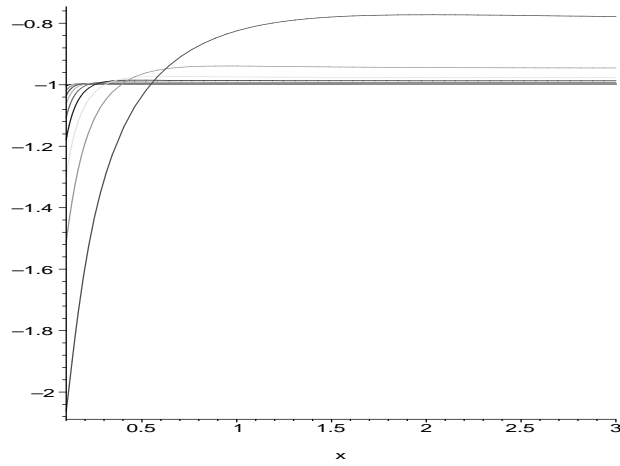


FIG. 31 – Graphes du rapport  $\frac{\log |v_m - 1|}{\log(\lambda_0 m)}$  en fonction de  $\frac{2\pi R}{\lambda_0}$ , à  $\lambda_0 m$  fixé. Les 19 courbes correspondent à  $\lambda_0 m$  fixé respectivement à  $4\pi R$ ,  $6\pi R$ , ...,  $40\pi R$ . Le rapport tend visiblement vers  $-1$  pour  $\lambda_0 < 1$  et lorsque  $\lambda_0 m$ .

Nous avons donc, en utilisant Cauchy-Schwarz :

$$\begin{aligned}
(73) &= O\left(\sum_{\omega \in \mathbb{Z}} \hat{b}_1\left(2\pi \frac{\omega}{L}\right) \hat{b}_2\left(2\pi \frac{-\omega - \frac{k_1 L}{\lambda_0} - \frac{k_2 L}{\lambda_0}}{L}\right) (v_{|\frac{k_1 L}{\lambda_0} + \omega|} - 1)L\right) \\
&= O\left(\sup_{\omega \in \mathbb{Z}} \frac{\|b_1\|_{L^2}}{1 + A|\omega|^p} \frac{\|b_2\|_{L^2}}{1 + A\left|-\omega - \frac{k_1 L}{\lambda_0} - \frac{k_2 L}{\lambda_0}\right|^p} (v_{|\frac{k_1 L}{\lambda_0} + \omega|} - 1)\right)
\end{aligned}$$

Dans le cas où  $|k_1| \geq |k_2|$ , nous avons  $|k_1| \geq \frac{1}{2}|k_1 - k_2| = \frac{|k_-|}{4\pi} \geq \frac{3}{2}$ . Dans le cas où la première partie de cette démonstration fournit une estimation sur  $v_*$ , c'est à dire lorsque  $|\frac{k_1 L}{\lambda_0} + \omega| \geq \frac{\pi R}{\lambda_0 r_0}$ , on a

$$\begin{aligned}
&\sup_{\omega \in \mathbb{Z}: |\frac{k_1 L}{\lambda_0} + \omega| \geq \frac{\pi R}{\lambda_0 r_0}} \frac{\|b_1\|_{L^2}}{1 + A|\omega|^p} \frac{\|b_2\|_{L^2}}{1 + A\left|-\omega - \frac{k_1 L}{\lambda_0} - \frac{k_2 L}{\lambda_0}\right|^p} (v_{|\frac{k_1 L}{\lambda_0} + \omega|} - 1) \\
&= O\left(\frac{\|b_1\|_{L^2} \|b_2\|_{L^2}}{1 + A\left(\frac{|k_+|}{\lambda_0}\right)^p} \frac{1}{k_-}\right)
\end{aligned}$$

Lorsqu'au contraire  $|k_1| \geq |k_2|$  et  $|\frac{k_1 L}{\lambda_0} + \omega| < \frac{\pi R}{\lambda_0 r_0}$ , nous avons

$$|\omega| > \frac{|k_1|L}{\lambda_0} - \frac{\pi R}{\lambda_0 r_0} \geq \frac{|k_-|L}{4\lambda_0\pi} - \frac{\pi R}{\lambda_0 r_0}$$

ce qui est plus grand que  $\frac{|k_-|L}{2\lambda_0}$  si nous imposons que  $K_0 \geq \frac{4\pi}{r_0}$ . Dans ce cas, nous avons tout de même l'estimation  $v_m \leq \frac{|\log \lambda_0|}{\lambda_0}$  fournie par la partie 2.3.3 (améliorable), ce qui montre que

$$\begin{aligned}
&\sup_{\omega \in \mathbb{Z}: |\frac{k_1 L}{\lambda_0} + \omega| < \frac{\pi R}{\lambda_0 r_0}} \frac{\|b_1\|_{L^2}}{1 + A|\omega|^p} \frac{\|b_2\|_{L^2}}{1 + A\left|-\omega - \frac{k_1 L}{\lambda_0} - \frac{k_2 L}{\lambda_0}\right|^p} (v_{|\frac{k_1 L}{\lambda_0} + \omega|} - 1) \\
&= O\left(\frac{\|b_1\|_{L^2}}{1 + A\left|\frac{|k_-|L}{2\lambda_0}\right|^p} \frac{\|b_2\|_{L^2}}{1} \frac{|\log \lambda_0|}{\lambda_0}\right)
\end{aligned}$$

ce qui donne l'estimation de l'énoncé.

Le cas  $|k_2| \geq |k_1|$  se fait de la même manière que  $|k_1| \geq |k_2|$ , il suffit de poser  $\omega' = -\omega - \frac{k_1 L}{\lambda_0} - \frac{k_2 L}{\lambda_0}$ . Ainsi, nous avons fini la démonstration pour  $S$  étant le disque et  $b_1$  et  $b_2$  deux fonctions quelconques.

### Début du cas général.

Nous nous intéressons donc au cas où  $k_+$  peut être petit. Dans ce cas, nous allons utiliser  $k_- = 2\pi(k_1 - k_2)$ . Ici, nous supposons que  $|k_-| > 6\pi$

pour simplifier la démonstration, qui est inspirée de la démonstration de la propriété 4.3; toutefois, le traitement de la diagonale ( $u$  petit) est plus difficile ici. Par exemple, nous serons amenés à manipuler la fonction  $\frac{\partial_- u}{u}$ , avec  $\partial_- := \partial_1 - \partial_2 = \frac{\partial}{\partial t} - \frac{\partial}{\partial t'}$ : cette fonction aura une singularité marquée pour  $t$  proche de  $t'$ .

### Utilisation du cas du cercle.

Pour compenser cette singularité, nous utiliserons simultanément les mêmes transformations (intégrations par parties, découpes d'intégrales) sur le cas du diffractant  $S$ , et sur le cas où :

- le diffractant  $S$  a été remplacé par  $\hat{S}$  qui est le disque de circonférence  $L$ ;
- la fonction  $u$  a été remplacée par  $\hat{u}$  définie par

$$\hat{u}(t, t') = 2L \sin \pi \frac{|t - t'|}{L};$$

- le noyau  $K$  a été remplacé par  $\hat{K}$ , défini sur le même modèle ce qui donne (68);
- l'opérateur  $A$  a aussi été remplacé par  $\hat{A}$ .

Ces quatre remplacements correspondent au transfert entre le cas  $S$  général et le cas  $S$  disque de circonférence  $L$ .

Grâce aux raisonnements pour le cas où  $S$  est un disque, nous savons déjà que

$$\int_0^L \int_0^L \hat{K}(t, t') b_1(t) e^{2i\pi \frac{k_1}{\lambda_0} t} b_2(t') e^{2i\pi \frac{k_2}{\lambda_0} t'} dt dt' \quad (77a)$$

est en (76). Pour montrer la propriété 4.4, il reste donc à montrer que

$$(73) - (77a) \text{ est en (76).}$$

### Intégrations par parties.

Prenons le produit de l'équation différentielle matricielle (75f) par  $\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}^T$  pour nous adapter au signe moins présent dans la définition de  $k_-$ . Avec la notation  $\partial_- := \partial_1 - \partial_2 = \frac{\partial}{\partial t} - \frac{\partial}{\partial t'}$ , nous avons alors :

$$\partial_- \begin{bmatrix} Y \\ uZ \end{bmatrix} = \frac{1}{\lambda_0} \begin{bmatrix} ik_- & -\frac{\partial_- u}{u} \\ u\partial_- u & ik_- \end{bmatrix} \begin{bmatrix} Y \\ uZ \end{bmatrix}$$

Le déterminant de cette matrice vaut  $\partial_- u^2 - k_-^2$ , et la propriété 4.2 et nous allons intégrer par parties avec la dérivation vectorielle suivante :

$$D : \begin{bmatrix} f \\ g \end{bmatrix} \mapsto (t, t') \mapsto \partial_- \left( \frac{1}{(\partial_- \bar{u})^2 - k_-^2} \begin{bmatrix} ik_- & -\bar{u}\partial_- \bar{u} \\ \frac{\partial_- \bar{u}}{\bar{u}} & ik_- \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} \right)$$

Notons que  $d_- = k_-^2 - (\partial_- \bar{u})^2 \geq k_-^2 - 16\pi^2 \geq \frac{5}{8}|k_-|^2$  (propriété 4.2) : ainsi, nous n'aurons pas de problème de petitesse de  $|(\partial_- \bar{u})^2 - k_-^2|$ , ni de problème relatifs à ses relations avec  $\partial_-$ .

En revanche, il faudra faire attention aux effets des dérivations  $\partial_-$  sur  $\frac{\partial_- \bar{u}}{\bar{u}}$  qui n'est pas continu en  $t = t'$ . Nous ne ferons que deux intégrations par parties, et encore seulement sur les sous-domaines  $A_X = \{(t, t') : |t - t'| > X\}$  et  $A_{\lambda_0} = \{(t, t') : |t - t'| > \lambda_0\}$ , pour un réel  $X$  vérifiant  $X < \lambda_0$ ; nous allons faire tendre  $X$  vers 0. Comme  $D^T \begin{bmatrix} Y \\ uZ \end{bmatrix} = \frac{1}{\lambda_0} \begin{bmatrix} Y \\ uZ \end{bmatrix}$ , ces deux intégrations par parties donnent

$$\begin{aligned} (73) &= \int \int \begin{bmatrix} \frac{1}{\lambda_0} f \\ \frac{1}{\lambda_0} g \end{bmatrix}^T \lambda_0 D^T \begin{bmatrix} Y \\ uZ \end{bmatrix} dt dt' \\ &= \lambda_0^{-1} \int \int_{\{(t, t') : |t - t'| < X\}} \begin{bmatrix} f \\ g \end{bmatrix}^T \begin{bmatrix} Y \\ uZ \end{bmatrix} dt dt' \\ &+ \int \int_{A_X} \begin{bmatrix} f \\ g \end{bmatrix}^T \frac{1}{\partial_- \bar{u}^2 - k_-^2} \begin{bmatrix} ik_- & \frac{\partial_- \bar{u}}{\bar{u}} \\ -\bar{u}\partial_- \bar{u} & ik_- \end{bmatrix} \partial_- \begin{bmatrix} Y \\ uZ \end{bmatrix} dt dt' \\ &= \lambda_0^{-1} \int \int_{A_X} \begin{bmatrix} Y \\ uZ \end{bmatrix}^T \begin{bmatrix} f \\ g \end{bmatrix} dt dt' \end{aligned} \quad (77b)$$

$$- \int \int_{A_X \setminus A_{\lambda_0}} \begin{bmatrix} Y \\ uZ \end{bmatrix}^T \partial_- \left( \frac{1}{\partial_- \bar{u}^2 - k_-^2} \begin{bmatrix} ik_- & -\bar{u}\partial_- \bar{u} \\ \frac{\partial_- \bar{u}}{\bar{u}} & ik_- \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} \right) dt dt' \quad (77c)$$

$$\pm 2 \int_{t' = t \pm X} \begin{bmatrix} Y \\ uZ \end{bmatrix}^T \frac{1}{\partial_- \bar{u}^2 - k_-^2} \begin{bmatrix} ik_- & -\bar{u}\partial_- \bar{u} \\ \frac{\partial_- \bar{u}}{\bar{u}} & ik_- \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} dt \quad (77d)$$

$$+ \lambda_0 \int \int_{A_{\lambda_0}} \begin{bmatrix} Y \\ uZ \end{bmatrix}^T D^2 \begin{bmatrix} f \\ g \end{bmatrix} dt dt' \quad (77e)$$

$$\pm 2\lambda_0 \int_{t' = t - (\pm \lambda_0)} \begin{bmatrix} Y \\ uZ \end{bmatrix}^T \frac{1}{\partial_- \bar{u}^2 - k_-^2} \begin{bmatrix} ik_- & -\bar{u}\partial_- \bar{u} \\ \frac{\partial_- \bar{u}}{\bar{u}} & ik_- \end{bmatrix} D \begin{bmatrix} f \\ g \end{bmatrix} dt \quad (77f)$$

où  $f$  et  $g$  sont comme dans la démonstration précédente des fonctions  $C^p$  indépendantes de  $\lambda_0$ .

**Estimations de comparaisons entre le cas général et le cas du disque.**

L'intégrale (77a) peut se découper exactement comme nous venons de le faire pour (73). Notons  $(\widehat{77b}) + (\widehat{77c}) + (\widehat{77d}) + (\widehat{77e}) + (\widehat{77f})$  cette décomposition ; notons  $\hat{D}$  la dérivation analogue à  $D$ , et  $\hat{Z}$ ,  $\hat{Y}$ ,  $\hat{v}$ ,  $\hat{f}$  et  $\hat{g}$  les analogues de  $Z$ ,  $Y$ ,  $v$ ,  $f$  et  $g$ , toutes ces fonction qui ont été définies après (77a).

Les dérivées partielles d'ordre au plus deux de  $u$  et de  $\hat{u}$  sont, d'après leurs expressions dans la preuve du lemme 4.3.1, égales lorsque  $t = t'$  : donc  $u - \hat{u} = O(u^3)$ .

Vu les développements limités (29d) et (29e), on voit que  $Y - \hat{Y} = O(u^2)$ ,  $Z - \hat{Z} = O(u\lambda_0)$  et  $uZ - \hat{u}\hat{Z} = O(u^3 \log(\frac{u}{\lambda_0})/\lambda_0)$  lorsque  $u \leq \lambda_0$ .

D'après les équivalents (32a) et (32b), lorsque  $u \geq \lambda_0$  on a

$$Y - \hat{Y} = O\left(\sqrt{\frac{\lambda_0}{u}} - \sqrt{\frac{\lambda_0}{\hat{u}}}\right) = O\left(\sqrt{\frac{\lambda_0}{u}}\left(1 - \sqrt{1 + \frac{u - \hat{u}}{\hat{u}}}\right)\right) = O(\sqrt{\lambda_0 u^3})$$

et

$$uZ - \hat{u}\hat{Z} = O(\sqrt{\lambda_0 u^5}).$$

$\hat{v}$  et  $v$  sont deux fonctions différentes. Il s'en suit que  $\hat{g}$  et  $g$  sont deux fonctions différentes. On a  $\hat{f} = f$  car ces deux fonctions ne dépendent que de  $b_1$  et de  $b_2$ .

Un premier gain de la comparaison ave le cas du disque est que le terme central (77b), à priori mauvais à cause du terme  $\lambda_0^{-1}$ , tend en fait vers 0 uniformément en  $X \in (0, \lambda_0)$ .

**Le terme (77b)–(77b) est suffisamment petit.**

Raisonnons d'abord pour  $X \in (0, \lambda_0)$ . Ensuite, nous ferons tendre  $X$  vers 0. Lorsque  $u < \lambda_0$ , on a  $Y - \hat{Y} = O(|\log \frac{u}{\hat{u}}|) = O(\bar{u}^2)$ , et  $uZ = O(\lambda_0)$ . Ainsi, on a

$$\begin{aligned} (77b) - (\widehat{77b}) &= \lambda_0^{-1} \int \int_{(t,t'); |t-t'| < X} (Y - \hat{Y})f + guZ - \hat{g}\hat{u}\hat{Z} \\ &= O\left((\|f\|_\infty + \|g\|_\infty + \|\hat{g}\|_\infty)\left(\frac{X^3}{\lambda_0} + X\right)\right) \end{aligned}$$

qui tend bien vers 0 lorsque  $X$  tend vers 0. Donc il existe  $X > 0$  tel que l'énoncé de la proposition est vérifié.

**Dériver  $\frac{1}{\bar{u}}$  dans le terme (77c)–(77c).**

Le terme (77c)–(77c) contient la dérivation  $\partial_-$ . Il y a un problème si elle dérive le dénominateur de  $\frac{\partial_- \bar{u}}{\bar{u}}$ , auquel cas on se retrouve avec  $-(\frac{\partial_- \bar{u}}{\bar{u}})^2$ , ce qui donne la contribution

$$-\iint_{A_X \setminus A_{\lambda_0}} uZ \frac{1}{\partial_- \bar{u}^2 - k_-^2} \left(\frac{\partial_- \bar{u}}{\bar{u}}\right)^2 f dt dt' + \iint_{A_X \setminus A_{\lambda_0}} \hat{u}\hat{Z} \frac{1}{\partial_- \bar{\hat{u}}^2 - k_-^2} \left(\frac{\partial_- \bar{\hat{u}}}{\bar{\hat{u}}}\right)^2 f dt dt' \quad (77g)$$

qui admet la décomposition « massacre » suivante :

$$\begin{aligned} & - \int \int_{A_X \setminus A_{\lambda_0}} (uZ - \hat{u}\hat{Z}) \frac{1}{\partial_- \bar{u}^2 - k_-^2} \left(\frac{\partial_- \bar{u}}{\bar{u}}\right)^2 f dt dt' \\ & - \int \int_{A_X \setminus A_{\lambda_0}} \hat{u}\hat{Z} \frac{(\partial_- \bar{\hat{u}} - \partial_- \bar{u})(\partial_- \bar{\hat{u}} + \partial_- \bar{u})}{(\partial_- \bar{u}^2 - k_-^2)(\partial_- \bar{\hat{u}}^2 - k_-^2)} \left(\frac{\partial_- \bar{u}}{\bar{u}}\right)^2 f dt dt' \\ & - \int \int_{A_X \setminus A_{\lambda_0}} \hat{u}\hat{Z} \frac{1}{\partial_- \bar{\hat{u}}^2 - k_-^2} \frac{(\partial_- \bar{u} - \partial_- \bar{\hat{u}})(\partial_- \bar{u} + \partial_- \bar{\hat{u}})}{\bar{u}^2} f dt dt' \\ & - \int \int_{A_X \setminus A_{\lambda_0}} \hat{u}\hat{Z} \frac{1}{\partial_- \bar{\hat{u}}^2 - k_-^2} (\partial_- \bar{\hat{u}})^2 \frac{(\bar{u} - \bar{\hat{u}})(\bar{u} + \bar{\hat{u}})}{(\bar{u}\bar{\hat{u}})^2} f dt dt' \end{aligned} \quad (77h)$$

On a  $u(t, t') \leq \lambda_0$ , lorsque  $(t, t') \in A_X \setminus A_{\lambda_0}$ , donc on a  $uZ - \hat{u}\hat{Z} = O(u^3 \log(\frac{u}{\lambda_0})/\lambda_0)$ ,  $\partial_- \bar{u} - \partial_- \bar{\hat{u}} = O(u^2)$ ,  $\frac{1}{\bar{u}} - \frac{1}{\bar{\hat{u}}} = O(u)$  et  $uZ = O(\lambda_0)$ , ce qui avec la décomposition (77h) de (77g) montre que (77g) est en

$$O\left(\|f\|_{L^\infty}(\lambda_0/d_- + \lambda_0^2/d_-^2 + \lambda_0^2/d_- + \lambda_0^2/d_-)\right) \leq O\left(\|f\|_{H^1(\partial S)}(\lambda_0/d_- + \lambda_0^2/d_-^2)\right).$$

**Dériver  $f$  est difficile dans le terme (77c)–(77c).**

Nous ne voulons pas utiliser  $\|\nabla f\|_\infty$  dans l'estimation. Il faut donc étudier plus particulièrement la contribution suivante :

$$\int \int_{A_X \setminus A_{\lambda_0}} \frac{Yik_- + uZ \frac{\partial_- \bar{u}}{\bar{u}}}{\partial_- \bar{u}^2 - k_-^2} \partial_- f dt dt' - \int \int_{A_X \setminus A_{\lambda_0}} \frac{\hat{Y}ik_- + \hat{u}\hat{Z} \frac{\partial_- \bar{\hat{u}}}{\bar{\hat{u}}}}{\partial_- \bar{\hat{u}}^2 - k_-^2} \partial_- f dt dt'$$

Cette contribution admet une décomposition en une forme similaire à (77h) ; En utilisant la même technique que pour (77h), et en utilisant  $f \in H^1(\partial S)$ ,  $\frac{\partial_- \bar{u}}{\bar{u}} - \frac{\partial_- \bar{\hat{u}}}{\bar{\hat{u}}} = O(u)$  et  $Y - \hat{Y} = O(u^2)$ , on montre que ce terme est en

$$O\left(\|f\|_{H^1(\partial S)}\left(\frac{\lambda_0^{5/2}}{d_-}k_- + \frac{\lambda_0^{5/2}}{d_-^2}k_- + \frac{\lambda_0^{3/2}}{d_-} + \frac{\lambda_0^{5/2}}{d_-^2} + \frac{\lambda_0^{5/2}}{d_-}\right)\right) \leq O\left(\|f\|_{H^1(\partial S)}\left(\frac{\lambda_0^{5/2}}{d_-}k_- + \frac{\lambda_0^{3/2}}{d_-}\right)\right)$$

où on a utilisé  $d_- > 1$  et  $\lambda_0 < 1$ .



**Dériver  $\partial_- u$  est difficile en présence de  $\frac{1}{\bar{u}}$  ou de  $Yik_-$  dans le terme (77c)–(77c).**

Il y a un problème plus petit si la dérivation  $\partial_-$  attaque un des  $\partial_- u$  en présence du terme  $\frac{1}{\bar{u}}$  ou du produit  $Yik_-$ , car on a  $\partial_-^2 \bar{u} - \partial_-^2 \hat{\bar{u}} = O(u)$  qui n'est pas très petit, et car  $Y$  a une singularité en log. Or la contribution concernée par ce problème est

$$\begin{aligned} & \int \int_{A_X \setminus A_{\lambda_0}} uZ \frac{1}{\partial_- \bar{u}^2 - k_-^2} \frac{\partial_-^2 \bar{u}}{\bar{u}} f dt dt' - \int \int_{A_X \setminus A_{\lambda_0}} \hat{u}\hat{Z} \frac{1}{\partial_- \hat{\bar{u}}^2 - k_-^2} \frac{\partial_-^2 \hat{\bar{u}}}{\hat{\bar{u}}} \hat{f} dt dt' \\ & - \int \int_{A_X \setminus A_{\lambda_0}} uZ \frac{2\partial_- \bar{u} \partial_-^2 \bar{u}}{(\partial_- \bar{u}^2 - k_-^2)^2} \frac{\partial_- \bar{u}}{\bar{u}} f dt dt' + \int \int_{A_X \setminus A_{\lambda_0}} \hat{u}\hat{Z} \frac{2\partial_- \hat{\bar{u}} \partial_-^2 \hat{\bar{u}}}{(\partial_- \hat{\bar{u}}^2 - k_-^2)^2} \frac{\partial_- \hat{\bar{u}}}{\hat{\bar{u}}} \hat{f} dt dt' \\ & - \int \int_{A_X \setminus A_{\lambda_0}} Y \frac{2\partial_-^2 \bar{u} \partial_- \bar{u}}{(\partial_- \bar{u}^2 - k_-^2)^2} ik_- f dt dt' + \int \int_{A_X \setminus A_{\lambda_0}} \hat{Y} \frac{2\partial_-^2 \hat{\bar{u}} \partial_- \hat{\bar{u}}}{(\partial_- \hat{\bar{u}}^2 - k_-^2)^2} ik_- \hat{f} dt dt'. \end{aligned}$$

Cette grosse expression se décompose en une forme similaire à (77h); En utilisant la même technique que pour (77h), et en utilisant  $\frac{\partial_- \bar{u}^2}{\bar{u}} - \frac{\partial_- \hat{\bar{u}}^2}{\hat{\bar{u}}} = O(u^2)$  et  $\partial_- \bar{u} \partial_-^2 \bar{u} ik_- - \partial_- \hat{\bar{u}} \partial_-^2 \hat{\bar{u}} ik_- = O(uk_-)$ , on obtient une majoration en

$$\begin{aligned} & O\left(\|f\|_{L^\infty} \left( \frac{\lambda_0^2}{d_-} + \frac{\lambda_0^3}{d_-^2} + \frac{\lambda_0^2}{d_-} + \frac{\lambda_0^3}{d_-} + \frac{\lambda_0^2}{d_-^2} + \frac{\lambda_0^3}{d_-^3} + \frac{\lambda_0^2}{d_-^2} + \frac{\lambda_0^3}{d_-^2} \right. \right. \\ & \left. \left. + \frac{\lambda_0^3}{d_-^2} k_- + \frac{\lambda_0^3}{d_-^3} k_- + \frac{\lambda_0^2}{d_-^2} k_- \right) \leq O(\lambda_0^2/d_- + \lambda_0^2 k_-/d_-^2) \end{aligned}$$

où nous avons utilisé  $\lambda_0 < 1$ ,  $k_- > 1$  et  $d_- > 1$ .

**Le reste du terme (77c)–(77c).**

Les contributions que nous n'avons pas abordées dans (77c)–(77c) sont

$$\begin{aligned} & - \int \int_{A_X \setminus A_{\lambda_0}} \begin{bmatrix} Y \\ uZ \end{bmatrix}^T \partial_- \left( \frac{1}{\partial_- \bar{u}^2 - k_-^2} \begin{bmatrix} 0 & -\bar{u} \partial_- \bar{u} \\ 0 & ik_- \end{bmatrix} \right) \begin{bmatrix} 0 \\ g \end{bmatrix} dt dt' + \hat{\phantom{}} \\ & + \int \int_{A_X \setminus A_{\lambda_0}} \begin{bmatrix} Y \\ uZ \end{bmatrix}^T \frac{1}{\partial_- \bar{u}^2 - k_-^2} \begin{bmatrix} 0 & -\bar{u} \partial_- \bar{u} \\ 0 & ik_- \end{bmatrix} \begin{bmatrix} 0 \\ \partial_- g \end{bmatrix} dt dt' - \hat{\phantom{}} \end{aligned}$$

où les deux termes  $\hat{\phantom{}}$  signalent l'analogie des intégrales précédentes mais pour  $\hat{S}$  au lieu de  $S$ . Ici, il n'est plus possible de faire jouer la symétrie pour obtenir des compensations comme dans (77h), parce que  $g - \hat{g}$  n'a pas de propriété particulière. On utilise alors les majorations (29d) et (29e), ... ce

qui donne

$$\begin{aligned} & \mathcal{O}\left(\|g\|_\infty\left(\frac{\lambda_0^2}{d_-^2} + \frac{\lambda_0}{d_-} + \frac{\lambda_0^2}{d_-^2}k_-\right) + \|g\|_{\mathbb{H}^1(c_S)}\left(\frac{\sqrt{\lambda_0^3}}{d_-} + \frac{\sqrt{\lambda_0^3}}{d_-}k_-\right)\right) \\ & \leq \mathcal{O}\left(\|g\|_{\mathbb{H}^1(c_S)}\left(\frac{\lambda_0}{d_-} + \frac{\sqrt{\lambda_0^3}}{d_-}k_-\right)\right) \end{aligned}$$

**Limite nulle du terme (77d)–(77d̂) si  $g = 0$ .**

Raisonnons d'abord pour  $X \in (0, \lambda_0)$ . Ensuite, nous ferons tendre  $X$  vers 0. Prouvons-le d'abord pour  $g$  nul.

Lorsque  $g = 0$ , la différence (77d)–(77d̂) admet elle aussi une décomposition « massacre » de la forme (77h). On en conclut qu'on a un terme en

$$\mathcal{O}\left(\|f\|_{L^\infty}\left(\frac{X^2}{d_-}k_- + \frac{\log \frac{X}{\lambda_0} X^2}{d_-^2}k_- + \frac{X^3 \log \frac{X}{\lambda_0}}{\lambda_0 d_-} + \frac{\lambda_0 X}{d_-^2} + \frac{\lambda_0 X}{d_-}\right)\right)$$

qui tend bien vers 0 lorsque  $X$  tend vers 0 : donc pour  $X$  suffisamment petit, l'énoncé est vérifié.

**Limite nulle du terme (77d) si  $f = 0$ .**

Lorsque  $f = 0$ , le terme (77d) se reformule en

$$(77d) = 2 \int_{t'=t \pm X} \begin{bmatrix} Y \\ uZ \end{bmatrix}^T \frac{1}{\partial_- \bar{u}^2 - k_-^2} \begin{bmatrix} 0 & -\bar{u} \partial_- \bar{u} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ g \end{bmatrix} dt \quad (77i)$$

$$\pm 2 \int_{t'=t \pm X} \begin{bmatrix} Y \\ uZ \end{bmatrix}^T \frac{1}{\partial_- \bar{u}^2 - k_-^2} \begin{bmatrix} 0 & 0 \\ 0 & ik_- \end{bmatrix} \begin{bmatrix} 0 \\ g \end{bmatrix} dt \quad (77j)$$

Nous pouvons jouer sur les symétries dans (77j) : on a  $Z(t, t') - Z(t', t) = \mathcal{O}(\bar{u}(t, t')Z(t, t')\frac{k_-}{\lambda_0})$ , et  $g(t, t') - g(t', t) = \mathcal{O}(f(t, t')u(t, t'))$  car  $g(t, t') = b_1(t)b_2(t')v(t, t')$  avec  $v$  qui est  $C^p$  d'après le lemme 4.3.2 ; de plus,  $u(t', t) = u(t, t')$  et  $\partial_- \bar{u}(t', t) = \partial_- \bar{u}(t, t')$ . Ainsi,

$$\begin{aligned} (77j) &= \mathcal{O}\left(2\frac{k_-^2}{\lambda_0} \int_{t'=t+X} |\bar{u}uZ| \frac{1}{|\partial_- \bar{u}^2 - k_-^2|} |g| dt'\right) \\ &+ \mathcal{O}\left(2k_- \int_{t'=t+X} |uZ| \frac{1}{|\partial_- \bar{u}^2 - k_-^2|} u|f| dt\right) \\ &= \|f\|_\infty \mathcal{O}\left(\frac{k_-^2 X}{\lambda_0 d_-} + \frac{X k_-}{d_-}\right) \end{aligned}$$

qui tend bien vers 0 lorsque  $X$  tend vers 0. Le terme (77i) tend aussi vers 0 lorsque  $X$  tend vers 0, car une estimation donne directement (77i) =  $O(\|g\|_{L^\infty}(\frac{X \log \frac{X}{\lambda_0}}{d_-}))$ .

### Faire tendre $X$ vers 0.

Dans la décomposition (77b)...(77f) de (73), les trois premiers termes dépendent de  $X$  ; pour mieux les contrôler, nous leur avons soustrait les termes correspondants de la décomposition (77b)...(77f) de (77a) : et ce contrôle est effectif, car nous avons montré page 166 que (77b)–(77b) tendait vers 0, et nous pouvons montrer avec les trois paragraphes précédents que (77d)–(77d) tend aussi vers 0 lorsque  $X$  tend vers 0 ; enfin, en regroupant les estimations que nous avons prouvées dans les pages 167 à 169, nous savons que (77c)–(77c) se majore comme prédit par l'énoncé de façon indépendante de  $X$ .

Ainsi, nous pouvons choisir  $X$  tel que dans la décomposition de (73)–(77a) analogue à la décomposition (77b)...(77f), les trois premiers termes se majorent comme prédit par l'énoncé de façon indépendante de  $X$ .

### La suite est identique.

Les expressions (77e)–(77e) et (77f)–(77f) se traitent de la même manière que les expressions (77c)–(77c) et (77d)–(77d). Afin d'abrèger, nous nous contentons de citer les cas les plus difficiles.

