Méthodes d'accéleration pour la résolution numérique en électrolocation et en chimie quantique

Philippe Laurent

Thèse de Doctorat

Philippe LAURENT

Mémoire présenté en vue de l'obtention du
grade de Docteur de l'École nationale supérieure des mines de Nantes
sous le label de l'Université de Nantes Angers Le Mans

École doctorale : Sciences et technologies de l'information et mathématiques

Discipline : Automatique, productique, section CNU 921
Spécialité : Automatique et mathématiques appliquées
Unité de recherche : Institut de Recherche en Communications et Cybernétique de Nantes (IRCCYN)

Soutenue le 26 octobre 2015
Thèse n° : 2015 EMNA0122

Méthodes d'accélération pour la résolution numérique en électrolocation et en chimie quantique

JURY

Rapporteurs :	**M. Alfio BORZI**, Professeur, Université de Würzburg
	M^me^ Laurence HALPERN, Professeur, Université Paris 13
Examinateur :	**M^me^ Virginie EHRLACHER**, Chargée de Recherche, École des Ponts ParisTech
Invités :	**M^me^ Christine CHEVALLEREAU**, Directeur de Recherche, École Centrale de Nantes
	M. Vincent LEBASTARD, Maître assistant associé, École des Mines de Nantes
Directeurs de thèse :	**M. Frédéric BOYER**, Professeur, École des Mines de Nantes
	M. Pierre ROUCHON, Professeur, École des Mines de Paris
Co-directeur de thèse :	**M. Julien SALOMON**, Maître de conférences, Université Paris-Dauphine

À Barbara
et à nos futurs enfants.

Remerciements

Cette thèse s'achève enfin ! Je n'y croyais plus ! Bien que sa durée sur le papier soit confortable, elle s'est réellement effectuée en un temps record. Bien qu'il y ait eu des moments difficiles, comme par exemple le vol de mon ordinateur et de mon disque dur de sauvegarde pour n'en citer qu'un, ce moment restera pour moi, au delà d'un travail, une recherche introspective forte qui m'a permis de mieux me connaître, à l'image d'un rite initiatique. D'un point de vue plus professionnel, la grande variété des outils et des méthodes découvertes et utilisées durant cette période, me permet d'affirmer que les mathématiques "c'est quand même bien fait !". Enfin, une thèse n'est pas grand chose sans les rencontres et les échanges effectués. C'est pourquoi il est de coutume, et je m'en réjouis, de remercier dans ces quelques lignes les personnes que l'on a été amené à rencontrer et qui nous ont accompagné. Je tiens d'avance à m'excuser des oublis éventuels, ils ne sont pas volontaires !

Tout d'abord je tiens à remercier Frédéric BOYER et Pierre ROUCHON qui m'ont donné la chance de pouvoir effectuer une thèse, ce qui n'était pas jouer d'avance, tout comme sa réalisation par la suite ! Ils m'ont mis le pied à l'étrier en me laissant une grande liberté sur les thématiques abordées. Ce sont de réels personnages, chacun à leur façon, fascinants par l'étendue de leurs connaissances. Merci d'avoir allumé l'étincelle.

Je tiens également à remercier A. BORZI, C. CHEVALLEREAU, V. EHRLACHER, L. HALPERN et V. LEBASTARD qui m'ont fait l'honneur d'accepter de faire partie de mon jury.

J'ai eu la chance d'être accueilli pendant deux années à l'école des Mines de Nantes. J'en conserve un souvenir ému, entre autres, grâce à l'équipe du projet ANGELS, le personnel du DAP et de façon plus générale de l'École. Je remercie aussi au passage Vincent et Mathieu pour leurs aides précieuses sur des sujets divers et variés et plus particulièrement dans l'utilisation de la BEM "maison". Des remerciements plus particuliers à Reda, Brahim, Younes, Tarek et Jérémie qui ont animé mes journées et mes nuits Nantaises (et Hawaïennes). Merci enfin à l'Université Paris-Dauphine de m'avoir accueilli par la suite.

Ma gratitude va aussi à P. B. GOSSIAUX, R. RABAH et S. TABIOU qui m'ont permis d'être leur chargé de TD à l'École des Mines de Nantes, mais aussi P. CARDALIAGUET et G. LEGENDRE à l'université Paris-Dauphine. Les divers échanges que l'on a pu avoir m'ont grandement enrichi.

Je suis reconnaissant envers J. SALOMON et G. TURINICI de m'avoir donné un sujet de stage en M2 et, par la suite, de m'avoir permis d'aller défendre nos travaux à Hawaï. Malheureusement, à notre insu, nous n'avons pas pu prolonger cette collaboration par un doctorat, j'espère que nous aurons d'autres occasions. Merci également de m'avoir aidé à trouver cette thèse.

Un grand merci à K. BEAUCHARD et O. KAVIAN d'avoir pris le temps de me recevoir pour me permettre de les questionner. Ces échanges m'ont permis de trouver des pistes fructueuses.

Je tiens à remercier l'École publique qui m'a permis de me construire et de me structurer. Je suis reconnaissant envers l'école primaire Eugénie Cotton, le collège Musselburg, le lycée Alain, l'université de Versailles Saint-Quentin-en-Yvelines, l'université Paris-Dauphine, l'École des Mines de Nantes et les personnes qui font vivre ces institutions. Les noms de Mme COHEN, M. FORMON, M. LEGANGNEUX, M. SILVERA, Mme TILLIER, M. ZIANI... me reviennent en mémoire, je les remercie (et les autres) de leur passion, leur patience et leur dévouement. Je suis maintenant moi-même enseignant, je tente de rendre à mon tour modestement, dans cette grande

chaîne de transmission, tout ce qui m'a été donné et de susciter, je l'espère, des vocations. Je tiens à remercier les classes de 2DE8, 2DE9, 1ES2 et 1ES4 du lycée Alain qui ont essuyé les plâtres, mais aussi aux classes qui ont/vont suivi/suivre...

Merci à mes amis de ne pas avoir coupé les ponts, même si nos rencontres ont été de plus en plus sporadiques. Leur soutien par leurs présences ou leurs pensées le jour de la soutenance à Nantes m'a porté ! Je remercie plus particulièrement Christophe et Pierre qui ont mis la main à la pâte, ce qui m'a remotivé dans les moments de flottements et de doutes.

Sans mes parents je serais peu de chose, merci à eux. Mes remerciements vont aussi, de façon plus générale, à toute ma famille. Je remercie plus particulièrement mon frère, David, de m'avoir détourné de la biologie pour m'orienter vers les mathématiques. Très jeune il a toujours essayé de me transmettre ce qu'il apprenait à l'école. Comme par exemple la notion de fonction qui, du haut de mes huit-neuf ans, m'a laissé perplexe pendant un certain temps ! Il a, de plus, toujours eu la bonne idée, durant mon adolescence, de m'offrir des livres de vulgarisation au lieu du dernier jeu vidéo à la mode. Après l'avoir longtemps maudit, l'âge faisant, je lui suis reconnaissant de m'avoir initié à cette discipline. Il a toujours été de bons conseils : pour la thèse il m'avait prévenu ! Au passage son fils (six ans), Arno la "patate", ne s'est pas contenté d'être un fin écouteur durant la soutenance, il a reproduit les schémas ! Attention à ne pas faire une thèse comme ton père et ton oncle !

Je tiens ici à déclarer mon amour à Barbara, qui malgré l'éloignement pendant deux ans m'a témoigné un amour sans faille. Elle a durant cette thèse toujours été bienveillante, repoussant les remarques et les remontrances à l'après-thèse, mince ça arrive ! Cette thèse et ces remerciements (je me dois d'être reconnaissant) seraient égrainés de fautes en tout genre sans ses innombrables relectures.

Enfin, je ne peux pas terminer ces remerciements sans ré-évoquer Guillaume LEGENDRE et Julien SALOMON. Sans eux cette thèse n'existerait pas ! ! Je pourrais écrire une thèse à leur sujet, c'est pourquoi ces courts remerciements sont inversement proportionnels à la reconnaissance et à l'affection sans bornes que j'ai pour eux !

Table des matières

1

Introduction

Sommaire

Cette thèse aborde deux thématiques différentes. La première est liée au sens électrique appliqué à la robotique. Cette thématique aborde le développement et l'analyse de méthodes liées à la robotique. La seconde traite d'un problème inverse en chimie quantique. Ce dernier travail est un prolongement de travaux de Master 2. Chacun de ces sujets sera introduit plus en détails dans ce chapitre introductif, dans l'ordre mentionné précédemment. Par la suite, les chapitres 2 et 3 présentent les travaux liés à la robotique, tandis que le chapitre 4, ceux se référant à la chimie quantique. La quasi totalité des résultats numériques présentés dans cette thèse a été réalisée à l'aide d'Octave[1].

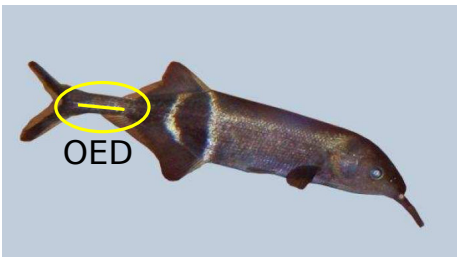
1.1 Généralités sur le sens électrique et le projet ANGELS

Dans cette section, on décrit le contexte dans lequel s'inscrit une part importante de la thèse. Sont présentées dans la première sous-section, une rapide présentation du sens-électrique et son application à la robotique. Dans la sous-section suivante, nous présentons le projet qui a donné lieu à cette thèse. Les différents dispositifs matériels créés durant cette période sont exposés dans cette partie.

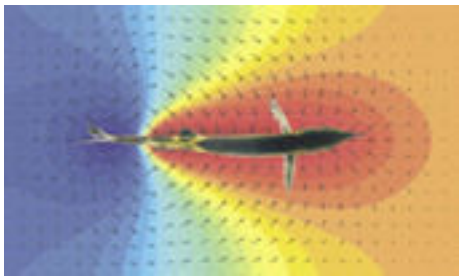
1.1.1 Sens électrique des poissons et biomimétisme

Cette section est inspirée de [14], [24] et [10]. La première citation est un livre de référence dans le domaine.

Connus depuis l'antiquité, utilisés comme prémisse de l'électrothérapie, les poissons électriques sont restés un mystère pour la communauté scientifique jusqu'au XVIII^{e} siècle. C'est à cette époque que J. Walsh, décrit les chocs produits par ces poissons comme étant de nature électrique. Les poissons dits fortement électriques, comme la torpille, le poisson-chat électrique ou l'anguille produisent des décharges pour se défendre ou s'attaquer à leurs proies. Celles des torpilles peuvent atteindre 200 volts et 30 ampères. Malgré cette découverte subsistait un mystère. Contrairement aux poissons fortement électriques, de petits poissons présents en Afrique et en Amérique du Sud, présentent une activité électrique effrénée ne leur permettant pas de se défendre. On dira de ces poissons qu'ils sont faiblement électriques (on utilisera l'acronyme PFE). Par exemple les Mormyridae, vus en photo 1.1, produisent, à l'aide de leur organe électrique de décharges (on utilisera l'acronyme OED), des décharges de l'ordre de 5 à 20 volts. Ceci posa un



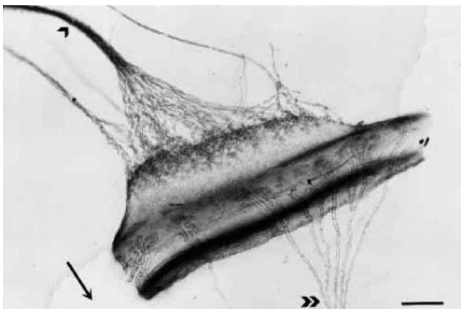
(a) L'OED est situé de part et d'autre de l'animal au niveau de la queue.



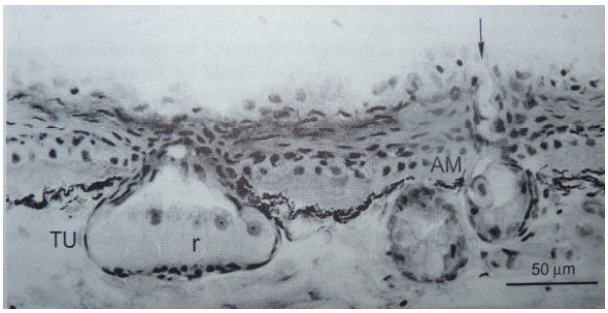
(b) Champ et potentiel électriques produits par le poisson.

FIGURE 1.1 – Poissons faiblement électriques Gnathonemus petersii, de la famille des Mormyridae. Ces images sont tirées de [18].

[1]Octave est un langage interprété de haut-niveau, destiné en premier lieu au calcul numérique. Octave est distribué sous *licence publique générale GNU*. L'adresse officiel est http://www.gnu.org/software/octave/.



(a) Photomontage tiré de [15] présentant l'EOD d'un poisson couteau.

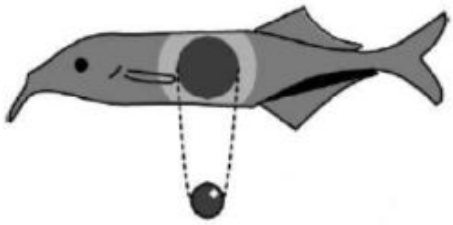


(b) Deux types d'électrorécepteurs présents à la surface de la peau d'un poisson couteau. Photo tirée de [14].

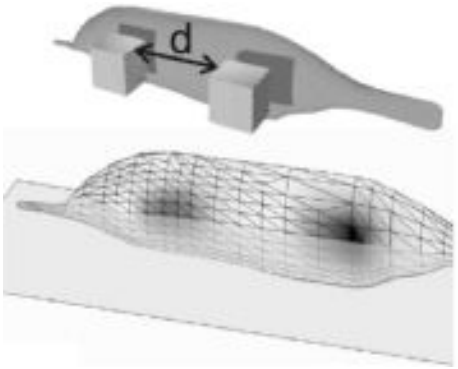
FIGURE 1.2 – Organes dédiés à l'électrolocation chez les PFE.

réel problème à C. Darwin dans sa théorie de l'évolution [17]. En effet, il ne pouvait pas expliquer la subsistance de cette fonction, d'apparence inutile. Il régla ce problème en affirmant que les PFE sont un des maillons dans l'évolution des poissons communs vers les poissons fortement électriques. Une centaine d'années plus tard, dans les années 50 H. Lissmann et K. Machin dans [33] ont montré expérimentalement que les PFE perçoivent leur environnement et communiquent entre eux grâce aux décharges électriques qu'ils produisent.

Dans son principe, l'électrolocation, ou sens électrique actif, est basée sur l'émission d'un champ électrique par l'organe émetteur l'EOD. De forte conductivité par rapport au milieu ambiant, le corps du poisson focalise les lignes de champ émises, les obligeant ainsi à traverser sa peau électro-sensible, créant en quelque sorte une bulle de perception électrique autour de son corps, ce qui est en figure 1.1b. Ainsi, grâce à une multitude de capteurs électrosensibles distribués sur son épiderme, le poisson fabrique une image électrique représentative de son environnement. Une illustration des organes intervenant dans l'electrolocalisation est présentée en 1.2.



(a) Schéma représentant l'image électrique produite par une sphère sur la peau d'un PFE. Schéma tiré de [18].



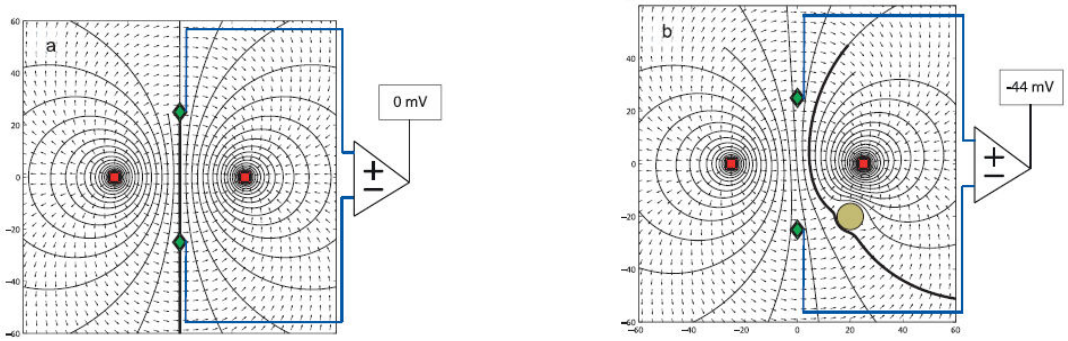
(b) Calcul numérique de l'image électrique de deux cubes sur un avatar numérique d'un PFE. Image tirée de [42].

FIGURE 1.3 – Représentation d'images électriques.

Dès lors qu'un objet entre dans cette bulle de perception, il déforme le champ électrique produit par le poisson et donc le champ appliqué sur sa peau. Ces variations par rapport à un signal sans objet lui permettent de déterminer la distance qui le sépare d'un objet, la forme et les paramètres électriques de ce dernier. Cette dernière découverte est due à Von der Emde *et al.* dans [18] paru

en 1998. Ces poissons vivent dans des eaux troubles. Ils auraient développé ce sens afin de pallier le manque de visibilité.

Parallèlement, dans les années 90, un changement de paradigme s'opéra en robotique. Celui-ci a, entre autres, été motivé par les difficultés à reproduire artificiellement certains mécanismes produits par l'intelligence humaine. Poursuivant l'intuition de Darwin, et s'appuyant sur une mathématisation plus poussée des sciences de la nature, les chercheurs ont observé que la nature, au fil du temps, a sélectionné des solutions souvent optimales pour un problème donné. La bio-inspiration consiste ainsi à transposer ou à s'inspirer de ces solutions pour le développement de dispositifs techniques. Ce principe dépasse bien évidement le cadre strict de la robotique et peut être appliqué à d'autres domaines. C'est dix ans plus tard que la première équipe de roboticiens, sous la direction de M. MacIver, s'intéressa au sens électrique. Dans [46] est présenté leur premier dispositif expérimental constitué de quatre électrodes disposées en losange, dont la figure 1.4 donne une version schématisée. Deux électrodes opposées sont polarisées entre elles, jouant ainsi le rôle de l'OED. La mesure est alors réalisée sur les deux électrodes restantes. C'est cette mesure qui permet de caractériser la présence d'un objet. Par la suite, en France, s'est dé-



(a) dispositif sans objet (b) dispositif avec objet

FIGURE 1.4 – Dispositif expérimental en losange. Ce graphique est tiré de [46].

veloppé autour de F. Boyer le projet RAAMO (acronyme signifiant *Robot Anguille Autonome en Milieu Opaque*), dont la photo 1.5 présente un résultat marquant. RAAMO est un robot doté du sens électrique imitant la nage de l'anguille. Le projet ANGELS (acronyme signifiant *ANGuilliform robot with ELectric Sense*), détaillé dans la sous-section suivante, a succédé au robot RAAMO. Ce

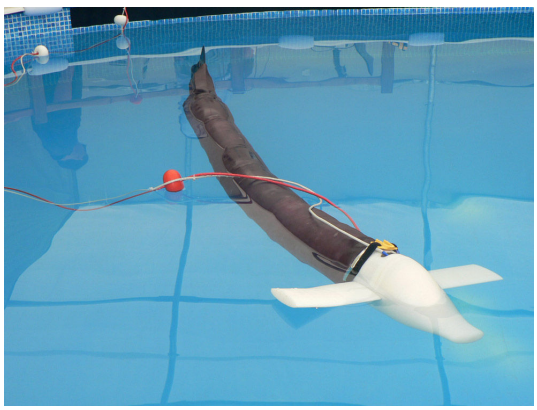


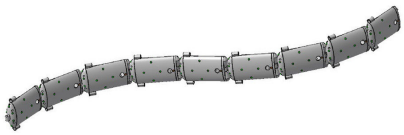
FIGURE 1.5 – Robot RAAMO.

projet part du constat suivant : bien que potentiellement très prometteur, le sens électrique reste encore à ce jour largement ignoré de la communauté de Robotique sous-marine. Dans ce cadre,

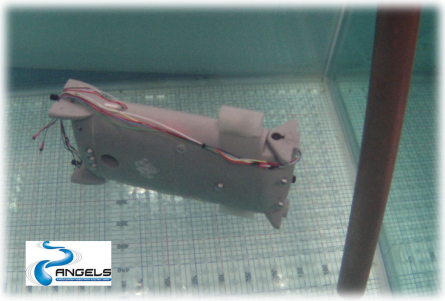
le robot Angels a été le premier robot sous-marin ayant navigué de manière autonome grâce au sens électrique[2].

1.1.2 Projet ANGELS

Le projet ANGELS, du nom du robot développé à cette occasion, est un projet européen FET (acronyme signifiant *Future and Emerging Technologies*) qui regroupa pendant trois ans, de 2009 à 2012, huit partenaires internationaux spécialistes en robotique, physique, électronique, mécatronique, biologie et automatique. En s'inspirant de l'anguille et des PFE, l'objectif fut de créer un robot aquatique doté d'une nage énergétiquement efficace et très manœuvrante et du sens électrique pour la perception. Ce système, sensoriel et moteur, a été choisi car il est tout particulièrement adapté aux grands déplacements dans des milieux confinés et turbides pour lesquels ce type de robot est dédié. De plus, le robot ANGELS est capable de se détacher en plusieurs modules autonomes, afin d'optimiser l'exploration ou d'évoluer dans des milieux plus difficiles d'accès. La figure 1.6 présente le robot. La finalité plus générale de ANGELS était de construire un proto-



(a) Représentation du robot ANGELS par un logiciel de CAO 3D, réalisé en début de projet.



(b) Premier module ANGELS réalisé. Il a été utilisé pour différents tests sur le sens électriques et la locomotion.

(c) Les neuf modules ANGELS assemblés en fin de projet.

FIGURE 1.6 – Ces photos représentent les deux morphologies possibles du robot ANGELS : module(s) seul 1.6b ou connectés 1.6a et 1.6c.

type de robot autonome capable de se reconfigurer, afin d'adapter sa morphologie à une situation donnée.

Cette thèse est en grande partie centrée sur des modèles et des méthodes attachés au mode de perception utilisé par le robot ANGELS, à savoir le sens électrique. C'est pourquoi nous présentons maintenant les dispositifs expérimentaux mis en place sur le site de l'École des Mines de Nantes. Ceux-ci ont vocation à étudier et développer des modèles liés au sens électrique et à son application en robotique. Avant d'être embarquées dans le robot, les méthodes sont testées dans un *réservoir* d'eau représenté sur la photo 1.7. Par la suite, les essais sur le robot autonome ont lieu dans une piscine dédiée. Il est possible d'installer deux morphologies de capteurs dans

[2] http://www.youtube.com/watch?v=0dVc3JrIds0

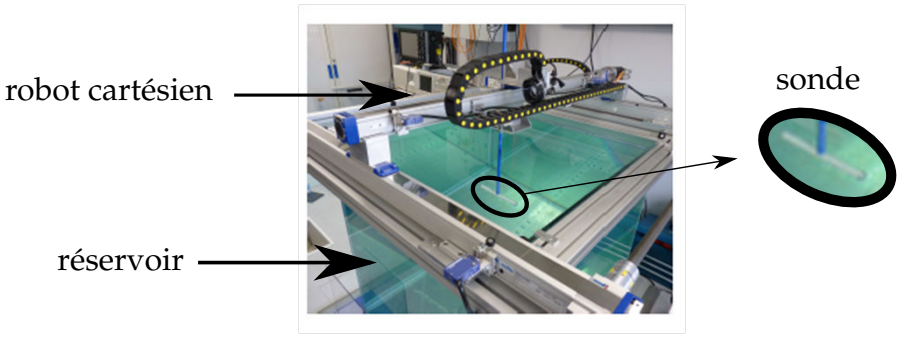


FIGURE 1.7 – Banc d'essai utilisé durant le projet ANGELS. Le réservoir peut contenir 1m^3 d'eau. Sur celui-ci est fixé un robot *cartésien trois axes*, c'est-à-dire qu'il se repère selon trois coordonnées x, y et θ. Les sondes sont attachées au robot cartésien à l'aide d'une canne, dans laquelle passe le câblage électrique qui relie l'ordinateur et les dispositifs de mesure. Ces installations ont été remplacées depuis par des dispositifs plus importants.

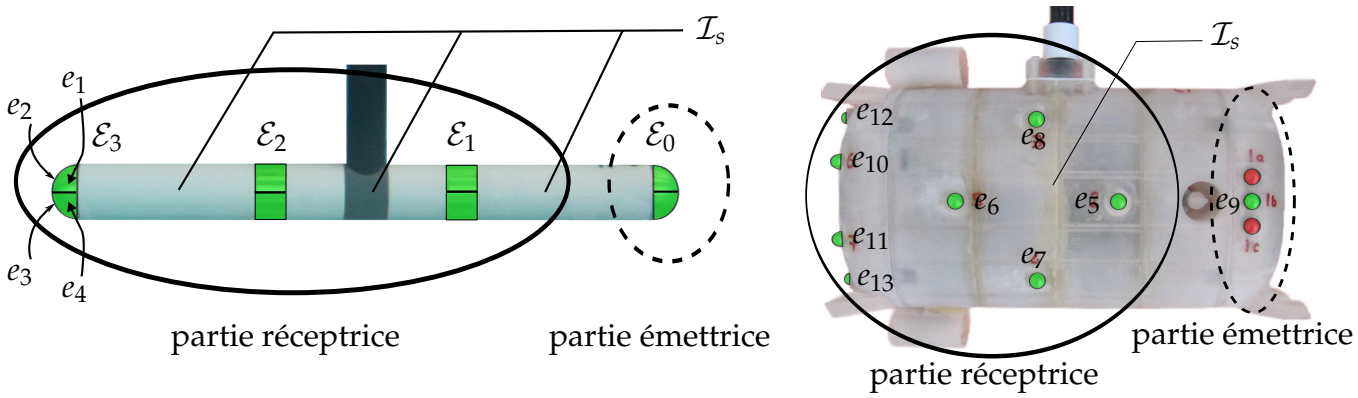
le réservoir : des sondes dites élancées, voir la photo 1.8a et les modules ANGELS, voir la photo 1.8b. La géométrie des sondes élancées étant particulièrement adaptée à l'électrolocation, elles sont un support privilégié pour les études et le développement technique. Une fois établis, nous transposons les résultats sur un module ANGELS dans le réservoir, puis dans le robot autonome. Nous ferons régulièrement référence à ces deux types de capteur par la suite. A l'heure actuelle les senseurs disposent d'un nombre maximum de 16 électrodes, ce qui reste très peu comparé aux quelques milliers d'électro-récepteurs cutanés du poisson. Néanmoins, comme le montre la figure 1.9, ce faible nombre d'électrodes peut être compensé par le mouvement pour générer une image électrique. L'un des prochains objectifs matériels est la création d'une véritable image électrique de l'environnement, ce qui amènera à densifier un nombre élevé d'électrodes de taille réduite. D'un point de vue pratique, cela pose un certain nombre de problèmes. Par exemple, si l'on dispose des électrodes de plus en plus proches les unes des autres, le risque de créer des courts-circuits augmente. Plusieurs solutions sont envisageables : ajouter des électrodes reliées non pas au générateur de tension, mais directement à la masse, qu'on appelle électrodes passives ; ou bien appliquer un jeu de tension adapté variant progressivement de l'une à l'autre.

1.2 La modélisation du sens électrique pour la robotique

Cette section est composée de trois parties. Dans la première sous-section, après avoir présenté un cadre pour étudier le sens électrique en robotique, nous présentons les problèmes direct et inverse qui y sont liés. Dans les sections suivantes nous présentons deux méthodes. La première est une méthode numérique appelée méthode des éléments de frontière permettant la résolution du problème direct. Elle est basée sur les formulations intégrales de frontière du problème. La seconde méthode, dite "des réflexions", a été introduite récemment pour l'étude du sens électrique dans [12] et permet de résoudre le problème direct à partir de sous-problèmes liés. Cette méthode peut, dans son esprit, être mise en rapport avec celle de Schwarz : là où cette dernière décompose un domaine, la méthode des réflexions décompose une frontière. Les éléments de frontière et la méthode des réflexions sont présentés ici pour introduire la section suivante consacrée aux algorithmes associés.

1.2.1 Modèle mathématique

Dans cette sous-section nous ne présentons pas le contenu physique ou biologique des équations présentées. Pour plus d'informations sur le sujet, on pourra se référer à [24] et [10]. On pourra



(a) Sonde élancée.

(b) Sonde ANGELS.

FIGURE 1.8 – Les figures 1.8a et 1.8b représentent les deux morphologies de sondes utilisables dans le réservoir. Chacune d'elles est composée de seize d'électrodes indiquées en vert. Elles sont notées $\{e_k\}_{k\in[\![1;16]\!]}$ et l'ensemble total de la surface qu'elles recouvrent est noté $\mathcal{E} := \cup_{k=1}^{16} e_k$. Elles sont généralement placées par symétrie à la surface du capteur, afin d'extraire des contributions axiales ou latérales des perturbations. Les électrodes en rouge sur ANGELS ne sont plus utilisées. Les surfaces isolantes sont regroupées sous la notation $\mathcal{I}_s$. On distingue deux zones sur la sonde : la partie *émettrice* en traits pleins et la partie *réceptrice* en pointillés. Elles sont toutes deux reliées au générateur de puissance. La différence de potentiels entre ces deux zones, crée la bulle de perception électrique, mimant ainsi l'activité du poisson. La mesure s'effectue alors sur la partie réceptrice. Sur la sonde élancée, les électrodes $\{e_k\}_{k\in[\![1;16]\!]}$ sont regroupées par quatre pour former des anneaux ou des hémisphères aux extrémités, que l'on note $\{\mathcal{E}_{k'}\}_{k'\in[\![0;3]\!]}$. Ce regroupement spécifique aux sondes élancées permet de définir plusieurs degrés de résolution. Les regroupements $\mathcal{E}_{k'}$ peuvent aussi être envisagés pour une vision plus globale, tandis que les e_k permettent de décrire plus finement la situation.

aussi consulter [52] pour une description d'un modèle utilisé pour la peau du poisson. Les notations introduites dans cette sous-section dépassent largement le contexte de l'électrolocation. Ce cadre recouvrant des applications plus larges, nous avons tout de même souhaité qu'il y figure. Une lecture rapide permettra au lecteur de se familiariser avec les éléments de langage et certaines notations.

Soit Ω un ouvert inclus dans $\mathbb{R}^{n_d}$, avec $n_d = 2$ ou $n_d = 3$. Les bords sont généralement supposés réguliers ou au moins lipschitziens. Considérons un ouvert, noté $\mathcal{B}$, représentant le robot[3]. Le domaine[4] $\mathcal{B}$ respecte, les propriétés suivantes

$$\mathcal{B} \subset \mathbb{R}^{n_d} \setminus \overline{\Omega} \quad \text{et} \quad \partial\mathcal{B} \cap \partial\Omega = \partial\mathcal{B}.$$

On considère de plus que Ω contient $p \in \mathbb{N}$ domaines. Soit $\mathcal{O} = \cup_{l=1}^{p} \mathcal{O}_l$ où les $\mathcal{O}_l$ sont appelés objets. Enfin on nommera domaine extérieur, milieu ambiant ou domaine mouillé, l'ouvert D_e qui est le complémentaire des objets dans Ω, c'est-à-dire $D_e := \Omega \setminus \cup_{l=1}^{p} \overline{\mathcal{O}_l}$. Quelque soit l'ensemble A considéré, on notera $\chi(A)$ la fonction indicatrice associée.

On peut alors définir l'équation régissant le milieu. Soit $\gamma_e \in \mathbb{R}^+$ et p réels $(\gamma_l)_{l\in[\![1;p]\!]}$ tels que

$$\forall l \in [\![1;p]\!], \quad 0 \leq \gamma_l \neq \gamma_e < +\infty.$$

Ces valeurs représentent respectivement la conductivité du domaine extérieur et celle des objets. Dans un cadre applicatif, on prend régulièrement la conductivité de l'eau sur D_e, c'est-à-dire

[3] Pour qualifier le robot, nous utiliserons aussi les termes : module ; sonde ; agent.

[4] Domaine : ouvert connexe borné.

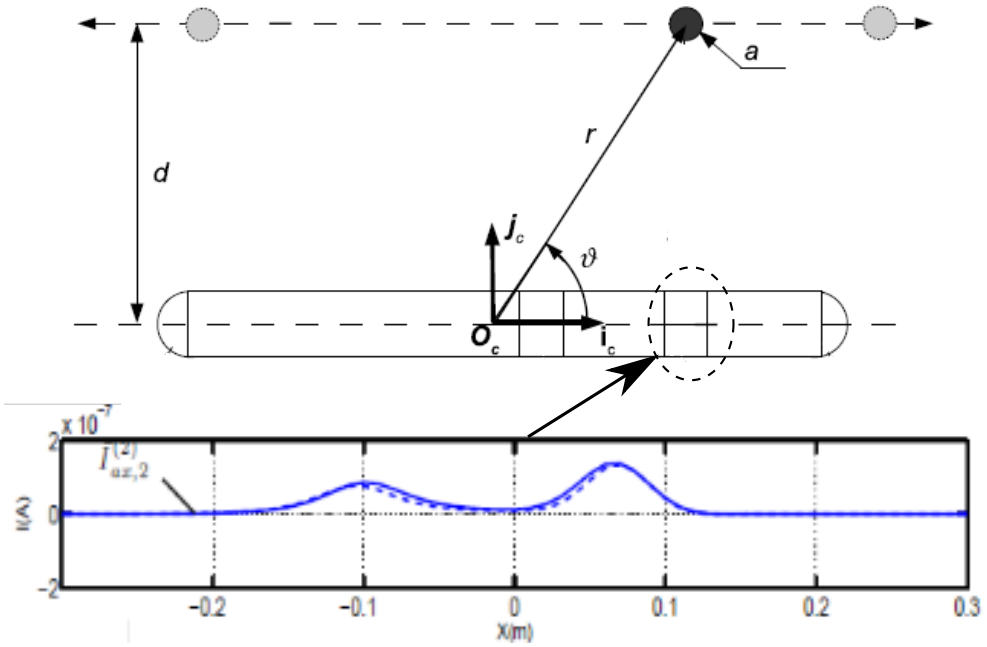


FIGURE 1.9 – Une sphère se déplace parallèlement à l'axe de la sonde élancée. Le courant est mesuré sur la bague $\mathcal{E}_2$ tout au long du déplacement de la sphère, créant ainsi une image électrique de cette dernière.

$\gamma_e = 0,04\ \mathrm{S}\cdot\mathrm{m}^{-1}$. Cependant, dans un cadre mathématique, on pose $\gamma_e = 1$ et ceci sans perte de généralité. On définit alors la conductivité dans tout le domaine Ω de la manière suivante

$$\sigma := \gamma_e \chi\left(D_e\right) + \sum_{l=1}^{p} \gamma_l \chi\left(\mathcal{O}_l\right).$$

L'équation régissant le milieu est définie comme suit

$$-\nabla\cdot\sigma\nabla u = 0 \quad \text{dans } \Omega, \tag{1.1}$$

où u représente le potentiel électrostatique. On appelle champ électrique le gradient du potentiel, c'est-à-dire ∇u. Nous pouvons remarquer plusieurs points.

1. La conductivité du milieu est considérée constante par morceaux. Cela permet d'éviter le mauvais conditionnement des problèmes étudiés pour d'autres types de conductivité. Pour plus de détails voir [31] et [10].

2. Suivant les besoins, il est possible de reformuler (1.1) à partir de conditions définies aux bords des objets. En effet

$$(1.1) \Leftrightarrow \begin{cases} -\Delta u = 0 & \text{dans } \Omega \setminus \left(\cup_{l=1}^{p} \partial\mathcal{O}_l\right), \\ [u]_{\partial\mathcal{O}_l} = 0 & \forall l \in [\![1;p]\!], \\ \left[\gamma \frac{\partial u}{\partial \nu}\right]_{\partial\mathcal{O}_l} = 0 & \forall l \in [\![1;p]\!]. \end{cases}$$

La notation ν fait référence à la normale unitaire extérieure au bord du domaine Ω ou orienté vers l'intérieur des objets. Nous avons également utilisé les notations suivantes

$$\forall l \in [\![1;p]\!], \forall x \in \partial\mathcal{O}_l, \quad u_\pm\left(x\right) := \lim_{t\to 0^+} u\left(x \pm t\nu_x\right) \quad \text{et} \quad \frac{\partial u}{\partial \nu^\pm}\left(x\right) := \nabla u_\pm\left(x\right)\cdot\nu_x,$$

ce qui nous permet de définir

$$\forall l \in [\![1;p]\!], \quad [u]_{\partial\mathcal{O}_l} := (u_- - u_+)|_{\partial\mathcal{O}_l} \text{ et } \left[\gamma\frac{\partial u}{\partial \nu}\right]_{\partial\mathcal{O}_l} := \left(\gamma_e\frac{\partial u}{\partial \nu^-} - \gamma_l\frac{\partial u}{\partial \nu^+}\right)\Big|_{\partial\mathcal{O}_l}.$$

Les crochets $[\cdot]$ dans la définition correspondent à ce que l'on appelle parfois "saut", c'est-à-dire la différence d'une quantité à l'interface de deux ouverts.

3. Soit $l \in [\![1;p]\!]$ donné. Si l'on considère le cas où $\gamma_l = 1$ comme possible, l'objet est alors de même conductivité que le milieu. Ce qui correspond à une situation sans l'objet $\mathcal{O}_l$. C'est pourquoi cette valeur est exclue. Lorsque $\gamma_l = 0$, nous sommes dans la situation où l'objet est un isolant parfait, c'est-à-dire que les conditions aux limites imposées sur le bord extérieur de l'objet sont des conditions de Neumann homogènes. Dans ce cas, on peut définir l'ouvert $\mathcal{O}_l$ de la même façon que $\mathcal{B}$ à partir de Ω. On ne l'ajoutera donc pas dans la définition de σ. En revanche, on définit comme mentionnée la condition au bord suivante

$$\left.\frac{\partial u}{\partial \nu^-}\right|_{\partial\mathcal{O}_l} = 0. \tag{1.2}$$

Enfin, lorsque γ_l devient très grand, nous nous rapprochons de la situation où l'objet est un conducteur parfait. Pour prendre en compte ce cas limite, il faut définir $\mathcal{O}_l$ de la même manière que $\mathcal{B}$. Les conditions au bord de l'objet sont alors les suivantes

$$\exists c_{o,l} \in \mathbb{R}, \quad u_-|_{\partial\mathcal{O}_l} = c_{o,l} \quad \text{et} \quad \int_{\partial\mathcal{O}_l}\frac{\partial u}{\partial \nu^-}\mathrm{d}S = 0 \tag{1.3}$$

où $c_{o,l}$ est une constante inconnue.

Définissons maintenant les conditions au bord de la sonde. On distingue trois types de secteurs à sa surface $\partial\mathcal{B}$: l'isolant $\mathcal{I}_s$; les conducteurs émetteurs $\mathcal{E}$ et les conducteurs flottants $\mathcal{F}$. Les conducteurs sont regroupés sous la notation $\mathcal{C} = \mathcal{E} \cup \mathcal{F}$ et par conséquent $\partial\mathcal{B} = \mathcal{I}_s \cup \mathcal{C}$. On découpe les conducteurs en n_c électrodes $\{e_k\}_{k\in[\![1;n_c]\!]}$ que l'on répartit en n_e électrodes émettrices et n_f électrodes flottantes. Il arrive que l'on regroupe certaines électrodes, comme l'indique la photo 1.8a dans le cas de la sonde élancée. La notation alors couramment utilisée est un indiçage de $\mathcal{E}$. Sur les secteurs isolants, on impose une condition de Neumann homogène, c'est-à-dire

$$\left.\frac{\partial u}{\partial \nu^-}\right|_{\mathcal{I}_s} = 0. \tag{1.4}$$

En effet la nature isolante de cette zone empêche aux courants de pénétrer, ainsi les charges sont obligées de circuler en surface. Les secteurs isolants assurent la continuité du potentiel sur le bord de la sonde entre les secteurs conducteurs évitant ainsi le court-circuit. À la surface des électrodes émettrices, nous imposons le potentiel ou plus précisément une différence de potentiel entre elles. Nous considérons donc des conditions de Dirichlet. Toutes les électrodes de ce type sont physiquement reliées au même générateur de puissance. La somme totale des courants à la surface de ce secteur doit par conséquent être nulle. Cela se traduit de la manière suivante

$$\forall e_k \subset \mathcal{E}, \exists c_{e,k} \in \mathbb{R}, \quad u_-|_{e_k} = c_{e,k} \quad \text{et} \quad \int_{\mathcal{E}}\frac{\partial u}{\partial \nu^-}\mathrm{d}S = 0, \tag{1.5}$$

où les $(c_{e,k})_{k\in[\![0;n_e]\!]}$ sont des constantes imposées. À la surface des conducteurs flottants le potentiel reste constant mais non déterminé. En revanche il est nécessaire d'imposer une certaine forme de

neutralité électrique. En effet, la somme des courants est nulle à la surface de chacune de ces électrodes. Ceci s'écrit

$$\forall e_k \subset \mathcal{F}, \exists c_{f,k} \in \mathbb{R}, \quad u_-|_{e_k} = c_{f,k} \quad \text{et} \quad \int_{e_k} \frac{\partial u}{\partial \nu^-} \mathrm{d}S = 0, \tag{1.6}$$

où les $\left(c_{f,k}\right)_{k\in[\![1;n_f]\!]}$ sont des constantes inconnues. Dans le cas où le bord de Ω n'est pas intégralement décrit, on impose généralement des conditions de Neumann homogènes sur $\partial\Omega \setminus \partial\mathcal{B}$, c'est-à-dire

$$\left.\frac{\partial u}{\partial \nu^-}\right|_{\partial\Omega\setminus\partial\mathcal{B}} = 0. \tag{1.7}$$

Cette partie du bord, quand elle existe, représente généralement le réservoir[5] dans lequel sont plongées les sondes. Elle peut aussi représenter un mur ou un très grand objet.

Dans le cas où Ω n'est pas borné, il faut prendre en compte des conditions d'évanescences du signal, c'est-à-dire d'écrire le comportement de u en l'infini. Les conditions choisies peuvent varier suivant le contexte, de la même manière que les conditions imposées sur le réservoir. Les conditions généralement retenues sont les suivantes : quelque soit la direction radiale ω, la fonction u vérifie

$$\begin{cases} \lim\limits_{r\to+\infty} r\,|u(r\omega)| < +\infty \\ \lim\limits_{r\to+\infty} r\frac{\partial u}{\partial \omega}(r\omega) = 0. \end{cases} \tag{1.8}$$

Le choix de cette hypothèse se justifie physiquement ainsi : on considère que le champ s'annule lorsqu'il est évalué loin des charges générées par la sonde.

Le système $((1.1) - (1.8))$ recouvre un grand nombre de cas pratiques. Le terme imposé sur le capteur fait généralement référence aux types de conditions choisies sur les électrodes et les valeurs prescrites le cas échéant. Par la suite les quantités suivantes

$$\forall e_k \subset \mathcal{E}, \quad I_k := \gamma_e \int_{e_k} \frac{\partial u}{\partial \nu^-}\, \mathrm{d}S, \tag{1.9}$$

$$\forall e_k \subset \mathcal{F}, \quad J_k := u|_{e_k}, \tag{1.10}$$

seront appelées *mesures*. En effet ces grandeurs $(I_k)_{k\in[\![1;n_e]\!]}$ et $(J_k)_{k\in[\![1;n_f]\!]}$ représentent respectivement les courants et les potentiels accessibles physiquement par mesure sur le robot. Précisons que dans la suite du document, nous n'indiquerons plus les symboles "+" et "-" sur les fonctions définies aux bords d'un ouvert, lorsque cela ne prête pas à confusion.

On peut à présent formuler les deux grands types de problèmes.

Problème 1.2.1 (problème direct)**.** *Étant donnée la géométrie du capteur, des objets et les paramètres électriques imposés, déterminer la mesure sur le capteur.*

Problème 1.2.2 (problème inverse)**.** *Étant donnée la géométrie du capteur ainsi que ses paramètres électriques imposés et mesurés, on cherche à déterminer les paramètres géométriques et électriques des objets.*

Le problème inverse 1.2.2 est généralement appelé problème d'électrolocation ou d'électrolocalisation.

[5] On parlera aussi de conteneur.

1.2.2 Représentation par intégrales de frontière et leurs versions numériques

Comme fil conducteur de cette sous-section, nous proposons le problème direct suivant.

Problème 1.2.3 (Problème de Dirichlet vers Neumann). *Soit Ω un domaine à bord régulier de $\mathbb{R}^{n_d}$. Quelque soit une fonction f admissible du bord de Ω, on cherche à déterminer $\left.\frac{\partial u}{\partial \nu}\right|_{\partial\Omega}$, où u est la solution du problème de Laplace, muni d'une condition de Dirichlet, suivant*

$$\begin{cases} -\Delta u = 0 & \text{dans } \Omega, \\ u = f & \text{sur } \partial\Omega. \end{cases} \tag{1.11}$$

Pour résoudre un tel problème, on en constitue généralement une formulation faible, que l'on résout numériquement à l'aide de la méthode des éléments finis (FEM). Nous nous proposons de présenter une alternative à cette approche.

Celle-ci repose sur la formulation par intégrale de frontière du problème, dont l'origine est liée à la théorie du potentiel. Celle-ci a été développée au début du siècle dernier, entre autres dans [25] et [19]. Lorsque l'on a affaire à des opérateurs différentiels, cette formulation est généralement possible. Celle-ci ne porte que sur des quantités définies au bord du domaine. Les avantages sont donc nombreux. On peut citer, entre autres, la facilité de prendre en compte des bords déformables et la résolution de problème extérieur, c'est-à-dire le cas où Ω est non borné. Les formulations au bord ou dans l'espace, ainsi que leurs méthodes numériques associées, n'ont pas à être mises en opposition. Au contraire, il est même possible de coupler les deux. Par exemple, la première peut traiter une situation locale dans un ouvert, tandis que l'autre décrit les influences en l'infini.

Construisons à présent une représentation intégrale pour le problème 1.2.3. Ce passage s'inspire de [9], [4] et [22]. Pour plus de détails, on se reportera à ces références et à leurs bibliographies. Cette formulation fait intervenir les deux ingrédients suivants.

1. Une formule de Green

$$\int_\Omega \left(\Delta v(y)u(y) - v(y)\Delta u(y)\right)\, \mathrm{d}V_y = \int_{\partial\Omega} \left(\frac{\partial v}{\partial \nu}(y)u(y) - v(y)\frac{\partial u}{\partial \nu}(y)\right)\, \mathrm{d}S_y, \tag{1.12}$$

   vraie pour toute fonction test v admissible.

2. La solution fondamentale ou fonction de Green G, à prendre au sens des distributions, vérifiant l'équation

$$\forall x \in \Omega, \forall y \in \Omega, \quad \Delta_y G\left(x,y\right) = \delta_x\left(y\right),$$

   et dont l'expression suivant la dimension n_d est

$$G\left(x,y\right) = \begin{cases} \frac{1}{2\pi}\ln|x-y| & \text{si } n_d = 2, \\ \frac{|x-y|^{2-d}}{(2-d)\omega_{n_d}} & \text{si } n_d > 2, \end{cases}$$

   où ω_{n_d} représente l'aire de l'hypersphère unité de dimension n_d.

Pour tout $x \in \Omega$, si l'on pose $v = G(x,\cdot)$, alors la formule (1.12) devient

$$u(x) - \int_\Omega G(x,y)\Delta_y u(y)\, \mathrm{d}V_y = \int_{\partial\Omega} \frac{\partial G}{\partial \nu}(x,y)u(y)\, \mathrm{d}S_y - \int_{\partial\Omega} G(x,y)\frac{\partial u}{\partial \nu}(y)\, \mathrm{d}S_y.$$

Comme on a supposé que u est harmonique dans Ω, on a

$$\forall x \in \Omega, \quad u(x) = \int_{\partial\Omega} \frac{\partial G}{\partial \nu}(x,y)u(y)\, \mathrm{d}S_y - \int_{\partial\Omega} G(x,y)\frac{\partial u}{\partial \nu}(y)\, \mathrm{d}S_y. \tag{1.13}$$

On peut aussi considérer u comme étant la solution d'une équation de Poisson. Dans ce cas le terme de volume dans l'expression précédente aurait été conservé. On définit ensuite les opérateurs de simple couche

$$\forall x \in \mathbb{R}^{n_d}, \quad \mathcal{S}_\Omega\left(v_1\right)(x) := \int_{\partial\Omega} G(x,y) v_1(y)\, \mathrm{d}S_y$$

et de double couche

$$\forall x \in \mathbb{R}^{n_d} \setminus \partial\Omega, \quad \mathcal{D}_\Omega\left(v_2\right)(x) := \int_{\partial\Omega} \frac{\partial G}{\partial \nu}(x,y) v_2(y)\, \mathrm{d}S_y,$$

où v_1 et v_2 sont deux fonctions admissibles du bord de Ω. Ce qui nous permet de réécrire (1.13) de la manière suivante

$$\forall x \in \Omega, \quad u(x) = \mathcal{D}_\Omega\left(u\right)(x) - \mathcal{S}_\Omega\left(\frac{\partial u}{\partial \nu}\right)(x). \tag{1.14}$$

La formulation précédente permet de déterminer la solution de l'équation (1.11) à l'intérieur de Ω. Deux problèmes subsistent dans cette expression.

1. La quantité $\frac{\partial u}{\partial \nu}$ inconnue. C'est même la quantité que nous allons en fait déterminer.
2. Nous ne cherchons pas à caractériser la solution u dans tout le domaine. Seules les valeurs $\partial\Omega$ nous intéressent, puisque la formule (1.13) montre qu'elles sont les seules nécessaires à la détermination complète de u.

L'équation (1.14) est donc sous-déterminée. Pour régler ce problème, nous faisons tendre $x \in \Omega$ vers le bord $\partial\Omega$ dans l'expression (1.14) ce qui donne

$$\forall x \in \partial\Omega, \quad u(x) = \left(\frac{1}{2}\mathcal{I}_\Omega + \mathcal{D}_\Omega\right)(u)(x) - \mathcal{S}_\Omega\left(\frac{\partial u}{\partial \nu}\right)(x). \tag{1.15}$$

Cette fois-ci l'opérateur de double couche est à prendre au sens de la valeur principale. La notation $\mathcal{I}_\Omega$ désigne l'opérateur identité défini sur le bord de Ω. Le coefficient devant l'identité est une constante liée localement à la régularité du bord. Dès lors que le bord n'est pas régulier, cette constante peut être amenée à changer. On peut réécrire (1.15) plus simplement

$$\frac{1}{2}\, u|_{\partial\Omega} = \mathcal{D}_\Omega\left(u|_{\partial\Omega}\right) - \mathcal{S}_\Omega\left(\left.\frac{\partial u}{\partial \nu}\right|_{\partial\Omega}\right). \tag{1.16}$$

Pour être tout à fait exhaustif la formule (1.15) peut être étendue à l'extérieur de l'ouvert Ω ainsi

$$\forall x \in \mathbb{R}^{n_d}, \quad \mathcal{D}_\Omega\left(u\right)(x) - \mathcal{S}_\Omega\left(\frac{\partial u}{\partial \nu}\right)(x) = \begin{cases} u(x) & \text{si } x \in \Omega, \\ \frac{u(x)}{2} & \text{si } x \in \partial\Omega, \\ 0 & \text{si } x \in \mathbb{R}^{n_d} \setminus \overline{\Omega}. \end{cases}$$

Pour résoudre le problème 1.2.3, nous rassemblons dans (1.16) les termes inconnus dans le membre de gauche et les données dans le membre de droite, ce qui donne l'équation linéaire suivante

$$\mathcal{S}_\Omega\left(\left.\frac{\partial u}{\partial \nu}\right|_{\partial\Omega}\right) = \left(\frac{1}{2}\mathcal{I}_\Omega - \mathcal{D}_\Omega\right)\left(u|_{\partial\Omega}\right). \tag{1.17}$$

Nous pouvons à présent résoudre numériquement le problème 1.2.3, en discrétisant l'équation (1.17). Pour résoudre un problème par formulations intégrales de frontière, on peut distinguer deux grands types de méthode :

Éléments de frontière : approche par collocation. Cette méthode consiste à discrétiser l'équation intégrale prise en un nombre fini de points dits de collocation.

