TÉCNICO LISBOA

A numerical study of the Dirac Spectrum and Transmission Problems employing the Method of Fundamental Solutions

Francisco Alves Bento

Thesis to obtain the Master of Science Degree in

Applied Mathematics and Computation

Supervisors: Pedro Ricardo Simão Antunes
Juha Hans Videman

Examination Committee

Chairperson: Prof. Pedro Lima
Members of the Committee: Prof. Hugo Tavares
Prof. Pedro Serranho

September 2023

“If Stavrogin believes, he does not believe that he believes. And if he does not believe, he does not believe that he does not believe.”

Fyodor Dostoyevsky
Demons

Acknowledgments

À primeira vista, escrever a secção de Agradecimentos quase que toca a formalidade, pequenos parágrafos, uma minuta, curtas linhas sem floreados ou sem a escrita técnica que premeia o resto do trabalho. Porém, como no fim de qualquer etapa, tem o condão de nos levar a reavaliar todo o nosso percurso e toda a gente que tornou este trabalho possível.

Assim, o papel de destaque final vai para os meus orientadores, os professores Juha Videman e Pedro Antunes. Estou-lhes muito grato por aceitarem fazer este percurso comigo, pela disponibilidade que sempre demonstraram, pelas dúvidas esclarecidas, pela experiência e conhecimento que comigo partilharam; e claro, pela motivação e confiança que me deram e tiveram em mim.

Não poderei deixar ainda de agradecer ao Centro de Análise Matemática, Geometria e Sistemas Dinâmicos (CAMGSD), pela bolsa que me foi atribuída e sob a qual tive a oportunidade de efetuar este trabalho.

Claro, aos meus colegas e amigos que me acompanharam, pelas discussões e conversas, pelo apoio que me deram, este que foi sem dúvida importantíssimo em vários pontos, o meu obrigado. Vários houveram que ficaram pelo caminho, outros tantos que nunca souberam que dele fizeram parte, e uns poucos que de um estreito carreiro fizeram áleas: a também esses o meu obrigado.

Mas quem nunca ficou pelo caminho foi a minha família, aos meus pais, à minha irmã, aos meus avós. Este trabalho não teria sido possível sem o vosso incondicional apoio e pelo que sempre me proporcionaram. O meu grande obrigado.

Abstract

This thesis delves into the application of the Method of Fundamental Solutions (MFS), a meshless technique, to address a duo of Partial Differential Equations (PDE) problems. Meshless methods provide an alternative to the standard mesh-based approaches, especially suited for intricate geometries. This study is centered on two focal points: firstly, the spectral analysis of the Dirac operator with infinite mass boundary conditions, investigated through large-scale simulations; secondly, the resolution of transmission problems involving the Poisson equation within polygonal and curved domains.

Within the MFS framework, the spectral behavior of the Dirac operator is explored, both verifying existing conjectures and postulating new ones. The study also covers transmission problems with the Poisson equation, utilizing singularity subtraction techniques to refine the accuracy of the method.

Consisting of six chapters, this thesis establishes foundational theory, rigorously introduces and implements the MFS, incorporating strategies to address inherent limitations. The results presented underscore the method's validity in addressing challenging PDE problems, showcasing the effectiveness of meshless methods.

Keywords

Meshless methods; Method of Fundamental Solutions; Dirac operator with infinite mass boundary conditions; Transmission problems with the Poisson equation; Numerical simulations; Singularity subtraction techniques.

Resumo

Esta dissertação investiga a aplicação do Método das Soluções Fundamentais (MSF em português), um método sem malha, adereçando dois distintos problemas em Equações de Derivadas Parciais (EDPs). Os métodos sem malha são uma alternativa aos clássicos métodos com malha e são particularmente adequados a geometrias mais complexas. Este estudo centra-se em dois pontos principais: primeiramente, na análise espetral do operador de Dirac com condições de fronteira de massa infinita, que foi investigado usando simulações de larga escala; em segundo lugar, na resolução de problemas de transmissão que envolvem a equação de Poisson, tanto em domínios poligonais e curvos.

Sob a estrutura do MSF, o comportamento espetral do operador de Dirac é explorado sistematicamente, verificando conjeturas existentes e postulando novos resultados. Este estudo cobre também problemas de transmissão com a equação de Poisson, utilizando técnicas de subtração de singularidade que permitem melhorar a precisão do método.

Tendo seis capítulos, esta tese estabelece teoria basilar, introduz e implementa o MSF rigorosamente, incorporando estratégias que abordam as suas inerentes limitações. Os resultados apresentados sublinham a validade do método em resolver problemas de EDPs complicados, mostrando assim a eficácia de métodos sem malha.

Palavras Chave

Métodos sem malha; Método das Soluções Fundamentais; Operador de Dirac com condições de fronteira de massa infinita; Problemas de Transmissão com a Equação de Poisson; Simulações numéricas; Técnicas de Subtração de Singularidade.

Contents

List of Figures

List of Algorithms

Listagens

Acronyms

BFGS	Broyden–Fletcher–Goldfarb–Shanno algorithm
BVP	Boundary Value Problem
MFS	Method of Fundamental Solutions
PDE	Partial Differential Equations
RBF	Radial Basis Functions
RMSE	Root Mean Squared Error

1

Introduction

Contents

1.1 On the applications of the Method of Fundamental Solutions

Partial differential equations (Partial Differential Equations (PDE)) serve as fundamental tools for modeling a wide spectrum of physical phenomena across scientific disciplines, ranging from engineering and physics to biology and finance. Given the complexity and infeasibility of deriving analytical solutions for many cases, accurate numerical solutions have become imperative. While well-established methods like finite differences and finite elements are available for a wide range of PDE, meshless methods offer an effective alternative, particularly for intricate geometries. This study delves into the Method of Fundamental Solutions (MFS), a meshless technique, investigating its applications in solving two distinctive problem domains: the spectral analysis of the Dirac operator under infinite mass boundary conditions, and transmission problems involving the Poisson equation within polygonal and curved domains.

Emerging in the latter part of the previous century, meshless methods provide an alternative to traditional mesh-based approaches, circumventing the challenges of mesh generation in complex geometries. Drawing inspiration from potential and integral equations theory, the more recent Method of Fundamental Solutions approximates solutions by exploiting the fundamental solutions of governing PDE. It has garnered attention for solving eigenvalue problems, as seen in [2], [3], and [4]. Its adaptability across diverse geometries renders it well-suited for handling intricate problem settings. Our aim is to uncover the MFS's capabilities in addressing various PDE challenges, thereby providing valuable insights into the systems under examination.

The spectral analysis of the Dirac operator under infinite mass boundary conditions, a problem classified as pivotal in shape optimization theory by [5], plays a critical role in our comprehension of quantum mechanics and quantum field theory. Our focus is centered on comprehending the spectral behavior of this operator. By employing the Method of Fundamental Solutions we strive to offer numerical insights that both validate existing conjectures and engender new ones.

Shifting our attention to transmission problems for the Poisson equation, which hold significance in fields like heat conduction, electromagnetism, and contact mechanics, we employ the MFS to investigate solutions within polygonal and curved domains. This study delves into the complexities introduced by interfaces and compatibility conditions in such scenarios. In order to increase the precision of the method, we incorporate methodologies to enhance the MFS, integrating singularity subtraction techniques to heighten accuracy, especially in proximity to domain corners. Importantly, this study marks the first instance of utilizing this technique with the Method of Fundamental Solutions for these specific problems.

In subsequent sections, comprising theoretical foundations and numerical methodologies, our objective is to present a straightforward perspective of the MFS's role in addressing complex PDE problems. This study not only furthers our understanding of meshless approaches but also bridges the gap between numerical simulation and theoretical research, serving as a source of new and challenging problems.

1.2 Thesis Overview

This thesis is structured into five distinct chapters, each organized as follows:

- Chapter 2 introduces foundational concepts in Functional Analysis and Partial Differential Equations. While the majority of these results can be found in classical references and are often covered in graduate courses, they serve as crucial underpinnings for the subsequent chapters. Notably, Chapter 4 draws heavily upon these concepts to establish the theoretical framework of the Method of Fundamental Solutions. This chapter not only presents and explains the method but also employs some different theoretical approaches to enhance its rigor.

- Chapter 3 serves as an introduction to the problems investigated within this thesis and is divided into three distinct sections. The initial section delves into the analysis of the Laplace operator, presenting established results and conducting a literature review. Although not directly connected with the study of the Dirac operator, the similarities between the two lead to the conjecture that significant findings of the Laplace operator could extend to the Dirac problem with infinite mass boundary conditions. This section also serves as a ground for the formulation of new conjectures concerning the Dirac operator's spectrum.

  The subsequent portion of this chapter focuses on the exploration of the Dirac Operator. It introduces the operator, elucidates some of its properties, and highlights its spectral characteristics. Additionally, a concise yet insightful proof demonstrates the absence of separable solutions in polar coordinates, extending what was previously known for cartesian coordinates. Recent conjectures postulated by field experts are presented, and novel conjectures, influenced by the prior analysis of the Laplace operator, are introduced. These conjectures subsequently become subjects of investigation using the MFS, enabling a comprehensive exploration of their validity and implications.