### Un terme $\frac{1}{\bar{u}}$ dérivé deux fois dans (77e).

Si les deux dérivations  $\partial_-$  dans (77e) sont appliquées au même terme  $\frac{1}{\bar{u}}$ , on a pour  $f = 0$  la contribution

$$\begin{aligned} & \lambda_0 \int \int_{A_{\lambda_0}} \begin{bmatrix} Y \\ uZ \end{bmatrix}^T \frac{1}{(\partial_- \bar{u}^2 - k_-^2)^2} \begin{bmatrix} -\bar{u}\partial_- \bar{u} & 0 \\ 2ik_- & -\bar{u}\partial_- \bar{u} \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} \left[ 2\left(\frac{\partial_- \bar{u}}{\bar{u}}\right)^3 - \frac{(\partial_-^2 \bar{u})\partial_- \bar{u}}{\bar{u}^2} \right] dt dt' - \hat{\imath} \\ &= O\left(\lambda_0^{3/2} \int \int_{A_{\lambda_0}} \frac{1}{\bar{u}} \begin{bmatrix} \sqrt{u} \\ u^{3/2} \end{bmatrix}^T \begin{bmatrix} u & 0 \\ k_- & u \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{1}{u^3} dt dt' / d_-^2\right) + \hat{\imath} \\ &= O\left(\lambda_0^{3/2} \int \int_{A_{\lambda_0}} u^{-3/2} dt dt' / d_-^2\right) \\ &= O(\lambda_0 / d_-^2) \end{aligned}$$

et pour  $f \neq 0$  et  $g = 0$ , on utilise les compensations du type  $Y - \hat{Y} = O(u^2)$  et  $\partial_- \bar{u} - \partial_- \hat{\bar{u}} = O(u)$  dans notre morceau de (77c)–(77c), ce qui donne  $O(\lambda_0 k_- / d_-^3)$ .

**Termes  $\frac{1}{u}$  non dérivés dans (77e).**

Si aucun des termes en  $\frac{1}{u}$  n'est dérivé, et comme tout le reste est dérivable sans peine c'est comme si rien n'était dérivé, on a une contribution en

$$\begin{aligned}
& O\left(\lambda_0^{3/2}(\|b_1\|_{W^{2,1}} + \|b_2\|_{W^{2,1}}) \sup_{(t,t') \in A_{\lambda_0}} \frac{1}{u} \left[ \frac{\sqrt{u}}{u^{3/2}} \right]^T \begin{bmatrix} |k_-| & u \\ \frac{1}{u} & |k_-| \end{bmatrix} \begin{bmatrix} |k_-| & u \\ \frac{1}{u} & |k_-| \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} / d_-^2\right) + \widehat{() } \\
&= O\left(\lambda_0^{3/2}(\|b_1\|_{W^{2,1}} + \|b_2\|_{W^{2,1}}) \sup_{(t,t') \in A_{\lambda_0}} \left(\frac{1}{u} \left[ \frac{\sqrt{u}}{u^{3/2}} \right]^T \begin{bmatrix} k_-^2 & |k_-|u \\ \frac{|k_-|}{u} & k_-^2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)^2 dt dt' / d_-^2\right) \\
&= O\left(\lambda_0^{3/2}(\|b_1\|_{W^{2,1}} + \|b_2\|_{W^{2,1}}) \sup_{(t,t') \in A_{\lambda_0}} [k_-^2(u^{-\frac{1}{2}} + u^{1/2}) + |k_-|(u^{1/2} + u^{-\frac{1}{2}})] dt dt' / d_-^2\right) \\
&= O\left((\|b_1\|_{W^{2,1}} + \|b_2\|_{W^{2,1}}) \lambda_0 k_-^2 / d_-^2\right).
\end{aligned}$$

Et grâce à des compensations du type  $Y - \widehat{Y} = O(\sqrt{\lambda_0 u^3})$ , on peut même obtenir  $O\left((\|b_1\|_{W^{2,1}} + \|b_2\|_{W^{2,1}}) \lambda_0^{3/2} k_-^2 / d_-^2 + (\|b_1\|_{W^1} + \|b_2\|_{H^1(c_S)}) \lambda_0 k_0 / d_-^2\right)$  en notant que ces compensations sont incompatibles avec  $\partial_-^3 \widehat{u}$  et avec  $g$ .

**Terme  $\frac{1}{u}$  dérivé dans (77f).**

Si  $f = 0$ , le même calcul donne

$$\begin{aligned}
& O\left(\lambda_0 \begin{bmatrix} 1 \\ \lambda_0 \end{bmatrix}^T \frac{1}{\lambda_0^2} \begin{bmatrix} \lambda_0 & 0 \\ k_- & \lambda_0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} / d_-^2\right) + \widehat{() } \\
&= O(\lambda_0 / d_-^2)
\end{aligned}$$

et pour  $f \neq 0$  et  $g = 0$ , on utilise les compensations du type  $Y - \widehat{Y} = O(u^2)$  et  $\partial_- \bar{u} - \partial_- \widehat{u} = O(u^2)$ , ce qui donne  $O(\lambda_0 k_- / d_-^3)$ . ■

Nous venons de prouver deux propriétés qui prédisent des décroissances lorsque respectivement  $|k_+|$  et  $|k_-|$  sont grands. Or nous avons pu nous assurer que  $\begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$  est loin de l'image de  $\frac{1}{2\pi} \nabla u$ , c'est à dire, d'après la propriété 4.2, que  $k_1^2 + k_2^2 > 9$  par exemple. Ceci produirait une dérivation similaire à  $D$  sauf que son dénominateur a des dérivées qui explosent.

Voici un premier exemple, simple, où nous pouvons utiliser ces propriétés.

Supposons que l'espace de discrétisation  $V_h$  est non pas celui des ondelettes ou celui de Malvar, mais juste celui des exponentielles  $e_\omega$  définies en (70), pour  $-\frac{L}{2h} < \omega \leq \frac{L}{2h}$ . Dans ce cas, on peut facilement prouver (60) et (61). Le nombre de degrés de liberté  $N$  par longueur d'onde s'écrit

$$N = \frac{\lambda_0}{h}. \quad (78)$$

**Théorème 4.5** *Quel que soit le diffractant pourvu qu'il n'ait pas de coin, il existe  $N$  (78) tel que pour tout  $\lambda_0 < \frac{1}{2}$ , la discrétisation à l'échelle  $h = \frac{\lambda_0}{N\|A^{-1}\|}$  du problème (21h) dans l'espace  $V_h$  engendré par les éléments  $(e_\omega)_{-\frac{L}{2h} < \omega \leq \frac{L}{2h}}$  de la base de Fourier (70) vérifie (60) et (61), indépendamment de  $\lambda_0$ , ce qui fait que le théorème 3.1 montre la stabilité de la discrétisation du problème (21h) dans  $V_h$  indépendamment de  $\lambda_0$ . Rappelons que  $\|A^{-1}\|$  est peut-être en  $O(1)$  (voir partie 4.1).*

■ Rappelons que  $\frac{L}{2h}$  est entier. Il suffit de raisonner pour  $\lambda_0 < \mu$ , avec  $\mu$  une constante au choix. En effet, la démonstration de l'existence de  $N$  convenant à  $\lambda_0 \in [\frac{1}{2}, \mu]$  est facile, vu que les théorèmes 2.27 et 2.28 montrent qu'une inégalité du type (65a) suffira à prouver ce théorème uniformément en  $\lambda_0 \in [\frac{1}{2}, \mu]$ ; et une inégalité du type (65a) se montre facilement avec la caractérisation en Fourier de l'espace  $H^1(\mathbb{R}/L\mathbb{Z})$ .

Soit  $f \in V_h$  de norme  $L^2(\partial S)$  valant 1 : on a donc

$$\sum_{-\frac{L}{2h} < \omega \leq \frac{L}{2h}} |\hat{f}(\omega)|^2 = L \quad (79a)$$

et

$$\hat{f}(\omega) = 0$$

si  $\omega \leq -\frac{L}{2h}$  ou  $\omega \geq \frac{L}{2h}$  (condition qu'on écrira  $|\omega - \frac{1}{2}| > \frac{L}{2h}$  ci-dessous, car  $\frac{L}{2h}$  est entier) et où  $\hat{f}$  est la transformée de Fourier  $L$ -périodique de  $f$ .

Écrivons maintenant (58) :

$$\begin{aligned} & \|(\text{Id} - P_h)(A - \text{Id})P_h f\|_{L^2} \\ &= \sum_{|\omega' - \frac{1}{2}| > \frac{L}{2h}} |\widehat{Af}(\omega') - \hat{f}(\omega')|^2 \\ &= \sum_{|\omega' - \frac{1}{2}| > \frac{L}{2h}} \left| \sum_{-\frac{L}{2h} < \omega \leq \frac{L}{2h}} \int \int e_{-\omega'}(t') K(t', t) e_\omega(t) dt' dt \hat{f}(\omega) \right|^2 \end{aligned} \quad (79b)$$

où la dernière somme est en fait la norme  $\ell^2$  d'un vecteur résultat d'un produit entre, d'une part, la matrice de terme général

$$\int \int e_{-\omega'}(t') K(t', t) e_\omega(t) dt dt' \quad (79c)$$

avec l'indice de ligne  $|\omega' - \frac{1}{2}| > \frac{L}{2h}$  et l'indice de colonne  $-\frac{L}{2h} < \omega \leq \frac{L}{2h}$ , et, d'autre part, le vecteur de terme général  $\hat{f}(\omega)$  qui a une norme  $\ell^2$  valant  $\sqrt{L}$  d'après (79a).

La dernière somme est alors plus petite que le produit de la norme  $\ell^2$  du vecteur par la norme d'opérateur  $\ell^2 \rightarrow \ell^2$  de la matrice (79c). Or les deux propriétés ci-dessus nous permettent de jauger la norme d'opérateur  $\ell^1 \rightarrow \ell^1$  et la norme d'opérateur  $\ell^\infty \rightarrow \ell^\infty$  de la matrice (79a).

Dans la propriété précédente 4.4, on a  $b_1 = 1$  et  $b_2 = 1$ ; leurs transformées de Fourier  $L$ -périodiques sont donc à support réduit à un point, et nous voyons dans la démonstration que le premier terme du  $O()$  est nul car  $k_-$  ne peut pas être nul d'après (79c), donc cette propriété nous donne une majoration du terme général (79c) en

$$\frac{2\pi\lambda_0^2|\frac{\omega+\omega'}{L}|}{(2\pi\lambda_0|\frac{\omega+\omega'}{L}|)^2 - 16\pi^2} L^{\frac{1}{2}} \quad (79d)$$

où le  $d_-$  de cette propriété 4.4 a été estimé à l'aide de la propriété 4.2.

La propriété pénultième 4.3 donne elle-même une majoration de (79c) en

$$|L \log \lambda_0|^{1/2} \lambda_0^{p-1/2} \left( \frac{1 + 2\pi\lambda_0|\frac{\omega-\omega'}{L}|}{(2\pi\lambda_0|\frac{\omega-\omega'}{L}|)^2 - 16\pi^2} \right)^p \quad (79e)$$

où le  $d_+$  de cette propriété 4.3 a été estimé à l'aide de la propriété 4.2. Pour l'exemple, supposons que  $p$  vaille 2.

Nous avons deux majorations (79d) et (79e), qu'il s'agit de combiner pour jauger la norme  $\ell^\infty \rightarrow \ell^\infty$  et la norme  $\ell^1 \rightarrow \ell^1$  de la matrice de terme général (79c).

Pour simplifier les calculs, nous commençons à exiger sur  $N$  (78) que

$$(2\pi\lambda_0|\frac{\omega+\omega'}{L}|)^2 - 16\pi^2 \text{ ou } (2\pi\lambda_0|\frac{\omega-\omega'}{L}|)^2 - 16\pi^2$$

soit dominé par son facteur contenant  $\omega$  et  $\omega'$  lorsqu'ils vérifient  $|\omega'| \geq N\frac{L}{2\lambda_0}$  et  $|\omega| \leq N\frac{L}{2\lambda_0}$ . Il suffit pour cela que

$$N \geq 6.$$

Dans ce cas, il y a grosso modo égalité entre (79d) et (79e) lorsque

$$|\omega - \omega'|^2 = \sqrt{|\log \lambda_0| L^3 / \lambda_0} |\omega + \omega'|$$

(on rappelle que  $p = 2$ ) ce qui, lorsque  $|\omega'| \geq N\frac{L}{2\lambda_0}$  et  $|\omega| \leq N\frac{L}{2\lambda_0}$ , est de la même efficacité subjective que

$$|\omega' - \omega| = c := \max\left(\frac{2L}{\lambda_0}, \left[\sqrt{\frac{|\log \lambda_0| L^3}{\lambda_0}}\right]^{1/2} \left[\frac{NL}{\lambda_0}\right]^{1/2}\right) = \max\left(\frac{2L}{\lambda_0}, \sqrt[4]{\frac{N^2 L^5 |\log \lambda_0|}{\lambda_0^3}}\right). \quad (79f)$$

Nous n'utiliserons donc (79d) que lorsque  $|\omega' - \omega| \leq c$ , et (79e) que lorsque  $|\omega' - \omega| > c$ .

Soit  $\mu > 0$  tel que pour tout  $\lambda_0 < \mu$ , on ait  $c < \frac{NL}{2\lambda_0}$ .

Alors,  $|\omega' - \omega| \leq c$  entraînerait

$$|\omega' + \omega| \geq 2|\omega'| - |\omega - \omega'| \geq |\omega'| + N\frac{L}{2\lambda_0} - c \geq |\omega'|. \quad (79g)$$

La matrice de terme général (79c) a grâce à (79d) et (79e) une norme  $\ell^\infty \rightarrow \ell^\infty$  en

$$O\left(\frac{\lambda_0 L^{1/2}}{2\pi \frac{\lambda_0}{L} \inf |\omega'|} 2c\right) \leq O\left(\sqrt[4]{\frac{\lambda_0 |\log \lambda_0|}{N^2}}\right)$$

plus

$$O\left(|L \log \lambda_0|^{1/2} \lambda_0^{3/2} \left(\frac{L}{2\pi \lambda_0}\right)^2 2 \int_c^\infty \frac{dx}{x^2}\right) \leq O\left(\sqrt[4]{\frac{\lambda_0 |\log \lambda_0|}{N^2}}\right).$$

La norme  $\ell^1 \rightarrow \ell^1$  admet elle aussi exactement les deux mêmes estimations ci-dessus.

Nous allons faire tendre  $N$  vers l'infini pour montrer qu'un certain  $N$  assez grand satisfait l'énoncé. La somme de ces deux expressions tend vers 0 uniformément en  $\lambda_0$  lorsque  $N$  tend vers l'infini : ainsi, par interpolation, la matrice de terme général (79c) a une norme  $\ell^2 \rightarrow \ell^2$  qui tend vers 0 uniformément en  $\lambda_0$  lorsque  $N \rightarrow \infty$ .

Nous avons prouvé (60). Les propriétés 4.3 et 4.4 n'exploitaient pas la définition de  $v$  fait en (75b) : ainsi, elles sont tout aussi bien valables pour le noyau adjoint  $K^*$  de l'opérateur adjoint  $A^*$  de Fredholm de seconde espèce défini de façon analogue à (23). Ainsi, (61) se démontre avec la même méthode. ■

Cette démonstration prouve aussi que lorsqu'on impose  $N = 10$  (ou  $8 + \varepsilon$ ), pour tout diffractant, il existe  $\mu$  petit tel que pour toute longueur d'onde  $\lambda_0 < \mu$ , la discrétisation du problème (21h) dans  $V_h$  est stable avec  $h$  donné par (78).

#### 4.4 Les ondelettes discrétisent correctement.

Nous allons prouver l'analogie du théorème 4.5 pour les ondelettes : on y obtiendra

$$h = \frac{1}{N} \lambda_0^{1 + \frac{1}{n_g}} \text{ et non } h = \frac{1}{N} \lambda_0,$$

où  $n_g$  est le nombre de moments nuls de l'ondelette.

**Propriété 4.6** Lorsque  $L$  est un multiple de  $h$ , et lorsque  $\omega \in \mathbb{Z}$ , on a

$$P_h(e_\omega)(t) = \overline{\hat{\phi}(2\pi \frac{h}{L}\omega)} \sum_{\omega' \in \omega + \frac{h}{L}\mathbb{Z}} \hat{\phi}(2\pi \frac{h}{L}\omega') e_{\omega'}(t) \quad (80)$$

et

$$\|P_h(e_\omega)\|_{L^2(\partial S)} = |\hat{\phi}(2\pi \frac{h}{L}\omega)| L^{1/2} \quad (81)$$

où  $P_h$  est le projecteur sur l'espace  $V_h$  défini en (65b), où l'exponentielle complexe  $e_\omega \in L^2(\partial S)$  est la fonction définie par  $e^{i\frac{2\pi}{L}\omega t}$  en (70), et où la fonction  $\phi$  (la fonction d'échelle mère) est la seule fonction  $C^\infty$  à support compact telle que, pour tout  $j$  tel que  $L2^j$  est plus grand que le diamètre de son support, on a

$$\psi_{j,m}^{(0)}(t) = 2^{j/2} \sum_{t' \in 2^j t - m + 2^j LZ} \phi(t') \quad (82a)$$

c'est à dire que  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  est une version déperiodisée de la fonction  $\psi_{0,0}^{(0)} : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}$ . La convention choisie pour la transformée de Fourier  $\hat{\phi}$  a été précisée en (67d).

■ La définition de  $V_h$  impose déjà que  $L$  soit entier, et que  $\frac{1}{h}$  est une puissance de deux négative. Posons donc  $h = 2^{-j}$  avec  $j$  pour lequel on a

$$\psi_{j,0}^{(0)}(t) = 2^{j/2} \sum_{t' \in 2^j t + 2^j LZ} \phi(t').$$

La fonction

$$e_c^{(j)} = L^{-1/2} 2^{-j/2} \sum_{l=1}^{L2^j} e_c(l2^{-j}) \psi_{j,l}^{(0)}$$

paramétrée par  $c \in \{1, 2, \dots, 2^j\}$  est le terme général d'une famille de  $V_{2^{-j}}$ ; c'est une base orthonormée, car la matrice de passage la couplant à la base canonique constituée des  $\psi_{j,*}^{(0)}$  est une matrice de Vandermonde de terme général  $L^{-1/2} 2^{-j/2} e_c(l2^{-j})$  qui est unitaire.

Donc le projecteur orthogonal sur  $V_{2^{-j}}$  appliqué à la fonction  $e_\omega$  s'écrit comme

$$P_{2^{-j}}(e_\omega)(t) = \sum_{c=1}^{L2^j} e_c^{(j)} \int_0^L e_\omega(t) \overline{e_c^{(j)}(t)} dt \quad (82b)$$



Ceci mène à ce calcul d'un coefficient de Fourier de la fonction périodique  $e_c^{(j)}$  :

$$\int_0^L \overline{e_c^{(j)}} e_\omega = \int_0^L \overline{e_c^{(j)}(t)} e_\omega(t) dt = L^{-1/2} 2^{-j/2} \int_0^L \sum_{l=1}^{L2^j} e_c(-l2^{-j}) \psi_{j,l}^{(0)}(t) e_\omega(t) dt$$

Lorsque  $c \notin \omega + L2^j\mathbb{Z}$ , le changement de variable  $t = t' + 2^{-j}$  dans la dernière intégrale montre que son produit contre  $1 - e_1(2^{-j})$  est nul : ainsi, dans la somme de (82b) seul le terme correspondant à  $c \in \omega + L2^j\mathbb{Z}$  est non nul. Découpons la dernière intégrale, ce qui donne

$$\int_0^L \overline{e_c^{(j)}} e_\omega = L^{-1/2} 2^{-j/2} \sum_{m=1}^{L2^j} \int_{\frac{m-1}{2^j}}^{\frac{m}{2^j}} \sum_{l=1}^{L2^j} \psi_{j,l}^{(0)}(t) e_c(-l2^{-j}) e_\omega(t) dt$$

Comme  $c \in \omega + L2^j\mathbb{Z}$ , le changement de variable  $t = t' + (m-1)2^{-j}$  montre que tous les termes de la somme  $\sum_{m=1}^{L2^j}$  sont égaux ; comme de plus les deux exponentielles se combinent, on a

$$\int_0^L \overline{e_c^{(j)}} e_\omega = L^{-1/2} 2^{-j/2} L2^j \int_0^{\frac{1}{2^j}} \sum_{l=1}^{L2^j} \psi_{j,l}^{(0)}(t) e_\omega(t - l2^{-j}) dt$$

Avec le changement de variable  $t' = 2^j t$  et la définition de  $\phi$ , on a

$$\begin{aligned} \int_0^L \overline{e_c^{(j)}} e_\omega &= L^{-1/2} 2^{-j/2} L2^j \int_0^1 \sum_{l=1}^{L2^j} 2^{j/2} \sum_{t'' \in t' - l + 2^j L\mathbb{Z}} \phi(t'') e_\omega(2^{-j}(t' - l)) 2^{-j} dt' \\ &= L^{1/2} \int_0^1 \sum_{l=1}^{L2^j} \sum_{t'' \in t' - l + 2^j L\mathbb{Z}} \phi(t'') e_\omega(2^{-j}t'') dt' \\ &= L^{1/2} \int_0^1 \sum_{l \in \mathbb{Z}} \phi(t' - l) e_\omega(2^{-j}(t' - l)) dt' \\ &= L^{1/2} \int_{\mathbb{R}} \phi(t) e_\omega(2^{-j}t) dt \end{aligned}$$

où nous avons utilisé le changement de variable  $t' = t + l$ .

Ainsi, nous avons montré que  $\int_0^L \overline{e_c^{(j)}} e_\omega$  est nul lorsque  $c \notin \omega + L2^j\mathbb{Z}$ , et que sinon il vaut  $L^{1/2} \hat{\phi}(2\pi \frac{\omega}{L2^j})$ . Ce résultat d'abord combiné avec (82b) nous donne

$$P_{2^{-j}}(e_\omega) = \overline{\hat{\phi}(2\pi \frac{\omega}{L2^j})} L^{1/2} e_c^{(j)}$$

pour  $\omega \in c + L2^j\mathbb{Z}$ , ce qui prouve (81) ; le même résultat combiné avec la formule d'inversion  $L$ -périodique de Fourier nous donne

$$e_c^{(j)} = L^{-1/2} \sum_{\omega' \in c + L2^j\mathbb{Z}} \hat{\phi}(2\pi \frac{\omega'}{L2^j}) e_{\omega'} \quad (82c)$$

Les deux dernières égalités prouvent finalement (80). ■

**Propriété 4.7** *Si  $n_g$  est le nombre de moments nuls de  $\psi_{j,0}^{(1)}$ , c'est à dire si  $m \mapsto g_m^{(1)}$  est orthogonal à  $\mathbb{R}_{n_g-1}[m]$ , alors on a*

$$\|(Id - P_h)(e_\omega)\|_{L^2} \leq C\left(\frac{h}{L}\omega\right)^{n_g}$$

■ La notation  $\|\cdot\|$  abrègera la norme  $\|\cdot\|_{L^2(\partial S)}$ . Nous rappelons que  $L$  est le périmètre du diffractant  $S$ .

Voici l'idée :  $\|P_h(e_\omega)\|$  est inférieur à  $\|e_\omega\| = L^{1/2}$  car  $P_h$  est un projecteur orthogonal. Or le terme correspondant à  $\omega' = \omega$  dans la somme de (80) sera de norme très proche de  $L^{1/2}$  (aussi bien que  $g^{(0)}(\pi \frac{L}{h}\omega)\|e_\omega\|$  pour  $\xi$  très petit) : donc les autres termes de cette somme seront petits.

La fonction  $\phi$  est de support compact, donc sa transformée de Fourier  $\hat{\phi}$  est  $C^\infty$ , et donc nous avons le choix entre deux possibilités, selon que 0 est un zéro d'ordre fini ou non de  $1 - |\hat{\phi}|^2$  :

- Soit  $1 - |\hat{\phi}(\xi)|^2$  admet un équivalent du type  $C\xi^{n'}$  pour un certain entier  $n'$  ;
- Soit  $1 - |\hat{\phi}(\xi)|^2$  est un  $o(\xi^{n'})$  pour tout entier  $n'$ .

L'équation d'échelle (65f) va nous permettre d'éliminer la seconde possibilité, et même de donner  $n'$ . Cette équation d'échelle donne avec la notation  $\phi$  vue en (82a) :

$$\sum_{m=-m_0}^{m=m_1} g_m^{(0)} \phi(x - m) = \phi\left(\frac{x}{2}\right)$$

Ce qui en Fourier devient

$$2\hat{g}(\xi)\hat{\phi}(\xi) = 2\hat{\phi}(2\xi) \tag{82d}$$

où la fonction  $\hat{g}$ , qui tient lieu de transformée de Fourier  $2\pi$ -périodique de  $g_m^{(0)}$ , se définit par

$$\hat{g}(\xi) = \frac{1}{2} \sum_{m=-m_0}^{m=m_1} g_m^{(0)} e^{im\xi}. \tag{82e}$$

Kahane-Lemarié [KL95] montrent que l'identité d'orthogonalité (65e) se traduit ici par

$$|\hat{g}(\xi)|^2 + |\hat{g}(\xi + \pi)|^2 = 1 \tag{82f}$$

et que l'orthogonalité de  $g_m^{(1)}$  à  $\mathbb{R}_{n-1}[X]$  se traduit par

$$\hat{g}(\xi + \pi) \sim C_g(|\xi|^n) \quad (82g)$$

avec une certaine constante  $C_g$ . Or on a, selon (82d) :

$$[|\hat{\phi}(2\xi)|^2 - |\hat{\phi}(\xi)|^2] = [|\hat{g}(\xi)|^2 - 1]|\hat{\phi}(\xi)|^2$$

La fonction  $\hat{\phi}$  est  $C^\infty$ , et Kahane-Lemarié [KL95] donne  $\hat{\phi}(0) = \int \phi = \pm 1$ , donc les deux crochets ci-dessus sont équivalents ; dans l'alternative ci-dessus sur la finitude du zéro de  $\hat{\phi}$ , seule la première hypothèse est valide ; (82f) et (82g) montre alors que

$$n' = 2n_g. \quad (82h)$$

Dans (80), un des termes de la somme, celui pour  $\omega' = \omega$ , a une norme valant  $|\hat{\phi}(2\pi\frac{h}{L}\omega)|^2 L$  ; le théorème de Pythagore montre alors que les autres termes (orthogonaux au terme  $\omega' = \omega$ ) de cette somme sont petits :

$$\begin{aligned} \overline{\|\hat{\phi}(2\pi\frac{h}{L}\omega)\|} \sum_{\omega' \in \omega + \frac{L}{h}\mathbb{Z}^*} \|\hat{\phi}(2\pi\frac{h}{L}\omega')e_{\omega'}\|^2 &= \|P_h e_\omega\|^2 - \|\hat{\phi}(2\pi\frac{h}{L}\omega)|^2 e_\omega\|^2 \\ &\leq (1 - |\hat{\phi}(2\pi\frac{h}{L}\omega)|^4)L \end{aligned}$$

En développant  $\|(\text{Id} - P_h)e_\omega\|$  avec (80), le théorème de Pythagore puis l'inégalité ci-dessus donnent :

$$\begin{aligned} \|(\text{Id} - P_h)e_\omega\|^2 &\leq \|(1 - |\hat{\phi}(2\pi\frac{h}{L}\omega)|^2)e_\omega\|^2 + (1 - |\hat{\phi}(2\pi\frac{h}{L}\omega)|^4)L \\ &= 2(1 - |\hat{\phi}(2\pi\frac{h}{L}\omega)|^2)L \end{aligned}$$

En utilisant (82h), le résultat vient naturellement. ■

**Théorème 4.8** *On suppose qu'on a un diffractant  $\mathbb{C}^{p+3}$  avec  $p \geq 1$  (donc sans coin). On suppose que la norme de l'inverse de l'opérateur  $A$  du problème (22) est bornée indépendamment de  $\lambda_0$  (voir la partie 4.1). On suppose que  $\hat{\phi}$  est en  $O((1 + \xi)^{-\frac{1}{p} - \varepsilon})$ , en plus de l'hypothèse déjà vue  $1 - |\hat{\phi}(\xi)|^2 = O(\xi^{2n_g})$ .*

*Alors il existe  $N$  tel que pour tout  $\lambda_0 < \frac{1}{2}$  la discrétisation du problème (21h) dans l'espace  $V_h$  engendré par les ondelettes est stable, pourvu que*

$$h = \frac{1}{N\|A^{-1}\|} \lambda_0^{1 + \frac{1}{n_g}} |\log \lambda_0|. \quad (83)$$

En améliorant le théorème 2.27 comme indiqué sous son énoncé, il devrait être possible de remplacer l'exposant  $\frac{1}{n_g}$  par  $\frac{1}{2n_g}$ .

■ Cette démonstration est similaire à la démonstration du théorème 4.5, sauf que nous devons insérer les propriétés 4.7 et 4.6.

Soit  $f \in V_h$  tel que  $\|f\| = 1$ , où par  $\|\cdot\|$  nous entendons  $\|\cdot\|_{L^2(\partial L)}$ . Pour montrer (60), nous devons montrer qu'il existe  $N$  tel que, dès que  $h$  satisfait (83), le nombre

$$\|(\text{Id} - P_h)(A - \text{Id})P_h f\|$$

est borné par  $\frac{1}{2}\|A^{-1}\|$ .

Soit  $\lambda$  le produit de  $\lambda_0$  par un nombre entre  $\frac{1}{2}$  et 1, de sorte à ce que  $\frac{\log \lambda}{\log 2} \in \mathbb{Z}$  avec  $\frac{1}{2}\lambda_0 < \lambda < \lambda_0$ .

Voici un premier découpage basé sur l'inégalité triangulaire, avec la base en Fourier (70), et où  $M$  est un entier à préciser ( $\frac{\lambda}{h}$  et  $\frac{L}{\lambda}$  sont supposés être entiers eux aussi) :

$$\begin{aligned} & L\|(\text{Id} - P_h)(A - \text{Id})P_h f\| \\ &= L\|(\text{Id} - P_h)(A - \text{Id})f\| \\ &= \left\| \sum_{\omega' \in \mathbb{Z}} (\text{Id} - P_h)e_{\omega'} \int e_{-\omega'}(t)(A - \text{Id})f(t)dt \right\| \\ &\leq \left\| \sum_{|\omega' - \frac{1}{2}| < M \frac{L}{2\lambda}} (\text{Id} - P_h)e_{\omega'} \int e_{-\omega'}(t)(A - \text{Id})f(t)dt \right\| \quad (84a) \end{aligned}$$

$$+ \left\| \sum_{|\omega' - \frac{1}{2}| > M \frac{L}{2\lambda}} e_{\omega'} \int e_{-\omega'}(t)(A - \text{Id})f(t)dt \right\| \quad (84b)$$

où nous avons utilisé que  $\|\text{Id} - P_h\|_{L^2(\partial S) \rightarrow L^2(\partial S)}$  vaut 1 comme c'est la norme d'un projecteur orthogonal.

### Pour les petites fréquences, une sorte d'inégalité de Jackson cachée.

Avant d'utiliser la propriété 4.7 pour (84a), nous devons transformer cette dernière en remarquant grâce à (80) que les éléments de la somme contenue dans (84a) sont orthogonaux entre eux à condition que

$$M \leq \frac{\lambda}{h}.$$

Le théorème de Pythagore donne alors

$$(84a)^2 = \sum_{|\omega' - \frac{1}{2}| < M \frac{L}{\lambda}} \|(\text{Id} - P_h)e_{\omega'}\|^2 \left| \int e_{-\omega'}(t)(A - \text{Id})f(t)dt \right|^2$$

$$\begin{aligned}
&\leq \sum_{|\omega' - \frac{1}{2}| < M \frac{L}{\lambda}} C(M \frac{h}{L} \omega')^{2n_g} \left| \int e_{-\omega'}(t)(A - \text{Id})f(t)dt \right|^2 \\
&\leq C(M \frac{h}{\lambda})^{2n_g} \|(A - \text{Id})f\|^2 \\
&\leq C(\frac{\lambda}{hM})^{-2n_g} \lambda^{-1} |\log \lambda|^2 \|f\|^2
\end{aligned}$$

où nous avons utilisé les théorèmes 2.27 (améliorable) et 2.28. Ainsi, le nombre

$$\left(\frac{\lambda}{hM}\right)^{-2n_g} \frac{1}{\lambda}$$

doit être petit indépendamment de  $\lambda$  (et donc de  $\lambda_0$ ): c'est ce qui sera à l'origine de l'exposant  $1 + \frac{1}{2n_g}$  dans (83).