Éléments finis de frontière. Cette méthode est une méthode de type Galerkin. Elle consiste à dériver de (1.17) une formulation variationnelle. Celle-ci est obtenue en multipliant les formulations intégrales de frontière par des fonctions tests appropriées et en intégrant le tout. Cette méthode s'apparente aux éléments finis.

Ces deux méthodes sont généralement désignées par l'acronyme anglais BEM, pour *Boundary Element Methods*. La méthode des éléments de frontière par collocation est la plus populaire des deux en raison de sa souplesse de mise en œuvre. Des versions en deux et trois dimensions ont été développées au sein du projet ANGELS. Elles servent de référence pour évaluer la précision des méthodes plus spécifiques. Pour les besoins du chapitre 3, nous présentons les étapes clés de sa construction en deux dimensions. L'exposé de ce passage est inspiré de [9] et du cours de H. Lemonnier [30], version détaillée de l'article [31]. Les notations sont en grande partie tirées de ce dernier. Les références couramment citées sur le sujet sont [21] et [13]. Cette liste n'étant pas exhaustive, on se référera à ces documents et à leurs bibliographies pour des informations complémentaires.

Pour constituer un maillage du bord de Ω, nous découpons sa frontière par des éléments linéaires, ceci afin d'approcher $\partial\Omega$ par un polygone. C'est pourquoi nous définissons un ensemble de $\mathrm{N}_\Omega \in \mathbb{N}$ points $\{P_i\}_{i\in[\![1;\mathrm{N}_\Omega]\!]}$ appartenant à $\partial\Omega$. Les éléments $\{E_i\}_{i\in[\![1;\mathrm{N}_\Omega]\!]}$ de notre maillage sont alors définis comme étant les segments d'extrémités $\{P_i\}_{i\in[\![1;\mathrm{N}_\Omega]\!]}$, c'est-à-dire

$$\forall i \in [\![1;\mathrm{N}_\Omega]\!], \quad E_i := \begin{cases} [P_i;P_{i+1}] & \text{si } i \neq \mathrm{N}_\Omega, \\ [P_i;P_1] & \text{sinon .} \end{cases}$$

On suppose que $u|_{\partial\Omega}$ et $\frac{\partial u}{\partial \nu}\Big|_{\partial\Omega}$ sont constantes sur chacun des éléments. Soit N_Ω points $\{M_i\}_{i\in[\![1;\mathrm{N}_\Omega]\!]}$ représentant les milieux des éléments et Id_Ω la matrice identité d'ordre N_Ω. On peut alors considérer la version discrétisée de l'équation (1.17)

$$\left(\frac{1}{2}\,\mathrm{Id}_\Omega - \mathrm{D}_\Omega\right) U_\Omega = S_\Omega \partial U_\Omega, \tag{1.18}$$

où

$$\forall i \in [\![1;\mathrm{N}_\Omega]\!], \quad U_{\Omega,i} := f(M_i) \quad \text{et} \quad \partial U_{\Omega,i} := \frac{\partial u}{\partial \nu}(M_i) \tag{1.19}$$

et les coefficients des opérateurs de simple et double couches sont donnés par

$$\forall (i,j) \in [\![1;\mathrm{N}_\Omega]\!]^2, \quad (\mathrm{S}_\Omega)_{i,j} := \int_{y\in E_j} G(M_i,y)\,\mathrm{d}S_y \;\text{ et }\; (\mathrm{D}_\Omega)_{i,j} := \int_{y\in E_j} \frac{\partial G}{\partial \nu}(M_i,y)\,\mathrm{d}S_y.$$

La donnée f du problème ne peut pas être évaluée en dehors de $\partial\Omega$ comme cela est pourtant indiqué dans (1.19). C'est pourquoi on choisit le point de $\partial\Omega$ situé au milieu des extrémités de l'élément comme valeur d'approximation. Pour ne pas alourdir les notations, nous n'avons pas introduit un nouveau point. Soit $(i,j) \in [\![1;\mathrm{N}_\Omega]\!]^2$. On définit la tangente t_j comme étant la tangente à l'élément E_j, orientée par la numérotation de ses extrémités et ceci dans le sens croissant. La normale ν_j est alors déduite de t_j en lui appliquant une rotation de $-\frac{\pi}{2}$. L'expression des coefficients $(\mathrm{S}_\Omega)_{i,j}$ et $(\mathrm{D}_\Omega)_{i,j}$ s'obtient ensuite simplement dans la base locale de l'élément E_j, centrée sur M_j et dont les vecteurs de base sont t_j et ν_j. On se référera à la figure 1.10 pour une description graphique de la situation. On note L la longueur de l'élément E_j. Dans la base locale de l'élément E_j, le point de collocation M_i a pour coordonnées

$$x_t = -M_jM_i \cdot t_j \quad \text{et} \quad x_\nu = -M_jM_i \cdot \nu_j.$$

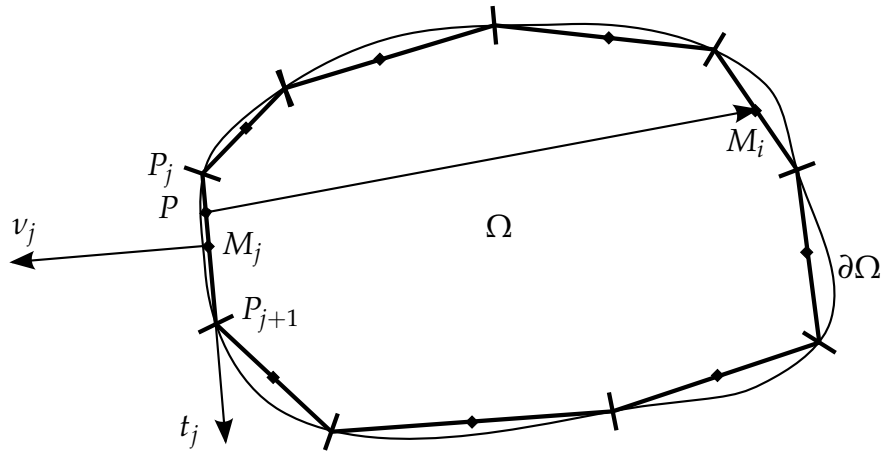


FIGURE 1.10 – Bord du domaine Ω approché par un polygone. Description des donnés géométriques principales de l'élément E_j.

En introduisant enfin les notations

$$x_1 = x_t - \frac{L}{2} \quad \text{et} \quad x_2 = x_t + \frac{L}{2},$$

on a

$$(\mathrm{S}_\Omega)_{i,j} := \begin{cases} \frac{-1}{4\pi}\left(x_1 \ln\left(x_1^2 + x_\nu^2\right) - x_2 \ln\left(x_2^2 + x_\nu^2\right) + 2x_\nu \arctan\left(\frac{x_\nu(x_1-x_2)}{x_1x_2+x_\nu^2}\right) + 2d\right) & \text{si } i \neq j, \\ \frac{1}{4\pi}\left(d \ln \frac{d^2}{4} - 2d\right) & \text{sinon} \end{cases}$$

et

$$(\mathrm{D}_\Omega)_{i,j} := \begin{cases} \frac{1}{2\pi} \arctan\left(\frac{x_\nu(x_1-x_2)}{x_1x_2+x_\nu^2}\right) & \text{si } i \neq j, \\ 0 & \text{sinon.} \end{cases}$$

Attention dans chacune de ces expressions, il faut tenir compte des signes respectifs du numérateur et du dénominateur pour évaluer l'arctangente[6]. Une fois le système (1.18) constitué, il ne reste plus qu'à inverser S_Ω pour obtenir une approximation de la solution du problème.

Des expressions similaires existent en trois dimensions. On peut bien évidemment considérer des maillages dont les points sont situés uniquement sur $\partial\Omega$ et des discrétisations d'ordre plus élevées. Dans ces cas, la quadrature nécessite généralement le recours à une méthode d'interpolation. Pour plus d'informations, on se référera à la bibliographie citée plus haut.

1.2.3 Méthode des réflexions

La méthode des réflexions constitue un point essentiel de cette thèse et est l'objet d'un chapitre dédié, le chapitre 2. Cette méthode est une méthode du type *diviser pour mieux régner*. Pour un problème direct donné, l'idée est de considérer des sous-problèmes plus élémentaires. En considérant alternativement la résolution de ces sous-problèmes, on reconstitue, en appliquant le principe de superposition, la solution du problème de départ. Elle est très similaire à la méthode de décomposition de domaine, dans le sens où elle aurait pu être appelée méthode de décomposition de frontières. Cette dernière est un classique de l'hydrodynamique à bas Reynolds. En effet, son origine est généralement attribuée à un travail de M. Smoluchowski dans [45]. Dans cet article, il l'emploie à déterminer les forces hydrodynamiques s'exerçant sur un ensemble de sphères chutant à l'intérieur d'un fluide visqueux. Sont généralement cités comme livres de références sur le

[6] Sous Octave il faut utiliser la fonction `atan2`.

sujet : le livre [20] de J. Happel et H. Brenner et le livre [26] de S. Kim et S. J. Karrila. Attention tout de même à une petite imprécision dans [20], sur laquelle nous reviendrons dans le chapitre 2. Une description détaillée de la méthode, de son historique et de la bibliographie associée est présentée dans le chapitre 2. L'article [12] ayant motivé ces travaux, nous présentons ici une première approche de la méthode à partir du modèle considéré dans cet article. Précisons tout de même que la convergence de ce cas particulier est encore un problème ouvert. Un problème s'en rapprochant, appelé problème mixte, a tout de même été traité. Pour plus de détails se référer au chapitre 2.

Nous nous plaçons dans la situation où Ω est le complémentaire dans $\mathbb{R}^{n_d}$ du capteur $\mathcal{B}$ et d'un objet $\mathcal{O}$ isolant, c'est-à-dire $\Omega := \mathbb{R}^{n_d} \setminus \overline{(\mathcal{B} \cup \mathcal{O})}$. Soit f une fonction constante par morceaux définie sur $\mathcal{E}$. Nous ne préciserons pas ici les espaces fonctionnels, ainsi que les conditions d'évanescence en l'infini, afin de ne pas encombrer cette première présentation. On se référera au chapitre 2 pour des informations à ce sujet. Nous cherchons à résoudre le problème direct 1.2.1 pour la condition f, c'est-à-dire que l'on veut déterminer la mesure $\frac{\partial u}{\partial \nu}\Big|_{\mathcal{E}}$, sachant que u est la solution du problème suivant

$$\begin{cases} -\Delta u = 0 & \text{dans } \Omega, \\ u = f & \text{sur } \mathcal{E}, \\ \frac{\partial u}{\partial \nu} = 0 & \text{sur } \mathcal{I}_s, \\ \frac{\partial u}{\partial \nu} = 0 & \text{sur } \partial\mathcal{O}. \end{cases} \tag{1.20}$$

Soit $\Omega_0 := \mathbb{R}^{n_d} \setminus \overline{\mathcal{B}}$ et $\Omega_1 := \mathbb{R}^{n_d} \setminus \overline{\mathcal{O}}$. Pour résoudre ce problème, nous considérons deux sous-problèmes liés au problème (1.20). Le premier est centré sur le capteur. Quelque soit f_0 et g_0 deux fonctions admissibles définies respectivement sur $\mathcal{E}$ et $\mathcal{I}_s$, on cherche à déterminer v la solution du problème suivant

$$\begin{cases} -\Delta v = 0 & \text{dans } \Omega_0, \\ v = f_0 & \text{sur } \mathcal{E}, \\ \frac{\partial v}{\partial \nu} = g_0 & \text{sur } \mathcal{I}_s. \end{cases} \tag{1.21}$$

Le deuxième problème est centré sur l'objet. Quelque soit g_1 une fonctions admissible définie sur $\partial\mathcal{O}$, on cherche à déterminer w la solution du problème suivant

$$\begin{cases} -\Delta w = 0 & \text{dans } \Omega_1, \\ \frac{\partial w}{\partial \nu} = g_1 & \text{sur } \partial\mathcal{O}. \end{cases} \tag{1.22}$$

La solution u est alors envisagée comme étant une somme de termes, alternativement solution du problème (1.21) et du problème (1.22) :

$$u = \sum_{m=0}^{+\infty} u_m.$$

Ces différentes solutions sont nommées réflexions. On en déduit par linéarité la mesure comme étant la somme des traces normales sur $\mathcal{E}$ du gradient de ces différentes réflexions, c'est-à-dire

$$\frac{\partial u}{\partial \nu}\Big|_{\mathcal{E}} = \sum_{m=0}^{+\infty} \frac{\partial u_m}{\partial \nu}\Big|_{\mathcal{E}}.$$

Expliquons maintenant comment sont construites ces différentes solutions. La première u_0, est la solution du problème (1.21), pour les conditions aux bords suivantes : $f_0 = f$ et $g_0 = 0$. Ceci revient à considérer le problème de départ sans objet. Ce problème est généralement appelé problème basal. Par rapport à u et aux données du problème, la solution u_0 est exacte sur le capteur.

En revanche les conditions aux bords de l'objet ne correspondent pas. Pour rééquilibrer cette situation, on considère la réflexion u_1 comme étant la solution du problème (1.22), pour la donnée $g_1 = -\frac{\partial u_0}{\partial \nu}\Big|_{\partial\mathcal{O}}$. Comme par construction

$$\frac{\partial u_0}{\partial \nu}\Big|_{\partial\mathcal{O}} + \frac{\partial u_1}{\partial \nu}\Big|_{\partial\mathcal{O}} = 0,$$

on en déduit que la fonction harmonique $u_0 + u_1$ est exacte sur $\partial\mathcal{O}$. Malheureusement, en raison de la contribution de u_1, elle ne l'est pas sur $\partial\mathcal{B}$. De la même façon pour rééquilibrer la situation, on considère la solution u_2 cette fois-ci du problème (1.21), pour les données $f_0 = -u_1|_{\mathcal{E}}$ et $g_0 = -\frac{\partial u_1}{\partial \nu}\Big|_{\mathcal{I}_s}$. On obtient donc une fonction harmonique $u_0 + u_1 + u_2$ exacte sur $\partial\mathcal{B}$, mais pas sur $\partial\mathcal{O}$. On réitère alors la démarche. Dès lors que les réflexions décroissent au fur et à mesure en norme, le procédé converge vers la solution. La méthode est illustrée par la figure 1.11.

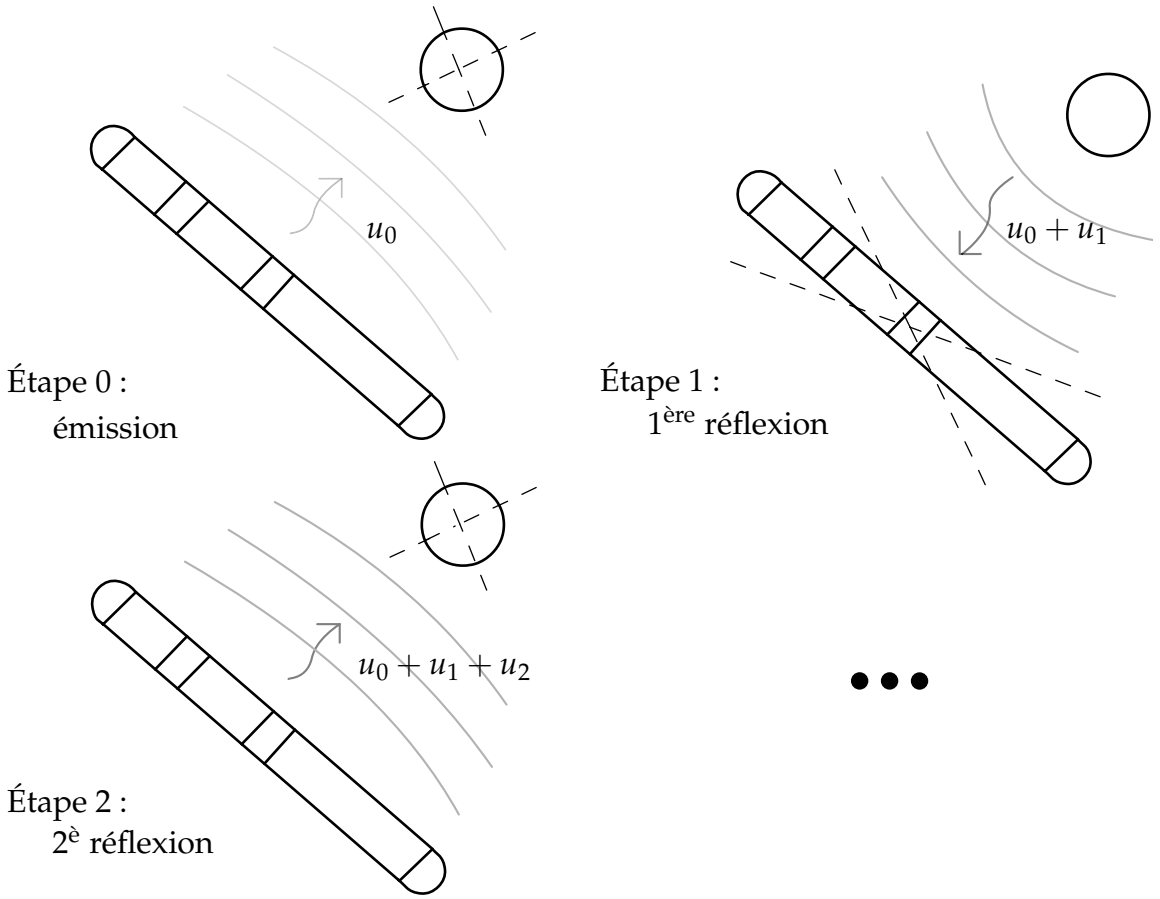


FIGURE 1.11 – Schéma descriptif de la méthode des réflexions tiré de [12].

Dans le chapitre 2, nous présentons une version séquentielle et parallèle de la méthode et ceci pour un nombre quelconque d'objets. La version séquentielle multi-objets semble être un algorithme original. Au-delà des aspects algorithmiques, le résultat central de ce chapitre est la mise en lien de la méthode des réflexions avec celle des projections alternées. Ceci nous permet d'en d'exploiter les différentes propriétés. Sont ensuite présentées des applications pour des conditions aux bords variées. Ce chapitre est une version avancée d'un futur article.

1.3 Résultats sur le sens électrique appliqué à la robotique

L'un des objectifs principaux du projet ANGELS est la commande du robot, afin d'assurer son autonomie. Il faut préciser que le cahier des charges stipule le fait que les algorithmes, développés à cette fin, doivent réaliser les calculs en temps réel et être faciles à mettre en œuvre par le matériel embarqué. Aux travers des choix et des investigations de l'équipe de Nantes, cette section présente une bibliographie sur ce point et un sujet connexe. Elle introduit également les travaux du chapitre 3.

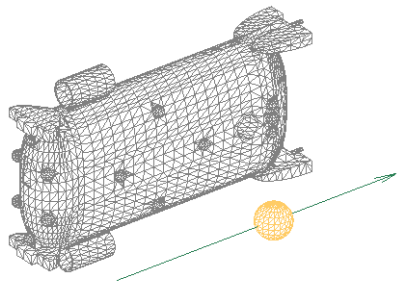
L'une des idées développées au sein de l'équipe est une commande réflexe dont on peut trouver une description dans [11]. Cette technique est de bas niveau car elle est directement liée à la mesure. Entre autres applications elle permet au robot de s'aligner avec un objet, de s'en approcher ou de s'en éloigner, ou encore d'en suivre les contours. Décrivons par exemple le cas de l'alignement lorsque un objet entre dans sa bulle de perception. Le capteur a une vitesse d'avancement constante et s'oriente suivant une différence gauche-droite de la mesure. Dès lors que deux électrodes sont placées de part et d'autre du robot et ceci de façon symétrique, cette différence lui permet d'ajuster régulièrement son cap. Une fois la différence annulée, il est orienté face à l'objet. Cette méthode est très efficace, mais nécessite en pratique la description de chaque scénario avant implémentation.

Une autre stratégie peut consister à résoudre le problème d'électrolocation. Si ce problème est résolu, le robot détient une carte locale de son environnement plus ou moins détaillée suivant la précision de la méthode. Ce sujet d'investigation a été dynamisé par l'article [18]. En effet, les auteurs, équipe de biologistes spécialistes du sens électrique chez le poisson rattachés au projet ANGELS, décrivent la capacité du poisson à percevoir la distance qui le sépare d'un objet, sa forme et ses paramètres électriques. Suite à cela, comme nous l'avions déjà mentionné, l'équipe de M. MacIver a présenté un algorithme d'électrolocation dans [46]. Celui-ci est basé sur une méthode probabiliste, appelée filtre particulaire. La reconstruction des objet est précise au millimètre, mais nécessite un temps de calcul trop important et une initialisation adéquate de la scène. De plus, elle est adaptée à un appareillage trop éloigné d'un véritable robot. Une méthode probabiliste a aussi été utilisée par l'équipe de Nantes, le filtre de Kalman. C'est une technique couramment utilisée en automatique, qui offre une alternative au filtre particulaire. Dans [29], la méthode est appliquée afin de reconstruire une petite sphère ou un mur. Pour une bibliographie des résultats plus poussés sur le sujet, chez les PFE et en robotique, le lecteur peut se référer à la bibliographie de [24] et de [10]. À l'heure actuelle, la bibliographie la plus fournie se situe dans une thématique connexe : la tomographie par impédance. Largement utilisée dans le domaine médical, cette méthode d'imagerie non intrusive consiste à mesurer la conductivité et la permittivité de certaines parties du corps à partir d'électrodes positionnées sur la peau du patient. Cette technique a des applications dans d'autres domaines comme la géophysique. Pour un descriptif détaillé des méthodes liées au domaine médical voir [4] et [1]. Récemment, l'article [31] de H. Lemonnier et J.F. Peytraud a tout particulièrement retenu notre attention. Cet article porte sur les liquides dysphasiques. L'objectif des auteurs est de reconstruire des inclusions présentes dans un fluide à partir d'électrodes positionnées autour d'un tube isolant. L'unique différence avec notre sujet d'étude est que ce travail porte sur un problème intérieur et non extérieur. L'une de nos perspectives est d'étudier cet article et, s'il est pertinent pour notre cas, de l'appliquer. Nous y reviendrons par la suite, dans le chapitre 3.

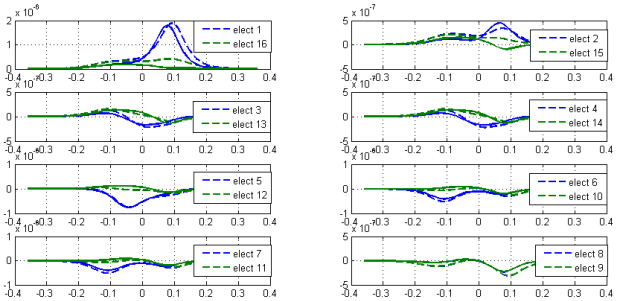
La résolution du problème d'électrolocation a généralement recours à un simulateur du problème direct. C'est pourquoi nous présentons ici différents modèles constitués pour les sondes de type élancé ou ANGELS. Rappelons que la démarche habituellement suivie par l'équipe est de concevoir des méthodes adaptées aux sondes élancées puis de les adapter à la sonde ANGELS dans le réservoir pour finalement le mettre en œuvre dans une implémentation dans un module ANGELS autonome. Le premier simulateur a avoir été développé est une BEM trois dimensions.

Nous renvoyons le lecteur intéressé à la sous-section 1.2.2 pour plus de détails sur la méthode. Ce simulateur très chronophage a été conçu pour les deux types de sondes afin d'évaluer le degré de précision des autres méthodes. Dans [23] est décrit la première réduction de modèle développée pour les sondes élancées et nommée modèle poly-sphériques. Cette méthode consiste à substituer chaque électrode par une sphère et à négliger les isolants. L'objet est généralement une petite sphère située à une distance raisonnable du capteur. Ce qui revient, peu ou prou, à développer au premier ordre la géométrie et les données électriques considérées à leur surface. Il est également possible de prendre en compte un objet très grand, comme un mur. On utilise alors la méthode des images, qui nous ramène à une situation constituée uniquement de petits objets. La méthode poly-sphérique est efficace, mais nécessite tout de même un calibrage du rayon des sphères représentant le capteur. Ce calibrage a été affiné dans [11]. A la suite de ces premiers travaux, comme présentée dans la sous-section précédente, l'équipe a eu recours à la méthode des réflexions. Dans cette méthode, chaque réflexion est approchée par un développement des quantités électriques et géométriques au premier ordre. Le problème est que ces développements sont surtout adaptés à des capteurs élancés et des petits objets. En effet, dans ce cas l'effet des zones isolantes présentes à la surface de la sonde peut être négligé en particulier lorsque celle-ci est soumise à un champ externe.

En revanche, il se trouve que pour la sonde ANGELS, la proportion d'isolants par rapport au conducteur n'est plus négligeable, ce qui entraîne des erreurs trop importantes sur les mesures quand l'objet est trop grand et/ou trop proche. Dans le cas contraire le simulateur fournit une approximation correcte. Cette situation est illustrée dans le graphique 1.12. Expliquons intuitive-



(a) Représentation graphique de la scène. La sphère isolante de rayon 2 cm, se déplace parallèlement au capteur, à une distance de 11 cm de l'axe central de ce dernier.



(b) Mesures effectuées, en ampères, sur les électrodes de la sonde. Les traits continus représentent les résultats BEM, tandis que les traits discontinus ceux du simulateur.

FIGURE 1.12 – Exemple illustrant le défaut en courant sur la sonde ANGELS lorsqu'il est calculé par le simulateur.

ment le problème posé par les parties isolantes. Le calcul est basé sur l'exemple de la sous-section 1.2.3 et la méthode des réflexions. Cette dernière permet de traiter chaque élément constitutif de la scène de manière indépendante en introduisant les interactions à partir des conditions de bord. Ceci donne lieu à un développement en série de la perturbation mesurée. Chacun des termes représente soit un retour induit par l'objet, soit une réaction de la sonde à ce dernier ; c'est-à-dire qu'il représente respectivement soit une solution de l'équation (1.22), soit une solution de l'équation (1.21). Intéressons nous au deuxième cas, en l'illustrant par la réaction au premier retour de

l'objet ; c'est-à-dire déterminer la solution u_2 de l'équation

$$\begin{cases} -\Delta u_2 = 0 & \text{dans } \Omega_0, \\ u_2 = -u_1 & \text{sur } \mathcal{E}, \\ \frac{\partial u_2}{\partial \nu} = -\frac{\partial u_1}{\partial \nu} & \text{sur } \mathcal{I}_s. \end{cases} \tag{1.23}$$

En développant au premier ordre les équations intégrales, *cf* sous-section 1.2.2, de ce dernier problème par rapport aux données géométriques de la sonde élancée, seuls les termes liés aux électrodes restent. Ce qui revient, en première approximation, à n'évaluer les données électriques qu'uniquement sur les électrodes et à négliger les isolants. Ce calcul revient à approcher le problème (1.23) par le suivant : déterminer la solution u_2' de l'équation

$$\begin{cases} -\Delta u_2' = 0 & \text{dans } \Omega_0, \\ u_2' = -u_1 & \text{sur } \mathcal{E}, \\ \frac{\partial u_2}{\partial \nu} = 0 & \text{sur } \mathcal{I}_s. \end{cases} \tag{1.24}$$

On constate donc que les isolants ne sont pris en compte qu'à travers la réaction de la sonde. Cette approximation est, comme mentionnée plus haut, valide dès lors que la surface $\mathcal{I}_s$ est négligeable par rapport à la surface $\mathcal{E}$. Pour cette raison, au cours du travail préliminaire, nous avons introduit une étape intermédiaire. Elle consistait à déterminer le potentiel induit par le capteur soumis à un champ externe, dans le cas où toute sa coque $\partial\mathcal{B}$ est considérée comme un isolant ; c'est-à-dire de considérer une approximation de la solution du problème suivant : déterminer la solution w_2 de l'équation

$$\begin{cases} -\Delta w_2 = 0 & \text{dans } \Omega_0, \\ \frac{\partial w_2}{\partial \nu} = -\frac{\partial u_1}{\partial \nu} & \text{sur } \partial\mathcal{B}. \end{cases} \tag{1.25}$$

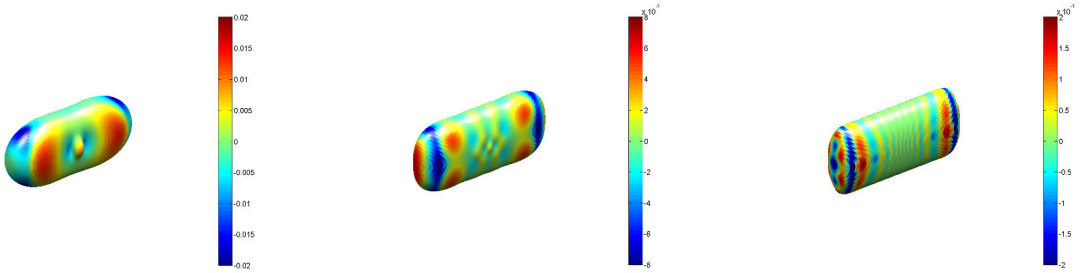
En sommant w_2 à la solution w_2' de l'équation

$$\begin{cases} -\Delta w_2' = 0 & \text{dans } \Omega_0, \\ w_2' = -u_1 - w_2 & \text{sur } \mathcal{E}, \\ \frac{\partial u_2}{\partial \nu} = 0 & \text{sur } \mathcal{I}_s, \end{cases} \tag{1.26}$$

équation du même type que (1.24), on retrouve bien la solution u_2 du problème (1.23).

Une façon d'approcher la solution du problème (1.25) consiste à décomposer la géométrie du capteur et les valeurs électriques sur la base des harmoniques sphériques. Pour illustrer cette étape, le graphique 1.13 présente la décomposition de la coque de la sonde ANGELS $\partial\mathcal{B}$ sur cette base. Le nombre d'harmoniques nécessaires pour reconstruire la forme, s'est avéré prohibitif. De plus, la solution de l'équation (1.25) ainsi obtenue, s'est avérée de mauvaise qualité.

Malgré ce mauvais résultat, nous avons persévéré dans cette direction, en ayant pour but d'appliquer un autre type de réduction à l'équation (1.25). Nous avons ainsi exploité l'idée d'une réduction de modèle par bases réduites. Quitte à employer ce type de méthode, nous avons fait le choix de revenir au problème direct complet. Ceci a donné lieu à la démarche et aux résultats présentés dans le chapitre 3. Ce chapitre propose de combiner une inversion géométrique et une réduction de modèle pour la résolution du problème direct. La première méthode permet de traiter les objets bornés ou non de la même façon. Tandis que la seconde, basée sur des pré-calculs, permet de diminuer significativement la taille des systèmes à résoudre une fois le simulateur implanté dans le robot. Précisons que dans ce chapitre certaines simplifications de modèles ont été mises en œuvre afin de simplifier la démarche mathématique. Ce chapitre constitue donc une première étape et ouvre des perspectives à plus long terme. Ces deux méthodes ont été sélectionnées car elles s'étendent naturellement à des situations plus réalistes.



(a) Avec 25 éléments de base.

(b) Avec 100 éléments de base.

(c) Avec 676 éléments de base.

FIGURE 1.13 – Approximations de la géométrie de la sonde ANGELS à partir des harmoniques sphériques. L'unité de l'erreur est le mètre. Le maximum de l'erreur absolue est respectivement environ égal à 20,3 mm, 8,8 mm et 2,8 mm. Le maximum de l'erreur relative est d'environ 31,9%, 8,1% et 2,6%.

1.4 Chimie quantique

Cette section constitue l'introduction des problématiques liées au chapitre 4. Ce dernier traite d'un problème inverse dans le domaine de la chimie quantique. Avant d'exposer cette problématique, il est intéressant d'évoquer le problème de contrôlabilité directe et d'en connaître son origine.

À l'apparition des premiers lasers, dans les années 60, s'est rapidement posée la question suivante : est-il possible de modifier la structure d'une ou d'un ensemble de molécules à partir d'une onde lumineuse bien choisie ? Cette question renvoie à un problème de contrôlabilité. Avant d'exposer plus formellement ce problème, posons les bases du modèle que nous étudierons tout au long de la partie dédiée à la question dans cette thèse.

Nous nous plaçons naturellement dans le cadre de la physique quantique puisque nous cherchons à décrire des systèmes physiques à l'échelle atomique, voire subatomique. Soit un temps $T > 0$ et un système de N particules isolées et sans contrainte. Pour représenter un système en physique classique, l'espace généralement le plus adapté est l'espace à trois dimensions $\mathbb{R}^3$. En physique quantique l'espace adéquat est l'*espace des configurations* $\Omega := \mathbb{R}^{3N}$. Comme dans cette thèse les variables de spin ne sont pas prises en compte, l'ensemble Ω décrit toutes les coordonnées possibles des particules de ce système dans un espace à trois dimensions, c'est-à-dire l'ensemble de toutes les configurations possibles de ce dernier. Pour décrire ce système de N particules, on définit une fonction complexe ψ, appelée *fonction d'onde*. Cette fonction appartient à l'espace de Hilbert $L^2(\Omega;\mathbb{C})$ et est de norme égale à un et ceci pour tout $t \in [0;T]$. Le module au carré de la fonction d'onde $|\psi(x,t)|^2$ évalué en $x \in \Omega$ et $t \in [0;T]$, correspond physiquement à la densité de probabilité que le système a de se trouver dans la configuration x à l'instant t. L'évolution de cette fonction est régie par l'équation de Schrödinger

$$\begin{cases} \imath \frac{\partial \psi}{\partial t}(x,t) = H\psi(x,t), \\ \psi(t=0) = \psi_{\text{init}}(x), \end{cases} \tag{1.27}$$

où $\psi(t=0) = \psi_{\text{init}}(x)$ est la condition initiale du système et l'opérateur H désigne l'opérateur appelé *Hamiltonien* du système. L'Hamiltonien est défini par

$$H = -\frac{\hbar^2}{2m}\Delta + V(\cdot,\cdot), \tag{1.28}$$

où $-\frac{\hbar^2}{2m}\Delta$ correspond à l'énergie cinétique du système, $\hbar = 6,63 \times 10^{-34}\text{J}\cdot\text{s}^{-1}$ est la constante de Planck, m est la masse de la particule étudiée et V est le potentiel électrostatique auquel est soumis

le système. Ce dernier est généralement pris indépendamment du temps. Dans ce cas particulier H l'est aussi et l'équation (1.27) admet alors une unique solution que l'on peut expliciter formellement de la manière suivante

$$\forall x \in \Omega, \forall t \in [0;T], \quad \psi(x,t) = e^{-\imath H}\psi_{\text{init}}(x). \tag{1.29}$$

L'opérateur, qui à une condition initiale ψ_{init} et un temps $t \in [0;T]$ associe la fonction d'onde $\psi(x,t)$, est appelé *propagateur*. On appelle *propagation* la résolution du problème (1.27).

Le système de particules est soumis à un laser dans les applications que nous traitons dans cette thèse. Il faut, par conséquent, le prendre en compte dans le modèle. Pour ce faire, l'équation (1.27) doit donc être modifiée. Le modèle que nous adoptons pour rendre compte de l'interaction est le suivant

$$\begin{cases} \imath\frac{\partial\psi}{\partial t}(x,t) = (H - \varepsilon(t)\mu(x))\,\psi(x,t), \\ \psi(t=0) = \psi_{\text{init}}(x), \end{cases} \tag{1.30}$$

où ε est une fonction scalaire de $L^2(0;T)$ et $\mu(x)$ est un opérateur sur $L^2(\Omega;\mathbb{C})$. Ces quantités représentent respectivement le rayonnement du laser et le moment dipolaire, c'est-à-dire l'opérateur rendant compte de l'effet du champ sur le système considéré.

Discutons brièvement les hypothèses qui ont été faites pour obtenir cette équation. Dans un modèle plus général, le terme d'interaction $\varepsilon(t)\mu(x)$ s'écrit sous la forme $\varphi(E(x,t))$, où $E(x,t)$ est un vecteur de $\mathbb{R}^3$ représentant un champ électrique appliqué au système. Pour aboutir à (1.30), nous avons donc fait trois simplifications. Tout d'abord, nous avons considéré le champ comme constant en espace, de sorte qu'il ne dépende plus de x. Cette hypothèse est raisonnable à l'échelle quantique, car les lasers modernes produisent des champs constants sur des sections assez larges. Ensuite, nous avons linéarisé φ en 0, en supposant le champ suffisamment petit. Cette simplification n'est pas toujours justifiée physiquement. De plus, du point de vue mathématique et pour ce qui concerne les algorithmes que nous construisons, elle n'est pas réellement nécessaire dans la mesure où nos méthodes s'adapteraient sans problème à une fonction non-linéaire. Nous la gardons cependant par souci de simplification de l'exposé, et parce qu'elle est acceptable pour les applications que nous considérons. Enfin, nous supposons que le rayonnement est polarisé si bien que le champ électrique produit est porté par un vecteur fixe. Cette hypothèse correspond très bien aux champs produits par des lasers. Sous cette hypothèse, le champ est donc représenté par une grandeur scalaire, que nous avons donc notée ici ε. Le système proposé est un modèle simplifié adapté aux applications pour lesquelles il est destiné. Précisons que l'on définit de la même manière les termes propagateur et propagation pour le problème (1.30) que pour le problème (1.27).

On peut alors énoncer le problème de contrôlabilité évoqué plus haut, à savoir le contrôle par un laser d'un système constitué de N particules.

Problème 1.4.1. *Étant donné l'Hamiltonien H, le moment dipolaire μ, un état initial ψ_{init} et un état cible ψ_{cible}, existe-t-il un champ ε et un temps T tel que l'état ψ_{init} soit conduit sur ψ_{cible} par le propagateur associé à (1.30) ? Si oui, comment calculer ces quantités ?*

Les réponses à cette question de contrôlabilité sont restées peu nombreuses jusqu'au début des années 90. C'est en 1995 que les premiers critères de contrôlabilité ont été donnés par V. Ramakrishna dans [41]. Ses résultats ont permis d'affirmer qu'à de rares exceptions près, tous les systèmes quantiques sont contrôlables. D'un point de vue pratique, c'est en 2001 que R. Lewis et *al.* dans [32] ont répondu compléter ces résultats par des expériences en laboratoire couronnées de succès. Par la suite, le contrôle quantique a connu des avancées significatives aussi bien du point de vue théorique que du point de vue pratique, voir [5, 51, 50, 40] et les références qu'elles contiennent. Des résultats ont été obtenus sur l'existence de contrôles [7, 8, 49, 2, 3] ou sur des moyens efficaces de calculer et de générer des champs lasers qui permettent d'atteindre des objectifs concernant

l'état des systèmes quantiques [53, 37, 39, 44, 35, 36]. D'un autre côté, la conception de champs lasers adaptés joue aussi un rôle majeur lorsque l'objectif est d'identifier quelques propriétés du système quantique à contrôler. C'est pour cela que certaines méthodes ont été développées afin d'identifier des caractéristiques de certains systèmes en dimension finie [28], ou de calculer des champs lasers permettant de les discriminer [6]. Pour plus d'informations et une présentation plus approfondie des concepts présentés ici, on consultera le chapitre intitulé *Control of quantum dynamics : concepts, procedures and future prospects* dans [16] paru en 2003, mais aussi [47], [43] et [48], ainsi que leurs bibliographies. Notons en outre que l'identification d'opérateurs pour l'équation de Schrödinger a déjà été étudiée dans la littérature. Pour un résultat théorique, où l'on ne considère pas d'interactions avec un laser, on pourra se référer à [34].

Indiquons aussi que d'autres domaines d'applications sont concernés par ces résultats. On peut, entre autres, citer la conception de portes logiques pour les futurs *ordinateurs quantiques* ou encore l'*imagerie par résonance magnétique* (IRM).

Suite à l'engouement suscité par le contrôle par laser en chimie quantique, un nouveau problème lié à l'identification des systèmes est donc apparu en chimie quantique. On peut l'énoncer sous la forme suivante.

Problème 1.4.2 (Problème inverse en contrôlabilité)**.** *Étant donné un état initial ψ_{init}, un état cible ψ_{cible} et un champ laser ε est-il possible de déterminer un Hamiltonien H et un moment dipolaire D de l'équation* (1.30) *de sorte que la propagation* (1.30) *associée conduise l'état initial ψ_{init} à l'état cible ψ_{cible} ? Si oui, comment calculer ces quantités ?*

Ceci revient à inverser le problème de contrôlabilité. Ici le laser n'est plus vu comme un moyen de contrôler le système de particules, mais comme un nouveau "réactif" venant s'ajouter aux composés chimiques classiques. Pour un champ laser donné, il est alors intéressant de connaître les systèmes susceptibles de "réagir" avec lui. Dans ce cadre, les particules, représentées par les opérateurs H et D dans l'équation (1.30) deviennent les inconnues du problème.

Dans le chapitre 4, qui apporte une contribution à cette problématique, nous nous concentrerons donc sur l'identification d'Hamiltoniens et de moments dipolaires pour des systèmes de petites dimensions. Du point de vue théorique, nous obtenons un résultat d'existence locale : nous prouvons que l'inversion est toujours possible au voisinage de quelques états particuliers. Comme sous-produit de ce résultat, nous mettons en valeur certaines caractéristiques du champ laser permettant cette identification. Poursuivant l'approche locale, nous présentons, dans un deuxième temps, le problème discrétisé en temps et des méthodes de point fixe pour le résoudre numériquement. En particulier, une méthode de Newton est proposée conjointement avec des méthodes de continuation pour résoudre notre problème dans des cas où le résultat local n'est pas vérifié. Précisons que le chapitre 4 est une version augmentée d'un acte de conférence [27]. Les algorithmes que nous avons conçus ont d'ores-et-déjà été utilisés par les chimistes, voir [38]. Les auteurs y proposent une alternative pour le choix de continuation et appliquent le tout à un cas pratique. Ces deux méthodes de continuation, ainsi qu'une originale, sont présentées dans le chapitre 4.

Références

[1] Geoffray ADDE. "Image Processing Methods Applied to the Magneto-Electro-Encephalography Inverse Problem". Theses. Ecole des Ponts ParisTech, 2005. URL: https://pastel.archives-ouvertes.fr/pastel-00001593.

[2] F. ALBERTINI and D. D'ALESSANDRO. "Notions of controllability for bilinear multilevel quantum systems". *IEEE TRANSACTIONS ON AUTOMATIC CONTROL* 48.8 (2003), pp. 1399–1403. ISSN: 0018-9286. DOI: 10.1109/TAC.2003.815027.

[3] C. ALTAFINI. "Controllability of quantum mechanical systems by root space decomposition of su(N)". *JOURNAL OF MATHEMATICAL PHYSICS* 43.5 (2002), pp. 2051–2062. ISSN: 0022-2488. DOI: 10.1063/1.1467611.

[4] H. AMMARI. *An Introduction to Mathematics of Emerging Biomedical Imaging*. Mathématiques et Applications. Springer, 2008. ISBN: 9783540795537.

[5] A. ASSION et al. "Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses". *SCIENCE* 282.5390 (1998), pp. 919–922. ISSN: 0036-8075. DOI: 10.1126/science.282.5390.919.

[6] L. BAUDOUIN and JP. PUEL. "Uniqueness and stability in an inverse problem for the Schrodinger equation". *INVERSE PROBLEMS* 18.6 (2002), pp. 1537–1554. ISSN: 0266-5611. DOI: 10.1088/0266-5611/18/6/307.

[7] K. BEAUCHARD. "Local controllability of a 1-D Schrodinger equation". *JOURNAL DE MATHEMATIQUES PURES ET APPLIQUEES* 84.7 (2005), pp. 851–956. ISSN: 0021-7824. DOI: 10.1016/j.matpur.2005.02.005.

[8] K. BEAUCHARD and C. LAURENT. "Local controllability of 1D linear and nonlinear Schrodinger equations with bilinear control". *JOURNAL DE MATHEMATIQUES PURES ET APPLIQUEES* 94.5 (2010), pp. 520–554. ISSN: 0021-7824. DOI: 10.1016/j.matpur.2010.04.001.

[9] M. BONNET. *Équations intégrales et éléments de frontière : applications en mécanique des solides et des fluides*. Collection Sciences et techniques de l'ingénieur. CNRS éd., 1995. ISBN: 9782212058208.

[10] T. BOULIER. "Modélisation de l'électro-localisation active chez les poissons faiblement électriques". Thèse de doctorat dirigée par Ammari, Habib Mathématiques appliquées Palaiseau, Ecole polytechnique 2013. PhD thesis. 2013, 1 vol. (146 p.) URL: http://www.theses.fr/2013EPXX0108.

[11] F. BOYER et al. "Underwater Reflex Navigation in Confined Environment Based on Electric Sense". *Robotics, IEEE Transactions on* 29.4 (2013), pp. 945–956. ISSN: 1552-3098. DOI: 10.1109/TRO.2013.2255451.

[12] F. BOYER et al. "Model for a Sensor Inspired by Electric Fish". *Robotics, IEEE Transactions on* 28.2 (2012), pp. 492–505. ISSN: 1552-3098. DOI: 10.1109/TRO.2011.2175764.

[13] C.A. BREBBIA and J. DOMINGUEZ. *Boundary Elements : An Introductory Course*. Sydney Grammar School Press, 1996. ISBN: 9781853123498.

[14] Theodore H. BULLOCK, Carl D. HOPKINS, and Arthur N. POPPER. *Electroreception*. Ed. by Richard R. FAY. Springer, 2005. ISBN: 9780387282756.

[15] A.A. CAPUTI. "The electric organ discharge of pulse gymnotiforms : the transformation of a simple impulse into a complex spatio-temporal electromotor pattern". *Journal of Experimental Biology* 202.10 (1999), pp. 1229–1241. URL: http://jeb.biologists.org/content/202/10/1229.abstract.

[16] P.G. CIARLET, C.L. BRIS, and J.L. LIONS. *Handbook of Numerical Analysis*. Computational Chemistry : Reviews of Current Trends vol. 10. North-Holland, 1990. ISBN: 9780444512482.

[17] C. DARWIN. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. John Murray, 1859.

[18] G. von der EMDE et al. "Electric fish measure distance in the dark". *Nature* 395 (1998). DOI: 10.1038/27655.

[19] N.M. GUNTHER. *La Théorie du potentiel et ses applications aux problèmes fondamentaux de la physique mathématique, par N. M. Gunther*. impr.-édit. Gauthier-Villars, 55, quai des Grands-Augustins, 1934.

[20] J. HAPPEL and H. BRENNER. *Low Reynolds number hydrodynamics with special applications to particulate media*. Vol. 1. Mechanics of fluids and transport processes. Martinus Nijhoff publishers, 1983. DOI: 10.1007/978-94-009-8352-6.

[21] J. L. HESS and A. M. O. SMITH. "Calculation of potential flow about arbitrary bodies". *Progress in Aerospace Sciences* 8 (1967), pp. 1–138. DOI: 10.1016/0376-0421(67)90003-6.

[22] G. HSIAO and W.L. WENDLAND. *Boundary Integral Equations*. Applied Mathematical Sciences. Springer Berlin Heidelberg, 2008. ISBN: 9783540685456. DOI: 10.1007/978-3-540-68545-6.

[23] B. JAWAD et al. "Sensor model for the navigation of underwater vehicles by the electric sense". *Robotics and Biomimetics (ROBIO), 2010 IEEE International Conference on*. 2010, pp. 879–884. DOI: 10.1109/ROBIO.2010.5723442.

[24] Brahim JAWAD. "Modélisation de l'électrolocation pour la bio-robotique". Thèse de doctorat dirigée par Boyer, F. Bio-mécanique et bio-ingénierie Nantes, Ecole des Mines 2012. PhD thesis. 2012. URL: http://www.theses.fr/2012EMNA0010.

[25] Oliver Dimon KELLOGG. *Foundations of potential theory*. Berlin [u.a.]: Springer, 1929. URL: http://eudml.org/doc/203661.

[26] S. KIM and S. J. KARRILA. *Microhydrodynamics : principles and selected applications*. Butterworth-Heinemann, 1991.

[27] P. LAURENT et al. "Control through operators for quantum chemistry". *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*. 2012, pp. 1663–1667. DOI: 10.1109/CDC.2012.6427030.

[28] C. LE BRIS et al. "Hamiltonian identification for quantum systems : Well-posedness and numerical approaches". *ESAIM-CONTROL OPTIMISATION AND CALCULUS OF VARIATIONS* 13.2 (2007), pp. 378–395. ISSN: 1262-3377. DOI: 10.1051/cocv:2007013.

[29] V. LEBASTARD et al. "Underwater robot navigation around a sphere using electrolocation sense and Kalman filter". *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. 2010, pp. 4225–4230. DOI: 10.1109/IROS.2010.5648929.

[30] H. LEMONNIER. *Résolution de l'équation de Laplace par la méthode des éléments de frontière*. 2004. URL: http://herve.lemonnier.sci.free.fr/BEM/Laplace.pdf.

[31] H. LEMONNIER and J.F. PEYTRAUD. "Is 2D impedance tomography a reliable technique for two-phase flow ?" *Nuclear Engineering and Design* 184.2-3 (1998), pp. 253–268. ISSN: 0029-5493. DOI: 10.1016/S0029-5493(98)00201-5. URL: http://www.sciencedirect.com/science/article/pii/S0029549398002015.

[32] Robert J. LEVIS, Getahun M. MENKIR, and Herschel RABITZ. "Selective Bond Dissociation and Rearrangement with Optimally Tailored, Strong-Field Laser Pulses". *Science* 292.5517 (2001), pp. 709–713. DOI: 10.1126/science.1059133. eprint: http://www.sciencemag.org/content/292/5517/709.full.pdf. URL: http://www.sciencemag.org/content/292/5517/709.abstract.

[33] H.W. LISSMANN and K.E. MACHIN. "The mechanism of object location in Gymnarchus niloticus and similar fish". *The Journal of Experimental Biology* 35 (1958), pp. 451–486.

[34] Y. MADAY and J. SALOMON. "A greedy algorithm for the identification of quantum systems". *Proceedings of the 48th IEEE Conference on Decision and Control* 5 (2009), pp. 375–379. ISSN: 0191-2216. DOI: 10.1109/CDC.2009.5400691.

[35] Y. MADAY, J. SALOMON, and G. TURINICI. "Monotonic parareal control for quantum systems". *SIAM JOURNAL ON NUMERICAL ANALYSIS* 45.6 (2007), pp. 2468–2482. ISSN: 0036-1429. DOI: 10.1137/050647086.

[36] Y. MADAY, J. SALOMON, and G. TURINICI. "Monotonic time-discretized schemes in quantum control". *NUMERISCHE MATHEMATIK* 103.2 (2006), pp. 323–338. ISSN: 0029-599X. DOI: 10.1007/s00211-006-0678-x.

[37] Y. MADAY and G. TURINICI. "New formulations of monotonically convergent quantum control algorithms". *JOURNAL OF CHEMICAL PHYSICS* 118.18 (2003), pp. 8191–8196. ISSN: 0021-9606. DOI: 10.1063/1.1564043.

[38] M. NDONG, J. SALOMON, and D. SUGNY. "Newton algorithm for Hamiltonian characterization in quantum control". *Journal of Physics A : Mathematical and Theoretical* 47.26 (2014), p. 265302. DOI: 10.1088/1751-8113/47/26/265302. URL: http://stacks.iop.org/1751-8121/47/i=26/a=265302.

[39] JP. PALAO and R. KOSLOFF. "Optimal control theory for unitary transformations". *PHYSICAL REVIEW A* 68.6 (2003). ISSN: 1050-2947. DOI: 10.1103/PhysRevA.68.062308.

[40] H. RABITZ et al. "Wither the future of controlling quantum phenomena ?" *SCIENCE* 288 (2000), pp. 824–828.

[41] Viswanath RAMAKRISHNA et al. "Controllability of molecular systems". *Phys. Rev. A* 51 (2 1995), pp. 960–966. DOI: 10.1103/PhysRevA.51.960. URL: http://link.aps.org/doi/10.1103/PhysRevA.51.960.