  Finally, the third section of this chapter centers on the Poisson transmission problem. It establishes the problem's context and its relationship with the Poisson equation when featuring a discontinuous source term. This section adopts a modern and rigorous approach, providing an analysis of the relationship between the transmission problem and the classical Poisson equation. This examination is important for the subsequent application of the MFS, and it will be needed to theoretically justify the use of the method in Chapter 4.

- Chapter 4 introduces the Method of Fundamental Solutions, and presents the various density proofs which justify this numerical method for the various problems, improving both in the rigor and details. It also presents convergence and stability results, the advantages and disadvantages of the method, and different ways to address them, specifically an enrichment technique using partic-

ular (angular) solutions responsible for singularity subtraction, and the Subspace Angle Technique presented in [6] is explained. Finally, its numerical implementation and a direct search algorithm used to find the eigenvalues are presented.

- In Chapter 5, we present our numerical findings, organized into two distinct sections. The initial section examines the Dirac operator with infinite mass boundary conditions, involving extensive large-scale simulations. Subsequently, outcomes for various domain shapes, including quadrilaterals, triangles, and smooth domains, are outlined, emphasizing the pertinent discoveries. This section concludes by addressing an unconstrained minimization problem aimed at identifying optimal shapes. The second section is dedicated to the transmission problem, focusing on achieved numerical errors and the utilization of enrichment techniques to enhance the method's accuracy in such scenarios. These results are enhanced with visual aids and concise tables summarizing our findings.

- Finally, Chapter 6 finish this work, presents the relevant conclusions of this thesis, and proposes some future work related to this research topic.

2

Some Preliminary Results

Contents

2.1 Some concepts on Banach Spaces

This section begins by introducing preliminary concepts on Banach Spaces, which play a crucial role in the subsequent numerical methods to be presented. For more details see [7] or [8]. Consider a field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). We say that a vector space E is a *normed space* if there exists a map $\|\cdot\|$ (called a *norm*) over $\mathbb{F}$ such that

1. $\|\alpha x\| = |\alpha|\|x\|, \alpha \in \mathbb{F}, \forall x \in E$;
2. $\|x + y\| \leq \|x\| + \|y\|, \forall x, y \in E$;
3. $\|x\| \geq 0, \ \forall x \in E$;
4. $\|x\| = 0 \iff x = 0$.

In particular, a pivotal notion is of *Banach spaces*, i.e., E is a Banach space if it is a complete normed space.

Definition 2.1.1. *Consider a linear operator $T : E \to F$, where E and F are Banach spaces with the associated norms $\|\cdot\|_E$ and $\|\cdot\|_F$, respectively. Then*

1. *The **nullspace** (also called the **kernel**) of T is a subset of E such that*

$$N(T) = \{x \in E : Tx = 0\}.$$

*Accordingly, the **range** (also called the **image**) of T is a subset of F such that*

$$R(T) = \{y \in F : \text{ there exists some } x \in E \text{ such that } y = Tx\}.$$

2. *T is said to be **bounded** (continuous) if there exists $C > 0$ such that $\|Tx\|_F \leq C\|x\|_E$. We define the norm of the operator T as*

$$\|T\| = \sup_{\substack{x \in E \\ x \neq 0}} \frac{\|Tx\|_F}{\|x\|_E}.$$

In this case, we write $T \in \mathcal{L}(E, F)$. If $E = F$, we write $T \in \mathcal{L}(E)$;

3. *The space of linear and continuous maps from E to $\mathbb{R}$ is the **dual space** of E denoted by $E^\star$. If $S \in E^\star$, its norm (the dual norm) is defined in the same manner as the operator norm above, i.e.*

$$\|S\| = \sup_{\substack{x \in E \\ x \neq 0}} \frac{|\langle S, x\rangle|}{\|x\|}$$

where $\langle S, x\rangle_{E^\star, E} = Sx$ and denotes de duality pairing between $E^\star$ and E. As we will see below, it generalizes the notion of inner product in inner product spaces. Whenever it is obvious what dual pairing is being considered we just write $\langle \cdot, \cdot \rangle$;

4. *Assuming that T is bounded, T is said to be* ***compact*** *if for any bounded sequence $(u_n)_{n\in\mathbb{N}} \subset E$ there exists a subsequence $(u_{n_k})_{k\in\mathbb{N}}$ such that $(Tu_{n_k})_{k\in\mathbb{N}}$ converges in F;*

5. *Assume that the domain of T, which we represent by $\mathrm{Dom}(T)$, is dense in E. We say that the linear operator $T^\star : \mathrm{Dom}(T^\star) \subset F^\star \to E^\star$ is the* ***adjoint*** *of T if*

$$\langle v, Tu\rangle_{F^\star,F} = \langle T^\star v, u\rangle_{E^\star,E^\star},\ \forall v \in \mathrm{Dom}(T^\star),$$

where the domain of $T^\star$ is defined by

$$\mathrm{Dom}(T^\star) = \{v \in F^\star : \exists c \geq 0 \text{ such that } |\langle v, Tu\rangle_{F^\star,F}| \leq c\|u\|,\ \forall u \in \mathrm{Dom}(T)\}.$$

The main result of this section concerns the dual of a normed space. In reality, we do not need to assume that E is a Banach space to present the next results. However, throughout this work, every normed space is also complete. We refer to Chapters 1 and 3 from [8].

Definition 2.1.2 (Reflexive space)**.** *Let E be a normed space and denote its dual by $E^\star$. The bidual space E^{**} is the dual of $E^\star$ with the associated norm*

$$\|\xi\| = \sup_{\substack{f\in E^\star \\ f\neq 0}} \frac{|\langle \xi, f\rangle|}{\|f\|}.$$

If the (canonical) map $J : E \to E^{**}$ defined by

$$\langle Jx, f\rangle_{E^{**},E^\star} = \langle f, x\rangle_{E^\star,E},\ \forall x \in E, \forall f \in E^\star$$

is surjective then E is said to be reflexive.

The definition above is an important detail in the justification of the Method of Fundamental Solutions. The main ingredient to justify this numerical method is the Hahn-Banach theorem.

Theorem 2.1.3 (Analytical form of Hahn-Banach Theorem)**.** *Let E be a normed space and $p : E \to \mathbb{R}$ a functional satisfying*

$$p(\lambda x) = \lambda p(x),\ \forall x \in E,\ \lambda > 0$$
$$p(x + y) \leq p(x) + p(y)$$

Let $G \subset E$ be a linear subspace and $g : G \to \mathbb{R}$ a linear functional such that

$$g(x) \leq p(x),\ \forall x \in G.$$

Then, there exists a linear functional $f : E \to \mathbb{R}$ that extends g to E, agrees with g on G, i.e, $f(x) = g(x),\ \forall x \in G$ and also satisfies

$$f(x) \leq p(x)\ \forall x \in E.$$

Remark 2.1.4. *The theorem mentioned above holds particular significance in Functional Analysis as it demonstrates that the dual $E^\star$ of a normed space E possesses interesting properties that warrant further study to gain a better understanding of the underlying space E. It can even be utilized to identify, although not uniquely, elements in both E and its dual $E^\star$ through duality pairing. This result bears a resemblance to the desirable properties exhibited by Hilbert spaces, which we will explore further in this work. It is beneficial to establish density when working with the dual pairing between a Banach space and its dual. However, it is important to note that the existence of the functional f is not explicitly provided, as the proof of Theorem 2.1.3 relies on the Axiom of Choice (Zorn's Lemma).*

Under some conditions, an interesting consequence of Theorem 2.1.3 is that two disjoint (and non-empty) convex sets can always be separated by a hyperplane in an infinite-dimensional space.

Definition 2.1.5. *Let E be a normed space, f a linear functional on E, and $c \in \mathbb{R}$. A hyperplane H is a subset of E of the form*

$$H = \{x \in E : \langle f, x\rangle = c\}.$$

Proposition 2.1.6. *Let H be a hyperplane defined by the equation $\langle f, x\rangle = c$, for some linear functional f and $c \in \mathbb{R}$. Then, H is closed if and only if f is continuous.*

Notice that if H is a closed hyperplane then the linear functional f that defines the hyperplane is an element of $E^\star$.

Definition 2.1.7. *Let A and B be two subsets of E. We say that a hyperplane H defined by the equation $\langle f, x\rangle = c$, for some linear functional f and $c \in \mathbb{R}$, strictly separates A and B if*

$$\langle f, x\rangle < c, \ \forall x \in A,$$
$$\langle f, x\rangle > c, \ \forall x \in B$$

Theorem 2.1.8 (Second geometric form of Hahn-Banach Theorem)**.** *Let A and B be two disjoint, non-empty, and convex subsets of E such that A is closed and B is compact. Then, there exists a closed hyperplane that strictly separates A and B, i.e, there exists $f \in E^\star$ and $c \in \mathbb{R}$ such that for every $a \in A$ and $b \in B$*

$$\langle f, a\rangle < c < \langle f, b\rangle.$$

The following Lemma is a consequence of Theorem 2.1.8, and it is a useful tool to prove that some linear subspace $M \subset E$ is dense (in E). We start by introducing the notion of orthogonality in Banach spaces concerning duality pairing.