### Un produit matrice-vecteur pour (84b).

Une inégalité triangulaire et une égalité de Parseval donnent

$$\begin{aligned}
(84b) &\leq \frac{1}{L} \sum_{Q \in \mathbb{Z}} \left\| \sum_{|\omega' - \frac{1}{2}| > M \frac{L}{2\lambda}} \sum_{|\omega - Q \frac{L}{h} - \frac{1}{2}| > Q \frac{L}{2h}} e_{\omega'} \int e_{-\omega'}(t)(A - \text{Id})e_{\omega} dt \hat{f}(\omega) \right\| \\
&= \sum_{Q \in \mathbb{Z}} \left( \sum_{|\omega' - \frac{1}{2}| > M \frac{L}{2\lambda}} \left| \sum_{|\omega - Q \frac{L}{h} - \frac{1}{2}| > Q \frac{L}{2h}} \int e_{-\omega'}(t)(A - \text{Id})e_{\omega} dt \right|^2 |\hat{f}(\omega)|^2 \right)^{1/2} (84c)
\end{aligned}$$

Le dernier terme à la puissance 1/2 ci-dessus est en fait la norme  $\ell^2$  d'un produit matrice-vecteur avec d'une part la matrice de terme général (79c) avec l'indice de ligne  $\omega'$  et l'indice de colonne  $\omega$  vérifiant  $|\omega' - \frac{1}{2}| > M \frac{L}{2\lambda}$  et  $|\omega - Q \frac{L}{h} - \frac{1}{2}| > Q \frac{L}{2h}$ , et avec d'autre part le vecteur de terme général  $\hat{f}(\omega)$ . Comme  $\omega'$  vérifie ici aussi  $|\omega' - \frac{1}{2}| > M \frac{L}{2\lambda}$ , on peut utiliser la même technique que pour le produit matrice-vecteur qui était contenu dans (79b).

### Majorer le terme général de la matrice.

Le terme général de (79c) admet ici aussi le majorant (79e); La propriété 4.4 donne également le majorant

$$\frac{F_{\omega - \omega'}}{2\pi \lambda_0 \left| \frac{\omega + \omega'}{L} \right| (|\omega - \omega'|^2 + 1)} + (79d) \quad (84d)$$

avec  $F_m$  qui est une suite indicée par  $m \in \mathbb{Z}$  vérifiant

$$\sum_{m \in \mathbb{Z}} |F_m|^2 = 1.$$

Nous utiliserons le majorant (84d) lorsque

$$|\omega - \omega'| < c_Q := \max\left(\frac{3L}{\lambda}, \sqrt[p]{L \sqrt{\frac{|\log \lambda|}{\lambda}} \left(\frac{|Q|}{h} + \frac{M}{\lambda}\right)}\right)$$

et nous utiliserons le majorant (79e) lorsque  $|\omega - \omega'| \geq c_Q$ .

Il suffit de raisonner pour  $\lambda < \mu$ , avec  $\mu$  une constante au choix, pour les mêmes raisons que dans la démonstration du théorème 4.5. Nous en profitons pour demander sur  $\mu$  que  $c_0 \leq \frac{ML}{2\lambda}$  dès que  $\lambda < \mu$ , ce qui est possible lorsque

$$M \geq 6.$$

De même que pour (79g), cette contrainte sur  $\mu$  fait que lorsque  $|\omega - \omega'| < c_Q$  et  $Q = 0$  on a

$$|\omega' + \omega| \geq 2|\omega'| - |\omega - \omega'| \geq M \frac{L}{\lambda} - c_0 \geq \frac{ML}{2\lambda}$$

et, lorsque  $|\omega - \omega'| < c_Q$  et  $|Q| \geq 1$  et en supposant  $M \geq 2$  et  $\frac{\lambda}{Mh} \geq \left(\frac{\lambda}{h}\right)^{\frac{1}{p}}$  :

$$\begin{aligned} |\omega' + \omega| &\geq 2|\omega| - |\omega - \omega'| \geq (2|Q| - 1) \frac{L}{h} - c_Q \\ &\geq \left(\frac{7}{4}|Q| - \frac{5}{4}\right) \frac{L}{h} + \frac{\lambda}{Mh} \left(\frac{|Q|}{2} + \frac{1}{2}\right) \frac{ML}{2\lambda} - \left(\frac{\lambda}{h}\right)^{\frac{1}{p}} \left(\frac{|Q|\lambda}{Mh} + 1\right)^{\frac{1}{p}} c_0 \\ &\geq \left(\frac{7}{4}|Q| - \frac{5}{4}\right) \frac{L}{h} \geq \frac{1}{8}|Q| \frac{L}{h} + \frac{h}{4\lambda M} |Q| M \frac{L}{2\lambda} \geq \frac{1}{8}|Q| \frac{L}{h} + M \frac{L}{4\lambda}. \end{aligned}$$

On remarque après coup que l'hypothèse  $|Q| \geq 1$  était inutile, ce qui donne finalement

$$|\omega - \omega'| < c_Q \quad \Rightarrow \quad |\omega' + \omega| \geq \frac{1}{8}|Q| \frac{L}{h} + M \frac{L}{4\lambda}.$$

### Norme d'opérateur de la matrice.

Ainsi, la norme  $\ell^1 \rightarrow \ell^1$  de la matrice, qui est égale à la plus grande norme  $\ell^1$  d'une colonne de la matrice, est, en utilisant (79d), (79e) puis la différence (84d)–(79d), la somme de termes en

$$\begin{aligned} O\left(\lambda \frac{\frac{\lambda}{8h}|Q| + \frac{M}{4}}{\left(\frac{\lambda}{8h}|Q| + \frac{M}{4}\right)^2 - 4} c_Q\right) &= O\left(\frac{\max\left(\frac{3L}{\lambda}, \sqrt[p]{L \sqrt{\frac{|\log \lambda|}{\lambda}} \left(\frac{|Q|}{h} + \frac{M}{\lambda}\right)}\right)}{\frac{|Q|}{h} + \frac{M}{\lambda}}\right) \\ &= O\left(\max\left(\frac{1}{|Q|\frac{\lambda}{h} + M}, \frac{|\lambda \log \lambda|^{\frac{1}{2p}}}{\left(\frac{|Q|\lambda}{h} + M\right)^{1-\frac{1}{p}}}\right)\right) \end{aligned}$$

plus

$$O\left(\sqrt{\lambda^{2p-1}L|\log\lambda|} \int_{c_Q}^{\infty} \left(\frac{L}{2\pi\lambda}\left(x + \frac{L}{2\pi\lambda}\right)\right)^p dx\right) = \begin{cases} O\left(\frac{|\lambda \log \lambda|^{\frac{1}{2p}}}{\left(\frac{|Q|\lambda}{h} + M\right)^{1-\frac{1}{p}}}\right) & \text{si } |Q| > \frac{h}{\lambda^{p-\frac{1}{2}}} \\ O\left(\sqrt{\lambda^{2p-3}|\log\lambda|}\right) & \text{sinon} \end{cases}$$

plus

$$O\left(\frac{\left(\sum_{\omega' \in \mathbb{Z}} |F_{\omega-\omega'}|^2\right)^{1/2}}{\frac{\lambda}{8h}|Q| + \frac{M}{4}} \left(\sum_{\omega' \in \mathbb{Z}} \frac{1}{(|\omega - \omega'|^2 + 1)}\right)^{1/2}\right) = O\left(\frac{1}{M + \frac{\lambda}{h}|Q|}\right)$$

où on a utilisé  $M \geq 8$ ,  $p = 2$ ,  $|\log \lambda| \lambda < 1$  et  $M \geq 2$  et  $c_Q \geq \frac{2L}{\lambda}$ ,  $c_Q = \frac{3L}{\lambda} \iff |Q| \sim \frac{h}{\lambda} \lambda^{3/2-p}$ , et l'inégalité de Cauchy-Schwarz. La norme  $\ell^\infty \rightarrow \ell^\infty$  de la matrice s'estime de la même manière, ce qui par interpolation donne cette estimation pour la norme  $\ell^2 \rightarrow \ell^2$ .

### Norme du vecteur.

Nous allons utiliser le résultat intermédiaire (82c). Comme  $e_m^{(j)}$  était une base orthonormée de  $V_h$  avec  $h = 2^j$ , l'élément  $f \in V_h$  s'écrit sous la forme

$$f = \sum_{m=1}^{\frac{L}{h}} a_m e_m^{(j)}$$

et alors (82c) donne

$$1 = \|f\|^2 = \sum_{m=1}^{\frac{L}{h}} |a_m|^2 = \sum_{|\omega - \frac{QL}{h} - \frac{1}{2}| < \frac{L}{2h}} \frac{|\hat{f}(\omega)|^2}{|\hat{\phi}(2\pi \frac{h\omega}{L})|^2} \quad (84e)$$

ce qui permet d'évaluer la norme de notre vecteur de terme général  $\hat{f}(\omega)$  avec l'indice  $\omega$  satisfaisant  $|\omega - \frac{QL}{h} - \frac{1}{2}| < \frac{L}{2h}$ . Grâce à la propriété 4.9,  $\hat{\phi}$  est en  $O((1 + \xi)^{-\frac{1}{p}-\varepsilon})$ , on voit alors que la norme  $\ell^2$  de ce vecteur est en

$$O\left(\left(\frac{1}{2} + |Q|\right)^{-\sigma_\phi}\right). \quad (84f)$$

avec  $\sigma_\phi > \frac{1}{p}$ .

**Produit de ces deux normes.**

Nous réinjectons ces résultats, ce qui donne :

$$\begin{aligned}
L(84b) &= (84c) \\
&= O\left(\sum_{Q \in \mathbb{Z}} \frac{|\lambda \log \lambda|^{\frac{1}{2p}}}{\left(\frac{|Q|\lambda}{h} + M\right)^{1-\frac{1}{p}}} \left(\frac{1}{2} + |Q|\right)^{-\sigma_\phi}\right) \\
&+ O\left(\sum_{\substack{|Q| < \frac{h}{\lambda^{p-\frac{3}{2}}}} \sqrt{\lambda^{2p-3} |\log \lambda|} \left(\frac{1}{2} + |Q|\right)^{-\sigma_\phi}\right) \\
&+ O\left(\sum_{Q \in \mathbb{Z}} \frac{1}{M + \frac{\lambda}{h}|Q|} \left(\frac{1}{2} + |Q|\right)^{-\sigma_\phi}\right) \\
&= O\left(|\lambda \log \lambda|^{\frac{1}{2p}} \min\left(\frac{\lambda}{h}, M\right)^{\frac{1}{p}-1}\right) + O\left(\sqrt{|\log \lambda|} \lambda^{\sigma_\phi(p-\frac{3}{2})} \left(\frac{h}{\lambda}\right)^{1-\sigma_\phi}\right) \\
&\quad + O\left(\min\left(\frac{\lambda}{h}, M\right)^{-1}\right) \\
&= O\left(M^{\frac{1}{p}-1}\right) + O\left(\lambda^{\sigma_\phi(p-\frac{3}{2})} \left(\frac{h}{\lambda}\right)^{1-\sigma_\phi}\right) + O\left(M^{-1}\right)
\end{aligned}$$

qui exige  $\sigma_\phi > \frac{1}{p}$  et  $\sigma_\phi < 1$  (cette dernière restriction étant toujours possible).

**Au bilan.**

Choisissons des entiers  $M$  et  $\frac{\lambda}{h}$  tels que la dernière expression ci-dessus est plus petite que  $\frac{1}{2} \inf_{\lambda_0 < \frac{1}{2}} \|A^{-1}\|^{-1}$ , et qui satisfassent aussi toutes les conditions exigées en chemin.

On a donc trouvé  $\mu > 0$  tel que pour tout  $\lambda_0 < 2\mu$ ,

$$\|(\text{Id} - P_h)(A - \text{Id})P_h\|_{L^2 \rightarrow L^2} = O\left(\left(\frac{\lambda_0}{hM}\right)^{-n_g} \lambda_0^{-1/2} + h \left|\frac{\log \lambda_0}{\lambda_0}\right|^{1/2} + \sqrt[4]{\frac{|\log \lambda_0| \lambda_0}{M^2}}\right).$$

Cela marche tout aussi bien pour le dual  $c_A^*$ , car les propriétés 4.3 et 4.4 se démontrent de la même façon lorsqu'on remplace le noyau  $K$  par le noyau adjoint  $K^*$  (ce qui change la définition de  $v$  faite en (75b)). Ainsi, (61) se démontre avec la même méthode. ■

**Propriété 4.9** *Si  $\sigma_\phi > 1$ , on a*

$$\hat{\phi}(\xi) \leq C(1 + \xi)^{-\sigma_\phi} \tag{84g}$$

*si  $\phi \in C^{\sigma_\phi-1}$  et si  $\phi$  est à support compact.*

La démonstration a été par exemple faite par Kahane-Lemarié [KL95], chapitre 6.



## 4.5 Malvar

Nous avons, pour les bases de Malvar, un théorème analogue au théorème 4.8. La démonstration est proposée dans [G05]. Cette démonstration est plus facile que la démonstration ci-dessus, à cause du découplage naturel entre la phase et l'amplitude des fonctions de la base de Malvar (67g). Ainsi, on peut appliquer les propriétés 4.4 et 4.3, et aussi 5.2. La démonstration est aussi plus longue car la base de Malvar est indexée par deux paramètres au lieu d'un.

## 5 Étude de matrices issues des systèmes

Lorsque la famille ordonnée  $\Lambda_h = \{w_1, w_2, w_3, \dots, w_{L/h}\}$  est une base du sous-espace vectoriel  $V_h$  choisi pour le problème numérique (24), on peut écrire ce problème numérique dans la base  $\Lambda_h$  sous la forme matricielle

$$Mu = v \quad (85)$$

où le vecteur  $u$  contient les coordonnées  $(u_m)_{1 \leq m \leq L/h}$  de la solution numérique  $J_h$  dans la base  $\Lambda_h$ , où le vecteur  $v$  contient les coordonnées du second membre  $L_\alpha$  de (24) dans la base orthogonale  $\Lambda_h$ , et où  $M$  est la matrice de taille  $\frac{L}{h} \times \frac{L}{h}$  de coefficient général

$$\int_0^L w_l(t)w_c(t)dt + \int_0^L \int_0^L K(t,t')w_l(t)w_c(t')dt \quad (86)$$

où  $l$  et  $c$  sont les numéros de ligne et de colonne, et où  $K$  (23) est le noyau de l'opérateur de Fredholm de seconde espèce  $A$  (55).

Cette matrice  $M$  est malheureusement pleine, comme nous allons le voir dans un premier exemple. Heureusement, nous pourrions dans certains cas utiliser une approximation  $\hat{M}$  de  $M$ , qui tient compte des erreurs dues au calcul par ordinateur de ses coefficients, et rendra possible l'annulation d'une proportion non négligeable de ses coefficients, ce qui accélèrera d'autant son produit avec un vecteur.

### 5.1 La justification de la discrétisation

Le théorème suivant (estimant une erreur d'interpolation) montre que si  $\hat{M}$  est suffisamment proche de  $M$ , la solution  $u = (u_1, u_2, \dots, u_{L/h})$  du système numérique  $\hat{M}u = v$  donne les coordonnées dans  $\Lambda_h$  d'un élément  $J_h \in V_h$  comme vu en partie 3.3.1 suffisamment proche de la solution exacte  $J$  du problème (22) :

**Théorème 5.1** *Considérons la matrice  $M$  du système (85) du problème discrétisé (24). Soit une autre matrice  $\hat{M}$ , qui est utilisée comme approximation de la matrice  $M$ . Reprenons les hypothèses du théorème 3.1, qui sont*

$$\|(\text{Id} - P_h)AP_h\| \leq \frac{1}{2}\|A^{-1}\|^{-1} \text{ et } \|P_hA(\text{Id} - P_h)\| \leq \frac{1}{2}\|A^{-1}\|^{-1} \quad (87)$$

où  $\|\cdot\|$  est la norme d'opérateur  $\|\cdot\|_{L^2(\partial S) \rightarrow (\partial S)}$ .

Si  $\hat{M}$  est suffisamment proche de  $M$  (si la norme d'opérateur  $\ell^2 \rightarrow \ell^2$  de la différence  $M - \hat{M}$ , qu'on note  $\varepsilon$ , est plus petite que  $\frac{1}{4}\|A^{-1}\|^{-1}$ ), alors on a

$$\|J_h - J\|_{L^2(\partial S)} \leq 3\|J - P_h J\|_{L^2(\partial S)} + 4\varepsilon\|A^{-1}\|\|J\|_{L^2(\partial S)}.$$

Lorsque de plus l'hypothèse de la propriété 3.2 est vraie, c'est à dire lorsque

$$\|(\text{Id} - P_h)(A - \text{Id})(\text{Id} - P_h)\|_{L^2(\partial S) \rightarrow L^2(\partial S)} \leq \frac{1}{2},$$

alors on a

$$\begin{aligned} \|J_h - J\|_{L^2(\partial S)} \leq & \left(4\varepsilon\|A^{-1}\| + 6\|(\text{Id} - P_h)AP_h\|_{L^2(\partial S) \rightarrow L^2(\partial S)}\right)\|J\|_{L^2(\partial S)} \\ & + 6\|P_h L_\alpha - L_\alpha\|_{L^2(\partial S)}, \end{aligned}$$

Notons que pour que prendre  $\varepsilon$  beaucoup plus petit que  $\|A^{-1}\|^{-2}$  ait un intérêt, il faut que la première inégalité de (87) ci-dessus soit vraie avec un nombre beaucoup plus petit que  $\frac{1}{2}$ .

■ Lorsque la matrice  $\hat{M}$  est utilisée au lieu de la matrice  $M$  de l'opérateur  $P_h A$  dans le système (59a), alors la solution  $J_h$  vérifie

$$P_h \hat{A} J_h = P_h L_\alpha = P_h A J$$

au lieu de (59a), où  $\hat{A}$  est l'opérateur de matrice  $\hat{M}$  dans la base  $A_h$ , que nous complétons par 0 dans l'orthogonal de l'espace  $V_h$  engendré par  $A_h$ . Alors on a  $\|P_h(A - \hat{A})P_h\| = \varepsilon$ , et l'inégalité (63a) dans la démonstration du théorème 3.1 devient

$$\begin{aligned} & \left(\frac{1}{\|A^{-1}\|} - c_h\right)\|J_h - P_h J\| \\ & \leq \|P_h A P_h (J_h - P_h J) - P_h (A - \hat{A}) P_h J_h\| \\ & \leq \|P_h A (\text{Id} - P_h) J\| + \|P_h (A - \hat{A}) P_h\| \|J_h\| \\ & \leq c_h^* \|(\text{Id} - P_h) J\| + \varepsilon (\|J_h - P_h J\| + \|P_h J - J\| + \|J\|) \end{aligned}$$

où la norme  $\|\cdot\|$  est tour à tour  $\|\cdot\|_{L^2(\partial S)}$ ,  $\|\cdot\|_{L^2(\partial S) \rightarrow L^2(\partial S)}$  ou  $\|\cdot\|_{\ell^2 \rightarrow \ell^2}$ . De cette inégalité, on tire

$$\|J_h - P_h J\|_{L^2(\partial S)} \leq \frac{(c_h^* + \varepsilon)\|J - P_h J\|_{L^2(\partial S)} + \varepsilon\|J\|_{L^2(\partial S)}}{\|A^{-1}\|^{-1} - c_h - \varepsilon}$$

La suite est une simple application de la propriété 3.2. ■

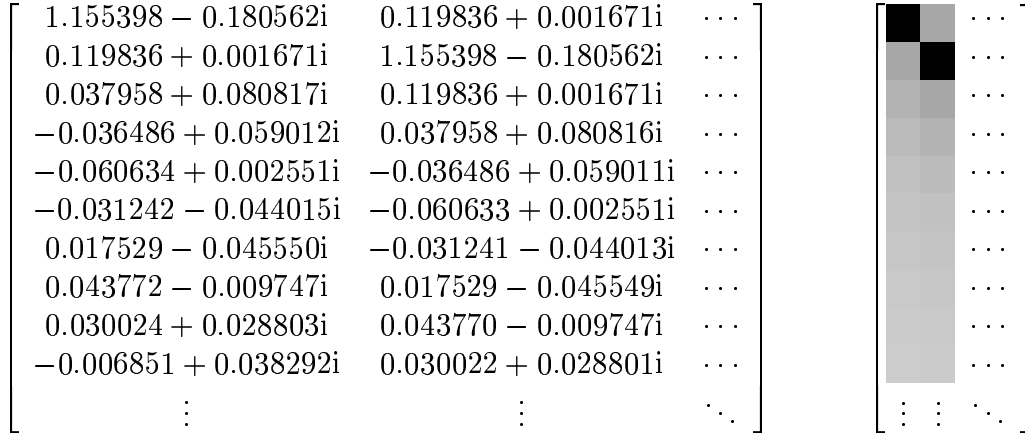


FIG. 32 – codage en gris des modules des premiers termes de la matrice de rigidité  $M$ .

### 5.1.1 Dans la base canonique des fonctions d'échelle.

Lorsque la famille  $A_h$  vaut (65b), qui est la base dite canonique pour l'espace de discrétisation  $V_h$  correspondant aux fonctions d'échelle avec  $h = 2^{-j}$ , les fonctions  $w_m$  sont définies par

$$w_m = \phi_{j,m}^{(0)}, \quad m = 1.. \frac{L}{h}.$$

La figure 32 présente un extrait des coefficients de la matrice  $M$ , lorsque le diffractant  $S$  est donné par la forme sans coin (34) qui a pour périmètre  $L = 16$ , avec la longueur d'onde  $\lambda_0 = \frac{1}{10}$ , avec le paramètre de CFIE (55) valant  $\alpha = 4$ , et avec le pas de discrétisation  $h = \frac{1}{64}$ .

Ces coefficients, comme tout résultat noté sous forme décimale dans cette partie 5, est indiqué avec toutes ses décimales significatives, en tenant compte de la précision du calcul utilisé.

Chacun de ces coefficients est le résultat d'un programme qui calcule les approximations successives

$$a_{j',0} = \int_0^L w_m(t)w_{m'}(t)dt + (h2^{-j'})^{-2} \sum_{t \in \frac{h}{2^{j'}}\mathbb{Z}} \sum_{t' \in \frac{h}{2^{j'}}\mathbb{Z}} K(t,t')w_m(t)w_{m'}(t')dt$$

pour  $j'$  valant successivement 0, 1, 2, 3 ... et qui calcule des versions accélérées de ces suites avec

$$a_{j',j''} = \frac{a_{j',j''-1} - 4^{-j''}a_{j'-1,j''-1}}{1 - 4^{-j''}}$$

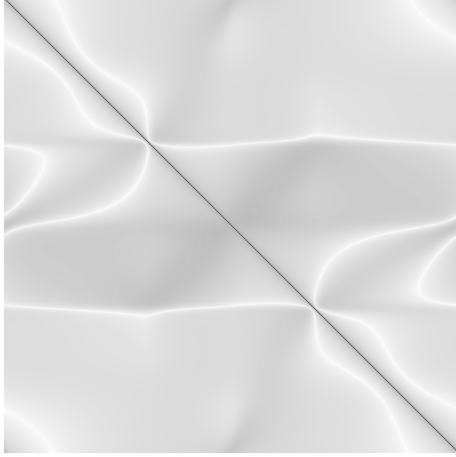


FIG. 33 – La matrice de rigidité  $M$  complète, pour la base de fonctions d'échelle (65b).

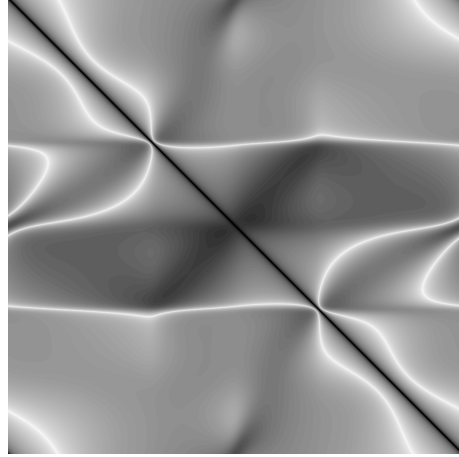
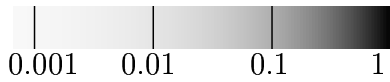


FIG. 34 – La même matrice, avec une échelle de gris différente.



pour  $j'$  et  $j''$  valant successivement 1, 2, 3, ...; lorsque  $a_{j'}$  est la valeur en  $4^{-j'}$  d'une fonction admettant un développement limité à un certain ordre en 0, alors  $a_{j',j''}$  se rapproche plus vite de la limite de  $a_{j'}$  que la suite  $a_j$  elle-même lorsque  $j' \rightarrow \infty$ , c'est le principe de l'accélération de Richardson. Le programme ayant calculé les coefficients ci-dessus par cette méthode accélérée s'arrête dès que  $|a_{j'+1,j''} - a_{j',j''}|$  est plus petit que  $10^{-6}$  pour un couple d'entier  $j'$  et  $j''$ ; expérimentalement, la précision ainsi obtenue est plus petite que  $1.5 \times 10^{-6}$ , et l'erreur moyenne est en  $0.6 \times 10^{-6}$ .

La figure 33 présente cette matrice  $M$  au complet, alors que la figure 32 n'en contenait qu'un extrait; nous avons omis les crochets. Elle contient  $1024 \times 1024$  coefficients, symbolisés chacun par un carré comme dans la figure 32; en fait, la figure 32 contient l'agrandissement du coin en haut à gauche de la figure 33; la figure 32 représente dix lignes et deux colonnes, et la figure 33 représente la matrice complète, soit  $\frac{L}{h} = 1024$  lignes et 1024 colonnes.

Pour mieux montrer la répartition des petits coefficients, nous avons ajouté la figure 34, avec une autre échelle de gris, plus foncée. Cette matrice semble être à diagonale dominante, toutefois un calcul rapide montre que les coefficients extradiagonaux, d'une taille médiane proche de  $10^{-2}$ , sont à  $1024 - 1 = 1023$  par colonnes; et que de plus, seuiller dans  $M$  les

2% coefficients les plus petits résulte en une erreur  $\|M - \hat{M}\|_{\ell^2 \rightarrow \ell^2}$  qui vaut  $\frac{1}{5}\|A^{-1}\|_{L^2 \rightarrow L^2}^{-1}$ , ce qui représente une erreur potentielle de 20% sur la solution selon le théorème 5.1.

Cette matrice foncée figure 34 montre aussi un épi-phénomène constaté par Canning [C94] : certains coefficients, qui sont plus petits que la précision  $10^{-6}$  avec laquelle ils ont été calculés, forment des courbes asymétriques et régulières, bien visibles figure 34. L'explication tient à ce que  $K(t, t')$  s'annule lorsque  $\frac{1}{\lambda_0}(s(t) - s(t'))$  est grand tout en formant l'angle (50) avec le vecteur  $s'(t)^\perp$  ; ici cet angle vaut  $\arccos -\frac{\alpha}{4\pi} \approx 109^\circ$ , et en effet l'ensemble des couples  $(l, c)$  pour lesquels le coefficient, ligne  $l$  et colonne  $c$  de la matrice figure 34, est petit, ressemble bien à

$$\left\{ (l, c) \in [1, \frac{L}{h}]^2 : s(hc) \text{ est sur l'une des deux demi-droites passant par } s(hl) \text{ et d'angle } 109^\circ \text{ avec } s(l)'^\perp \right\}$$

d'après la trajectoire de  $s$  visible en (69a) (avec  $s(0)$  à l'extrémité basse de la figure 23 et nous rappelons que  $s$  paramétrise  $\partial S$  dans le sens direct).

Canning [C94] en tirait la conclusion qu'une optimisation de la valeur de  $\alpha$ , par exemple faite par Kress [K85] et Amini [A93], n'est pas suffisante si on veut diminuer la taille des termes dans cette matrice. Canning suggère alors de combiner les éléments de la base  $A_h$  entre eux afin d'augmenter la petitesse d'un grand nombre de termes de cette matrice : c'est ce que nous ferons en parties 5.2.1 et 5.2.2.

Cette matrice est très peu compressible : seulement 0.4% des coefficients les plus petits sont facultatifs pour garantir moins de 8% d'erreur sur la solution  $J$ . Ceci s'explique par le fait que  $\|A^{-1}\|_{L^2 \rightarrow L^2}^{-1}$ , qui d'après le théorème 5.1 permet de calculer ce pourcentage, est très petit, il vaut à peu près 0.053.

### 5.1.2 Les problèmes de résonances

On voit numériquement que  $\|A^{-1}\|_{L^2 \rightarrow L^2}^{-1} \approx 0.053$ , qui est un nombre petit. Cette petitesse affecte directement les théorèmes 3.1 et 5.1, ce qui veut dire qu'il faudra d'autant plus augmenter le rapport  $\frac{\lambda_0}{h}$  afin que le théorème 4.8 fonde l'application numérique.

La petitesse de  $\|A^{-1}\|^{-1}$  est due à l'existence d'une cavité ouverte à l'intérieur du diffractant, représentée en (34) ; cette cavité est le siège de réverbérations. Nous avons pu calculer les vecteurs propres de  $A^*A$ , et en particulier le vecteur de plus petite valeur propre.

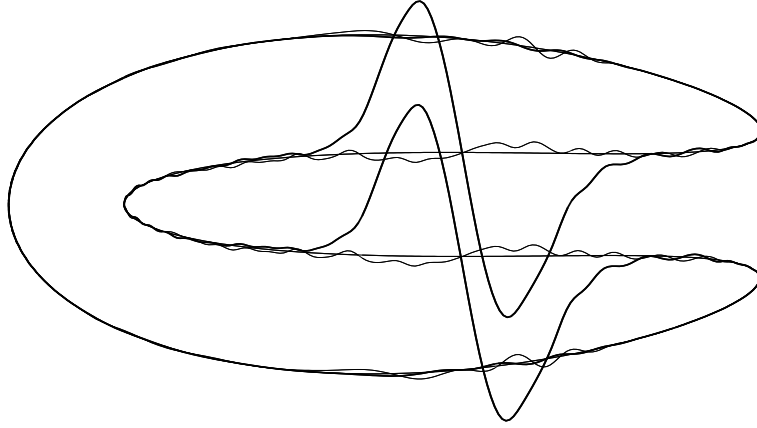


FIG. 35 – La fonction  $f$  du cas d'égalité de  $\|Af\| \geq \|A^{-1}\|^{-1}\|f\|$  lorsque  $\lambda_0 = \frac{1}{10}$ , et pour le diffractant sans coin (34) de périmètre  $L$ .

La figure 35 présente une fonction  $f$  telle que  $\|Af\| \approx \|A^{-1}\|^{-1}\|J\|$ , avec la partie réelle de  $f$  et le bord du diffractant  $S$  en gras, et la partie imaginaire de  $f$  en trait fin ; lorsque les graphes des parties réelles ou imaginaires sont très proches du bord du diffractant (qu'on voit ici de biais), cela veut dire que leurs valeurs absolues sont très proches de 0. Formellement, cette figure contient trois graphes tridimensionnels paramétrisés par  $t$ , avec les deux premières dimensions (horizontales) données par les deux coordonnées de  $s(t)$  pour les trois graphes, et la troisième dimension (verticale) donnée par respectivement 0,  $\mathcal{R}ef(t)$  et  $\mathcal{I}mf(t)$ .

La figure 36 présente (en omettant le signe) la fonction  $E = G \star f$ , avec  $G$  le noyau de Helmholtz (28). Cette fonction, ainsi qu'on le voit, est définie à la fois dans  $S$  et à l'extérieur de  $S$ . La partie droite de  $S$  est omise, car la fonction  $f$  présente la même symétrie que le diffractant  $S$ .

Nous voyons un déphasage (approximativement de  $i$ ) entre  $f$  et  $G \star f$ , qui est dû au rapport entre la longueur d'onde et la largeur de la cavité.

La fonction  $E^{inc}$  correspondant au cas où la solution  $J$  de (22) vaut  $f$  peut se lire dans la restriction de  $G \star f$  au diffractant  $S$ , car on a  $E^{inc} = -G \star f$  dans  $S$  d'après la propriété 2.20. Ainsi, nous pouvons avoir une idée du champ incident  $E^{inc}$  qu'il faut appliquer pour que  $J$  contienne autant de phénomènes de résonance que visible sur la figure 36. Pour permettre au lecteur de mieux évaluer cette fonction  $E^{inc}$ , la figure 37 est la même que la figure 36, sauf qu'une échelle logarithmique de gris montre mieux les valeurs de faible amplitude (qui n'ont pas été calculées avec la précision requise).