[42] D. ROTHER et al. "Electric images of two low resistance objects in weakly electric fish". *Biosystems* 71.1-2 (2003), pp. 171–179. URL: http://ampere.iie.edu.uy/publicaciones/2003/RMCGCB03.

[43] J. SALOMON. "Contrôle en chimie quantique : conception et analyse de schémas d'optimisation". Thèse de doctorat dirigée par Maday, Yvon Mathématiques appliquées Paris 6 2005. PhD thesis. 2005, 1 vol. (VIII–155 f.) URL: http://www.theses.fr/2005PA066354.

[44] J. SALOMON. "Convergence of the time-discretized monotonic schemes". *ESAIM - Mathematical Modeling and Numerical Analysis* 41.1 (2007), pp. 77–93. ISSN: 0764-583X. DOI: 10.1051/m2an:2007008.

[45] M. SMOLUCHOWSKI. "Über die Wechselwirkung von Kugeln, die sich in einer zähen Flüssigkeit bewegen". *Bull. Int. Acad. Sci. Cracovie, Cl. Sci. Math. Nat., Sér. A Sci. Math.* (1911), pp. 28–39. DOI: 10.1038/27655.

[46] J. SOLBERG, K. LYNCH, and M. MACIVER. "Active Electrolocation for Underwater Target Localization". *The International Journal of Robotics Research* 27 (2008), pp. 529–548. DOI: 10.1177/0278364908090538.

[47] G. TURINICI. "Analyse de méthodes numeriques de simulation et contrôle en chimie quantique". Thèse de doctorat dirigée par Maday, Yvon Mathématiques appliquées Paris 6. PhD thesis. 2000, 1 vol. (183 p.) URL: http://www.theses.fr/2000PA066463.

[48] G. TURINICI. "Lumière singulière". *La Recherche* hors-série n° 14 (2015), pp. 52–57.

[49] G. TURINICI and H. RABITZ. "Wavefunction controllability for finite-dimensional bilinear quantum systems". *JOURNAL OF PHYSICS A-MATHEMATICAL AND GENERAL* 36.10 (2003), pp. 2565–2576. ISSN: 0305-4470. DOI: 10.1088/0305-4470/36/10/316.

[50] WS. WARREN, H. RABITZ, and M. DAHLEH. "Coherent control of quantum dynamics - the dream is alive". *SCIENCE* 259.5101 (1993), pp. 1581–1589. ISSN: 0036-8075. DOI: 10.1126/science.259.5101.1581.

[51] TC. WEINACHT, J. AHN, and PH. BUCKSBAUM. "Controlling the shape of a quantum wavefunction". *NATURE* 397.6716 (1999), pp. 233–235. ISSN: 0028-0836. DOI: 10.1038/16654.

[52] R. WILLIAMS, B. RASNOW, and C. ASSAD. "Hypercube Simulation of Electric Fish Potentials". *Distributed Memory Computing Conference, 1990., Proceedings of the Fifth*. 1990, pp. 470–477. DOI: 10.1109/DMCC.1990.555422.

[53] G. von WINCKEL, A. BORZI, and S. VOLKWEIN. "A globalized Newton method for the accurate solution of a dipole quantum control problem". *SIAM JOURNAL ON SCIENTIFIC COMPUTING* 31.6 (2009), pp. 4176–4203. ISSN: 1064-8275. DOI: 10.1137/09074961X.

2

Méthode des réflexions

Sommaire

2.1 Introduction

En 1911, M. Smoluchowski [40] introduisit une méthode itérative pour calculer les forces hydrodynamiques s'exerçant sur un nombre arbitraire de sphères tombant dans un fluide visqueux non borné. Nommée plus tard *méthode des réflexions*, cette technique a été mise à l'honneur dans des livres de référence sur l'hydrodynamique à bas Reynolds, tels ceux de J. Happel et H. Brenner [26], de Kim et Karrila [31] ou de J. K. G. Dhont [19]. On la retrouve aussi dans plusieurs articles traitant du mouvement de particules plongées dans un fluide visqueux, dont voici une liste non exhaustive [32, 29, 12, 28, 43]. Généralement, les auteurs traitent un problème linéaire muni de conditions aux bords extérieurs associées à plusieurs "objets", tels que des particules dans le cas d'une suspension. La méthode consiste à résoudre et sommer les solutions des problèmes aux bords associés à un seul objet, appelées *réflexions*. D'un point de vue physique, on peut visualiser son principe en *"imaginant une perturbation initiale se réfléchissant sur les bords impliqués et produisant un effet moindre à chaque réflexion"* [26, chapitre 8], d'où son nom.

De façon plus formelle, la méthode suit une stratégie *diviser pour régner*, qui suppose que chaque problème à un objet est, d'une certaine façon, plus simple à résoudre qu'un problème à plusieurs objets. En ce sens, elle admet, au moins dans l'idée, plusieurs similitudes avec les méthodes de décomposition de domaine du type Schwarz (voir [24] par exemple), qui essayent de résoudre un problème (généralement interne) de valeurs aux bords en le divisant en problèmes de valeurs aux bords sur différents sous-domaines se chevauchant partiellement ou non et en itérant leurs résolutions de façon cyclique afin d'harmoniser la solution entre les sous-domaines adjacents. Il y a cependant une différence conceptuelle importante entre les deux approches. En effet, chacun des problèmes à un objet à résoudre est défini sur un domaine, possiblement non borné, qui est le complément de l'objet considéré (*i.e.* l'intérieur de son complémentaire) et, en tant que tel, contient le domaine sur lequel le problème principal est défini. La méthode des réflexions doit donc plutôt être vue comme une méthode de décomposition de frontières.

La méthode des réflexions repose historiquement sur des calculs faits à la main, utilisant des formules explicites et malléables pour les réflexions, et fut donc appliquée pour un petit nombre d'objets identiques de formes simples, pour lesquels on dispose de formes analytiques des lois de Stokes et Faxén, et pour lesquels le champ de vecteurs vitesse peut être approché par troncature. Avec le temps, cette restriction, couplée avec le fait que la convergence peut être assez lente (plus encore quand les objets sont proches les uns des autres), demandant donc un nombre important de termes dans la série tronquée pour obtenir une approximation valable, la rendit moins intéressante. Cependant, comme la méthode en elle-même ne repose par sur la façon dont les problèmes aux bords sont résolus, il devint clair, avec l'apparition des ordinateurs, que des méthodes de discrétisation numérique pouvaient être utilisées, et même combinées[1] avec les méthodes analytiques susmentionnées, pour résoudre efficacement des problèmes impliquant plusieurs objets de types, de formes et de tailles différents, ainsi que de positions arbitraires.

Malgré un grand nombre d'articles récents concernant les applications de cette méthode [28, 37, 10, 43], il apparaît que la méthode a rarement été étudiée d'un point de vue mathématique. On doit citer les travaux précurseurs de Luke [36], dans lesquels la convergence d'une variante de la méthode appliquée à la résolution du *problème de mobilité*[2] dans un domaine borné est théoriquement étudiée. En formulant le problème dans un espace de Hilbert approprié et en interprétant

[1] On peut par exemple regarder l'article [10] dans lequel différentes méthodes de résolution sont utilisées selon la nature du problème à un objet considéré.

[2] En hydrodynamique à bas Reynolds, la méthode des réflexions est typiquement utilisée pour résoudre deux types de problèmes d'interaction hydrodynamiques entre des particules : les problèmes résistif et ceux de mobilité, voir par exemple [31]. Ces deux types de problèmes sont des problèmes aux bords basés sur les équations de Stokes pour un fluide newtonien incompressible. Dans les problèmes du premier type, les forces et moments de torsion sont déterminés pour des particules dont la vitesse dans le fluide ambiant est connue, alors que dans le second type la vitesse des particules est inconnue.

la méthode en utilisant des opérateurs de projection orthogonale (indépendamment des travaux contemporains et similaires de Lions sur la méthode de Schwarz [35]), Luke prouva que la méthode converge toujours, de façon linéaire sous une condition géométrique sur les sous-espaces associées aux projections. De plus, dans [41], la version "classique" de la méthode est analysée quand elle est appliquée à la résolution d'un problème de Dirichlet dans un domaine tridimensionnel non borné complémentaire à un ensemble de sphères. Recourant à des techniques analytiques, Traytak établit dans ce cas des conditions de convergence nécessaires et suffisantes exprimées en termes de rayons des sphères et distances entre leurs centres. En dépit de ces deux résultats importants, une théorie mathématique de la méthode des réflexions semble toujours manquante.

De plus, on pourra ajouter qu'un certain nombre d'actes manqués ont amené à des incompréhensions de la part des communautés scientifiques utilisant cette méthode. Comme écrit plus haut, il existe plusieurs formes de la méthode des réflexions dans la littérature. Ainsi, quand Ichiki et Brady affirment dans [28] que *"tandis que la convergence de la méthode des réflexions a été prouvée pour le problème de mobilité, la convergence pour le problème de résistif est une question ouverte"* et donnent un *contre-exemple* à la convergence de la méthode dans le cas de configurations impliquant plus que deux objets, ils dévalorisent d'une certaine manière leur découverte en ne remarquant pas que la méthode des réflexions utilisée dans leur travail diffère de celle introduite et étudiée dans [36]. De plus, alors que la propriété de convergence de la méthode dépend du type de problème aux bords que l'on essaye de résoudre, les conclusions basées sur la seule nature du problème sont parfois incorrectement posées. Par exemple, Luke [36] admet que la preuve de convergence pour le problème de mobilité n'est pas valable pour le problème résistif, et dans [43] les auteurs écrivent à partir de considérations numériques : *"Clairement les deux problèmes sont philosophiquement équivalents : mais numériquement nous trouvons de façon empirique que leurs convergences se comportent de façon très différentes. Pour des sphères bien séparées, il n'y a de problème de convergence dans aucun des problèmes, mais si les sphères sont suffisamment proches, le problème résistif semble converger moins bien que le problème de mobilité."* Cependant, la convergence pour les deux problèmes peut être prouvée avec les mêmes outils mathématiques. De même, le fait que les problèmes aux bords sont résolus dans un domaine borné ou non a été avancé comme modifiant la théorie de convergence de la méthode. Dans [36], on peut lire : *"Une des restrictions les plus frappantes dans l'hypothèse de la preuve de convergence est que le contenant [...] est borné. L'intuition physique suggère qu'il n'y a pas de raison pour que la présence de murs limitant le contenant pourrait aider la méthode des réflexions. Mais une preuve rigoureuse reste manquante."* Une fois de plus, la convergence de la méthode peut être démontrée pour un problème aux bords dans un domaine borné aussi bien que dans un domaine non borné.

Notre but est donc double. D'abord, nous voulons proposer un cadre théorique unifiant l'analyse des différentes formes existantes de la méthode, pour pouvoir expliquer les cas de divergence observés dans certaines applications, et fournir, dans certains cas, des résultats de convergences. De cette manière, nous espérons clarifier les incompréhensions qui pourraient rester quant à cette méthode, et, peut être, la proposer à une audience plus large.

Le chapitre s'organise de la façon suivante. Utilisant une analogie bien connue entre l'hydrodynamique et l'électrostatique, nous commençons par illustrer les présentations diverses de la méthode des réflexions trouvées dans la littérature sur une série d'exemples de problèmes impliquant un opérateur de Laplace. Cet exposé nous permet de définir de façon plus générale les deux variantes de la méthode : la variante historique, et celle introduite dans [36]. Ensuite, en étendant un résultat de Luke, on exploite l'utilisation de la méthode dans un espace de Hilbert pour analyser les deux formes de la méthode. Des résultats de convergence conditionnelle des méthodes itératives pour la résolution de systèmes linéaires sont donnés pour la forme parallèle, et une variante assurant la convergence inconditionnelle dans certaines applications (celles impliquant des projecteurs orthogonaux) est proposée. Pour la forme séquentielle, la convergence

inconditionnelle est obtenue (quand les projecteurs sont orthogonaux). Finalement, des exemples d'applications sont traités et quelques résultats numériques sont présentés.

2.2 Présentation de la méthode

Nous commençons par une description de la méthode au travers de son application à un problème d'écoulement de Stokes résistif. D'un point de vue mathématique, ce problème revient à résoudre l'équation de Stokes intérieur ou extérieur pour des conditions aux bords de Dirichlet. Toutefois, pour simplifier notre présentation, nous exploitons une analogie qui existe entre l'hydrodynamique à bas Reynolds (aussi dit écoulement de Stokes) et l'électrostatique, rappelée par Luke dans [36]. Ceci revient à remplacer l'équation de Stokes par celle de Laplace, la vitesse du fluide est alors remplacée par le potentiel électrostatique.

2.2.1 Un analogue électrostatique au problème résistif : la forme parallèle

Nous considérons donc l'analogue électrostatique au problème résistif en hydrodynamique. Celui-ci est composé d'un système d'équations décrivant le potentiel électrostatique u dans un conteneur Ω relié à la terre, contenant N objets conducteurs O_j, avec $j \in [\![1;N]\!]$. Sur le bord de ces N objets, c'est-à-dire sur ∂O_j pour $j \in [\![1;N]\!]$, on définit respectivement une fonction scalaire U_j. Le problème est de trouver le potentiel u solution de

$$\begin{cases} -\Delta u = 0 & \text{dans } \Omega \setminus \left(\cup_{j=1}^{N} O_j\right), \\ u = U_j & \text{sur } \partial O_j,\ j \in [\![1;N]\!]. \end{cases} \tag{2.1}$$

Le domaine Ω peut être borné ou non. Dans le second cas, il est indispensable d'imposer des conditions d'évanescences du potentiel en l'infini (voir la section 2.4.3). Dans ce qui suit, on suppose que le conteneur est borné et qu'il est muni d'une condition aux bords de Dirichlet homogène,

$$u = 0 \quad \text{sur } \partial\Omega.$$

Pour approcher la solution, la méthode des réflexions consiste à sommer des solutions intermédiaires $u_j^{(k)}$, pour $j \in [\![1;N]\!]$ et $k \in \mathbb{N}^*$, solutions respectives des problèmes à un seul objet de la forme suivante

$$\forall i \in [\![1;N]\!], \quad \begin{cases} -\Delta u_i^{(k+1)} = 0 & \text{dans } \Omega \setminus O_i, \\ u_i^{(k+1)} = -\sum_{j=1}^{i-1} u_j^{(k)} - \sum_{j=i+1}^{N} u_j^{(k)} & \text{sur } \partial O_i, \\ u_i^{(k+1)} = 0 & \text{sur } \partial\Omega, \end{cases} \tag{2.2}$$

chacun de ces sous-problèmes n'utilisent que la donnée initiale U_i, soit

$$\forall i \in [\![1;N]\!], \quad \begin{cases} -\Delta u_i^{(1)} = 0 & \text{dans } \Omega \setminus O_i, \\ u_i^{(1)} = U_i & \text{sur } \partial O_i, \\ u_i^{(1)} = 0 & \text{sur } \partial\Omega. \end{cases}$$

Ainsi l'approximation de la solution à la fin du $k^{\text{è}}$ *cycle* est définie par

$$u^{(k)} = \sum_{\ell=1}^{k} \sum_{i=1}^{N} u_i^{(\ell)} \quad \text{dans } \Omega \setminus \left(\cup_{j=1}^{N} O_j\right). \tag{2.3}$$

Bien que la présentation, l'équation sous-jacente et les conditions aux limites peuvent différer, la forme présentée de la méthode est la version historique, celle qui apparaît dans l'article original de Smoluchowski [40]. Le livre de Happel et Brenner [26] ou celui de Kim et Karrila [31], ou, plus récemment, l'article de Traytak [41] y font référence. Nous observons que les données pour chacun des problèmes à un objet à un cycle donné dépendent uniquement des quantités calculées au cours du cycle précédent ; sauf pour le premier, où elles correspondent aux données du problème de départ. La résolution de ces sous-problèmes peut ainsi être effectuée simultanément, d'où la mise en rapport avec la méthode de Jacobi pour la résolution itérative de systèmes d'équations linéaires. Ceci nous conduit à appeler cette procédure, la *forme parallèle* de la méthode des réflexions.

Concernant la convergence de cette version de la méthode, Happel et Brenner indiquent, dans leur livre page 236, que : *"il doit être souligné qu'aucune preuve rigoureuse n'existe sur la converge des itérations vers la solution souhaitée"*. Le fait est, qu'un exemple de divergence pour une configuration particulière est présenté par Ichiki et Brady dans [28]. Ils considèrent le problème résistif en présence de particules sphériques rigides dans un domaine non borné. Dans ce cas, les solutions du problème pour des particules simples sont estimées par des développements multipolaires tronqués. De plus, les auteurs conjecturent que la divergence observée est sans rapport avec l'ordre de troncature des développements. Dans un autre article, Traytak [41] a analysé la méthode appliquée à la résolution d'un problème de Laplace avec des conditions de Dirichlet en domaine non borné contenant un ensemble de sphères. Il obtient des conditions nécessaires et suffisantes de convergence, ce qui lui permet d'exposer des cas simples de divergence lorsque le nombre de sphères est supérieur ou égal à huit.

Fait intéressant, on peut remarquer que la technique de *décomposition de frontières*, introduite par Balabane dans [5] (voir aussi [42]) pour résoudre un problème aux bords pour l'équation de Helmholtz en domaine non borné contenant un ensemble de diffuseurs disjoints, est presque identique à cette forme de la méthode des réflexions. Dans le même article, une condition suffisante pour la convergence, en fonction de la fréquence, des diamètres et des aires des diffuseurs, ainsi que les distances entre ces derniers, est établie. Le lien entre le principe de la méthode des réflexions et l'approche de la diffusion multiple des ondes connues sous le nom de modèle de Foldy-Lax [22, 33, 34] peut aussi être signalé [3].

2.2.2 Un analogue électrostatique au problème de mobilité : une forme séquentielle

Une autre version de la méthode des réflexions a été proposée par J. Happel et H. Brenner dans [26, chapitre 6]. Malheureusement cette dernière n'est pas consistante pour plus de deux objets. La version corrigée de cette dernière sera appelée *version séquentielle* par la suite. La première méthode séquentielle consistante semble avoir été proposée par Luke dans [36] pour résoudre le problème de mobilité en hydrodynamique à N objets. Ce dernier reconnaît tout de même avoir pris certaines libertés dans la construction de la méthode. Néanmoins ces différentes méthodes séquentielles peuvent être reliées dans le cadre présenté dans la section suivante.

Présentons à présent l'analogue électrostatique du problème de mobilité : étant donnés N scalaires Q_j, pour $j \in [\![1; N]\!]$, (correspondant aux charges des différents conducteurs O_j), nous cherchons le champ u (le potentiel électrique) et les N scalaires c_j, avec $j \in [\![1; N]\!]$, solutions du

[3] Pour voir cela, on peut par exemple comparer la hiérarchie des différents niveaux d'approximation, allant de l'approximation de Born jusqu'au modèle de Foldy-Lax, dans [11], en considérant une suite d'approximations produites par la méthode des réflexions

système suivant

$$\begin{cases} -\Delta u = 0 & \text{dans } \Omega \setminus \left(\cup_{j=1}^{N} O_j\right), \\ u = 0 & \text{sur } \partial\Omega, \\ u = c_j & \text{sur } \partial O_j,\ j \in [\![1;N]\!], \\ \displaystyle\int_{\partial O_j} \frac{\partial u}{\partial \boldsymbol{n}}\,\mathrm{d}S = Q_j, & j \in\in [\![1;N]\!]. \end{cases} \tag{2.4}$$

Afin d'expliciter les champs auxiliaires, on réinterprète[4] cette forme de la méthode en explicitant les suites de champs $\left(u_i^{(k)}\right)_{k\in\mathbb{N}^*}$ et de scalaires $\left(c_i^{(k)}\right)_{k\in\mathbb{N}^*}$, pour $i \in [\![1;N]\!]$, telles que

$$\forall k \in \mathbb{N}^*,\ \forall i \in [\![1;N]\!],\quad \begin{cases} -\Delta u_i^{(k+1)} = 0 \quad \text{dans } \Omega \setminus O_i, \\ u_i^{(k+1)} = 0 \quad \text{sur } \partial\Omega, \\ u_i^{(k+1)} = c_i^{(k)} - \sum_{j=1}^{i-1} u_j^{(k+1)} - \sum_{j=i+1}^{N} u_j^{(k)} \quad \text{sur } \partial O_i, \\ \displaystyle\int_{\partial O_i} \frac{\partial u_i^{(k+1)}}{\partial \boldsymbol{n}}\,\mathrm{d}S = -\sum_{j=1}^{i-1} \int_{\partial O_i} \frac{\partial u_j^{(k+1)}}{\partial \boldsymbol{n}}\,\mathrm{d}S - \sum_{j=i+1}^{N} \int_{\partial O_i} \frac{\partial u_j^{(k)}}{\partial \boldsymbol{n}}\,\mathrm{d}S = 0, \end{cases} \tag{2.5}$$

avec

$$\forall i \in [\![1;N]\!],\quad \begin{cases} -\Delta u_i^{(1)} = 0 \quad \text{dans } \Omega \setminus O_i, \\ u_i^{(1)} = 0 \quad \text{sur } \partial\Omega, \\ u_i^{(1)} = c_i^{(1)} - \sum_{j=1}^{i-1} u_j^{(1)} \quad \text{sur } \partial O_i, \\ \displaystyle\int_{\partial O_i} \frac{\partial u_i^{(1)}}{\partial \boldsymbol{n}}\,\mathrm{d}S = Q_i - \sum_{j=1}^{i-1} \int_{\partial O_i} \frac{\partial u_j^{(1)}}{\partial \boldsymbol{n}}\,\mathrm{d}S = Q_i, \end{cases} \tag{2.6}$$

l'approximation de la solution du problème (2.4) au $k^{\text{è}}$ cycle est donnée par (2.3). Il est à noter que $\forall k \in \mathbb{N}^*$, nous avons utilisé que $\int_{\partial O_i} \frac{\partial u_j^{(k)}}{\partial \boldsymbol{n}}\,\mathrm{d}S = 0$ si $j \neq i$, puisque par construction $u_j^{(k)}$ est harmonique dans un voisinage de O_i. Ici, les données pour chacun des problèmes à un objets (2.5) à un cycle donné, dépendent des quantités calculées durant le cycle précédent et l'actuel (ainsi que les données du problème à résoudre). Ceci implique un calcul séquentiel de chacune des solutions, ce qui est dans l'esprit de la méthode de Gauss–Seidel, qui est aussi une façon itérative de résoudre un système linéaire. Nous appelons donc naturellement cette variante, la *forme séquentielle* de la méthode des réflexions. Un résultat de convergence inconditionnelle pour cette variante de la méthode est prouvé dans [36] (voir paragraphe 2.4.2 pour un résumé). Cependant, nous n'avons pas trouvé d'autres utilisations de sa méthode dans la littérature. De plus, il indique lui-même qu'il n'a pas réussi à démontrer sa convergence pour le problème résistif.

2.3 Cadre mathématique

Nous résumons à présent les applications des différentes formes de la méthode des réflexions par un problème aux bords abstrait. Soulignons que ce cadre mathématique unique permet une utilisation plus large de la méthode. La preuve de la consistance de la méthode sous ces différentes formes est alors aisée. Cependant, les preuves de convergence dépendront fortement des opérateurs sous-jacents impliqués. En conséquence, nous restons délibérément vague sur le cadre

[4] Précisons que l'algorithme introduit dans [36] est présenté sous la forme d'une méthode de correction de sous-espaces (voir paragraphe 2.4.1). Par souci de cohérence, nous le modifions afin de le faire correspondre avec l'algorithme que nous présentons par la suite. Ce dernier correspond à la présentation de la forme parallèle de la méthode dans la précédente sous-section 2.2.1 et est dans l'esprit de [26, chapitre 6].

fonctionnel à ce stade. Il sera précisé dans un cas particulier dans la section suivante et dans quelques exemples d'applications connexes.

Étant donné un entier naturel non nul N, nous considérons un domaine connexe et régulier $\Omega \subseteq \mathbb{R}^d$, et une famille de sous-domaines connexes et réguliers $O_j \subset \Omega$, pour $j \in [\![1;N]\!]$. Nous sommes intéressés par la résolution du problème aux bords suivant : *trouver une fonction u, dans un espace fonctionnel approprié, telle que*

$$\begin{cases} \mathcal{L}u = 0 & \text{dans } \Omega \setminus \left(\cup_{j=1}^{N} \overline{O_j}\right) \\ \mathcal{B}_j u = b_j, & j \in [\![1;N]\!], \end{cases} \tag{2.7}$$

où $\mathcal{L}$ et $\left(\mathcal{B}_j\right)_{j\in[\![1;N]\!]}$ sont des opérateurs linéaires et où chacune des fonctions b_j, pour $j \in [\![1;N]\!]$, est une donnée appartenant respectivement à l'ensemble image d'un opérateur $\mathcal{B}_j$. Dans les applications mentionnées précédemment, la fonction u correspond physiquement à un champ de vitesse ou à un potentiel électrique. L'opérateur $\mathcal{L}$ décrit son comportement dans le domaine. C'est généralement un opérateur elliptique (opérateur de Stokes ou de Laplace), mais pas nécessairement comme on peut le voir dans [5] où la méthode est utilisée pour un opérateur de Helmholtz dans un domaine non borné. Soit $j \in [\![1;N]\!]$, l'opérateur $\mathcal{B}_j$ est un opérateur différentiel agissant sur le bord ∂O_j du $j^{\text{è}}$ objet. Celui-ci vérifie certaines conditions d'admissibilités à l'égard de $\mathcal{L}$, afin d'obtenir un problème bien posé. Cet opérateur correspond généralement à un opérateur de trace, comme par exemple $\gamma_{|_{\partial O_j}}(\cdot) := (\cdot)|_{\partial O_j}$. Par conséquent la fonction b_j est une donnée associée à un type de conditions aux bords définie par l'opérateur $\mathcal{B}_j$, ce qui correspond dans l'exemple précédent à une condition de Dirichlet.
Précisons que si le domaine Ω est borné, on peut ajouter au système une condition aux bords homogène sur $\partial\Omega$, sinon il est nécessaire d'ajouter une condition d'évanescence en l'infini, comme dans le cas d'un problème extérieur. Dans tous les cas, le système complet doit conduire à un problème bien posé dans un espace fonctionnel adéquat.

Pour résoudre ce problème en utilisant la version parallèle de la méthode des réflexions, on considère, pour commencer, un cycle d'initialisation qui peut s'écrire

$$u^{(1)} = \sum_{i=1}^{N} u_i^{(1)}{}_{|_{\Omega\setminus\left(\cup_{j=1}^{N}\overline{O_j}\right)}}, \tag{2.8}$$

où les champs auxiliaires initiaux $u_i^{(1)}$, pour $i \in [\![1;N]\!]$, satisfont

$$\forall i \in [\![1;N]\!], \quad \begin{cases} \mathcal{L}u_i^{(1)} = 0 & \text{sur } \Omega \setminus \overline{O_i}, \\ \mathcal{B}_i u_i^{(1)} = b_i. \end{cases} \tag{2.9}$$

L'initialisation est suivie par une phase itérative, dans laquelle on résout à la $k^{\text{è}}$ étape, avec $k \in \mathbb{N}^*$, le sous-problème suivant

$$\forall i \in [\![1;N]\!], \quad \begin{cases} \mathcal{L}u_i^{(k+1)} = 0 & \text{sur } \Omega \setminus \overline{O_i}, \\ \mathcal{B}_i u_i^{(k+1)} = -\mathcal{B}_i \left(\sum_{j=1}^{i-1} u_j^{(k)} + \sum_{j=i+1}^{N} u_j^{(k)}\right), \end{cases} \tag{2.10}$$

où le champ $u_j^{(k)}$ a été calculé au cycle précédent. Notons $u^{(k)}$ l'approximation de la solution au cycle précédent. Ainsi la nouvelle approximation est définie par la formule de mise à jour suivante

$$u^{(k+1)} = u^{(k)} + \sum_{i=1}^{N} u_i^{(k+1)}{}_{|_{\Omega\setminus\left(\cup_{j=1}^{N}\overline{O_j}\right)}}. \tag{2.11}$$

Par conséquent la formule d'approximation de la solution après le $k^{\text{è}}$ cycle est donnée par la double somme suivante

$$\forall k \in \mathbb{N}^*, \quad u^{(k)} = \sum_{\ell=1}^{k} \sum_{i=1}^{N} u_i^{(\ell)}{}_{|_{\Omega \setminus \left(\cup_{j=1}^{N} \overline{O_j}\right)}}.$$

Si relation de récurrence (2.11) suggère explicitement un ordre de sommation (sommation suivant l'indice i à chaque cycle), la somme peut être d'abord effectuée suivant l'indice ℓ, voir notamment [5].

Considérons à présent la forme séquentielle de la méthode des réflexions. Tandis que l'initialisation est toujours donnée par (2.8), les sous-problèmes à résoudre sont maintenant remplacés par

$$\forall i \in [\![1; N]\!], \quad \begin{cases} \mathcal{L} u_i^{(1)} = 0 \quad \text{dans } \Omega \setminus \overline{O_i}, \\ \mathcal{B}_i \left(\sum_{j=1}^{i} u_j^{(1)} \right) = b_i. \end{cases} \tag{2.12}$$

et ceux de la phase itérative s'écrivent

$$\forall i \in [\![1; N]\!], \quad \begin{cases} \mathcal{L} u_i^{(k+1)} = 0 \quad \text{dans } \Omega \setminus \overline{O_i}, \\ \mathcal{B}_i u_i^{(k+1)} = -\mathcal{B}_i \left(\sum_{j=1}^{i-1} u_j^{(k+1)} + \sum_{j=i+1}^{N} u_j^{(k)} \right), \end{cases} \tag{2.13}$$

La mise à jour du champ est identique à (2.11).

Poursuivant les travaux de Luke [36], nous reformulons la méthode en terme d'opérateur de projection. Pour ce faire, il est commode d'étendre les solutions des problèmes aux bords précédents, afin qu'elles soient définie sur l'ensemble Ω tout entier. Pour le problème (2.7), on considère donc l'extension $\tilde{u}$ de u solution du problème suivant

$$\begin{cases} \mathcal{L} \tilde{u} = 0 \quad \text{dans } \Omega \setminus \left(\cup_{j=1}^{N} \partial O_j \right) \\ \mathcal{B}_j \tilde{u} = b_j \end{cases} \tag{2.14}$$

ainsi que des conditions de transmission à travers la frontière des objets, afin de coupler les problèmes intérieurs au problème extérieur. Il y a généralement plusieurs façons de choisir ces conditions, puisque cette construction théorique n'a en pratique aucune conséquence sur la solution de (2.7). Il suffit simplement de les choisir de telle sorte que le problème soit bien posé sur Ω. Notons cependant que dans certains cas, la condition de transmission imposée à $\tilde{u}$ peut ne pas suffire à définir de façon unique la solution à l'intérieur de chacun des objets. Ainsi des contraintes supplémentaires doivent être alors ajoutées. Par exemple, lorsque l'opérateur $\mathcal{L}$ correspond à l'opérateur de Laplace et que l'opérateur $\mathcal{B}_i$ est lié à une condition aux limites de Neumann, voir la section 2.4.2, une condition sur la moyenne de la solution peut alors être considérée. Par souci de simplicité, nous utiliserons dans le reste du chapitre la même notation pour les solutions aux problèmes initiaux et leurs extensions.

La solution de ce problème prolongé est recherchée, éventuellement dans un sens faible, dans un espace de fonctions H. Nous désignons par V le sous-espace de H contenant les solutions obtenues lorsque les données du problème prennent toutes les valeurs possibles. De même, les solutions des sous-problèmes à un objet peuvent être étendues. Pour tout entier $i \in [\![1; N]\!]$, on note V_i le sous-espace de H contenant les solutions à un tel sous-problème quand la donnée associée au bord $\partial\Omega_i$ prend toutes les valeurs possibles. Nous considérons alors l'opérateur P_i de H dans V_i tel que

$$P_i \ : \ v \mapsto w, \tag{2.15}$$

où w est la solution du problème suivant

$$\begin{cases} \mathcal{L} w = 0 \quad \text{dans } \Omega \setminus \partial O_i, \\ \mathcal{B}_i w = \mathcal{B}_i v, \end{cases} \tag{2.16}$$

plus une condition de transmission à travers ∂O_i. Il est très simple de vérifier que cet opérateur est bien défini et linéaire, pour tout entier $i \in [\![1;N]\!]$, dès lors que les sous-problèmes associés aux extensions sont bien posés, et qu'il est idempotent, c'est-à-dire que ${P_i}^2 = P_i \circ P_i = P_i$.

L'introduction de ces opérateurs permet l'obtention de formules d'erreur sur lesquelles se fondent le cadre de notre analyse de la méthode dans la section suivante. En effet, pour la forme parallèle de la méthode , nous avons le résultat suivant.

Théorème 2.3.1. *Supposons que le problème étendu* (2.14) *admet une unique solution u, telle que sa restriction à $\Omega \setminus \left(\cup_{j=1}^N \overline{O_j}\right)$ résout le problème initial* (2.7)*, et que les suites $(u_i^{(k)})_{k\in\mathbb{N}^*,\, i\in[\![1;N]\!]}$ et $(u^{(k)})_{k\in\mathbb{N}^*}$ générées par la version parallèle de la méthode soient définies de manière unique. Alors, on a la formule d'erreur suivante*

$$\forall k \in \mathbb{N}^*, \quad u^{(k)} - u = -\left(Id - \sum_{i=1}^N P_i\right)^k u. \tag{2.17}$$

Démonstration. En raison de la définition de la méthode, on peut facilement démontrer par induction que la suite $(u^{(k)})_{k\in\mathbb{N}^*}$ satisfait la propriété suivante

$$\forall k \in \mathbb{N}^*, \quad \forall i \in [\![1;N]\!],\ \mathcal{B}_i\left(u^{(k)} + u_i^{(k+1)}\right) = b_i,$$

qui, compte tenu de la définition de l'opérateur de projection P_i et du problème aux bords satisfait par u, peut se traduire

$$\forall k \in \mathbb{N}^*, \quad \forall i \in [\![1;N]\!],\ P_i\left(u^{(k)} + u_j^{(k+1)}\right) = P_i u.$$

Par linéarité et en utilisant (2.10), il s'en suit que

$$\forall k \in \mathbb{N}^*,\ \forall i \in [\![1;N]\!], \quad P_i\left(u^{(k)} - u\right) = P_i\left(\sum_{j=1}^{i-1} u_j^{(k)} + \sum_{j=i+1}^{N} u_j^{(k)}\right),$$

de sorte qu'en l'appliquant à (2.11), on obtient

$$\begin{aligned} \forall k \in \mathbb{N}^*, \quad u^{(k+1)} &= u^{(k)} + \sum_{i=1}^N u_i^{(k+1)} \\ &= u^{(k)} - \sum_{i=1}^N P_i\left(\sum_{j=1}^{i-1} u_j^{(k)} + \sum_{j=i+1}^{N} u_j^{(k)}\right) \\ &= u^{(k)} - \sum_{i=1}^N P_i\left(u^{(k)} - u\right). \end{aligned} \tag{2.18}$$

On obtient donc

$$\forall k \in \mathbb{N}^*, \quad u^{(k+1)} - u = \left(Id - \sum_{i=1}^N P_i\right)\left(u^{(k)} - u\right), \tag{2.19}$$

et par conséquent, par application récursive de cette identité,

$$\forall k \in \mathbb{N}^*, \quad u^{(k)} - u = \left(Id - \sum_{i=1}^N P_i\right)^{k-1}\left(u^{(1)} - u\right).$$

Nous concluons en utilisant $u^{(1)} = \left(\sum_{i=1}^N P_i\right) u$, ce qui est en accord avec (2.8) et (2.9). □

Quiconque est familier avec les méthodes de décomposition de domaine remarquera la forte similarité entre la formule d'erreur (2.17) et celle de la méthode de Schwarz additive (voir [24] par exemple) appliquée à N sous-domaines, lorsqu'elle est utilisée sur le système continu. Pour souligner ce fait, il convient également de remarquer que la relation de récurrence (2.19) est vérifiée pour $k = 0$ si l'on pose $u^{(0)} = 0$. Ainsi, bien qu'un choix particulier lors de la phase d'initialisation de la méthode des réflexions soit effectué, il apparaît que toute fonction harmonique dans $\Omega \setminus \left(\cup_{j=1}^{N} \overline{\partial O_j}\right)$ peut être utilisée en théorie comme initialisation. Nous profitons de cette observation au paragraphe 2.4.1.

Nous en déduisons une formule sur la propagation de l'erreur pour la version séquentielle de la méthode.

Théorème 2.3.2. *Supposons que le problème étendu* (2.14) *admet une unique solution u, telle que sa restriction à $\Omega \setminus \left(\cup_{j=1}^{N} \overline{O_j}\right)$ soit la solution du problème* (2.7)*, et que les suites $(u_i^{(k)})_{k\in\mathbb{N}^*,\, i\in[\![1;N]\!]}$ et $(u^{(k)})_{k\in\mathbb{N}^*}$ soient définies par la version séquentielle de la méthode des réflexions de manière unique. Alors, on a la formule d'erreur suivante*

$$\forall k \in \mathbb{N}^*, \quad u^{(k)} - u = -{E_N}^k u, \tag{2.20}$$

où, pour tout $i \in [\![1; N]\!]$, l'opérateur de propagation d'erreur E_i est défini par

$$E_i = (Id - P_i) \dots (Id - P_1). \tag{2.21}$$

Démonstration. La preuve que nous présentons ici consiste à démontrer successivement les trois égalités suivantes :

1. $\forall k \in \mathbb{N}^*, \quad u^{(k)} = \sum_{\ell=0}^{k-1} {E_N}^\ell u^{(1)}$,
2. $u^{(1)} = (Id - E_N)\, u$,
3. $\forall k \in \mathbb{N}^*, \quad u^{(k)} = \left(Id - {E_N}^k\right) u$.

Étape 1. En appliquant l'opérateur P_i et d'après (2.13), l'itération (2.10) peut être reformulée de la manière suivante

$$\forall k \in \mathbb{N}^*, \quad u_i^{(k+1)} = -P_i \left(\sum_{j=1}^{i-1} u_j^{(k+1)} + \sum_{j=i+1}^{N} u_j^{(k)} \right). \tag{2.22}$$

De cette façon, un cycle de la méthode des réflexions se lit comme un produit de N projecteurs.

Considérons maintenant la suite $(v_i^{(\ell)})_{\ell\in\mathbb{N}^*,\, i\in[\![1;N]\!]}$ définie par

$$v_i^{(1)} = \sum_{j=1}^{i} u_j^{(1)} \tag{2.23}$$

$$\forall \ell \in \mathbb{N}^*, \quad v_i^{(\ell+1)} = \sum_{j=1}^{i} u_j^{(\ell+1)} + \sum_{j=i+1}^{N} u_j^{(\ell)}. \tag{2.24}$$

Une propriété cruciale de cette suite est que

$$\forall \ell \in \mathbb{N}^*,\ \forall i \in [\![1; N]\!], \quad v_i^{(\ell+1)} = E_i {E_N}^{\ell-1} u^{(1)}. \tag{2.25}$$

En effet, en ajoutant $\sum_{j=1}^{i-1} u_j^{(\ell+1)} + \sum_{j=i+1}^{N} u_j^{(\ell)}$ dans chacun des membres de (2.22) et en utilisant le fait que $u_i^{(\ell)} = P_i(u_i^{(\ell)})$, on obtient

$$\begin{aligned}\forall \ell \in \mathbb{N}^*, \quad \sum_{j=1}^{i} u_j^{(\ell+1)} + \sum_{j=i+1}^{N} u_j^{(\ell)} &= \sum_{j=1}^{i-1} u_j^{(\ell+1)} + \sum_{j=i+1}^{N} u_j^{(\ell)} - P_i\left(\sum_{j=1}^{i-1} u_j^{(\ell+1)} + \sum_{j=i+1}^{N} u_j^{(\ell)}\right) \\ &= \sum_{j=1}^{i-1} u_j^{(\ell+1)} + \sum_{j=i}^{N} u_j^{(\ell)} - P_i\left(\sum_{j=1}^{i-1} u_j^{(\ell+1)} + \sum_{j=i}^{N} u_j^{(\ell)}\right).\end{aligned}$$

En utilisant la définition de $v_i^{(\ell+1)}$, on peut réécrire la dernière égalité ainsi

$$\forall \ell \in \mathbb{N}^*, \quad \forall i \in [\![2; N]\!],\ v_i^{(\ell+1)} = (Id - P_i) v_{i-1}^{(\ell+1)}$$

ou

$$\forall \ell \in \mathbb{N}^*, \quad v_1^{(\ell+1)} = (Id - P_1) v_N^{(\ell)},$$

sinon. Ces relations de récurrence, combinées avec (2.23) et (2.8), nous permettent d'obtenir (2.25).

D'autre part, d'après (2.23), (2.24) et (2.11), on obtient

$$\forall \ell \in \mathbb{N}^*, \quad u^{(\ell+1)} - u^{(\ell)} = v_N^{(\ell+1)}.$$

En raison de (2.25), on obtient

$$\forall \ell \in \mathbb{N}^*, \quad u^{(\ell+1)} - u^{(\ell)} = {E_N}^{\ell} u^{(1)}.$$

Sommer ces égalités de $\ell = 0$ à k, pour $k \in \mathbb{N}^*$, permet de terminer la preuve de cette étape 1.

Étape 2. Soit i tel que $1 \leq i \leq N$. D'après (2.12), $u_i^{(1)}$ et $u - \sum_{j=1}^{i-1} u_j^{(1)}$ satisfont la même condition aux bords de ∂O_i. En utilisant la définition du projecteur P_i, on obtient

$$\forall i \in [\![1; N]\!], \quad u_i^{(1)} = P_i\left(u - \sum_{j=1}^{i-1} u_j^{(1)}\right).$$

Supposons maintenant que

$$\forall i \in [\![1; N-1]\!], \quad \sum_{j=1}^{i} u_j^{(1)} = (Id - E_i)\, u. \tag{2.26}$$

On constate alors que

$$\begin{aligned}\sum_{j=1}^{i+1} u_j^{(1)} &= P_{i+1}\left(u - \sum_{j=1}^{i} u_j^{(1)}\right) + (Id - E_i)\, u \\ &= P_{i+1}\left(Id - (Id - E_i)\right) u + (Id - E_i)\, u \\ &= P_{i+1} E_i u + (Id - E_i)\, u \\ &= (Id - E_{i+1})\, u.\end{aligned}$$

Par conséquent, (2.26) est aussi valide pour $i+1$ et ainsi pour tout $i \in [\![1; N]\!]$.
L'étape 2 est alors prouvée en prenant $i = N$ dans (2.26) et en combinant le résultat avec (2.8).

Étape 3. Il résulte des deux premières étapes que

$$\forall k \in \mathbb{N}^*, \quad u^{(k)} = \sum_{\ell=0}^{k-1} {E_N}^{\ell} (Id - E_N)\, u,$$

ce qui nous conduit à l'égalité désirée comme $\sum_{\ell=0}^{k-1} {E_N}^{\ell} (Id - E_N) = Id - {E_N}^{k}$. □

Comme précédemment, on peut remarquer que la formule d'erreur (2.20) est très proche de celle de la méthode de Schwarz appliquée à N sous-domaines, du point de vue du problème continu.

2.4 Analyse de la convergence

Compte tenu des formules de propagation d'erreur obtenues dans la section précédente, il est clair que l'erreur de la méthode se comporte comme une suite géométrique. Un premier résultat général de convergence peut donc être trivialement établi.

Théorème 2.4.1. *Une condition nécessaire et suffisante sur les opérateurs P_i, pour $i \in [\![1; N]\!]$, pour que la méthode des réflexions converge est la suivante :*

- $\rho\left(\sum_{i=1}^{N} P_i\right) < 2$ *pour la version parallèle,*
- $\rho\left(E_N\right) < 1$ *pour la version séquentielle,*

où $\rho(\cdot)$ représente le rayon spectral.

Un revers conséquent de la généralité de ce résultat est qu'il est difficile à démontrer dans des cas pratiques. Dans certaines références (voir [5, 41] pour la forme parallèle ou la section 2.4.3 pour la séquentielle), il est adapté afin d'obtenir une condition suffisante pour la sommabilité de la série.

Dans certaines applications, en particulier lorsque le problème à résoudre est de type elliptique, les opérateurs P_i peuvent être vus comme des projecteurs *orthogonaux*. Ceci entraîne des résultats de convergence plus forts. Nous allons donc distinguer les cas pour lesquels les opérateurs de projections sont *orthogonaux*, des cas où ils ne le sont pas. Pour illustrer cet aspect important, plusieurs exemples pratiques sont abordés, y compris celui présenté dans [36].

2.4.1 Résultats théoriques dans le cas orthogonal

Dans cette sous-section, inspirée par les travaux de Luke [36] sur la forme séquentielle de la méthode, mais aussi par l'approche de Xu et Zikatanov dans [46], nous faisons un certain nombre d'hypothèses supplémentaires afin d'étoffer le cadre mathématique précédent. Premièrement, l'espace H est à présent un espace de Hilbert. Nous supposons de plus que le problème peut être reformulé sous une forme variationnelle équivalente, par l'intermédiaire d'une forme bilinéaire continue $a(\cdot, \cdot)$ de $H \times H$ dans $\mathbb{R}$, induite par l'opérateur $\mathcal{L}$, définissant ainsi un produit scalaire sur H. Dans ce contexte, l'espace V est le sous-espace de H dont les éléments sont les solutions (faibles) du problème quand le terme source est associé à tous les objets. On définit aussi les espaces V_i qui sont les espaces de solutions pour les sous-problèmes où le terme source ne fait intervenir que le $i^{\text{è}}$ objet. Il est clair que les espaces V_i, pour tout $i \in [\![1; N]\!]$, sont des sous-espaces fermés de V et que nous avons généralement, en raison du fait que les objets soient disjoints, que $V = \overline{\sum_{i=1}^{N} V_i}$.

On peut montrer que les opérateurs de projection $(P_i)_{i \in [\![1;N]\!]}$ précédemment introduits sont orthogonaux s'ils satisfont la condition suivante

$$\forall u \in H, \quad \forall v \in V_i,\ a(P_i u, v) = a(u, v).$$

Quand c'est le cas, on peut alors introduire les projecteurs orthogonaux définis pour tout $i \in [\![1; N]\!]$ par $P_{M_i} = Id - P_i$. Ces derniers projettent H sur $M_i = V_i^{\perp}$ avec $i \in [\![1; N]\!]$.

Forme parallèle

Comme vue ci-dessus, la méthode des réflexions ne pourra pas converger si le rayon spectral de l'opérateur $\sum_{i=1}^{N} P_i$ est plus grand que deux. Le fait que les projecteurs soient orthogonaux n'améliore généralement pas ce résultat. On pourra néanmoins consulter l'article [9] qui présente certains résultats sur le spectre d'une somme d'opérateurs de projection orthogonale. Toutefois, on peut observer que la formule d'erreur (2.17) correspond à celle d'une méthode itérative[5] du type Richardson stationnaire appliquée à la résolution du système linéaire

$$\left(\sum_{i=1}^{N} P_i\right) u = u^{(1)},$$

impliquant l'opérateur parfois dénommé *l'opérateur de Schwarz additif* $\sum_{i=1}^{N} P_i$ et où nous avons fait le choix $u^{(0)} = 0$. Il apparaît alors que la solution u peut être obtenue de façon avantageuse par une méthode de Krylov, comme celle du gradient conjugué (lorsque les opérateurs de projection sont symétriques), dont la vitesse de convergence est déterminée par $\sum_{i=1}^{N} P_i$. Une telle approche a été utilisée dans [28] pour résoudre un problème résistif pour lequel la méthode des réflexions est en défaut.

Cependant, une manière d'obtenir la convergence inconditionnelle, tout en conservant le caractère parallèle de la méthode, est de modifier légèrement l'algorithme de telle sorte que la formule de récurrence de la méthode devient

$$\forall k \in \mathbb{N}^*, \quad u^{(k+1)} - u = \left(Id - \frac{1}{N}\sum_{i=1}^{N} P_i\right)\left(u^{(k)} - u\right) = \left(\frac{1}{N}\sum_{i=1}^{N}(Id - P_i)\right)\left(u^{(k)} - u\right),$$

à la place de (2.19), ainsi que $u^{(1)} = \frac{1}{N}\sum_{i=1}^{N} P_i u$, ce qui donne

$$\forall k \in \mathbb{N}^*, \quad u^{(k)} - u = -\left(\frac{1}{N}\sum_{i=1}^{N}(Id - P_i)\right)^k u.$$

En effet, il est bien connu, depuis les travaux de Cimmino [13] sur la *méthode des projections moyennées*[6] et son extension aux ensembles convexes par Auslender [3], que pour tout sous-espace (linéaire) fermé $M_1, \ldots, M_N$ de H, on a le résultat de convergence simple suivant

$$\forall v \in H, \quad \lim_{k\to+\infty} \left\| \left(\sum_{i=1}^{N} \alpha_i P_{M_i}\right)^k v - P_{\cap_{i=1}^{N} M_i} v \right\|_H = 0,$$

où les scalaires $(\alpha_i)_{i\in[\![1;N]\!]}$ sont les poids des combinaisons convexes, choisis de telle sorte que $\forall i \in [\![1;N]\!]$, $\alpha_i > 0$ et $\sum_{i=1}^{n} \alpha_i = 1$. Dans le cadre actuel, on a, $\forall i \in [\![1;N]\!]$, $\alpha_i = \frac{1}{N}$ et $Id - P_i = P_{M_i}$, qui est l'opérateur de projection orthogonale sur M_i. Ceci implique la convergence de la méthode parallèle modifiée en utilisant le fait que $\cap_{i=1}^{N} M_i = \cap_{i=1}^{N} V_i^{\perp} = V^{\perp}$, puisque u appartient à V.

D'un point de vue pratique, la mise à jour de la formule pour approcher la solution est maintenant

$$u^{(k+1)} = u^{(k)} + \frac{1}{N}\sum_{i=1}^{N} u_i^{(k+1)},$$

[5] En effet, dans sa forme la plus simple, qui est celle sans pré-conditionnement, la méthode de Richardson stationnaire pour la résolution de système linéaire $Ax = b$ est définie par la relation de récurrence suivante

$$\forall k \in \mathbb{N}, \quad x^{(k+1)} = (I - A)\, x^{(k)} + b,$$

l'initialisation $x^{(0)}$ étant choisie de façon arbitraire.

[6] Méthode itérative pour résoudre des systèmes d'équations linéaires par une approche géométrique

le champ $u_i^{(k+1)}$ est déterminé par la résolution du sous-problème (2.10) dont la seconde équation est remplacée par

$$\mathcal{B}_i(u_i^{(k+1)}) = -\mathcal{B}_i\left(\frac{1}{N}\sum_{j=1}^{i-1} u_j^{(k)} + \frac{1-N}{N}\,u_i^{(k)} + \frac{1}{N}\sum_{j=i+1}^{N} u_j^{(k)}\right).$$

Nous pouvons remarquer que l'interprétation des champs auxiliaires comme étant des "réflexions" est encore valide pour cette version modifiée, puisque l'on a

$$\forall k \in \mathbb{N}^*,\ \forall i \in [\![1;N]\!], \quad \mathcal{B}_i\left(u^{(k)} + u_i^{(k+1)}\right) = b_i.$$

Forme séquentielle

Quand les projecteurs sont orthogonaux, la forme séquentielle de la méthode des réflexions est étroitement liée à la *méthode (cyclique) des projections alternées* (acronyme MPA), qui est un algorithme itératif pour déterminer le projeté orthogonal d'un point de H dans $\cap_{i=1}^N M_i$, en considérant les projections successives de l'itéré courant dans les sous-espaces M_i pour $i \in [\![1;N]\!]$. Cette méthode trouve des applications dans la résolution de systèmes linéaires (comme la méthode de Kaczmarz) ou des équations aux dérivées partielles (comme la méthode de Schwarz), dans l'approximation de fonctions multivariées, dans la construction numérique d'applications conformes, en probabilités et en statistiques (méthode ordinaire des moindres carrés), dans les méthodes multi-grilles, dans la restauration d'images et en tomodensitométrie. Le lecteur intéressé pourra trouvé plus de détails dans l'état de l'art [16] de Deutsch. Sa conception repose sur le résultat ci-dessous[7], d'abord démontré par Von Neumann en 1933, mais non publié jusqu'en 1949, dans le cas où $N = 2$ lorsqu'il travaillait sur la théorie des opérateurs [38] et étendu plus tard par Halperin dans [25] pour tout entier $N \geq 2$.