Definition 2.1.9. *Let E be a Banach Space and M be a linear subspace of E. We define the orthogonal of M in E in respect to the duality pairing as*

$$M^\perp = \{\psi \in E^\star : \langle \psi, \varphi\rangle = 0, \ \forall \varphi \in M\}.$$

Accordingly, if $N \subset E^\star$ is a linear subspace, its orthogonal is defined as

$$N^\perp = \{\varphi \in E : \langle \psi, \varphi \rangle = 0,\ \forall \psi \in N\}$$

Lemma 2.1.10. *Let M and N be in the same conditions as the definition above. Then*

$$(M^\perp)^\perp = \overline{M}$$

and

$$\overline{N} \subset (N^\perp)^\perp.$$

In particular, if E is a reflexive Banach space then

$$(N^\perp)^\perp = \overline{N}.$$

Proof. Since $(M^\perp)^\perp \subset E$ and $(N^\perp)^\perp \subset E^\star$ are closed sets, by definition $M \subset (M^\perp)^\perp$ and $N \subset (N^\perp)^\perp$ the inclusions

$$\overline{M} \subseteq (M^\perp)^\perp, \qquad \overline{N} \subseteq (N^\perp)^\perp$$

follow. To check that $(M^\perp)^\perp \subseteq \overline{M}$ we argue by contradiction. Let $x_0 \in (M^\perp)^\perp$ such that $x_0 \notin \overline{M}$. Then, by Theorem 2.1.8 there exists a hyperplane with equation $\langle f, x \rangle = c$ for some $f \in E^\star$ and $c \in \mathbb{R}$ that strictly separates the sets $\{x_0\}$ and $\overline{M}$ (both are obviously non-empty convex sets). In particular,

$$\langle f, x \rangle_{E^\star, E} < c < \langle f, x_0 \rangle_{E^\star, E},\ \forall x \in M.$$

Since M is a linear subspace, then $\langle f, x \rangle = 0,\ \forall x \in M$ since, otherwise, given any $x \in M$ we would have that

$$\alpha \langle f, x \rangle_{E^\star, E} = \langle f, \alpha x \rangle_{E^\star, E} < c,\ \forall \alpha \in \mathbb{R}$$

which can only be possible if $\langle f, x \rangle_{E^\star, E} = 0$. Therefore, $f \in M^\perp$ and $\langle f, x_0 \rangle_{E^\star, E} > 0$ but that is a contradiction since, by hypothesis, $x_0 \in (M^\perp)^\perp$ and $\langle f, x_0 \rangle_{E^\star, E} = 0,\ \forall f \in M^\perp$.

To prove that $(N^\perp)^\perp = \overline{N}$, we use the same type of argument. Let $f_0 \in (N^\perp)^\perp$ such that $f_0 \notin \overline{N}$. Once again, there exists a hyperplane with equation $\langle \xi, f \rangle = c$ for some $\xi \in E^{**}$ and $c \in \mathbb{R}$ that strictly separates $\{f_0\}$ and $\overline{N}$, that is

$$\langle \xi, f \rangle < c < \langle \xi, f_0 \rangle,\ \forall f \in N.$$

Since N is a linear subspace we can also conclude that $\langle \xi, f \rangle = 0, \forall f \in N$ and $\langle \xi, f_0 \rangle > 0$. In order to get a contradiction, like in the case above, since E is a reflexive space, then the canonical map J defined in (2.1.2) is surjective, and we can write

$$\langle \xi, f_0 \rangle_{E^{**}, E^\star} = \langle J x_0, f_0 \rangle_{E^{**}, E^\star} = \langle f_0, x_0 \rangle_{E^\star, E}$$

for some $x_0 \in E$. If we can prove that $\langle f_0, x_0 \rangle_{E^\star, E} = 0$, then the contradiction follows. Since $f_0 \in (N^\perp)^\perp$, then $\langle f_0, x_0 \rangle_{E^\star, E} = 0$ if $x_0 \in N^\perp$, i.e, $\langle f, x_0 \rangle_{E^\star, E} = 0, \forall f \in N$. Let $f \in N$. Then, by reflexivity,

$$\langle f, x_0 \rangle_{E^\star, E} = \langle J x_0, f \rangle_{E^{\star\star}, E^\star} = \langle \xi, f \rangle_{E^{\star\star}, E^\star} = 0, \ \forall f \in N$$

as we saw above (N is a linear subspace). The desired result follows. □

Remark 2.1.11. *The result above will be useful to justify the Method of Fundamental Solutions. To prove that some subset N of a Banach Space E is dense it will suffice to show that its orthogonal $N^\perp$ only contains the trivial element, which belongs to both E and $E^\star$.*

2.2 Some concepts on Hilbert Spaces

In this section, we introduce some complementary results in Hilbert spaces. Once again, for more details, see [7], [8], or [9]. Consider the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). We say that a vector space H is an *inner product space* (or a Pre-Hilbert space) if there exists a map $(\cdot, \cdot)$ (called an *inner product*) over $\mathbb{F}$ such that

1. $(x, y) = \overline{(y, x)}$ (The bar denotes complex conjugation if $\mathbb{F} = \mathbb{C}$);
2. $(x + y, z) = (x, z) + (y + z)$;
3. $(\alpha x, y) = \alpha(x, y)$, $\alpha \in \mathbb{F}$;
4. $(x, x) \geq 0, \ \forall x \in H$;
5. $(x, x) = 0 \iff x = 0$.

Given $x, y \in H$, we say that x and y are orthogonal (denoted by $x \perp y$) if $(x, y) = 0$. Accordingly, given $E, F \subset H$, if $x \perp y$ for every $x \in E, y \in F$ then we say that E and F are orthogonal, $E \perp F$. We also denote by $E^\perp$ the set of all $y \in H$ that are orthogonal to every $x \in E$, i.e, $E^\perp = \{y \in H : (x, y) = 0, \ \forall x \in E\}$ which we call the orthogonal complement of E. We recall that every inner product space is also a normed space, where the inner product induces the norm

$$\|x\| = \sqrt{(x, x)}$$

satisfying the Cauchy-Schwarz inequality

$$|(x, y)| \leq \|x\| \, \|y\| , \ x, y \in H.$$

Finally, if the normed space is complete for the induced norm, then we say that it is a Hilbert space. In what follows, H will always denote a Hilbert space.

Example 2.2.1. *A very classical Hilbert space, which is going to be used throughout all of this work, is the space of square-integrable real-valued functions in an open and bounded subset Ω of $\mathbb{R}^d$, which is denoted by $L^2(\Omega)$ with the inner product given by*

$$(f,g)_{L^2(\Omega)} = \int_\Omega f(x)g(x)dx.$$

If considering complex-valued functions, the inner product is given by

$$(f,g)_{L^2(\Omega)} = \int_\Omega f(x)\overline{g}(x)dx,$$

where a bar over an expression represents the complex conjugate of the scalar (function).

While working within the framework of Hilbert spaces, proving the density of a subspace $M \subset H$ is more intuitive and can be derived straightforwardly. It may be of interest to the reader to compare the following results with Definition (2.1.9) and Lemma 2.1.10.

Theorem 2.2.2. *Consider a closed subspace $M \subset H$. Then,*

$$H = M \oplus M^\perp.$$

In other words, every $u \in H$ admits a unique decomposition $u = v + w$, where $v \in M$ and $w \in M^\perp$.

Corollary 2.2.3. *Consider a subspace $M \subset H$. Then M is dense in H if and only if $M^\perp = \{0\}$.*

Proof. Let $T = \overline{M}$. We want to prove that $T = H$. Using Theorem 2.2.2, it suffices to check that $T^\perp = \{0\}$. Since the inner product is continuous, then $T^\perp = M^\perp = \{0\}$.

On the other hand, since by definition T is closed, by Theorem 2.2.2 we have that $H = T \oplus T^\perp = T \oplus \{0\} = T$ as we wished. □

A surprising property in Hilbert spaces is the fact that every linear and continuous function $T : H \to \mathbb{F}$ can be *represented* by some unique element in H. In what follows we assume that $\mathbb{F} = \mathbb{R}$.