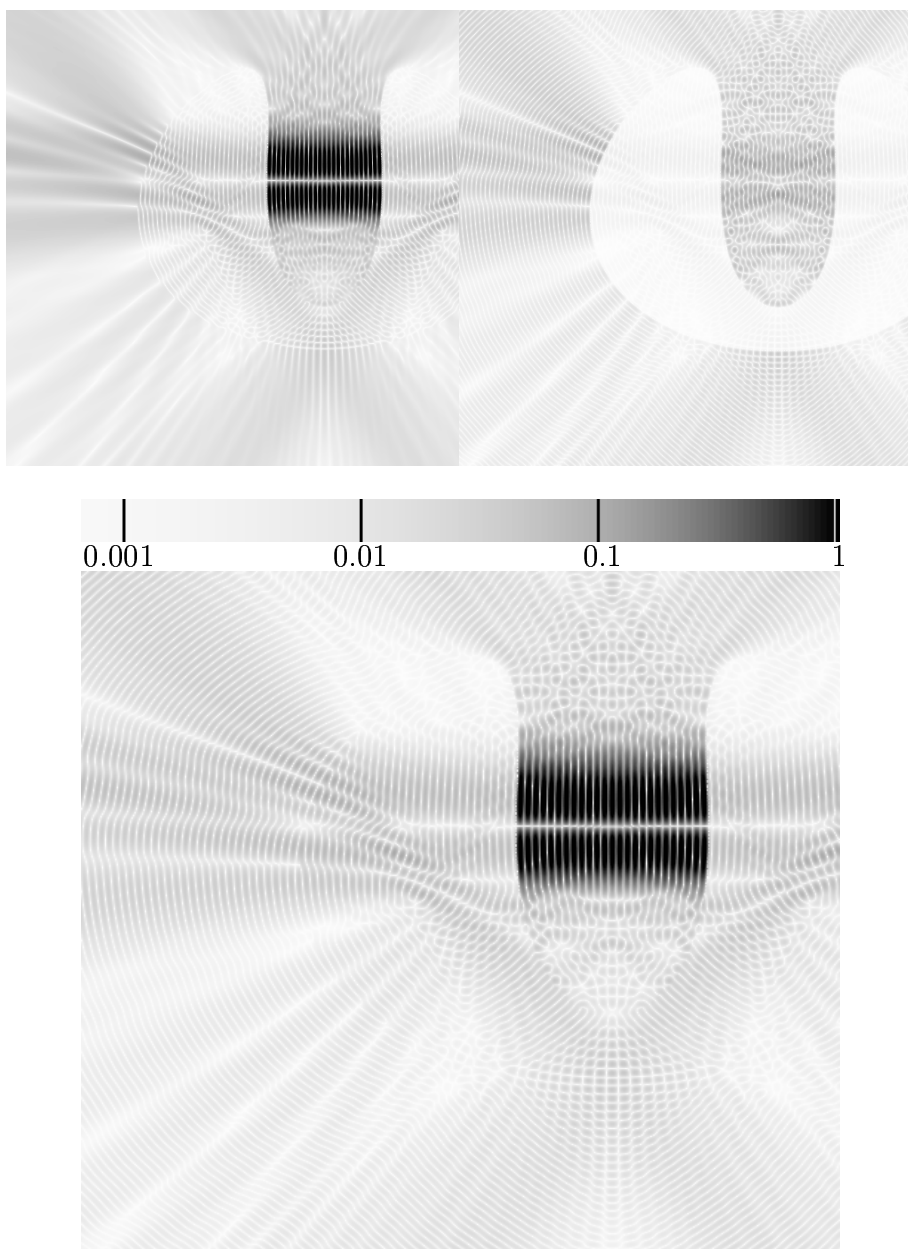


FIG. 36 – *Module, et parties réelle et imaginaire de  $E = f \star G$ .*



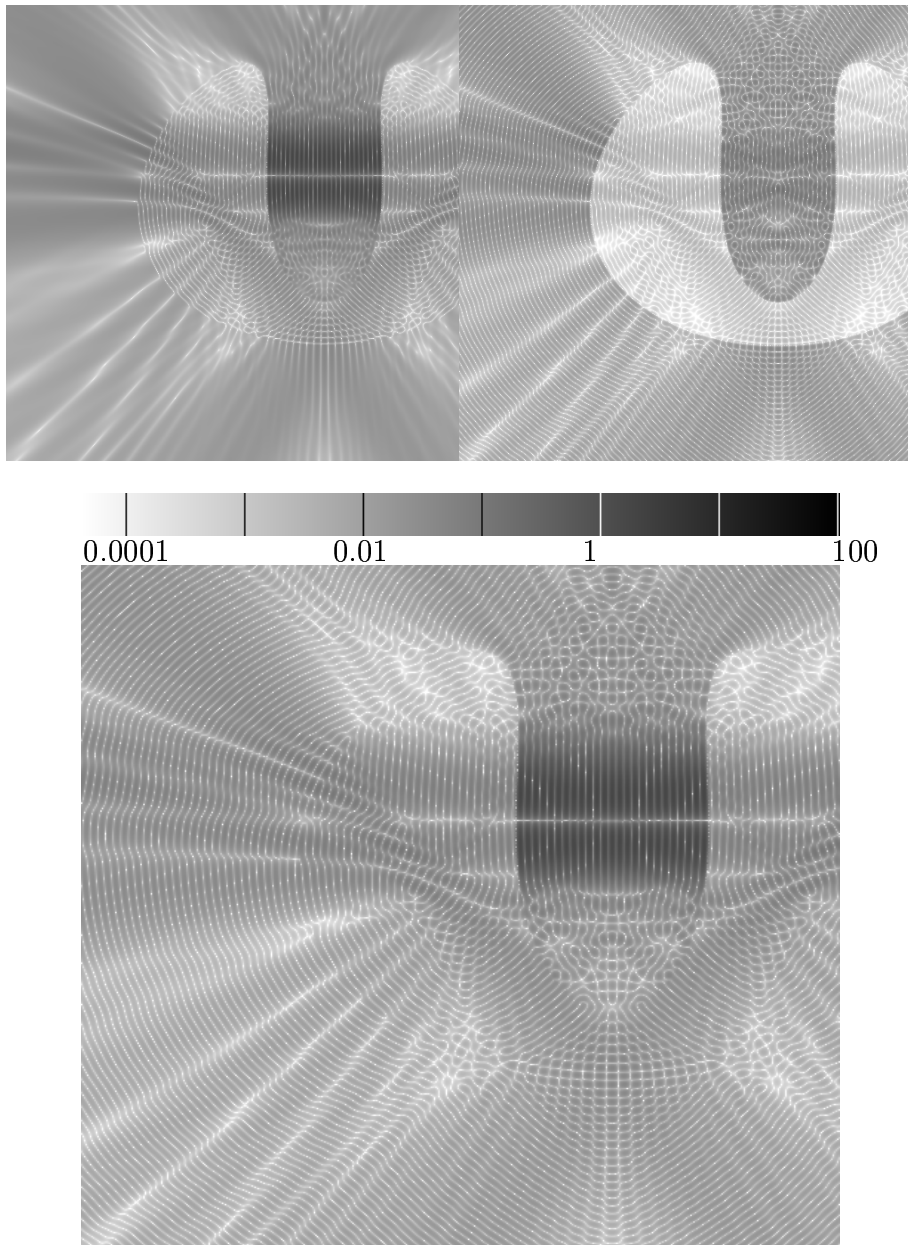


FIG. 37 – Module, et parties réelle et imaginaire de  $E = f \star G$ , en échelle de gris logarithmique.

Maintenant que nous avons vu analysé physiquement la raison pour laquelle  $\|A^{-1}\|^{-1}$  vaut seulement 0.053 à peu près, nous allons montrer l'espace  $V_h$  des ondelettes satisfait les hypothèses du théorème 5.1, même si la petitesse de  $\|A^{-1}\|^{-1}$  rend ces hypothèses d'autant plus difficiles.

### 5.1.3 Hypothèses du théorème fondant l'application numérique.

D'après la démonstration du théorème 4.8 montrant que  $h$  proportionnel à  $\lambda_0^{1+1/2n_g}$ , les deux normes

$$\|(\text{Id} - P_h)AP_h\| \text{ et } \|P_hA(\text{Id} - P_h)\|$$

intervenant dans les hypothèses du théorème 3.1 qui fonde l'application numérique sont en  $h^{1/2}$  à  $\lambda_0$  fixé. Donc on en déduit que

$$\begin{aligned} \|(P_{h/2^j} - P_{2h/2^j})AP_h\| &= \|(P_{h/2^j} - P_{2h/2^j})(\text{Id} - P_{2h/2^j})AP_{2h/2^j}P_h\| \\ &\leq \|P_{h/2^j} - P_{2h/2^j}\| \|(\text{Id} - P_{2h/2^j})AP_{2h/2^j}\| \|P_h\| = O(2^{-j/2}) \end{aligned}$$

et comme on a

$$\|(\text{Id} - P_h)AP_h\| \leq \sum_{j=1}^{\infty} \|(P_{h/2^j} - P_{2h/2^j})AP_h\|$$

on en tire l'idée qu'il suffit d'étudier les premiers termes de la dernière somme, ce qui est fait dans la table 4 (pour plus de précision, nous remplaçons les inégalités citées dans les théorèmes 3.1 et 5.1 par leur version plus précise vue dans les démonstration respectives).

On en tire le sentiment que  $\|(\text{Id} - P_h)AP_h\| \approx 0.03$  et  $\|P_hA(\text{Id} - P_h)\| \approx 0.03$  pour la discrétisation choisie ci-dessus dans la partie 5.1.1. C'est juste, mais cela suffit à satisfaire les hypothèses du théorème 3.1 lorsque  $\|A^{-1}\|^{-1} \approx 0.053$ .

Ces hypothèses nous permettent également, via le théorème 5.1, de remplacer par 0 un certain nombre de termes dans la matrice  $M$  du système (85). Si on note  $\hat{M}$  la matrice résultat de ces remplacements, la condition sur  $\hat{M}$  est que la norme de ces termes supprimés doit vérifier

$$\|\hat{M} - M\| \leq \frac{2\%}{100\%} \|A^{-1}\|^{-1} \quad (88)$$

qui garantit que l'erreur résultant  $\|J_h - J\|$  entre la solution  $J$  et notre approximation  $J_h$  est de l'ordre que 8%.

La table 4 semble montrer que  $h = \frac{1}{10}\lambda_0^{\frac{9}{8}}$  semble convenir pour une erreur globale de 3%. La table 5 montre que, dans le cas où le diffractant  $S$  présente des coins, le théorème 3.1 ne suffit pas à justifier une application numérique dans les ondelettes.

$\lambda_0$	1	1	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{20}$
$\ A^{-1}\ ^{-1}$	0.21	0.21	0.053	0.053	0.075
$\ A\ $	3.38	3.38	4.89	4.89	5.77
$h$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{128}$
$\ (P_{\frac{h}{2}} - P_h)AP_h\ $	0.027	0.005	0.015	0.007	0.015
$\ (P_{\frac{h}{4}} - P_{\frac{h}{2}})AP_h\ $	.0014	.0006	.0016		
$\ (\text{Id} - P_h)AP_h\ $	<0.03	<.006	<0.02	<0.01	<0.02
$\ P_hA(P_{\frac{h}{2}} - P_h)\ $	0.012	0.005	0.015	0.007	0.015
$\ P_hA(P_{\frac{h}{4}} - P_h)\ $	.0013	.0006	.0016		
$\ P_hA(\text{Id} - P_h)\ $	<0.02	<.006	<0.02	<0.01	<0.02
$\ (P_{\frac{h}{2}} - P_h)(A - \text{Id})(P_{\frac{h}{2}} - P_h)\ $	0.08	0.04	0.10	0.05	0.1044
$\ (P_{\frac{h}{4}} - P_{\frac{h}{2}})(A - \text{Id})(P_{\frac{h}{2}} - P_h)\ $	0.005	.0026	0.007		
$\ (P_{\frac{h}{2}} - P_h)(A - \text{Id})(P_{\frac{h}{4}} - P_{\frac{h}{2}})\ $	0.005	.0026	0.007		
$\ (P_{\frac{h}{4}} - P_{\frac{h}{2}})(A - \text{Id})(P_{\frac{h}{4}} - P_{\frac{h}{2}})\ $	0.04	.0199	0.05		
$\ (\text{Id} - P_h)(A - \text{Id})(\text{Id} - P_h)\ $	<0.15	<0.1	<0.2	<0.1	<0.2
$\ (P_{\frac{h}{2}} - P_h)L_\alpha\ /\ J\ $	0.02	.0008	0.005	$2 \cdot 10^{-6}$	0.001
$\ (P_{\frac{h}{4}} - P_{\frac{h}{2}})L_\alpha\ /\ J\ $	.0008	$10^{-5}$	$2 \cdot 10^{-6}$		
$\ (\text{Id} - P_{\frac{h}{2}})L_\alpha\ /\ J\ $	0.02	0.001	$3 \cdot 10^{-6}$	<.001	<.002
$\ P_hJ - J\ /\ J\  \leq$	0.035	0.008	0.03	0.012	0.03
$\varepsilon$ utile à partir de	0.002	.0005	.001	.0002	.0015
qui garantit $\ J - J_h\ /\ J\  \leq$	0.05	.01	.08	.02	

Table 4 – Vérification numérique que  $h = \frac{1}{8}\lambda_0^{17/16}$  convient, pour le diffractant (34) sans coin.

$\lambda_0$	1	1	1	$\frac{1}{10}$	$\frac{1}{10}$
$\ A^{-1}\ ^{-1}$	0.075	0.075	0.075	0.009	
$\ A\ $	3.45	3.45	3.45	4.99	
$h$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{64}$	$\frac{1}{128}$
$\ (P_{\frac{h}{2}} - P_h)AP_h\ $	0.27	0.25	0.23	0.36	0.27
$\ (P_{\frac{h}{4}} - P_{\frac{h}{2}})AP_h\ $	0.21			0.25	
$\ (\text{Id} - P_h)AP_h\ $					
$\ P_h A(P_{\frac{h}{2}} - P_h)\ $	0.11	0.11	0.11	0.12	0.11
$\ P_h A(P_{\frac{h}{4}} - P_h)\ $	0.055			0.054	
$\ P_h A(\text{Id} - P_h)\ $					
$\ (P_{\frac{h}{2}} - P_h)(A - \text{Id})(P_{\frac{h}{2}} - P_h)\ $	0.13	0.12	0.12	0.16	0.13
$\ (P_{\frac{h}{4}} - P_{\frac{h}{2}})(A - \text{Id})(P_{\frac{h}{2}} - P_h)\ $	0.12			0.12	
$\ (P_{\frac{h}{2}} - P_h)(A - \text{Id})(P_{\frac{h}{4}} - P_{\frac{h}{2}})\ $	0.10			0.10	
$\ (P_{\frac{h}{4}} - P_{\frac{h}{2}})(A - \text{Id})(P_{\frac{h}{4}} - P_{\frac{h}{2}})\ $	0.12			0.13	
$\ (\text{Id} - P_h)(A - \text{Id})(\text{Id} - P_h)\ $					

Table 5 – Mesures pour le diffractant (33) avec des coins, qui montrent que la discrétisation en ondelettes ne marchera pas en présence de coins.

## 5.2 Choix de la base de paquets d'ondelettes

Afin de pouvoir négliger plus que 0.4% des coefficients de cette matrice tout en garantissant une erreur de moins de 8%, nous devons diminuer drastiquement la taille de la plupart des coefficients. Or la matrice montrée par la figure 34 est dans un certain sens assez régulière, même si cette figure ne montre que le module des coefficients de cette matrice.

Les ondelettes sont connues pour leurs capacités à compresser les matrices assez régulières, c'est à dire à faire en sorte que la plupart des coefficients d'une matrice soient très petits. Pour cela nous allons considérer une autre base que (65b) pour l'espace de discrétisation  $V_h$ , ce qui permettra d'obtenir plus que 0.4% de coefficients seuillables.

### 5.2.1 Bases d'ondelettes, Paquets d'ondelettes

Nous allons donc procéder à la transformation en ondelettes de cette matrice  $M$  de la figure 34, ce qui revient à la multiplier à gauche par la matrice orthogonale  $P_{L/h}$  (65h), et à droite par sa transposée  $P_{L/h}^T$ . La figure 38 représente la matrice obtenue par la première *étape* de cette transformation, avec le même système de niveaux de gris que pour la matrice dans la base canonique (65b) vue en figure 34.

Remarquons que la matrice  $P_{L/h}$  (65h), que nous venons d'utiliser, était la matrice de passage entre, d'une part, la base (65b) choisie pour  $\Lambda_h$  dans la sous-section précédente 5.1.1 pour construire le système (85), et d'autre part, la base (65g) : ainsi, la matrice représentée par la figure 38 est *aussi* la matrice  $M$  du système (85) lorsque la base  $\Lambda_h$  est la base (65g).

Cette base se partage clairement en deux groupes, celui avec  $n = 0$  et celui avec  $n = 1$  :

$$(65g) = \{\psi_{j-1,m}^{(0)}, m \in \mathbb{Z}/L2^{j-1}\mathbb{Z}\} \cup \{\psi_{j-1,m}^{(1)}, m \in \mathbb{Z}/L2^{j-1}\mathbb{Z}\}$$

Comme nous avons procédé ci-dessus à deux multiplications, la matrice a une structure assez visible de  $2 \times 2 = 4$  blocs de taille égales, dans la figure 38.

On remarque que les lignes et colonnes relatives à des éléments en  $\psi_{j-1,*}^{(1)}$  de cette base contiennent des coefficients plus petits : ces lignes et colonnes sont relatives aux fonctions  $\psi_{j-1,*}^{(1)}$  qui sont des ondelettes ; cette petitesse, expliquée dans la section 4.3, est due à ce que leur contenu fréquentiel est au-delà de  $\frac{2\pi}{\lambda}$ , d'après la représentation graphique de  $|\hat{\psi}_{0,0}^{(1)}|$  (67c), les propriétés de la section 4.3, et l'identité

$$|\hat{\psi}_{j-1,m}^{(1)}(\xi)| = 2^{-j} |\hat{\psi}_{0,0}^{(1)}(\xi)|.$$

(la transformation de Fourier utilisée pour  $\hat{\psi}_{*,*}^{(1)}$  ci-dessus a été définie par (67d) et c'est la transformée de Fourier des fonctions  $\mathbb{R} \rightarrow \mathbb{R}$ ; donc  $\hat{\psi}_{*,*}^{(1)}$  est ici une version déperiodisée de la fonction  $L$ -périodique  $\psi_{*,*}^{(1)}$ .)

Cette première matrice de passage est un peu décevante: en effet, on ne peut enlever que 43% des coefficients pour une tolérance de 8% sur la solution. Nous allons donc faire plus qu'une étape de transformation en ondelette, ce qui changera la matrice de passage.

Nous avons vu, au total, quatre matrices de passage dans la section 3.2.2: ainsi, il existe au moins trois autres bases possibles pour choisir  $\Lambda_h$  dans le système (85). Les figures 39, 40 et 41 présentent les trois matrices résultantes. Nous avons vu à la section 3.2.2 qu'il existe de l'ordre de  $2^{\frac{L}{2h}}$  matrices de passage possibles, dont les quatre matrices de passage ne sont qu'un échantillon, ainsi il y a beaucoup d'autres figures possibles.

Une première observation est que ces trois autres matrices semblent contenir moins de gros coefficients que dans la figure 38: en effet, selon (88) il est possible d'enlever respectivement 56%, 67% et 60% coefficients en faisant une erreur sur la solution de moins 8%.

Présentons maintenant des arguments expliquant pourquoi ces pourcentages sont meilleurs que le 0.4% vu pour la matrice en base de fonctions d'échelle de la figure 34.

La première de ces trois matrices, figure 39, est la transformation en ondelettes en quatre étapes de la matrice figure 34. Son résultat est assez décevant; c'est parce que la phase dans la matrice figure 34, visible sur la figure 42, n'est pas très régulière; à la lumière de la forme (51) du noyau  $K$  présent dans l'opérateur  $A$ , nous voyons que la fonction de Hankel  $H_0^{ank}$  intervient et provoque des oscillations dans les coefficients de cette matrice qui, grâce aux équivalents (32b) et (32a) de  $H_0^{ank}$ , se situe à des périodicités valant à peu près  $\frac{\lambda_0}{h}$  qui fait 6.4 coefficients.

La seconde de ces matrices, figure 40, est la transformation en paquets d'ondelettes de cet opérateur dans une base de paquets d'ondelettes équilibrée; d'après les graphes des transformées de Fourier  $|\hat{\psi}_{0,0}^{(n)}|$  vues partie 3.2.2 des éléments de la base (67a), cette base est une espèce de transformée de Fourier à fenêtre, qui est plus adaptée à la transformation d'un phénomène oscillatoire comme celui-ci.

La troisième et dernière de ces matrices, figure 41, est la transformation en paquets d'ondelettes de cet opérateur, dans une base de paquets d'ondelettes générée au hasard dans la partie (3.2.2) juste pour exemple. Et pourtant ce choix, arbitraire, est presque aussi efficace que le plus efficace des choix ci-dessus, qui était celui de la figure 40.

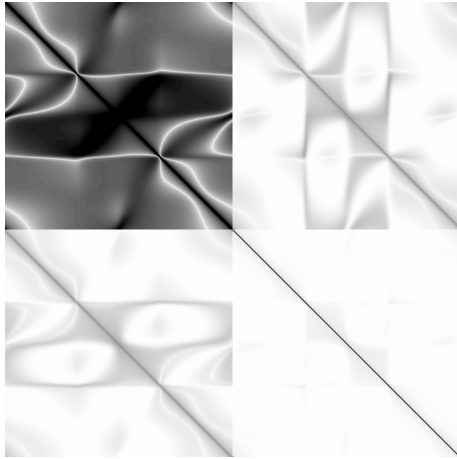


FIG. 38 – *La matrice de rigidité  $M$  pour la base d'ondelettes (65g) (avec  $j' = 1$  échelle représentée).*

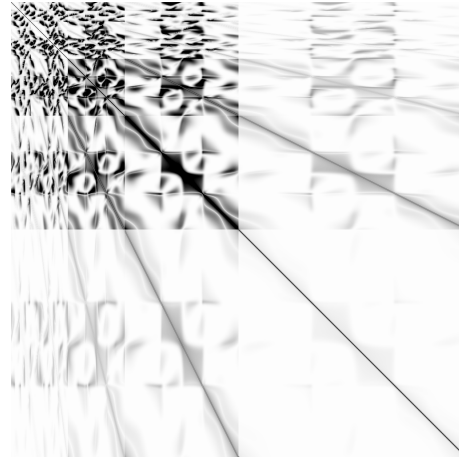


FIG. 39 – *La matrice de rigidité  $M$  pour la base d'ondelettes (65j) (avec  $j' = 4$  échelles représentées).*

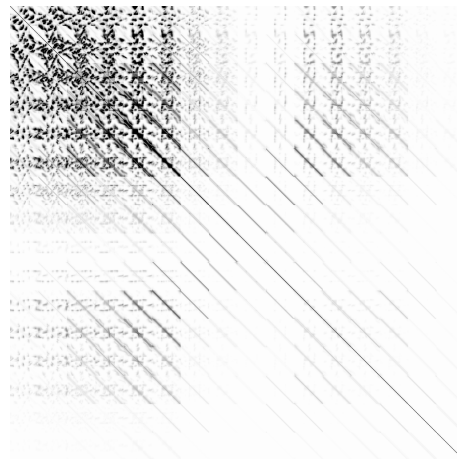


FIG. 40 – *La matrice de rigidité  $M$  pour la base de paquets d'ondelettes type Fourier (67a).*

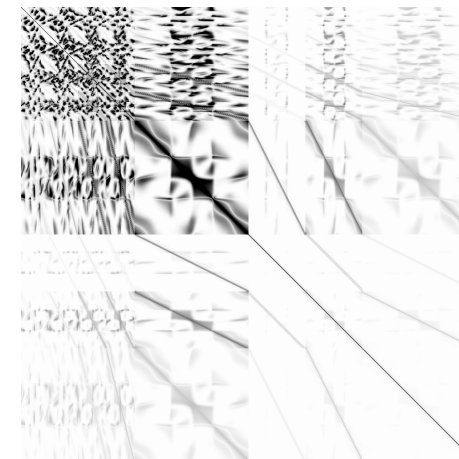


FIG. 41 – *La matrice de rigidité  $M$  pour (67f), qui est une base de paquets d'ondelettes quelconque.*

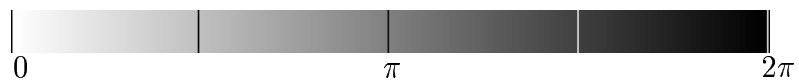
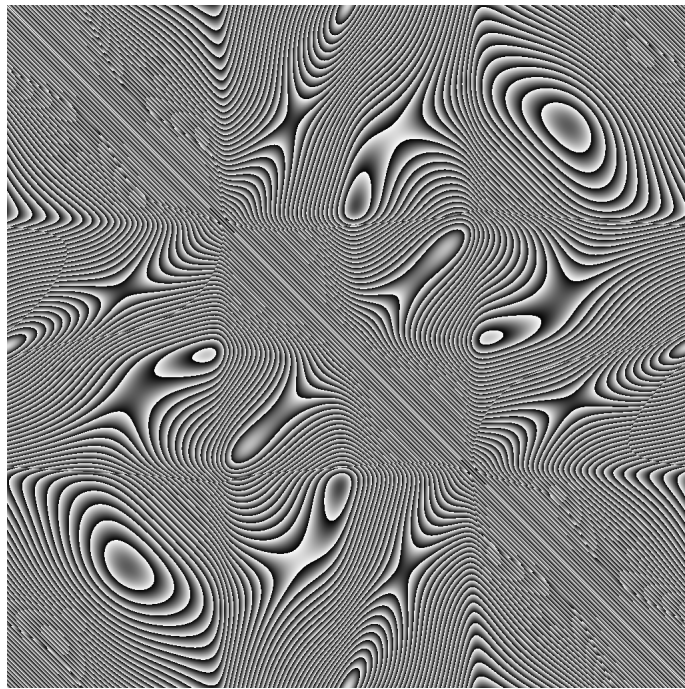


FIG. 42 – *Argument des coefficients de la matrice de rigidité  $M$  lorsque la base  $\Lambda_h$  est la base canonique de fonctions d'échelle (65b) lorsque  $S$  est le diffractant sans coin (34), et dans le cas  $\lambda_0 = \frac{1}{10}$ .*



### 5.2.2 Stratégies possibles

Le fait que la matrice de la figure 41 fasse presque aussi bien que les autres matrices, avec une base de paquets d'ondelettes pourtant choisie sans précaution aucune, signifie qu'il existe sans doute une base de paquets d'ondelettes plus efficace encore que les quatre bases utilisées dans les figures 34, 38, 39 et 40. Il existe de l'ordre de  $\approx 2^{\frac{L}{2n}}$  bases possibles, aussi le choix de la base de paquets d'ondelettes est difficile. Il existe plusieurs stratégies.

#### Stratégie Top-down.

Une première stratégie consiste à partir de la matrice de passage très simple (65h), et à y ajouter des produits dans chaque sous-blocs, progressivement, en ne validant chaque ajout que lorsqu'il augmente la part seuillable de la matrice. Ainsi, nous obtenons en résultat une matrice de passage qui ressemble à (67e) par sa dissymétrie, et qui admet la représentation plus compacte visible figure 47 avec la convention de la figure 21.

- Les noeuds les plus raffinés ne sont pas du tout évalués: donc cette stratégie n'évalue pas toutes les bases de paquets d'ondelettes possibles.
- Le coût est relativement faible: il faut calculer  $\varepsilon = \|M - \hat{M}\|$  à chaque noeud essayé.
- Le résultat dépend de l'ordre avec lequel les suppressions sont tentées.

Le coût est à peu près en  $N^{\frac{50}{h^2}}$  avec  $N$  le nombre de noeuds trouvés dans l'arbre, grâce au module ARPACK++ présenté page 145.

La matrice correspondante est représentée figure 43, et elle est seuillable à 70% pour une erreur sur la solution garantie inférieure à 8%. Nous avons toutefois deux validations de cette stratégie, par les deux stratégies suivantes.

#### La stratégie Top-Down non-standard

. La figure 44 présente une transformation dans une base de paquets d'ondelettes *bidimensionnelle* (non tensorisée), que nous n'avons pas décrite dans cette thèse. Nous avons vu dans la remarque 3.3 qu'à chaque base de paquets d'ondelettes correspond une partition finie du segment unité semi-ouvert en segments semi-ouverts de longueur *dyadique* (c'est à dire de la forme  $2^{-j}$  avec  $j$  entier). De même, à chaque base de paquets d'ondelettes bidimensionnelle correspond une partition finie du carré unité semi-ouvert

$$\{(x,y) \in \mathbb{R}^2 : 0 \leq x < 1, 0 \leq y < 1\}$$

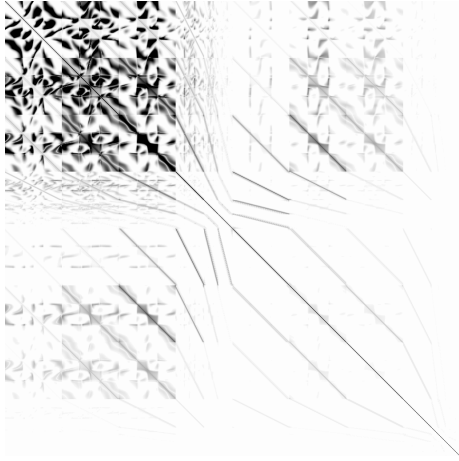


FIG. 43 – *La matrice de rigidité pour la base d'ondelettes trouvée par la stratégie Top-down, qui part de la figure 34.*

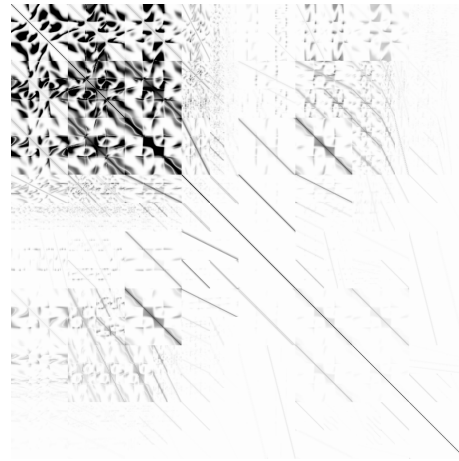


FIG. 44 – *Compression optimale de l'image bidimensionnelle de la matrice  $M$  dans une base de paquets d'ondelettes bidimensionnelle.*

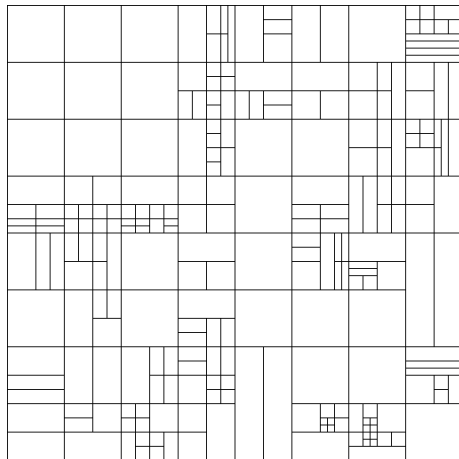


FIG. 45 – *La structure de la transformation en paquets d'ondelettes bidimensionnelle pour la figure 44.*

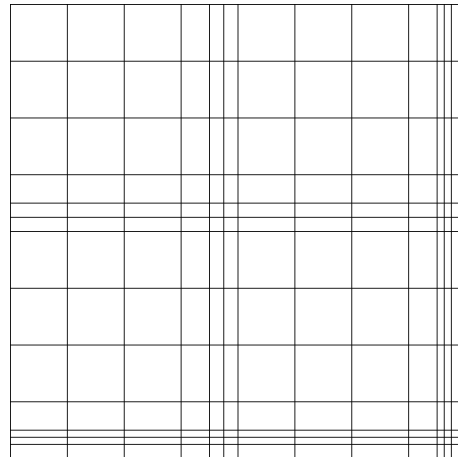


FIG. 46 – *La structure de la matrice de rigidité pour la base d'ondelettes trouvée par la stratégie Top-down, qui part de la figure 34.*

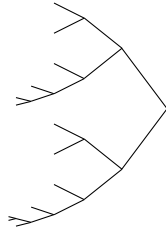


FIG. 47 – L’arbre correspondant à la stratégie *Top-Down*, qui a été construit en ajoutant les noeuds, deux par deux, lorsque la partie seuillable de  $M$  en augmentait.

en rectangles de longueur et largeur dyadiques, similairement semi-ouverts. Ainsi, la figure 44 résulte de la transformation en paquets d’ondelettes bidimensionnelle dessinée à la figure 45.