Théorème 2.4.2. *Soit $M_1, \ldots, M_N$ des sous-espaces fermés de l'espace de Hilbert H. Alors, pour tout v dans H, on a*

$$\lim_{k\to+\infty} \|(P_{M_N}P_{M_{N-1}}\ldots P_{M_1})^k v - P_{\cap_{i=1}^N M_i} v\|_H = 0.$$

Ainsi, la suite d'opérateurs $((P_{M_N}P_{M_{N-1}}\ldots P_{M_1})^k)_{k\in\mathbb{N}}$ converge simplement vers $P_{\cap_{i=1}^N M_i}$. Ce résultat donne lieu à la MPA, qui consiste, pour tout v dans H, à construire une suite $(v^{(k)})_{k\in\mathbb{N}}$ définie par

$$w^{(0)} = v \quad \text{et} \quad \forall k \in \mathbb{N},\ w^{(k+1)} = P_N \ldots P_1 w^{(k)},$$

ce qui revient à appliquer de manière récursive les projecteurs orthogonaux à l'itéré courant.

Grâce au théorème 2.3.2, on peut facilement déduire que la convergence de la version séquentielle de la méthode des réflexions résulte de la convergence de la MPA.

Taux de convergence

Remarquons que le résultat ci-dessus ne donne aucune information sur le taux de convergence de la méthode. Nous traitons maintenant cet aspect du problème en rappelant le résultat de dichotomie suivant pour la MPA, voir [17, théorème 6.4] et [7, théorème 1.4] pour le cas $N = 2$.

Théorème 2.4.3. *Soit $M_1, \ldots, M_N$ des sous-espaces fermés d'un espace de Hilbert H. Alors l'une des deux assertions suivantes est vérifiée*

[7] Notons que ce résultat reste valable lorsque les sous-espaces fermés sont remplacés par des sous-espaces affines fermés ayant une intersection non vide, voir [18, section 4].

1. *Si* $\sum_{i=1}^{N} M_i^{\perp}$ *est fermée, alors la suite* $((P_{M_N} P_{M_{N-1}} \dots P_{M_1})^k)_{k\in\mathbb{N}}$ *converge linéairement*[8] *vers* $P_{\cap_{i=1}^{N} M_i}$, *autrement dit, il existe des constantes* $C > 0$ *et* $\alpha \in [0,1)$ *telles que*

$$\forall k \in \mathbb{N}, \quad \left\| (P_{M_N} P_{M_{N-1}} \dots P_{M_1})^k - P_{\cap_{i=1}^{N} M_i} \right\|_{\mathcal{L}(H)} \leq C\,\alpha^k.$$

2. *Si* $\sum_{i=1}^{N} M_i^{\perp}$ *n'est pas fermée, alors la suite* $((P_{M_N} P_{M_{N-1}} \dots P_{M_1})^k)_{k\in\mathbb{N}}$ *converge arbitrairement vite vers* $P_{\cap_{i=1}^{N} M_i}$, *autrement dit,*

   (i) la suite $((P_{M_N} P_{M_{N-1}} \dots P_{M_1})^k)_{k\in\mathbb{N}}$ *converge simplement vers* $P_{\cap_{i=1}^{N} M_i}$,

   (ii) pour toute fonction à valeurs réelles ϕ, *définie sur l'ensemble des entiers naturels, qui converge vers* 0, *il existe un élément* v *de* H *tel que*

$$\forall k \in \mathbb{N}, \quad \left\| (P_{M_N} P_{M_{N-1}} \dots P_{M_1})^k v - P_{\cap_{i=1}^{N} M_i} v \right\|_{H} \geq \phi(k).$$

Précisons qu'il existe des conditions équivalentes au résultat ci-dessus basées sur la notion d'angle entre sous-espaces[9]. Tout d'abord dans le cas où $N = 2$, il est connu[10] que $c(M_1, M_2) < 1$, où $c(M_1, M_2)$ représente le cosinus de l'angle[11] entre les sous-espaces M_1 et M_2, si et seulement si $M_1 + M_2$ est fermé, si et seulement si $M_1{}^{\perp} + M_2{}^{\perp}$ est fermé, si et seulement si $(M_1 \cap (M_1 \cap M_2)^{\perp}) + (M_2 \cap (M_1 \cap M_2)^{\perp})$ est fermé. Dans le cas où $N \geq 3$, une généralisation de la notion d'angle de Friedrichs entre sous-espaces est introduite dans [4] et il est montré dans le même article que la méthode converge linéairement si et seulement si $c(M_1, \dots, M_N) < 1$, ce qui est une condition plus faible que la condition suffisante, basée sur [18, théorème 2.1] et [4, théorème 4.1], sur le cosinus $c_{ij} = c_0(M_i \cap (\cap_{\ell=1}^{N} M_\ell)^{\perp}, M_j \cap (\cap_{\ell=1}^{N} M_\ell)^{\perp})$ de l'angle de Dixmier entre deux sous-espaces M_i et M_j qui doit être strictement inférieur à un.

Dans le cas de la convergence linéaire de la MPA, des bornes pour l'erreur peuvent être déduites. Rappelons quelques résultats existants sur le sujet. Dans le cas où $N = 2$, on a

$$\left\| (P_{M_2} P_{M_1})^k - P_{M_1 \cap M_2} \right\|_{\mathcal{L}(H)} \leq c(M_1, M_2)^{2k-1},$$

où $c(M_1, M_2)$ désigne le cosinus de l'angle entre les sous-espaces M_1 et M_2. Ce résultat est dû à Aronszajn (voir [2, section 12]) et a été redécouvert plusieurs fois par la suite. Cette borne est optimale, voir [30, théorème 2]. Dans le cas où $N \geq 3$, des bornes supérieures de l'erreur ont été données par Smith, Solomon et Wagner [39, théorème 2.2], par Kayalar et Weinert [30, théorème 3], et aussi par Deutsch et Hundal [18, théorème 2.7]. Toutefois dans ce cas, aucune borne d'erreur

[8] Certains auteurs parlent de convergence *uniforme*, voir [4].

[9] Notamment dans [36], la convergence linéaire de la méthode des réflexions a été établie en utilisant des conditions sur l'écart existant entre des paires de sous-espaces concernés, à savoir plus précisément le sinus des angles entre ces paires de sous-espaces.

[10] Voir par exemple [15]

[11] Selon la définition de Friedrichs [23], si M_i et M_j sont des sous-espaces fermés d'un espace de Hilbert H, l'angle entre M_i et M_j est l'angle compris dans l'intervalle $[0, \frac{\pi}{2}]$ dont le cosinus est défini par

$$c(M_i, M_j) = \sup \left\{ |\langle x, y \rangle| \;|\; x \in M_i \cap (M_i \cap M_j)^{\perp},\ \|x\| \leq 1,\ y \in M_j \cap (M_i \cap M_j)^{\perp},\ \|y\| \leq 1 \right\}.$$

Une autre notion est celle de l'angle minimal entre M_i et M_j, donnée par Dixmier dans [20], qui est l'angle compris dans l'intervalle $[0, \frac{\pi}{2}]$ dont le cosinus est défini par

$$c_0(M_i, M_j) = \sup \left\{ |\langle x, y \rangle| \;|\; x \in M_i,\ \|x\| \leq 1,\ y \in M_j,\ \|y\| \leq 1 \right\}.$$

ne peut-être optimale si elles ne dépendent que des angles entre les différents sous-espaces, voir [18, exemple 3.7]. Plus récemment, Xu et Zikatanov ont obtenu dans [46] l'égalité suivante

$$\left\| P_{M_N} P_{M_{N-1}} \dots P_{M_1} - P_{\cap_{i=1}^N M_i} \right\|_{\mathcal{L}(H)}^2 = \frac{c_0}{1+c_0},$$

où

$$c_0 = \sup_{\|v\|=1} \inf_{\sum_{i=1}^N v_i = v} \sum_{i=1}^N \left\| (Id - P_{M_i}) \sum_{j=i+1}^N v_j \right\|^2, \quad v \in \sum_{j=1}^N M_j{}^{\perp} \quad \text{et} \quad v_j \in M_j{}^{\perp}.$$

Précisons que les différentes manières existantes pour accélérer la MPA, par relaxation ou symétrisation par exemple, ont déjà été proposées et étudiées, voir [8] et les références associées. De telles techniques d'accélération peuvent être directement appliquées à la version séquentielle de la méthode des réflexions.

Enfin, comme les projections parallèles correspondent aux projections alternées dans l'espace produit, une variante adéquate du théorème ci-dessus existe pour la méthode des projections moyennées et est une conséquence des résultats présentés dans [6, 7].

Interprétation en tant que méthode de correction de sous-espaces

L'utilisation de champs auxiliaires dans les deux formes de la méthode des réflexions empêche *a priori* d'obtenir une relation de récurrence "directe" entre $u^{(k+1)}$ et $u^{(k)}$ uniquement en terme de projection. Cependant, on peut remarquer que la formule de propagation d'erreur (2.20) de la version séquentielle ressemble à celle de la *méthode de correction successive de sous-espaces* (MCSS), comme l'a définie Xu dans [44, 45], c'est-à-dire la méthode où l'on a choisi d'initialiser par zéro.

Cette méthode est une généralisation d'une famille plus large et diversifiée d'algorithmes itératifs utilisés pour résoudre des systèmes d'équations linéaires, comme par exemple la méthode de Gauss–Seidel, les méthodes de décomposition multi-grilles et la méthode de décomposition de domaine. Dans [46], un lien entre la MPA et la MCSS est établi. Cette relation peut être exploitée afin de voir la méthode (séquentielle et parallèle) sous un nouveau jour, ce qui amènerait à une description plus générale de ces algorithmes.

Afin de s'en rendre compte, partons comme précédemment de $u^{(0)} = 0$ afin de réécrire la formule de propagation d'erreur (2.20) sous la forme suivante

$$\forall k \in \mathbb{N}, \quad u^{(k)} - u = {E_N}^k \left(u^{(0)} - u \right).$$

Rappelant que pour tout $i \in [\![1; N]\!]$, l'opérateur de projection orthogonale dans le sous-espace affine $M_i + u$ est défini par

$$\forall v \in H, \quad P_{M_i+u} v = P_{M_i}(v-u) + u,$$

nous trouvons le résultat suivant

$$\forall k \in \mathbb{N}, \quad u^{(k)} - u = (P_{M_N+u} P_{M_{N-1}+u} \dots P_{M_1+u})^k u^{(0)} - u,$$

ce qui implique, par induction, la relation de récurrence suivante

$$\begin{aligned} \forall k \in \mathbb{N}, \quad u^{(k+1)} &= P_{M_N+u} P_{M_{N-1}+u} \dots P_{M_1+u} u^{(k)} \\ &= (Id - P_N(\cdot - u))(Id - P_{N-1}(\cdot - u)) \dots (Id - P_1(\cdot - u)) u^{(k)}. \end{aligned} \tag{2.27}$$

La version parallèle de la méthode peut aussi être interprétée comme une version parallèle de la méthode de correction de sous-espaces. Cette famille de méthode comprend par exemple la méthode de Jacobi pour la solution des systèmes linéaires, puisque (2.18) donne directement

$$\forall k \in \mathbb{N}, \quad u^{(k+1)} = \left(Id - \sum_{i=1}^{N} P_i(\cdot - u) \right) u^{(k)}. \tag{2.28}$$

En notant que le problème aux bords sous-jacent associé à $w = P_i(v-u)$ est

$$\begin{cases} \mathcal{L}w = 0 \quad \text{dans } \Omega \setminus \overline{O_i}, \\ \mathcal{B}_i(w) = \mathcal{B}_i(v-u) = \mathcal{B}_i(v) - b_i, \end{cases}$$

nous observons que la nécessité de considérer des champs auxiliaires peut alors être levée, en considérant des projecteurs prenant en compte les données du problème à chaque cycle et non uniquement au cycle d'initialisation. D'un point de vue pratique, une telle forme nécessite moins de mémoire pour sa mise en œuvre. De plus, même si les suites $(u^{(k)})_{k\in\mathbb{N}^*}$ générées par (2.27) ou (2.28) coïncident avec celles construites par les algorithmes "classiques" quand $u^{(0)} = 0$, tous les choix d'initialisation sont possibles[12] dans V avec les "variantes".

2.4.2 Exemples pratiques dans le cas orthogonal

A la vue des résultats précédents, l'orthogonalité des projecteurs P_i est une condition suffisante pour la convergence de la méthode, dans sa forme séquentielle ou dans la version modifiée de sa forme parallèle. Dans cette section, nous montrons sur quelques exemples comment prouver cette propriété. L'idée consiste à écrire une étape de la méthode comme une projection de Galerkin. Nous nous concentrons ainsi sur des exemples qui utilisent l'équation de Laplace et rappelons les résultats de Luke dans [36] pour l'équation de Stokes.

L'équation de Laplace

Dans la suite, nous admettons que les bords ∂O_j et $\partial\Omega$ sont lisses, *i.e.* $\mathcal{C}^2$, donc que les champs considérés sont $\mathcal{C}^1$ sur les clôtures de chaque composante connexe de $\Omega \setminus \left(\cup_{j=1}^{N} \partial O_j\right)$. Dans la suite, nous utiliserons souvent les deux formules suivantes. D'abord, la formule d'intégration de Green qui affirme que, pour w et z dans $H_0^1(\Omega)$, et $\Delta z = 0$ dans $\Omega \setminus \left(\cup_{j=1}^{N} \partial O_j\right)$, on a

$$\int_\Omega \nabla w \cdot \nabla z = \sum_{j=1}^{N} \int_{\partial O_j} [w \partial_n z], \tag{2.29}$$

où $[\cdot]$ dénote le saut à travers la surface d'intégration. La seconde formule est la formule de représentation de Green, selon laquelle tout champ $w \in H_0^1(\Omega)$ satisfaisant $\Delta w = 0$ dans $\Omega \setminus \left(\cup_{j=1}^{N} \partial O_j\right)$ peut être représenté par

$$w(x) = \sum_{j=1}^{N} \int_{\partial O_j} G(x,y)[\partial_n w(y)] - \partial_n G(x,y)[w(y)] d\sigma_y, \tag{2.30}$$

où G est la fonction de Green du problème. On rappelle qu'en notant $\|\cdot\|$ la norme euclidienne usuelle sur $\mathbb{R}^2$ ou $\mathbb{R}^3$, cette fonction est définie par $G := \varphi + F$ avec

$$\forall (x,y) \in \Omega^2, \quad F(x,y) = \begin{cases} -\frac{1}{2\pi} \log(\|x-y\|) & \text{si } \Omega \subset \mathbb{R}^2, \\ \frac{1}{4\pi} \frac{1}{\|x-y\|} & \text{si } \Omega \subset \mathbb{R}^3, \end{cases}$$

[12] Il est à noter que le choix de $u^{(0)} = u$ conduit trivialement à la convergence après le premier cycle.

et φ vérifie $\Delta\varphi(x,\cdot) = 0$ sur Ω et $\varphi(x,y) = -F(x,y)$, $\forall x,y \in \Omega \times \partial\Omega$. Plus de détails peuvent être trouvés dans [21]. Finalement, notons que ces formules et tous les résultats de cette section restent vrais dans des domaines non bornés sur lesquels $G = F$. Dans ces cas, on pourra travailler avec des espaces de Sobolev à poids. On considère un tel cadre dans la section 2.4.3.

Conditions aux bords de Dirichlet On s'attaque d'abord au problème présenté dans la section 2.2.1.

Théorème 2.4.4. *Soit* $\mathcal{L} = -\Delta$, $\mathcal{B}_j = \gamma_{|\partial O_j}$,

$$H = \left\{ w \in H_0^1(\Omega),\ \mathcal{L}u = 0 \text{ sur } \Omega \setminus \left(\cup_{j=1}^N \partial O_j\right),\ \forall j \in [\![1;N]\!], [\mathcal{B}_j w] = 0 \right\}$$

muni du produit scalaire usuel de $H_0^1(\Omega)$. *Pour tout* $i \in [\![1;N]\!]$, *l'opérateur* P_i *défini par (2.15–2.16) est le projecteur orthogonal sur* $V_i = \{w \in H, \mathcal{L}w = 0 \text{ sur } \Omega \setminus \partial O_i,\ [v] = 0 \text{ sur } \partial O_i\}$. *De plus, on a* $H = \sum_{j=1}^N V_j$.

Démonstration. On fixe i tel que $1 \leq i \leq N$ et on démontre le résultat pour P_i. Soit $b_i = \mathcal{B}_i u$, on commence par prouver que les champs u et $P_i u$ vérifient :

$$\int_\Omega \nabla v \nabla u = \sum_{j=1}^N \int_{\partial O_j} [\partial_n v] b_j \qquad \forall v \in H, \tag{2.31}$$

$$\int_\Omega \nabla v \nabla P_i u = \sum_{j=1}^N \int_{\partial O_j} [\partial_n v] b_j \qquad \forall v \in V_i. \tag{2.32}$$

L'équation (2.31) s'obtient en combinant (2.29) avec le fait que $v \in H$, ce qui implique que $\Delta v = 0$ sur $\Omega \setminus \left(\cup_{j=1}^N \partial O_j\right)$. Pour avoir (2.32), notons que du fait de (2.15–2.16), u et $P_i u$ ont les mêmes conditions aux bords de Dirichlet sur ∂O_i. Comme $v \in V_i$, $\Delta v = 0$ sur $\Omega \setminus \partial O_i$, v est alors régulier sur tous les ∂O_j, pour $j \neq i$. Ainsi $[\partial_n v] = 0$ pour $j \neq i$. On utilise ensuite que $V_i \subset H$ et l'orthogonalité provient de l'orthogonalité de Galerkin :

$$\int_\Omega \nabla v \nabla (u - P_i u) = 0 \qquad \forall v \in V_i.$$

Prouvons maintenant que $H = \sum_{j=1}^N V_j$. Utilisant le fait que $[u(y)] = 0$ sur $\cup_{j=1}^N \partial O_j$ dans l'équation (2.30), on a :

$$u(x) = \sum_{j=1}^N \int_{\partial O_j} G(x,y)[\partial_n u(y)] d\sigma_y.$$

On obtient la décomposition $u = \sum_{j=1}^N u_j$ en écrivant que

$$u_i(x) = \int_{\partial O_i} G(x,y)[\partial_n u(y)] d\sigma_y = \int_{\partial O_i} \varphi(x,y)[\partial_n u(y)] d\sigma_y + \int_{\partial O_i} F(x,y)[\partial_n u(y)] d\sigma_y.$$

Dans cette dernière équation, le premier terme est régulier et n'intervient pas dans les valeurs de saut de u_i. Le second terme correspond à l'unique solution w du problème de transmission (voir [14]) sur $\mathbb{R}^p$

$$\begin{cases} -\Delta w = 0 & \text{dans } \mathbb{R}^p \setminus \partial O_i, \\ [w] = 0 & \text{sur } \partial O_i, \\ [\partial_n w] = [\partial_n u(y)] & \text{sur } \partial O_i. \end{cases}$$

et donc $[u_i] = [w] = 0$. La régularité de $\partial\Omega$ et la définition de G impliquant que $u_i \in H_0^1(\Omega)$, on en conclut que $u_i \in V_i$. □

Remarque 2.4.5. *Comme on n'utilise pas le fait que $\Delta u = 0$ dans la preuve, on obtient en fait que P_i est le projecteur orthogonal de $H_0^1(\Omega)$ dans V_i.*

Conditions aux bords de Neumann A notre connaissance, ce cas n'a pas encore été considéré dans la littérature. On pose $\mathcal{B}_j = v \mapsto \partial_n v_{|\partial O_j}$. Comme conséquence de la formulation (2.14), les champs s'étendent dans les objets de telle sorte que la trace de $\partial_n u$ soit continue. Ce prolongement nécessite alors d'être précisé.

En effet, les problèmes intérieurs aux objets sont mal posés et sont soumis à une condition de compatibilité. Pour surpasser cette difficulté, on définit deux sous-espaces de $H_0^1(\Omega)$:

$$H := \left\{ v \in H_0^1(\Omega), -\Delta v = 0 \text{ dans } \Omega \setminus \left(\cup_{j=1}^N \partial O_j\right),\ \forall j \in [\![1;N]\!],\ [\mathcal{B}_j v] = 0,\ \int_{\partial O_j} [v] = 0,\ \int_{\partial O_j} \mathcal{B}_j v = 0 \right\}, \tag{2.33}$$

$$V_i := \left\{ v \in H, -\Delta v = 0 \text{ dans } \Omega \setminus \partial O_i,\ [\mathcal{B}_j v] = 0,\ \int_{\partial O_i} [v] = 0,\ \int_{\partial O_i} \mathcal{B}_i v = 0 \right\}. \tag{2.34}$$

Bien sûr on a $V_i \subset H$. La condition $\int_{\partial O_i} \mathcal{B}_i v = \int_{\partial O_j} \partial_n v = 0$ correspond à la condition usuelle de compatibilité associée à un problème de Neumann pur. Celle-ci est satisfaite dans le cadre des deux versions de l'algorithme présentées dans la section 2.3 : en effet, la projection est toujours appliquée à des champs appartenant à H, voir en particulier les problèmes (2.10) et (2.13). L'unicité de l'extension intérieure est garantie par la contrainte supplémentaire $\int_{\partial O_j} [v] = 0$.

Théorème 2.4.6. *Soit H défini par (2.33), muni du produit scalaire usuel de $H_0^1(\Omega)$ et soit $\mathcal{B}_j = v \mapsto \partial_n v_{|\partial O_j}$, $\mathcal{L} = \Delta$. Pour tout $i \in [\![1;N]\!]$, l'opérateur P_i défini sur H par (2.15–2.16) est le projecteur orthogonal sur le sous-espace V_i, défini par (2.34). De plus, on a $H = \sum_{j=1}^N V_j$.*

Démonstration. On continue de noter $b_i = \mathcal{B}_i u$ et on prouve que les champs u et $P_i u$ vérifient :

$$\int_\Omega \nabla v \nabla u = \sum_{j=1}^N \int_{\partial O_j} [v] b_j \qquad \forall v \in H, \tag{2.35}$$

$$\int_\Omega \nabla v \nabla P_i u = \sum_{j=1}^N \int_{\partial O_j} [v] b_j \qquad \forall v \in V_i. \tag{2.36}$$

L'équation (2.35) s'obtient en utilisant (2.29) avec $\Delta u = 0$ dans $\Omega \setminus \left(\cup_{j=1}^N \partial O_j\right)$. Comme $v \in V_i$ dans l'équation (2.36), $\Delta v = 0$ dans $\Omega \setminus \partial O_i$, donc v est régulier sur tous les ∂O_j, pour $j \neq i$. Donc $[v] = 0$ pour $j \neq i$, et (2.36). L'orthogonalité de P_i s'en suit.

Pour prouver que $H = \sum_{j=1}^N V_j$, on utilise que $[\partial_n u(y)] = 0$ sur $\cup_{j=1}^N \partial O_j$ pour simplifier l'équation (2.30). On obtient :

$$u(x) = -\sum_{j \in [\![1;N]\!]} \int_{\partial_n O_j} \partial_n G(x,y)[u(y)] d\sigma_y.$$

On a donc la décomposition $u = \sum_{j \in [\![1;N]\!]} u_j$ en posant :

$$u_i(x) = -\int_{\partial O_i} \partial_n G(x,y)[u(y)] d\sigma_y.$$

Avec le même type de raisonnement que dans le cas de Dirichlet, on trouve que $u_i \in H_0^1(\Omega)$, $[\partial_n u_i] = [\partial_n u] = 0$ et $[u_i] = [u]$ ce qui implique que $\int_{\partial O_i} [u_i] = \int_{\partial O_i} [u] = 0$. Par conséquent $u_i \in V_i$. □

Problème de mobilité avec conditions aux bords des quatre types Ce problème correspond à celui présenté en section 2.2.2. Posons $\mathcal{B}_j = v \mapsto v_{|\partial O_j} - m_j(v)$, où $m_j(v) = \frac{1}{\int_{\partial O_j} 1} \int_{\partial O_j} v$. Nous définissons deux sous-espaces de $H_0^1(\Omega)$:

$$H := \left\{ v \in H_0^1(\Omega), -\Delta v = 0 \text{ dans } \Omega \setminus \left(\cup_{j=1}^N \partial O_j\right),\ \forall j \in [\![1;N]\!],\ [\mathcal{B}_j v] = 0,\ \int_{\partial O_j} [v] = 0,\ \int_{\partial O_j} [\partial_n v] = \int_{\partial O_j} \partial_n v = 0 \right\}, \tag{2.37}$$

$$V_i := \left\{ v \in H, -\Delta v = 0 \text{ dans } \Omega \setminus \partial O_i,\ [\mathcal{B}_i v] = 0,\ \int_{\partial O_i} [v] = 0,\ \int_{\partial O_i} [\partial_n v] = \int_{\partial O_i} \partial_n v = 0 \right\}, \tag{2.38}$$

de sorte que $V_i \subset H$. Comme dans le cas de Neumann, les contraintes $\int_{\partial O_j}[v] = 0$ et $\int_{\partial O_j}[\partial_n v] = 0$, garantissent le fait que les problèmes à l'intérieur des objets soient bien posés.

Théorème 2.4.7. *Soit H définie en (2.37) que l'on munit du produit scalaire usuel de $H_0^1(\Omega)$ et soit $\mathcal{B}_j = v \mapsto v_{|\partial O_j} - m_j(v)$ et $\mathcal{L} = \Delta$. Pour tout $i \in [\![1;N]\!]$, l'opérateur P_i défini sur H par (2.15–2.16) projette orthogonalement H sur le sous-espace V_i, défini par (2.38). De plus, nous avons $H = \sum_{j=1}^N V_j$.*

Démonstration. Utilisant (2.29) comme dans le cas de Dirichlet, nous avons

$$\begin{aligned} \int_\Omega \nabla v \nabla u &= \sum_{j=1}^N \int_{\partial O_j} [\partial_n v] b_j \qquad \forall v \in H, \\ \int_\Omega \nabla v \nabla P_i u &= \sum_{j=1}^N \int_{\partial O_j} [\partial_n v] b_j \qquad \forall v \in V_i. \end{aligned}$$

En complément des preuves précédentes, nous avons fait usage de l'argument suivant $\int_{\partial O_j}[\partial_n v] m_j(u) = m_j(u) \int_{\partial O_j}[\partial_n v] = 0$, ce qui implique que $\int_{\partial O_j}[\partial_n v] u = \int_{\partial O_j}[\partial_n v](u - m_j(u)) = \int_{\partial O_j}[\partial_n v] b_i$. L'orthogonalité de P_i s'ensuit.

Il reste à prouver que $H = \sum_{j=1}^N V_j$. Comme dans le cas de Dirichlet, nous utilisons que $[u] = 0$ sur $\cup_{j=1}^N \partial O_j$ dans l'équation (2.30) et nous obtenons

$$u(x) = \sum_{j=1}^N \int_{\partial O_j} G(x,y)[\partial_n u(y)] d\sigma_y.$$

Nous obtenons la décomposition $u = \sum_{j=1}^N u_j$ en posant

$$u_i(x) = \int_{\partial O_i} G(x,y)[\partial_n u(y)] d\sigma_y.$$

Le fait que $u_i \in V_i$ s'obtient à l'aide du même type d'arguments que dans les démonstrations précédentes. □

Problème de mobilité pour les équations de Stokes

Dans [36], Luke propose une variante de la méthode des réflexions (la version séquentielle) et a prouvé sa convergence lorsqu'elle est appliquée à la résolution d'un système d'équations modélisant le mouvement d'une suspension dans un fluide. Dans ce paragraphe, nous rappelons et résumons ses résultats.

Nous appellerons Ω le *conteneur* (un ouvert borné, connexe et lisse de $\mathbb{R}^3$), $(O_i \subset \Omega)_{i \in [\![1;N]\!]}$ les *particules rigides* (qui sont aussi des ouverts bornés, connexes et lisses de $\mathbb{R}^3$, tel qu'il n'y ait

pas de recouvrement), $\cup_{j=1}^{N} O_j$ la *phase solide* et $\Omega \setminus \overline{\left(\cup_{j=1}^{N} O_j\right)}$ la *phase liquide*. le mouvement des particules rigides, sans inertie, soumis à des forces et des moments de torsions extérieurs est décrit par le champ de vitesse du fluide $\boldsymbol{u}$ et de pression p qui vérifie le problème suivant

$$\mu\,\Delta \boldsymbol{u} = \boldsymbol{\nabla} p \quad \text{dans } \Omega \setminus \left(\cup_{j=1}^{N} \partial O_j\right) \tag{2.39}$$

$$\operatorname{div} \boldsymbol{u} = 0 \quad \text{dans } \Omega \setminus \left(\cup_{j=1}^{N} \partial O_j\right) \tag{2.40}$$

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{sur } \partial\Omega \tag{2.41}$$

$$[\boldsymbol{u}]_{|\partial O_j} = 0, \quad j \in [\![1; N]\!], \tag{2.42}$$

$$\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{U}_j + \boldsymbol{\Omega}_j \times (\boldsymbol{x} - \boldsymbol{x}_j), \quad \boldsymbol{x} \in \partial O_j,\ j \in [\![1; N]\!], \tag{2.43}$$

$$\int_{\partial O_j} [\boldsymbol{\sigma}(\boldsymbol{u})]_{|\partial O_j} \cdot \boldsymbol{n}\, \mathrm{d}S = \boldsymbol{F}_j, \quad j \in [\![1; N]\!], \tag{2.44}$$

$$\int_{\partial O_j} (\boldsymbol{x} - \boldsymbol{x}_j) \times [\boldsymbol{\sigma}(\boldsymbol{u})]_{|\partial O_j} \cdot \boldsymbol{n}\, \mathrm{d}S = \boldsymbol{\tau}_j, \quad j \in [\![1; N]\!], \tag{2.45}$$

où μ représente la viscosité du fluide, les points $\boldsymbol{x}_i$ sont les centres de masse des particules, $\boldsymbol{\sigma}(\boldsymbol{u})$ est le tenseur des contraintes, $\boldsymbol{\sigma}(\boldsymbol{u}) = -p\,\mathbf{I} + \mu\,\left(\boldsymbol{\nabla}\boldsymbol{u} + (\boldsymbol{\nabla}\boldsymbol{u})^T\right)$ et la vitesse $\boldsymbol{U}_i$ et la vitesse angulaire $\boldsymbol{\Omega}_i$ de la $i^{\text{è}}$ particule sont les inconnues à déterminer en même temps que l'écoulement du fluide, tandis que le total des forces $\boldsymbol{F}_i = m_i \boldsymbol{g}$ et les moments de torsion $\boldsymbol{\tau}_i$ sur les particules sont donnés [13]. Rappelons que l'écoulement a été prolongé à l'intérieur des particules, c'est pourquoi l'équation de Stokes est aussi vérifiée à l'intérieur et que le champ d'écoulement est continu à travers les frontières.

Introduisons à présent l'espace

$$F = \left\{ \boldsymbol{v} \in (H_0^1(\Omega))^3 \mid \operatorname{div} \boldsymbol{v} = 0 \text{ dans } \mathscr{D}'(\Omega) \right\}$$

et la forme bilinéaire

$$a(\boldsymbol{u}, \boldsymbol{v}) = \mu \int_{\Omega} \boldsymbol{\nabla}\boldsymbol{u} : \boldsymbol{\nabla}\boldsymbol{v}\, \mathrm{d}\boldsymbol{x},$$

pour toutes fonctions $\boldsymbol{u}$ et $\boldsymbol{v}$ dans F. On peut alors définir l'ensemble des solutions faibles de l'équation de Stokes associées aux forces agissant sur le fluide au niveau des frontières des particules en posant

$$\forall K \subset \Omega, \quad \mathscr{F}(K) = \{\boldsymbol{f} \in (H^{-1}(\Omega))^3 \mid \operatorname{supp} \boldsymbol{f} \subset K\},$$

$$H = \left\{ \boldsymbol{u} \in F \mid \exists \boldsymbol{f} \in \mathscr{F}(\cup_{j=1}^{N} \partial O_j),\ \text{tel que } \forall \boldsymbol{v} \in F,\ a(\boldsymbol{u}, \boldsymbol{v}) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{(H^{-1}(\Omega))^3, (H_0^1(\Omega))^3} \right\},$$

et

$$\forall i \in [\![1; N]\!], \quad H_i = \left\{ \boldsymbol{u} \in F \mid \exists \boldsymbol{f} \in \mathscr{F}(\partial O_i),\ \text{tel que } \forall \boldsymbol{v} \in F,\ a(\boldsymbol{u}, \boldsymbol{v}) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{(H^{-1}(\Omega))^3, (H_0^1(\Omega))^3} \right\}.$$

On peut montrer que les éléments de H sont des fonctions de $(H^1(\Omega))^3$ qui vérifient les équations ((2.39)–(2.41)) au sens faible et que les opérateurs de projection $Id - P_i$ de H dans

$$M_i = \left\{ \boldsymbol{u} \in H_i \mid \int_{\partial O_i} [\boldsymbol{\sigma}(\boldsymbol{u})]_{|\partial O_i} \cdot \boldsymbol{n}\, \mathrm{d}S = 0 \text{ et } \int_{\partial O_i} (\boldsymbol{x} - \boldsymbol{x}_i) \times [\boldsymbol{\sigma}(\boldsymbol{u})]_{|\partial O_i} \cdot \boldsymbol{n}\, \mathrm{d}S = 0 \right\},$$

sont orthogonaux par rapport au produit scalaire de H définie par la forme bilinéaire $a(\cdot, \cdot)$. À partir de cela, Luke prouve la convergence linéaire de la méthode en montrant que la somme $\sum_{i=1}^{N} M_i^{\perp}$ est un sous-espace fermé de H.

[13] Par exemple, pour les suspensions plongées dans un champ gravitationnel uniforme, on a $\boldsymbol{F}_i = m_i \boldsymbol{g}$, où m_i est la masse de la particule et $\boldsymbol{g}$ est l'accélération de la pesanteur, et $\boldsymbol{\tau}_i = \boldsymbol{0}$.

2.4.3 Un exemple dans le cas non orthogonal

Nous nous concentrons maintenant sur le problème incluant des objets munis de conditions aux bords de Dirichlet et d'autres munis de condition de Neumann. Au vu de nos connaissances actuelles, nous ne sommes pas capables de prouver que les projecteurs associés à cette situation sont non orthogonaux. Par contre, la divergence de la méthode peut être observée numériquement dans certains tests. Cependant, on peut établir une condition suffisante sur la géométrie pour assurer la convergence de la méthode. Cette section est dédiée à ce résultat. Soit $\Omega = \mathbb{R}^3$ et H l'espace de Sobolev à poids défini dans [1] de la manière suivante

$$H = \left\{ u : \Omega \to \mathbb{R} \mid \int_\Omega \frac{u^2(x)}{1+\|x\|^2} dx < +\infty, \nabla u \in L^2(\Omega) \right\}.$$

Dans ce cadre la fonction de Green est explicite et coïncide avec la solution fondamentale. En particulier, si le champ $u \in H$ satisfait $-\Delta u = 0$ dans $\Omega \setminus \left(\cup_{j=1}^N \partial O_j\right)$, il admet la représentation intégrale suivante

$$u(x) = \frac{1}{4\pi} \sum_{1\le j\le N} \int_{\partial O_j} \left(\frac{n\cdot(x-y)}{\|x-y\|^3}[u(y)] - \frac{1}{\|x-y\|}[\partial_n u(y)] \right) d\sigma_y, \tag{2.46}$$

pour tout $x \in \Omega \setminus \left(\cup_{j=1}^N \partial O_j\right)$. La dérivée normale de u est alors représentée par

$$\partial_n u(x) = -\frac{1}{4\pi} \sum_{1\le j\le N} \int_{\partial O_j} \left(\frac{1}{\|x-y\|^3} \left(1 + 3\frac{(n\cdot(x-y))^2}{\|x-y\|^2}\right)[u(y)] + \frac{n\cdot(x-y)}{\|x-y\|^3}[\partial_n u(y)] \right) d\sigma_y \tag{2.47}$$

Nous étendons le champ à l'intérieur des objets de la même manière que dans (2.14), ce qui permet d'assurer la continuité des conditions aux bords, associées aux données, à travers le bord des objets. Par conséquent, $[u_i] = 0$ (respectivement $[\partial_n u_i] = 0$) pour les objets de type Dirichlet (respectivement Neumann). Dans un tel cas et pour $y \in \partial O_i$, nous notons $u_i(y)$ (respectivement $\partial_n u_i(y)$) la valeur de la trace extérieure de u_i au point y (respectivement $\partial_n u_i$). De la même manière, $u_{i|\partial O_i}$ (respectivement $\partial_n u_{i|\partial O_i}$) fait référence à la trace extérieure de u (respectivement $\partial_n u_i$) sur le bord de ∂O_i.

Supposons que le $i^{\text{è}}$ objet est associé à des conditions de Dirichlet. On note $c_i > 0$ le carré de la constante de continuité associée à l'opérateur de Dirichlet-vers-Neumann, qui est la constante satisfaisant l'inégalité suivante

$$\|\partial_n u_i\|^2_{L^2(\partial O_i)} \le c_i \|b_i\|^2_{L^2(\partial O_i)}, \tag{2.48}$$

où u_i est la solution de (2.14), avec O_i l'unique objet, $\mathcal{B}_j : v \mapsto v_{|\partial O_i}$, $\mathcal{L} = -\Delta$. Si le $i^{\text{è}}$ objet est associé à des conditions de Neumann, alors $c_i > 0$ le carré de la constante de continuité associée à l'opérateur Neumann-vers-Dirichlet, qui est la constante satisfaisant l'inégalité suivante

$$\|u_i\|^2_{L^2(\partial O_i)} \le c_i \|b_i\|^2_{L^2(\partial O_i)}. \tag{2.49}$$

où u_i est la solution de (2.14), avec O_i l'unique objet, $\mathcal{B}_j = v \mapsto \partial_n v_{|\partial O_i}$, $\mathcal{L} = \Delta$. Sans perte de généralité, nous supposons que $c_i \ge 1$.

Nous nous restreindrons dans la suite à la forme séquentielle de l'algorithme. Des résultats similaires peuvent être obtenus pour la version parallèle en utilisant la méthode proposée dans [5].

Commençons par une estimation technique. Nous associons aux termes u_i^k de la suite définie par (2.12) et (2.13), les termes

$$a_{i,m}^{(k)} := \left\| u_i^{(k+1)} \right\|^2_{L^2(\partial O_m)} + \left\| \partial_n u_i^{(k+1)} \right\|^2_{L^2(\partial O_m)}.$$

Lemme 2.4.8. *On note S_i la surface du $i^{\grave{e}}$ objet. Soit $m \in \mathbb{N}$, tel que $1 \leq m \leq N$ et $m \neq i$. Pour tout $m \neq i$, on a*

$$a_{i,m}^{(k+1)} \leq \kappa_{i,m} \left(\sum_{j=1}^{i-1} a_{j,i}^{(k+1)} + \sum_{j=i+1}^{N} a_{j,i}^{(k)} \right), \tag{2.50}$$

où

$$\kappa_{i,m} = \frac{(N-1) S_i S_m c_i}{\pi^2} \left(\min \left(\frac{1}{16 d_{i,m}^2}, \frac{1}{d_{i,m}^6} \right) + \frac{1}{16 d_{i,m}^4} \right), \tag{2.51}$$

avec $d_{i,m} = \min_{(x,y) \in \partial O_i \times \partial O_m} (\|x - y\|)$.

Démonstration. Nous distinguons dans la démonstration les objets de type Dirichlet de ceux de type Neumann.

Objets de type Dirichlet : Supposons que i représente l'indice associé à un objet de type Dirichlet. En raison du problème (2.13), nous avons

$$u_{i|\partial O_i}^{(k+1)} = - \sum_{j=1}^{i-1} u_{j|\partial O_i}^{(k+1)} - \sum_{j=i+1}^{N} u_{j|\partial O_i}^{(k)}.$$

Par conséquent, pour tout $y \in \partial O_i$

$$\left[\partial_n u_i^{(k+1)}(y) \right] = \partial_n u_i^{(k+1)}(y) + \sum_{j=1}^{i-1} \partial_n u_j^{(k+1)}(y) + \sum_{j=i+1}^{N} \partial_n u_j^{(k)}(y).$$

D'autre part, en raison du problème (2.46), on a pour tout $x \in \Omega$

$$u_i^{(k+1)}(x) = - \frac{1}{4\pi} \int_{\partial O_i} \frac{1}{\|x - y\|} \left[\partial_n u_i^{(k+1)}(y) \right] d\sigma_y$$

si bien que

$$\begin{aligned} u_i^{(k+1)}(x) \quad = \quad & - \frac{1}{4\pi} \int_{\partial O_i} \frac{1}{\|x - y\|} \partial_n u_i^{(k+1)}(y) d\sigma_y - \frac{1}{4\pi} \sum_{j=1}^{i-1} \int_{\partial O_i} \frac{1}{\|x - y\|} \partial_n u_j^{(k+1)}(y) d\sigma_y \\ & - \frac{1}{4\pi} \sum_{j=i+1}^{N} \int_{\partial O_i} \frac{1}{\|x - y\|} \partial_n u_j^{(k)}(y) d\sigma_y. \end{aligned}$$

Supposons maintenant que $x \in \partial O_m$, avec $m \neq i$. En appliquant l'inégalité de Cauchy-Schwarz, le membre de droite peut être estimé par

$$\left| u_i^{(k+1)}(x) \right| \leq \frac{\sqrt{S_i}}{4\pi d_{i,m}} \left(\left\| \partial_n u_i^{(k+1)} \right\|_{L^2(\partial O_i)} + \sum_{j=1}^{i-1} \left\| \partial_n u_j^{(k+1)} \right\|_{L^2(\partial O_i)} + \sum_{j=i+1}^{N} \left\| \partial_n u_j^{(k)} \right\|_{L^2(\partial O_i)} \right).$$

En utilisant l'inégalité (2.48), $c_i \geq 1$, en élevant au carré et en intégrant sur ∂O_m cette dernière inégalité, nous obtenons

$$\left\| u_i^{(k+1)} \right\|_{L^2(\partial O_m)}^2 \leq \frac{(N-1) S_i S_m c_i}{16 \pi^2 d_{i,m}^2} \left(\sum_{j=1}^{i-1} a_{j,i}^{(k+1)} + \sum_{j=i+1}^{N} a_{j,i}^{(k)} \right). \tag{2.52}$$

En appliquant cette fois-ci (2.47), nous répétons cette opération avec $\partial_n u_i^{(k)}$ à la place de $u_i^{(k)}$. La seule différence dans les calculs concerne le terme $\frac{1}{d_{i,m}}$ qui est remplacé par $\frac{1}{d_{i,m}^2}$. Nous obtenons alors

$$\left\| \partial_n u_i^{(k+1)} \right\|_{L^2(\partial O_m)}^2 \leq \frac{(N-1) S_i S_m c_i}{16 \pi^2 d_{i,m}^4} \left(\sum_{j=1}^{i-1} a_{j,i}^{(k+1)} + \sum_{j=i+1}^{N} a_{j,i}^{(k)} \right). \tag{2.53}$$

En sommant (2.52) et (2.53), nous observons que (2.50) est valide pour les objets de type Dirichlet.

Objets de type Neumann : Nous renouvelons le raisonnement précédent. Supposons que i représente l'indice associé à un objet de type Neumann. Dans ce cas, (2.13) implique que

$$\partial_n u_{i|\partial O_i}^{(k+1)} = -\sum_{j=1}^{i-1} \partial_n u_{j|\partial O_i}^{(k+1)} - \sum_{j=i+1}^{N} \partial_n u_{j|\partial O_i}^{(k)}$$

ce qui donne pour tout $y \in \partial O_i$

$$\left[u_i^{(k+1)}(y)\right] = u_i^{(k+1)}(y) + \sum_{j=1}^{i-1} u_j^{(k+1)}(y) + \sum_{j=i+1}^{N} u_j^{(k)}(y) + \lambda_i^{(k+1)}.$$

où $\lambda_i^{(k+1)}$ est un réel choisi après le calcul de $u_i^{(k+1)}$ de telle sorte que $\int_{\partial O_i} \left[u_i^{(k+1)}\right] = 0$. D'autre part, d'après (2.46), on a pour tout $x \in \Omega$

$$u_i^{(k+1)}(x) = \frac{1}{4\pi} \int_{\partial O_i} \frac{n \cdot (x-y)}{\|x-y\|^3} \left[u_i^{(k+1)}(y)\right] d\sigma_y$$

de sorte que

$$\begin{aligned} u_i^{(k+1)}(x) \;=\; & \frac{1}{4\pi} \int_{\partial O_i} \frac{n \cdot (x-y)}{\|x-y\|^3} u_i^{(k+1)}(y) d\sigma_y + \frac{1}{4\pi} \sum_{j=1}^{i-1} \int_{\partial O_i} \frac{n \cdot (x-y)}{\|x-y\|^3} u_j^{(k+1)}(y) d\sigma_y \\ & + \frac{1}{4\pi} \sum_{j=i+1}^{N} \int_{\partial O_i} \frac{n \cdot (x-y)}{\|x-y\|^3} u_j^{(k)}(y) d\sigma_y + \frac{\lambda_i^{(k+1)}(N-1)}{4\pi} \int_{\partial O_i} \frac{n \cdot (x-y)}{\|x-y\|^3} d\sigma_y. \end{aligned}$$

En admettant que $x \notin O_i$, ce dernier terme s'annule.
Supposons maintenant que $x \in \partial O_m$, avec $m \neq i$. En appliquant l'inégalité de Cauchy-Schwarz, le membre de droite peut être estimé par

$$\left|u_i^{(k+1)}(x)\right| \leq \frac{\sqrt{S_i}}{4\pi d_{i,m}^2} \left(\left\|u_i^{(k+1)}\right\|_{L^2(\partial O_i)} + \sum_{j=1}^{i-1} \left\|u_j^{(k+1)}\right\|_{L^2(\partial O_i)} + \sum_{j=i+1}^{N} \left\|u_j^{(k)}\right\|_{L^2(\partial O_i)} \right).$$

En appliquant l'inégalité (2.49), $c_i \geq 1$, en élevant au carré et en intégrant sur ∂O_m cette dernière inégalité, nous obtenons

$$\left\|u_i^{(k+1)}\right\|_{L^2(\partial O_m)}^2 \leq \frac{(N-1) S_i S_m c_i}{16\pi^2 d_{i,m}^4} \left(\sum_{j=1}^{i-1} a_{j,i}^{(k+1)} + \sum_{j=i+1}^{N} a_{j,i}^{(k)} \right). \tag{2.54}$$

En utilisant (2.47), nous pouvons répéter cette procédure avec $\partial_n u_i^{(k)}$ à la place de $u_i^{(k)}$. La seule différence dans les calculs concerne le terme $\frac{1}{d_{i,m}^2}$ qui est remplacé par $\frac{4}{d_{i,m}^3}$. Nous obtenons alors

$$\left\|\partial_n u_i^{(k+1)}\right\|_{L^2(\partial O_m)}^2 \leq \frac{(N-1) S_i S_m c_i}{\pi^2 d_{i,m}^6} \left(\sum_{j=1}^{i-1} a_{j,i}^{(k+1)} + \sum_{j=i+1}^{N} a_{j,i}^{(k)} \right). \tag{2.55}$$

En sommant (2.52) et (2.53), nous observons que (2.50) est encore vraie pour les objets de type Neumann. □

Nous obtenons alors un critère de convergence.

Théorème 2.4.9. *Définissons* $\gamma := \max_{i,m\in[\![1;N]\!],\ m\neq i} \kappa_{i,m}$, *avec* $\kappa_{i,m}$ *donné par* (2.51) *et supposons que*

$$N\gamma < 1, \tag{2.56}$$

alors la suite $(u_i^{(k)})_{i\in[\![1;N]\!],\ k\mathbb{N}}$ *converge dans* H.

Démonstration. Nous commençons par définir la suite $(\alpha_n)_{n\in\mathbb{N},n\geq 1}$ par $\alpha_{kN+i} = \max_{m\in[\![1;N]\!],\ m\neq i}\left(a_{j,m}^{(k)}\right)$. D'après (2.50), les termes de cette suite satisfont, pour $i > N$

$$\alpha_i \leq \gamma \sum_{j=1}^{N} \alpha_{i-j}.$$

D'après (2.56), ceci implique que

$$\max_{j\in[\![1;N]\!]}(\alpha_{i+j}) \leq N\gamma \max_{j\in[\![1;N]\!]}(\alpha_{i-N+j}),$$

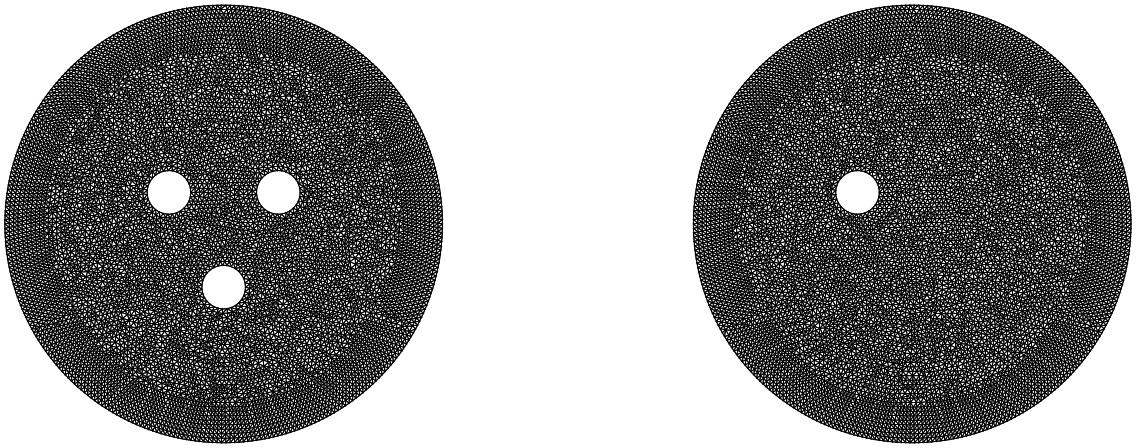
En appliquant encore une fois (2.56), nous pouvons conclure que $(\alpha_n)_{n\in\mathbb{N},n\geq 1}$ converge. Par conséquent, la série $\sum_k u_i^{(k)}$ converge sur les bords, ainsi elle converge aussi dans H. Sa limite satisfait (2.14). On conclut la preuve grâce à l'unicité de la solution. □

2.5 Test numérique

Ce premier test numérique concerne le taux de convergence de la méthode. Il est inspiré du contre-exemple de convergence de la méthode parallèle trouvé dans [28]. En dimension deux, nous considérons un domaine borné, formé d'un disque centré en l'origine et de rayon égal à 10. Ce disque contient lui même trois disques de rayon égal à 1, chacun d'eux étant centré sur l'un des sommets d'un triangle équilatéral centré en l'origine. Nous considérons le paramètre l représentant la longueur des côtés du triangle. Nous étudions ici, lorsque l varie, la convergence ou la divergence de différentes formes de la méthode, pour un problème de Laplace muni de conditions aux bords de Dirichlet.