Theorem 2.2.4 (Riesz Representation Theorem)**.** *Let $T : H \to \mathbb{R}$ be a linear and continuous functional. Then, there exists a unique $u \in H$ such that*

$$Tv = (u,v),\ \forall v \in H.$$

Moreover, let $H^\star$ be the dual space of H, that is, the space of all linear and continuous functions from H to $\mathbb{R}$. Then the map $H^\star \mapsto H$ is an isometric isomorphism (which we denote by $\approxeq$) where

$$\|u\|_H = \|T\|_{H^\star}.$$

Remark 2.2.5. *Notice how the inner product in Hilbert spaces has replaced the duality pairing defined for Banach spaces. In fact, in a certain sense, Riesz Representation Theorem 2.2.4 allows us to make*

a stronger statement regarding the dual spaces of Hilbert spaces and work in a more natural framework without ever resorting to the Hahn-Banach Theorem 2.1.3. For example, it is interesting to observe that the definition of orthogonality and orthogonal subspaces in Hilbert spaces (via the inner product) and Banach spaces (via the duality pairing) are essentially the same, with the distinction being an isomorphism between the Hilbert space and its dual. However, in certain cases, working with the definition of orthogonality in Banach spaces can be more useful as it allows for better generalization, such as when proving the density of a closed subspace.

A more general and useful result in our work is the Lax-Milgram Theorem.

Definition 2.2.6. *We say that a bilinear form $a : H \times H \to \mathbb{R}$ is continuous and coercive if*

- $|a(u,v)| \leq C\|u\| \, \|v\|, \ \forall u, v \in H$
- $a(u,u) \geq \alpha \|u\|^2, \ \forall u \in H$

respectively.

Theorem 2.2.7 (Lax-Milgram Theorem)**.** *Let $a(u,v)$ be a bilinear, continuous, and coercive bilinear form on H. If T is a linear and continuous functional in H, then there exists a unique $u \in H$ such that*

$$a(u,v) = T(v), \ \forall v \in H.$$

In this section, the well-known *Spectral Theorem* is presented. With this in mind, we will now present key results and concepts (without proof) that will provide the necessary foundation for stating the theorems.

Definition 2.2.8. *A sequence $(e_n)_{n\in\mathbb{N}} \in H$ is a Hilbert basis of H if*

1. $(e_n, e_m) = \begin{cases} 1, & n = m \\ 0, & n \neq m \end{cases}$;
2. $\overline{\operatorname{span}\{(e_n)_{n\in\mathbb{N}}\}} = H$.

In a sense, a Hilbert basis resembles a basis in a finite-dimensional vector space.

Proposition 2.2.9. *Let $(e_n)_{n\in\mathbb{N}}$ be a Hilbert basis of H. Then, for every $u \in H$, we can write*

$$u = \sum_{k\in\mathbb{N}} (u, e_k) e_k \quad \textit{and} \quad \|u\|^2 = \sum_{k\in\mathbb{N}} |(u, e_k)|.$$

The last equality is known as Parseval's identity.

The proposition above is particularly interesting because it allows us to express every element of H in terms of a countable basis. The following result guarantees the existence of a Hilbert basis if certain conditions are met.

Definition 2.2.10. *We say that H is a separable Hilbert space if there exists a countable subset $M \subset H$ such that $\overline{M} = H$.*

Theorem 2.2.11. *Every separable Hilbert space admits an orthonormal Hilbert basis.*

We now present some properties that our operators must satisfy in order to state the Spectral Theorem.

Definition 2.2.12. *Consider a linear operator $T : H_1 \to H_2$, where H_1 and H_2 are Hilbert spaces.*

1. *Assume that $H = H_1 = H_2$ and $T \in \mathcal{L}(H)$. We say that $T^\star$ is the **adjoint** of T if*

$$(y, Tx) = (T^\star y, x),\ \forall x, y \in H;$$

*If $T = T^\star$, i.e, if T and $T^\star$ domains (and their image) coincide we say that T is **self-adjoint**.*[1]

2. *As above, let $T \in \mathcal{L}(H)$. We say that λ is an **eigenvalue** of T if $N(T - \lambda I) \neq \{0\}$. In that case, we say that $\lambda \in \sigma(T)$ where $\sigma(T)$ is called the **spectrum** of T*[2]*. We also say that u is an **eigenvector** associated with the eigenvalue λ if $u \in N(T - \lambda I) \setminus \{0\}$.*

We can derive some important properties regarding the spectrum of a compact operator and the spectrum of self-adjoint operators.

Proposition 2.2.13. *Let H be a Hilbert space and consider a compact operator $T \in \mathcal{L}(H)$. Then,*

- $0 \in \sigma(T)$*;*
- *one of the following holds:*
  - $\sigma(T) = \{0\}$*;*
  - $\sigma(T) \setminus \{0\}$ *is a finite set;*
  - $\sigma(T) \setminus \{0\}$ *is a sequence converging to* 0*.*

Proposition 2.2.14. *Let H be a Hilbert space and consider a self-adjoint operator $T \in \mathcal{L}(H)$. In these conditions, $\sigma(T)$ is real and eigenvectors corresponding to distinct eigenvalues are orthogonal.*

It is now possible to state one of the main results of this section.

[1] The existence and uniqueness of $T^\star$ may not be obvious, but if follows from Riesz Representation Theorem 2.2.4. In any case, notice the similarities between this definition and the one given in Definition (2.1.1). In fact, in Banach spaces, the existence of the adjoint also comes from the Hahn-Banach Theorem 2.1.3!

[2] Remarkably, in the infinite-dimensional case, the set of eigenvalues $EV(T)$ may not coincide with the spectrum $\sigma(T)$. $T - \lambda I$ may fail to be invertible even if $T - \lambda I$ is injective.

Theorem 2.2.15 (Spectral Theorem for compact and self-adjoint operators)**.** *Let H be a separable Hilbert space of infinite dimension and let $T \in \mathcal{L}(H)$ be a compact self-adjoint operator. Then, H admits a Hilbert basis $(e_n)_{n\in\mathbb{N}}$ such that*

$$Te_n = \lambda_n e_n$$

for $\lambda_n \in \mathbb{R}$, $\lambda_n \to 0$ as $n \to \infty$, where λ_n can be assumed to be a decreasing sequence.

2.3 Lebesgue and Sobolev Spaces

In this section, we apply the results and concepts from the previous ones. Besides the usual references already presented, we recommend [10], [11], and [12]. Let $\Omega \subset \mathbb{R}^d$ be an open set. Consider a *multi-index* $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$, where $|\alpha| = \alpha_1 + \cdots + \alpha_d$. Given a function defined in Ω, we denote its partial derivatives of order α by

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}}.$$

As usual, we denote the space of test functions with compact support in Ω by

$$\mathcal{D}(\Omega) = C_0^\infty(\Omega) = \{\varphi \in C^\infty(\Omega) : \operatorname{supp}\varphi \text{ is compact in } \Omega\}.$$

Definition 2.3.1 (Lebesgue spaces)**.** *Let $1 \leq p \leq \infty$. We define the Lebesgue space (L^p space)*

$$L^p(\Omega) = \left\{u : \Omega \to \mathbb{R} : u \text{ is measurable and } \int_\Omega |f|^p < \infty\right\}$$

with the associated norm

$$\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f|^p\right)^{\frac{1}{p}}.$$

If $p = \infty$ we set

$$L^\infty(\Omega) = \left\{u : \Omega \to \mathbb{R} : u \text{ is measurable and } \exists C > 0 : |f(x)| \leq C \text{ a.e on } \Omega\right\}$$

with the associated norm

$$\|f\|_{L^\infty(\Omega)} = \inf\{C : |f(x)| \leq C \text{ a.e on } \Omega\}$$

Definition 2.3.2. *We say that $f \in L^1_{loc}(\Omega)$ if f is integrable in every compact $K \subset \Omega$, i.e.*

$$f\chi_K \in L^1(\Omega),\ \forall K \subset \Omega \text{ compact.}$$

This definition can be extended accordingly to every $L^p(\Omega)$ space, with $1 \leq p \leq \infty$.

Before introducing the notion of weak derivative in Sobolev spaces, it will be useful to dive into some Distribution theory. Consider the space of test functions $\mathcal{D}(\Omega)$ above. While we are not interested to define a topology in this space, we want to define linear and continuous functionals acting on $\mathcal{D}(\Omega)$. For now, it suffices to define (sequential) convergence in $\mathcal{D}(\Omega)$.

Definition 2.3.3. *Let $(\varphi_n)_n \in \mathcal{D}(\Omega)$ and $\varphi \in \mathcal{D}(\Omega)$. If*

1. $\forall n \in \mathbb{N}$ *there exists a compact $K \subseteq \Omega$ such* $\operatorname{supp} \varphi_k \subseteq K$*;*

2. $\forall \alpha \in \mathbb{N}_0^d \ \lim_n \|D^\alpha \varphi_n - D^\alpha \varphi\|_{L^\infty(\Omega)} = 0$

then we say that φ_n converges to φ in $\mathcal{D}(\Omega)$.