Toute matrice  $M$  dans une base de paquets d’ondelettes de  $V_h$  est aussi une transformation dans une base de paquets d’ondelettes bidimensionnelle de la matrice initiale dans la figure 34. En particulier, la matrice de la figure 43 obtenue par la stratégie *Top-Down* est aussi la transformation dans la base de paquets d’ondelettes correspondant au découpage figure 46.

Ainsi, la recherche d’une base de paquets d’ondelettes pour  $V_h$  maximisant la partie seuillable de la matrice revient à optimiser la seuillabilité de la matrice lorsque cette matrice parcourt un sous-ensemble de l’ensemble des matrices résultant de transformations en paquets d’ondelettes bidimensionnelles. La matrice correspondant à la base optimale de paquets d’ondelettes pour  $V_h$  est donc moins bien seuillable que la matrice correspondant à la transformation optimale de paquets d’ondelettes bidimensionnelle. Cette dernière matrice fournit donc une *borne supérieure* de ce qu’on peut faire avec une base de paquets d’ondelettes pour  $V_h$ . Cette dernière matrice est plus facile à trouver même s’il faut chercher parmi un nombre encore plus grand de transformations en paquets d’ondelettes bidimensionnelles : si ces transformations sont beaucoup plus nombreuses que le nombre de base de paquets d’ondelettes, chaque transformation est voisine de beaucoup plus de transformations au sens où il suffit de couper un rectangle. Ainsi, la recherche de cette dernière matrice est plus facile car il y a plus de chance qu’un minimum local soit le minimum global. La matrice de la figure 44 est donc susceptible d’être un minimum global.

Et en effet, la matrice de la figure 43 a à peu près le même découpage que la matrice de la figure 44, sauf à certains endroits où les coefficients sont tous petits ; et la matrice de la figure 43 est une matrice (non-standard) de laquelle on peut enlever 73% des coefficients tout en garantissant que l’erreur

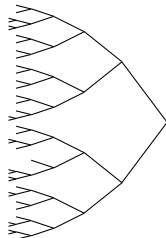


FIG. 48 – *L'arbre correspondant au paquet d'ondelette contenant toutes les subdivisions de la transformation en paquets d'ondelettes bidimensionnelle optimale figure 44.*

sur la solution  $J$  est de moins de 8%. La matrice de la figure 43 est seuillable à 70%, ce qui est très proche de 73% : ainsi, nous avons vu une raison pour laquelle la stratégie Top-Down est efficace.

Toutefois, la stratégie non-standard ne donne elle-même pas de résultat utilisable.

D'une part, la recherche d'une base de paquets d'ondelettes de  $V_h$  qui ait les mêmes subdivisions que celles présentes dans la transformation de paquets d'ondelettes bidimensionnelle est loin d'être satisfaisante. Une telle transformation résulterait en la base représentée figure 48. La transformation en paquets d'ondelettes bidimensionnelle figure 49 résultant de la tensorisation de la base 48 par elle-même donne la matrice figure 51 qui n'est seuillable qu'à 60% pour une tolérance de 8% sur la solution  $J$  : c'est bien moins que les 70% permis par la matrice figure 43 de la stratégie précédente Top-down pour la même tolérance.

D'autre part, il existe un intermédiaire entre les découpages figures 49 et 45 : nous pouvons utiliser deux bases, une base de départ et une base d'arrivée, en multipliant  $M$  à gauche et à droite non pas par la même matrice de passage mais avec deux matrices de passages différentes. Les découpages résultants sont tensorisés, mais non symétriques, comme le découpage de la figure 50. Ce dernier découpage est en fait le découpage le plus grossier qui soit tensorisé (les lignes de découpes sont discontinues) et plus fin que le découpage de la figure 45. Toutefois, l'utilisation de deux bases au lieu d'une ne mène pas à des résultats très satisfaisant : la matrice  $M$  qui en résulte n'est seuillable qu'à 62% pour une tolérance de 8% sur la solution  $J$  : c'est bien moins que les 70% permis par la stratégie précédente Top-down pour la même tolérance.

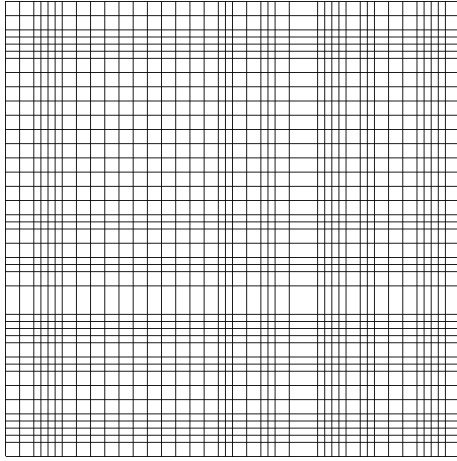


FIG. 49 – *Dessin de la base de paquets d'ondelettes de  $V_h$  recouvrant le dessin de la transformation en paquets d'ondelettes bidimensionnelle de la figure 44.*

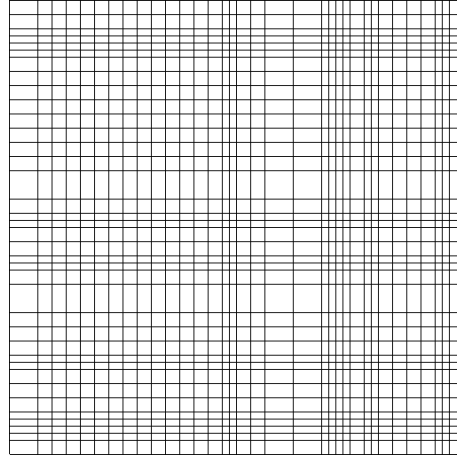


FIG. 50 – *Dessin de la transformation en paquets d'ondelettes tensorisée et recouvrant le dessin de la transformation en paquets d'ondelettes bidimensionnelle de la figure 44.*



FIG. 51 – *La matrice de rigidité dans la base de paquets d'ondelettes de  $V_h$  recouvrante de la figure 49.*

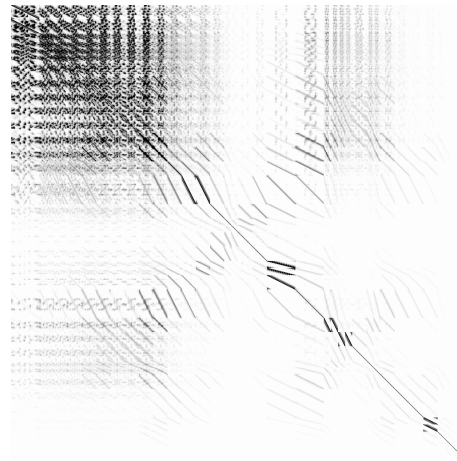


FIG. 52 – *La transformation en paquets d'ondelettes tensorisée recouvrante de la figure 50 de la matrice originelle figure 34.*

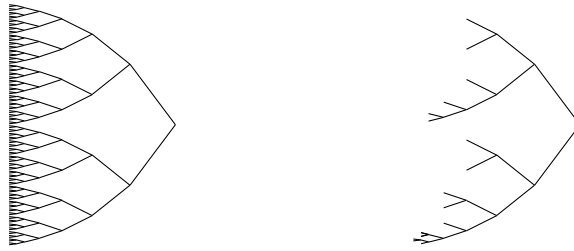


FIG. 53 – *Départ et arrivée de la stratégie Bottom-up.*

### Stratégie Bottom-up.

Cette autre stratégie, par exemple citée Coifman-Wickerhauser [CW92], consiste à construire, peu à peu, cette base de paquets d'ondelettes, mais en partant d'une base d'ondelette type Fourier donnant la figure 40 et la matrice de passage figure 18, puis en supprimant petit à petits les facteurs de produits dans cette dernière matrice de passage, en ne validant chaque suppression que lorsqu'elle contribue à augmenter le pourcentage de termes de la matrice supprimables au sens (88) qui garantit que l'erreur entre  $J$  et notre approximation  $J_h$  est de l'ordre que 8% (théorème 5.1).

Dans la représentation par arbre, cela revient à supprimer les enfants d'un noeud n'ayant pas de petit-enfant, progressivement, en validant chaque suppression par un calcul de norme d'opérateur par la méthode ARPACK++ citée page 145. C'est la stratégie duale de la stratégie Top-down.

- Les noeuds les plus grossiers ne sont pas évalués.
- Cette méthode est assez coûteuse, comme on part d'un arbre avec beaucoup de noeuds.
- Le résultat dépend de l'ordre avec lequel les suppressions sont tentées.

La figure 53 présente l'arbre de la base de paquets d'ondelettes de départ, et celle d'arrivée, qui compte moins de noeuds que celle de départ. Les matrices correspondantes sont représentées sur les figures 54 et 55. Le résultat, sur la figure 55, est compressible à 70% pour une tolérance maximum de 8% sur la solution  $J$ . C'est la même compressibilité que pour la stratégie Top-Down; de plus, nous remarquons que les arbres ne diffèrent que tout en bas, et que cette différence ne concerne que des coefficients très petits (en bas ou à droite des matrices), qui sont seuillés, et donc dont la découpe n'influe pas sur le résultat : cette différence ne peut donc pas avoir d'effet.

Ainsi, cette stratégie Bottom-up, qui a pour défaut de ne pas évaluer les noeuds les plus grossiers, trouve le même résultat que la stratégie Top-Down, qui avait pour défaut de ne pas évaluer les noeuds les plus fins. C'est la seconde validation de la stratégie Top-Down.

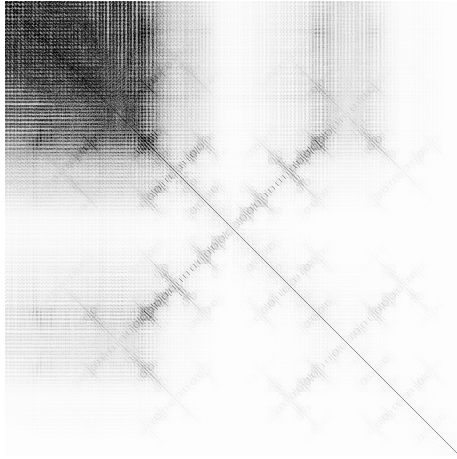


FIG. 54 – *La matrice de rigidité pour une base de paquets d'ondelettes type Fourier plus raffinée qu'en (67a).*

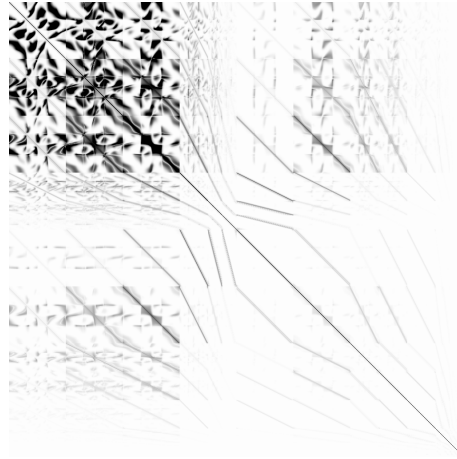


FIG. 55 – *La matrice de rigidité pour la base d'ondelettes trouvée par la stratégie Bottom-up, qui part de la figure 54.*

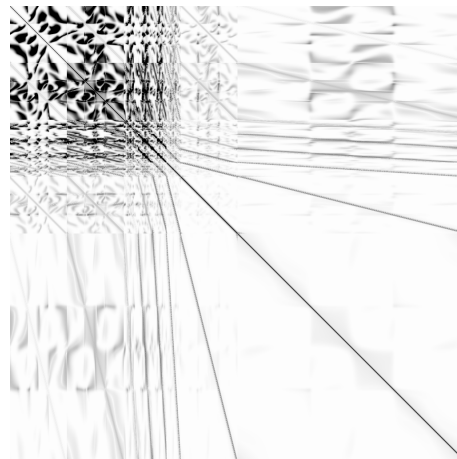


FIG. 56 – *La matrice de rigidité pour une base d'ondelettes de la stratégie Predefined-Wavelet-Packet.*

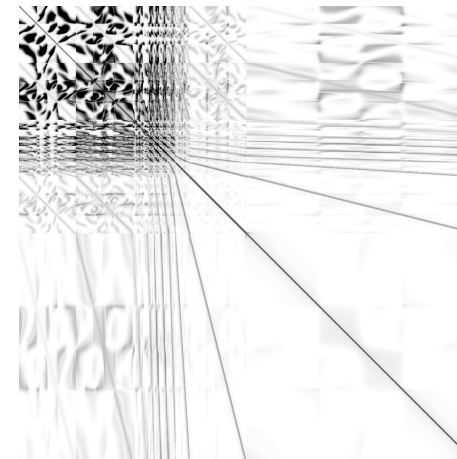


FIG. 57 – *La matrice de rigidité pour une autre base d'ondelettes de la stratégie Predefined-Wavelet-Packet.*



FIG. 58 – Les deux arbres fréquemment directs de la stratégie *Predefined-Wavelet-Packet* utilisés dans les figures 56 et 57.

Pour information, la matrice complète de la figure 54 qui correspond à la base de paquets d'ondelettes donnée à la figure 53 n'est seuillable qu'à 36% si on veut garantir une tolérance de 8% sur la solution : donc il faut faire attention à ne pas trop raffiner lorsqu'on choisit sa base de paquets d'ondelettes. Nous mentionnons la stratégie suivante, publiée par Deng-Ling [DL99b], qui s'avèrera peu efficace, un peu parce qu'elle découpe trop là où il y a des coefficients importants.

### Stratégie *Predefined-Wavelet-Packet*.

Deng-Ling [DL99b] propose de n'insérer de noeuds, dans l'arbre binaire représentant la base de paquets d'ondelettes choisie, qu'à l'endroit correspondant aux éléments des bases de paquets d'ondelettes de fréquence à peu près  $\frac{2\pi}{\lambda_0}$ , c'est à dire aux fonctions  $\psi_{j',*}^{(n)}$  telles que  $n \approx 2^{1-j'}/\lambda_0 - \frac{1}{2}$  (Deng-Ling garde en moyenne deux  $n$  par valeur de  $j'$ ). Cette base de paquets d'ondelettes est donc connue à priori : ainsi, Deng-Ling a appelé cette base *Predefined Wavelet Packet*, d'où le nom de cette stratégie.

Lorsque  $h = \frac{1}{64}$  et  $\lambda_0 = \frac{1}{10}$ , la figure 53 présente deux arbres constitués de noeuds pour lesquels  $|n - 2^{1-j'}/\lambda_0 + \frac{1}{2}| < 2$ . Les figures 56 et 57 présentent les matrices correspondantes. Ces matrices sont respectivement seuillables à 60% et 58%, en garantissant une tolérance de 8% sur la solution.

Attention, la présentation et les résultats numériques de Deng-Ling [DL99b] ne tiennent pas du tout compte du remplacement (66) qui est nécessaire pour que ce critère  $n + \frac{1}{2} \approx 2^{1-j'}/\lambda_0$  ait un sens (voir la remarque 5.8).

Ainsi, leurs applications numériques utilisent en réalité les arbres de la figure 59, ce qui donne et les matrices résultantes (non représentées) sont seuillables à respectivement 57% et 59%.





FIG. 59 – Les deux arbres réellement utilisés par Deng-Ling qui a ignoré le remplacement (66).

$h$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$
$1000\varepsilon$	21	10.5	4.2	2.1	1.0	10.5	4.2	2.1	1.0	0.42	0.21	0.10
Milliers de coefficients	4.8	6.2	7.9	9.0	9.9	7.2	9.8	12.7	14.5	16.9	19.4	21.3
Base de paquets d'ondelettes												
$100 \frac{\ J - J_h\ }{\ J\ } \leq$	17	10	6	5	4	6	3	2	1.4	1.1	0.95	0.89

Table 6 – Stratégie Top-Down pour  $\lambda_0 = 1$  lorsque  $S$  est donné par (34).

## 5.3 Position des coefficients importants de cette matrice

### 5.3.1 Chiffres de compression

Nous avons donc utilisé la stratégie Top-Down, qui a donné les résultats des tables 6 et 7, dans le cas du diffractant (34) sans coin. Essentiellement, l'arbre généralement construit est un arbre de paquets d'ondelettes bien équilibré pour les fréquences plus petites que  $\frac{3\pi}{\lambda_0}$  à peu près, puis moins raffiné pour les fréquences supérieures. Nous remarquons que la matrice  $a$  a à peu près  $\lambda_0^{-1.4}$  coefficients nécessaires dans la matrice de masse pour la précision garantie de 2% sur le résultat.

### 5.3.2 Éviter le calcul des coefficients seuillés

Le bilan du calcul fait jusqu'ici est assez lourd :

- Le calcul de la matrice  $M_0$  dans la base (65b) de  $V_h$  contient  $h^{-2}$  intégrales doubles, dont  $16h^{-1}$  d'entre elles prennent de l'ordre de  $10^6$  opérations (les intégrales (86) pour lesquelles les supports de  $w_m$  et












$h$	$\frac{1}{64}$	$\frac{1}{64}$	$\frac{1}{64}$	$\frac{1}{64}$	$\frac{1}{64}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{128}$	$\frac{1}{128}$	$\frac{1}{128}$	$\frac{1}{128}$
$1000\varepsilon$	5.3	3.18	2.12	1.59	1.06	0.53	2.12	1.06	0.53	.212	.106
Milliers de coefficients	216	245	269	286	310	394	315	396	464	529	649
Base de paquets d'ondelettes											
$100 \frac{\ J - J_h\ }{\ J\ } \leq$	25	16	12	10	8	6.5	6.7	4.0	2.7	2.0	1.7

Table 7 – *Stratégie Top-Down pour  $\lambda_0 = \frac{1}{10}$  lorsque  $S$  est donné par (34).*

- $w_{m'}$  s'intersectent), et dont les autres ne prennent en moyenne que « seulement » de l'ordre de  $10^3$  opérations.
- La stratégie Top-Down demande ensuite de calculer une norme matricielle une fois par feuille essayée; le coût de calcul de cette norme matricielle, par ARPACK++ (page 145) avec des méthodes d'Arnoldi, est tout de même d'une cinquantaine de produits matrice-vecteur.
  - De plus, il faut calculer à partir de  $M_0$  la matrice  $M$  dans les bases de paquets d'ondelettes essayées au cours de la stratégie Top-Down; à cause de la nature par blocs des matrices de passages vue en partie (3.2.2), le coût cumulatif de ces transformations n'est qu'en  $h^{-2} |\log h|$  au plus.
  - Enfin, il faut estimer les normes d'opérateur intervenant dans les hypothèses des théorèmes 3.1 et 5.1, ce qui en pratique, avec la méthode utilisée pour les tables 4 et 5, coûte approximativement quatre fois plus cher que le premier alinéa ci-dessus.

Pour essayer d'aller plus vite, nous pourrions dans un premier temps économiser sur le second alinéa. Les résultats des autres stratégies, ainsi que les diverses matrices montrées dans les sous-sections précédentes, nous suggèrent de construire une nouvelle stratégie, similaire à la stratégie Predefined-Wavelet-Packet, mais en raffinant toutes les fréquences plus petites que  $\frac{2\pi}{\lambda_0}$  plutôt que de ne raffiner que cette dernière fréquence: ces fréquences correspondent à tout le bloc en haut à gauche de toutes les matrices illustrées dans les figures précédentes.

Les figures 56 et 57 montrent que la stratégie Predefined-Wavelet-Packet a tendance à découper en de multiples blocs la partie de la matrice contenant l'essentiel des coefficients; nous allons voir que ce sur-découpage est néfaste. La figure 49 montre que la stratégie Non-standard qui lui correspond a elle aussi tendance à sur-découper.

Les figures 40 et 54 ne correspondent pas à proprement parler à des stratégies viables ; néanmoins, la seconde d'entre elle a également tendance à sur-découper ce bloc. Curieusement, l'autre est celle que nous avons vu seuitable à 67%, ce qui est supérieur à toutes les autres figures que nous venons de rappeler (qui sont respectivement seuitable à 60%, 58%, 60% et 36%).

Examinons mieux cette matrice, illustrée figure 40 : dans les 36 blocs en haut à gauche, nous voyons que le dessin de chaque bloc est différent. Dans une certaine mesure, ces dessins sont assez complémentaires si on essaye de les superposer, c'est à dire que (comme nous allons le voir ci-dessous) si un pixel ligne  $l$  et colonne  $c$  (avec  $|l - c| \gg 1$ ) dans un bloc donné est noir, il y a de fortes chances que ses homologues, ligne  $l$  colonne  $c$  mais dans les autres blocs, soient plus souvent blanc que noirs. En effet, pour  $l = 64$  et  $c = 42$ , considérons le coefficient ligne  $l$  colonne  $c$  dans le sixième bloc de la deuxième rangée, bloc que la figure 60 présente seul. D'après (86) et d'après l'orthogonalité des paquets d'ondelettes, ce coefficient vaut

$$\begin{aligned} & \int_0^L \psi_{\frac{|\log h|}{\log 2}-4,c-1}^{(\kappa-1)} \psi_{\frac{|\log h|}{\log 2}-4,l-1}^{(g-1)} dt + \int_0^L \int_0^L \psi_{\frac{|\log h|}{\log 2}-4,c-1}^{(\kappa-1)}(t) K(t,t') \psi_{\frac{|\log h|}{\log 2}-4,l-1}^{(g-1)}(t') dt dt' \\ &= \delta_\kappa^g \delta_c^l + \int_0^L \int_0^L \psi_{2,41}^{(1)}(t) K(t,t') \psi_{2,63}^{(5)}(t') dt dt' \\ &= 0 + \int_0^L \int_0^L \psi_{0,0}^{(1)}(4(t-41)) K(t,t') \psi_{0,0}^{(5)}(4(t'-63)) dt dt' \end{aligned}$$

qui contient deux fonctions  $\psi_{2,41}^{(1)}$  et  $\psi_{2,63}^{(5)}$  qui d'après les graphes de  $\psi_{0,0}^{(*)}$  dessinés en (67b) ont le gros de leurs valeurs localisé comme indiqué sur la figure 61. Et en effet, nous voyons que ce coefficient, dans la ligne  $l = 64$  et la colonne  $c = 42$  de son bloc, qui en fait vaut à peu près  $0.013 - 0.013i$ , est plus grand que tout coefficient dans la ligne  $l = 64$  et la colonne  $c = 42$  mais dans un autre bloc, selon les figures 62 et 63.

Or le coefficient dans la ligne  $l = 64$  et la colonne  $c = 42$  dans le  $\kappa$ -ième bloc de la  $g$ -ième rangée vaut

$$\begin{aligned} & \int_0^L \int_0^L \psi_{2,41}^{(g-1)}(t) K(t,t') \psi_{2,63}^{(\kappa-1)}(t') dt dt' \\ &= \int_0^L \int_0^L \psi_{0,0}^{(g-1)}(4(t-41)) K(t,t') \psi_{0,0}^{(\kappa-1)}(4(t'-63)) dt dt' \end{aligned} \quad (89a)$$

qui, comme nous allons le voir en utilisant l'argument de l'angle d'émission fréquentiel (67i) donné par Canning [C95] et la localisation fréquentielle des paquets d'ondelettes (67c), est petit lorsque  $g \neq 2$ , ou lorsque  $\kappa \neq 6$ .

Détaillons ce raisonnement. Pour que le coefficient (89a) soit gros, il est nécessaire d'après (67i) que la longueur d'onde dominante  $\lambda_{\psi_{0,0}^{(g-1)}}$  de la fonction  $\psi_{4,41}^{(g-1)}$  soit de l'ordre de  $\frac{\lambda_0}{\cos \theta'}$  avec  $\theta'$  l'angle (de droites) sous lequel

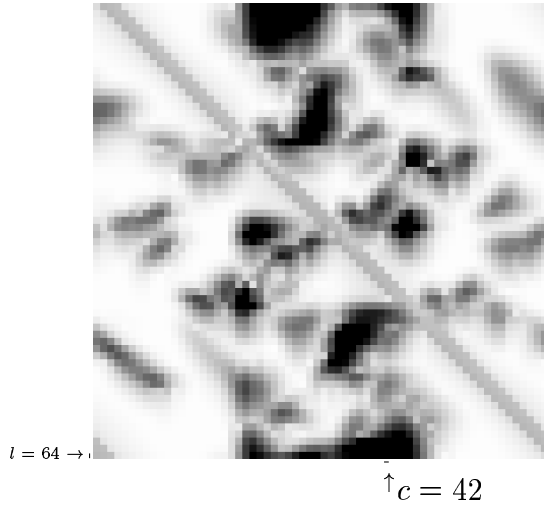


FIG. 60 – Place du coefficient ligne 64 colonne 42 dans ce sixième bloc de la seconde rangée de la figure 40.

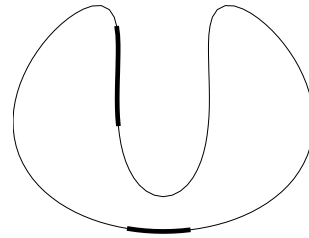


FIG. 61 – Les supports des fonctions impliquées dans ce coefficient pointé figure 60.

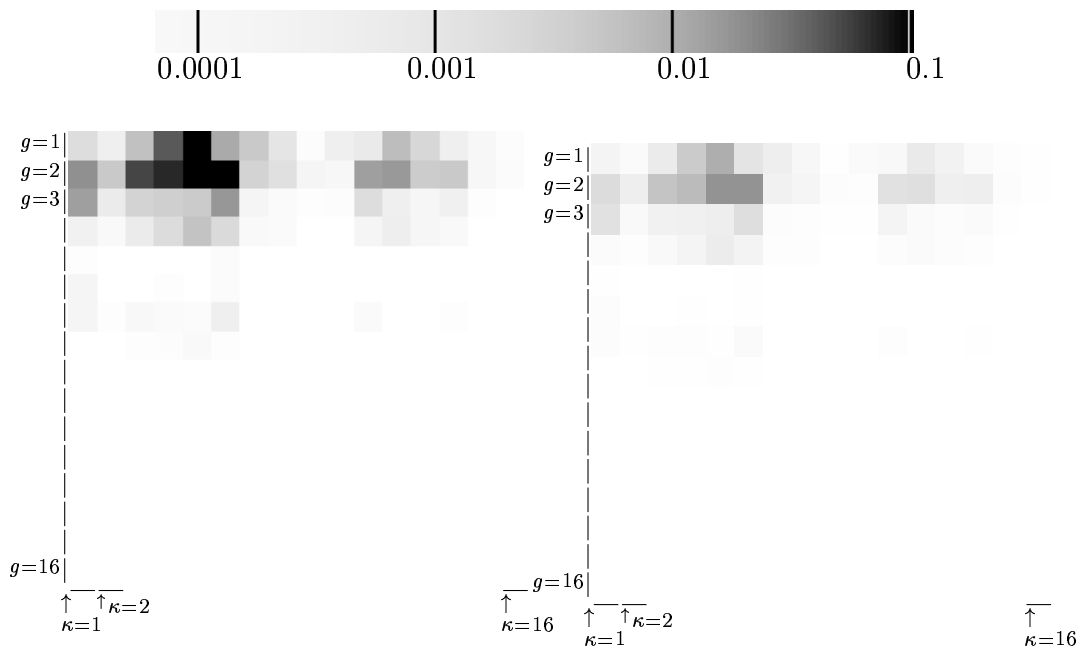


FIG. 62 – La matrice de la figure 40 dans lequel chacun des 256 blocs a été remplacé par un carré de la couleur de la taille de son coefficient ligne 64 colonne 42.

FIG. 63 – La figure 62 mais avec l'échelle de gris ci-dessous.



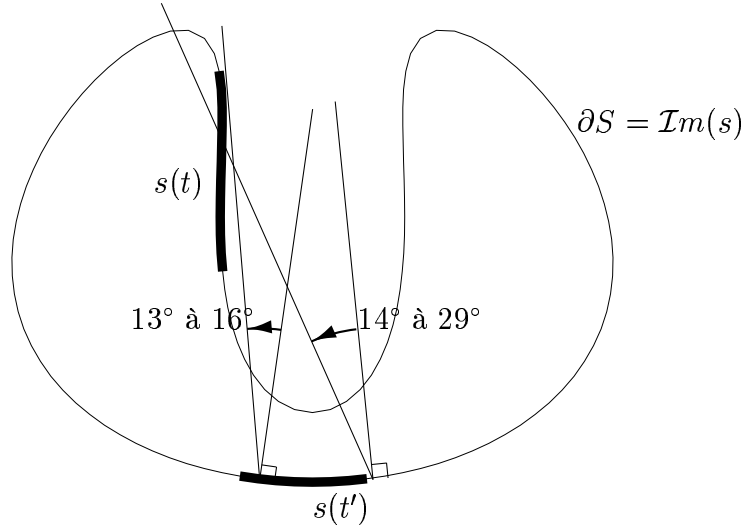


FIG. 64 – L'intervalle  $[13^\circ, 29^\circ]$  de l'angle  $\theta'$  sous lequel la fonction en  $t'$  voit la fonction en  $t$  dans (89a), lorsque le diffractant  $S$  est donné par (34).

le support de  $\psi_{2,63}^{(\kappa-1)}$  voit le support de  $\psi_{2,41}^{(g-1)}$  par rapport à la normale à  $\partial S$ . Nous voyons sur la figure 64 que cet angle  $\theta'$  a un cosinus entre  $\cos 29^\circ \approx 0.87$  et  $\cos 13^\circ \approx 0.97$ : nous voyons que les  $g$  qui conviennent sont ceux tels que  $\hat{\psi}_{0,0}^{(g-1)}(\xi)$  est grand en  $\xi$  quelque part entre  $\frac{2\pi}{2^2}(\lambda_0/.97)^{-1} \approx 13.7$  et  $\frac{2\pi}{2^2}(\lambda_0/.87)^{-1} \approx 15.2$ , ce qui en lisant les décomposition de Fourier des fonctions en  $\psi_{0,0}^{(*)}$  dans (67c) sélectionne  $g = 4$ ,  $g = 5$  et  $g = 6$ .

Ainsi, nous avons compris pourquoi, lorsque  $g \notin \{4, 5, 6\}$ , (89a) était petit. De même, la mesure de l'angle de sortie analogue  $\theta$  montre que (89a) est petit si  $\kappa \notin \{1, 2\}$ . C'est pourquoi les figures 62 et 63 sont concentrées à quelques termes concentrés sur les  $2 \times 3 = 6$  carrés désignés par ces trois valeurs de  $g$  et ces deux valeurs de  $\kappa$ .

De la même manière, il est possible de regarder les mêmes regroupements, pour  $(l, c)$  différent de (64, 42). Le même argument basé sur (89a) permettra d'expliquer un regroupement semblable dans l'analogie de la figure 62 mais pour  $(l, c)$  différent de (64, 42). Lorsqu'on regroupe tous ces analogues, en mettant sur la même ligne ceux avec le même  $l$  et sur la même colonne ceux qui ont le même  $c$ , nous trouvons la figure 65. La figure 65 contient la figure 63 dans son bloc dernière ligne ( $l = 64$ ) et dans la colonne  $c = 42$ .