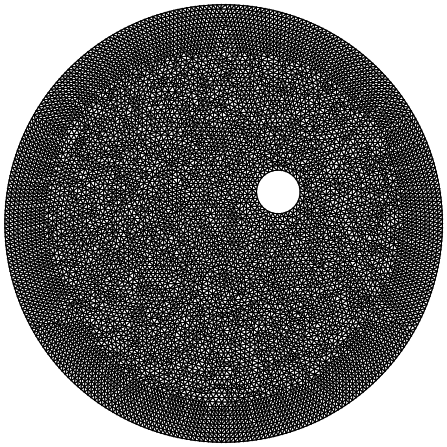
Nous avons utilisé Freefem++ [27] pour réaliser ces simulations. Tous les problèmes aux bords sont discrétisés par la méthode des éléments finis, à l'aide d'un maillage composé d'environ cinq éléments par unité de longueur. Ceci permet, entre autres, une mesure directe de l'erreur globale. On remarquera que ce type de méthode numérique n'est clairement pas le plus adapté à l'utilisation de la méthode des réflexions. Cela peut en effet être observé sur la figure 2.1, présentant les maillages respectivement utilisés pour le problème général et pour chacun des sous-problèmes correspondants. L'effort de calcul nécessaire pour résoudre les sous-problèmes est plus élevé que pour le problème initial, en raison du remplissage des disques. Ceci souligne le fait qu'en tant que méthode de décomposition de frontière, la méthode des réflexions est mieux adaptée aux méthodes de discrétisation fondées sur des formules de représentation intégrale de frontières, comme la méthode des éléments de frontières, voir 1.2.2.

Le tableau 2.1 indique que le nombre de cycles requis pour que l'approximation produite par les différentes versions de la méthode des réflexions s'approche de façon satisfaisante de la solution du problème complet et ceci pour différentes valeurs de la longueur l. Le critère d'arrêt ici choisi est l'erreur relative en norme H^1 entre deux approximations de la solution à un cycle d'écart qui doit être inférieur ou égal à 10^{-12}. Précisons que le cas où l est égale à 2 correspond au cas où les disques se touchent. Puisque nous voulons que les objets soient disjoints, nous partons du cas où l est égale à $2,1$ jusqu'au cas où l est égale à $10,1$ avec un pas constant de longueur $0,5$.

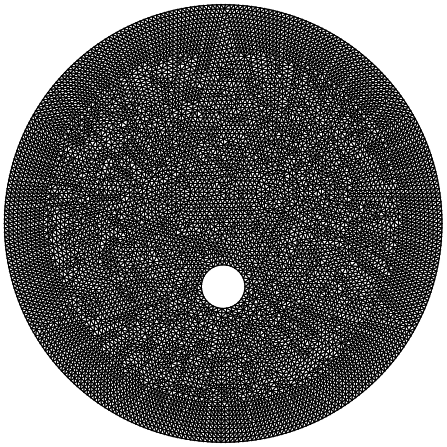


(a)

(b)



(c)



(d)

FIGURE 2.1 – Maillages pour le problème général (figure 2.1a) et les trois sous-problèmes liés (figures 2.1b, 2.1c et 2.1d), avec $l = 5$.

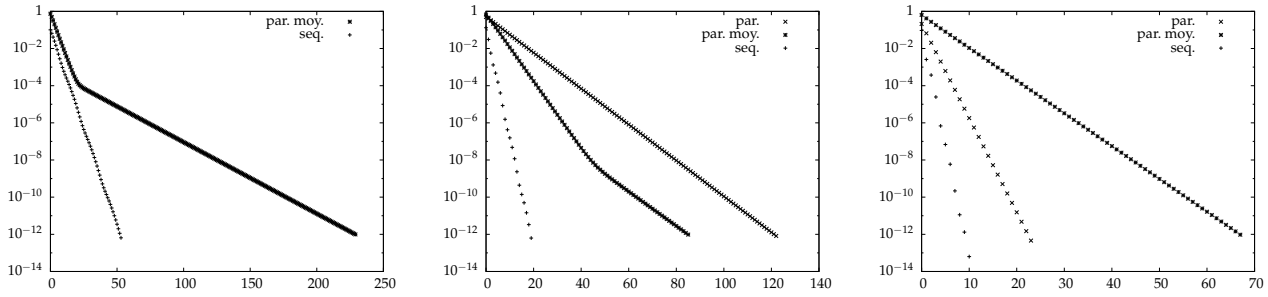
	2.1	2.6	3.1	3.6	4.1	4.6	5.1	5.6	6.1	6.6	7.1	7.6	8.1	8.6	9.1	9.6	10.1
par.	–	–	–	–	269	121	79	59	47	39	33	29	26	23	21	19	17
par. moy.	679	212	142	111	93	82	79	72	68	67	67	67	67	67	67	67	67
seq.	174	53	35	27	22	19	17	15	14	13	12	11	10	10	9	8	8

TABLE 2.1 – Nombre de cycles nécessaires, lorsque l varie, pour obtenir la convergence numérique des différentes formes de la méthode des réflexions. L'absence de résultat indique que la méthode a divergé.

La propriété de convergence inconditionnelle des versions séquentielle et parallèle moyennée est clairement observée, ainsi que le fait que la vitesse de convergence ralentie lorsque la distance entre les objets diminue. Pour la version parallèle de la méthode, la vitesse diminue de façon similaire, jusqu'au moment où elle diverge, ce qui est en accord avec le fait que cette version converge sous conditions.

Pour terminer, la figure 2.2 représente l'erreur relative de la méthode en fonction du nombre de cycles, pour trois valeurs différentes du paramètre l. Ce graphique confirme la convergence linéaire de la méthode. On peut observer de plus, que la version séquentielle est celle qui converge le plus rapidement. Remarquons que nous n'avons pas d'explication sur le brusque changement

de pente dans l'évolution de l'erreur pour la forme parallèle moyennée quand $l = 2,6$ ou $l = 4,6$. Néanmoins, ce défaut semble être lié à un problème de discrétisation, comme il apparaît plus tôt (respectivement plus tard) lorsque la taille des mailles augmente (respectivement diminue). Curieusement, les deux autres formes de la méthode ne sont pas affectées par ce phénomène.



(a) $l = 2.6$ (b) $l = 4.6$ (c) $l = 8.6$

FIGURE 2.2 – Erreur relative en norme H^1 de la méthode en fonction du nombre de cycles pour trois valeurs données de la longueur l.

Références

[1] C. AMROUCHE, V. GIRAULT, and J. GIROIRE. "Dirichlet and neumann exterior problems for the n-dimensional laplace operator an approach in weighted sobolev spaces". *J. Math. Pures et Appl.* 76.1 (1997), pp. 55 –81. DOI: 10.1016/S0021-7824(97)89945-X.

[2] N. ARONSZAJN. "Theory of reproducing kernels". *Trans. Amer. Math. Soc.* 68.3 (1950), pp. 337–404. DOI: 10.1090/S0002-9947-1950-0051437-7.

[3] A. AUSLENDER. *Optimisation : méthodes numériques*. Masson, 1976.

[4] C. BADEA, S. GRIVAUX, and V. MÜLLER. "The rate of convergence in the method of alternating projections". *St. Petersburg Math. J.* 23.3 (2012), pp. 413–434. DOI: 10.1090/S1061-0022-2012-01202-1.

[5] M. BALABANE. "Boundary decomposition for Helmholtz and Maxwell equations 1 : disjoint sub-scatterers". *Asymptotic Anal.* 38.1 (2004), pp. 1–10.

[6] H. H. BAUSCHKE and J. M. BORWEIN. "On projection algorithms for solving convex feasibility problems". *SIAM Rev.* 38.3 (1996), pp. 367–426. DOI: 10.1137/S0036144593251710.

[7] H. H. BAUSCHKE, F. DEUTSCH, and H. HUNDAL. "Characterizing arbitrarily slow convergence in the method of alternating projections". *Internat. Trans. Oper. Res.* 16.4 (2009), pp. 413–425. DOI: 10.1111/j.1475-3995.2008.00682.x.

[8] H. H. BAUSCHKE et al. "Accelerating the convergence of the method of alternating projections". *Trans. Amer. Math. Soc.* 355.9 (2003), pp. 3433–3461. DOI: 10.1090/S0002-9947-03-03136-2.

[9] P. E. BJØRSTAD and J. MANDEL. "On the spectra of sums of orthogonal projections with applications to parallel computing". *BIT* 31.1 (1991), pp. 76–88. DOI: 10.1007/BF01952785.

[10] F. BOYER et al. "Model for a sensor inspired by electric fish". *IEEE Trans. Robot.* 28.2 (2012), pp. 492–505. DOI: 10.1109/TRO.2011.2175764.

[11] M. CASSIER and C. HAZARD. "Multiple scattering of acoustic waves by small sound-soft obstacles in two dimensions : mathematical justification of the Foldy–Lax model". *Wave Motion* 50.1 (2013), pp. 18–28. DOI: 10.1016/j.wavemoti.2012.06.001.

[12] S. B. CHEN and H. J. KEH. "Electrophoresis in a dilute dispersion of colloidal spheres". *AIChE J.* 34.7 (1988), pp. 1075–1085. DOI: 10.1002/aic.690340704.

[13] G. CIMMINO. "Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari". *La Ricerca Scientifica* 16.9 (1938), pp. 326–333.

[14] R. DAUTRAY and J.-L. LIONS. *Mathematical analysis and numerical methods for science and technology. vol. 4. , integral equations and numerical methods*. Berlin, New York, Paris: Springer, 1990.

[15] F. DEUTSCH. "Rate of convergence of the method of alternating projections". *Parametric Optimization and Approximation*. Ed. by B. BROSOWSKI and F. DEUTSCH. Vol. 72. International series of numerical mathematics. Birkhäuser-Verlag, 1985, pp. 96–107.

[16] F. DEUTSCH. "The method of alternating orthogonal projections". *Approximation theory, spline functions and applications*. Ed. by S. P. SINGH. Vol. 356. NATO ASI Series. Springer Netherlands, 1992, pp. 105–121. DOI: 10.1007/978-94-011-2634-2_5.

[17] F. DEUTSCH and H. HUNDAL. "Slow convergence of sequences of linear operators II : arbitrarily slow convergence". *J. Approx. Theory* 162.9 (2010), pp. 1717–1738. DOI: 10.1016/j.jat.2010.05.002.

[18] F. DEUTSCH and H. HUNDAL. "The rate of convergence for the method of alternating projections, II". *J. Math. Anal. Appl.* 205.2 (1997), pp. 381–405. DOI: 10.1006/jmaa.1997.5202.

[19] J. K. G. DHONT. *An introduction to dynamics of colloids*. Vol. 2. Studies in interface science. Elsevier, 1996.

[20] J. DIXMIER. "Étude sur les variétés et les opérateurs de Julia, avec quelques applications". *Bull. Soc. Math. France* 77 (1949), pp. 11–101.

[21] L. C. EVANS. *Partial differential equations*. Graduate studies in mathematics. Providence (R.I.): American Mathematical Society, 1998.

[22] L. L. FOLDY. "The multiple scattering of waves. I. General theory of isotropic scattering by randomly distributed scatterers". *Phys. Rev.* 67.3-4 (1945), pp. 107–119. DOI: 10.1103/PhysRev.67.107.

[23] K. FRIEDRICHS. "On certain inequalities and characteristic value problems for analytic functions and for functions of two variables". *Trans. Amer. Math. Soc.* 41.3 (1937), pp. 321–364. DOI: 10.1090/S0002-9947-1937-1501907-0.

[24] M. J. GANDER. "Schwarz methods over the course of time". *Electron. Trans. Numer. Anal.* 31 (2008), pp. 228–255.

[25] I. HALPERIN. "The product of projection operators". *Acta Sci. Math. (Szeged)* 23.1-2 (1962), pp. 96–99.

[26] J. HAPPEL and H. BRENNER. *Low Reynolds number hydrodynamics with special applications to particulate media*. Vol. 1. Mechanics of fluids and transport processes. Martinus Nijhoff publishers, 1983. DOI: 10.1007/978-94-009-8352-6.

[27] F. HECHT. "New development in FreeFem++". *J. Numer. Math.* 20.3-4 (2012), pp. 251–265. DOI: 10.1515/jnum-2012-0013.

[28] K. ICHIKI and J. F. BRADY. "Many-body effects and matrix inversion in low-Reynolds-number hydrodynamics". *Phys. Fluids* 13.1 (2001), pp. 350–353. DOI: 10.1063/1.1331320.

[29] R. B. JONES. "Hydrodynamic interaction of two permeable spheres I : The method of reflections". *Phys. A* 92.3-4 (1978), pp. 545–556. DOI: 10.1016/0378-4371(78)90150-4.

[30] S. KAYALAR and H. L. WEINERT. "Error bounds for the method of alternating projections". *Math. Control Signals Systems* 1.1 (1988), pp. 43–59. DOI: 10.1007/BF02551235.

[31] S. KIM and S. J. KARRILA. *Microhydrodynamics : principles and selected applications*. Butterworth-Heinemann, 1991.

[32] G. J. KYNCH. "The slow motion of two or more spheres through a viscous fluid". *J. Fluid Mech.* 5.2 (1959), pp. 193–208. DOI: 10.1017/S0022112059000155.

[33] M. LAX. "Multiple scattering of waves". *Rev. Mod. Phys.* 23.4 (1951), pp. 287–310. DOI: 10.1103/RevModPhys.23.287.

[34] M. LAX. "Multiple scattering of waves. II. The effective field in dense systems". *Phys. Rev.* 85.4 (1952), pp. 621–629. DOI: 10.1103/PhysRev.85.621.

[35] P. L. LIONS. "On the Schwarz alternating method. I". *First international symposium on domain decomposition methods for partial differential equations*. Ed. by R. GLOWINSKI et al. SIAM, 1988, pp. 1–42.

[36] J. H. C. LUKE. "Convergence of a multiple reflection method for calculating Stokes flow in a suspension". *SIAM J. Appl. Math.* 49.6 (1989), pp. 1635–1651. DOI: 10.1137/0149099.

[37] C. METTOT and E. LAUGA. "Energetics of synchronized states in three-dimensional beating flagella". *Phys. Rev. E* 84.6 (2011), p. 061905. DOI: 10.1103/PhysRevE.84.061905.

[38] J. von NEUMANN. "On rings of operators. Reduction theory". *Ann. Math. (2)* 50.2 (1949), pp. 401–485. DOI: 10.2307/1969463.

[39] K. T. SMITH, D. C. SOLOMON, and S. L. WAGNER. "Practical and mathematical aspects of the problem of reconstructing objects from radiographs". *Bull. Amer. Math. Soc.* 83.6 (1977), pp. 1227–1270. DOI: 10.1090/S0002-9904-1977-14406-6.

[40] M. SMOLUCHOWSKI. "Über die Wechselwirkung von Kugeln, die sich in einer zähen Flüssigkeit bewegen". *Bull. Int. Acad. Sci. Cracovie, Cl. Sci. Math. Nat., Sér. A Sci. Math.* (1911), pp. 28–39.

[41] S. D. TRAYTAK. "Convergence of a reflection method for diffusion-controlled reactions on static sinks". *Phys. A Statist. Mech. Appl.* 362.2 (2006), pp. 240–248. DOI: 10.1016/j.physa.2005.03.061.

[42] H. WANG and J. LIU. "On decomposition method for acoustic wave scattering by multiple obstacles". *Acta Math. Sci.* 33.1 (2013), pp. 1–22. DOI: 10.1016/S0252-9602(12)60191-X.

[43] H. J. WILSON. "Stokes flow past three spheres". *J. Comput. Phys.* 245 (2013), pp. 302–316. DOI: 10.1016/j.jcp.2013.03.020.

[44] J. XU. "Iterative methods by space decomposition and subspace correction". *SIAM Rev.* 34.4 (1992), pp. 581–613. DOI: 10.1137/1034116.

[45] J. XU. "The method of subspace corrections". *J. Comput. Appl. Math.* 128.1-2 (2001), pp. 335–362. DOI: 10.1016/S0377-0427(00)00518-5.

[46] J. XU and L. ZIKATANOV. "The method of alternating projections and the method of subspace corrections in Hilbert space". *J. Amer. Math. Soc.* 15.3 (2002), pp. 573–597. DOI: 10.1090/S0894-0347-02-00398-3.

3

Bases réduites pour le problème direct en électrolocation

Sommaire

3.1 Introduction

Ce chapitre concerne les problématiques présentées en introduction dans la section 1.3. En vue de la résolution du problème d'électrolocation et de la commande du robot ANGELS, nous cherchons donc à concevoir un algorithme léger et rapide pour résoudre le problème direct. Plus précisément, notre approche est pensée pour être embarquée sur le robot et lui permettre de réagir en temps réel. Faisant suite à d'autres modèles présentés dans la section 1.3, nous proposons ici une alternative qui, une fois pleinement développée, devrait être plus performante. Notre démarche est décrite en section 3.2. Résumons-en d'ores-et-déjà les quatre principales étapes.

- Une transformation géométrique : l'inversion géométrique. Nous tirons de cette transformation deux avantages : on se ramène à une situation bornée et les objets non bornés, comme les murs, sont alors perçus bornés.

- Une méthode pour discrétiser : la méthode des éléments de frontière (BEM). En plus de la description faite dans ce chapitre, on donne dans l'introduction plus de détails sur les grandes lignes de cette méthode, voir sous-section 1.2.2. Notons que, dans un premier temps, nous avions résolu le problème par une méthode d'éléments finis (FEM). Cette démarche n'ayant pas été satisfaisante nous l'avons abandonnée. Pour autant, nous avons réalisé des tests de comparaison entre FEM et BEM. Ceux-ci sont présentés dans la section 3.3 et montrent l'intérêt de notre approche par BEM.

- Une méthode de réduction de modèle : la méthode des bases réduites par décomposition orthogonale propre (POD). Le problème de la BEM est qu'elle génère des matrices pleines de grande dimension. Le temps de constitution du système et de sa résolution est, au regard de contraintes de calcul en temps réel, prohibitif. La méthode de réduction que nous construisons permet une chute de la dimension des systèmes à résoudre.

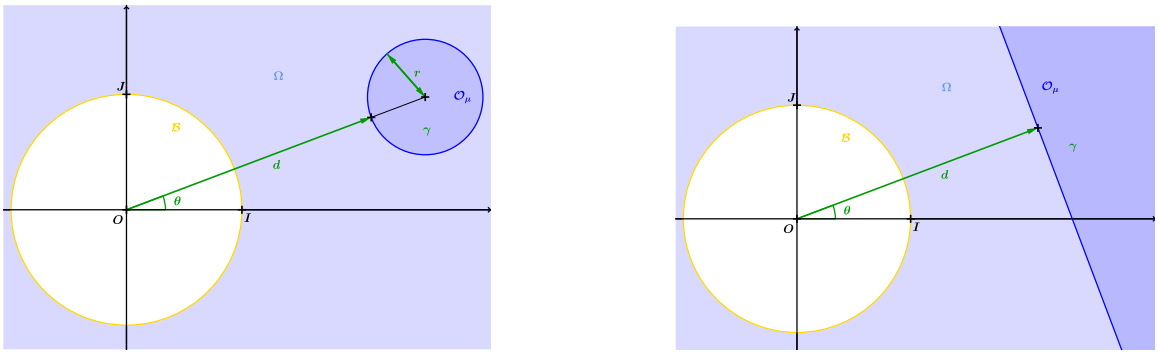
- En complément de la méthode des bases réduites : la méthode d'interpolation empirique (EIM). Certains termes du système ont une dépendance compliquée aux paramètres, ce qui pose des problèmes lors de la réduction. Il est alors indispensable de les interpoler.

Dans la section 3.3, cette démarche est testée numériquement. Enfin la section 3.4 présente des perspectives à donner à ce travail.

Présentons à présent le problème qui servira de fil conducteur au chapitre. Précisons que le problème est présenté ici sous une forme volontairement simplifiée. Les simplifications principales consistent à considérer une géométrie de capteur simple, en dimension deux, munie de conditions aux bords facilitant la démarche mathématique. Pour autant, l'ensemble de ce qui suit peut être aisément transposé à la dimension trois, pour des capteurs dont la géométrie et les conditions aux bords sont plus réalistes. Les différentes notations utilisées dans ce chapitre sont présentées dans la sous-section 1.2.1 ou introduites au fur et à mesure suivant les besoins. De plus, rappelons que nous n'indiquerons pas les symboles "+" et "-" sur les fonctions définies aux bords d'un ouvert, hormis pour celles qui ne sont pas continues à l'interface de l'objet.

On se donne donc un repère orthonormé (O, I, J) du plan. L'enveloppe extérieure du robot $\mathcal{B}$ considérée est ici délimitée par le cercle $\partial\mathcal{B}$ centré en l'origine O et de rayon 1. Ainsi, le domaine Ω, dit domaine extérieur, est le complémentaire dans $\mathbb{R}^2$ du disque unité $\overline{\mathcal{B}}$. Afin de définir une famille d'ouverts susceptibles de représenter l'objet perturbant la scène, nous définissons un ensemble $\mathscr{P}_\mu$ de vecteurs μ de paramètres le caractérisant, soit

$$\mathscr{P}_\mu = \left\{ (r, d, \theta, \gamma) \in \overline{\mathbb{R}}_+^* \times \mathbb{R}^3 \mid 0 < r \leq +\infty,\ 1 < d < +\infty,\ 0 \leq \theta \leq \frac{\pi}{2},\ 0 \leq \gamma < +\infty \right\}. \quad (3.1)$$



(a) Bord de l'objet matérialisé par un cercle, c'est-à-dire le cas où $r < +\infty$.

(b) Bord de l'objet matérialisé par une droite, c'est-à-dire le cas où $r = +\infty$.

FIGURE 3.1 – Représentation schématique de la géométrie du problème 3.1.3.

Ainsi pour $\mu \in \mathscr{P}_\mu$ fixé, l'objet $\mathcal{O}_\mu$ est l'ouvert de bord

$$\partial\mathcal{O}_\mu = \begin{cases} \left\{(x,y) \in \mathbb{R}^2 \mid (x-(d+r)\cos\theta)^2 + (y-(d+r)\sin\theta)^2 = r^2\right\} & \text{si } r < +\infty; \\ \left\{(x,y) \in \mathbb{R}^2 \mid -\cos\theta\, x - \sin\theta\, y + d = 0\right\} & \text{sinon,} \end{cases}$$

c'est-à-dire le disque de rayon r et de centre $((d+r)\cos\theta, (d+r)\sin\theta)$ ou la droite de normale unitaire $(-\cos\theta, -\sin\theta)$ et dont la distance minimale à l'origine est d. Comme la définition inclut le cas $r = +\infty$, qui correspond au cas de la droite, nous considérons une classe d'objets d'échelle très variée. Le paramètre angulaire θ variant dans l'intervalle $[0, \frac{\pi}{2}]$, nous ne traitons que les cercles-droites dont le centre est situé dans le premier quadrant. Les autres situations peuvent être prises en compte en exploitant les symétries géométriques et électriques du capteur. En effet, ce dernier dispose au minimum d'une symétrie géométrique et électrique gauche/droite.

Remarque 3.1.1. *Nous aurions pu de plus couvrir le cas des conteneurs, c'est-à-dire les disques contenant le capteur. L'ensemble de paramètres, dans ce cas plus général, aurait alors été*

$$\begin{aligned} &\left\{(r,d,\theta,\gamma) \in \overline{\mathbb{R}}_+^* \times \mathbb{R}^3 \mid 0 < r \leq +\infty,\ 1 < d < +\infty,\ -\pi < \theta \leq \pi,\ 0 \leq \gamma < +\infty\right\} \\ \cup\ &\left\{(r,d,\theta,\gamma) \in \mathbb{R}^4 \mid d < -1, |d| \leq r,\ -\pi < \theta \leq \pi,\ 0 \leq \gamma < +\infty\right\}. \end{aligned} \tag{3.2}$$

Remarque 3.1.2. *Précisons qu'il est commode en dimension deux d'utiliser la définition en termes de nombres complexes des cercles-droites, ce qui présente le double avantage de n'utiliser qu'une équation pour les représenter et simplifie les calculs pour l'inversion géométrique présentée en section 3.2.1. Comme ces calculs ne seront pas détaillés, nous rappelons rapidement la définition de $\partial\mathcal{O}_\mu$ que l'on obtient de cette façon :*

$$\forall \mu \in \mathscr{P}_\mu,\ \partial\mathcal{O}_\mu = \left\{z \in \mathbb{C} \mid \frac{1}{r} z\bar{z} + \left(\frac{d}{r} + 1\right)(\bar{\omega} z + \omega \bar{z}) + d\left(\frac{d}{r} + 2\right) = 0\right\}, \tag{3.3}$$

avec $\omega = -cos\theta - \imath \sin\theta$. Ainsi $\partial\mathcal{O}_\mu$ est soit le cercle de centre $-(d+r)\omega$ et de rayon r, soit la droite de vecteur normal unitaire ω et dont la distance minimale à l'origine est d.

Dans ce modèle et sans perte de généralité, la conductivité du milieu est égale à 1 et celle de l'objet est égale à γ. La figure 3.1 résume les éléments géométriques qui précèdent. Nous pouvons maintenant formuler le problème qui nous intéresse dans ce chapitre.

Problème 3.1.3 (Problème de Dirichlet extérieur avec un objet). *Pour $\mu \in \mathscr{P}_\mu$, déterminer la mesure $\frac{\partial u}{\partial \nu}\Big|_{\partial\mathcal{B}}$, où u est la solution du problème suivant*

$$\begin{cases} -\Delta u = 0, & \text{sur } \Omega \setminus \partial\mathcal{O}_\mu, \\ u = f, & \text{sur } \partial\mathcal{B}, \\ [u]_{\partial\mathcal{O}_\mu} = 0, & \text{sur } \partial\mathcal{O}_\mu, \\ \left[\gamma \frac{\partial u}{\partial \nu}\right]_{\partial\mathcal{O}_\mu} = 0, & \text{sur } \partial\mathcal{O}_\mu, \\ \lim_{R\to\infty} \sup_{|x|\leq R} |u(x)| < \infty, \end{cases} \tag{3.4}$$

avec pour $(x,y) \in \partial\mathcal{B}$, $f(x,y) := \frac{x}{\sqrt{x^2+y^2}}$. Rappelons que ν fait référence à la normale unitaire extérieure au bord du domaine $\Omega \setminus \overline{\mathcal{O}_\mu}$, c'est-à-dire dirigée vers l'intérieur du capteur et de l'objet.

3.2 Diverses transformations et méthodes mises en œuvre

L'objectif de cette section est de présenter les quatre éléments constitutifs de la démarche que l'on va suivre pour résoudre numériquement le problème 3.1.3. Les contraintes que l'on se fixe sont la rapidité de la résolution et la facilité de mise en œuvre sur un matériel informatique rudimentaire. Dans la première sous-section est proposée une première transformation : l'inversion géométrique. Cette transformation ramène notre problème à une situation bornée, où la prise en compte des objets est réalisée de la même manière et ceci de la plus petite à la plus grande échelle d'objets possible. Dans la deuxième sous-section est évoquée la discrétisation du problème. La méthode employée s'appelle la méthode de collocation, qui s'inscrit plus généralement dans les méthodes des éléments finis de frontière. Ces méthodes sont basées sur les représentations intégrales de frontière de notre problème. Ensuite dans la troisième sous-section, nous introduisons la résolution de notre problème discrétisé par la méthode de bases réduites basée sur une décomposition orthogonale propre. Cette méthode a pour but de diminuer la complexité de notre problème mettant à profit la paramétrisation définissant les objets. Dans notre cas, cette dernière méthode doit être compléteé par une méthode d'interpolation présentée dans la dernière sous-section.

3.2.1 Inversion géométrique

Commençons par définir l'inversion géométrique dans le plan.

Définition 3.2.1 (Inversion géométrique, ou transformation dite de Kelvin). *Pour tout point M de $\mathbb{R}^2$ distinct de l'origine O, il existe un unique point M' de $\mathbb{R}^2$ tel que :*

- *les points O, M et M' soient alignés ;*
- *$OM \times OM' = 1$.*

On peut ainsi définir l'inversion de centre O comme étant l'application de $\mathbb{R}^2 \setminus \{O\}$ dans lui-même qui, à un point M, associe l'unique point M' correspondant aux caractéristiques précédentes.

Rappelons que cette transformation

- est conforme ; c'est-à-dire que les angles sont conservés (mais pas nécessairement leurs orientations) ce qui préserve l'harmonicité ;
- laisse stable l'ensemble des hypersphères-hyperplans.

Pour traiter numériquement (3.4), nous commençons par inverser géométriquement le problème 3.1.3. Le fait que l'inversion conserve l'harmonicité entraîne que nous serons amenés à traiter le même type de problème après inversion géométrique. La stabilité des hypersphères-hyperplans induit que les objets conserveront leurs propriétés géométriques. L'intégralité de la démarche que nous présentons ici peut être appliquée en dimension trois. La plupart des résultats de cette sous-section sont classiques. Pour l'inversion de l'équation de Laplace se référer à [1] (Sous-section 7.1.2, page 310).

Pour pouvoir utiliser pleinement l'inversion géométrique, il faut étendre la définition précédente pour pouvoir inclure l'inversion de l'origine O. On introduit pour cela la notation suivante

$$\forall \alpha \in]-\pi;\pi],\ \infty := \lim_{R\to+\infty} (R\cos(\alpha), R\sin(\alpha)) \quad \text{et} \quad \overline{\mathbb{R}^2} := \mathbb{R}^2 \cup \{\infty\}.$$

La définition de l'infini n'est donc pas dépendant de la direction. Présentons maintenant l'inversion géométrique complétée sous une forme explicite prolongeant la définition 3.2.1.

Définition 3.2.2. *L'inversion est l'application* $\mathscr{I} : \overline{\mathbb{R}^2} \to \overline{\mathbb{R}^2}$ *définie par*

$$\forall\, (x,y) \in \mathbb{R}^2,\ \mathscr{I}\,(x,y) = (\xi,\eta) = \frac{1}{R^2}\,(x,y)\,,$$

où $R = \sqrt{x^2+y^2}$, *et prolongée par* $\mathscr{I}\,(\infty) = O$ *et* $\mathscr{I}\,(O) = \infty$.

Propriétés 3.2.3. *L'inversion* $\mathscr{I}$ *présentée dans la définition 3.2.2, vérifie les propriétés suivantes.*

1. *Elle est bijective.*
2. *C'est une involution.*
3. *On appelle cercle d'inversion, le cercle* $\mathscr{C}\,(O,1)$ *(cercle unité centré en l'origine). Il est invariant point par point.*
4. *Les droites passant par* O *sont des invariants globaux.*
5. *Plus généralement, l'ensemble des cercles-droites est stable par inversion.*

La dernière propriété indique que l'image, par l'inversion, d'un cercle ou d'une droite, est un cercle ou une droite. Mais il est parfaitement envisageable que l'image d'un cercle (respectivement d'une droite) soit une droite (respectivement une cercle). Voici par conséquent un descriptif des situations possibles :

- L'image d'une droite passant par O est elle-même.
- L'image d'une droite ne passant pas par O est un cercle passant par O.
- L'image d'un cercle passant par O est une droite ne passant pas par O.
- L'image d'un cercle ne passant pas par O est un cercle ne passant pas par O.

Nous allons maintenant expliciter les modifications induites par la transformation $\mathscr{I}$ sur le problème 3.1.3. En utilisant la définition et les propriétés 3.2.3, on montre facilement que

- $\mathscr{I}\,(\Omega) = \mathcal{B}$, car nous avons fait correspondre le bord du capteur avec le cercle $\mathscr{C}\,(O,1)$;
- $\mathscr{I}\,(\partial\mathcal{B}) = \partial\mathcal{B}$, où l'égalité est ici entendue point par point.

Décrivons l'effet de l'inversion sur l'objet. Pour faire suite à la remarque 3.1.1, où l'on présentait l'ensemble de paramètres (3.2) le plus général, l'image par $\mathscr{I}$ de tous ces ouverts sont les disques strictement inclus dans le disque unité centré en O. Comme notre ensemble de paramètres $\mathscr{P}_\mu$ est plus petit, nous devons éliminer un certain nombre de disques. Comme les centres de l'ouvert et du disque image doivent être alignés, le centre du disque image est nécessairement situé dans le premier quadrant. On en conclut aussi que le paramètre angulaire θ ne change pas. Comme nous ne prenons pas en compte les conteneurs, aucun des disques images ne doit contenir l'origine ou uniquement sur son bord. Ainsi l'ensemble des paramètres $\mu' = (r', d', \theta', \gamma')$, noté $\mathscr{P}_{\mu'}$ qui correspond au système inversé, peut être défini de la manière suivante

$$\mathscr{P}_{\mu'} = \left\{ (r', d', \theta', \gamma') \in \mathbb{R}^4 \mid 0 < r' < 0,5,\ 0 \leq d' < 1 - 2r',\ 0 \leq \theta' \leq \frac{\pi}{2},\ 0 \leq \gamma' < +\infty \right\}. \quad (3.5)$$

Les paramètres géométriques r', d' et θ' sont reliés à r, d et θ de la manière suivante

$$r' = \frac{1}{d\,(d/r + 2)}, \quad d' = \frac{1}{r(d/r + 2)} \quad \text{et } \theta' = \theta,$$

ce qui nous permet de définir l'ouvert $\mathcal{O}_{\mu'}$, pour un paramètre μ' donné, délimité par le cercle $\partial\mathcal{O}_{\mu'}$ de la manière suivante

$$\forall \mu' \in \mathscr{P}_{\mu'}, \quad \partial\mathcal{O}_{\mu'} = \left\{ (\xi, \eta) \in \mathbb{R}^2 \mid (\xi - (d' + r')\cos\theta)^2 + (\eta - (d' + r')\sin\theta)^2 = r'^2 \right\}.$$

Remarque 3.2.4. *Ces résultats et leurs généralisations au cas étendu (3.2) se retrouvent aisément en utilisant la définition complexe de l'application $\mathscr{I}$ et les équations correspondantes des cercles-droites.*

Pour ce qui concerne la dimension trois, il suffit de considérer le plan passant par le centre des deux sphères, ce qui nous ramène à la situation précédente. Ainsi ces formules restent également valables dans ce contexte.

Nous appliquons maintenant l'inversion aux fonctions définies sur Ω et aux équations du problème 3.1.3. Une fonction u définie sur Ω est représentée par une fonction $\tilde{u}$ définie sur $\mathcal{B}$ par

$$\forall (x, y) \in \Omega, \quad \tilde{u}(\xi, \eta) = u(x, y), \ \text{ où } (\xi, \eta) = \mathscr{I}(x, y).$$

Comme l'inversion est une application conforme, le caractère harmonique d'une fonction est conservé, si bien que l'on peut montrer que

$$\Delta\tilde{u}(\xi, \eta) = R^4 \Delta u(x, y) = 0,$$

mais aussi

$$\forall (\xi, \eta) \in \partial\mathcal{B}, \quad \tilde{u}(\xi, \eta) = u(\xi, \eta) = f(\xi, \eta),$$

$$\forall (\xi, \eta) \in \partial\mathcal{B} \cup \partial\mathcal{O}_{\mu'}, \quad \frac{\partial\tilde{u}}{\partial\nu'^-}(\xi, \eta) = -R^2 \frac{\partial u}{\partial\nu^-}(x, y)$$

$$\forall (\xi, \eta) \in \partial\mathcal{O}_{\mu'}, \quad \tilde{u}(\xi, \eta) = u\left(\mathscr{I}^{-1}(\xi, \eta)\right) \quad \text{et} \quad \frac{\partial\tilde{u}}{\partial\nu'^+}(\xi, \eta) = -R^2 \frac{\partial u}{\partial\nu^+}(x, y),$$

si l'on pose $(\xi, \eta) = \mathscr{I}(x, y)$ et $R = \sqrt{x^2 + y^2}$.

En utilisant les notations introduites dans la sous-section 1.2.1, pour ce qui concerne les sauts de $\tilde{u}$ et de $\frac{\partial\tilde{u}}{\partial\nu'}$, nous avons

$$[\tilde{u}]_{\partial\mathcal{O}_{\mu'}} = (\tilde{u}_- - \tilde{u}_+)|_{\partial\mathcal{O}_{\mu'}} \quad \text{et} \quad \left[\gamma \frac{\partial\tilde{u}}{\partial\nu'}\right]_{\partial\mathcal{O}_{\mu'}} = \left(\frac{\partial\tilde{u}}{\partial\nu'^-} - \gamma\frac{\partial\tilde{u}}{\partial\nu'^+}\right)\Big|_{\partial\mathcal{O}_{\mu'}}$$

et par conséquent

$$[\tilde{u}]_{\partial\mathcal{O}_{\mu'}} = 0 \quad \text{et} \quad \left[\gamma\frac{\partial\tilde{u}}{\partial\nu'}\right]_{\partial\mathcal{O}_{\mu'}} = \left[\gamma\frac{\partial u}{\partial\nu}\right]_{\partial\mathcal{O}_{\mu}} = 0.$$

En posant $\gamma' = \gamma$, on retrouve les même conditions de transmission pour l'objet, à un facteur près, dans les domaines inversés ou non géométriquement.

On peut alors considérer le problème suivant résultant de l'inversion géométrique du problème 3.1.3.

Problème 3.2.5 (Problème de Dirichlet intérieur avec un objet). *On cherche pour tout $\mu' \in \mathscr{P}_{\mu'}$ à déterminer la mesure $\frac{\partial\tilde{u}}{\partial\nu'}\Big|_{\partial\mathcal{B}}$, où $\tilde{u}$ est la solution du problème suivant*

$$\begin{cases} -\Delta\tilde{u} = 0, & \text{sur } \mathcal{B}\setminus\partial\mathcal{O}_{\mu'}, \\ \tilde{u} = u = f, & \text{sur } \partial\mathcal{B}, \\ [\tilde{u}]_{\partial\mathcal{O}_{\mu'}} = 0, & \text{sur } \partial\mathcal{O}_{\mu'}, \\ \left[\gamma\frac{\partial\tilde{u}}{\partial\nu'}\right]_{\partial\mathcal{O}_{\mu'}} = 0, & \text{sur } \partial\mathcal{O}_{\mu'}. \end{cases} \tag{3.6}$$

Remarque 3.2.6. *En utilisant le même type de résultat que précédemment, on montre que*

$$\frac{\partial\tilde{u}}{\partial\nu'}\Big|_{\partial\mathcal{B}}(\xi,\eta) = -R^2\,\frac{\partial u}{\partial\nu}\Big|_{\partial\mathcal{B}}(x,y).$$

Or comme pour tout $(x,y) \in \partial\mathcal{B}$, $R = 1$, on en déduit que

$$\frac{\partial u}{\partial\nu}\Big|_{\partial\mathcal{B}}(x,y) = -\frac{\partial\tilde{u}}{\partial\nu'}\Big|_{\partial\mathcal{B}}(\xi,\eta)$$

Ainsi, pour un vecteur de paramètre μ donné et son vecteur μ' associé, les problèmes 3.1.3 et 3.2.5 donnent lieu à la même solution au signe près.

3.2.2 Formulation par intégrales de frontière et méthode de colocation

Dans nos recherches, le recours à une inversion géométrique a d'abord été motivé par la volonté de se ramener à un domaine borné. Dans ce cadre, nous disposons d'outils pratiques et efficaces tel Freefem++ par exemple permettant la résolution des problèmes variationnels par FEM. Dans ce contexte, la solution $\frac{\partial\tilde{u}}{\partial\nu'}$ du problème (3.2.5) n'est cependant pas directement accessible. En effet, il est nécessaire de résoudre un premier système linéaire pour déterminer $\tilde{u}$ dans tout le domaine d'étude, puis d'en résoudre un second pour obtenir notre solution $\frac{\partial\tilde{u}}{\partial\nu'}$ sur le bord du domaine. L'application de la procédure détaillée dans ce chapitre ne semblant pas concluante lorsqu'elle est traitée par FEM, nous nous sommes rabattus sur la BEM. Nous présenterons tout de même certains résultats numériques obtenus par FEM dans la section 3.3 pour illustrer le bénéfice de ce changement d'approche.

Comme nous avons pu le voir dans la sous-section 1.2.3 de l'introduction, la mise en œuvre numérique de la BEM repose sur une formulation sous forme d'intégrales de frontière. Pour établir une telle représentation dans notre cas, il suffit d'en établir une dans $\mathcal{B}\setminus\overline{\mathcal{O}_{\mu'}}$ et une dans $\mathcal{O}_{\mu'}$, puis de les recoller. Comme le bord de l'ouvert $\mathcal{B}\setminus\overline{\mathcal{O}_{\mu'}}$ est $\partial\mathcal{B}\cup\partial\mathcal{O}_{\mu'}$, on a

$$\int_{\partial\mathcal{B}\cup\partial\mathcal{O}_{\mu'}}\left(\frac{\partial G}{\partial\nu'}(x,y)\tilde{u}(y) - G(x,y)\frac{\partial\tilde{u}}{\partial\nu'}(y)\right)\mathrm{d}S_y = \begin{cases} \tilde{u}(x) \text{ si } x \in \mathcal{B}\setminus\overline{\mathcal{O}_{\mu'}}, \\ \frac{1}{2}\tilde{u}(x) \text{ si } x \in \partial\mathcal{B}\cup\mathcal{O}_{\mu'}, \\ 0 \text{ si } x \notin \overline{\mathcal{B}\setminus\overline{\mathcal{O}_{\mu'}}}, \end{cases}$$

ce qui nous permet d'écrire

$$\int_{\partial\mathcal{B}}\left(\frac{\partial G}{\partial\nu'}(x,y)\tilde{u}(y)-G(x,y)\frac{\partial\tilde{u}}{\partial\nu'}(y)\right)\,\mathrm{d}S_y+\int_{\partial\mathcal{O}_{\mu'}}\left(\frac{\partial G}{\partial\nu'}(x,y)\tilde{u}(y)-G(x,y)\frac{\partial\tilde{u}}{\partial\nu'^-}(y)\right)\,\mathrm{d}S_y=\begin{cases}\tilde{u}(x) & \text{si } x\in\mathcal{B}\setminus\overline{\mathcal{O}_{\mu'}},\\ \frac{1}{2}\tilde{u}(x) & \text{si } x\in\partial\mathcal{B}\cup\partial\mathcal{O}_{\mu'},\\ 0 & \text{si } x\notin\overline{\mathcal{B}\setminus\overline{\mathcal{O}_{\mu'}}},\end{cases}\tag{3.7}$$

où G est la fonction de Green associée au problème sur $\mathcal{B}\setminus\overline{\mathcal{O}_{\mu'}}$.

Nous procédons de la même manière pour l'ouvert $\mathcal{O}_{\mu'}$. Le signe change à cause de l'orientation opposée de la normale. Nous obtenons :

$$\int_{\partial\mathcal{O}_{\mu'}}\left(G(x,y)\frac{\partial\tilde{u}}{\partial\nu'^+}(y)-\frac{\partial G}{\partial\nu'}(x,y)\tilde{u}(y)\right)\,\mathrm{d}S_y=\begin{cases}\tilde{u}(x) & \text{si } x\in\mathcal{O}_{\mu'},\\ \frac{1}{2}\tilde{u}(x) & \text{si } x\in\partial\mathcal{O}_{\mu'},\\ 0 & \text{si } x\notin\overline{\mathcal{O}_{\mu'}},\end{cases}\tag{3.8}$$

En effectuant l'opération (3.7) $+\,\gamma$(3.8) et en considérant les conditions aux bords du problème (3.2.5)

$$[\tilde{u}]_{\partial\mathcal{O}_{\mu'}}=0\quad\text{et}\quad\left[\gamma\frac{\partial\tilde{u}}{\partial\nu'}\right]_{\partial\mathcal{O}_{\mu'}}=0,$$

on obtient alors le système suivant

$$\int_{\partial\mathcal{B}}\left(\frac{\partial G}{\partial\nu'}(x,y)\tilde{u}(y)-G(x,y)\frac{\partial\tilde{u}}{\partial\nu'}(y)\right)\,\mathrm{d}S_y+(1-\gamma)\int_{\partial\mathcal{O}_{\mu'}}\frac{\partial G}{\partial\nu'}(x,y)\tilde{u}(y)\,\mathrm{d}S_y=\begin{cases}\frac{1}{2}\tilde{u}(x) & \text{si } x\in\partial\mathcal{B},\\ \frac{1+\gamma}{2}\tilde{u}(x) & \text{si } x\in\partial\mathcal{O}_{\mu'}.\end{cases}\tag{3.9}$$

Pour simplifier la compréhension et le passage au système discrétisé, nous introduisons quelques notations et conventions. Nous réutilisons les notations $\mathcal{S}_\Omega$ et $\mathcal{D}_\Omega$ introduites dans la sous-section 1.2.3 respectivement pour les opérateurs de simple et double couches. Dans notre cadre, nous indiquons que l'intégration a lieu sur le bord d'un ouvert Ω_1 et que l'évaluation a lieu sur le bord d'un ouvert Ω_2 par les notations $\mathcal{S}_{\Omega_1,\Omega_2}$ et $\mathcal{D}_{\Omega_1,\Omega_2}$. Dans le cas où $\Omega_1=\Omega_2$, nous notons comme précédemment $\mathcal{S}_{\Omega_1}$ et $\mathcal{D}_{\Omega_1}$. Par exemple, l'opérateur noté $\mathcal{D}_{\mathcal{B},\mathcal{O}_{\mu'}}$ est défini par

$$\forall x\in\partial\mathcal{O}_{\mu'},\quad\mathcal{D}_{\mathcal{B},\mathcal{O}_{\mu'}}(v_1)(x):=\int_{\partial\mathcal{B}}\frac{\partial G}{\partial\nu'}(x,y)v_1(y)\,\mathrm{d}S_y,$$

tandis que l'opérateur $\mathcal{D}_{\mathcal{O}_{\mu'}}$ est quant à lui défini par

$$\forall x\in\partial\mathcal{O}_{\mu'},\quad\mathcal{D}_{\mathcal{O}_{\mu'}}(v_2)(x):=\int_{\partial\mathcal{O}_{\mu'}}\frac{\partial G}{\partial\nu'}(x,y)v_2(y)\,\mathrm{d}S_y,\tag{3.10}$$

où les fonctions v_1 et v_2 sont des fonctions admissibles respectivement des bords $\partial\mathcal{B}$ et $\partial\mathcal{O}_{\mu'}$. On peut alors réécrire le système (3.9) de la manière suivante

$$\begin{cases}\mathcal{D}_{\mathcal{B}}\left(\tilde{u}|_{\partial\mathcal{B}}\right)(x)-\mathcal{S}_{\mathcal{B}}\left(\left.\frac{\partial\tilde{u}}{\partial\nu'}\right|_{\partial\mathcal{B}}\right)(x)+(1-\gamma)\mathcal{D}_{\mathcal{O}_{\mu'},\mathcal{B}}\left(\tilde{u}|_{\partial\mathcal{O}_{\mu'}}\right)(x)=\frac{1}{2}\,\tilde{u}|_{\partial\mathcal{B}}(x) & \text{si } x\in\partial\mathcal{B},\\ \mathcal{D}_{\mathcal{B},\mathcal{O}_{\mu'}}\left(\tilde{u}|_{\partial\mathcal{B}}\right)(x)-\mathcal{S}_{\mathcal{B},\mathcal{O}_{\mu'}}\left(\left.\frac{\partial\tilde{u}}{\partial\nu'}\right|_{\partial\mathcal{B}}\right)(x)+(1-\gamma)\mathcal{D}_{\mathcal{O}_{\mu'}}\left(\tilde{u}|_{\partial\mathcal{O}_{\mu'}}\right)(x)=\frac{1+\gamma}{2}\,\tilde{u}|_{\partial\mathcal{O}_{\mu'}}(x) & \text{si } x\in\partial\mathcal{O}_{\mu'}.\end{cases}$$

En regroupant les inconnues et les données du problème (3.2.5) respectivement dans le membre de gauche et dans celui de droite, nous obtenons :

$$\begin{cases} \mathcal{S}_{\mathcal{B}}\left(\frac{\partial\tilde{u}}{\partial\nu'}\Big|_{\partial\mathcal{B}}\right)(x) + (\gamma-1)\mathcal{D}_{\mathcal{O}_{\mu'},\mathcal{B}}\left(\tilde{u}|_{\partial\mathcal{O}_{\mu'}}\right)(x) = \left(\mathcal{D}_{\mathcal{B}} - \frac{1}{2}\mathcal{I}_{\mathcal{B}}\right)(\tilde{u}|_{\partial\mathcal{B}})(x) & \text{si } x \in \partial\mathcal{B}, \\ \mathcal{S}_{\mathcal{B},\mathcal{O}_{\mu'}}\left(\frac{\partial\tilde{u}}{\partial\nu'}\Big|_{\partial\mathcal{B}}\right)(x) + \left((\gamma-1)\mathcal{D}_{\mathcal{O}_{\mu'}} + \frac{1+\gamma}{2}\mathcal{I}_{\mathcal{O}_{\mu'}}\right)\left(\tilde{u}|_{\partial\mathcal{O}_{\mu'}}\right)(x) = \mathcal{D}_{\mathcal{B},\mathcal{O}_{\mu'}}(\tilde{u}|_{\partial\mathcal{B}})(x) & \text{si } x \in \partial\mathcal{O}_{\mu'}, \end{cases}$$

où les opérateurs $\mathcal{I}_{\mathcal{B}}$ et $\mathcal{I}_{\mathcal{O}_{\mu'}}$ désignent l'opérateur identité agissant respectivement sur les fonctions de $\partial\mathcal{B}$ et $\partial\mathcal{O}_{\mu'}$. Sous forme matricielle, ces équations s'écrivent

$$\begin{pmatrix} \mathcal{S}_{\mathcal{B}} & (\gamma-1)\mathcal{D}_{\mathcal{O}_{\mu'},\mathcal{B}} \\ \mathcal{S}_{\mathcal{B},\mathcal{O}_{\mu'}} & (\gamma-1)\mathcal{D}_{\mathcal{O}_{\mu'}} + \frac{1+\gamma}{2}\mathcal{I}_{\mathcal{O}_{\mu'}} \end{pmatrix} \begin{pmatrix} \frac{\partial\tilde{u}}{\partial\nu'}\Big|_{\partial\mathcal{B}} \\ \tilde{u}|_{\partial\mathcal{O}_{\mu'}} \end{pmatrix} = \begin{pmatrix} \mathcal{D}_{\mathcal{B}} - \frac{1}{2}\mathcal{I}_{\mathcal{B}} \\ \mathcal{D}_{\mathcal{B},\mathcal{O}_{\mu'}} \end{pmatrix} \tilde{u}|_{\partial\mathcal{B}}. \tag{3.11}$$

Pour résoudre le problème (3.2.5) quel que soit le paramètre, il ne nous reste plus qu'à appliquer la BEM, méthode présentée dans ses grandes lignes dans la sous-section 1.2.2 de l'introduction.

Soit $\mathrm{N}_{\mathcal{B}}$ et $\mathrm{N}_{\mathcal{O}_{\mu'}}$ le nombre de nœuds considérés pour discrétiser respectivement $\partial\mathcal{B}$ et $\partial\mathcal{O}_{\mu'}$. Commençons par définir les données géométriques discrétisées du capteur. On note

$$\forall k \in [\![1;\mathrm{N}_{\mathcal{B}}]\!], \quad \alpha_k := -\pi + (k-1)\tfrac{2\pi}{\mathrm{N}_{\mathcal{B}}-1}, \qquad \beta_k := \begin{cases} \frac{\alpha_k+\alpha_{k+1}}{2} & \text{si } k \neq \mathrm{N}_{\mathcal{B}}, \\ \frac{\alpha_k+\alpha_1}{2} & \text{sinon}, \end{cases}$$

$$P_k := (\cos(\alpha_k), \sin(\alpha_k)) \in \partial\mathcal{B}, \qquad M_k := \begin{cases} \frac{P_k+P_{k+1}}{2} & \text{si } k \neq \mathrm{N}_{\mathcal{B}}, \\ \frac{P_k+P_1}{2} & \text{sinon}, \end{cases}$$

$$E_k := \begin{cases} [P_k; P_{k+1}] & \text{si } k \neq \mathrm{N}_{\mathcal{B}}, \\ [P_k; P_1] & \text{sinon}, \end{cases}$$

où les $(E_k)_{k\in[\![1;\mathrm{N}_{\mathcal{B}}]\!]}$ et les $(P_k)_{k\in[\![1;\mathrm{N}_{\mathcal{B}}]\!]}$ représentent respectivement les éléments et les nœuds du maillage de $\partial\mathcal{B}$, où les $(M_k)_{k\in[\![1;\mathrm{N}_{\mathcal{B}}]\!]}$ correspondent aux milieux des éléments et où les angles $(\alpha_k)_{k\in[\![1;\mathrm{N}_{\mathcal{B}}]\!]}$ et $(\beta_k)_{k\in[\![1;\mathrm{N}_{\mathcal{B}}]\!]}$ permettent de construire respectivement les familles de points $(P_k)_{k\in[\![1;\mathrm{N}_{\mathcal{B}}]\!]}$ et $(M_k)_{k\in[\![1;\mathrm{N}_{\mathcal{B}}]\!]}$. On définit les mêmes types de données pour l'objet

$$\forall k \in [\![1;\mathrm{N}_{\mathcal{O}_{\mu'}}]\!], \quad \alpha'_k := \pi - (k-1)\tfrac{2\pi}{\mathrm{N}_{\mathcal{O}_{\mu'}}-1},$$
$$P'_k := \left((d'+r')\cos(\theta) + r'\cos(\alpha'_k), (d'+r')\sin(\theta) + r'\sin(\alpha'_k)\right) \in \partial\mathcal{O}_{\mu'},$$

où les $(E'_k)_{k\in[\![1;\mathrm{N}_{\mathcal{O}_{\mu'}}]\!]}$, les $(M'_k)_{k\in[\![1;\mathrm{N}_{\mathcal{O}_{\mu'}}]\!]}$ et les $(\beta'_k)_{k\in[\![1;\mathrm{N}_{\mathcal{O}_{\mu'}}]\!]}$ sont construites de manière analogue. On peut remarquer que l'intervalle $]-\pi;\pi]$ n'est pas discrétisé dans le même sens afin de construire le maillage de $\partial\mathcal{B}$ et de $\partial\mathcal{O}_{\mu'}$. Cet ordre prend en compte l'orientation des normales.