Definition 2.3.4 (Space of distributions)**.** *The dual space of $D(\Omega)$, denoted by $\mathcal{D}^\star(\Omega)$, is called the space of distributions, and we say that $T \in \mathcal{D}^\star$ is a distribution.*

A very illustrative example, with consequences when defining the duality pairing in Sobolev spaces, is that any locally integrable function $u \in L^1_{\text{loc}}(\Omega)$ defines a distribution. It is easy to prove that the operator T_u defined by

$$\begin{aligned} T_u : \mathcal{D}(\Omega) &\to \mathbb{R} \\ \varphi &\mapsto \int_\Omega u\varphi \end{aligned}$$

is linear and continuous. Therefore, one can give meaning to the action of a distribution over a test function, whenever the distribution is induced by a locally integrable function u. In this case, we can write

$$\langle u, \varphi \rangle_{D^\star(\Omega), D(\Omega)} = \int_\Omega u\varphi$$

where the duality pairing $\langle \cdot, \cdot \rangle_{D^\star(\Omega), D(\Omega)}$ can be seen as a generalization of the $L^2(\Omega)$ inner product.

Definition 2.3.5 (Sobolev Spaces)**.** *For $k \in \mathbb{N}$ and $1 \leq p \leq \infty$ we define the Sobolev space*

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega), \forall \alpha \in \mathbb{N}_0^d : |\alpha| \leq k\}$$

with the associated norms

- $1 \leq p < \infty$*,*
$$\|u\|_{W^{k,p}(\Omega)} \coloneqq \Big(\sum_{|\alpha| \leq k} \|D^\alpha u\|^p_{L^p(\Omega)} \Big)^{\frac{1}{p}};$$
- $p = \infty$*,*
$$\|u\|_{W^{k,p}(\Omega)} \coloneqq \max_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)}.$$

We say that $D^\alpha u$ is the weak derivative of order α of $u \in L^p(\Omega)$ if $D^\alpha u \in L^p(\Omega)$. The operator D^α is well-defined as a distribution and satisfies

$$\int_\Omega D^\alpha u \varphi = (-1)^{|\alpha|} \int_\Omega u D^\alpha \varphi, \ \forall \varphi \in \mathcal{D}(\Omega).$$

Throughout this work, we are mostly concerned with Sobolev spaces with $p = 2$. In this case, we set $H^k(\Omega) \coloneqq W^{k,2}(\Omega)$ which is a Hilbert space for the inner product[3]

$$(u, v) \coloneqq \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v)_{L^2(\Omega)}.$$

One of the main tools used in the last section of this chapter is the following embedding theorem that relates the topologies of $H^1(\Omega)$ and $L^2(\Omega)$.

Theorem 2.3.6 (Rellich Theorem)**.** *Assume that Ω is a bounded Lipschitz domain. Then, the embedding $H^1(\Omega) \to L^2(\Omega)$ is compact, i.e, given the bounded sequence $(u_n)_{n\in\mathbb{N}} \subset H^1(\Omega)$ there exists a convergent subsequence $(u_{n_k})_{k\in\mathbb{N}} \subset L^2(\Omega)$.*

While we have only defined Sobolev spaces for $k \in \mathbb{N}$, it is possible to define fractional Sobolev spaces with a real exponent $s \in \mathbb{R}_0^+$. In particular, such a generalization can be made using Fourier Transforms if $p = 2$. Below we give an equivalent definition for a function $u \in H^k(\mathbb{R}^d)$.

Lemma 2.3.7. *Let $u \in L^2(\mathbb{R}^d)$. Then*

$$u \in H^k(\mathbb{R}^d) \iff (1 + |\xi|^k)\hat{u} \in L^2(\mathbb{R}^d)$$

where $\hat{u} = \mathcal{F}u$ denotes the Fourier Transform of u, given by $\mathcal{F}u(\xi) = \hat{u}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi x\xi} u(x) d\xi$.

The above characterization of the $H^k(\mathbb{R}^d)$ space motivate the following definition

Definition 2.3.8. *Let $s \in \mathbb{R}$. We define*

$$H^s(\mathbb{R}^d) = \{u \in L^2(\mathbb{R}^d) : (1 + |\xi|^2)^{\frac{s}{2}} \hat{u} \in L^2(\mathbb{R}^d)\}$$

with the norm

$$\|u\|_{H^s(\mathbb{R}^d)} = \|(1 + |\xi|^2)^{\frac{s}{2}} \hat{u}\|_{L^2(\mathbb{R}^2)}.$$

The definition above only holds for Sobolev spaces defined over the whole space $\mathbb{R}^d$. In this work, we are mostly concerned with the behavior over a bounded set $\Omega \subset \mathbb{R}^d$. Unfortunately, there are multiple definitions of fractional Sobolev spaces over a bounded set that may not agree between themselves if the boundary of Ω is not smooth enough (if the boundary fails to be parameterized by a continuous function). In any case, the following definition suffices for this work, see [13], [14] or [15].

Definition 2.3.9. *Let $\Omega \subset \mathbb{R}^d$ be a bounded set with a Lipschitz boundary. We define*

$$\mathring{H}^s(\Omega) = \{v \in H^d(\mathbb{R}^d) : \operatorname{supp} v \subset \overline{\Omega}\}$$

and

$$H^s(\Omega) = H^s(\mathbb{R}^d) \setminus \mathring{H}^s(\mathbb{R}^d \setminus \Omega).$$

[3] Observe that if $k = 0$ then $H^0(\Omega) = L^2(\Omega)$

The norm of a function $u \in H^s(\Omega)$ is given by

$$\|u\|_{H^s(\Omega)} = \inf\{\|\tilde{u}\|_{H^s(\mathbb{R}^d)} : \tilde{u} \in H^s(\mathbb{R}^d),\ \tilde{u}_{|\Omega} = u\}.$$

Fractional Sobolev spaces are important for our study because they are deeply related to the boundary behavior of a given function. For example, if $u \in H^1(\Omega)$ and Ω is bounded and a Lipschitz domain, then $u_{|\partial\Omega} \in H^{\frac{1}{2}}(\partial\Omega)$. However, the statement above must be defined rigorously: not only u is only defined in the open set Ω, but u is only defined *almost everywhere* and the Lebesgue measure of $\partial\Omega$ is zero. Intuitively, we consider the continuous extension of functions from Ω to the boundary $\partial\Omega$ which is only possible if the domain is regular *enough*. This is done in the context of trace theory, c.f [16], [17] or [18].

Theorem 2.3.10. *Let Ω be a bounded set with a Lipschitz boundary. Then, there exists a linear and continuous mapping called the trace operator*

$$\gamma_0 : H^1(\Omega) \to H^{\frac{1}{2}}(\partial\Omega)$$

that admits a bounded right inverse represented by γ_0^{-1}.

In particular, if $u \in H^2(\Omega)$, then $\frac{\partial u}{\partial x_j} \in H^1(\Omega)$ for $j = 1, \dots, d$ and the operator

$$\begin{aligned} \gamma_1 : H^2(\Omega) &\to L^2(\partial\Omega) \\ u &\mapsto \frac{\partial u}{\partial n} = \gamma_0(\nabla u) \cdot n \end{aligned}$$

is linear and continuous[4]*.*

The result above is stated in a very weak form since we are working in Lipschitz domains. If Ω is a smooth domain, then $\gamma_1 : H^2(\Omega) \to H^{\frac{1}{2}}(\partial\Omega)$, c.f [10]. However, it should be noted that the normal derivative operator γ_1 cannot be defined if we only assume that $u \in H^1(\Omega)$. If such a continuous operator $\mathcal{N}$ existed, then $\mathcal{N}\varphi = 0$ for every $\varphi \in \mathcal{D}(\Omega)$. By continuity, this would imply $\mathcal{N}u = 0$ for all $u \in H^1(\Omega)$, leading to a contradiction.

An important (closed) subspace of $H^1(\Omega)$ is the kernel of the trace operator γ_0,

$$\ker \gamma_0 = \{u \in H^1(\Omega) : \gamma_0 u = 0\} =: H_0^1(\Omega)$$

which can be equivalently defined as the closure of $\mathcal{D}(\Omega)$ in the $H^1(\Omega)$ norm, i.e, $H_0^1(\Omega) := \overline{\mathcal{D}(\Omega)}^{H^1(\Omega)}$. $H_0^1(\Omega)$ is of major importance in the Dirichlet Laplacian problem, since the functions in $H_0^1(\Omega)$ "vanish" on $\partial\Omega$. Next, we state an important result to be used when studying the spectrum of the Dirichlet Laplacian.

Theorem 2.3.11 (Poincaré inequality)**.** *Let Ω be a bounded set. Define $W_0^{1,p}(\Omega) := \overline{\mathcal{D}(\Omega)}^{W^{1,p}(\Omega)}$. Then, there exists $C > 0$ such that*

$$\|u\|_{L^p(\Omega)} \leq C\|\nabla u\|_{L^p(\Omega)},\ \forall u \in W^{1,p}(\Omega).$$

[4] We take γ_0 as an element wise operator, where $\gamma_0(\nabla u) = \left(\gamma_0(\frac{\partial u}{\partial x_1}), \dots, \gamma_0(\frac{\partial u}{\partial x_d})\right)$.

Finally, we shed some light on the dual space of Sobolev spaces (with fractional exponent). When working over the whole space $\mathbb{R}^d$ one can prove the following result, c.f [19] or [20].