Cette figure représente en fait la matrice  $M$  de notre système (85) lorsque la base  $\Lambda_h$  choisie vaut

$$\{\psi_{j-4,m}^{(n)}, n \in \{0, 1, \dots, 2^4 - 1\}, m \in \mathbb{Z}/L2^{j-4}\mathbb{Z}\} \quad (89b)$$

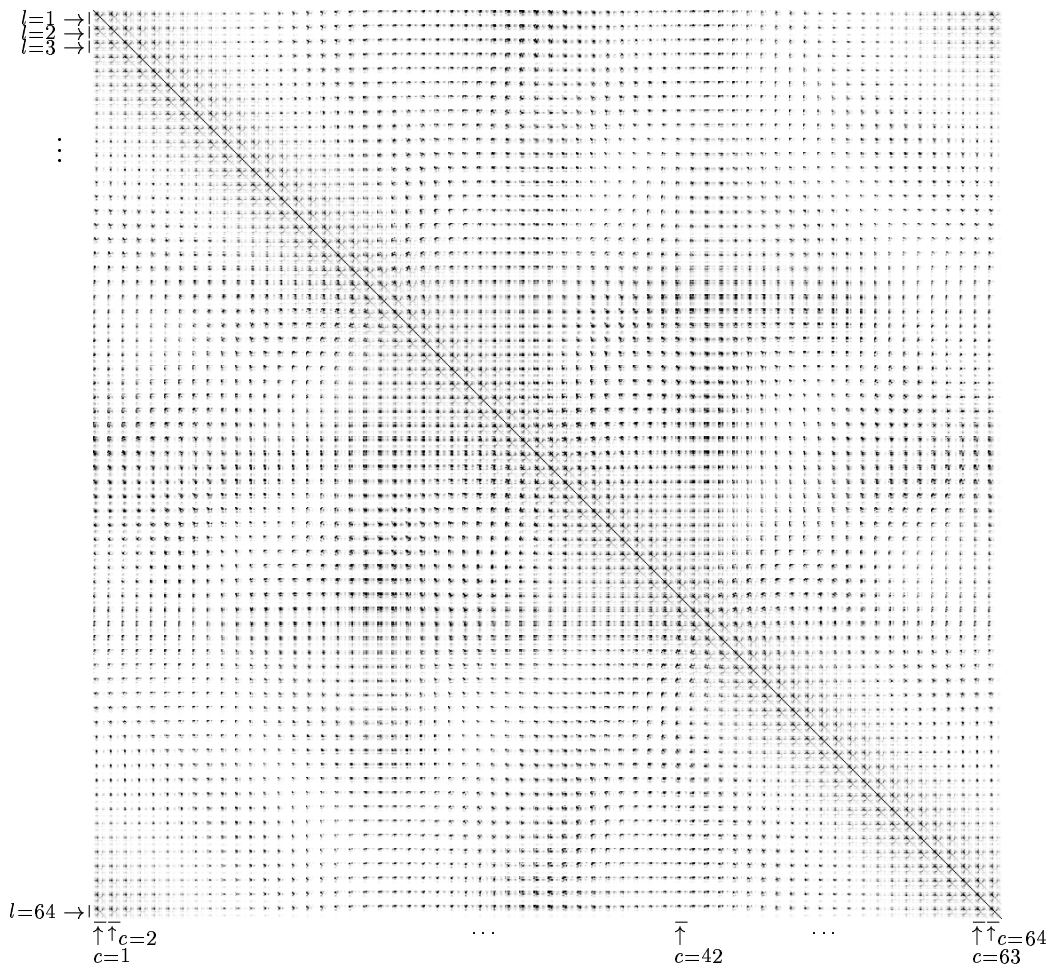


FIG. 65 – Agrégation du cas ligne  $l$  colonne  $c$  de la figure 62, pour  $l$  et  $c$  entre 1 et 64.



qui est très similaire à (67a) sauf que l'ordre n'est pas le même : la base (67a) commence par

$$\psi_{j-4,0}^{(0)}, \psi_{j-4,1}^{(0)}, \dots, \psi_{j-4,63}^{(0)}, \psi_{j-4,0}^{(1)}, \psi_{j-4,1}^{(1)}, \dots, \psi_{j-4,63}^{(1)}, \psi_{j-4,0}^{(2)}, \psi_{j-4,1}^{(2)}, \dots$$

tandis que la base (89b) commence par

$$\psi_{j-4,0}^{(0)}, \psi_{j-4,0}^{(1)}, \dots, \psi_{j-4,0}^{(15)}, \psi_{j-4,1}^{(0)}, \psi_{j-4,1}^{(1)}, \dots, \psi_{j-4,1}^{(15)}, \psi_{j-4,2}^{(0)}, \psi_{j-4,2}^{(1)}, \dots$$

La première regroupe les termes de la base par leur similarité fréquentielle, et la seconde par leur position spatiale. Ainsi, la figure 65 présente en fait la même chose que la figure 40, sauf que les coefficients  $y$  sont d'abord regroupés sur des critères spatiaux.

Nous allons maintenant justifier l'argument de l'angle de sortie (67i). Attention, la justification suivante n'est pas rigoureuse (la propriété 5.6 fournira une justification rigoureuse).

D'après les représentations des fonctions  $\psi_{0,0}^{(*)}$  en (67b) et de leur décomposition de Fourier en (67c), nous voyons que les fonctions  $\psi_{j',c-1}^{(\kappa-1)}(t)$  et  $\psi_{j',l-1}^{(g-1)}(t')$  se comportent en gros comme les fonctions

$$b_1(t)e^{2i\pi\frac{k_1}{\lambda_0}t} \quad \text{et} \quad b_2(t')e^{2i\pi\frac{k_2}{\lambda_0}t'}$$

avec

$$k_1 = \pm 2^{j'-1}\lambda_0\left(\kappa - \frac{1}{2}\right) \quad \text{et} \quad k_2 = \pm 2^{j'-1}\lambda_0\left(g - \frac{1}{2}\right), \quad (89c)$$

et avec  $b_1$  et  $b_2$  qui sont deux fonctions  $L$ -périodiques de supports de longueur  $O(2^{j'})$  et centrés approximativement en  $2^{-j'}c$  et en  $2^{-j'}l$  respectivement ; la propriété 5.6 montrera dans quelle mesure ces choix de  $k_1$  et de  $k_2$  faits en (89c) sont bons et discutables à la fois. Nous allons supposer, pour l'instant, que le support de  $b_1$  et de  $b_2$  ne dépend pas trop de  $\kappa$  ou de  $g$  ; cette hypothèse sera confirmée par la propriété 5.6.

Ainsi, nous voyons que le coefficient (89a) de la matrice  $M$  est de la forme (73) pour lequel la propriété 4.3 montre une décroissance en  $d_+^{-1}$  où  $d_+$ , défini dans cette propriété, est plus petit que la distance entre le vecteur  $\begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$  et la diagonale épaissie de l'amplitude de variation de  $\frac{1}{2\pi}\nabla u$ .

Or dans le cas où  $c = 42$  et  $c = 64$ ,  $b_1$  et  $b_2$  ont des supports qui sont à peu près indiqués sur la figure 64, et la figure 66 présente la trajectoire de  $\nabla u(t, t')$  pour  $t$  et  $t'$  restreints aux supports de respectivement  $b_1$  et  $b_2$ . Comme premier signe de cohérence, nous voyons que cette trajectoire est, en ordonnée, située entre  $-2\pi \sin 13^\circ$  et  $-2\pi \sin 29^\circ$ , où ces deux angles avaient

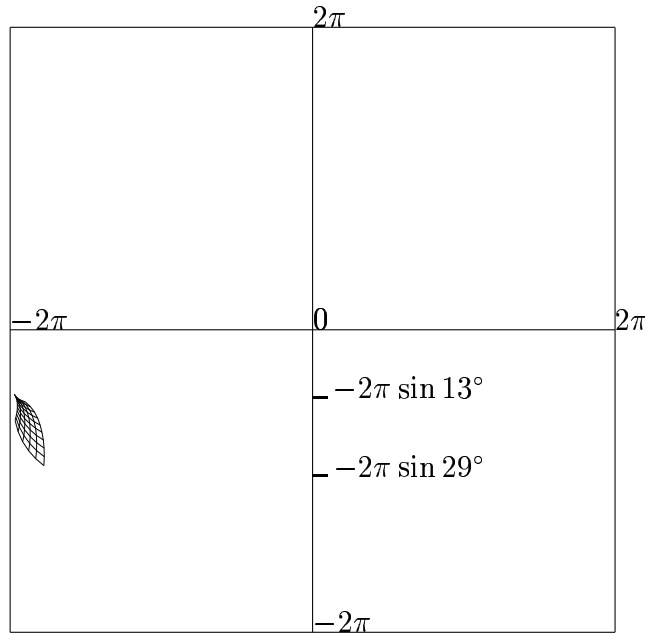


FIG. 66 – La trajectoire de  $\nabla u(t, t')$  avec  $u$  défini par (74), restreint à  $t \in \text{supp}b_1$  et  $t' \in \text{supp}b_2$ .

été mesurés sur la figure 64. Et nous voyons que le maximum du paramètre  $(\frac{\partial u}{\partial t} + \frac{\partial u}{\partial t'})$  qui entre dans la définition de  $d_+$  est entre 7.6 et 8.6.

La propriété 4.3 prédit donc que  $|k_1 + k_2|$  ne doit pas être entre  $\frac{7.6}{2\pi} \approx 1.21$  et  $\frac{8.6}{2\pi} \approx 1.37$ . Vu la définition de  $k_1$  et de  $k_2$  en (89c), nous voyons que, si (89a) est de la forme (73), alors (89a) est grand lorsque

$$|\kappa - g| \quad \text{ou} \quad |\kappa + g - 1| \quad \text{est proche de} \quad 2\lambda_0[1.21, 1.37]$$

et comme l'intervalle ci-dessus est approximable par  $[6.0, 6.9]$ , nous voyons que cela désigne les carrés indiqués par les cercles, et leurs voisins plus ou moins immédiats, dans la figure 67, avec l'hypothèse importante que le support de  $b_1$  et celui de  $b_2$  ne doivent pas trop dépendre de  $\kappa$  ou de  $g$ .

Nous remarquons que les gros termes sont effectivement près des coefficients pointés par la propriété 4.3. Néanmoins, nous avons signalé juste après la fin de la démonstration de cette propriété qu'il aurait suffi de s'assurer que  $\begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$  est loin de l'image de  $\frac{1}{2\pi}\nabla u$ , ce qui ici est un critère qui sélectionne très exactement les six coefficients cerclés les plus hauts lorsqu'on compare les définitions de  $k_1, k_2$  faites en (89c) avec des changements de signes, et les figures 63 et 66.



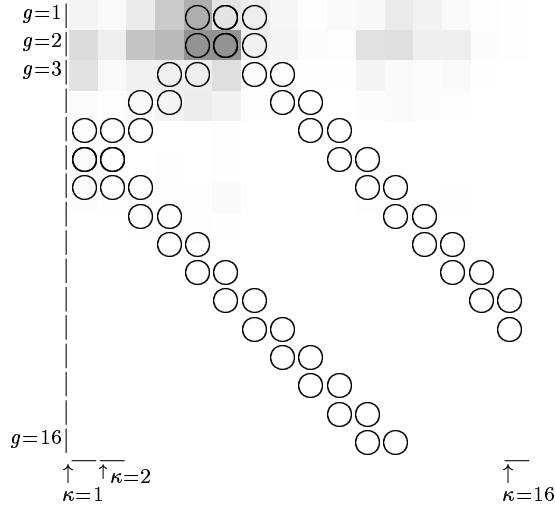


FIG. 67 – Environs où la propriété permet de détecter les gros termes de la figure 63.

Nous allons démontrer qu'en effet c'est possible, mais seulement lorsque les supports de  $b_1$  et de  $b_2$  ne s'intersectent pas, ce qui est une hypothèse nouvelle par rapport aux propriétés 4.4 et 4.3. Incidemment, nous allons démontrer le critère angulaire (67i). Un homologue de cette propriété avec un plan de démonstration a été trouvé par mon directeur de thèse dans Bradie-Coifman-Grossman [BCG93] (et ébauché par Averbuch-Braverman-Coifman-Israeli-Sidi [ABC00]) pour un opérateur de noyau purement oscillatoire du type (67h) et non pour l'opérateur intégral de Helmholtz de noyau (89a).

**Propriété 5.2** *En l'absence de coin dans le diffractant, si  $s$  est  $C^{p+3}$ , et si les supports de  $b_1$  et de  $b_2$  ne s'intersectent pas, l'intégrale (73) est en*

$$O\left(\frac{\lambda_0^p}{d'^2 d^p} (\|b_1\|_{L^1} \|b_2\|_{L^1} (\frac{1}{d^p} + \frac{1}{d^p}) + \|b_1\|_{L^1} \|b_2\|_{W^{p,1}} + \|b_1\|_{W^{p,1}} \|b_2\|_{L^1})\right)$$

où  $p \geq 1$ , où

$$d = \sqrt{\min(k_1^2 + k_2^2, \inf_{t:b_1(t) \neq 0} \inf_{t':b_2(t') \neq 0} (k_1 - \frac{1}{2\pi} \frac{\partial u}{\partial t})^2 + (k_2 - \frac{1}{2\pi} \frac{\partial u}{\partial t'})^2)}$$

et où  $d'$  est la distance

$$d' = \inf_{t:b_1(t) \neq 0} \inf_{t':b_2(t') \neq 0} \|s(t) - s(t')\|$$

mesurée dans  $\partial S$  entre le support de  $b_1$  et le support de  $b_2$ , et cette distance  $d'$  doit être strictement positive. Nous nous restreignons au cas où  $\lambda_0 < \frac{1}{2}$ .

■ Nous rappelons que  $u(t, t') = 2\pi|s(t) - s(t')|$  (74). La définition de  $K$  en (51), et les développements limités (54d) et (54e) (avec  $n = p + 1$ ) montrent que  $K$  est une combinaison linéaire (à coefficients indépendants de  $u$  et de  $\lambda_0$ ) de ces termes :

$$\frac{1}{\lambda_0} \log\left(\frac{1}{1 + \frac{\lambda_0^2}{u^2}}\right), \quad \lambda_0^{m-1}(\lambda_0^2 + u^2)^{-\frac{m}{2}}, \quad \lambda_0^{m-\frac{3}{2}} e^{i\frac{1}{\lambda_0}u} (\lambda_0^2 + u^2)^{-\frac{m}{2} + \frac{1}{4}}, \quad (89d)$$

$$\frac{\partial}{\partial t} u(u^2)^{-\frac{1}{2}}, \quad \lambda_0^{m-1}(\lambda_0^2 + u^2)^{-\frac{m}{2}} \frac{\partial}{\partial t} u, \quad \lambda_0^{m-\frac{3}{2}} e^{i\frac{1}{\lambda_0}u} (\lambda_0^2 + u^2)^{-\frac{m}{2} + \frac{1}{4}} \frac{\partial}{\partial t} u, \quad (89e)$$

$$O(\lambda_0^{p+\frac{1}{2}}(\lambda_0^2 + u^2)^{-\frac{p}{2} - \frac{3}{4}}), \quad O(\lambda_0^{p+\frac{1}{2}}(\lambda_0^2 + u^2)^{-\frac{p}{2} - \frac{3}{4}} \frac{\partial}{\partial t} u); \quad (89f)$$

Comme  $1 \leq m \leq p$ , comme  $u$  est minoré par  $d'$ , comme  $(u^2)^{-\frac{1}{2}} \nabla(u^2)$  est borné car  $\nabla u$  est borné, comme  $u^2$  a une norme  $C^n$  en  $O(1)$ , on a, pour  $n \geq 1$ , et uniformément en  $\lambda_0 < \frac{1}{2}$  :

$$\begin{aligned} \left\| \frac{1}{\lambda_0} \log\left(\frac{1}{1 + \frac{\lambda_0^2}{u^2}}\right) \right\|_{C^n} &= \lambda_0 O(d'^{-2}) + \lambda_0 \left\| \frac{1}{1 + \frac{\lambda_0^2}{u^2}} (u^2)^{-2} \nabla(u^2) \right\|_{C^{n-1}} \\ &= \lambda_0 O(d'^{-2}) + \lambda_0 \left\| \frac{1}{u^2 + \lambda_0^2} \frac{1}{u^2} \nabla(u^2) \right\|_{C^{n-1}} \\ &= \lambda_0 O(d'^{-1-n} (\lambda_0^2 + d'^2)^{-1} + d' (\lambda_0^2 + d'^2)^{-1 - \frac{n}{2}}) \\ &= O(d'^{-2-n}) \end{aligned}$$

où on a utilisé que  $|\nabla(u^2)| \leq 2u \leq 2\sqrt{u^2 + \lambda_0^2}$ , et que une dérivée  $q$ -ième de  $\frac{1}{u^2 + \lambda_0^2} \frac{1}{u^2}$  est le produit de  $\frac{1}{u^2 + \lambda_0^2} \frac{1}{u^2}$  par des termes du genre  $\frac{\partial^m(u^2)}{u^2 + \lambda_0^2} m$ ,  $\frac{\partial^m(u^2)}{u^2} m$ , qui sont ici indicés avec des multi-entiers  $m$  dont la somme cumulée doit faire  $q$  dans chaque produit issu du développement de cette dérivée  $q$ -ième. La même technique de décomposition donne aussi les trois estimations suivantes :

$$\begin{aligned} \|\lambda_0^{m-1}(\lambda_0^2 + u^2)^{-\frac{m}{2}}\|_{C^n} &= \lambda_0^{m-1} O((\lambda_0^2 + d'^2)^{-\frac{m}{2} - \frac{n}{2}} + (\lambda_0^2 + d'^2)^{-\frac{m}{2} - 1} d'^{-n}) \\ &= O((\lambda_0^2 + d'^2)^{-\frac{m+1}{2}} + (\lambda_0^2 + d'^2)^{-\frac{3}{2}} d'^{-n}) \\ &= O(d'^{-n-1}) \end{aligned}$$

$$\left\| \frac{1}{2} \frac{\partial}{\partial t} \log(u^2) \right\|_{C^n} = O(d'^{1-n})$$

$$\left\| \frac{1}{2} \lambda_0^{m-1} (\lambda_0^2 + u^2)^{-\frac{m}{2}} (u^2)^{-\frac{1}{2}} \frac{\partial}{\partial t} (u^2) \right\|_{C^n} = \lambda_0^{m-1} O((\lambda_0^2 + d'^2)^{-\frac{m}{2}} d'^{-n}) = O(d'^{-n-1})$$

Donc (73) est la somme d'un élément venant des restes (89f) et dont la norme se majorant par

$$\lambda_0^{p+\frac{1}{2}} \int \int C(\lambda_0^2 + u(t, t')^2)^{-\frac{p}{2} - \frac{3}{4}} b_1(t) b_2(t) dt dt' = O(\lambda_0^{p+\frac{1}{2}} d'^{\frac{1}{2}-p} \|b_1\|_1 \|b_2\|_1)$$

où  $\|\cdot\|_1 = \|\cdot\|_{L^1(\mathbb{R}/L\mathbb{Z})}$ , et de plusieurs termes du genre

$$\int \int e^{i\frac{1}{\lambda_0}(u(t,t')-2\pi k_1 t-2\pi k_2 t')} v(t,t') dt dt' \quad (89g)$$

avec  $u(t,t')$  valant  $2\pi|s(t) - s(t')|$  ou éventuellement 0 (pour les premiers éléments de (89d)), et où  $v$  a une norme  $W^{n,1}$  en

$$O\left(\|b_1\|_1 \|b_2\|_1 / d'^n + \|b_1\|_1 \|b_2\|_{W^{n,1}} + \|b_1\|_{W^{n,1}} \|b_2\|_1 / d'^2\right)$$

uniformément en  $\lambda_0$ , où nous avons utilisé des inégalités de normes en

$$\|b\|_{W^{q,1}}^p \leq \|b\|_1^{p-q} \|b\|_{W^{p,1}}^q$$

tirées de l'interpolation d'espaces normés vue en figure 1.

Notons  $k$  le vecteur  $2\pi \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$ , et  $z$  le vecteur  $\begin{bmatrix} t \\ t' \end{bmatrix}$ . On a donc à intégrer par parties en intégrant ceci :  $e^{i\frac{1}{\lambda_0}(u-\langle k|z \rangle)}$ . Or on a

$$\nabla e^{i\frac{1}{\lambda_0}(u-\langle k|z \rangle)} = \frac{i}{\lambda_0} (\nabla u - k) e^{i\frac{1}{\lambda_0}(u-\langle k|z \rangle)} \quad (89h)$$

et

$$-i\lambda_0 \left\langle \frac{\nabla u - k}{\|\nabla u - k\|^2} \middle| \nabla \right\rangle e^{i\frac{1}{\lambda_0}(u-\langle k|z \rangle)} = e^{i\frac{1}{\lambda_0}(u-\langle k|z \rangle)} \quad (89i)$$

Donc on utilise, sur le reste de l'intégrande, la dérivation à coefficients non constants

$$D : f \mapsto \left\langle \nabla \middle| f \frac{\nabla u - k}{\|\nabla u - k\|^2} \right\rangle \quad (89j)$$

qui se base sur la fonction

$$\frac{\nabla u - k}{\|\nabla u - k\|^2} \quad (89k)$$

et ses dérivées, contrôlées par le lemme suivant :

**Lemme 5.2.1** *On a, dans  $\mathbb{C}$  :*

$$\frac{d^{p_1}}{dt^{p_1}} \frac{d^{p_2}}{dt'^{p_2}} \left( \frac{1}{H(t,t')} \right) = \frac{P}{H(t,t')^{p_1+p_2+1}} \quad (89l)$$

avec  $H(t,t') = \frac{\partial}{\partial t} u(t,t') + 2\pi k_1 - i \frac{\partial}{\partial t'} u(t,t') - i 2\pi k_2$  où  $P$  est un polynôme homogène en les  $\frac{d^{q_1}}{dt^{q_1}} \frac{d^{q_2}}{dt'^{q_2}} H(t,t')$  avec  $q_1 \leq p_1$  et  $q_2 \leq p_2$ . De plus,  $P$  est de degré  $p_1 + p_2$ .

■ Nous faisons une récurrence sur  $p_1 + p_2$ , initiée sans problème. Un argument générique montre qu'il suffit d'étudier l'effet de  $\frac{d}{dt}$  sur (89ℓ), ce qui donne :

$$\frac{d}{dt} \frac{d^{p_1}}{dt^{p_1}} \frac{d^{p_2}}{dt^{p_2}} \left( \frac{1}{H} \right) = \frac{H \frac{d}{dt} P - (p_1 + p_2 + 1) P \frac{\partial}{\partial t} H}{H^{p_1 + p_2 + 1} H} \quad (89m)$$

Comme

$$H \frac{d}{dt} P = \sum_{q_1, q_2} H \frac{\partial}{\partial X_{q_1, q_2}} P \frac{\partial^{q_1 + 1}}{\partial t^{q_1 + 1}} \frac{\partial^{q_2}}{\partial t'^{q_2}} H,$$

on voit que le numérateur de (89m) est bien un polynôme homogène de degré  $p_1 + p_2 + 1$ . Notons que ce lemme est une application de la formule de Faa di Bruno. ■

Comme  $(u^2)^{-\frac{1}{2}} \nabla(u^2)$  est borné, comme  $u^2$  a une norme  $C^n$  en  $O(1)$ , alors  $u$  a une norme  $C^n$  en  $1 + d^{1-n}$  ; comme la partie réelle et la partie imaginaire de  $\frac{1}{H}$  sont les composantes du vecteur (89k), on voit que ce vecteur a ses dérivées partielles d'ordre  $p$  bornées par  $C_p(\|\nabla u - k\|^{-p-1} + \|\nabla u - k\|^{-1})d'^{-p}$ .

Itérons une intégration par parties  $p$  fois (89g) en utilisant (89j), ce qui donne :

$$\iint \left( -i\lambda_0 \left\langle \frac{\nabla u - k}{\|\nabla u - k\|^2} \middle| \nabla \right\rangle \right)^p e^{i\frac{1}{\lambda_0}(u - \langle k|z \rangle)} v(z) dz = (i\lambda_0)^p \iint e^{i\frac{1}{\lambda_0}(u - \langle k|z \rangle)} D^p v(z) dz \quad (89n)$$

Comme  $D^p v$  est une somme de termes contenant seulement  $p$  dérivations (à coefficients non constants), chaque dérivation va s'appliquer sur au plus un terme  $\frac{\nabla u - k}{\|\nabla u - k\|^2}$  qui est dans la dérivation (89j), et donc, avec le lemme ci-dessus, on voit que

$$|(89g)| \leq C(p, s) \left( \frac{\|v\|_{W^{p,1}}}{\inf_z \|\nabla u(z) - k\|^p} + \frac{\|v\|_{L^1}}{\inf_z \|\nabla u(z) - k\|^{2p}} \right) \lambda_0^p \quad (89o)$$

avec  $C(p, s)$  dépendant seulement de  $p$  et du diffractant  $s$ . Ceci montre le résultat annoncé. ■

Cette propriété 5.2 permet donc de comparer les figures 63 et 66, à symétries près. Elle montre, pour la figure , que le coefficient ligne  $g$  et de colonne  $\kappa$  dans la figure 63 est important, d'après les définitions de  $k_1$  et  $k_2$  en (89c), si l'un des quatre points de coordonnées

$$\begin{bmatrix} \pm 2^{j'} \pi \lambda_0 (\kappa - \frac{1}{2}) \\ \pm 2^{j'} \pi \lambda_0 (g - \frac{1}{2}) \end{bmatrix}$$

est sur le graphe de  $\nabla u$  ou à proximité, ou proche du coin en haut à droite de cette matrice.

Comme on a  $2^j \pi \lambda_0 \approx 1.25$ , nous voyons clairement que la symétrie selon l'axe vertical de la figure 66 permet d'expliquer la taille des six coefficients cerclés les plus hauts de la figure 67.

Nous allons maintenant essayer de prédire l'évolution de la figure 63 lorsque  $l$  ou  $c$  prennent des valeurs différentes de 64 et de 42, afin de prédire la position des gros coefficients dans la matrice de la figure 65. Grâce à la propriété précédente, il suffit pour cela de comprendre la trajectoire de  $\nabla u$ , dessinée à la figure 66 pour le cas  $l = 64$  et  $c = 42$ .

Canning [C95] remarque le critère angulaire (67i) pour cela, et nous allons le montrer :

**Propriété 5.3** *L'image de  $\nabla u(t, t')$  pour  $t$  et  $t'$  restreints aux supports des fonctions respectivement  $b_1$  et  $b_2$  est restreinte à un rectangle. En abscisse, elle est constituée des valeurs possibles de  $2\pi \langle s'(t) | \frac{s(t) - s(t')}{\|s(t) - s(t')\|} \rangle$ , et en ordonnée elle est constituée des valeurs possibles de  $2\pi \langle s'(t') | \frac{s(t') - s(t)}{\|s(t') - s(t)\|} \rangle$ , pour  $t \in \text{supp}b_1$  et  $t' \in \text{supp}b_2$ .*

■ La démonstration est très simple, elle se fait à la lecture des dérivées de  $u$ , défini en (74), qui sont calculées en 75a. ■

Une bonne observation permet de dire que la figure 66 est essentiellement un arc d'ellipse épaissi, comme expliqué par Bradie-Coifman-Grossman [BCG93], et comme nous allons le montrer rigoureusement dans la propriété suivante :

**Propriété 5.4** *L'image de  $\nabla u(t, t')$  pour  $t$  et  $t'$  restreint au supports des fonctions respectivement  $b_1$  et  $b_2$  est constitué d'un arc d'ellipse épaissi.*

- cette ellipse est centrée sur l'origine ;
- les droites de pente  $-1$  sont des axes de cette ellipse ;
- la droite verticale d'abscisse  $2\pi$  est une tangente de cette ellipse ;
- $\sqrt{\frac{2|e|}{1+|e|}}$  est l'excentricité de cette ellipse, avec  $e = \langle s'(t_0) | s'(t'_0) \rangle$  où  $t_0$  et  $t'_0$  sont à peu près au milieu des supports respectifs de  $b_1$  et de  $b_2$  ;
- l'ellipse passe par le point de coordonnées  $[\frac{1}{-e}]$  ;
- l'arc d'ellipse est délimité en abscisse et en ordonnée comme selon la propriété 5.3.

Lorsque les morceaux concernés de  $\partial S$  sont courbes, il faut ajouter à cet arc d'ellipse tous les points situés à une distance inférieure à

$$\pi \|s''\|_{\infty} m \left(1 + 2\sqrt{2} \frac{m}{d'}\right)$$

où  $m = \max(|\text{supp}b_1|, |\text{supp}b_2|)$  est la taille maximale du support de  $b_1$  ou de  $b_2$ .

■ Soit  $\hat{u}$  une version de  $u$  construite en linéarisant  $s$ : soit  $\hat{u} = 2\pi \|\hat{s}(t) - \bar{s}(t')\|$  où  $\hat{s}(t) = s(t_0) + (t - t_0)s'(t_0)$  et où  $\bar{s}(t') = s(t'_0) + (t' - t'_0)s'(t'_0)$ . Avant de montrer que  $\nabla \hat{u}$  décrit parfaitement l'ellipse décrite dans l'énoncé, cherchons à évaluer la différence entre  $\nabla u$  et  $\nabla \hat{u}$ . On a :

$$\begin{aligned} & \|\hat{s}(t) - s(t)\| \\ &= \|\hat{s}(t_0) + (t - t_0)\hat{s}'(t_0) - s(t_0) - (t - t_0)s'(t_0) - \int_{t_0}^t (t - t'')s''(t'')dt''\| \\ &\leq \|s''\|_{\infty} \frac{(t - t_0)^2}{2} \end{aligned}$$

donc  $\|\bar{s}(t') - s(t')\| \leq \|s''\|_{\infty} \frac{(t' - t'_0)^2}{2}$  et

$$\|\hat{u}(t, t') - u(t, t')\| \leq \pi \|s''\|_{\infty} ((t - t_0)^2 + (t' - t'_0)^2)$$

De plus,

$$\|\hat{s}'(t) - s'(t)\| = \|\hat{s}'(t_0) - s'(t_0) - \int_{t_0}^t s''(t'')dt''\| \leq \|s''\|_{\infty} \frac{|t - t_0|}{2}$$

et  $\|\bar{s}'(t') - s'(t')\| \leq \|s''\|_{\infty} \frac{|t' - t'_0|}{2}$ , et  $u(t, t')$  est minoré par  $u_0$ , donc

$$\begin{aligned} & \|\nabla u(t, t') - \nabla \hat{u}(t, t')\| \\ &= 4\pi^2 \left\| \begin{bmatrix} \langle s'(t) - \hat{s}'(t) | \frac{\hat{s}(t) - \bar{s}(t')}{\hat{u}(t, t')} \rangle \\ \langle -s'(t') + \bar{s}'(t') | \frac{\hat{s}(t) - \bar{s}(t')}{\hat{u}(t, t')} \rangle \end{bmatrix} + \right. \\ & \quad \left. \begin{bmatrix} \langle s'(t) | \frac{s(t) - \hat{s}(t)}{u(t, t')} - \frac{s(t') - \bar{s}(t')}{u(t, t')} + \frac{\hat{s}(t) - \bar{s}(t')}{\hat{u}(t, t')} \left( \frac{\hat{u}(t, t')}{u(t, t')} - 1 \right) \rangle \\ \langle -s'(t') | \frac{s(t) - \hat{s}(t)}{u(t, t')} - \frac{s(t') - \bar{s}(t')}{u(t, t')} + \frac{\hat{s}(t) - \bar{s}(t')}{\hat{u}(t, t')} \left( \frac{\hat{u}(t, t')}{u(t, t')} - 1 \right) \rangle \end{bmatrix} \right\| \\ &\leq 4\pi^2 \left\| \begin{bmatrix} \|s'(t) - \hat{s}'(t)\| \\ \|-s'(t') + \bar{s}'(t')\| \end{bmatrix} \right\| \times \frac{1}{2\pi} + \\ & \quad 4\pi^2 \left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\| \left( \left\| \frac{s(t) - \hat{s}(t)}{u(t, t')} \right\| + \left\| \frac{s(t') - \bar{s}(t')}{u(t, t')} \right\| + \frac{1}{2\pi} \frac{|\hat{u}(t, t') - u(t, t')|}{u(t, t')} \right) \\ &\leq \|s''\|_{\infty} \left( \pi \sqrt{(t - t_0)^2 + (t' - t'_0)^2} + 4\pi^2 \sqrt{2} \frac{(t - t_0)^2 + (t' - t'_0)^2}{u(t, t')} \right) \\ &\leq \pi \|s''\|_{\infty} m \left(1 + 2\sqrt{2} \frac{m}{d'}\right) \end{aligned} \tag{89p}$$

De plus, nous avons

$$\frac{1}{2\pi} \nabla \hat{u}(t, t') = \begin{bmatrix} \langle \hat{s}'(t) | \frac{\hat{s}(t) - \bar{s}(t')}{|\hat{s}(t) - \bar{s}(t')|} \rangle \\ \langle -\bar{s}'(t') | \frac{\hat{s}(t) - \bar{s}(t')}{|\hat{s}(t) - \bar{s}(t')|} \rangle \end{bmatrix} = \begin{bmatrix} \langle s'(t_0) | \frac{\hat{s}(t) - \bar{s}(t')}{|\hat{s}(t) - \bar{s}(t')|} \rangle \\ \langle -s'(t'_0) | \frac{\hat{s}(t) - \bar{s}(t')}{|\hat{s}(t) - \bar{s}(t')|} \rangle \end{bmatrix}$$

On voit que ce sont les produits scalaires d'un vecteur unitaire  $\frac{\hat{s}(t) - \bar{s}(t')}{|\hat{s}(t) - \bar{s}(t')|}$  avec deux vecteurs unitaires  $s'(t_0)$  et  $-s'(t'_0)$ . Ces deux vecteurs forment un angle dont le cosinus a été noté  $-e$ . Ainsi,  $\frac{1}{2\pi} \nabla u(t, t')$  est de la forme  $\begin{bmatrix} \cos(\gamma - \eta) \\ \cos(\gamma + \eta) \end{bmatrix}$ , avec  $\gamma$  dépendant de  $(t, t')$  et  $\eta$  tel que  $\cos(2\eta) = -e$ .