Passons à présent à la discrétisation de l'équation (3.11). Commençons par la discrétisation des fonctions. On introduit

$$\forall k \in [\![1;\mathrm{N}_{\mathcal{B}}]\!], \quad \left(\widetilde{U}_{\mathcal{B}}\right)_k := f(\cos(\beta_k), \sin(\beta_k)) = \cos(\beta_k) \text{ et } \left(\partial\widetilde{U}_{\mathcal{B}}\right)_k := \frac{\partial\tilde{u}}{\partial\nu'}\Big|_{\partial\mathcal{B}}(M_k)$$

$$\forall k \in [\![1;\mathrm{N}_{\mathcal{O}_{\mu'}}]\!], \quad \left(\widetilde{U}_{\mathcal{O}_{\mu'}}\right)_k := \tilde{u}(M'_k),$$

où la fonction f représente la condition au bord du capteur imposée dans le problème 3.2.5. La discrétisation des opérateurs est réalisée via les matrices

$$\forall (l,k) \in [\![1;\mathrm{N}_{\mathcal{B}}]\!]^2, \quad (\mathrm{S}_{\mathcal{B}})_{l,k} := \int_{y\in E_k} G(M_l, y)\,\mathrm{d}S_y \text{ et } (\mathrm{D}_{\mathcal{B}})_{l,k} := \int_{y\in E_k} \frac{\partial G}{\partial\nu'}(M_l, y)\,\mathrm{d}S_y,$$

$$\forall (l,k) \in [\![1;\mathrm{N}_{\mathcal{O}_{\mu'}}]\!] \times [\![1;\mathrm{N}_{\mathcal{B}}]\!],$$

$$\left(\mathrm{S}_{\mathcal{B},\mathcal{O}_{\mu'}}\right)_{l,k} := \int_{y\in E_k} G\left(M'_l,y\right)\mathrm{d}S_y \text{ et } \left(\mathrm{D}_{\mathcal{B},\mathcal{O}_{\mu'}}\right)_{l,k} := \int_{y\in E_k} \frac{\partial G}{\partial \nu'}\left(M'_l,y\right)\mathrm{d}S_y,$$

$$\forall (l,k) \in [\![1;\mathrm{N}_{\mathcal{B}}]\!] \times [\![1;\mathrm{N}_{\mathcal{O}_{\mu'}}]\!], \quad \left(\mathrm{D}_{\mathcal{O}_{\mu'},\mathcal{B}}\right)_{l,k} := \int_{y\in E'_k} \frac{\partial G}{\partial \nu'}\left(M_l,y\right)\mathrm{d}S_y.$$

Arrêtons-nous sur la discrétisation de l'opérateur $\mathcal{D}_{\mathcal{O}_{\mu'}}$, dont la définition est rappelée en (3.10). Il est simple de montrer par changement de variables que

$$\forall \mu' \in \mathscr{P}_{\mu'},\ \forall x \in \partial\mathcal{O}_{\mu'}, \quad \mathcal{D}_{\mathcal{O}_{\mu'}}(v)(x) = \int_{y\in\mathscr{C}(O,1)} \frac{\partial G}{\partial \nu'}\left(x,y\right) v(y)\,\mathrm{d}S_y,$$

où $\mathscr{C}(O,1)$ est le cercle de centre O et de rayon 1 et la fonction v est une fonction admissible de $\mathscr{C}(O,1)$. Nous avons noté le bord d'intégration $\mathscr{C}(O,1)$ et non $\partial\mathcal{B}$ pour indiquer que la normale considérée est rentrante dans le domaine obtenu après inversion. On définit alors

$$\forall \mu' \in \mathscr{P}_{\mu'}, \quad \mathcal{D}_{\mathcal{O}'} := \mathcal{D}_{\mathcal{O}_{\mu'}}$$

et sa version discrétisée

$$\forall (l,k) \in [\![1;\mathrm{N}_{\mathcal{O}_{\mu'}}]\!]^2, \quad \left(\mathrm{D}_{\mathcal{O}'}\right)_{l,k} := \int_{y\in E'_k} \frac{\partial G}{\partial \nu'}\left(M'_l,y\right)\mathrm{d}S_y,$$

dont les nœuds, extrémités des éléments $\left(E'_k\right)_{k\in[\![1;\mathrm{N}_{\mathcal{O}_{\mu'}}]\!]}$, sont situés sur $\mathscr{C}(O,1)$ dans ce contexte.

Ceci nous amène finalement à considérer la résolution du système suivant

$$\begin{pmatrix} \mathrm{S}_{\mathcal{B}} & (\gamma-1)\mathrm{D}_{\mathcal{O}_{\mu'},\mathcal{B}} \\ \mathrm{S}_{\mathcal{B},\mathcal{O}_{\mu'}} & (\gamma-1)\mathrm{D}_{\mathcal{O}'} + \frac{1}{2}(1+\gamma)\,\mathrm{Id}_{\mathcal{O}_{\mu'}} \end{pmatrix} \begin{pmatrix} \partial\widetilde{U}_{\mathcal{B}} \\ \widetilde{U}_{\mathcal{O}_{\mu'}} \end{pmatrix} = \begin{pmatrix} \mathrm{D}_{\mathcal{B}} - \frac{1}{2}\,\mathrm{Id}_{\mathcal{B}} \\ \mathrm{D}_{\mathcal{B},\mathcal{O}_{\mu'}} \end{pmatrix} \widetilde{U}_{\mathcal{B}}, \tag{3.12}$$

où $\mathrm{Id}_{\mathcal{B}}$ et $\mathrm{Id}_{\mathcal{O}_{\mu'}}$ sont respectivement les matrices identités de taille $\mathrm{N}_{\mathcal{B}}$ et $\mathrm{N}_{\mathcal{O}_{\mu'}}$.

Remarque 3.2.7. *Tout ce qui a été présenté dans cette section s'étend aisément à une scène composée de plusieurs objets. Ceci sera discuté dans la dernière section 3.4.*

3.2.3 Méthode des bases réduites

Cette sous-section s'appuie sur les références [9], [3] et [5]. Pour plus d'informations on se référera à la bibliographie de ces articles. Enfin [5] est une revue des résultats sur la méthode et les méthodes attachées. Si l'on définit pour tout $\mu' \in \mathscr{P}_{\mu'}$

$$A(\mu') := \begin{pmatrix} \mathrm{S}_{\mathcal{B}} & (\gamma-1)\mathrm{D}_{\mathcal{O}_{\mu'},\mathcal{B}} \\ \mathrm{S}_{\mathcal{B},\mathcal{O}_{\mu'}} & (\gamma-1)\mathrm{D}_{\mathcal{O}'} + \frac{1}{2}(1+\gamma)\,\mathrm{Id}_{\mathcal{O}_{\mu'}} \end{pmatrix}, X_{\mu'} := \begin{pmatrix} \partial\widetilde{U}_{\mathcal{B}} \\ \widetilde{U}_{\mathcal{O}_{\mu'}} \end{pmatrix} \text{ et } b(\mu') := \begin{pmatrix} \mathrm{D}_{\mathcal{B}} - \frac{1}{2}\,\mathrm{Id} \\ \mathrm{D}_{\mathcal{B},\mathcal{O}_{\mu'}} \end{pmatrix} \widetilde{U}_{\mathcal{B}},$$

alors on peut réécrire l'équation (3.12) sous la forme

$$A(\mu')X_{\mu'} = b(\mu'). \tag{3.13}$$

À partir de maintenant, on suppose que les solutions de cette équation ont un degré de précision suffisant et sont qualifiées d'exactes.

On cherche donc à résoudre pour un grand nombre de paramètres $\mu' \in \mathscr{P}_{\mu'}$ un système d'ordre $N_{\mathcal{B}} + N_{\mathcal{O}_{\mu'}}$ très grand. Pour un paramètre donné, la résolution d'un système de la forme (3.12)

peut être très chronophage et nécessiter un matériel informatique très performant. Comme nous sommes amenés dans les applications à le résoudre un très grand nombre de fois, cette stratégie peut s'avérer compliquée à mettre en œuvre, voire impossible. Par exemple en dimension trois l'application de la BEM pour résoudre notre problème à μ' donné, avec une géométrie constituée d'environ dix mille éléments, prend environ dix minutes, temps de constitution du système compris et ceci pour un ordinateur standard. On est bien loin du temps réel et de la facilité de mise en œuvre exigée par le capteur autonome. Il est par conséquent indispensable de trouver une alternative. C'est pour répondre à ces exigences que l'on propose l'utilisation d'une méthode de base réduite. Ce type de méthode appartient de manière plus générale aux méthodes dites de réduction de modèle. Cette méthode se décompose en deux temps.

Procédure hors-ligne : On s'autorise à résoudre un certain nombre de fois le système de grande dimension (3.12). À partir de ces solutions, on constitue une base dite "réduite" et une matrice P de changement de base associée. Il reste alors à déterminer comment optimiser le choix des paramètres μ', afin d'extraire les comportements dominants des solutions. Le système (3.12) est par conséquent réduit, c'est-à-dire ramené à un système de petite dimension, en lui appliquant le changement de base P. On stocke enfin des éléments constitutifs d'un petit système que l'on va être amené à résoudre.

Procédure en ligne : Dans cette phase, tous les calculs effectués sont de petites dimensions. Pour un nouveau μ', on assemble le système réduit, puis on le résout. Il ne reste plus qu'à appliquer le changement de base à la solution du petit système pour obtenir celle du grand.

Il faut de plus préciser que pour que cette approche soit efficace la matrice $A(\mu')$ et le vecteur $b(\mu')$ doivent vérifier une certaine condition dite de séparabilité que l'on présente ci-après. Dans le cas contraire, il est nécessaire de recourir à une méthode d'interpolation. Cette procédure vient alors compléter la procédure de base réduite. Elle respecte aussi la décomposition hors-ligne/en ligne du calcul. Nous présentons cette alternative à la sous-section 3.2.4.

Commençons par la description de la procédure hors-ligne. Soit $E := \mathbb{R}^{N_{\mathcal{B}}+N_{\mathcal{O}_{\mu'}}}$ l'ensemble dans lequel sont définies les solutions. Cet ensemble de très grande taille aurait été de dimension infinie, si l'on avait considéré l'équation (3.11) et non sa version discrétisée (3.13). On définit ensuite l'espace des solutions possibles de l'équation (3.13)

$$\mathscr{S}' := \left\{ X_{\mu'} | \forall \mu' \in \mathscr{P}_{\mu'}, X_{\mu'} \text{ est la solution de l'équation (3.13)} \right\} \subset E.$$

Comme E, l'ensemble $\mathscr{S}'$ est de très grande taille. Soit un entier $n \ll N_{\mathcal{B}} + N_{\mathcal{O}_{\mu'}}$. Nous cherchons à construire un espace de dimension n qui approche $\mathscr{S}'$. On introduit par conséquent une grille dite de construction

$$\mathscr{P}'_c := \left\{ \mu'_1, \ldots, \mu'_n \right\} \subset \mathscr{P}_{\mu'},$$

appelée "training set" ou "training grid" en anglais. Les solutions de l'équation (3.12) obtenues pour les paramètres appartenant à $\mathscr{P}'_c$ sont appelées instantanées ou échantillons, "snapshot" en anglais. On note $\mathscr{S}'_c$ l'espace engendré par les combinaisons linéaires des instantanées

$$\mathscr{S}'_c := \text{vec}\left(X_{\mu'_1}, \ldots, X_{\mu'_n} \right).$$

Dans le cas où $\mathscr{S}'$ est de dimension infinie, $\mathscr{S}'_c$ constitue une première approximation de dimension finie. Plus n sera grand, plus la grille de construction sera fine, meilleure sera l'approximation de $\mathscr{S}'$ par $\mathscr{S}'_c$. Une manière théorique de quantifier la qualité de cette approximation consiste à considérer la notion d'épaisseur de Kolmogorov, définie par la suite de réels

$$\forall n \in [\![1; N_{\mathcal{B}} + N_{\mathcal{O}_{\mu'}}]\!], \quad d_n\left(\mathscr{S}'\right) := \inf_{\substack{Y \subset E \\ dim Y = n}} \sup_{X_{\mu'} \in \mathscr{S}'} \left\| X_{\mu'} - \mathrm{P}_Y X_{\mu'} \right\|,$$

où P_Y projette orthogonalement E sur l'ensemble Y. À n fixé, d_n représente l'erreur la plus grande que l'on puisse produire par le meilleur choix de sous-espace de dimension n dans E. Cette estimation est théorique. En pratique, nous présentons dans la suite un indicateur permettant de juger de la qualité de l'approximation d'un sous espace donné.

Dans la sous-section 3.3 nous explicitons le choix de la grille de construction. Une fois la grille de construction fixée, nous cherchons à extraire de $\mathscr{S}_c'$ une famille libre à la fois petite et représentative de $\mathscr{S}_c'$. Par abus de langage, nous appelons cette famille "base réduite" associée à $\mathscr{S}_c'$. Une méthode pour construire cette base réduite consiste à utiliser la méthode de décomposition orthogonale propre (POD), aussi nommée méthode d'analyse en composantes principales. Cette méthode s'apparente à une généralisation de la méthode des moindres carrés, dans le sens où nous ne cherchons pas seulement à décrire une direction principale dans un nuage de points, mais un ensemble de directions classées par ordre d'importance. Cette méthode est très bien décrite et documentée dans [5], c'est pourquoi nous n'en n'évoquons ici que les grandes lignes.

Commençons par définir l'opérateur R de corrélation empirique

$$\forall X \in E, \quad RX := \frac{1}{n} \sum_{i=1}^{n} \left\langle X_{\mu_i'}, X \right\rangle X_{\mu_i'},$$

où $\langle \cdot, \cdot \rangle$ est le produit scalaire de E défini de manière classique par

$$\forall (X_1, X_2) \in E^2, \quad \langle X_1, X_2 \rangle := {}^t X_1 X_2.$$

On peut bien évidemment choisir un produit scalaire plus adapté au problème. Ceci constitue une amélioration dont nous discuterons dans la dernière section. L'opérateur R est un opérateur compact auto-adjoint et il existe un ensemble orthonormal $\{\varphi_i\}_{i=1}^{n'}$ constitué de $n' \leq n$ vecteurs propres de R, associées aux valeurs propres $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n'} > 0$, tel que

$$\forall X \in E, \quad RX := \sum_{i=1}^{n'} \lambda_i \langle \varphi_i, X \rangle \varphi_i.$$

Nous notons alors

$$\forall N \in [\![1; n']\!], \quad \mathrm{POD}_N \left(\left\{ X_{\mu_i'} \right\} \right) := \{\varphi_i\}_{i=1}^{N}$$

la base de taille $N \leq n'$ ainsi obtenue. La POD permet d'obtenir la meilleure approximation par rapport au carré de l'erreur et l'erreur peut être estimée par $|\lambda_{N+1}|$. Pour déterminer cette base $\{\varphi_i\}$, une méthode consiste à minimiser itérativement en N le deuxième membre de l'égalité suivante

$$\inf_{\substack{Y \subset E \\ dim Y = n}} \frac{1}{n} \sum_{i=1}^{n} \left\| X_{\mu_i'} - P_Y X_{\mu_i'} \right\|^2 = \frac{1}{n} \sum_{i=1}^{n} \left\| X_{\mu_i'} - P_{\mathrm{POD}_N} X_{\mu_i'} \right\|^2 = \sum_{i=N+1}^{n'} \lambda_i, \tag{3.14}$$

où P_{POD_N} est l'opérateur projetant orthogonalement E sur POD_N. Comme critère d'arrêt on peut soit choisir de fixer un nombre d'itération N_{BR} au départ, soit choisir ce nombre en cours de calcul de telle sorte que $\lambda_{N_{\mathrm{BR}}+1}$ soit plus petite qu'une tolérance ρ_{BR} donnée. On obtient alors la matrice de changement de base P en concaténant les N_{BR} premiers vecteurs propres sélectionnés. Ainsi $P := (\varphi_1 \ \ldots \ \varphi_{N_{\mathrm{BR}}})$ est une matrice constituée de $N_{\mathcal{B}} + N_{\mathcal{O}_{\mu'}}$ lignes et de N_{BR} colonnes.

Une procédure plus efficace numériquement consiste à résoudre un problème aux valeurs propres sur la matrice de Gram des instantanées. Pour ce faire, on concatène les instantanées afin de former une matrice

$$M := \left(X_{\mu_1'} \ \ldots \ X_{\mu_n'} \right).$$

La matrice de Gram des instantanées est la matrice composée de tous les différents produits scalaires entre les colonnes de la matrice M possible, c'est-à-dire

$$K_{X'} := {}^t M M.$$

Il se trouve que les vecteurs propres et les valeurs propres de la matrice $\frac{1}{n}K_{X'}$ sont reliés à ceux de l'opérateur R. En effet, φ est un vecteur propre unitaire de R associé à la valeur propre $\lambda > 0$ que l'on représente de la manière suivante

$$\varphi = \frac{1}{\sqrt{n\lambda}} \sum_{i=1}^{n} a_i X_{\mu'_i}, \text{ avec } \mathbf{a} = (a_i)_{i\in[\![1;n]\!]} \in (\ker K_{X'})^{\perp}$$

si et seulement si $\mathbf{a}$ est un vecteur propre unitaire de $\frac{1}{n}K_{X'}$ associé à la valeur propre $\lambda > 0$. Il est alors simple de constituer une base. On se donne par conséquent une tolérance ρ_{BR}, afin de sélectionner les vecteurs propres $\{\mathbf{a}_k\}$ de $\frac{1}{n}K_{X'}$ dont les valeurs propres $\{\lambda_k\}$ associées sont strictement supérieure à ρ_{BR}. On note N_{BR} le nombre de valeurs propres et vecteurs propres associés, ainsi sélectionnés. Pour déterminer $\text{POD}_{N_{\text{BR}}}\left(\left\{X_{\mu'_i}\right\}\right)$,

$$\forall i \in [\![1; N_{\text{BR}}]\!], \quad \varphi_i := \frac{1}{\sqrt{n\lambda_i}} M\mathbf{a}_i.$$

On obtient la matrice de changement de base P comme précédemment. La procédure pour calculer P par cette deuxième méthode est résumée dans l'algorithme 1.

Algorithme 1 : Procédure hors-ligne de construction du changement de base pour la méthode de bases réduites par POD

Entrées : Soit $\left\{X_{\mu'_i}\right\}_{i=1}^{n}$ un ensemble d'instantanées et ρ_{BR} une tolérance.
Sorties : La matrice de changement de base P.
début
Définir $M := \left(X_{\mu'_1} \ \dots \ X_{\mu'_n}\right)$;
Calculer $K_{X'} := {}^tMM$;
Extraire les valeurs propres $\{\lambda_i\}_{i\in[\![1;n]\!]}$ et les vecteurs propres $\{\mathbf{a}_i\}_{i\in[\![1;n]\!]}$ de $K_{X'}$;
Définir N_{BR} comme étant la plus grande valeur $N \in [\![1; n]\!]$, tel que $\lambda_N \geq \rho_{\text{BR}}$;
Calculer $P := M\,(\mathbf{a}_1 \dots \mathbf{a}_{N_{\text{BR}}})\,\text{diag}\left(\frac{1}{\sqrt{n\lambda_1}}, \dots, \frac{1}{\sqrt{n\lambda_{N_{\text{BR}}}}}\right)$;
fin

Remarque 3.2.8. *Cette méthode est pertinente dans un contexte de bases réduites, dès lors que nous sommes capables de calculer une très grande quantité d'échantillons en un temps raisonnable. Ceci n'est pas encore tout à fait acceptable, car trop coûteux pour un problème réaliste. Ces résultats doivent donc être considérés comme une première étape. D'autres approches théoriques et techniques existent pour réduire les calculs de bases réduites en phase hors-ligne. Elles seront exposées dans la section 3.4. Faute de temps nous ne les avons pas mis en œuvre.*

Nous utilisons à présent la matrice P pour projeter l'équation (3.13) sur l'espace de petite dimension $\mathscr{S}'_{\text{red}} := \text{vec}\left(\text{POD}_{N_{\text{BR}}}\left(\left\{X_{\mu'_i}\right\}\right)\right)$, ce qui donne

$$A_{\text{red}}(\mu')X_{\text{red},\mu'} = b_{\text{red}}(\mu'), \tag{3.15}$$

où nous avons défini $A_{\text{red}}(\mu')$ et $b_{\text{red}}(\mu')$ de la manière suivante

$$A_{\text{red}}(\mu') := {}^t\text{P}A(\mu')\text{P} \text{ et } b_{\text{red}}(\mu') := {}^t\text{P}b(\mu').$$

Pour tout $\mu' \in \mathscr{P}_{\mu'}$, les solutions des équations (3.13) et (3.15) donnent lieu à l'approximation

$$\mathrm{P} X_{\mathrm{red},\mu'} \approx X_{\mu'},$$

avec un niveau de précision lié à la tolérance ρ_{BR}.

Intéressons-nous à l'assemblage effectif de $A_{\mathrm{red}}(\mu')$ et $b_{\mathrm{red}}(\mu')$. Pour ce faire, nous définissons deux matrices P_1 et P_2, respectivement de taille $\mathrm{N}_{\mathcal{B}} \times \mathrm{N}_{\mathrm{BR}}$ et $\mathrm{N}_{\mathcal{O}_{\mu'}} \times \mathrm{N}_{\mathrm{BR}}$, de telle sorte que $\mathrm{P} = \begin{pmatrix} \mathrm{P}_1 \\ \mathrm{P}_2 \end{pmatrix}$. On peut alors ré-exprimer

$$\begin{aligned} A_{\mathrm{red}}(\mu') &= \begin{pmatrix} {}^t\mathrm{P}_1 & {}^t\mathrm{P}_2 \end{pmatrix} A(\mu') \begin{pmatrix} \mathrm{P}_1 \\ \mathrm{P}_2 \end{pmatrix} \\ &= {}^t\mathrm{P}_1 S_{\mathcal{B}} \mathrm{P}_1 + (\gamma-1){}^t\mathrm{P}_1 D_{\mathcal{O}_{\mu'},\mathcal{B}} \mathrm{P}_2 + {}^t\mathrm{P}_2 S_{\mathcal{B},\mathcal{O}_{\mu'}} \mathrm{P}_1 + (\gamma-1){}^t\mathrm{P}_2 D_{\mathcal{O}_{\mu'}} \mathrm{P}_2 + \tfrac{1}{2}(1+\gamma){}^t\mathrm{P}_2 \mathrm{P}_2 \\ &=: S_{\mathcal{B},\mathrm{red}} + (\gamma-1) D_{\mathcal{O}_{\mu'},\mathcal{B},\mathrm{red}} + S_{\mathcal{B},\mathcal{O}_{\mu'},\mathrm{red}} + (\gamma-1) D_{\mathcal{O}_{\mu'},\mathrm{red}} + \tfrac{1}{2}(1+\gamma){}^t\mathrm{P}_2 \mathrm{P}_2 \end{aligned} \tag{3.16}$$

et

$$\begin{aligned} b_{\mathrm{red}}(\mu') &= \begin{pmatrix} {}^t\mathrm{P}_1 & {}^t\mathrm{P}_2 \end{pmatrix} b(\mu') \\ &= {}^t\mathrm{P}_1 D_{\mathcal{B}} \widetilde{U}_{\mathcal{B}} - \tfrac{1}{2}{}^t\mathrm{P}_1 \widetilde{U}_{\mathcal{B}} + {}^t\mathrm{P}_2 D_{\mathcal{B},\mathcal{O}_{\mu'}} \widetilde{U}_{\mathcal{B}} \\ &=: D_{\mathcal{B},\mathrm{red}} - \tfrac{1}{2}{}^t\mathrm{P}_1 \widetilde{U}_{\mathcal{B}} + D_{\mathcal{B},\mathcal{O}_{\mu'},\mathrm{red}}, \end{aligned} \tag{3.17}$$

où l'identification des définitions en dernière ligne se fait terme à terme. Dans cette équation, on distingue deux types de termes :

termes séparables : Ce sont les termes de la forme $\Theta(\mu')\, A$, où A est une matrice ne dépendant pas du paramètre μ' et où $\Theta(\mu')$ est un scalaire. Dans les expressions (3.16) et (3.17) les termes séparables sont

$$S_{\mathcal{B},\mathrm{red}},\ (\gamma-1) D_{\mathcal{O}_{\mu'},\mathrm{red}},\ \frac{1}{2}(1+\gamma){}^t\mathrm{P}_2 \mathrm{P}_2,\ D_{\mathcal{B},\mathrm{red}} \ \text{ et } \ -\frac{1}{2}{}^t\mathrm{P}_1 \widetilde{U}_{\mathcal{B}}.$$

Une fois assemblées, ces matrices de petite taille ne dépendent pas de μ'. On peut alors les stocker.

termes non séparables : Ce sont les termes dont on ne peut extraire le paramètre lors de l'intégration. Dans les expressions (3.16) et (3.17) les termes non séparables sont

$$(\gamma-1) D_{\mathcal{O}_{\mu'},\mathcal{B},\mathrm{red}},\ S_{\mathcal{B},\mathcal{O}_{\mu'},\mathrm{red}} \ \text{ et } \ D_{\mathcal{B},\mathcal{O}_{\mu'},\mathrm{red}}.$$

Une nouvelle question se pose : comment traiter les termes non séparables ? Ceci fait l'objet de la sous-section suivante.

Dès lors qu'une matrice peut être exprimée comme une somme de termes séparables, on dit qu'elle est séparable. Il est d'usage dans la littérature d'appeler affinement décomposable ou dépendance affine cette propriété. Cette utilisation peut être trompeuse, les coefficients $\Theta(\mu')$ pouvant être non linéaires par rapport à μ'. C'est pourquoi nous préférons l'utilisation du mot séparable.

Pour finir supposons que tous les termes de notre équation soient de paramètre séparable. Dans ce cas, la procédure en ligne consiste à

1. Calculer les coefficients $\Theta(\mu')$ dépendant de μ'.
2. Assembler les termes de l'équation : multiplier les petites matrices stockées par leur coefficient et sommer le tout.
3. Résoudre le système.
4. Multiplier la solution par la matrice de passage P, pour obtenir l'approximation.

3.2.4 Méthode d'interpolation empirique

Dans le cas où les opérateurs intervenant dans la méthode de base réduite ne sont pas paramétriquement séparables, il est nécessaire de faire appel à une méthode d'interpolation. Une méthode couramment utilisée dans ce contexte est la méthode d'interpolation empirique (EIM). La présentation de la méthode donnée ici est inspirée de [3]. Pour plus d'informations sur la méthode, on se référera aux références [2], [4], [7] et [5], ainsi qu'à leurs bibliographies.

Comme nous l'avons déjà dit les termes non séparables de l'équation (3.15) présentée dans la sous-section précédente sont les matrices $(\gamma-1)D_{\mathcal{O}_{\mu'},\mathcal{B},\mathrm{red}}$, $S_{\mathcal{B},\mathcal{O}_{\mu'},\mathrm{red}}$ et $D_{\mathcal{B},\mathcal{O}_{\mu'},\mathrm{red}}$. Si l'on revient au début de notre démarche chacune de ces matrices est respectivement associée aux opérateurs suivants

$$\forall x \in \partial\mathcal{B}, \quad \mathcal{D}_{\mathcal{O}_{\mu'},\mathcal{B}}(\cdot)(x) = \int_{\partial\mathcal{O}_{\mu'}} \frac{\partial G}{\partial \nu'}(x,y)(\cdot)(y)\,\mathrm{d}S_y,$$

$$\forall x \in \partial\mathcal{O}_{\mu'}, \quad \mathcal{S}_{\mathcal{B},\mathcal{O}_{\mu'}}(\cdot)(x) = \int_{\partial\mathcal{B}} G(x,y)(\cdot)(y)\,\mathrm{d}S_y \text{ et } \mathcal{D}_{\mathcal{B},\mathcal{O}_{\mu'}}(\cdot)(x) = \int_{\partial\mathcal{B}} \frac{\partial G}{\partial \nu'}(x,y)(\cdot)(y)\,\mathrm{d}S_y,$$

auxquels après discrétisation, nous n'avons fait qu'appliquer des opérations linéaires. De plus, ce sont tous des opérateurs à noyau, dans notre cas, une fonction de Green ou sa dérivée. L'idée est par conséquent d'interpoler ces noyaux sur les différentes grilles, c'est-à-dire sur la grille de discrétisation des bords de la géométrie du problème et sur la grille de construction de la base réduite. Cette interpolation doit séparer les variables d'espaces de celles des paramètres. Ceci permet d'obtenir par linéarité une interpolation de nos matrices de petite dimension et par conséquent à les exprimer comme une somme de termes séparables. On voit alors que l'on peut distinguer deux temps dans ce calcul, ce qui est en accord avec la méthode de base réduite. Résumons :

Procédure hors-ligne : En plus de ce qui est fait pour le cas décomposable, on construit des bases de fonctions permettant l'interpolation des noyaux. On intègre ensuite ces différentes bases, que l'on projette sur l'espace réduit.

Procédure en ligne : On détermine les coefficients permettant l'interpolation à partir des matrices de petite dimension, calculées lors de la phase hors-ligne. Une fois calculés, ces termes peuvent être ajoutés aux termes séparables pré-calculés, nous permettant ainsi d'obtenir l'équation (3.15) et de la résoudre.

Pour fixer les idées et afin de ne pas alourdir le propos, nous présentons uniquement la méthode pour le terme $S_{\mathcal{B},\mathcal{O}_{\mu'},\mathrm{red}}$. La même procédure s'applique aux deux autres termes. On considère par conséquent la fonction de Green

$$G : (x,y) \mapsto G(x,y) = \frac{1}{2\pi} \ln \|x-y\|,$$

et la fonction g que l'on définit de la manière suivante

$$\forall \vartheta = (\beta,\beta') \in]-\pi;\pi]^2, \forall \mu'_k = (r',d',\theta',\gamma') \in \mathscr{P}'_c, \quad g(\mu'_k,\vartheta) := G(x,y),$$

où les variables x et y sont reliées à μ'_k et à ϑ de la manière suivante

$$x = \left((d'_k + r'_k)\cos\theta_k + r'_k\cos\beta', (d'_k + r'_k)\sin\theta_k + r'_k\sin\beta'\right) \quad \text{et} \quad y = (\cos\beta, \sin\beta).$$

Soit un entier $\mathrm{N_{EIM}} \leq n$, où n correspond au nombre d'instantanés sélectionnés. Lors de la procédure hors-ligne sont construits : une matrice triangulaire inférieure B d'ordre $\mathrm{N_{EIM}}$ dont les coefficients diagonaux sont tous égaux à 1 ; une base de fonctions $\{q_k\}_{k\in[\![1;\mathrm{N_{EIM}}]\!]}$; un ensemble de points $\{\vartheta_k\}_{k\in[\![1;\mathrm{N_{EIM}}]\!]}$, habituellement appelés points magiques et un ensemble de paramètres

$\left\{\mu_k^{\mathrm{EIM}}\right\}_{k\in[\![1;\mathrm{N_{EIM}}]\!]}$. L'objectif est d'interpoler la fonction g et ceci par séparation de variables. Nous cherchons par conséquent à obtenir une expression de la forme suivante

$$\forall \mu_k' \in \mathcal{P}_c', \forall \vartheta \in]-\pi;\pi]^2, \quad g(\mu_k', \vartheta) \approx g_{\mathrm{N_{EIM}}}\left(\mu_k', \vartheta\right) := \sum_{m=1}^{\mathrm{N_{EIM}}} \lambda_m\left(\mu_k'\right) q_m\left(\vartheta\right), \tag{3.18}$$

où pour $\mu_k' \in \mathcal{P}_c'$ donné, les coefficients $\left(\lambda_m\left(\mu_k'\right)\right)_{[\![1;\mathrm{N_{EIM}}]\!]}$ sont déterminés par la résolution de l'équation suivante

$$B\begin{pmatrix} \lambda_1\left(\mu_k'\right) \\ \vdots \\ \lambda_{\mathrm{N_{EIM}}}\left(\mu_k'\right) \end{pmatrix} = \begin{pmatrix} g\left(\mu_k', \vartheta_1\right) \\ \vdots \\ g\left(\mu_k', \vartheta_{\mathrm{N_{EIM}}}\right) \end{pmatrix}. \tag{3.19}$$

La méthode est construite de telle sorte que l'interpolation soit exacte pour $\left\{\mu_k^{\mathrm{EIM}}\right\}_{k\in[\![1;\mathrm{N_{EIM}}]\!]}$, c'est à dire

$$\forall k \in [\![1;\mathrm{N_{EIM}}]\!], \forall \vartheta \in]-\pi;\pi]^2, \quad g_{\mathrm{N_{EIM}}}\left(\mu_k^{\mathrm{EIM}}, \vartheta\right) = g(\mu_k^{\mathrm{EIM}}, \vartheta).$$

L'intégralité de la procédure est résumée dans l'algorithme 2. À chaque itération la fonction de

Algorithme 2 : Procédure EIM hors ligne : construction de la base $\{q_k\}_{k\in[\![1;d]\!]}$ et des points magiques

Entrées : soit une tolérance ρ_{EIM} et une grille de construction $\mathcal{P}_c'$.
Sorties : $\mathrm{N_{EIM}}$, $\{q_k\}_{k\in[\![1;\mathrm{N_{EIM}}]\!]}$, $\{\vartheta_k\}_{k\in[\![1;\mathrm{N_{EIM}}]\!]}$, $\left\{\mu_k^{\mathrm{EIM}}\right\}_{k\in[\![1;\mathrm{N_{EIM}}]\!]}$ et B.

début

- Déterminer $\mu_1^{\mathrm{EIM}} := \underset{\mu_k'\in\mathcal{P}_c'}{\operatorname{argmax}} \left\|g(\mu_k', \cdot)\right\|_{\mathrm{L}^\infty(\Omega)}$;
- Déterminer $\vartheta_1 := \underset{\vartheta\in]-\pi;\pi]^2}{\operatorname{argmax}} \left|g(\mu_1^{\mathrm{EIM}}, \vartheta)\right|$;
- Définir $q_1(\cdot) := \frac{g(\mu_1^{\mathrm{EIM}}, \cdot)}{g(\mu_1^{\mathrm{EIM}}, \vartheta_1)}$;
- Définir $B_{1,1} := 1$;
- Calculer $\forall \mu_k' \in \mathcal{P}_c', \forall \vartheta \in]-\pi;\pi]^2, \quad g_1\left(\mu_k', \vartheta\right)$;
- Déterminer Erreur $:= \underset{\mu_k', \vartheta}{\operatorname{argmax}} \left|g_1\left(\mu_k', \vartheta\right)\right|$;
- Définir $m := 2$;
- **tant que** Erreur $\geq \rho_{\mathrm{EIM}}$ **faire**
  - Déterminer $\mu_m^{\mathrm{EIM}} := \underset{\mu_k'\in\mathcal{P}_c'}{\operatorname{argmax}} \left\|g_{m-1}(\mu_k', \cdot)\right\|_{\mathrm{L}^\infty(\Omega)}$;
  - Déterminer $\vartheta_m := \underset{\vartheta\in]-\pi;\pi]^2}{\operatorname{argmax}} \left|g_{m-1}(\mu_m^{\mathrm{EIM}}, \vartheta)\right|$;
  - Définir $q_m(\cdot) := \frac{g_{m-1}(\mu_m^{\mathrm{EIM}}, \cdot)}{g_{m-1}(\mu_m^{\mathrm{EIM}}, \vartheta_m)}$;
  - Définir $\forall l \in [\![1;m]\!], \ B_{m,l} := q_m(\vartheta_l)$;
  - Calculer $\forall \mu_k' \in \mathcal{P}_c', \forall \vartheta \in]-\pi;\pi]^2, \quad g_{m+1}\left(\mu_k', \vartheta\right)$;
  - Erreur $\leftarrow \underset{\mu_k', \vartheta}{\operatorname{argmax}} \left|g_{m+1}\left(\mu_k', \vartheta\right)\right|$;
  - $m \leftarrow m + 1$;
- **fin**
- Définir $\mathrm{N_{EIM}} := m - 1$;

fin

base rajoutée est celle qui maximise l'erreur, c'est pourquoi nous parlons de méthode gloutonne.

Une fois les éléments d'interpolation du noyau déterminés, nous les utilisons pour interpoler $S_{\mathcal{B},\mathcal{O}_{\mu'},\mathrm{red}}$. On peut remarquer dans l'algorithme 2 que les éléments de la base fonctionnelle $\{q_k\}_{k\in[\![1;\mathrm{N_{EIM}}]\!]}$ sont des combinaisons linéaires des $\left\{g\left(\mu_k^{\mathrm{EIM}},\cdot\right)\right\}_{k\in[\![1;\mathrm{N_{EIM}}]\!]}$. Il existe donc une matrice de coefficients $(\eta_{l,k})_{(l,k)\in[\![1;\mathrm{N_{EIM}}]\!]^2}$ telle que

$$\forall k \in [\![1;\mathrm{N_{EIM}}]\!], \quad q_k = \sum_{l=1}^{\mathrm{N_{EIM}}} \eta_{l,k} g\left(\mu_l^{\mathrm{EIM}},\cdot\right).$$

À partir de cette constatation, on construit une base d'opérateurs $\left\{S_{\mu_k^{\mathrm{EIM}},\mathrm{red}}\right\}_{k\in[\![1;\mathrm{N_{EIM}}]\!]}$ à partir de l'expression précédente et par linéarité des différentes opérations. Ceci conduit à

$$\forall k \in [\![1;\mathrm{N_{EIM}}]\!], \quad S_{\mu_k^{\mathrm{EIM}},\mathrm{red}} = \sum_{l=1}^{\mathrm{N_{EIM}}} \eta_{l,k} S_{\mathcal{B},\mathcal{O}_{\mu_l^{\mathrm{EIM}}},\mathrm{red}}.$$

Durant la procédure en ligne, pour un $\mu'_k \in \mathscr{P}'_c$ donné, on peut alors approcher $S_{\mathcal{B},\mathcal{O}_{\mu'_k},\mathrm{red}}$ de la manière suivante

$$S_{\mathcal{B},\mathcal{O}_{\mu'_k},\mathrm{red}} \approx \sum_{m=1}^{\mathrm{N_{EIM}}} \lambda_m\left(\mu'_k\right) S_{\mu_m^{\mathrm{EIM}},\mathrm{red}}.$$

3.3 Résultats numériques

Dans cette section, nous testons numériquement notre approche. Nous commençons par présenter les valeurs numériques utilisées, puis des résultats illustrant l'intérêt des techniques mises en œuvre. Comme annoncé à la sous-section 3.2.2, nous comparerons les résultats obtenus par BEM et FEM pour un même jeu de paramètres en fin de section. Dans la suite, nous notons par $[\![a;k;b]\!]$, l'ensemble de k points uniformément répartis entre a et b, c'est-à-dire les points

$$\forall m \in [\![0;k-1]\!], \quad x_m := a + m\frac{b-a}{k-1}.$$

De plus, un crochet ouvert signifie que l'on exclut le point de bord correspondant.

Les capacités de l'ordinateur ayant effectué les simulations étant limitées, le nombre de nœuds au bord du capteur et de l'objet le sont également. On choisit $\mathrm{N}_{\mathcal{B}} = \mathrm{N}_{\mathcal{O}_{\mu'}} = 256$, pour un total de 512 nœuds. Pour construire la grille de construction $\mathscr{P}'_c$, nous commençons par mailler uniformément $\mathscr{P}_\mu$. Soit $\mathrm{N}_r = \mathrm{N}_d = \mathrm{N}_\theta = 7$ et $\mathrm{N}_\gamma = 10$. On définit la grille de construction intermédiaire $\mathscr{P}_c \subset \mathscr{P}_\mu$ de la manière suivante

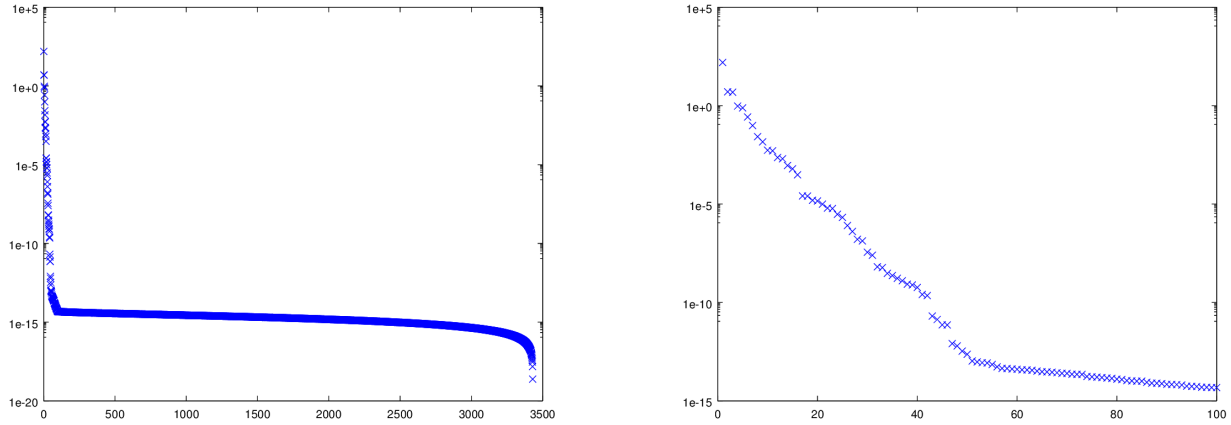
$$\forall \mu_k = (r_k, d_k, \theta_k, \gamma_k) \in \mathscr{P}_c, \quad r_k \in [\![0,5;\mathrm{N}_r - 1;5]\!] \cup \{+\infty\}, \qquad d_k \in [\![1,1;\mathrm{N}_d;2,6]\!],$$
$$\theta_k \in \left[\!\left[0;\mathrm{N}_\theta;\tfrac{\pi}{2}\right]\!\right], \qquad \gamma_k \in \left[\!\left[0,1;\tfrac{\mathrm{N}_\gamma}{2};1\right]\!\right] \cup \left[\!\left[1;\tfrac{\mathrm{N}_\gamma}{2}+1;10\right]\!\right].$$

La grille $\mathscr{P}'_c$ est alors obtenue par inversion des paramètres de la grille $\mathscr{P}_c$. On distingue deux types de paramètres :

1. les paramètres intervenants dans les termes non séparables : r', d' et θ. On peut utiliser une grille de $\mathrm{N}_{ns} := \mathrm{N}_r \times \mathrm{N}_d \times \mathrm{N}_\theta = 343$ triplets de valeurs pour ces paramètres.
2. le paramètre intervenant dans les termes séparables ou partiellement séparables : γ. Il y a $\mathrm{N}_s := \mathrm{N}_\gamma = 10$ jeux de valeurs pour ce type de paramètre.

Il y a au total $n := \mathrm{N}_s \times \mathrm{N}_{ns} = 3430$ quadruplets dans $\mathscr{P}'_c$. Au moins 343 quadruplets sont redondants. En effet lorsque $\gamma_k = 1$, nous nous retrouvons dans la situation sans objet.

Nous présentons dans le graphique 3.2 les différentes valeurs propres de la matrice de Gram associée à la grille de construction $\mathscr{P}'_c$. La décroissance quasi-linéaire est un signe encourageant.



(a) données brutes. (b) zoom.

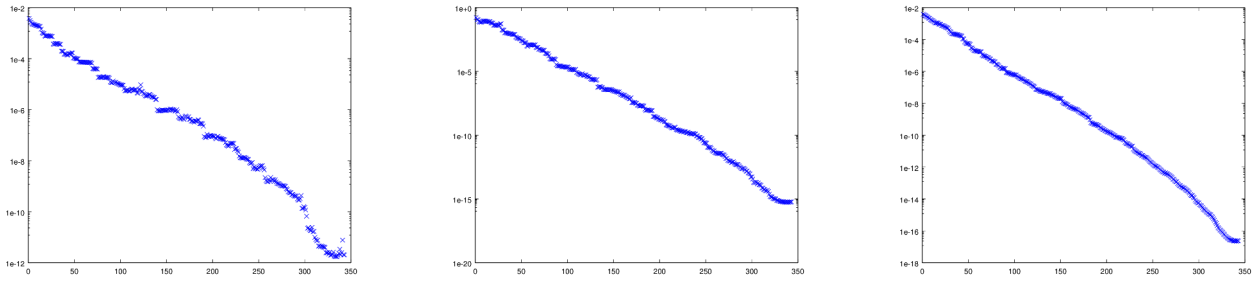
FIGURE 3.2 – Valeurs propres de la matrice de Gram, ordonnées dans le sens décroissant. Une échelle semi-log a été utilisée sur l'axe des ordonnées.

Si l'on prend une tolérance $\rho_{BR} = 10^{-12}$, on obtient que $\mathrm{N_{BR}} = 48$. Ceci entraîne une diminution de près de 91% de la taille du système étudié. Cette diminution sera d'autant plus importante que la valeur $\mathrm{N}_{\mathcal{B}} + \mathrm{N}_{\mathcal{O}_{\mu'}}$ sera grande. La précision des solutions réduites n'en sera que meilleure. Le produit scalaire utilisé pour la POD est le produit scalaire standard sur $\mathbb{R}^{\mathrm{N}_{\mathcal{B}}+\mathrm{N}_{\mathcal{O}_{\mu'}}}$, et ce pour des raisons de simplicité. Un produit scalaire plus adapté pourra être considéré dans des travaux ultérieurs à cette thèse.

Passons à présent aux résultats numériques liés à la méthode d'interpolation des opérateurs non séparables. Ces résultats sont présentés dans le graphique 3.3 ci-dessous. Nous indiquons comme exemple, le calcul effectué pour l'erreur de la matrice $S_{\mathcal{B},\mathcal{O}_{\mu'_m}}$

$$\forall l \in [\![1;\mathrm{N}_{ns}]\!], \quad \sum_{m=1}^{\mathrm{N}_{ns}} \frac{\max \left| S_{\mathcal{B},\mathcal{O}_{\mu'_m}} - S^l_{\mathcal{B},\mathcal{O}_{\mu'_m}} \right|}{\mathrm{N}_{ns}}, \tag{3.20}$$

où $S^l_{\mathcal{B},\mathcal{O}_{\mu'_m}}$ représente l'approximation d'ordre l de la matrice $S_{\mathcal{B},\mathcal{O}_{\mu'_m}}$. Pour une tolérance $\rho_{EIM} = 10^{-12}$ pour les trois opérateurs, nous obtenons une valeur de $\mathrm{N_{EIM}}$ comprise entre 250 et 343. Dans nos tests, ce résultat semble relativement indépendant de $\mathrm{N}_{\mathcal{B}} + \mathrm{N}_{\mathcal{O}_{\mu'}}$. Précisons que les opérateurs considérés dans l'interpolation sont cherchés à γ fixé, comme ils ne dépendent que des paramètres non séparables. Par conséquent $\mathrm{N_{EIM}}$ est nécessairement inférieur à $\mathrm{N}_{ns} \leq n$. Notons que les opérateurs non séparables sont invariants par rotation de l'objet, puisque leurs noyaux sont déterminés à partir de distances ou d'angles entre le bord de l'objet et celui du capteur. Un travail ultérieur consistera à mettre à profit cette invariance pour diminuer la taille de $\mathscr{P}_c$, ou bien, à taille constante, permettra de raffiner cet ensemble par rapport aux autres paramètres. Cette remarque pourra être prise en compte lors de l'application de la méthode à un capteur plus réaliste. Comme précisé en sous-section 1.3, des commandes réflexes ont été développées pour



(a) Erreur commise pour interpoler $S_{\mathcal{B},\mathcal{O}_{\mu'}}$ (b) Erreur commise pour interpoler $D_{\mathcal{B},\mathcal{O}_{\mu'}}$ (c) Erreur commise pour interpoler $D_{\mathcal{O}_{\mu'}\mathcal{B}}$

FIGURE 3.3 – Erreur d'interpolation par EIM, en fonction du nombre d'éléments de base, pour les matrices non séparables. Échelle semi-log selon l'axe des ordonnées.

le projet ANGELS, celles-ci permettent, entre autres, au capteur de s'aligner avec un objet. De plus le simulateur direct a pour finalité d'être utilisé pour la commande ou le problème inverse. Ainsi coupler la méthode de bas niveau avec celle des bases réduites, réglerait le problème des invariants géométriques non pris en compte. Cela aurait pour effet de pouvoir considérer une grille plus fine pour les autres paramètres.

Comparons maintenant les solutions obtenues par BEM avec celles obtenues par bases réduites. Cette comparaison est illustrée par le graphique 3.4. On pose $\mathrm{N_{BR}} = 80$ et $\mathrm{N_{EIM}} = 250$. L'erreur est déterminée de la manière suivante

$$\forall l \in [\![1;10;\mathrm{N_{BR}}]\!], \forall m \in [\![1;10;\mathrm{N_{EIM}}]\!], \quad \sum_{k=1}^{n} \frac{\max \left| X_{\mu'_k} - X^{l,m}_{\mathrm{red},\mu'_k} \right|}{n}, \tag{3.21}$$

où la solution réduite $X^{l,m}_{\mathrm{red},\mu'_k}$ a été calculée à partir d'un système réduit de dimension l et a nécessité trois interpolations EIM d'opérateurs non séparables. Chacune de ces interpolations fait intervenir une base de m opérateurs. Nous avons par conséquent fait le choix d'utiliser le même nombre d'opérateurs de base pour chacun des opérateurs non séparables. Ceci permet de conserver une certaine lisibilité sur le graphique. Nous remarquons que la propriété de reproduction est

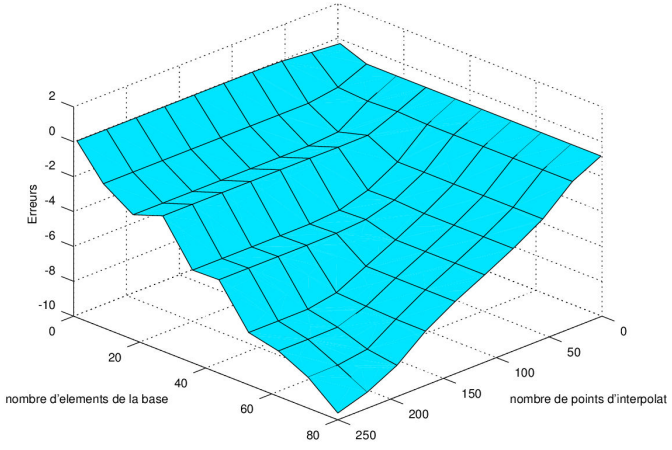


FIGURE 3.4 – Erreur commise entre les solutions exactes BEM et les solutions calculées par bases réduites couplées EIM. Échelle semi-log selon l'axe des ordonnées.

bien vérifiée. En effet, lorsque $l = \mathrm{N_{BR}}$ et que $m = \mathrm{N_{EIM}}$ nous sommes dans l'ordre de grandeur de l'erreur numérique. Comme dit précédemment, au delà d'être encourageant, ce résultat incite à l'utilisation de produits scalaires adaptés au problème lors des différentes étapes de construction. De plus, la décroissance de la courbe pour une taille de système petite dimension fixée semble plus rapide que dans les graphiques de 3.3. En effet, la construction de la base incorpore à chaque itération l'élément qui maximise l'erreur. Or cette erreur porte sur l'estimation du noyau et non de l'opérateur lui même. Une autre procédure gloutonne, orientée problème, pourrait améliorer significativement cette construction, en éliminant plus de termes. Pour plus d'informations sur ce type de stratégie, on consultera [8].