Theorem 2.3.12. *Let $s \in \mathbb{R}$ and $u \in H^s(\mathbb{R}^d)$. Then, any linear and continuous functional $T \in (H^s(\mathbb{R}^d))^\star$ that acts in $H^s(\mathbb{R}^d)$ can be uniquely represented by some $v \in H^{-s}(\mathbb{R}^d)$ and the duality pairing is given by*

$$\langle T, u\rangle_{(H^s(\mathbb{R}^d))^\star, H^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d} \hat{u}\hat{v} = \int_{\mathbb{R}^d} (1+|\xi|^2)^{\frac{s}{2}}\hat{u}(1+|\xi|^2)^{-\frac{s}{2}}\hat{v} \leq \|u\|_{H^s(\Omega)}\|v\|_{H^{-s}(\Omega)}$$

where $\hat{u}$ and $\hat{v}$ represents the Fourier Transform of u and v, respectively. In particular, the dual of $H^s(\mathbb{R}^d)$ is isomorphic to $H^{-s}(\mathbb{R}^d)$.

Remark 2.3.13. *Considering the inclusion $\mathcal{D}(\Omega) \subset H^s(\Omega)$, it is straightforward to observe that $(H^s(\Omega))^\star \subset D^\star(\Omega)$ which implies that $L^2(\Omega) \subset (H^s(\Omega))^\star \approxeq H^{-s}(\Omega)$. Consequently, based on the preceding theorem, we note that Sobolev spaces with negative exponent s can be regarded as spaces of distributions with the following inclusions:*

$$H^s(\Omega) \subset L^2(\Omega) \subset H^{-s}(\Omega)$$

In this case, one can consider $L^2(\Omega)$ as a pivot space, and by virtue of the identification of $L^2(\Omega)$ with its dual, the duality pairing $\langle\cdot,\cdot\rangle_{H^{-s}(\Omega),H^s(\Omega)}$ and the $L^2(\Omega)$ inner product coincide for every $u \in H^s(\Omega)$ and

$$\langle v, u\rangle_{H^{-s}(\Omega),H^s(\Omega)} = \int_\Omega uv$$

whenever it makes sense, i.e, when $v \in L^2(\Omega)$.

Theorem 2.3.10 allows to give some meaning to space $H^{-\frac{1}{2}}(\Omega)$. Let $u \in H^2(\Omega)$ and $v \in H^{\frac{1}{2}}(\partial\Omega)$. Then, $\gamma_0^{-1}v \in H^1(\Omega)$ and using Green's formula (see in the section below)

$$\langle \gamma_1 u, v\rangle_{H^{-\frac{1}{2}}(\partial\Omega),H^{\frac{1}{2}}(\partial\Omega)} = \int_\Omega \Delta u\gamma_0^{-1}v + \int_\Omega \nabla u\cdot\nabla\gamma_0^{-1}v$$

we have $\gamma_1 u \in H^{-\frac{1}{2}}(\Omega)$.

2.4 Spectral Decomposition of the Laplace Operator

In this section, we make a brief study regarding the Laplace Operator $-\Delta = -\sum_{n=1}^d \frac{\partial^2}{\partial x_n^2}$ in a bounded domain $\Omega \subset \mathbb{R}^d$ with Lipschitz boundary. Firstly, we recall the Divergence Theorem, e.g. [21].

Theorem 2.4.1 (Divergence Theorem)**.** *Let $\Omega \subset \mathbb{R}^d$ defined as above. Then,*

$$\int_\Omega \operatorname{div}\phi dx = \int_{\partial\Omega} \phi\cdot\mathbf{n}d\sigma,$$

where $\mathbf{n}$ denotes the exterior unitary normal.

A main consequence of the Divergence Theorem are the well-known *Green's Formulas*, with major importance in this work.

Corollary 2.4.2 (Green's Formulas)**.** *In this same conditions of the Theorem 2.4, let* $u, v \in H^2(U)$*. Then,*

1. $\int_\Omega \Delta u dx = \int_{\partial\Omega} \frac{\partial u}{\partial n} d\sigma$;
2. $\int_\Omega \Delta uv dx = -\int_\Omega \nabla u \cdot \nabla v dx + \int_{\partial\Omega} \frac{\partial u}{\partial n} v d\sigma$;
3. $\int_\Omega \Delta uv - u\Delta v dx = \int_{\partial\Omega} \frac{\partial u}{\partial n} v - \frac{\partial v}{\partial n} u d\sigma$.

The study of the spectrum of the following equation is of major importance throughout this work and will be studied in the following chapter. For now, we will only state and prove a classical result which can also be found in numerous textbooks, see [8] [9], [22] or [23]. While we assume null Dirichlet boundary conditions, we notice that the Neumann case is analogous.

Definition 2.4.3. *Consider the Helmholtz equation with null Dirichlet boundary conditions*

$$\begin{cases} -\Delta u(x) = \lambda u(x), & x \in \Omega, \\ u(x) = 0, & x \in \partial\Omega, \end{cases} \tag{2.1}$$

where $\Delta = \sum_{i=0}^{d} \frac{\partial^2}{\partial x_i^2}$*. Then,* $\lambda \in \mathbb{C}$ *is an eigenvalue of the equation* (2.1) *if there exists an eigenfunction* $u \neq 0$ *belonging to the function spaces* $C^2(\Omega) \cap C(\overline{\Omega})$*.*

Theorem 2.4.4. *There exists a Hilbert basis* $(u_n)_{n\in\mathbb{N}}$ *of* $L^2(\Omega)$ *consisting of eigenfunctions* u_n *of* $-\Delta$*, i.e, for each* $n \in \mathbb{N}$ *there exists a pair eigenvalue/eigenfunction* (λ_n, u_n) *such that*

$$-\Delta u_n = \lambda_n u_n$$

where the sequence of eigenvalues can be ordered in increasing order and $\lambda_n \to \infty,\ n \to \infty$*. In particular, define* $E_n = \operatorname{span}\{u_1, \ldots, u_n\}$ *and the Rayleigh Quotient*

$$R(u) = \frac{\|\nabla u\|^2_{L^2(\Omega)}}{\|u\|^2_{L^2(\Omega)}}.$$

Then,

$$\lambda_n = \min_{\substack{u \in E_{n-1}^\perp \\ u \neq 0}} R(u) = \max_{\substack{u \in E_n \\ u \neq 0}} R(u).$$

Proof. For each $f \in L^2(\Omega)$, we consider the problem

$$\begin{cases} -\Delta u(x) = f, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega \end{cases}$$

with the associated variational form

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega fv, \ \forall v \in H_0^1(\Omega).$$

Using Lax-Milgram Theorem 2.2.7, it is straightforward to prove that the variational form above admits a unique weak solution $u \in H_0^1(\Omega)$ and the operator

$$T : L^2(\Omega) \to L^2(\Omega)$$
$$f \mapsto u$$

is well-defined. To prove that T is a compact operator, we use Poincaré and Cauchy-Schwarz inequalities and notice that

$$\alpha \|u\|^2_{H^1(\Omega)} \leq \int_\Omega |\nabla u|^2 = \int_\Omega fu \leq \|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} \|u\|_{H^1(\Omega)} \implies \|u\|_{H^1(\Omega)} \leq C\|f\|_{L^2(\Omega)}$$

where $\alpha, C > 0$. The above result can be written as

$$\|Tf\|_{H^1(\Omega)} \leq C\|f\|_{L^2(\Omega)}, \ \forall f \in L^2(\Omega)$$

and by Theorem 2.3.6 T is a compact operator. To check that T is self-adjoint it suffices to consider the weak variational form of the null Dirichlet boundary problems

$$-\Delta u = f \qquad -\Delta v = g$$

for $f, g \in L^2(\Omega)$ and apply Green's formulas. It is also easy to see that $(Tf, f)_{L^2(\Omega)} \geq 0, \forall f \in L^2(\Omega)$ since

$$\int_\Omega (Tf) f = \int_\Omega uf = \|\nabla u\|^2_{L^2(\Omega)} \geq 0.$$

Applying the Spectral Theorem 2.2.15 to T, there exists a Hilbert basis $(u_n)_{n\in\mathbb{N}}$ such that

$$Tu_n = \mu_n u_n$$

for $\mu_n \in \mathbb{R}$, $\mu_n \to 0$ as $n \to \infty$. In particular, taking $f = \lambda_n u_n$, where $\lambda_n = \frac{1}{\mu_n}$, one can write

$$-\Delta u_n = \lambda_n u_n,$$

or in the integral form

$$\int_\Omega |\nabla u|^2 = \lambda_n \int_\Omega u^2,$$

with $\lambda_1 \leq \lambda_2 \leq \cdots \to \infty$. To check the variational form of the eigenvalues λ_n, let $u \in E_{n-1}^{\perp}$. Then,