L'équation de l'ellipse décrite dans l'énoncé est

$$\left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial u}{\partial t'}\right)^2 + 2e \left(\frac{\partial u}{\partial t}\right) \left(\frac{\partial u}{\partial t'}\right) = 4\pi^2(1 - e^2) \quad (89q)$$

et en effet, on voit que les coordonnées de  $\frac{1}{2\pi} \nabla \hat{u}(t, t')$  vérifient :

$$\begin{aligned} & \frac{1}{4\pi^2} \left( \frac{\partial}{\partial t} \hat{u}(t, t') \right)^2 + \frac{\partial}{\partial t'} \hat{u}(t, t')^2 + 2e \frac{\partial}{\partial t} \hat{u}(t, t') \frac{\partial}{\partial t'} \hat{u}(t, t') \\ &= \cos(\gamma - \eta)^2 + \cos(\gamma + \eta)^2 + 2e \cos(\gamma - \eta) \cos(\gamma + \eta) \\ &= 2 \cos^2 \gamma \cos^2 \eta + 2 \sin^2 \gamma \sin^2 \eta - 2 \cos(2\eta) (\cos^2 \gamma \cos^2 \eta - \sin^2 \gamma \sin^2 \eta) \\ &= 2 \cos^2 \gamma \cos^2 \eta + 2(1 - \cos^2 \gamma) \sin^2 \eta - 2 \cos(2\eta) (\cos^2 \gamma \cos^2 \eta + (\cos^2 \gamma - 1) \sin^2 \eta) \\ &= \cos^2 \gamma [2 \cos^2 \eta - 2 \sin^2 \eta - 2 \cos(2\eta) (\cos^2 \eta + \sin^2 \eta)] + [2 \sin^2 \eta + 2 \cos(2\eta) \sin^2 \eta] \\ &= 0 + 2 \sin^2 \eta [1 + \cos(2\eta)] = 1 - \cos(2\eta)^2 = 1 - e^2 \end{aligned}$$

■

La propriété 5.2 ne permet pas d'expliquer le contenu des blocs de la matrice de la figure 65 ligne  $l$  colonne  $c$  lorsque  $l$  et  $c$  sont trop proches, car la distance  $d'$  qui intervient dans la propriété 5.2 devient nulle. Nous devons alors recourir à la propriété 4.3. Voici une propriété donnant un minorant sous forme close de la constante  $d_+$  qui intervenait dans la propriété 4.3 :

**Propriété 5.5** *Dans la propriété 4.3, le nombre*

$$d_+ = \min_{1 \leq q \leq 2} \min_{t, t'} \left| k_+^2 - \left( \frac{\partial u}{\partial t} + \frac{\partial u}{\partial t'} \right)^2 \right|^q$$

est tel que

$$\left( \frac{\partial u}{\partial t} + \frac{\partial u}{\partial t'} \right) \leq \pi m^2 \|s''\|_\infty^2, \quad (89r)$$

avec  $m$  la longueur du support de  $b_1^2 + b_2^2$ , sur laquelle nous supposons que  $m \|s''\|_\infty \leq 2$ .

■ Reprenons la notation  $\hat{u}$  vue dans la démonstration de la propriété 4.3. Montrons que

$$|\partial_+ \bar{u}| \leq \pi m^2 \|s''\|_\infty^2. \quad (89s)$$

Or

$$\begin{aligned} \frac{\bar{u}(t, t') \partial_+ \bar{u}(t, t')}{4\pi^2} &= \langle s'(t) - s'(t') | s(t) - s(t') \rangle = \int_t^{t'} \int_t^{t'} \langle s''(a) | s'(b) \rangle da db \\ &= \int_t^{t'} \int_t^{t'} \langle s''(a) | s'(a) \rangle da db + \int_t^{t'} \int_t^{t'} \int_a^b \langle s''(a) | s''(c) \rangle dc da db \\ &\leq \int_t^{t'} \int_t^{t'} 0 da db + \|s''\|_\infty^2 \left| \int_t^{t'} \int_t^{t'} \int_a^b da db dc \right| \leq \|s''\|_\infty^2 \frac{1}{3} |t - t'|^3 \end{aligned}$$

De plus,

$$\begin{aligned} (t' - t)^2 - \frac{\bar{u}(t, t')^2}{4\pi^2} &= \int_t^{t'} \int_t^{t'} \langle s'(a) | s'(a) - s'(b) \rangle da db = \int_t^{t'} \int_t^{t'} \int_a^b \langle s'(a) | s''(c) \rangle dc da db \\ &= \int_t^{t'} \int_t^{t'} \int_a^b \langle s'(c) | s''(c) \rangle dc da db - \int_t^{t'} \int_t^{t'} \int_a^b \int_a^c \langle s''(e) | s''(c) \rangle de dc da db \\ &\leq \int_t^{t'} \int_t^{t'} \int_a^b 0 dc da db + \|s''\|_\infty^2 \left| \int_t^{t'} \int_t^{t'} \int_a^b \int_a^c de dc da db \right| = \|s''\|_\infty^2 \frac{1}{12} (t - t')^4 \end{aligned} \quad (89t)$$

Et donc, lorsque  $m \leq \frac{2}{\|s''\|_\infty}$ , on a  $(t' - t)^2 - \frac{\bar{u}(t, t')^2}{4\pi^2} \leq \frac{1}{3} (t' - t)^2$ , donc  $\bar{u}(t, t') > \frac{4\pi}{3} |t - t'|$ , et donc  $\partial_+ \bar{u}(t, t') < \pi \|s''\|_\infty^2 |t - t'|^2$ . ■

Nous sommes maintenant en mesure de prédire exactement la position des coefficients importants dans la matrice permutée de la figure 65, et donc par conséquent la position des coefficients importants dans toute matrice de rigidité  $M$  faite dans une base de paquets d'ondelettes, à ceci près que nous devons être capables de valider la définition de  $k_1$ ,  $k_2$ ,  $b_1$  et  $b_2$  dans (89c). Nous allons faire cette validation dans la partie suivante; cette validation expliquera aussi la présence de coefficients, plus petits mais pas si petits, dans des endroits non prévus par les trois propriétés ci-dessus, dans la figure 63.

### 5.3.3 la localisation du Pic en Fourier des paquets d'ondelettes

Nous devons maintenant expliquer, comment, dans un coefficient ligne  $l$  colonne  $c$  d'un bloc ligne  $g$  colonne  $\kappa$  de la matrice de rigidité  $M$  (85) résultant de la discrétisation du problème de Helmholtz intégral (22), il peut



exister des fonctions  $b_1$  et  $b_2$  de supports de longueur  $O(2^{j'})$  et centrés approximativement en  $2^{-j'}c$  et en  $2^{-j'}l$  respectivement, et des fréquences  $k_1$  et  $k_2$  telles que

$$b_1(t)e^{2i\pi\frac{k_1}{\lambda_0}t} \quad \text{et} \quad b_2(t')e^{2i\pi\frac{k_2}{\lambda_0}t}$$

modélise suffisamment bien les deux fonctions intervenant dans la définition (86) de ce coefficient, avec  $b_1$  et  $b_2$  indépendants de  $g$  et de  $\kappa$ .

Fixons  $c$  et  $l$ , nous travaillons dans la matrice d'un système  $M$  dans une base de paquets d'ondelettes du type

$$\{\psi_{j',m}^{(n)}, m \in \mathbb{Z}/L2^{j'}\mathbb{Z}, n \in \{0,1,\dots,2^{j-j'}-1\}\} \quad (89u)$$

qui est le cas général de la base  $V_h$  (67a) approchant une famille de Fourier à fenêtre.

**Propriété 5.6** *Il existe deux fonctions  $b_1$  et  $b_2$  ne dépendant que de  $c$  et de  $l$  telles que les paquets d'ondelettes*

$$\psi_{j',c-1}^{(\kappa-1)} \quad \text{et} \quad \psi_{j',l-1}^{(g-1)} \quad (89v)$$

paramétrés par

$$l, c \in \{1, 2, \dots, L2^{j'}\} \quad \text{et} \quad \kappa, g \in \{1, 2, \dots, 2^{j-j'}\}$$

soient de la forme

$$\frac{2^{-j'}}{2\pi} \int_{-\infty}^{\infty} e^{-2^{-j'}\xi c} \hat{\psi}_{0,0}^{(\kappa-1)}(2^{-j'}\xi) e^{i\xi t} b_1(t) d\xi$$

et

$$\frac{2^{-j'}}{2\pi} \int_{-\infty}^{\infty} e^{-2^{-j'}\xi l} \hat{\psi}_{0,0}^{(g-1)}(2^{-j'}\xi) e^{i\xi t} b_1(t') d\xi.$$

L'entier  $j$  est donné par le pas de discrétisation  $h = 2^{-j}$ . Ces deux fonctions  $b_1$  et  $b_2$  ont une norme  $L^1$  en  $O(2^{-j'})$ , elles ont leur support séparé de

$$d' = 2^{-1-j'} \min(|c-l|, |c-l| - L2^{j'}) - 2^{-j'}(m_0 + m_1) \quad (89w)$$

où  $m_0 + m_1$  défini en (65f) est la longueur du filtre utilisé; les fonctions  $b_1$  et  $b_2$  ont une norme  $W^{p,1}$  avec  $p$  entier positif non nul en

$$\left(\frac{1}{2} \min(|c-l|, |c-l| - L2^{j'}) - (m_0 + m_1)\right)^{1-p}$$

Les deux paquets d'ondelettes (89v) sont ceux qui interviennent dans le coefficient ligne  $l$  colonne  $c$  du bloc ligne  $g$  colonne  $\kappa$  de la matrice de rigidité  $M$  dans la base (89u).

■ Nous savons que le support de  $\psi_{0,0}^{(n)}$  est contenu dans  $[-m_0 - m_1, m_0 + m_1]$ , car la longueur de leur support est de  $m_0 + m_1$  et leur support contient 0 (démonstration par récurrence sur  $\frac{\log n}{\log 2}$ ). Ainsi (89w) est la moitié de la distance minimale entre les supports des paquets d'ondelettes (89v), qui sont

$$s_1 = 2^{-j'}[c - m_0 - m_1, c + m_0 + m_1] \quad \text{et} \quad s_2 = 2^{-j'}[l - m_0 - m_1, l + m_0 + m_1].$$

Nous pouvons donc allonger les deux intervalles ci-dessus en les rallongeant à gauche et à droite de  $d'/2$ , ce qui donne deux nouveaux intervalles  $s'_1$  et  $s'_2$ . Ensuite, nous pouvons trouver deux fonctions  $b_1$  et  $b_2$  qui ont pour support  $s'_*$  et qui valent 1 sur  $s_*$ , qui satisfont aux propriétés annoncées ci-dessus dans l'énoncé. Enfin, nous multiplions la formule d'inversion de Fourier des fonctions (89v) par  $b_1$  et  $b_2$  respectivement. ■

Il faut maintenant expliquer le remplacement 66. La figure 68 montre la restriction à  $[0, 2^4\pi]$  des transformées de Fourier de  $\psi_{0,0}^{(n)}$  pour  $n$  allant de 0 (première ligne) à 7. Nous avons déjà vu ces graphes, organisés non en niveaux de gris mais plus traditionnellement, en (67c), sans la restriction à  $[0, 2^4\pi]$ .

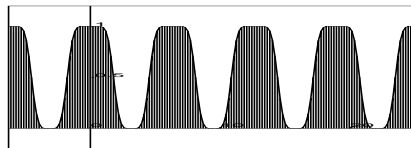
La figure 69 présente, avec la même notation, ces quatre fonctions sans le remplacement 66 : nous voyons bien que ce remplacement est nécessaire pour établir l'ordre correct fréquentiel des paquets d'ondelettes.

D'après  $\hat{g}$  défini en (82e), la définition faite en (67a) de la fonction  $\psi_{*,*}^{(*)}$  montre qu'on a la relation de récurrence suivante, pour  $n$  un entier :

$$\begin{aligned} \hat{\psi}_{0,0}^{(4n)}(2\xi) &= \hat{g}(\xi)\hat{\psi}_{0,0}^{(2n)}(2\xi) \\ \hat{\psi}_{0,0}^{(4n+1)}(2\xi) &= \hat{g}(\xi + \pi)\hat{\psi}_{0,0}^{(2n)}(2\xi) \\ \hat{\psi}_{0,0}^{(4n+2)}(2\xi) &= \hat{g}(\xi + \pi)\hat{\psi}_{0,0}^{(2n+1)}(2\xi) \\ \hat{\psi}_{0,0}^{(4n+3)}(2\xi) &= \hat{g}(\xi)\hat{\psi}_{0,0}^{(2n+1)}(2\xi) \end{aligned} \quad (89x)$$

et oublier la permutation (66) revient à échanger  $4n + 2$  et  $4n + 3$  ci-dessus.

Seulement,  $\hat{g}$  est en fait, pour l'ondelette de Daubechies de longueur 16 :



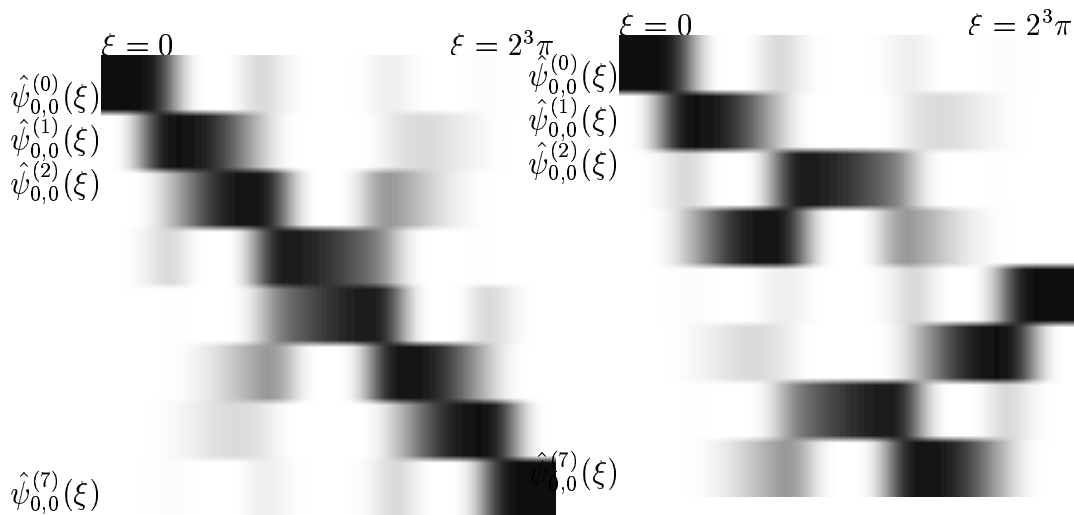


FIG. 68 – *Transformée de Fourier des premiers paquets d'ondelettes.*

FIG. 69 – *Transformée de Fourier des premiers paquets d'ondelettes, si on oublie le remplacement (66).*



Ce qui est proche de  $m_A$ , fonction indicatrice de  $[-\frac{\pi}{2}, \frac{\pi}{2}] + 2\pi\mathbb{Z}$  dessinée ici :

$$\begin{array}{c}
 m_A(\xi) \\
 \uparrow \\
 \begin{array}{c}
 \square \quad \square \quad \square \\
 -2\pi \quad \quad 2\pi \quad \quad \xi
 \end{array}
 \end{array}
 \tag{90}$$

Cette proximité nous permet de montrer que les fonctions  $\psi_{0,0}^{(*)}$  sont ordonnées correctement en fréquence :

**Propriété 5.7** Si  $\hat{g} = m_A$ ,  $\hat{\psi}_{0,0}^{(n)}$  est la fonction indicatrice de  $[-(n+1)\pi, -n\pi] \cup [n\pi, (n+1)\pi]$ .

■ Toutes les fonctions en jeu sont paires, donc on peut se restreindre à  $\mathbb{R}_+$ . Montrons d'abord que le support de  $\mathcal{F}(\psi^n)$  est de la forme  $[m\pi, (m+1)\pi]$  avec  $m \in \mathbb{N}$ , par récurrence sur  $n$ .

**Cas  $n = 0$  :**

Nous avons alors

$$\hat{\psi}_{0,0}^{(0)}\xi = \prod_{j \in \mathbb{N}} \hat{g}(\xi 2^{-j-1}).$$

Soit  $\xi \in \mathbb{R}_+$ . Si  $\xi < \frac{\pi}{2}$ , on voit bien que  $\hat{\psi}_{0,0}^{(0)}$  est un produit valant 1. Sinon, soit  $j$  l'entier tel que  $2^j < \frac{\xi}{\pi} \leq 2^{j+1}$  : le terme d'indice  $j$  du produit ci-dessus est nul.

**Autre cas :**

Soit  $n'$  l'entier immédiatement inférieur à  $\frac{2n+1}{4}$  :  $n' < n$ , donc l'hypothèse de récurrence donne la propriété pour  $n'$  correspondant :  $\xi \mapsto \mathcal{F}(\psi^{n'})\left(\frac{\xi}{2}\right)$  a pour support  $[2n'\pi, 2(n'+1)\pi]$ . Sur ce support et si  $n'$  est pair,  $\xi \mapsto \hat{g}\left(\frac{\xi}{2}\right)$  vaut 1 sur  $[2n'\pi, (2n'+1)\pi]$  et 0 sur  $[(2n'+1)\pi, 2(n'+1)\pi]$ , et c'est l'inverse pour  $\xi \mapsto \hat{g}\left(\frac{\xi}{2} + \pi\right)$  : donc (89x) permet de conclure. Si  $n'$  est impair,  $\xi \mapsto \hat{g}\left(\frac{\xi}{2}\right)$  vaut 0 sur  $[2n'\pi, (2n'+1)\pi]$  et 1 sur  $[(2n'+1)\pi, 2(n'+1)\pi]$ , et c'est l'inverse pour  $\xi \mapsto \hat{g}\left(\frac{\xi}{2} + \pi\right)$  : donc (89x) permet ici aussi de conclure. ■

Supposons un instant que la fonction  $2\pi$ -périodique  $\hat{g}$  soit la fonction indicatrice  $m_B$  de l'ensemble  $[0, \frac{\pi}{2}] + \pi\mathbb{Z}$  :



nous voyons dans (89x) que prendre cette valeur  $m_B$  au lieu de  $m_A$  pour  $\hat{g}$  consiste précisément à oublier de faire le remplacement 66 ;

**Remarque 5.8** Si  $\hat{g} = m_B$ ,  $\mathcal{F}(\psi^n)$  est la fonction indicatrice de  $[-(p(n) + 1)\pi, -p(n)\pi] \cup [p(n)\pi, (p(n) + 1)\pi]$  où  $p(n)$  est un entier,  $p$  étant une permutation de Frank Gray.

Voici les premières valeurs de cette permutation de Frank Gray :

$n$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$p(n)$	0	1	3	2	6	7	5	4	12	13	15	14	10	11	9	8	24	25	27	26	30
$p^{-1}(n)$	0	1	3	2	7	6	4	5	15	14	12	13	8	9	11	10	31	30	28	29	24

Cette permutation est connue sous le nom de permutation Binary Reflected Gray Code (BRGC).

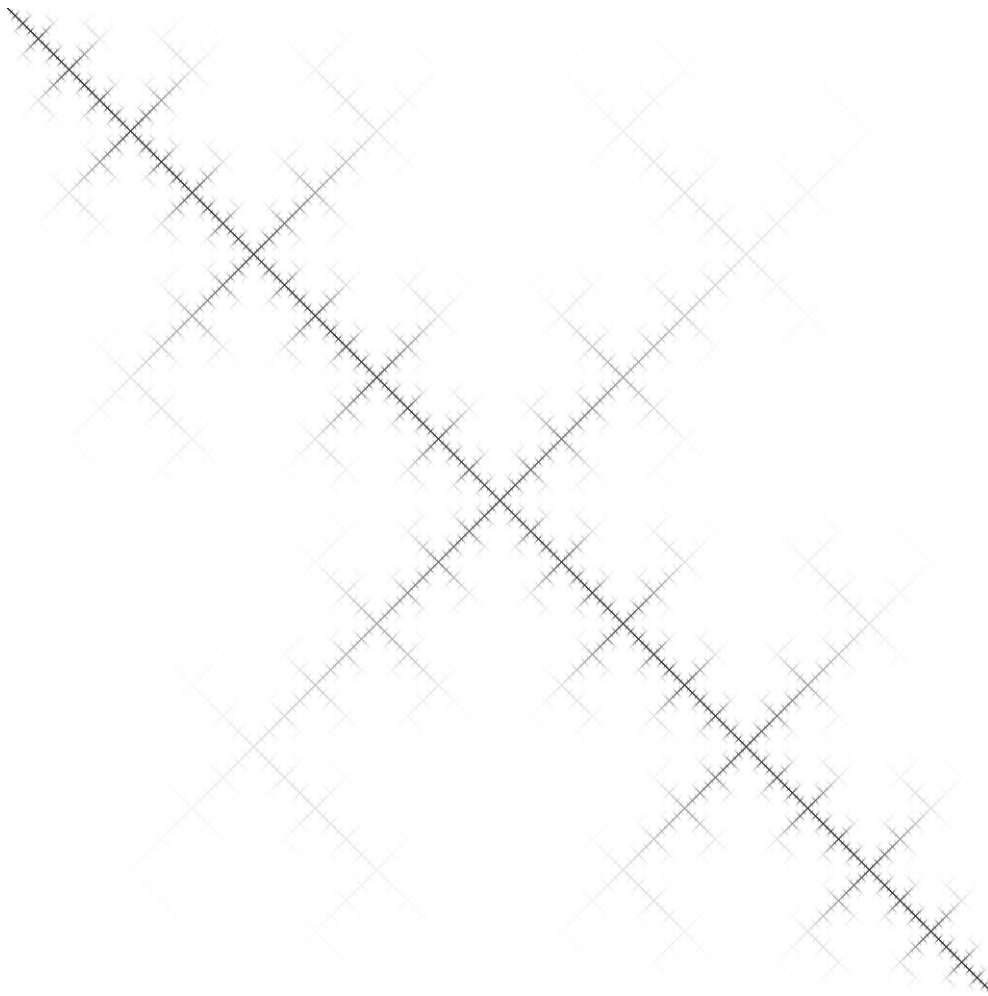


FIG. 70 – *Transformée de Fourier de plus de paquets d'ondelettes.*



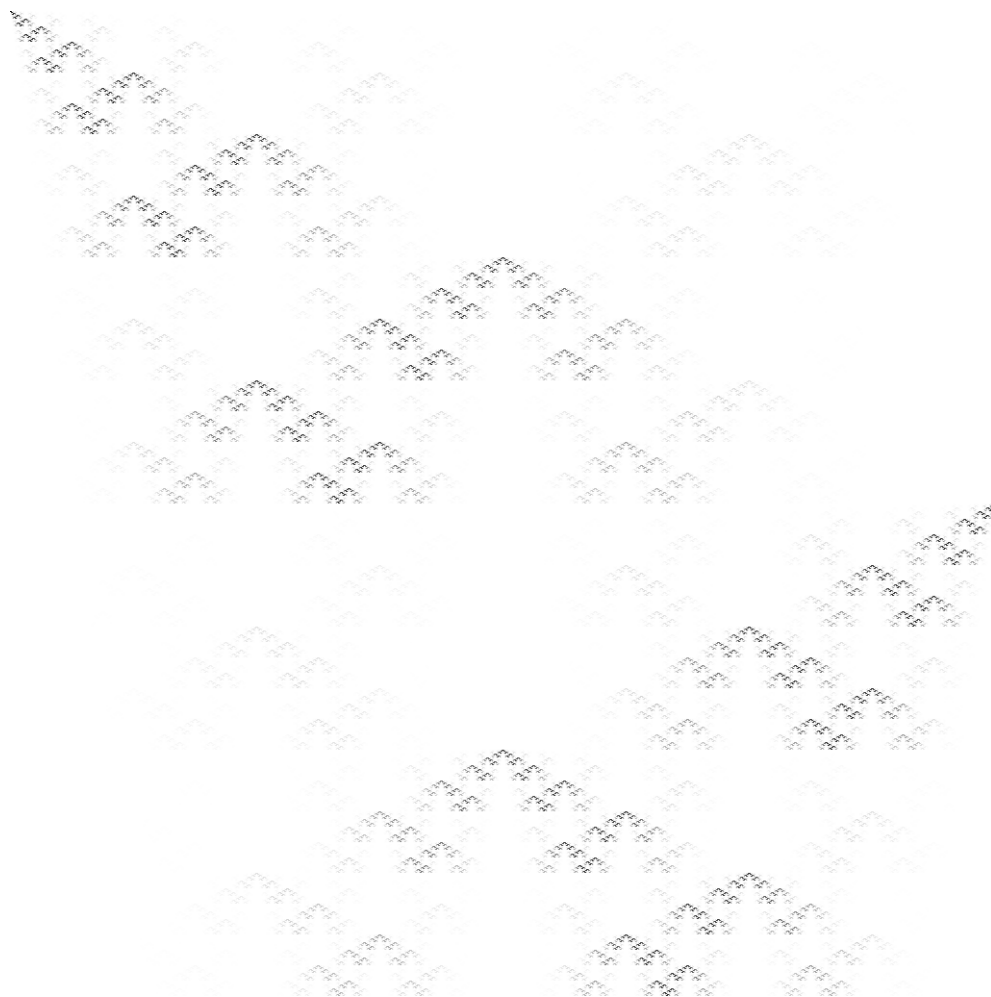


FIG. 71 – *Transformée de Fourier de plus de paquets d'ondelettes si on oublie le remplacement 66.*



Nous comprenons maintenant que le remplacement (66) est précisément lié à la forme pas tout à fait régulière de  $m_A$  et donc de  $\hat{g}$ .

Les figures 70 et 71 présentent une généralisation des figures 68 et 69 :

L'étude de la figure 70 montre que le graphe de  $\psi_{0,0}^{(m2^i)}$  avec  $m$  entier impair et  $i$  entier est un axe horizontal de symétrie locale de cette figure au voisinage de sa fréquence principale d'autant plus respecté que  $i$  est grand, ce que nous allons en partie montrer ci-dessous.

**Propriété 5.9** *S'il existe un entier  $m$  impair et un entier  $i$  tel que deux entiers  $n$  et  $n'$  sont entre  $m2^i$  et  $(m+1)2^i$ , et tels que  $n+n' = (2m+1)2^i$ , alors  $\hat{\psi}_{0,0}^{(n)}$  peut être grand vers la fréquence dominante de  $\hat{\psi}_{0,0}^{(n')}$ , très marquée si  $\frac{n'-n}{2^i}$  est plus petit que  $\frac{1}{2}$ .*

■ L'analyse de (89x) montre que  $\psi_{0,0}^{(n)}$  et  $\psi_{0,0}^{(n')}$  sont définis par les produits

$$\hat{\psi}_{0,0}^{(n)}\xi = \prod_{j \in \mathbb{N}} \hat{g}((\xi + a_j\pi)2^{-j-1}).$$

et

$$\hat{\psi}_{0,0}^{(n')}\xi = \prod_{j \in \mathbb{N}} \hat{g}((\xi + a'_j\pi)2^{-j-1}).$$

avec  $a_j$  et  $a'_j$  deux suites à valeur dans  $\{0,1\}$ ; la contrainte  $n+n' = (2m+1)2^i$  montre (par récurrence sur  $n - 2^i m$ ) que  $a_j = a'_j$  sauf pour  $j = k$  avec  $k$  vérifiant

$$2^{i-k} < m < 2^{i-k+1}.$$

Ainsi, ces deux produits sont identiques sauf que  $\hat{g}((\xi + \pi)2^{-k-1})$  est à changer en  $\hat{g}(\xi 2^{-k-1})$ ; or nous voyons sur la représentation de  $\hat{g}$  ci-dessus que ces deux fonctions sont sensiblement non-nulles en  $\pi(n + \frac{1}{2})$  et en  $\pi(n' + \frac{1}{2})$ , surtout si  $\frac{n'-n}{2^i}$  est plus petit que  $\frac{1}{2}$ ; de plus  $\hat{g}(\xi)/\hat{g}(\xi + \pi)$  est entre .1 et 10. sur  $[\frac{3}{8}\pi, \frac{5}{8}\pi]$ : ainsi,  $\psi_{0,0}^{(n)}/\psi_{0,0}^{(n')}$  est entre .1 et 10. en  $\pi(n + \frac{1}{2})$  et en  $\pi(n' + \frac{1}{2})$  si  $\frac{n'-n}{2^i}$ . Comme nous avons suggéré à la propriété 5.7 que  $\psi_{0,0}^{(n')}(\pi(n' + \frac{1}{2}))$  était non négligeable, nous voyons que  $\psi_{0,0}^{(n)}(\pi(n' + \frac{1}{2}))$  est également non négligeable, et même ces deux fonctions ont à peu près la même transformée de Fourier sur la bande de fréquence  $[\frac{8m+3}{8}2^i, \frac{8m+5}{8}2^i]$  puisque leur rapport est entre .1 et 10., ce qui explique que la représentation de la transformée de Fourier de  $\psi_{0,0}^{(m2^i)}$  avec  $m$  entier impair et  $i$  entier soit un axe horizontal de symétrie locale de cette figure d'autant plus respecté que  $i$  est grand. ■

### 5.3.4 Une approche numériquement économique et mathématiquement justifiée.

Comme précisé au paragraphe 3.2.1, il est tentant, lorsque le diffractant n'a pas de coins, d'éviter les doubles intégrales, c'est à dire d'éviter d'avoir dans  $V_h$  d'autres fonctions que des masses de Dirac.

Voici une approche que nous proposons : au lieu de discrétiser le problème numérique (24), nous observons que dans la discrétisation en Fourier justifiée par le théorème 4.5, il est possible d'approcher les termes de la matrice, avec la tolérance  $\varepsilon$  autorisée par le théorème 5.1.

Pour les approcher, nous pouvons mettre en évidence l'opérateur  $\hat{A}$  du cas du diffractant circulaire, comme dans les démonstrations des 2.14, 2.28, et de la propriété 4.4, et ne discrétiser que la différence entre l'opérateur  $A$  du cas d'un diffractant sans coin et l'opérateur  $A$  du cas du diffractant circulaire. Cette différence a pour intérêt que le noyau  $\frac{1}{\lambda_0}(\hat{Y} - Y)$  se comporte bien mieux que  $A$  (mais il faut tout de même  $h = \lambda_0^{1+\frac{1}{2p}}$  pour un certain  $p$  étant l'ordre de la méthode d'intégration, et que la matrice de l'opérateur  $\hat{A}$  est diagonale et ces coefficients diagonaux sont donnés par la formule close (71).

Nous pouvons même rendre cette matrice seuillable : divisons à gauche et à droite cette matrice par la matrice diagonale contenant  $\int e_m e_m^{(j)}$  à la  $m$ -ième ligne, ce qui fait un nouvel opérateur  $\hat{A}$ , continu, voir (82c). Discrétisons ce nouvel opérateur en ondelettes : ce tour de passe-passe fait qu'il suffit d'une FFT bidimensionnelle pour calculer la matrice du système résultant de cette nouvelle discrétisation (qui n'introduit exceptionnellement pas d'erreur). Ensuite, comme cet opérateur  $\hat{A}$  est aussi bien fait que  $A$ , l'analogie des propriétés 4.3 et 5.2 sont aussi valides pour  $\hat{A}$ .

Ainsi, nous obtenons une nouvelle méthode mathématiquement justifiée, et qui devrait avoir une très bonne complexité, surtout en profitant du seuillage prédictif construit à la partie précédente. Cette méthode n'a pas été mise en œuvre dans ce manuscrit de thèse, mais devrait être essayée d'ici la soutenance.