Le dernier test numérique que nous présentons correspond au graphique 3.5. Ici nous comparons des solutions exactes et des solutions bases réduites pour des paramètres aléatoires. Nous avons construit une grille aléatoire $\mathscr{P}_a \subset \mathscr{P}_\mu$, appelée "grille test" dans la littérature, constituée de cent jeux de paramètres, que nous avons ensuite inversé pour obtenir la grille $\mathscr{P}'_a$. Les solutions exactes et bases réduites sont calculées à partir de la grille $\mathscr{P}_a$. Les paramètres étant choisis ici

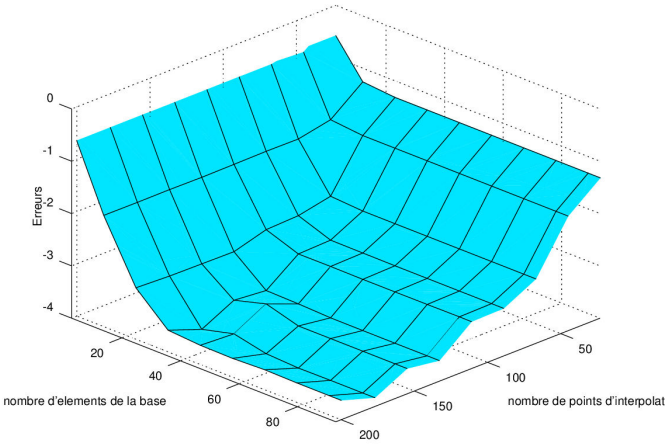


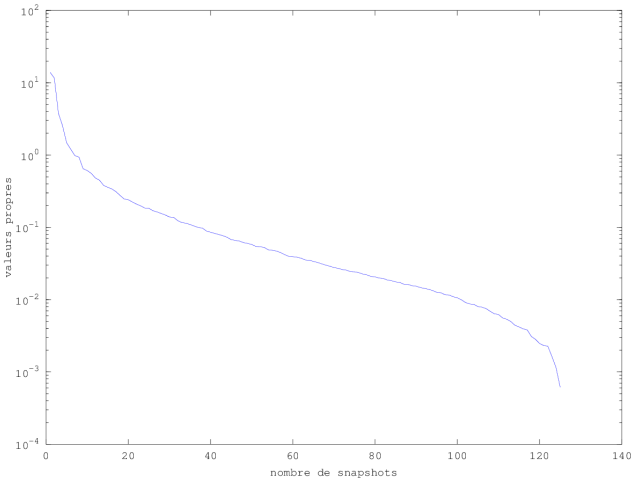
FIGURE 3.5 – Erreur commise entre des solutions exactes calculées sur une grille aléatoire et la solution calculée par BR couplée EIM. Échelle semi-log selon l'axe des ordonnées.

de façon aléatoire, ce test se rapproche d'une situation réaliste et permet de juger de la précision de l'approche. On constate sur le graphique que l'erreur stagne à partir de quarante éléments de base, ce qui confirme l'ordre de grandeur d'une base correcte, estimé en début de section avec la figure 3.2. L'EIM nécessite quant à elle plus d'éléments de base, ceci n'étant pas réellement un problème vue la rapidité de cette méthode. Le résultat reste tout de même mitigé étant donné que la courbe ne semble pas pouvoir passer en dessous de 10^{-4}. Ce résultat pourra sans doute être amélioré par les pistes indiquées précédemment. Une démarche prospective pourrait être aussi de considérer une grille pour un objet situé dans le domaine inversé géométriquement et non une grille ayant subie l'inversion.

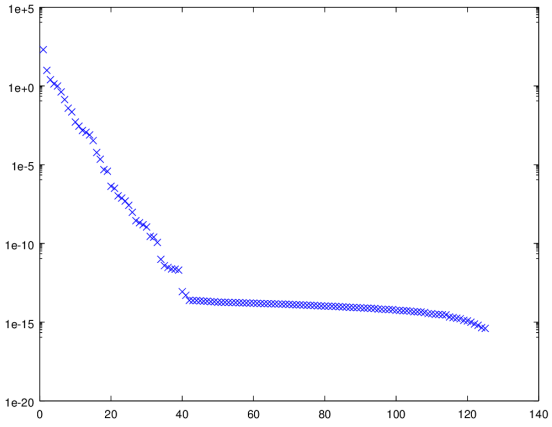
Comme mentionné plus haut, nous allons comparer l'application de la procédure par BEM et par FEM. Nous considérons un même jeu de paramètres pour calculer $\frac{\partial \tilde{u}}{\partial \nu}$ en BEM et $\tilde{u}$ en FEM. Cette grille est obtenue par inversion de la gille de construction suivante

$$\forall \mu_k = (r_k, d_k, \theta_k, \gamma_k) \in \mathscr{P}_c, \quad r_k \in [\![0,5; \mathrm{N}_r - 1; 5]\!] \cup \{+\infty\}, \quad d_k \in [\![1,1; \mathrm{N}_d; 2,1]\!], \quad \theta_k \in \left[\!\left[0; \mathrm{N}_\theta; \tfrac{\pi}{2}\right]\!\right], \quad \gamma_k = 10,$$

où $\mathrm{N}_r = 5$, $\mathrm{N}_d = 5$, $\mathrm{N}_\theta = 5$ et $\mathrm{N}_\gamma = 1$. Les deux graphiques 3.6a et 3.6b ci-dessous présentent les différentes valeurs des valeurs propres des matrices de Gram associées respectivement aux



(a) Valeurs propres de la matrice de Gram, ordonnées dans le sens décroissant, pour des solutions u calculées à partir de la FEM.



(b) Valeurs propres de la matrice de Gram, ordonnées dans le sens décroissant, pour des solutions $\frac{\partial u}{\partial \nu}$ calculées à partir de la BEM.

FIGURE 3.6 – Élément de comparaison entre les deux méthodes. Échelle semi-log selon l'axe des ordonnées.

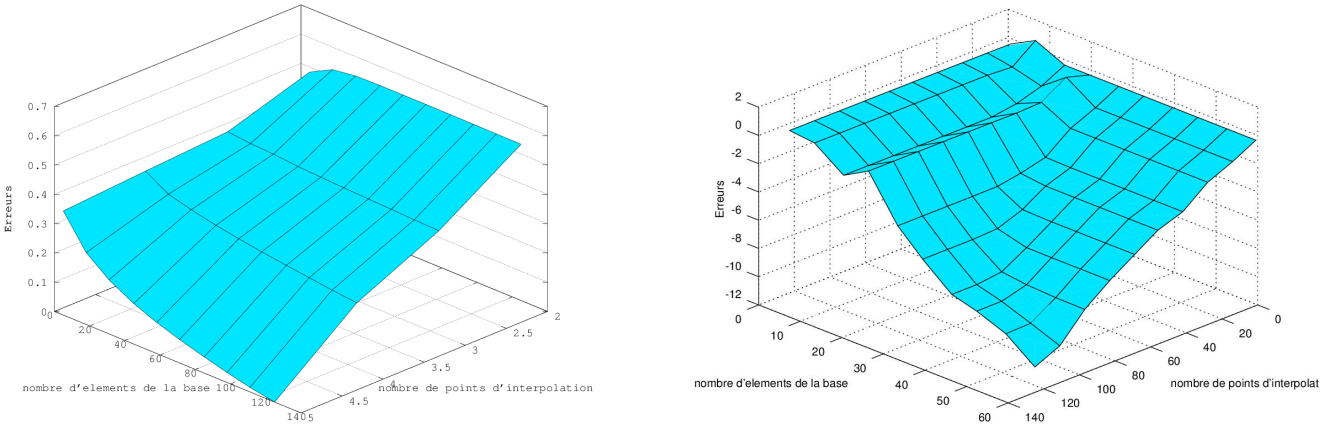
solutions calculées par FEM et par BEM. Ce graphique permet de comprendre notre choix de changer le mode de résolution. En effet, par BEM les valeurs des valeurs propres tombent dans l'erreur numérique environ à partir de la quarantième valeur propre. Tandis que par FEM, même en considérant les 125 valeurs, on ne descend pas en dessous de 10^{-3}.

Nous ne présentons pas ici de résultat pour comparer l'interpolation des termes non séparables rencontrés dans les méthodes par FEM et BEM. Il y a en effet un autre obstacle, cette fois-ci théorique, en faveur de la BEM. L'EIM exige au moins que les fonctions à interpoler soient continues, ce qui est le cas en BEM, mais pas en FEM. En BEM, nous intégrons et évaluons les opérateurs non séparables sur deux bords différents, ce qui rend les noyaux très réguliers. Dans le cas de la FEM, le noyau à interpoler est une fonction indicatrice. L'interpolation pose donc problème. C'est pour cela que nous avons utilisé une méthode d'interpolation standard pour l'interpolation en FEM. Nous avons fait appel à la fonction `interpn` d'Octave.

Comparons brièvement la méthode de bases réduites par FEM et par BEM. Nous nous référons ici au graphique 3.7. L'erreur est calculée comme dans (3.21). Les deux courbes respectent la propriété de reproduction. Malheureusement l'erreur par la méthode FEM est énorme tant que nous n'avons pas atteint le coin de la grille. Attention, le nombre de points d'interpolations sur le graphique 3.7a est à élever au cube. Cette puissance correspond au nombre de paramètres non séparables. L'interpolation à l'aide `interpn` sur Octave se fait à partir d'une grille régulière de paramètres suivant les trois dimensions des paramètres non séparables. En plus du mauvais résultat, cette méthode d'interpolation est très longue. Les bénéfices apportés par notre approche par BEM sont par conséquent significatifs.

3.4 Perspectives

Pour prolonger ces travaux, il faudrait dans un premier temps concevoir une méthode de type éléments finis de frontière pour résoudre le problème (3.2.5), ce qui n'est pas le cas dans l'approche que nous avons présentée ici, puisque nous utilisons une méthode d'éléments de frontière par collocation. L'avantage d'une telle formulation est qu'elle repose sur une interpolation et non sur des valeurs calculées analytiquement des fonctions de Green et de leurs dérivées. Ceci per-



(a) Méthode utilisée : FEM.

(b) Méthode utilisée : BEM.

FIGURE 3.7 – Élément de comparaison entre les deux méthodes. Sur chacun des graphiques est présentée l'erreur commise entre la solution exacte et la solution calculée par BR couplée avec une méthode d'interpolation pour les termes non séparables. Échelle semi-log selon l'axe des ordonnées.

mettrait de caractériser plus simplement le produit scalaire et la norme induite du problème. Cette méthode devra, de plus, être orientée vers la résolution de problèmes en trois dimensions et ceci pour une géométrie et des conditions plus réalistes pour le capteur. La forme du capteur et le positionnement des électrodes pourront être étudiés pour être plus adaptés à l'inversion géométrique. Comme nous l'avons évoqué dans la section précédente, les invariants géométriques et électriques devront être mieux exploités afin de simplifier l'interpolation des opérateurs de simple et double couches.

Comme nous l'avons dit précédemment, la POD pour être efficace nécessite un très grand nombre d'instantanés. Une fois le cadre fonctionnel et géométrique mieux maîtrisés, une méthode de bases réduites utilisant un estimateur a posteriori et un algorithme glouton (*greedy* en anglais) pourront être envisagés. Cette alternative permettrait d'estimer en petite dimension l'erreur commise entre l'approximation petite dimension et la solution BEM et aurait l'avantage d'une sélection itérative des instantanés à ajouter à la base. Par conséquent, ceux-ci seraient uniquement calculés s'ils étaient sélectionnés. La constitution de la base serait alors accélérée et la grille de construction considérée pourra être affinée.

Une troisième extension possible consisterait à prendre en compte une situation multi-objets. Deux alternatives sont possibles.

1. Combiner la méthode des réflexions et la méthode des bases réduites.

2. Pour $m \in \mathbb{N}^*$, réduire le système à m objets suivants

$$A_m(\mu')X_{m,\mu'} = b_m(\mu'), \tag{3.22}$$

où

$$A_m(\mu') := \begin{pmatrix} \mathrm{S}_{\mathcal{B}} & (\gamma_1-1)\mathrm{D}_{\mathcal{O}_{1,\mu'},\mathcal{B}} & (\gamma_2-1)\mathrm{D}_{\mathcal{O}_{2,\mu'},\mathcal{B}} & \dots & (\gamma_m-1)\mathrm{D}_{\mathcal{O}_{m,\mu'},\mathcal{B}} \\ \mathrm{S}_{\mathcal{B},\mathcal{O}_{1,\mu'}} & (\gamma_1-1)\mathrm{D}_{\mathcal{O}_{1,\mu'}} + \frac{1}{2}(\gamma_1+1)\,\mathrm{Id}_{\mathcal{O}_{1,\mu'}} & (\gamma_2-1)\mathrm{D}_{\mathcal{O}_{2,\mu'},\mathcal{O}_{1,\mu'}} & \dots & (\gamma_m-1)\mathrm{D}_{\mathcal{O}_{m,\mu'},\mathcal{O}_{1,\mu'}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{S}_{\mathcal{B},\mathcal{O}_{1,\mu'}} & (\gamma_1-1)\mathrm{D}_{\mathcal{O}_{1,\mu'},\mathcal{O}_{m,\mu'}} & (\gamma_2-1)\mathrm{D}_{\mathcal{O}_{2,\mu'},\mathcal{O}_{m,\mu'}} & \dots & (\gamma_m-1)\mathrm{D}_{\mathcal{O}_{m,\mu'}} + \frac{1}{2}(\gamma_m+1)\,\mathrm{Id}_{\mathcal{O}_{m,\mu'}} \end{pmatrix},$$

$$X_{m,\mu'} := \begin{pmatrix} \partial\widetilde{U}_{\mathcal{B}} \\ \widetilde{U}_{\mathcal{O}_{1,\mu'}} \\ \vdots \\ \widetilde{U}_{\mathcal{O}_{m,\mu'}} \end{pmatrix} \quad \text{et} \quad b_m(\mu') := \begin{pmatrix} \mathrm{D}_{\mathcal{B}} - \frac{1}{2}\,\mathrm{Id} \\ \mathrm{D}_{\mathcal{B},\mathcal{O}_{1,\mu'}} \\ \vdots \\ \mathrm{D}_{\mathcal{B},\mathcal{O}_{m,\mu'}} \end{pmatrix} \widetilde{U}_{\mathcal{B}}.$$

Si possible, cette réduction doit s'opérer à partir d'une matrice de changement de base constituée à partir d'une situation comportant très peu d'objets. Il est à noter que beaucoup de termes sont redondants du point de vue de l'interpolation. Par exemple, la procédure hors ligne réalisée pour le terme non séparable $\mathrm{S}_{\mathcal{B},\mathcal{O}_{\mu'}}$ dans la sous-section 3.2.4, peut être directement réappliquée aux termes

$$\mathrm{S}_{\mathcal{B},\mathcal{O}_{1,\mu'}},\ \mathrm{S}_{\mathcal{B},\mathcal{O}_{2,\mu'}},\ \ldots,\ \mathrm{S}_{\mathcal{B},\mathcal{O}_{m,\mu'}}.$$

Il est seulement nécessaire de construire une base en plus, pour les termes de la forme $\mathrm{D}_{\mathcal{O}_{l,\mu'},\mathcal{O}_{k,\mu'}}$. Dès lors que le nombre d'objets reste raisonnable, une telle démarche peut être entreprise.

En parallèle de la résolution base réduite directe, un travail sur le problème inverse pourra être mené. Dans [6], H. Lemonnier et J.F. Peytraud ont traité un problème analogue à celui que l'on a obtenu après inversion géométrique. Dans leur article, ils proposent une méthode de résolution du problème inverse basé sur la BEM. Cette méthode semble fonctionner même en situation bruitée. Ceci constituerait par conséquent une première piste pour aborder le problème d'électrolocation. Enfin, une fois la méthode testée par BEM, il serait intéressant du point de vue mathématique de savoir comment se comporte cette méthode dès lors qu'elle est appliquée à partir du système réduit.

Références

[1] Kendall E. ATKINSON. *The Numerical Solution of Integral Equations of the Second Kind*. Cambridge Books Online. Cambridge University Press, 1997. ISBN: 9780511626340. DOI: 10.1017/CBO9780511626340.

[2] Maxime BARRAULT et al. "An 'empirical interpolation' method : application to efficient reduced-basis discretization of partial differential equations". *Comptes Rendus Mathematique* 339.9 (2004), pp. 667–672. ISSN: 1631-073X. DOI: 10.1016/j.crma.2004.08.006. URL: http://www.sciencedirect.com/science/article/pii/S1631073X04004248.

[3] Fabien CASENAVE. "Méthodes de réduction de modèles appliquées à des problèmes d'aéroacoustique résolus par équations intégrales". Thèse de doctorat dirigée par Ern, Alexandre Mathématiques Paris Est 2013. PhD thesis. 2013. URL: www.theses.fr/2013PEST1076.

[4] Jens L. EFTANG and Benjamin STAMM. "Parameter multi-domain 'hp' empirical interpolation". *International Journal for Numerical Methods in Engineering* 90.4 (2012), pp. 412–428. ISSN: 1097-0207. DOI: 10.1002/nme.3327. URL: http://dx.doi.org/10.1002/nme.3327.

[5] B. HAASDONK. *Reduced Basis Methods for Parametrized PDEs – A Tutorial Introduction for Stationary and Instationary Problems*. Tech. rep. Chapter to appear in P. Benner, A. Cohen, M. Ohlberger and K. Willcox : "Model Reduction and Approximation for Complex Systems", Springer. 2014. URL: http://www.simtech.uni-stuttgart.de/publikationen/prints.php?ID=938.

[6] H. LEMONNIER and J.F. PEYTRAUD. "Is 2D impedance tomography a reliable technique for two-phase flow ?" *Nuclear Engineering and Design* 184.2-3 (1998), pp. 253–268. ISSN: 0029-5493. DOI: 10.1016/S0029-5493(98)00201-5. URL: http://www.sciencedirect.com/science/article/pii/S0029549398002015.

[7] Yvon MADAY and Olga MULA. "A generalized empirical interpolation method : application of reduced basis techniques to data assimilation". *Analysis and Numerics of Partial Differential Equations*. Ed. by SPRINGER. Springer INdAM Series. Springer, Jan. 2013, pp. 221–235. DOI: 10.1007/978-88-470-2592-9_13. URL: http://hal.upmc.fr/hal-00812913.

[8] C. PRUD'HOMME et al. "Reliable real-time solution of parametrized partial differential equations : Reduced-basis output bound methods". *Journal of Fluids Engineering-Transactions of The ASME* 124.1 (2002), pp. 70–80. ISSN: 0098-2202. DOI: 10.1115/1.1448332.

[9] Karen VEROY et al. "A posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations". *Proceedings of the 16th AIAA computational fluid dynamics conference* 3847 (2003), pp. 23–26. DOI: 10.2514/6.2003-3847.

4

Contrôlabilité inverse en chimie quantique

Sommaire

4.1 Introduction et notations

Ce chapitre a été introduit à la section 1.4. Il constitue une version augmentée d'un acte de conférence [6]. Nous considérons donc le problème inverse 1.4.2. Présentons notre démarche. Après avoir fixé une discrétisation en espace de l'équation de Schrödinger (1.30), nous donnons dans la section 4.2 une formulation mathématique du problème 1.4.2 vis-à-vis de cette discrétisation. Dans la section 4.3, nous présentons un résultat local pour ce problème. La section 4.4 décrit l'algorithme permettant de résoudre numériquement le problème d'identification. Nous concluons par des tests numériques dans la section 4.5.

Introduisons quelques notations concernant certains types de matrices que nous utilisons tout au long du chapitre. Soit $N_d \in \mathbb{N}^*$, on note $\mathbb{C}^{N_d,N_d}$ et $\mathbb{R}^{N_d,N_d}$ l'ensemble des matrices carrées d'ordre N_d à coefficients complexes et réels respectivement. Nous notons :

matrice unitaire :	$\mathcal{U}$	$=$	$\left\{M \in \mathbb{C}^{N_d,N_d} \mid M^*M = MM^* = \mathrm{Id}\right\}$;
matrice orthogonale :	$\mathcal{O}$	$=$	$\left\{M \in \mathbb{R}^{N_d,N_d} \mid {}^tMM = M{}^tM = \mathrm{Id}\right\}$;
matrice antihermitienne :	$\mathcal{H}$	$=$	$\left\{M \in \mathbb{C}^{N_d,N_d} \mid M^* = M\right\}$;
matrice hermitienne :	$\mathcal{A}_\mathcal{H}$	$=$	$\left\{M \in \mathbb{C}^{N_d,N_d} \mid M^* = -M\right\}$;
matrice symétrique :	$\mathcal{S}$	$=$	$\left\{M \in \mathbb{R}^{N_d,N_d} \mid {}^tM = M\right\}$;
matrice antisymétrique :	$\mathcal{A}$	$=$	$\left\{M \in \mathbb{R}^{N_d,N_d} \mid {}^tM = -M\right\}$;
matrice symétrique à diagonale nulle :	$\mathcal{S}_0$	$=$	$\mathcal{S} \cap \left\{M \in \mathbb{R}^{N_d,N_d} \mid \forall k \in [\![1; N_d]\!], M_{k,k} = 0\right\}$;
matrices unitaires symétriques :	$\mathcal{U}/\mathcal{O}$	$=$	$\left\{M \in \mathcal{U} \mid {}^tM = M\right\}$;

où ${}^tM$ désigne la matrice transposée de M, M^* son adjoint (*i.e.* $M^* = {}^t\overline{M}$) et Id représente la matrice identité. Pour simplifier, nous avons omis la dépendance à N_d des espaces matriciels précédents. Dans la suite, nous notons respectivement $\Re z$ et $\mathrm{Im}\, z$ la partie réelle et la partie imaginaire d'un nombre complexe z. Les symboles "$\cdot$", "$\otimes$" et "$\oplus$" représentent respectivement le produit de Hadamard, aussi dit de Schur,

$$\forall\, (A,B) \in \mathbb{C}^{m,n} \times \mathbb{C}^{m,n}, \quad A \cdot B = \begin{pmatrix} a_{1,1}b_{1,1} & \cdots & a_{1,n}b_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1}b_{m,1} & \cdots & a_{m,n}b_{m,n} \end{pmatrix} \in \mathbb{C}^{m,n},$$

le produit de Kronecker

$$\forall\, (A,B) \in \mathbb{C}^{m,n} \times \mathbb{C}^{p,q}, \quad A \otimes B = \begin{pmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{m,1}B & \cdots & a_{m,n}B \end{pmatrix} \in \mathbb{C}^{mp,nq};$$

il s'agit d'un cas particulier du produit tensoriel ; et la somme de Kronecker

$$\forall\, (A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{m,m}, \quad A \oplus B = A \otimes \mathrm{Id}_m + \mathrm{Id}_n \otimes B.$$

Enfin, rappelons quatre résultats importants sur les matrices unitaires.

Propriétés 4.1.1. 1. *Pour toute matrice antihermitienne $W \in \mathcal{A}_\mathcal{H}$, il existe une matrice hermitienne $H \in \mathcal{H}$, une matrice symétrique $S \in \mathcal{S}$ et une matrice antisymétrique $A \in \mathcal{A}_\mathcal{H}$ telle que $W = \imath H = A + \imath S$.*

2. *Pour toute matrice unitaire $U \in \mathcal{U}$, il existe une matrice antihermitienne $W \in \mathcal{A}_\mathcal{H}$ telle que $U = e^W$.*

3. *Pour toute matrice unitaire et symétrique $U \in \mathcal{U}/\mathcal{O}$, il existe une matrice symétrique $S \in \mathcal{S}$ telle que $U = e^{\imath S}$.*

4. *Pour toute matrice unitaire $U \in \mathcal{U}$, il existe une matrice diagonale réelle $D \in \mathbb{R}^{N_d,N_d}$ et une matrice antihermitienne $W_1 \in \mathcal{A}_\mathcal{H}$ telle que $U = e^{W_1}e^{\imath D}e^{W_1^*}$. De façon plus générale, il existe une matrice symétrique $S \in \mathcal{S}$ et une matrice anti-hermitienne $W_2 \in \mathcal{A}_\mathcal{H}$ telle que $U = e^{W_2}e^{\imath S}e^{W_2^*}$.*

4.2 Présentation du problème et premiers résultats

Dans cette section, nous présentons le problème inverse en dimension finie dans la sous-section 4.2.1. Celle-ci est suivie par la sous-section 4.2.2 traitant de quelques types de solutions particulières et utile aussi bien pour la suite de l'exposé que pour certaines applications.

4.2.1 Problème général

Dans la section 1.4 de l'introduction nous avons introduit le problème inverse en dimension infinie. Or, dans certains domaines d'applications l'équation (1.30) est de dimension finie. De plus, il est généralement plus simple de commencer l'étude d'un problème dans ce cadre, sachant que les systèmes de dimensions infinies de ce type se laissent, dans de nombreux cas, bien approcher par des systèmes de dimensions finies. C'est pourquoi nous présentons une version en dimension finie du problème 1.4.2. Nous ne rentrerons ni dans les détails de l'existence, ni dans ceux liés à la réduction de modèle des systèmes de dimensions infinies. L'idée étant, pour ces derniers, de constituer une base d'un espace de dimension finie approchant correctement l'ensemble des solutions de l'équation (1.30). En pratique, les chimistes utilisent souvent les N_d premiers comme base. Cette base peut, par exemple, être constituée des N_d premiers vecteurs propres de l'Halmiltonien. Une fois cette base choisie, nous ne faisons que réexprimer les éléments constitutifs de (1.30) dans cette même base. Pour plus d'informations à ce sujet, on consultera le chapitre intitulé *Control of quantum dynamics : concepts, procedures and future prospects* dans [4].

Supposons donc cette base choisie et commençons par modifier l'équation (1.30) en conséquence. Soit un temps $T > 0$, on considère une trajectoire $t \mapsto U(t) \in \mathcal{U}$[1], appelée *propagateur*, dont la dynamique sur l'intervalle de temps $[0;T]$ est régie par l'équation de Schrödinger suivante

$$\imath \dot{U}(t) = (H_0 + \varepsilon(t)\mu)\, U(t), \tag{4.1}$$

où $H_0 \in \mathcal{S}$ représente l'Hamiltonien, $\varepsilon \in L^2((0;T);\mathbb{R})$ correspond au champ électrique du laser et $\mu \in \mathcal{S}$ est la matrice associée au moment dipolaire du système. Précisons le lien entre la fonction d'onde ψ solution de l'équation (1.30) et le propagateur U solution de l'équation (4.1). L'ensemble de matrices $t \mapsto U(t)$ résulte de la concaténation des trajectoires $t \mapsto \psi(\cdot,t)$ lorsque $\psi(\cdot,t=0)$ parcourt la base orthonormée choisie. On se donne au temps $t = 0$ un état initial $U_{\text{init}} \in \mathcal{U}$,

$$i.e. \quad U(0) = U_{\text{init}}. \tag{4.2}$$

Sous les hypothèses précédentes, le problème $((4.1)-(4.2))$ est généralement bien posé. Cette démonstration se fait au cas par cas suivant la régularité de ε. Dès lors que ce propagateur existe, il est unique et dépend continûment de l'initialisation. On a de plus la propriété :

$$\forall t \in [0;T], \quad \|U(t)\|_F = \|U_{\text{init}}\|_F\,^2,$$

où l'on note $\|\cdot\|_F$ la norme de Frobenius induite par le produit scalaire

$$(A,B) \in \mathbb{C}^{N_d,N_d} \times \mathbb{C}^{N_d,N_d} \mapsto tr\,(A^*B)\,.$$

[1]Dès lors que le propagateur U existe : l'égalité $\frac{\mathrm{d}}{\mathrm{d}t}\left(U(t)^*U(t)\right) = 0$ vraie pour tout $t \in [0;T]$ et l'initialisation U_{init} prise dans l'ensemble des matrices unitaires, induisent que $t \mapsto U(t) \in \mathcal{U}$.

[2]Comme $U_{\text{init}} \in \mathcal{U}$, on a : $\forall t \in [0;T], \quad \|U(t)\|_F = \sqrt{N_d}$.

Il est à noter que sans perte de généralité, nous pouvons toujours nous ramener au problème où U_{init} est égale à l'identité. En effet, le propagateur $t \mapsto V(t) \in \mathcal{U}$, défini de la manière suivante

$$\forall t \in [0;T],\ V(t) := U(t)U_{\text{init}}^{*},$$

est le propagateur solution du problème $((4.1)-(4.2))$ pour $V(t=0) = \text{Id}$. Un tel propagateur est appelé *propagateur fondamental*.

On notera que lorsque les matrices H_0 et μ commutent,

$$i.e. \quad [H_0, \mu] := H_0\mu - \mu H_0 = 0,$$

la solution de ce problème est explicitement donnée par

$$\forall t \in [0;T], \quad U(t) = e^{-\imath\left(t\,H_0 + \int_0^t \varepsilon(s)\mathrm{d}s\,\mu\right)} U_{\text{init}}. \tag{4.3}$$

Malheureusement pour des applications pertinentes en pratique, on ne peut généralement pas faire cette hypothèse.

Formalisons à présent le problème de contrôlabilité en dimension finie étudié.

Problème 4.2.1. *Soit un temps $T > 0$ et un champ laser $\varepsilon \in L^2((0;T);\mathbb{R})$ fixé. On cherche à déterminer un couple $(H_0, \mu) \in \mathcal{S} \times \mathcal{S}$ de sorte que le propagateur fondamental, solution de l'équation (4.1), atteigne un état cible $U_{cible} \in \mathcal{U}$ donné au temps $t = T$,*

$$\text{i.e.} \quad U(T) = U_{cible}. \tag{4.4}$$

En introduisant l'application φ, définie par

$$\begin{array}{rccl} \varphi: & \mathcal{S}\times\mathcal{S} & \to & \mathcal{U}, \\ & (H_0,\mu) & \mapsto & \varphi(H_0,\mu) = U(T), \end{array}$$

on peut reformuler le problème 4.2.1 comme étant celui de la surjectivité de φ. Nous parlerons indifféremment des deux formulations par la suite.

Précisons que l'Hamiltonien H_0 est ici recherché comme étant une matrice symétrique et non une matrice hermitienne. C'est une situation particulière qui est naturelle pour le problème considéré, l'Hamiltonien étant la somme d'un opérateur cinétique et d'un potentiel, tous deux réels. Pour les mêmes raisons, on suppose que le moment dipolaire μ est réel et, quand cela est possible, que ses éléments diagonaux sont tous nuls. Cette condition supplémentaire est motivée par le désir d'identifier un unique couple (H_0, μ). En effet sous ces hypothèses, le nombre d'équations obtenues

$$\#\mathcal{S} + \#\mathcal{S}_0 = \frac{N_d(N_d+1)}{2} + \frac{N_d(N_d-1)}{2} = N_d^2,$$

est égal au nombre d'inconnues $\#\mathcal{U} = N_d^2$. Au vu de l'inclusion $\mathcal{S}\times\mathcal{S}_0 \subset \mathcal{S}\times\mathcal{S}$, dès lors que l'on a identifié un couple appartenant à $\mathcal{S}\times\mathcal{S}_0$ solution de l'équation (4.4), le problème 4.2.1 est résolu.

Ce problème est lié à des problèmes inverses en contrôle quantique, comme mentionné dans la section 1.4 de l'introduction. Rappelons que contrairement aux travaux précédents, nous ne cherchons pas ici à concevoir un champ laser nous permettant d'identifier un couple (H_0, μ), mais plutôt à étudier les propriétés des champs ε rendant l'équation (4.4) résoluble et nous permettant de déterminer numériquement les opérateurs solutions H_0 et μ correspondants.

4.2.2 Quelques cas particuliers

Nous présentons ici deux résultats. Le premier montre que l'application φ est surjective sur plus de la moitié de l'espace image $\mathcal{U}$ et ceci quel que soit le champ laser.

Proposition 4.2.2. *Soit un champ laser $\varepsilon \in L^2((0;T);\mathbb{R})$ et une cible $U_{cible} \in \mathcal{U}/\mathcal{O}$, alors l'application φ est surjective.*

Démonstration. Ce résultat repose sur la formule (4.3). Comme on suppose que $U_{\text{cible}} \in \mathcal{U}/\mathcal{O}$, d'après la proposition 3 de la propriété 4.1.1, il existe $S \in \mathcal{S}$ tel que $U_{\text{cible}} = e^{\imath S}$. Soit $\mu \in \mathcal{S}_0$, tel que $[\mu, S] = 0$, alors

$$\forall k \in \mathbb{Z}, \quad (H_0, \mu) = \left(-\frac{1}{T} \left(S + \int_0^T \varepsilon(t) \mathrm{d}t \, \mu + 2k\pi \, \mathrm{Id} \right), \mu \right),$$

est un couple solution du problème 4.2.1. Il existe toujours une telle matrice μ : il suffit, par exemple, de choisir la matrice nulle. □

Bien qu'encourageant, ce résultat est, somme toute, insuffisant. Dans la section suivante nous allons étendre localement ce résultat, en montrant que l'application φ est surjective sur un voisinage de presque tout élément de $\mathcal{U}/\mathcal{O}$.

Le second résultat, qui implique la surjectivité de φ moyennant un choix particulier de champs laser, trouve des applications en résonance magnétique nucléaire (RMN), domaine dans lequel il est possible de générer des champs constants. On pourra consulter [5] pour plus d'information sur le sujet.

Proposition 4.2.3. *Soit un temps intermédiaire $0 < T_{int} < T$, deux constantes réelles distinctes ε_1, ε_2 et un champ ε défini de la manière suivante*

$$\forall t \in [0;T], \quad \varepsilon(t) := \begin{cases} \varepsilon_1 & si\ t \in [0;T_{int}[, \\ \varepsilon_2 & sinon, \end{cases}$$

Sous ces hypothèses et pour toute cible $U_{cible} \in \mathcal{U}$, l'application φ est surjective.

Ce résultat repose sur le lemme suivant.

Lemme 4.2.4. *Soit $U \in \mathcal{U}$, alors il existe $(S_1, S_2) \in \mathcal{S}^2$, tel que*

$$U = \exp(\imath S_2) \exp(\imath S_1).$$

Démonstration. D'après la proposition 4 de la propriété 4.1.1, il existe $P \in \mathcal{U}$ et $(\alpha_i)_{i \in [\![1;N_d]\!]} \in \mathbb{R}^{N_d}$, tel que $U = P \mathrm{diag}\left((e^{\imath \alpha_i})_{i \in [\![1;N_d]\!]} \right) P^*$. Soit un vecteur quelconque $(\beta_i)_{i \in [\![1;N_d]\!]} \in \mathbb{R}^{N_d}$, on définit

$$U_2 := P \mathrm{diag}\left(\left(e^{-\imath \beta_i} \right)_{i \in [\![1;N_d]\!]} \right) {}^t P \ \text{ et } \ U_1 := \overline{P} \mathrm{diag}\left(\left(e^{\imath (\alpha_i + \beta_i)} \right)_{i \in [\![1;N_d]\!]} \right) P^*.$$

Les matrices U_1 et U_2 appartiennent clairement à $\mathcal{U}/\mathcal{O}$; il existe donc, d'après la proposition 3 de la propriété 4.1.1, un couple $(S_1, S_2) \in \mathcal{S}^2$, tel que

$$\exp(\imath S_1) = U_1 \quad \text{et} \quad \exp(\imath S_2) = U_2.$$

On vérifie enfin aisément que

$$U = \exp(\imath S_2) \exp(\imath S_1).$$

□

Démonstration de la proposition 4.2.3. Au vu de la formule (4.3), on a que

$$\forall t \in [0;T], \quad U(t) := \begin{cases} e^{-\imath\, t(H_0+\varepsilon_1\mu)}, & \text{si } t \in [0;T_{int}[, \\ e^{-\imath\,(t-T_{int})(H_0+\varepsilon_2\mu)} e^{-\imath\, T_{int}(H_0+\varepsilon_1\mu)}, & \text{sinon,} \end{cases}$$

dès lors que le couple (H_0, μ) existe. Ainsi

$$U(T) = e^{-\imath\,(T-T_{int})(H_0+\varepsilon_2\mu)} e^{-\imath\, T_{int}(H_0+\varepsilon_1\mu)},$$

sous cette condition.

Soit $U_{\text{cible}} \in \mathcal{U}$. Alors, d'après le lemme 4.2.4, il existe $(S_1, S_2) \in \mathcal{S}^2$, tel que

$$U_{\text{cible}} = \exp(\imath\, S_2) \exp(\imath\, S_1).$$

La matrice U_{cible} admet au moins un antécédent par φ, dès lors que le système

$$\begin{cases} S_1 &= -T_{int}(H_0 + \varepsilon_1\mu), \\ S_2 &= -(T - T_{int})(H_0 + \varepsilon_2\mu), \end{cases}$$

admet un couple solution $(H_0, \mu) \in \mathcal{S} \times \mathcal{S}$. Les solutions de ce système sont

$$H_0 = \frac{1}{T_{int}(T - T_{int})(\varepsilon_1 - \varepsilon_2)} \left((T - T_{int})\varepsilon_2 S_1 - T_{int}\varepsilon_1 S_2\right)$$

et

$$\mu = \frac{1}{T_{int}(T - T_{int})(\varepsilon_1 - \varepsilon_2)} \left((T_{int} - T) S_1 + T_{int} S_2\right).$$

□

Remarque 4.2.5. *Attention, dans la proposition précédente, le couple solution (H_0, μ) appartient à $\mathcal{S} \times \mathcal{S}_0$ si et seulement si*

$$\text{diag}(S_2) = \frac{T - T_{int}}{T_{int}} \text{diag}(S_1).$$

Si ce n'est pas le cas, il appartient simplement à $\mathcal{S} \times \mathcal{S}$. Pour conclure cette section, indiquons qu'il est possible d'utiliser un développement de Magnus pour obtenir des résultats approchés de ce problème. Celui-ci permet d'exprimer le propagateur U comme étant l'exponentielle d'une série de matrice dont les termes sont constitués à partir de ε et de commutateurs construits à partir des matrices H_0 et μ. Plus généralement, le développement de Magnus et celui de Fer sont utilisés afin d'approcher les solutions d'équations différentielles linéaires de matrices par des exponentielles de matrices. Pour plus d'informations, consulter [3] et sa bibliographie. Pour exploiter un tel type de développement, on peut, par exemple, supposer que H_0 et μ ont un petit commutateur. Ici, nous ne rentrons pas plus dans les détails. En effet, les travaux dans ce sens n'ont pas encore donné de résultats probants, contrairement à ceux mentionnés par la suite.

4.3 Résultats de contrôlabilité locale

Nous commençons ici par linéariser le problème, et obtenons du même coup une formulation intégrale de ce dernier. Ces résultats sont consignés dans la sous-section 4.3.1. Les inconnues étant des matrices dans les différentes équations considérées, nous adaptons donc dans la sous-section 4.3.2 ces dernières pour que les inconnues soient alors des vecteurs. Cette modification consiste simplement en une vectorisation, notion en lien avec la tensorisation, qui implique cependant dans notre cas des modifications techniques pour notre problème. Dans la sous-section 4.3.3 est présenté un résultat local sur la surjectivité de φ, qui est formalisé par la proposition 4.2.2.

4.3.1 Formulation intégrale du problème

Dans cette sous section, nous présentons quelques résultats théoriques sur l'inversion locale de l'équation (4.4) dont la finalité est une extension, valable localement, de la proposition 4.2.2. Afin de l'obtenir, nous appliquons des techniques issues du calcul différentiel. Pour ce faire, nous introduisons le plan tangent $\mathcal{A}_{H_0,\mu}$ à φ, au point $(H_0,\mu)\in\mathcal{S}^2$, qui est l'ensemble matriciel

$$\mathcal{A}_{H_0,\mu}=\left\{M\in\mathbb{C}^{N_d,N_d}\mid M^*U(T)+U(T)^*M=0\right\}.$$

On considère ensuite la différentielle de φ, au point $(H_0,\mu)\in\mathcal{S}^2$, définie par

$$\begin{array}{rccc}\mathrm{d}\varphi\,(H_0,\mu):&\mathcal{S}\times\mathcal{S}&\to&\mathcal{A}_{H_0,\mu},\\&(\delta H_0,\delta\mu)&\mapsto&\delta U(T),\end{array}$$

où $\delta U(T)$ est la solution au temps $t=T$ de l'équation linéarisée de Schrödinger

$$\left\{\begin{array}{rcl}\imath\delta\dot{U}(t)&=&(H_0+\varepsilon(t)\mu)\,\delta U(t)+(\delta H_0+\varepsilon(t)\delta\mu)\,U(t),\\ \delta U(0)&=&0,\end{array}\right. \tag{4.5}$$

le propagateur fondamental U étant solution de l'équation (4.1). Nous allons prouver que φ est une application surjective en utilisant le fait que $\mathrm{d}\varphi$ l'est aussi, ce que nous présentons dans le théorème suivant.

Théorème 4.3.1. *Soit $U_{cible}\in\mathcal{U}$, tel qu'il existe un couple $(H_0,\mu)\in\mathcal{S}^2$ vérifiant $\varphi\,(H_0,\mu)=U_{cible}$. Supposons que $\mathrm{d}\varphi\,(H_0,\mu)$ est une application surjective, c'est-à-dire que*

$$\forall V'\in\mathcal{A}_{H_0,\mu},\exists\,(\delta H_0,\delta\mu)\in\mathcal{S}\times\mathcal{S},\quad \mathrm{d}\varphi\,(H_0,\mu)\,(\delta H_0,\delta\mu)=V'$$

alors φ est localement surjective d'un voisinage de (H_0,μ) sur un voisinage de U_{cible}.

Cette stratégie est motivée par le résultat suivant.

Théorème 4.3.2 (Inversion généralisée sur les variétés)**.** *Soit X un espace de Banach et Y une variété différentiable de dimension finie. Nous considérons une fonction continue F d'un voisinage $V\subset X$ de $x_0\in X$ vers la variété Y, avec $F(x_0)=f$ qui est différentiable en x_0 et tel que la différentielle $DF(x_0)$ est surjective sur le plan tangent Tf_0 à Y en f_0. Alors, il existe un voisinage W de f_0 dans Y et $\beta>0$ tel que pour tout $f\in W$ il existe $x\in V$, avec $\|x-x_0\|\leq\beta$ et $F(x)=f$.*

Ce théorème est un résultat classique correspondant à une version du théorème d'inversion locale dédiée à la surjectivité. Pour plus d'informations, on se référera à [2]. Le théorème 4.3.1 nous invite à étudier la surjectivité de $\mathrm{d}\varphi$, c'est-à-dire à étudier la surjectivité au point $(H_0,\mu)\in\mathcal{S}^2$, pour tout $V'\in\mathcal{A}_{H_0,\mu}$ de l'équation linéaire

$$\mathrm{d}\varphi\,(H_0,\mu)\,(\delta H_0,\delta\mu)=V', \tag{4.6}$$

où $(\delta H_0,\delta\mu)\in\mathcal{S}^2$ est le couple solution. Avant de présenter le résultat annoncé en début de section, nous proposons une reformulation du problème de surjectivité de la différentielle. Ceci repose sur une formulation intégrale explicite de $\delta U(T)$.

Proposition 4.3.3. *Pour tout $V'\in\mathcal{A}_{H_0,\mu}$, montrer que (4.6) admet une solution est équivalent à montrer que pour tout $V\in\mathcal{H}$, l'équation*

$$\int_0^T U(t)^*(\delta H_0+\varepsilon(t)\delta\mu)U(t)\,\mathrm{d}t=V, \tag{4.7}$$

possède une solution ; les matrices V' et V étant reliées par $V=\imath U(T)^*V'$.

Démonstration. Tout d'abord, pour tout couple $(H_0, \mu) \in \mathcal{S}^2$ et $(\delta H_0, \delta\mu) \in \mathcal{S}^2$, on a que

$$\varphi\left(H_0,\mu\right)^* \mathrm{d}\varphi\left(H_0,\mu\right)\left(\delta H_0,\delta\mu\right) = U(T)^*\delta U(T) = \int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\left(U(t)^*\delta U(t)\right)\,\mathrm{d}t.$$

Un simple calcul permet de montrer que

$$\forall t \in [0,T], \quad \frac{\mathrm{d}}{\mathrm{d}t}\left(U(t)^*\delta U(t)\right) = -\imath U(t)^*\left(\delta H_0 + \varepsilon(t)\delta\mu\right)U(t)$$

et ainsi d'obtenir l'expression

$$V = \int_0^T U(t)^*\left(\delta H_0 + \varepsilon(t)\delta\mu\right)U(t)\,\mathrm{d}t,$$

où l'on a posé $V := \imath U(T)^*\delta U(T)$. Comme le second membre de l'égalité est hermitien, il en va de même pour V. Il existe donc une bijection entre $V' := \delta U(T) \in \mathcal{A}_{H_0,\mu}$ et $V \in \mathcal{H}$; ce qui nous permet de conclure. □

Cette reformulation de l'équation (4.6) est l'étape centrale du résultat et constitue la brique essentielle de l'algorithme de résolution locale présenté dans la section suivante.

4.3.2 Vectorisation du problème

Afin de ramener le système matriciel (4.7) à la résolution d'un système linéaire, nous introduisons des outils algébriques liés à la notion de vectorisation. La vectorisation d'une matrice consiste à concaténer les colonnes d'une matrice $A \in \mathbb{C}^{m,n}$, permettant ainsi d'obtenir un vecteur

$$\mathrm{vec}(A) = {}^t\left(a_{1,1},\ldots,a_{m,1},a_{1,2},\ldots,a_{m,2},\ldots,a_{1,n},\ldots,a_{m,n}\right)$$

appartenant à $\mathbb{C}^{m,n}$. Une fois vectorisée, l'équation (4.7) devient

$$\int_0^T W(t)^*\mathrm{d}t\ \mathrm{vec}\left(\delta H_0\right) + \int_0^T \varepsilon(t)W(t)^*\mathrm{d}t\ \mathrm{vec}\left(\delta\mu\right) = \mathrm{vec}\left(V\right), \tag{4.8}$$

où la trajectoire de matrices unitaires $t \mapsto W(t) := \overline{U\left(t\right)} \otimes U\left(t\right)$ est le propagateur fondamental de l'équation de Schrödinger tensorisée suivante

$$\imath\dot{W}(t) = \left(\tilde{H}_0 + \varepsilon(t)\tilde{\mu}\right)W(t), \tag{4.9}$$

avec $\tilde{H}_0 := (-H_0) \oplus H_0$ et $\tilde{\mu} := (-\mu) \oplus \mu$. Prenant la partie réelle et imaginaire de (4.9), on obtient le système

$$\begin{pmatrix} \Re(\mathrm{vec}\left(V\right)) \\ \mathrm{Im}\left(\mathrm{vec}\left(V\right)\right) \end{pmatrix} = \begin{pmatrix} \Re\left(\int_0^T W(t)^*\mathrm{d}t\right) & \Re\left(\int_0^T \varepsilon(t)W(t)^*\mathrm{d}t\right) \\ \mathrm{Im}\left(\int_0^T W(t)^*\mathrm{d}t\right) & \mathrm{Im}\left(\int_0^T \varepsilon(t)W(t)^*\mathrm{d}t\right) \end{pmatrix} \begin{pmatrix} \mathrm{vec}\left(\delta H_0\right) \\ \mathrm{vec}\left(\delta\mu\right) \end{pmatrix}, \tag{4.10}$$

de taille N_d^4. Précisons que la matrice $\int_0^T W(t)^*\mathrm{d}t$ est hermitienne, ainsi les matrices $\Re\left(\int_0^T W(t)^*\mathrm{d}t\right)$ et $\mathrm{Im}\left(\int_0^T W(t)^*\mathrm{d}t\right)$ sont respectivement symétrique et antisymétrique réelles. La nature des autres blocs du système (4.10) est conditionnée par le choix du champ ε.

Il est possible de réduire la taille du système (4.10), en profitant de la symétrie de δH_0, $\delta\mu$ et $\Re(V)$ et de l'anti-symétrie de $\mathrm{Im}\left(V\right)$. En ne considérant que la partie triangulaire inférieure des matrices précédentes, nous nous ramenons au système, de taille $N_d^2 + N_d$, suivant

$$\begin{pmatrix} \Re(\mathrm{v}\left(V\right)) \\ \mathrm{Im}\left(\mathrm{v}\left(V\right)\right) \end{pmatrix} = \begin{pmatrix} L\Re\left(\int_0^T W(t)^*\mathrm{d}t\right)D & L\Re\left(\int_0^T \varepsilon(t)W(t)^*\mathrm{d}t\right)D \\ L\mathrm{Im}\left(\int_0^T W(t)^*\mathrm{d}t\right)D & L\mathrm{Im}\left(\int_0^T \varepsilon(t)W(t)^*\mathrm{d}t\right)D \end{pmatrix} \begin{pmatrix} \mathrm{v}\left(\delta H_0\right) \\ \mathrm{v}\left(\delta\mu\right) \end{pmatrix}. \tag{4.11}$$

La notation v représente la vectorisation de la partie triangulaire inférieure d'une matrice A appartenant à $\mathbb{C}^{n,n}$,

$$i.e. \quad \mathrm{v}(A) = {}^t\left(a_{1,1}, \ldots, a_{n,1}, a_{2,2}, \ldots, a_{n,2}, \ldots, a_{n-1,n-1}, a_{n,n-1}, a_{n,n}\right) \in \mathbb{C}^{\frac{n(n+1)}{2}}.$$

Les matrices d'élimination L et d'expansion D permettent de relier les vectorisations vec et v, d'une matrice A, dès lors qu'elle est symétrique réelle, c'est-à-dire

$$\forall A \in \mathcal{S}, \quad L\mathrm{vec}(A) = \mathrm{v}(A) \text{ et } D\mathrm{v}(A) = \mathrm{vec}(A).$$

Pour plus de détails sur les outils liés à la vectorisation voir [9] et [10].

Il est encore possible de réduire le système (4.11), en éliminant la partie diagonale de $\delta\mu$. Ceci est pertinent dès lors que l'on cherche $\delta\mu$ dans $\mathcal{S}_0$. Le système résultant est alors d'ordre N_d^2. Dans ce cas nous avons fait correspondre l'ensemble où nous recherchons les solutions, à savoir $\mathcal{S} \times \mathcal{S}_0$, avec l'ensemble $\mathcal{H}$ où nous choisissons le second membre. Ceci nous permet d'expliciter un inverse dans le cas où le système précédent est de rang maximum. L'intérêt mathématique du choix du sous-ensemble $\mathcal{S} \times \mathcal{S}_0$ apparaît clairement ici.

4.3.3 Résultats locaux

On se donne dans cette sous-section une cible $U_{\text{cible}} \in \mathcal{U}/\mathcal{O}$. Ceci nous permet de définir $V_0 \in \mathcal{O}$, comme étant la matrice qui diagonalise la matrice U_{cible} de la manière suivante

$$U_{\text{cible}} = {}^tV_0 e^{i\Lambda} V_0,$$

où Λ est la matrice diagonale dont les coefficients diagonaux réels sont notés $\lambda_a \in]-\pi;\pi]$ et ceci pour tout $a \in [\![1; N_d]\!]$. L'application φ étant surjective par la proposition 4.2.2, on choisit $H_0 \in \mathcal{S}$ et $\mu \in \mathcal{S}$ de la manière suivante

$$U_{\text{cible}} = \varphi\left(H_0 := -\frac{1}{T}{}^tV_0\Lambda V_0, \mu := 0\right).$$

On peut donc considérer la trajectoire de matrice unitaire U définie par

$$\forall t \in [0;T], \quad U(t) = {}^tV_0 e^{i\Lambda\frac{t}{T}} V_0$$

qui vérifie bien la condition $\varphi(H_0, \mu) = U(t = T) = U_{\text{cible}}$.