$$\begin{aligned}
\|\nabla u\|_{L^2(\Omega)}^2 = (\nabla u, \nabla u)_{L^2(\Omega)}^2 &= \Big(\sum_{m \geq n} (u, u_m)_{L^2(\Omega)} \nabla u_m, \nabla u \Big)_{L^2(\Omega)} \\
&= \sum_{m \geq n} (u, u_m)_{L^2(\Omega)} (\nabla u_m, \nabla u)_{L^2(\Omega)} \\
&= \sum_{m \geq n} \lambda_m (u, u_m)_{L^2(\Omega)} (u_m, u)_{L^2(\Omega)} \\
&\geq \lambda_n \sum_{m \geq n} \left| (u, u_m)_{L^2(\Omega)} \right|^2 \\
&= \lambda_n \|u\|_{L^2(\Omega)}^2
\end{aligned}$$

where we used the bilinearity of the inner product, the fact that the sequence λ_n is non-decreasing, and Parseval's identity. It is easy to check that the equality is only attained if and only if u is in the eigenspace of λ_k. This proves that

$$\lambda_n = \min_{\substack{u \in E_{n-1}^{\perp} \\ u \neq 0}} R(u).$$

The other case is analogous. □

Remark 2.4.5. *Observe that* (2.4.4) *only guarantees that the eigenfunctions* u_n *belong to* $H_0^1(\Omega)$*. In order to achieve the regularity stated in Definition* (2.4.3)*, some conditions on* Ω *should be imposed: for example, if* Ω *is an open set of class* C^2*. If* Ω *is smooth, then* $u_n \in C^\infty(\overline{\Omega})$*.*

Corollary 2.4.6 (Homogeneity)**.** *Let* $\alpha > 0$*. Consider the set*

$$\alpha\Omega = \{\alpha x \in \mathbb{R}^d : x \in \Omega\},$$

i.e, $\alpha\Omega$ *is a dilation of* Ω *by a factor of scale* α*. Then, for all* $n \in \mathbb{N}$*,*

$$\alpha^2 \lambda_n(\alpha\Omega) = \lambda_n(\Omega),$$

where $\lambda_n(\alpha\Omega)$ *is the* n*-esim eigenvalue of* (2.1) *on the domain* $\alpha\Omega$ *(and analogously for* $\lambda_n(\Omega)$*).*

Proof. The proof is an easy consequence of the variational description above. Let $\varphi(x) = \alpha x$ and $\alpha\Omega = \phi(\Omega)$. Then,

$$\lambda_n(\alpha\Omega) = \min_{\substack{u \in E_{n-1}^{\perp} \\ u \neq 0}} \frac{\int_{\varphi(\Omega)} |\nabla u(x)|^2 dx}{\int_{\varphi(\Omega)} u(x)^2 dx} = \min_{\substack{u \in E_{n-1}^{\perp} \\ u \neq 0}} \frac{\int_{\Omega} |\nabla u(\alpha x)|^2 dx}{\int_{\Omega} u(\alpha x)^2 dx},$$

via a change of variables. Let $v(x) = u(\alpha x)$. Then,

$$\nabla v(x) = \alpha \nabla u(\alpha x)$$

and

$$\alpha^2\lambda_n(\alpha\Omega) = \min_{\substack{u\in E_{n-1}^{\perp}\\ u\neq 0}} \frac{\int_\Omega |\alpha\nabla u(\alpha x)|^2 dx}{\int_\Omega u(\alpha x)^2 dx} = \min_{\substack{u\in E_{n-1}^{\perp}\\ u\neq 0}} \frac{\int_\Omega |\nabla v(x)|^2 dx}{\int_\Omega v(x)^2 dx} = \lambda_n(\Omega).$$

□

3

From Spectral Theory and Shape Optimization to the Poisson Transmission Problem

Contents

Until otherwise indicated, let $\Omega \subset \mathbb{R}^d$ be an open and bounded domain with C^2 boundary, with $d \geq 2$.

Let $p(x)$ be a polynomial in the variables $x = (x_1, \ldots, x_d)$, and $p(\partial)$ be the partial differential operator obtained by substituting $\frac{\partial}{\partial x_i}$ for x_i in $p(x)$. We start by introducing the definition of a fundamental solution of a partial differential operator:

Definition 3.0.1. *Consider the polynomial $p(\partial)$. A distribution $\Phi \in \mathcal{D}^\star(\mathbb{R}^n)$ is said to be the fundamental solution of the partial differential operator $p(\partial)$ if*

$$p(\partial)\Phi = \delta,$$

where δ is the Dirac Delta distribution.

In particular, given Φ satisfying the above conditions, we have that $p(\partial)\Phi(x) = 0$ for $x \in \mathbb{R}^d \setminus \{0\}$. Then, it is easy to see that the fundamental solution of a partial differential operator is not unique: if v is such that $p(\partial)v(x) = 0$ for all $x \in \mathbb{R}^d$, then $p(\partial)(\Phi + v) = \delta$. However, the fundamental solutions given below are chosen because they exhibit an important radial behavior, which is needed for the numerical method presented in Chapter 4.

An important result in this context is the Malgrange-Ehrenpreis theorem, which is also based on the Hahn-Banach Theorem 2.1.3.

Theorem 3.0.2 (Malgrange-Ehrenpreis)**.** *Every constant partial differential operator $p(\partial)$, has a fundamental solution $\Phi \in \mathcal{D}^\star(\mathbb{R}^d)$.*

Proof. See [24]. □

Below the fundamental solution of the Laplace operator and some major results concerning Laplace and Helmholtz equations are presented.

3.1 The Laplace Operator

In what follows, consider the Laplace operator associated with the well-known Laplace equation

$$-\Delta u = 0. \tag{3.1}$$

Throughout this first part, we are mostly concerned about its spectrum which is associated with the Helmholtz equation

$$-(\Delta + \lambda)u = 0 \iff -\Delta u = \lambda u. \tag{3.2}$$

Remark 3.1.1. *In some literature, it is common to write the Helmholtz equation as $-\Delta u = k^2 u$, where k is known as an eigenfrequency, and $k^2 = \lambda$. This terminology is a consequence of the fact that the Helmholtz equation can be derived from the wave equation, where a constant c^2 (where $c \in \mathbb{R}$) is used.*

By Theorem 3.0.2 we know that both equations (3.1) and (3.2) admit fundamental solutions that are given below.

Proposition 3.1.2. *The function* $\Phi : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ *given by*

$$\Phi(x) = \begin{cases} -\frac{1}{2\pi} \log |x| \,, & d = 2 \\ \frac{1}{(n-2)|\partial B_1|} \frac{1}{|x|^{d-2}}, & d > 2 \end{cases}$$

is the fundamental solution of equation (3.1)*, where* $|\partial B_1|$ *denotes the boundary measure of the unitary ball.*

Proposition 3.1.3. *The function* $\Phi_\lambda : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ *given by*

$$\Phi_\lambda(x) = \begin{cases} \frac{i}{4} H_0^{(1)}(\sqrt{\lambda}\|x\|), & d = 2 \\ \frac{e^{i\sqrt{\lambda}\|x\|}}{4\pi\|x\|}, & d > 2 \end{cases}$$

is the fundamental solution of equation (3.2)*, where* $H_0^{(1)}$ *is the Hankel function of the first kind and order 0, given by*

$$H_0^{(1)}(x) = J_0(x) + iY_0(x),$$

where J_0 *and* Y_0 *are the Bessel functions of the first and second kind with order zero, respectively.*

If one considers the eigenfrequency form of the Helmholtz equation, then the fundamental solution in Proposition 3.1.3 would change accordingly.

3.1.1 Some shape optimization results

In this subchapter, some important results regarding shape optimization are presented. More precisely, we are interested in problems of the form

$$\min\{F(\lambda_1(\Omega), \dots, \lambda_k(\Omega)) : |\Omega| = c, \Omega \subset \mathbb{R}^d\}, \tag{3.3}$$

where F is a function of the first k eigenvalues of the Helmholtz equation (3.2) and $c > 0$. We point the reader to [25] and [26] for more details.

Theorem 3.1.4 (Faber-Krahn inequality)**.** *Let* B *be a ball of volume* c*. Then, among all domains* Ω *of volume* c *we have that,*

$$\lambda_1(B) = \min\{\lambda_1(\Omega) : |\Omega| = c\}.$$

In particular, as proved by Krahn in [27], the corresponding isoperimetric inequality

$$\lambda_1(\Omega) \geq \left(\frac{C_d}{c}\right)^{\frac{d}{2}} j_{d/2-1,1},$$

where C_d *is the volume of the* d*-dimensional unit ball and* $j_{p,1}$ *is the first positive zero of the Bessel function* J_p*, holds.*

Theorem 3.1.4 is a classic form of an isoperimetric inequality, conjectured for the first time by Lord Rayleigh. More recently, a reverse of the Faber-Krahn inequality was proven in [28].