Comme le filtre en ondelette peut en fait fonctionner progressivement, on peut diminuer encore la quantité mémoire utilisée. Par rapport à une méthode calculant la première matrice dans la famille de Fourier.

Ce serait la première méthode rapide et justifiée asymptotiquement lorsque  $\lambda_0 \rightarrow 0$ , qui viendrait faire concurrence aux méthodes multipôles qui, Amini-Profit [AP00] en témoigne, n'ont pas encore de théorème analogue au théorème 4.8.



## 6 Conclusion

Ainsi, nous avons fait la comparaison et la justification technique de deux approches pour résoudre les problèmes de Helmholtz.

En résultat, nous obtenons un moyen de prévoir rapidement quel est, par exemple, le profil radar d'un objet à section connue, c'est à dire le problème direct. Il semble intéressant de savoir si, à profil radar connu, il est possible de déduire la section ; Kress [K89] explique à quel point ce problème est mal posé ; toutefois, c'est un domaine de recherche très actif, où on trouve par exemple Badia et Ha-Duong [EH01].

Les résultats de la section 4.3 confirment l'heuristique « prendre  $N$  degrés de liberté par longueur d'onde » en se basant sur une analyse en Fourier, en prenant particulièrement compte de ce qui se passe asymptotiquement lorsque la longueur d'onde tend vers 0. Ils peuvent être généralisés dans un contexte tridimensionnel ; l'équation différentielle analogue de (29c) est plus simple car l'harmonique cylindrique  $H_0^{ank}$  est remplacée par son analogue sphérique  $r \mapsto \frac{e^{ir}}{r}$  qui vérifie une équation différentielle de degré plus petit.

Lorsqu'il y a des coins, Djaoua [D81], Costabel [C83], Costabel-Stephan [CS85] [CS87] [CS93], Stephan-Suri [SS89], Elschner [E93], Kress-Sloan [KS93] ou Schneider [S93] [S03] proposent de prendre des espaces de discrétisation qui ne sont plus invariants par translation. Pour pouvoir continuer à travailler avec des paquets d'ondelettes afin d'avoir une bonne analyse fréquentielle intéressante avec  $\lambda_0$  petit, nous pourrions ajouter une contrainte de nullité en les coins des fonctions dans  $V_h$ , puis d'ajouter à ceci un vecteur propre du cas d'un angle à bords droits infini ou du cas du carré, s'il est analytiquement résoluble.

### 6.1 Annexe

**Propriété 6.1**  $H_m^{ank'} \mathcal{R}e H_m^{ank} - H_m^{ank} \mathcal{R}e H_m^{ank'} = \frac{2i}{\pi r}$

■ Soit  $H = H_m^{ank}$ , et  $Y = H' \mathcal{R}e H - H \mathcal{R}e H'$  le membre droit. On a

$$\begin{aligned} Y' &= H'' \mathcal{R}e H - H \mathcal{R}e H'' + H' \mathcal{R}e H' - H' \mathcal{R}e H' \\ &= -\left(\frac{1}{r} H' + \left(1 - \frac{m^2}{r^2}\right) H\right) \mathcal{R}e H + H \mathcal{R}e \left(\frac{1}{r} H' + \left(1 - \frac{m^2}{r^2}\right) H\right) + 0 - 0 \\ &= -\frac{1}{r} Y \end{aligned}$$

et donc  $Y = \frac{C}{r}$  avec  $C$  une constante. Recherchons  $C$ . On a  $\mathcal{R}eY = 0$ ; de plus,

$$r\mathcal{I}mY = r\mathcal{I}mH'\mathcal{R}eH - r\mathcal{I}mH\mathcal{R}eH' = r|H|^2\mathcal{I}m\frac{H'}{H}$$

qui, d'après (42) et (43), tend vers  $\frac{2}{\pi}$ : CQFD. ■

## 6.2 Bibliographie

- [ABC93] B. Alpert, G. Beylkin, R. Coifman et V. Rokhlin, Wavelet-like bases for the fast solution of second-kind integral equations, *SIAM J. Sci. Comput.* **14** 159-184, 1993.
- [A93] S. Amini, Boundary integral solution of the exterior Helmholtz problem, *Comp. Mech.* **13** 2-11, 1993.
- [A99] S. Amini, On boundary integral operators for the Laplace and the Helmholtz equations and their discretisations, *Eng. Anal. Bound. Elem.* **23** 4, pp. 327-337, 1999.
- [AK95] S. Amini and S. M. Kirkup, Solution of Helmholtz equation in the exterior domain by elementary boundary integral methods, *J. of Comput. Phys.* **118** 208-221, 1995.
- [AP00] S. Amini and A. Profit, Analysis of the truncation errors in the fast multipole method for scattering problems. Proceedings of the 8th International Congress on Computational and Applied Mathematics, ICCAM-98 (Leuven), *J. Comput. Appl. Math.* **115** 1-2 23-33, 2000.
- [AW83] D. N. Arnold and W. L. Wendland, On the asymptotic convergence of collocation methods, *Mathematics of Computation* **41** 164 349-381, 1983.
- [ABC00] A. Averbuch, E. Braverman, R. Coifman, M. Israeli and A. Sidi, Efficient computation of oscillatory integrals via adaptive multiscale local Fourier bases, *A. C. H. A.* **9** 19-53, 2000.
- [BL96] Z. Baharav and Y. Leviatan, Impedance matrix compression using adaptively constructed basis functions, *IEEE Trans. Antennas* **44** 1231-1238, 1996.
- [B89] C. A. Balanis, *Advanced engineering electromagnetics*, New York : Wiley, 1989.
- [B81] A. Bamberger, *Analyse numérique des équations hyperboliques linéaires*, photocopié de cours, École Polytechnique, 1983.
- [BH87] A. Bamberger and T. Ha Duong, Diffraction d'une onde acoustique par une paroi absorbante: nouvelles équations intégrales, *Math. Meth. in the Appl. Sci.* **9** 431-454, 1987.
- [BCK02] J.-D. Benamou, F. Castella, T. Katsaounis and B. Perthame, High frequency limit of the Helmholtz Equations, *Rev. Mat. Iberoamericana* **18** 1 187-209, 2002.
- [BCR91] G. Beylkin, R. Coifman et R. Rokhlin, A fast algorithm for the evaluation of Legendre expansions, *SIAM J. Sci. Statis. Comput.* **12** 158-179, 1991.
- [BBJ94] E. Bleszynski, M. Bleszynski and T. Jaroszewicz, Fast integral-equation solver for electromagnetic scattering problems, *IEEE AP-S Int.*

*Symp.* digest, Seattle, WA, June, 416-419, 1994.

[BCG93] B. Bradie, R. Coifman et A. Grossmann, Fast Numerical Computations of Oscillatory Integrals related to Acoustic Scattering I, *A. C. H. A.* **1** (1993), pp 94-99.

[BW65] H. Brakhage und P. Werner, Über das Dirichletsche Außenraumproblem für die Helmholtzsche Schwingungsgleichung, *Arch. Math.* **16** 325-329, 1965.

[BCS02] A. Buffa, M. Costabel and C. Schwab, Boundary element methods for Maxwell's equations on non-smooth domains, *Numer. Math.* **92** no. 4, 679-710, 2002.

[BM71] A. J. Burton and G. F. Miller, The application of integral equation methods to the numerical solution of some exterior boundary-value problems, *Proc. Roy. Soc. Lond. A* **323** 201-210, 1971.

[C90] F. X. Canning, The impedance matrix localization (IML) method for moment-method calculations, *IEEE trans antennas* **32** 18-30, 1990.

[C91] F. X. Canning, Interaction matrix localization (IML) permits solution of larger scattering problems, *IEEE Trans. Magn.* **27** 4275-4278, 1991.

[C92a] F. X. Canning, Sparse approximation for solving integral equations with oscillatory kernels, *SIAM J. Sci. Stat. Comput.* **13** 71-87, 1992.

[C93a] F. X. Canning, Preconditioned Iterative Solution of IML/Moment Method Problems, *IEEE Trans. Magn.* **29** 1946-1949, 1993.

[C93b] F. X. Canning, Improved impedance matrix localization method, *IEEE Trans. Antennas* **41** 659-667, 1993.

[C94] F. X. Canning, Fast integral solutions using geometrical-theory-of-diffraction-like matrices, *Radio Science* **29** 993-1008, 1994.

[C95] F. X. Canning, Solution of impedance matrix localization form of moment method problems in five iterations, *Radio Science* **30** 1371-1384, 1994.

[C96] F. X. Canning, *Fast solvers for wave equations*, Smith et Al, vol. V. of ordinary and partial differential equations, proceedings of the 13th Dundee conference, June 25-28, 1996.

[CR03] F. X. Canning and K. Rogovin, A universal matrix solver for integral-equation-based problems, *IEEE Trans. Antennas* **45** 1239-1245, 2003.

[CR98] F. X. Canning and K. Rogovin, Fast direct solution of standard moment-method matrices, *IEEE antennas* **40** 15-25, 1998.

[CS96] F. X. Canning and J. F. Scholl, diagonal preconditioners for the EFIE using a wavelet basis, *IEEE Trans. Inform. Theory* **44** 1239-1245, 1996.

[C84a] G. A. Chandler, Numerical analysis of the boundary integral method, Mathematical Programming and Numerical Analysis Workshop, S. A. Gustafson and R. S. Womersley, *Proceedings of the centre for mathematical analysis, Australian national university* **6**, 211-230 1984.

- [C84b] G. A. Chandler, Galerkin's method for boundary integral equations on polygonal domains, *J. Austral. Math. Soc. Ser. B* **26** 1-13, 1984.
- [CS90] G. A. Chandler and I. H. Sloan, Spline qualocation methods for boundary integral equations, *Numer. Math.* **58** 537-567, 1990.
- [CP99] H.-L. Chen et S.-L. Peng, A quasi-wavelet algorithm for second kind boundary integral equations, *Adv. Comp. Math.* **11** 355-375, 1999.
- [CXZ99] Z. Chen, Y. Xu and J. Zhao, The discrete Petrov-Galerkin method for weakly singular integral equations, *J. Integ. Equ. Appl.* **11** 1-35, 1999.
- [C92b] C. K. Chui, On cardinal spline-wavelets, *Wavelets and their application*, M. B. Riskai, Ed MA: Jones Bartlet Publ., 453-470, 1992.
- [CD93] A. Cohen and I. Daubechies, On the instability of arbitrary biorthogonal wavelet packets, *SIAM J. math. anal* **24** 1340-1354, 1993.
- [CW92] R. R. Coifman and M. V. Wickerhauser, Entropy-based algorithms for best basis selection, *IEEE Trans. Inform. Theory* **38** 713-718, 1992.
- [CM91] R. R. Coifman et Y. Meyer, Remarques sur l'analyse de Fourier à fenêtre, *C. R. Acad. Sci. Paris I* **312** 259-261, 1991.
- [CMM82] R. R. Coifman, A. McIntosh and Y. Meyer, L'intégrale de Cauchy définit un opérateur borné sur  $L^2$  pour les courbes Lipschitziennes, *Ann. of Math.* **116** 361-387, 1982.
- [CMW92] R. R. Coifman, Y. Meyer and V. Wickerhauser, Size properties of wavelet packets, *Wavelets and their application*, M. B. Riskai, Ed MA: Jones bartlet Publ., 453-470, 1992.
- [CRW93] R. Coifman, V. Rokhlin, and S. Wandzura, The fast multipole method for the wave equation: a pedestrian prescription, *IEEE Trans. Antennas* **35** 7-12, 1993.
- [CK83] D. Colton and R. Kress, *Integral equation methods in scattering theory*, John Wiley & Sons, New York, 1983.
- [CK92] D. Colton and R. Kress, *Inverse acoustic and electromagnetic scattering theory*, Springer-Verlag, 1992.
- [C83] M. Costabel, Boundary integral operators on curved polygons, *Ann. Mat. Pura Appl.* **4** 133 305-326, 1983.
- [C88] M. Costabel, Boundary integral operators on Lipschitz domains: Elementary results, *SIAM J. Math. Anal.* **19** 613-626, 1988.
- [CD93] M. Costabel, M. Dauge, General edge asymptotics of solutions of second-order elliptic boundary value problems. I, II, *Proc. Roy. Soc. Edinburgh A* **123** 109-184, 1993.
- [CES93] M. Costabel, V. J. Ervin, E. P. Stephan, Quadrature and collocation methods for the double layer potential on polygons, *Z. Anal. Anwendungen* **12** 699-707, 1993.

- [CS85] M. Costabel and E. Stephan, Boundary integral equations for mixed boundary value problems in polygonal domains and Galerkin approximation, *Banach. Cent. Publ.* **15** 175-251, 1985.
- [CS87] M. Costabel and E. Stephan, On the convergence of collocation methods for boundary integral equation on polygons, *Mathematics of Computation* **49** 180 461-478, 1987.
- [CS93] M. Costabel and E. P. Stephan, Quadrature and collocation methods for the double layer potential on polygons, *Z. Anal. Anwend.* **12** 4 699-707, 1993.
- [DPS93] W. Dahmen, S. Prössdorf, and R. Schneider, Wavelet approximation methods for pseudodifferential equations II: Matrix compression and fast solution, *Advances in Computational Mathematics*, **1** 259-335, 1993.
- [DL85] R. Dautray et J.-L. Lions, *Analyse mathématique et calcul numérique*, Masson, 1985.
- [D99] E. Darve, Méthodes multipôles rapides: résolution des équations de Maxwell par des formulations intégrales, Université Paris-6, 1999.
- [DHS02] W. Dahmen, H. Harbrecht and R. Schneider, Compression techniques for boundary integral equations - optimal complexity estimates, *IGPM Report* **218**, RWTH Aachen, 2002.
- [DL02] H. Deng et H. Ling, An efficient preconditioner for electromagnetic integral equations using predefined wavelet packet basis, *IEEE Trans. Antennas* **50** 11 1633-1640, 2002.
- [D81] M. Djaoua, A method of calculation of lifting flows around two-dimensional corner-shaped bodies, *Mathematics of Computation* **36** 154 405-425, 1981.
- [DL97] H. Deng et H. Ling, Moment matrix sparsification using adaptive wavelet packet transform, *Electron. Lett.* **33** 1127-1128, 1997.
- [DL99a] H. Deng et H. Ling, Fast solution of electromagnetic integral equations using adaptive wavelet packet transform, *IEEE Trans. Antennas* **47** 674-682, 1999.
- [DL99b] H. Deng et H. Ling, On a class of predefined wavelet packet bases for efficient representation of electromagnetic integral equations, *IEEE Trans. Antennas* **47** 12, Dec 1999, pp. 674-682.
- [EH01] A. El Badia and T. Ha-Duong, Determination of point wave sources by boundary measurements, *Inverse problems* **17**, 1127-1139, 2001.
- [E93] J. Elschner, On the exponential convergence of some boundary element methods for Laplace's equation in nonsmooth domains, *Boundary value problems and integral equations in nonsmooth domains*, Luminy 1993, Lecture notes in Pure and Applied Mathematics **167** Dekker, New-York, 1995.

- [E73] B. I. Eskin, *Boundary problems for elliptic pseudodifferential operators* (in Russ.), Nauka, Moscow, 1973.
- [FJL77] E. B. Fabes, M. Jodeit, J. E. Lewis, Double layer potentials for domains with corners and edges, *Indiana Univ. Math. Journal* **26** 95-114, 1977.
- [G76] D. Gaier, Integralgleichungen erster Art und konforme Abbildungen, *Math. Zeitschr.* **147** 113-139, 1976.
- [G98] W. L. Golik, Wavelet packet for fast solution of electromagnetic integral equations, *IEEE Trans. Antennas* **46** 618-624, 1998.
- [G00] W. L. Golik, Sparsity and conditioning of impedance matrices obtained with semi-orthogonal and bi-orthogonal wavelet bases, *IEEE Trans. Antennas* **48** 473-481, 2000.
- [G05] D. Goujot, *N degrees of freedom per wavelength are enough in the Malvar Basis at high frequencies*, en préparation, 2004.
- [G85] P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, 1985.
- [G92] P. Grisvard, *Singularities in boundary value problems*, Masson, Paris, 1992.
- [GR87] L. Greengard and V. Rokhlin, A fast algorithm for particle simulation, *J. Comput. Phys.* **73** 325-348, 1987.
- [GNP73] J. P. Grégoire, J. C. Nédélec et J. Planchard, Calcul des fréquences propres d'un résonateur acoustique, *bull. Dire. Étud. Rech. EDF C 2* 5-23, 1973.
- [GNP74] J. P. Grégoire, J. C. Nédélec et J. Planchard, Problèmes relatifs à l'équation d'Helmholtz, *bull. Dire. Étud. Rech. EDF C 2* 16-33, 1974.
- [GS01] R. D. Grigorieff and I. H. Sloan, On qualocation and collocation methods for singular integral equations with piecewise continuous coefficients, using continuous splines on quasi-uniform meshes, Problems and methods in mathematical physics (Chemnitz, 1999), 146-161, *Oper. Theory Adv. Appl.* **121** Birkhäuser, Basel, 2001.
- [GS70] V. Guillemin, S. Sternberg, *Geometric asymptotics*, mathematical surveys and monographs **14**, AMS, 1970.
- [GYO00] N. Guan, K. Yashiro and S. Ohkawa, On a choice of wavelet base in the wavelet transform approach, *IEEE Trans. Antennas* **48** 1086-1090, 2000.
- [H86] T. Ha-Duong, A system of boundary integral equations for the problem of acoustic scattering by an absorbing wall, *Proceedings of the 8th international conference on boundary element method, Tokyo*, Springer-Verlag, 1986.
- [HN86] W. Hackbusch and Z. P. Nowak, A multi-level discretisation and solution method for potential flow problems in three dimensions, *Notes on*

*Numerical Fluid Mechanics* **14** 71-89, 1986.

[HN89] W. Hackbusch and Z. P. Nowak, On the fast matrix multiplication in the boundary element method by panel clustering, *Numer. Math.* **54** 463-491, 1989.

[H68] R. F. Harrington, field computation by moment methods, New York, *Macmillan*, 1968; *IEEE Press*, 1993.

[H90] R. F. Harrington, Origin and development of the method of moments for field computation, *IEEE antennas* **47** 674-682, 1990.

[HE81] R. E. Hockney and J. W. Eastwood, *Computer simulation using particles*, McGraw-Hill, New York, 1981.

[HKW80] G. C. Hsiao, P. Kopp and W. L. Wendland, A Galerkin collocation method for some integral equations of the first kind, *Computing* **25** 89-130, 1980.

[HW77] G. C. Hsiao, P. Kopp and W. L. Wendland, A finite element method for some integral equations of the first kind, *J. of Math. Anal. and Appl.* **58** 449-481, 1977.

[I98] F. Ihlenburg, *Finite element analysis of acoustic scattering*, Applied mathematical sciences 132, Springer, 1998.

[IB95a] F. Ihlenburg and I. Babuška, Finite element solution to the Helmholtz equation with high wave numbers part I: the h-version of the FEM, *Comput. Math. Appli.* **39** 9-37, 1995.

[IB95b] F. Ihlenburg and I. Babuška, Dispersion analysis and error estimation of Galerkin finite element methods for the numerical computation of waves, *Int. J. Numer. Meth. Eng.* **38** pp 3745-3774, 1995.

[IB97] F. Ihlenburg and I. Babuška, Finite element solution to the Helmholtz equation with high wave numbers – part II: the h-p version of the FEM, *SIAM J. Numer. Anal.* **34** 315-358, 1995.

[JK81a] D. S. Jerison and C. E. Kenig, The Dirichlet problem in nonsmooth domains, *Ann. of Math.* **113** 367-382, 1981.

[JK81b] D. S. Jerison and C. E. Kenig, The Neumann problem on Lipschitz domains, *Bull. Amer. Math. Soc.* **4** 203-207, 1981

[KL95] J.-P. Kahane, P.-G. Lemarié-Rieusset, *Fourier series and Wavelets*, Studies in the development of modern mathematics **3**, Gordon and Breach, 1995.

[K67] O. D. Kellogg, , *Foundations of potential theory*, Springer-Verlag, Berlin, 1967.

[KA82] R. E. Kleinman and T. S. Angell, Radiation condition and uniqueness, *Modern mathematical methods in diffraction theory and its applications in engineering*, Methoden und Verfahren der mathematischen Physik, Erhard Meister, 1982.



- [KJR03] P. Kolm, S. Jiang and V. Rokhlin, Quadruple and octuple layer potentials in two dimensions I: Analytical apparatus, *Appl. Comput. Harmon. Anal.* **14** 1 47-74, 2003.
- [K85] R. Kress, Minimizing the condition number of integral operators in acoustic and electromagnetic scattering, *Q. Jl. Mech. Appl. Math* **38** 323-341, 1985.
- [K89] R. Kress, *Linear Integral Equations*, Applied Mathematical Sciences, Springer, Berlin 1989.
- [KS93] R. Kress and I. H. Sloan, On the numerical solution of a logarithmic integral equation of the first kind for the Helmholtz equation, *Numer. Math.* **66** 199-214, 1993.
- [KS83] R. Kress, W. T. Spassov, *Ressemble à [A99]* On the condition number of boundary integral operators for the exterior Dirichlet problem for the Helmholtz equation, *Numer. Math.* **42** 77-95, 1983.
- [LM68] J.-L. Lions et E. Magenes, *Problèmes aux limites non homogènes et applications*, Dunod, Paris.
- [M96] G. Matviyenko, Optimized local trigonometric bases, *Appl. Comput. Harmon. Anal.* **3** 301-323, 1996.
- [M53] S. G. Mikhlin, *Linear Integral Equations*, Hindustan Publ, Dehli, 1953.
- [M59] S. G. Mikhlin, *Integral equations*, Pergamon Press LTD, 1959 (translated in 1962).
- [N66] J. Nečas, Les méthodes directes en théorie des équations elliptiques, 1966 Prague, édition française 1967 par Masson.
- [N01] J.-C. Nédélec, *Acoustic and electromagnetic equations*, Springer, 2001.
- [N92] J.-C. Nédélec, Quelques propriétés des dérivées logarithmiques des fonctions de Hankel, *C. R. Acad. Sci. Paris I* **314** 507-510, 1992.
- [NG97] R. D. Nevels and J. C. Goswami, Semi-orthogonal versus orthogonal wavelet basis sets for solving integral equations, *IEEE Trans. Antennas* **45** 9, sept. 1997.
- [PTG03] G. W. Pan, Y. V. Tretiakov and B. Gilbert, Smooth local cosine based Galerkin method for scattering problems, *IEEE Trans. Antennas* **51** 6 1177-1184, 2003.
- [PMR99] A. F. Peterson, R. Mittra, S. L. Ray, *Computational Methods for Electromagnetics*, IEEE Press, 1999.
- [PS97] T. Petersdorff and C. Schwab, Fully discrete multiscale Galerkin BEM, *Multiscale Wavelet Methods for PDE's*, W. Dahmen, A. J. Kurdila, and P. Oswald, Eds. New York: Academic, 287-345, 1997.
- [PTV92] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C/C++: The art of computer programming*, Cam-

bridge University Press, 1992.

[PS94] S. Prössdorf and J. Saranen, A fully discrete approximation method for the exterior Neumann problem of the Helmholtz equation, *Zeit. für Anal. Anw.* **13** 683-695, 1994.

[R43] F. Rellich, Über das asymptotische Verhalten der Lösungen von  $\Delta u + \lambda u = 0$  in unendlichen Gebieten, *Jber. Deutschen Math. Verein.* **53** 57-64, 1943.

[R83] V. Rokhlin, Solution of acoustic scattering problems by means of second kind integral equations, *Wave Motion* **5** 257-272, 1983.

[R93] V. Rokhlin, Diagonal forms of translation operators for the Helmholtz equation in three dimensions, *A. C. H. A.* **1** 82-93, 1993.

[R85] V. Rokhlin, Rapid solution of integral equations of scattering theory in two dimensions, Yale Research Report YALEU/DSC/RR-440, Nov. 1985; *J. Comp. Phys.*, **86** pp. 414-439, 1990.

[R87] W. Rudin, *Real and complex analysis*, McGraw-Hill, 1987.

[S85] E. Schnack, Stress analysis with a combination of HSM and BEM, *The mathematics of finite elements and applications V*, J. Whiteman ed, Academic Press, London, 273-282, 1985.

[S93] R. Schneider, Optimal convergence rates of adaptive algorithms for finite element multiscale methods, *Boundary value problems and integral equations in nonsmooth domains* 269-284, Luminy 1993, Lecture notes in Pure and Applied Mathematics **167** Dekker, New-York, 1995.

[S03] C. Schneider, Stability of qualocation methods for elliptic boundary integral equations. *J. Integral Equations Appl.* **15** 203-216, 2003.

[S61] L. Schwartz, *Méthodes mathématiques pour les sciences physiques*, Hermann, 1961.

[S64] E. Shamir, Reduced Hilbert transforms and singular integral equations, *J. Analyse Math.* **12** 227-305, 1964.

[SBL99] Y. Shifman, Z. Baharav and Y. Leviatan, Analysis of truncated periodic array using two-stage wavelet-packet transformations for impedance matrix compression, *IEEE Trans. Antennas* **47** 630-636, 1999.

[S00] I. H. Sloan, Qualocation, Numerical analysis 2000, Vol. VI, Ordinary differential equations and integral equations. Edited by Christopher Baker, Giovanni Monegato, John Pryce and Guido Vanden Berghe. *J. Comput. Appl. Math.* **125** 461-478, 2000.

[S91] I. H. Sloan, Error analysis of boundary integral equations, *Acta numerica* **1** 287-339, 1991.

[SW98] I. H. Sloan and W. L. Wendland, Qualocation methods for elliptic boundary integral equations, *Numer. Math.* **79** 451-483, 1998.

[SW99] I. H. Sloan and W. L. Wendland, Spline qualocation methods for variable-coefficient elliptic equations on curves, *Numer. Math.* **83** 497-533,

1999.

[S12] A. Sommerfeld, Die Greensche Function der Schwingungsgleichung, *Jber. Deutschen Math. Verein.* **21** 309-353, 1912.

[SLC97] J. M. Song, C. C. Lu, and W. C. Chew, Multi-level fast-multipole algorithm for electromagnetic scattering for large complex objects, *IEEE Trans. Antennas* **45** 1488-1493, 1997.

[SL93] B. Z. Steinberg and Y. Leviathan, A multiresolution study of two-dimensional scattering by metallic cylinders, *IEEE Trans. Antennas* **41** 200-207, 1993.

[SL96] B. Z. Steinberg and Y. Leviathan, On the use of wavelet expansions in the method of moments, *IEEE Trans. Antennas* **44** 572-579, 1996.

[SS89] E. P. Stephan and M. Suri, On the convergence of the p-version of the boundary element Galerkin method, *Math. Comp.* **52** 31-48, 1989.

[S56] G. G. Stokes, On the numerical calculation of a class of definite integrals and infinite series, *Camb. Philos. Trans.* **9** 166-187, 1856.

[T79] R. Temam, *Navier-Stokes equations*, Elsevier, 1979.

[T93] I. Terrasse, *Résolution mathématique et numérique des équations de Maxwell stationnaires par une méthode de potentiels retardés*, thèse de doctorat de l'École Polytechnique, soutenue devant S. Mas-Gallic, A. Bachelot, P. Joly, J. P. Martinaud, J. C. Nedelec (directeur), J. Rauch, P. A. Raviart et J. Tourneur, 1993.

[TB97] L. N. Trefethen and D. Bau, *Numerical linear algebra*, Philadelphia, PA: SIAM, 1997.

[V84] G. Verchota, Layer potentials and regularity for the Dirichlet problem for Laplace's equation in Lipschitz domains, *J. Funct. Anal.* **59** 572-611, 1984.

[V93] N. Ja. Vilenkin and A. U. Klimyk, *Representation of Lie groups and special functions* Mathematics and its Applications (russian ed), 1993.

[W18] G. N. Watson, The limits of applicability of the principle of stationary phase, *Proc. Camb. Philos. Soc.* **19** 49-55, 1918.

[W22] G. N. Watson, *A treatise on the theory of Bessel functions*, Cambridge University Press 1995.

[W56] C. H. Wilcox, A generalization of theorems of Rellich and Atkinson, *Proc. Amer. Math. Soc.* **7** 271-276, 1956.

[W90] W. L. Wendland, Boundary element methods for elliptic problems, *Mathematical theory of finite and boundary element methods* Schatz, Thomée, Wendland, Birkhäuser, Basel, 219-276, 1990.

[W94] M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*, A. K. Peters, Natick Mass., 1994.

[WC95] R. L. Wagner, W. C. Chew, A study of Wavelet for the Solution of Electromagnetic Integral Equation, *IEEE Trans. Antennas* **43** 8, 1995.

- [WW92] D. S. Wang and G. Welland, Modeling of electromagnetic scattering using wavelet techniques for geometry modeling and expansion functions, *URSI Radio Sci. Meet. Dig.*, Detroit, MI, 1992, p. 277.
- [YS88] Y. Yan and I. H. Sloan, On integral equations of the first kind with logarithmic kernels, *J. Integ. Eq. Appl.* **1** 4 549-579, 1988.
- [Z95] B. Zhou, *Méthode des équations intégrales pour la résolution des problèmes de diffraction à hautes fréquences*, thèse de l'Université d'Orsay soutenue devant T. Abboud, A. Bachelot, M. Balabane, B. Lebeau, J. C. Nédélec, P. A. Raviart et J. C. Saut, 1995.

<b>Localisation des énoncés et des figures</b>					
				Figure 11	70
Property I.1	11	Théorème 2.17	83	Figure 12	73
Notations I.2	11	Remarque 2.18	84	Figure 13	75
Property I.3	11	Propriété 2.19	85	Figure 14	89
Property I.5	11	Propriété 2.20	92	Figure 15	109
Definition I.6	12	Remarque 2.22	95	Figure 16	120
Theorem I.7	12	Propriété 2.23	96	Figure 17	121
Theorem I.8	13	Propriété 2.24	98	Figure 18	127
Property I.9	14	Propriété 2.25	102	Figures 19 et 20	130
Remark I.10	15	Théorème 2.27	105	Figure 21	131
Theorem I.12	17	Théorème 2.28	110	Figure 22	135
Definition I.13	19	Théorème 2.29	115	Figure 23	145
Theorem I.16	20	Théorème 2.30	115	Figure 24	147
Assumption I.17	21	Théorème 3.1	117	Figure 25	148
Assumption I.18	21	Propriété 3.2	118	Figure 26	149
Property I.19	22	Remarque 3.3	129	Figure 27	150
Property I.21	23	Propriété 4.1	146	Figure 28	151
Definition I.22	24	Propriété 4.2	154	Figure 29	152
Definition I.27	26	Propriété 4.3	155	Figures 30 et 31	162
Definition I.32	27	Propriété 4.4	159	Figure 32	187
Property I.33	27	Théorème 4.5	172	Figures 33 et 34	188
Property I.36	28	Propriété 4.6	175	Figure 35	190
Property I.37	32	Propriété 4.7	177	Figure 36	191
Property I.38	33	Théorème 4.8	178	Figure 37	192
Property I.39	34	Propriété 4.9	183	Figures 38 à 41	198
Definition I.40	35	Théorème 5.1	185	Figure 42	199
Property I.42	36	Propriété 5.2	216	Figures 43 à 46	201
Property I.43	37	Propriété 5.3	220	Figure 47	202
Property I.44	37	Propriété 5.4	220	Figure 48	203
Theorem I.45	38	Propriété 5.5	222	Figures 49 à 52	204
Theorem I.46	40	Propriété 5.6	224	Figure 53	205
Theorem I.47	40	Propriété 5.7	226	Figures 54 à 57	206
Propriété 2.1	55	Remarque 5.8	227	Figure 58	207
Définition 2.4	63	Propriété 5.9	230	Figure 59	208
Propriété 2.7	64	Propriété 6.1	232	Figures 60 à 63	211
Définition 2.9	67	Figure 1	15	Figure 64	212
Propriété 2.10	67	Figure 2	16	Figure 65	213
Remarque 2.11	68	Figures 3 à 5	31	Figure 66	215
Propriété 2.12	69	Figure 6	34	Figure 67	216
Théorème 2.13	71	Figure 7	54	Figures 68 et 69	226
Théorème 2.14	76	Figures 8 et 9	60	Figure 70	228
Propriété 2.16	81	Figure 10	65	Figure 71	229
Index : voir <a href="http://www.maths.univ-evry.fr/pages_perso/goujot/these.txt">http://www.maths.univ-evry.fr/pages_perso/goujot/these.txt</a>					