Théorème 4.3.4. *L'application φ est localement surjective d'un voisinage de $\left(H_0 = -\frac{1}{T}{}^tV_0\Lambda V_0, \mu = 0\right)$ sur un voisinage de U_{cible}.*

Démonstration. Pour prouver ce résultat, il nous suffit d'appliquer le théorème 4.3.1, c'est-à-dire de montrer la surjectivité de l'application $d\varphi(H_0, \mu)$. Pour tout $V' \in \mathcal{A}_{H_0,\mu}$, l'identification d'un couple $(\delta H_0, \delta\mu) \in \mathcal{S}^2$ équivaut d'après la proposition 4.3.3 à l'identification d'un couple solution de l'équation (4.7) pour un $V \in \mathcal{H}$ correspondant à V'. Par vectorisation, nous nous ramenons à (4.8)

$$\int_0^T W(t)^* \, \mathrm{d}t \, \mathrm{vec}(\delta H_0) + \int_0^T \varepsilon(t) W(t)^* \, \mathrm{d}t \, \mathrm{vec}(\delta\mu) = \mathrm{vec}(V).$$

Dès lors que la matrice $\int_0^T W(t)^* \, \mathrm{d}t$ est inversible, il est possible d'identifier un couple $(\delta H_0, \delta\mu)$, prouvant ainsi la surjectivité de $d\varphi(H_0, \mu)$ sur le plan tangent et par suite la surjectivité locale de φ.

Étudions donc l'inversibilité de la matrice $\int_0^T W(t)^* \, \mathrm{d}t$. Comme

$$\int_0^T W(t)^* \, \mathrm{d}t = \left({}^tV_0 \otimes {}^tV_0\right) \int_0^T e^{\imath\Lambda\frac{t}{T}} \otimes e^{-\imath\Lambda\frac{t}{T}} \, \mathrm{d}t \, (V_0 \otimes V_0)$$

et étant donné que $V_0 \otimes V_0 \in \mathcal{O}$ et que ${}^tV_0 \otimes {}^tV_0 = {}^t(V_0 \otimes V_0)$, on en déduit que l'inversibilité de $\int_0^T W(t)^* \, \mathrm{d}t$ repose sur celle de $\int_0^T e^{\imath\Lambda\frac{t}{T}} \otimes e^{-\imath\Lambda\frac{t}{T}} \, \mathrm{d}t$. Or

$$\int_0^T e^{\imath\Lambda\frac{t}{T}} \otimes e^{-\imath\Lambda\frac{t}{T}} \, \mathrm{d}t = \int_0^T e^{\imath\delta\Lambda\frac{t}{T}} \, \mathrm{d}t,$$

où

$$\delta\Lambda := \Lambda \oplus (-\Lambda) = \mathrm{diag}\left(\delta\lambda_{1,1}, \delta\lambda_{1,2}, \ldots, \delta\lambda_{1,N_d}, \delta\lambda_{2,1}, \ldots, \delta\lambda_{N_d,N_d}\right),$$

avec pour tout $(a,b) \in [\![1; N_d]\!]^2$, $\delta\lambda_{a,b} = \lambda_a - \lambda_b$. Un simple calcul permet de vérifier que

$$\forall (a,b) \in [\![1; N_d]\!]^2, \quad \int_0^T e^{\imath\delta\lambda_{a,b}\frac{t}{T}} \, \mathrm{d}t = T\mathrm{sinc}\left(\frac{\delta\lambda_{a,b}}{2}\right) e^{\imath\frac{\delta\lambda_{a,b}}{2}}.$$

Enfin, comme $T > 0$ et que $\delta\lambda_{a,b} \in]-2\pi; 2\pi[$, on en déduit que le module $T\left|\mathrm{sinc}\left(\frac{\delta\lambda_{a,b}}{2}\right)\right|$ est différent de zéro, ce qui entraîne l'inversibilité de la matrice $\int_0^T e^{\imath\delta\Lambda\frac{t}{T}} \, \mathrm{d}t$ et notre résultat par la même occasion. □

Pour aller plus loin nous explicitons, sous une certaine condition, un inverse de $\mathrm{d}\varphi\,(H_0, 0)$.

Théorème 4.3.5. *Si l'on suppose que pour tout $(a,b) \in [\![1; N_d]\!]^2$, tel que $a \neq b$, on a*

$$\hat{\varepsilon}^i_{a,b} \;:=\; \mathrm{Im}\left(\int_0^T \varepsilon(t) e^{\imath\frac{\delta\lambda_{a,b}}{T}(t-\frac{T}{2})} dt\right) \neq 0, \tag{4.12}$$

alors l'inverse de l'application $\mathrm{d}\varphi(H_0, 0)$ est donnée par

$$\begin{array}{rcl} \mathcal{A}_{H_0,\mu} & \to & \mathcal{S} \times \mathcal{S}, \\ V' & \mapsto & (\delta H_0, \delta\mu)\,, \end{array}$$

où les matrices δH_0 et $\delta\mu$ sont respectivement définies par

$$\delta H_0 := {}^tV_0 \delta\tilde{H}_0 V_0 \quad et \quad \delta\mu := {}^tV_0 \delta\tilde{\mu} V_0,$$

et les coefficients $h_{a,b}$ et $m_{a,b}$ des matrices $\delta\tilde{H}_0$ et $\delta\tilde{\mu}$ sont donnés par

$$\left\{\begin{array}{rcll} m_{a,b} & = & \dfrac{\cos\left(\frac{\delta\lambda_{a,b}}{2}\right) \mathrm{Im}\, v_{a,b} - \sin\left(\frac{\delta\lambda_{a,b}}{2}\right) \Re v_{a,b}}{\hat{\varepsilon}^i_{a,b}}, & \\ h_{a,b} & = & \dfrac{\cos\left(\frac{\delta\lambda_{a,b}}{2}\right) \Re v_{a,b} + \sin\left(\frac{\delta\lambda_{a,b}}{2}\right) \mathrm{Im}\, v_{a,b} - \hat{\varepsilon}^r_{a,b}\, m_{a,b}}{T\mathrm{sinc}\left(\frac{\delta\lambda_{a,b}}{2}\right)}, & si\ a \neq b, \\ m_{a,a} & \in & \mathbb{R}, & \\ h_{a,a} & = & \frac{1}{T}\left(v_{a,a} - \int_0^T \varepsilon(t)\, \mathrm{d}t\, m_{a,a}\right), & si\ a = b. \end{array}\right. \tag{4.13}$$

Ici, pour tout $(a,b) \in [\![1; N_d]\!]^2$, les coefficients $v_{a,b}$ sont ceux de la matrice $\widetilde{V} := V_0 V^t V_0$ et

$$\hat{\varepsilon}^r_{a,b} := \Re\left(\int_0^T \varepsilon(t) e^{\imath\frac{\delta\lambda_{a,b}}{T}(t-\frac{T}{2})} dt\right).$$

Démonstration. Pour prouver ce théorème, nous poursuivons la démonstration du théorème 4.3.4 précédent à partir de l'équation (4.8), dont on rappelle ici l'expression

$$\int_0^T W(t)^* \mathrm{d}t \ \mathrm{vec}\,(\delta H_0) + \int_0^T \varepsilon(t) W(t)^* \mathrm{d}t \ \mathrm{vec}\,(\delta\mu) = \mathrm{vec}\,(V)\,.$$

Comme on a montré, dans la démonstration du théorème 4.3.4, que

$$\int_0^T W(t)^* \ \mathrm{d}t = {}^t\,(V_0 \otimes V_0) \left(T\mathrm{sinc}\left(\frac{\delta\Lambda}{2}\right) e^{\imath\frac{\delta\Lambda}{2}} \right) (V_0 \otimes V_0)\,,$$

on a de la même façon

$$\int_0^T \varepsilon(t) W(t)^* \ \mathrm{d}t = {}^t\,(V_0 \otimes V_0) \int_0^T \varepsilon(t) e^{\imath\frac{\delta\Lambda}{T}t} \ \mathrm{d}t\,(V_0 \otimes V_0)\,,$$

on obtient ainsi que l'équation (4.8) est équivalente à

$$T\mathrm{sinc}\left(\frac{\delta\Lambda}{2}\right) e^{\imath\frac{\delta\Lambda}{2}} \mathrm{vec}\left(\widetilde{\delta H_0}\right) + \int_0^T \varepsilon(t) e^{\imath\frac{\delta\Lambda}{T}t} \ \mathrm{d}t \ \mathrm{vec}\left(\widetilde{\delta\mu}\right) = \mathrm{vec}\left(\widetilde{V}\right), \tag{4.14}$$

où l'on a posé

$$\widetilde{\delta H_0} := V_0 \delta H_0 {}^t V_0 \in \mathcal{S}, \quad \widetilde{\delta\mu} := V_0 \delta\mu {}^t V_0 \in \mathcal{S} \quad \text{et} \quad \widetilde{V} := V_0 V {}^t V_0 \in \mathcal{H}.$$

Notons respectivement $v_{a,b}$, $h_{a,b}$ et $m_{a,b}$, avec $(a,b) \in [\![1; N_d]\!]^2$ les coefficients des matrices $\widetilde{V}$, $\widetilde{\delta H_0}$ et $\widetilde{\delta\mu}$. Déterminer les coefficients des matrices $\widetilde{V}$, $\widetilde{\delta H_0}$ et $\widetilde{\delta\mu}$ à partir de l'égalité (4.14) donne lieu à deux cas.

Cas où $a = b$: comme

$$(4.14) \Rightarrow T\, h_{a,a} + \int_0^T \varepsilon(t) \ \mathrm{d}t \ m_{a,a} = v_{a,a},$$

on a par conséquent, quel que soit $m_{a,a} \in \mathbb{R}$, le résultat suivant

$$h_{a,a} = \frac{1}{T}\left(v_{a,a} - \int_0^T \varepsilon(t) \ \mathrm{d}t \ m_{a,a} \right).$$

Cas où $a \neq b$: comme

$$\begin{aligned} (4.14) \quad &\Rightarrow \quad T\mathrm{sinc}\left(\tfrac{\delta\lambda_{a,b}}{2}\right) e^{\imath\frac{\delta\lambda_{a,b}}{2}}\, h_{a,b} + \textstyle\int_0^T \varepsilon(t) e^{\imath\frac{\delta\lambda_{a,b}}{T}t} \ \mathrm{d}t \ m_{a,b} = v_{a,b}, \\ &\Rightarrow \quad T\mathrm{sinc}\left(\tfrac{\delta\lambda_{a,b}}{2}\right) h_{a,b} + \widehat{\varepsilon}\left(\tfrac{\delta\lambda_{a,b}}{T}\right) m_{a,b} = e^{-\imath\frac{\delta\lambda_{a,b}}{2}} v_{a,b}, \end{aligned}$$

où l'on a défini $\widehat{\varepsilon}\left(\frac{\delta\lambda_{a,b}}{T}\right) = \int_0^T \varepsilon(t) e^{\imath\frac{\delta\lambda_{a,b}}{T}(t-\frac{T}{2})} dt = \hat{\varepsilon}^r_{a,b} + i\hat{\varepsilon}^i_{a,b}$. En prenant la partie réelle et la partie imaginaire de la dernière égalité, on obtient le système suivant

$$\left\{ \begin{array}{rcl} T\mathrm{sinc}\left(\frac{\delta\lambda_{a,b}}{2}\right) h_{a,b} + \hat{\varepsilon}^r_{a,b}\, m_{a,b} &=& \cos\left(\frac{\delta\lambda_{a,b}}{2}\right) \Re v_{a,b} + \sin\left(\frac{\delta\lambda_{a,b}}{2}\right) \mathrm{Im}\, v_{a,b}, \\ \hat{\varepsilon}^i_{a,b}\, m_{a,b} &=& \cos\left(\frac{\delta\lambda_{a,b}}{2}\right) \mathrm{Im}\, v_{a,b} - \sin\left(\frac{\delta\lambda_{a,b}}{2}\right) \Re v_{a,b}, \end{array} \right.$$

On conclut en supposant que $\hat{\varepsilon}^i_{a,b} \neq 0$, ce qui nous donne le résultat

$$\left\{ \begin{array}{rcl} h_{a,b} &=& \dfrac{\cos\left(\frac{\delta\lambda_{a,b}}{2}\right) \Re v_{a,b} + \sin\left(\frac{\delta\lambda_{a,b}}{2}\right) \mathrm{Im}\, v_{a,b} - \hat{\varepsilon}^r_{a,b}\, m_{a,b}}{T\mathrm{sinc}\left(\frac{\delta\lambda_{a,b}}{2}\right)}, \\ m_{a,b} &=& \dfrac{\cos\left(\frac{\delta\lambda_{a,b}}{2}\right) \mathrm{Im}\, v_{a,b} - \sin\left(\frac{\delta\lambda_{a,b}}{2}\right) \Re v_{a,b}}{\hat{\varepsilon}^i_{a,b}}. \end{array} \right.$$

□

Ce théorème donne une première indication sur les conditions nécessaires à l'identification du couple (H_0, μ). La condition (4.12) est directement liée au champ laser. Il s'agit d'une condition de non-résonance pour commander le système. Poussons plus loin la description de la condition (4.12). Un simple calcul permet de montrer que le laser ε doit être choisi de telle sorte que la fonction $\varepsilon_{\not g,T/2}\left(\frac{T\cdot}{\delta\lambda_{a,b}}\right)$ ne soit pas orthogonale à la fonction sinus dans $L^2\left(\left(-\frac{\delta\lambda_{a,b}}{2};\frac{\delta\lambda_{a,b}}{2}\right);\mathbb{R}\right)$. Ici on a définit $\varepsilon_{\not g,T/2}$ de la manière suivante

$$\forall t \in \left[-\frac{T}{2};\frac{T}{2}\right], \quad \varepsilon_{\not g,T/2}(t) := \frac{\varepsilon\left(\frac{T}{2}+t\right)-\varepsilon\left(\frac{T}{2}-t\right)}{2}.$$

Enfin pour terminer cette section, nous présentons un corollaire du théorème précédent. Celui-ci donne une condition supplémentaire permettant de déterminer un couple solution appartenant à $\mathcal{S}\times\mathcal{S}_0$ et non à $\mathcal{S}\times\mathcal{S}$. Cette condition porte cette fois-ci sur la matrice V_0 opérant le changement de base diagonalisant la cible.

Corollaire 4.3.6. *On reprend les hypothèses et notations du théorème 4.3.5. Si l'on ajoute la condition*

$$\det\left({}^tV_0\cdot{}^tV_0\right)\neq 0, \tag{4.15}$$

alors l'inverse de l'application $\mathrm{d}\varphi(H_0,0)$ *est donnée par*

$$\begin{array}{rcl}\mathcal{A}_{H_0,\mu} & \to & \mathcal{S}\times\mathcal{S}_0,\\ V' & \mapsto & (\delta H_0,\delta\mu),\end{array}$$

où les coefficients de δH_0 *et* $\delta\mu_0$ *sont donnés par les formules (4.13). De plus, on a*

$$\mathrm{diag}\left(\widetilde{\delta\mu}\right) := \left({}^tV_0\cdot{}^tV_0\right)^{-1}K \tag{4.16}$$

où le vecteur K est défini de la manière suivante

$$\forall c\in[\![1;N_d]\!], \quad K_c := \sum_{(a,b)\in[\![1;N_d]\!]^2\setminus\{(a,b)|a=b\}} (V_0)_{a,c}\,(V_0)_{b,c}\,m_{a,b}$$

Démonstration. Comme on cherche un couple solution dans $\mathcal{S}\times\mathcal{S}_0$, on a par conséquent que $\mathrm{diag}\,(\delta\mu)=0$. En ne considérant que les termes diagonaux de l'égalité ${}^tV_0\widetilde{\delta\mu}V_0=\delta\mu$, on trouve aisément le résultat. □

4.4 Résolution locale du problème de contrôlabilité

Dans cette section, nous présentons un algorithme nous permettant de résoudre l'équation (4.4). La stratégie suivie est une adaptation directe des preuves et résultats précédents : nous considérons des approximations locales obtenues par une méthode de Newton. Dans notre approche, une étape cruciale consiste à obtenir une discrétisation temporelle appropriée de l'équation (4.1). Ceci fait l'objet de la première sous-section.

4.4.1 Discrétisation temporelle

Pour simuler numériquement l'équation (4.1), nous introduisons la discrétisation temporelle suivante : soit $N_T \in \mathbb{N}$, notons $\Delta T = \dfrac{T}{N_T}$ le pas et pour $n \in [\![0; N_T]\!]$ par U_n et ε_n l'approximation de $U(n\Delta T)$ et de $\varepsilon(n\Delta T)$. Afin de conserver le caractère unitaire de la trajectoire $t \mapsto U(t)$ au niveau discret, nous utilisons la méthode de la règle du trapèze, qui donne lieu au schéma de type Crank-Nicolson suivant

$$\imath \frac{U_{n+1} - U_n}{\Delta T} = (H_0 + \varepsilon_n \mu) \frac{U_{n+1} + U_n}{2}.$$

L'itération associée est donc

$$(\mathrm{Id} + L_n) U_{n+1} = (\mathrm{Id} - L_n) U_n,$$

où $L_n = \frac{\imath \Delta T}{2}(H_0 + \varepsilon_n \mu)$.

Détaillons maintenant les effets des variations respectives δH_0 et $\delta\mu$ de H_0 et μ sur la trajectoire discrétisée $(U_n)_{n \in [\![0;N_T]\!]}$. Nous avons

$$\begin{aligned}
(\mathrm{Id} + L_n)\delta U_{n+1} + \delta L_n U_{n+1} &= (\mathrm{Id} - L_n)\delta U_n - \delta L_n U_n, \\
\delta L_n (U_{n+1} + U_n) &= (\mathrm{Id} - L_n)\delta U_n - (\mathrm{Id} + L_n)\delta U_{n+1}, \\
(U_{n+1} + U_n)^* \delta L_n (U_{n+1} + U_n) &= -2(U_{n+1}^* \delta U_{n+1} - U_n^* \delta U_n),
\end{aligned}$$

où $\delta L_n = \frac{\imath \Delta T}{2}(\delta H_0 + \varepsilon_n \delta\mu)$, ce qui donne finalement

$$U_{n+1}^* \delta U_{n+1} - U_n^* \delta U_n = -\imath \Delta T U_{n+1/2}^* (\delta H_0 + \varepsilon_n \delta\mu) U_{n+1/2},$$

où $U_{n+1/2} = \dfrac{U_{n+1} + U_n}{2}$. Dès lors que la valeur initiale est fixée, on obtient ainsi

$$U_{N_T}^* \delta U_{N_T} = -\imath \Delta T \sum_{n=0}^{N_T - 1} U_{n+1/2}^* (\delta H_0 + \varepsilon_n \delta\mu) U_{n+1/2}. \tag{4.17}$$

Ce résultat peut être interprété comme une discrétisation de l'équation (4.7). Nous insistons sur le fait qu'un tel résultat semble spécifique à la méthode du trapèze. Il n'existe pas, à notre connaissance, d'autres méthodes numériques permettant d'obtenir une discrétisation de (4.7) où les variations δH_0 et $\delta\mu$ sont explicites.

4.4.2 Méthodes de point fixe

Au niveau discret, par abus de notation, nous conservons φ pour désigner l'application

$$\begin{aligned}
\varphi : \quad \mathcal{S} \times \mathcal{S} &\to \mathcal{U}, \\
(H_0, \mu) &\mapsto U_{N_T}.
\end{aligned}$$

Le problème de contrôlabilité 4.2.1 peut être envisagé, pour $U_{\text{cible}} \in \mathcal{U}$ donnée, comme étant la recherche de zéros de l'application non linéaire F, définie comme suit

$$\forall\, (H_0, \mu) \in \mathcal{S}^2, \quad F\,(H_0, \mu) := \varphi\,(H_0, \mu) - U_{\text{cible}}.$$

Nous présentons maintenant une méthode itérative pour déterminer des zéros de F.

Une méthode de Newton

Pour résoudre l'équation $F(H_0, \mu) = 0$, l'application de la méthode de Newton consiste en la relation de récurrence

$$\begin{aligned} \mathrm{d}F\left(H_0^{(k)}, \mu^{(k)}\right) \cdot \left(\delta H_0^{(k)}, \delta\mu^{(k)}\right) &= \mathrm{d}\varphi\left(H_0^{(k)}, \mu^{(k)}\right) \cdot \left(\delta H_0^{(k)}, \delta\mu^{(k)}\right), \\ &= -\left(\varphi\left(H_0^{(k)}, \mu^{(k)}\right) - U_{\text{cible}}\right), \end{aligned} \tag{4.18}$$

où k représente le numéro de l'itération et $\delta H_0^{(k)} = H_0^{(k+1)} - H_0^{(k)}$, $\delta\mu^{(k)} = \mu^{(k+1)} - \mu^{(k)}$.

Dans notre cas, il vient

$$\delta U_{N_T}^{(k)} = U_{\text{cible}} - U_{N_T}^{(k)}.$$

En utilisant (4.17), on peut réécrire cette dernière équation de la manière suivante

$$\Delta T \sum_{n=0}^{N_T-1} \left(U_{n+1/2}^{(k)}\right)^* \left(\delta H_0^{(k)} + \varepsilon_n \delta\mu^{(k)}\right) U_{n+1/2}^{(k)} = \imath\left(\left(U_{N_T}^{(k)}\right)^* U_{\text{cible}} - \mathrm{Id}\right),$$

dans laquelle nous rappelons que les inconnues sont $\delta H_0^{(k)}$ et $\delta\mu^{(k)}$. Cette équation n'a généralement pas de solution, puisque le membre de gauche appartient à $\mathcal{H}$, ce qui n'est pas le cas pour le membre de droite. Afin de résoudre ce problème, nous remplaçons $\imath\left(\left(U_{N_T}^{(k)}\right)^* U_{\text{cible}} - \mathrm{Id}\right)$ par une approximation du premier ordre $S^{(k)} \in \mathcal{H}$. Deux choix sont ici possibles

$$\exp\left(-\imath S^{(k)}\right) := \left(U_{N_T}^{(k)}\right)^* U_{\text{cible}} \tag{4.19}$$

$$S^{(k)} := \imath\frac{\left(U_{N_T}^{(k)}\right)^* U_{\text{cible}} - U_{target}^* U_{N_T}^{(k)}}{2}. \tag{4.20}$$

Dans nos tests numériques, le même comportement a été observé quel que soit le choix effectué. Nous utiliserons généralement la deuxième alternative, qui peut être vue comme une projection de $S^{(k)}$ sur $\mathcal{H}$.

Remarque 4.4.1. *La méthode précédente peut être quelque peu simplifiée pour obtenir une procédure dans laquelle les matrices sont assemblées de temps en temps au cours des itérations. Au lieu de mettre à jour, à chaque itération, le couple (H_0, μ) dans le terme $\mathrm{d}\varphi(H_0, \mu)$ de la formule (4.18), on peut garder une approximation constante $\left(H_0^{ref}, \mu^{ref}\right)$ de la solution. On note $\left(U_n^{ref}\right)_{n\in[\![0;N_T]\!]}$ la suite d'états correspondante. L'itération se lit alors*

$$\Delta T \sum_{n=0}^{N_T-1} \left(U_{n+1/2}^{ref}\right)^* \left(\delta H_0^{(k)} + \varepsilon_n \delta\mu^{(k)}\right) U_{n+1/2}^{ref} = S^{(k)},$$

où $S^{(k)}$ est définie dans la section précédente, voir (4.19) ou (4.20). Cette technique de réduction de complexité est classique pour la méthode de Newton.

Implémentation de l'algorithme itératif

La méthode précédente nécessite la résolution de systèmes linéaires qui ne sont pas donnés explicitement dans nos formulations. Pour combler cette lacune, nous expliquons ici comment assembler les matrices, c'est-à-dire comment réécrire l'équation

$$\Delta T \sum_{n=0}^{N_T-1} U_{n+1/2}^* \left(\delta H_0 + \varepsilon_n \delta\mu\right) U_{n+1/2} = S,$$

en terme d'un système linéaire. Une première étape consiste à réécrire la dernière équation de la manière suivante

$$\Delta T \left(\sum_{n=0}^{N_T-1} M_{U_{n+1/2}} \right) \mathrm{vec}\left(\delta H_0\right) + \Delta T \left(\sum_{n=0}^{N_T-1} \varepsilon_n M_{U_{n+1/2}} \right) \mathrm{vec}\left(\delta \mu\right) = \mathrm{vec}\left(S\right), \tag{4.21}$$

avec

$$M_{U_{n+1/2}} = {}^{t}U_{n+1/2} \otimes U^{*}_{n+1/2}.$$

Ici nous avons utilisé la même technique que celle employée pour passer de l'équation (4.7) à l'équation (4.8), à savoir la vectorisation présentée dans la sous-section 4.3.2. On sépare ensuite les parties réelle et imaginaire de l'équation (4.21). Pour terminer on utilise les matrices d'élimination et d'expansion, afin de réduire le système, en éliminant les coefficients de δH_0 et $\Re S$ situés strictement au dessus de leur diagonale et ceux de $\delta\mu$ et $\mathrm{Im}\, S$ situés sur ou au-dessus de leur diagonale. Finalement, le système résultant est de taille N_d^2.

4.5 Méthodes de continuations pour des résultats globaux

La procédure proposée dans la précédente section ne converge que si l'initialisation est choisie suffisamment proche de la solution. Nous présentons à présent plusieurs techniques de continuation qui permettent de globaliser notre méthode de Newton, c'est-à-dire d'étendre son champ d'application.

4.5.1 Généralités sur la méthode de continuation

La solution du problème 4.2.1 ne nous est, en général, pas accessible, hormis dans certains cas particuliers comme ceux présentés dans la sous-section 4.2.2. L'idée de la méthode de continuation, est de ne plus uniquement considérer le problème

$$F\left(H_0, \mu\right) = 0,$$

mais plutôt une famille de problèmes

$$G\left(\left(H_0, \mu\right), \theta\right) = 0,$$

paramétrée par un réel θ prenant ses valeurs dans l'intervalle $[0;1]$. Ce paramètre est appelé *paramètre de continuation*. Ce dernier relie de façon continue un problème que l'on sait résoudre

$$G\left(\left(H_0, \mu\right), \theta = 0\right) = 0;$$

au problème d'intérêt

$$G\left(\left(H_0, \mu\right), \theta = 1\right) := F\left(H_0, \mu\right) = 0.$$

Se ramener à une situation favorable, quand $\theta = 0$, revient à déformer le problème à résoudre et ceci en relâchant une donnée. La sous-section suivante présente une continuation sur le champ ε ; tandis que la troisième introduit une modification de la cible U_{cible}.

La mise en œuvre de la méthode est très simple. Soit $\delta\theta := \frac{1}{N_\theta}$ le pas de la méthode de continuation, avec $N_\theta \in \mathbb{N}^*$. Nous conservons toutes les notations introduites précédemment dans ce chapitre. On commence par résoudre le problème de départ

$$G\left(\left(H_0, \mu\right), 0\right) = 0,$$

dont la solution est notée $\left(H_0^{(0)}, \mu^{(0)}\right)$. Cette solution constitue le couple permettant l'initialisation de la procédure. Une méthode de point fixe, analogue à celle décrite dans la sous-section 4.4.2, peut alors être appliquée afin de résoudre le problème

$$G\left(\left(H_0, \mu\right), \theta = \delta\theta\right) = 0.$$

L'algorithme consiste ensuite à répéter cette procédure en résolvant itérativement, pour $k \in [\![2; N_\theta]\!]$, le problème

$$G\left(\left(H_0, \mu\right), \theta = k\delta\theta\right) = 0,$$

avec $\left(H_0^{(k-1)}, \mu^{(k-1)}\right)$ comme estimation initiale.

4.5.2 Continuations par déformation du champ

Comme mentionné précédemment, plusieurs méthodes existent pour résoudre le problème de contrôlabilité lorsque le terme ε dans l'équation (4.1) est inconnu et que H_0 et μ sont données. Pour plus d'informations sur le sujet, on se référera à [7, 8, 12]. Basée sur ce constat, la méthode de continuation que nous proposons est la suivante : étant donné un couple initial $\left(H_0^{(0)}, \mu^{(0)}\right)$, il faut déterminer un champ ε^0 de telle sorte que $U_{N_T}^{(0)}$, l'état final associé au couple $\left(H_0^{(0)}, \mu^{(0)}\right)$, approche correctement la cible U_{cible}. Par la suite, nous définissons un champ interpolé $\varepsilon^\theta = (1-\theta)\varepsilon^0 + \theta\varepsilon$. Notre algorithme consiste alors à résoudre itérativement le problème de contrôle par opérateur associé au champ $\varepsilon^{k\delta\theta}$ et avec $\left(H_0^{(k-1)}, \mu^{(k-1)}\right)$ comme estimation initiale. Précisons que l'initialisation $\left(H_0^{(0)}, \mu^{(0)}\right)$ est ici choisie arbitrairement. Cette approche a été publiée dans [6].

4.5.3 Continuations par déformation de la cible

Une alternative à la méthode précédente consiste à réaliser une continuation en effectuant la transformation sur l'état cible. Ceci revient à considérer le problème 4.2.1, en prenant une cible paramétrée U_{cible}^θ plutôt qu'une cible fixe. Plusieurs paramétrages $\theta \mapsto U^\theta$ sont possibles, nous en présentons trois dans la suite de la section. Comme la procédure itérative ne change pas et celle-ci ayant déjà été décrite précédemment, nous n'y revenons plus ici. Nous nous contentons de présenter les paramétrages de cible considérés.

Décomposition logarithmique

Cette procédure a été publiée dans [11]. Elle repose sur le paramétrage suivant

$$U_{\text{cible}}^\theta = \exp(\imath\varphi_S(\theta)S + \varphi_A(\theta)A),$$

où (S, A) est un couple composé d'une matrice symétrique et d'une matrice anti-symétriques associé à l'état cible U_{cible} du problème original par $U_{\text{cible}} = \exp(\imath S + A)$. Les fonctions $\varphi_S(\theta)$ et $\varphi_A(\theta)$ sont continues et seulement soumises aux contraintes suivantes

$$\varphi_S(0) = 1,\ \varphi_A(0) = 0,\ \varphi_S(1) = 1 \text{ et } \varphi_A(1) = 1.$$

Nous nous appuyons ici sur les assertions 1 et 2 des propriétés 4.1.1. La procédure d'extraction du couple (S, A) à partir de U_{cible} est très simple. En Octave, elle peut être codée de la manière suivante

```
A = real(logm(Ucible));A=.5*(A-A');
S = imag(logm(Ucible));S=.5*(S+S');
```

Décomposition en deux exponentielles de matrices symétriques

Cette fois-ci, on considère le paramétrage suivant

$$U_{\text{cible}}^{\theta} = \exp(\imath S)\exp(\imath\theta\widetilde{S}),$$

où $\left(S, \widetilde{S}\right)$ est un couple de matrices symétriques associé à l'état cible U_{cible} du problème original par $U_{\text{cible}} = \exp(\imath S)\exp(\imath\widetilde{S})$. Nous nous appuyons ici sur la décomposition des matrices unitaires proposée dans le lemme 4.2.4. La procédure d'extraction du couple $\left(S, \widetilde{S}\right)$ à partir de U_{cible} peut être codée en Octave de la manière suivante

```
% décomposition en Ucible=exp(i.S1)exp(i.S2)
iH = logm(Ucible);
[P,iD] = eig(iH);%norm(P*iD*P'-iH)=1e-15;
V = conj(P)*P';
% S1 : matrice symétrique réelle entrant dans la décomposition
S1 = imag(logm(conj(V)));
S1 = (S1 + transpose(S1))/2;
% S2 : matrice symétrique réelle entrant dans la décomposition
S2 = imag(logm(V*Ucible));
S2 = (S2 + transpose(S2))/2;
```

Dans le code, le couple de matrices $\left(S, \widetilde{S}\right)$ est noté $(S1, S2)$.

Décomposition par changement de base unitaire

Le dernier paramétrage que nous proposons est donné par

$$U_{\text{cible}}^{\theta} = \exp(\theta W)\exp(\imath S)\exp(\theta W^*),$$

où (W, S) est un couple de matrices hermitiennes et symétriques associé à l'état cible U_{cible} du problème original par $U_{\text{cible}} = \exp(W)\exp(\imath S)\exp(W^*)$. Nous nous appuyons ici sur l'assertion 4 de la proposition 4.1.1. La procédure d'extraction du couple (A, S) à partir de U_{cible} peut être codée en Octave de la manière suivante

```
% décomposition en Ucible=exp(W)exp(iS)exp(W')
[C,D] = eig(Ucible);
W = logm(C);W=.5*(W-W');
S = imag(logm(D));
```

Notons qu'ici, la matrice S n'est pas seulement symétrique, mais diagonale. D'autres choix de matrices S peuvent être considérés. Notons également que l'on peut envisager d'autres décompositions de ce type en changeant S par une matrice qui lui est semblable par changement de base orthogonale.

De nombreuses variantes de ces trois méthodes peuvent être envisagées. Ainsi un grand choix de procédures existe. L'intérêt de toutes ces décompositions est de fournir une initialisation explicite de la méthode de continuation. En effet, pour $\theta = 0$, nous voyons que dans les trois cas, une solution explicite existe au problème 4.2.1. En effet, il suffit de prendre par exemple $(H_0, \mu) := (-S/T, 0)$, comme mentionné dans la démonstration de la proposition 4.2.2. C'est

l'avantage principal de cette approche par rapport à la méthode de continuation par interpolation de champs proposée à la section précédente. Cette dernière nécessitait le calcul d'un champ ε^0 dans une étape préliminaire. Dans notre méthode de continuation, par déformation de la cible, nous n'avons pas besoin de phase de pré-calculs. De plus, si le pas de déformation est suffisamment petit, le résultat d'existence locale 4.3.4 garantit que la première déformation à partir de la solution explicite de l'initialisation possède bien une solution.

Enfin, retenons que l'avantage pratique des deux dernières solutions par rapport à la première est la possibilité de différencier exactement par rapport à la variable de continuation θ. Nous profitions de ce résultat dans la dernière approche proposée dans la sous-section 4.6.3.

4.6 Résultats numériques

Dans cette dernière section, nous présentons quelques résultats numériques obtenus à partir des algorithmes présentés dans les sections précédentes. Nous utiliserons $\varepsilon(t) := \sin(t)$, pour tout $t \in [0; T]$, comme champ laser dans l'équation (4.1), sauf en sous-section 4.6.2. Les autres données numériques sont $N_d = 5$, $T_0 = 10$, $N_T = 10^2$, $T = 2\pi T_0$ et $\Delta T = \frac{T}{N_T}$.

4.6.1 Méthode de Newton

Nous testons en premier lieu l'algorithme de Newton, en choisissant pour cela aléatoirement un couple (H_0, μ) dont les coefficients sont compris dans l'intervalle $[-1; 1]$, ce qui nous permet de déterminer l'état final U_{N_T} correspondant. Ensuite, nous initions la procédure de Newton avec une initialisation $(H_0 + \Delta H_0, \mu + \Delta\mu)$ où le couple $(\delta H_0, \delta\mu)$ est aussi choisi aléatoirement. Un exemple est présenté dans le tableau suivant.

Iterations	$\log_{10}\left(\left\|H_0^{(k)} - H_0\right\|_F\right)$	$\log_{10}\left(\left\|\mu^{(k)} - \mu\right\|_F\right)$
1	$-1,579029$	$-1,358376$
2	$-3,003599$	$-2,865026$
3	$-4,339497$	$-4,122528$
4	$-8,234980$	$-8,179398$
5	$-13,963299$	$-14,029020$
6	$-14,022486$	$-14,131066$

Nous retrouvons le couple (H_0, μ) en partant d'un couple tiré aléatoirement avec une erreur de 10%. La convergence numérique est obtenue après six itérations. Notons aussi que la convergence quadratique de la méthode de Newton est bien observée.

4.6.2 Méthode de continuation par déformation du champ

Dans un second test, nous utilisons la méthode de continuation présentée dans la sous-section 4.5.2 pour résoudre numériquement un problème qu'il n'est pas possible de traiter avec la méthode de la sous-section 4.4.2. Soit une cible U_{cible} obtenue en appliquant le champ laser ε au système et un couple (H_0, μ) choisi au hasard. Nous cherchons alors les matrices H_0' et μ' permettant de résoudre le problème de contrôlabilité associé au champ laser $\cos(3t)$ et à la cible U_{cible}. L'application directe de l'algorithme de Newton de la sous-section 4.4.2 ne fonctionne pas dans ce cas. Par conséquent l'algorithme ne converge pas. La méthode de continuation nous permet de résoudre ce problème numériquement. En prenant un pas $\delta\theta = \frac{1}{4}$ et en effectuant dix itérations de la méthode de Newton à chaque boucle. On obtient alors un couple (H_0', μ') satisfaisant. Cet exemple a été reproduit pour plusieurs couples (H_0, μ) choisis aléatoirement.

4.6.3 Méthode de continuation par déformation de la cible

Passons maintenant aux approches par déformation de la cible. Précisons qu'elles ont toutes été testées, même si elles ne sont pas utilisées dans les exemples qui suivent. Signalons que des tests complémentaires à ceux présentés ici sont donnés dans [11]. Les paramètres utilisés restent ceux indiqués en début de section.

En pratique, même pour $\theta = 0$, il est numériquement nécessaire de faire quelques itérations de Newton. En effet, le couple $(H_0, \mu) = (-S/T, 0)$ n'est pas la solution exacte du problème dès lors que l'on a effectué une discrétisation en temps. Celle-ci introduit alors une déviation par rapport à la trajectoire du problème continue en temps.

Ne disposant pas de résultat théorique d'existence, nos tests consistent à appliquer nos méthodes de continuation sur des états cibles tirés au hasard. Pour tester l'algorithme, nous construisons une série d'états cibles arbitraires de deux manières.

État cible complètement arbitraire. Nous tirons au hasard un couple $(S, A) \in \mathcal{S} \times \mathcal{A}$ et posons

$$U_{\text{cible}} = \exp(\imath S + A).$$

État cible issue d'une propagation. Nous tirons au hasard un couple $(H_0, \mu) \in \mathcal{S} \times \mathcal{S}$ et choisissons l'état final après propagation comme état cible. Dans cette approche, le couple solution (H_0, μ) est donc connu.

Approche par décomposition logarithmique

Un premier test consiste à tester l'approche par décomposition logarithmique. Les fonctions $\varphi_S(\theta)$ et $\varphi_A(\theta)$ sont définies selon les formules suivantes

$$\varphi_S(\theta) = \frac{a}{2}\sin(\pi\theta) + 1 \tag{4.22}$$

$$\varphi_A(\theta) = \frac{1}{2}(1 - \cos(\pi\theta)), \tag{4.23}$$

où a est un paramètre arbitraire. L'état cible est tiré au hasard suivant l'alternative de l'état cible complètement arbitraire. Nous observons dans ce cas un certain nombre d'échecs de la méthode. La figure 4.1 rend compte d'un de ces échecs. Nous voyons sur cette figure l'apparition d'une sorte de frontière d'inversibilité. À l'approche de cette zone, la jacobienne utilisée dans Newton devient singulière, et la méthode finit par exploser. Nous ne savons pas dans ce cas si la non-inversibilité est juste une conséquence de la méthode de Newton, inappropriée dans ce cas, ou bien si une obstruction plus profonde à l'inversibilité est ici révélée.

Approche par changement de base unitaire

Dans une deuxième série de tests, nous utilisons exclusivement la troisième décomposition. Ici, nous considérons un grand nombre d'états cibles et comptons le nombre de succès obtenus par cette approche. Sur vingt cas testés, nous observons l'aboutissement de la méthode dans 10% des cas lorsque l'état cible est complètement arbitraire et 25% lorsque l'état cible est issu d'une propagation. Dans tous ces cas, une simple application de la méthode de Newton présentée en section 4.4.2 conduit à une divergence. Ces résultats sont donc mitigés. Signalons tout de même que ces tests sont très théoriques et que l'application de la méthode à des cas pratiques en chimie est systématiquement couronnée de succès.

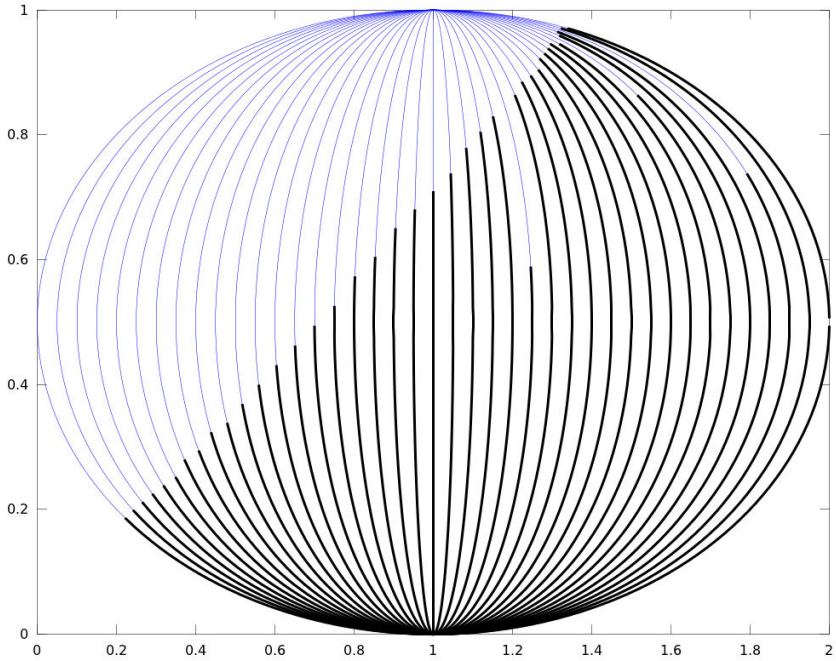


FIGURE 4.1 – Différents chemins sont testés pour la continuation, en faisant varier le paramètre a de -2 à 2 dans les formules ((4.22) − (4.23)). Les courbes bleues indiquent les trajectoires suivies par $\varphi_S(\theta)$ et $\varphi_A(\theta)$. Les courbes noires indiquent les parties de ces trajectoires où la méthode a fonctionné, c'est-à-dire où les itérations de Newton ont convergé.

Généralisation de la méthode de continuation

Dans ce troisième test numérique, nous considérons une généralisation de notre méthode de continuation suivant l'algorithme présenté dans le livre [1] de E.L. Allgower et K. Georg. Il s'agit toujours, dans ce cas, de décrire notre problème sous la forme d'une équation de la forme

$$G\left(\left(H_0, \mu\right), \theta\right) = 0,$$

où G s'applique sur un espace de dimension $N_d^2 + 1$ et arrive dans un espace de dimension N_d^2. Ceci a déjà été décrit dans la sous-section 4.5.1. L'idée supplémentaire ici est de suivre une courbe paramétrée $s \mapsto \left(\left(H_0(s), \mu(s)\right), \theta(s)\right)$ en partant du point d'initialisation de nos méthodes de continuation. À chaque étape, nous calculons la différentielle $G'(s)$ et faisons avancer les trois variables dans une direction appartenant à son noyau, qui est de dimension au moins égale à 1. Cette étape de prédiction est suivie de notre boucle de Newton qui modifie les trois variables pour les ramener sur la variété $G\left(\left(H_0, \mu\right), \theta\right) = 0$. Celle-ci peut ici s'interpréter comme une étape de correction. L'algorithme s'arrête lorsque la courbe croise l'hyperplan $\theta = 1$.

Cette approche n'étant pas encore tout a fait maîtrisée, nous ne la détaillons pas ici d'avantage. On trouvera une description complète et précise dans la référence indiquée. Signalons tout de même pour résumer qu'elle permet de généraliser nos méthodes de continuation en ne forçant pas la variable θ à aller vers la valeur 1 avec une vitesse prescrite.

Nous avons repris l'exemple associé à la figure 4.1, avec comme continuation l'approche par changement de base unitaire. Dans ce cas, nous obtenons la solution. La figure 4.2 montre l'évolution de θ au cours de l'algorithme. Nous voyons que dans ce cas, le problème est résolu en une soixantaine de pas de la méthode généralisée. Le problème constaté dans la première série de test est donc lié à la méthode elle-même : en forçant l'évolution de θ, l'algorithme est conduit à croiser une variété où la jacobienne devient singulière, c'est-à-dire la variété $\mathrm{d}G = 0$. A contrario,

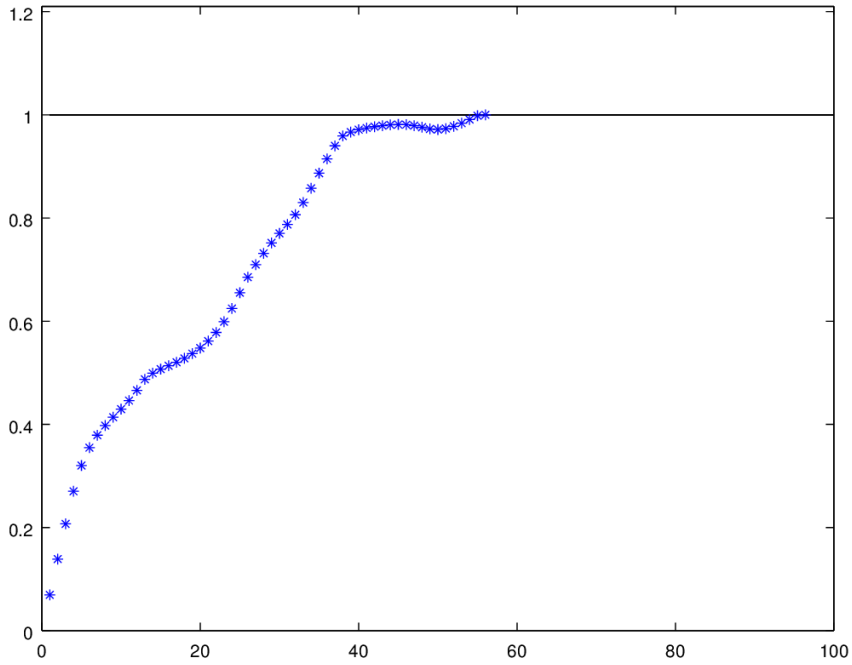


FIGURE 4.2 – Évolution des valeurs de θ au cours des itérations de la méthode généralisée.

la méthode généralisée permet de contourner cette variété en préservant le rang de la jacobienne. La question de savoir si une telle approche conduit toujours au croisement de l'hyperplan $\theta = 1$ est ouverte.

Références

[1] E.L. ALLGOWER and K. GEORG. *Numerical Continuation Methods : An Introduction*. Springer Series in Computational Mathematics. Springer Berlin Heidelberg, 2012. ISBN: 9783642612572.

[2] A. AVEZ. *Calcul différentiel*. Maîtrise de mathématiques pures. Masson, 1983. ISBN: 9782225790799.

[3] S BLANES et al. "Magnus and Fer expansions for matrix differential equations : the convergence problem". *Journal of Physics A : Mathematical and General* 31.1 (1998), p. 259. URL: http://stacks.iop.org/0305-4470/31/i=1/a=023.

[4] P.G. CIARLET, C.L. BRIS, and J.L. LIONS. *Handbook of Numerical Analysis*. Computational Chemistry : Reviews of Current Trends vol. 10. North-Holland, 1990. ISBN: 9780444512482.

[5] G. DRIDI et al. "A discrete-pulse optimal control algorithm with an application to spin systems". *Preprint HAL, hal-01186082* (2015).

[6] P. LAURENT et al. "Control through operators for quantum chemistry". *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*. 2012, pp. 1663–1667. DOI: 10.1109/CDC.2012.6427030.

[7] Y. MADAY, J. SALOMON, and G. TURINICI. "Monotonic time-discretized schemes in quantum control". *NUMERISCHE MATHEMATIK* 103.2 (2006), pp. 323–338. ISSN: 0029-599X. DOI: 10.1007/s00211-006-0678-x.

[8] Y. MADAY and G. TURINICI. "New formulations of monotonically convergent quantum control algorithms". *JOURNAL OF CHEMICAL PHYSICS* 118.18 (2003), pp. 8191–8196. ISSN: 0021-9606. DOI: 10.1063/1.1564043.

[9] Jan R. MAGNUS and H. NEUDECKER. "The Elimination Matrix : Some Lemmas and Applications". *SIAM Journal on Algebraic Discrete Methods* 1.4 (1980), pp. 422–449. DOI: 10.1137/0601049.

[10] J.R. MAGNUS and H. NEUDECKER. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley Series in Probability and Statistics : Texts and References Section. Wiley, 1999. ISBN: 9780471986331.

[11] M. NDONG, J. SALOMON, and D. SUGNY. "Newton algorithm for Hamiltonian characterization in quantum control". *Journal of Physics A : Mathematical and Theoretical* 47.26 (2014), p. 265302. DOI: 10.1088/1751-8113/47/26/265302. URL: http://stacks.iop.org/1751-8121/47/i=26/a=265302.

[12] G. von WINCKEL, A. BORZI, and S. VOLKWEIN. "A globalized Newton method for the accurate solution of a dipole quantum control problem". *SIAM JOURNAL ON SCIENTIFIC COMPUTING* 31.6 (2009), pp. 4176–4203. ISSN: 1064-8275. DOI: 10.1137/09074961X.

Thèse de Doctorat

Philippe LAURENT

Méthodes d'accélération pour la résolution numérique en électrolocation et en chimie quantique

Acceleration methods for numerical solving in electrolocation and quantum chemistry

Résumé
Cette thèse aborde deux thématiques différentes.
On s'intéresse d'abord au développement et à l'analyse de méthodes pour le sens électrique appliqué à la robotique. On considère en particulier la méthode des réflexions permettant, à l'image de la méthode de Schwarz, de résoudre des problèmes linéaires à partir de sous-problèmes plus simples. Ces deniers sont obtenus par décomposition des frontières du problème de départ. Nous en présentons des preuves de convergence et des applications. Dans le but d'implémenter un simulateur du problème direct d'électrolocation dans un robot autonome, on s'intéresse également à une méthode de bases réduites pour obtenir des algorithmes peu coûteux en temps et en place mémoire.
La seconde thématique traite d'un problème inverse dans le domaine de la chimie quantique. Nous cherchons ici à déterminer les caractéristiques d'un système quantique. Celui-ci est éclairé par un champ laser connu et fixé. Dans ce cadre, les données du problème inverse sont les états avant et après éclairage. Un résultat d'existence locale est présenté, ainsi que des méthodes de résolution numériques.

Mots clés
Électrolocation ; méthode des réflexions ; équations intégrales de frontière ; éléments de frontière (BEM) ; inversion géométrique ; méthode des bases réduites ; méthode de décomposition orthogonale propre (POD) ; méthode d'interpolation empirique (EIM) ; chimie quantique ; problème inverse ; méthode de continuation.

Abstract
This thesis tackle two different topics.
We first design and analyze algorithms related to the electrical sense for applications in robotics. We consider in particular the method of reflections, which allows, like the Schwartz method, to solve linear problems using simpler sub-problems. These ones are obtained by decomposing the boundaries of the original problem. We give proofs of convergence and applications. In order to implement an electrolocation simulator of the direct problem in an autonomous robot, we build a reduced basis method devoted to electrolocation problems. In this way, we obtain algorithms which satisfy the constraints of limited memory and time resources.
The second topic is an inverse problem in quantum chemistry. Here, we want to determine some features of a quantum system. To this aim, the system is ligthed by a known and fixed Laser field. In this framework, the data of the inverse problem are the states before and after the Laser lighting. A local existence result is given, together with numerical methods for the solving.

Key Words
Electrolocation ; reflections method ; boundary integral equations ; boundary element methods (BEM) ; geometric inversion ; reduced basis methods ; proper orthogonal decomposition (POD) ; empirical interpolation method (EIM) ; quantum chemistry ; inverse problem ; continuation method.