Theorem 3.1.5. *Let Ω be a bounded convex domain of $\mathbb{R}^d$ and denote the inradius of Ω (radius of the largest ball contained within the domain) by ρ_Ω. Then,*

$$\lambda_1(\Omega) \leq \frac{|\partial\Omega|}{d\rho_\Omega|\Omega|}\lambda_1(\mathbb{D})$$

While Faber-Krahn inequality deals with the first eigenvalue of the Laplace operator, other results have been uncovered for the second and third eigenvalues:

Theorem 3.1.6 (Krahn-Szegő)**.** *The domain that minimizes the quantity*

$$\min\{\lambda_2(\Omega) : |\Omega| = c\}$$

consists of two equal and disjoint balls of volume $\frac{c}{2}$.

A result regarding the topology of a given domain Ω and its connection with the minimization of each eigenvalue λ_k was given by Wolf and Keller in [29]. Roughly speaking, the result states that if $\Omega_k^\star$ is not connected and minimizes λ_k, then each connected component must be a minimizer for a lower eigenvalue.

Theorem 3.1.7 (Wolf-Keller)**.** *Let $\Omega_k^\star$ be the union of two disjoint domains, each of them with positive volume. Then,*

$$\lambda_k^\star = (\lambda_i^\star)^{\frac{d}{2}} + (\lambda_{k-i}^\star)^{\frac{d}{2}} = \min_{1\leq j\leq \frac{k-1}{2}}((\lambda_j^\star)^{\frac{d}{2}} + (\lambda_{k-j}^\star)^{\frac{d}{2}}),$$

where i is a value of $j \leq \frac{k-1}{2}$ that minimizes $\lambda_i^\star + \lambda_{k-i}^\star$. Furthermore,

$$\Omega_k^\star = \left[\left(\frac{\lambda_i^\star}{\lambda_k^\star}\right)^{\frac{1}{2}}\Omega_i^\star \bigcup \left(\frac{\lambda_{k-i}^\star}{\lambda_k^\star}\right)^{\frac{1}{2}}\Omega_{k-i}^\star\right]$$

The generalization of the Theorems (3.1.4) and (3.1.6) becomes harder to prove for high-order eigenvalues. Bucur and Henrot proved in [30] that there exists a domain Ω that minimizes λ_3, and it is conjectured to be the disk[1]. In [31], Bucur was able to assert the existence of, at least, one solution to problem (3.3).

Theorem 3.1.8 (Bucur)**.** *For every $k \in \mathbb{N}$ problem*

$$\min\{\lambda_k(\Omega) : |\Omega| = c\}$$

has at least one solution. Moreover, every solution is bounded and has a finite perimeter.

[1] Notice that Ω_3 must be connected in dimension 2. Otherwise, by Theorem 3.1.7, Ω_3 would be the union of the domains that minimize λ_1 and λ_2 (see Theorems (3.1.4) and (3.1.6)), where one can explicitly compute $\lambda_3 = \lambda_1 + \lambda_2 \approx 51.504$. However, this would be a contradiction since the eigenvalue of the unit disk is $\lambda_3(\mathbb{D}) \approx 46.125$ when considering unitary measure $c = 1$. For three dimensions the result is the same, but for $d \geq 4$ one cannot conclude anything.

In [32], Mazzoleni and Pratelli were able to generalize the above results for (*quasi-*)open sets and for a general functional F considered in (3.3) (see the reference for more details and definition of a quasi-open set).

Theorem 3.1.9 (Mazzoleni-Pratelli)**.** *Let $k \in \mathbb{N}$ and suppose that $F : \mathbb{R}^k \to \mathbb{R}$ in (3.3) is lower semi-continuous, increasing in each variable. Then, among the quasi-open sets, there exists a bounded minimizer Ω for the problem (3.3). More precisely, a minimizer Ω is contained in a cube of side R, where R depends on k and on the dimension of the space d, but not on F.*

In this work, we are also interested in studying the eigenvalues of polygonal domains, such as triangles and quadrilaterals. An important result concerning these domains is the fact that they are invariant under *Steiner symmetrizations* (in the sense that the Steiner symmetrization of a triangle or a quadrilateral is still a triangle or a quadrilateral, respectively), which preserves their area while decreasing the perimeter and its first eigenvalue. This type of transformation enables us to state the following result:

Theorem 3.1.10 (Pólya-Szégő)**.** *The quantity λ_1 is minimized for equilateral triangles among all triangles, where the inequality*

$$\frac{4\pi^2}{\sqrt{3}} \leq \lambda_1(T)$$

holds for every triangle T of area 1. Analogously, λ_1 is minimized for the square among all quadrilaterals.

However, an analogous result for the n-side polygon is still an open problem conjectured by Pólya and Szégő in [33].

Conjecture 3.1.11. *Let $n \geq 5$ and consider the class of n-side polygons. Then, the regular n-side polygon has the least first eigenvalue among all n-side polygons with fixed area.*

Very recently, Bogosel and Bucur proved in [34] that Conjecture (3.1.11) can be reduced to a finite number of certified numerical computations with machine precision and performed them for $k = 5, 6, 7, 8$.

Related to triangles, we can also cite the recent work of Serrano and Orriols in [35], which was based on the previous work of Antunes and Freitas [4] who conjectured that the first three eigenvalues are enough to define the shape of a triangle (such result resembles the famous Marc Kac question if one can "hear the shape of a drum" in [36], that is, if given the frequencies produced by a drum one could identify the drum's shape, which has proven to be false, see [37] for more details). In any case, Serano and Orriols were able to show that not any three eigenvalues suffice to fully characterize the shape of a triangle.

Theorem 3.1.12 (Serrano-Orriols)**.** *There exist two triangles T_A and T_B not isometric to each other such that $\lambda_i(T_A) = \lambda_i(T_B)$, for $i = 1, 2, 4$.*

Other important results in this area are related to the ratio between the first and the second eigenvalues.

Theorem 3.1.13 (Ashbaugh-Benguria)**.** *The solution to the maximization problem*

$$\max\left\{\frac{\lambda_2(\Omega)}{\lambda_1(\Omega)} : \Omega \subset \mathbb{R}^d, \Omega \textit{ open}\right\}$$

is the ball. In particular, it can be shown that

$$\frac{\lambda_2(\Omega)}{\lambda_1(\Omega)} \leq \frac{j^2_{\frac{d}{2},1}}{j^2_{\frac{d}{2}-1,1}}.$$

Remark 3.1.14. *Note that in the Theorem above we do not make any constraint regarding the volume of our domain (except finite measure). This is a consequence of the homogeneity proven in Corollary* (2.4.6) *and the fact that we are now considering a ratio between two eigenvalues in the same domain.*

3.2 The Dirac Operator

Motivated by recent advancements in nuclear, and molecular physics and the discovery of very interesting electrical, mechanical, and thermal properties, Dirac materials have resurged a lot of attention in the Dirac equation. Presented by Paul Dirac in its 1928 article [38], Dirac equation was able to successfully merge the famous Schrödinger equation with special relativity, while explaining the weird phenomena that today is known as *spin*, while predicting the existence *antimatter*. Capable of describing the relativistic dynamics of spin-$\frac{1}{2}$ particles (like the electron), one can determine the energy states by studying the spectrum of the Hamiltonian (Dirac) operator $\hat{H}$ in $L^2(\Omega, \mathbb{C}^2)$ for $\Omega \subset \mathbb{R}^2$,

$$\hat{H}\mathbf{u} = E\mathbf{u} \quad \text{with} \quad \hat{H} = -i(\sigma \cdot \nabla) + (m + V(x))\,\mathbb{I}_2, \tag{3.4}$$

where ∇ is the gradient operator, m is the particle mass, E its energy, $V(x)$ is some external potential and $\mathbf{u} \in L^2(\Omega, \mathbb{C}^2)$ is a two-component spinor. One of the major problems regarding the study of Dirac's equation is the fact that, unlike Schrödinger's equation, it has a matrix structure that is given by the Pauli matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

which can be incorporated in $\sigma = (\sigma_1, \sigma_2)$.

Setting the potential $V(x) = 0$ and considering $\Omega \subset \mathbb{R}^2$ be a bounded and open domain with C^3 boundary, we can rewrite equation (3.4) in the form

$$\begin{bmatrix} m & -i(\partial_1 - i\partial_2) \\ -i(\partial_1 + i\partial_2) & -m \end{bmatrix} \begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix} = E \begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix} \tag{3.5}$$

where we let $\mathbf{u}(x) = \begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix}$. In particular, we are interested in studying it under the so-called *infinite-mass boundary conditions*. We point the reader to [39], [40], and [41] for more details about this type of

boundary conditions and the results below. For a point $x \in \Gamma = \Gamma$, we denote by $\mathbf{n}(x) = \big(n_1(x), n_2(x)\big)^T$ the outward unit vector to Ω, and define the operator domain

$$D(\hat{H}) = \{u \in H^1(\Omega, \mathbb{C}^2) : u_2 = i(n_1 + in_2)u_1 \text{ on } \Gamma\}.$$

Let $\tau(x) = \big(n_2(x), -n_1(x)\big)^T$ be the unit tangent vector to some point $x \in \Gamma$ such that $(\tau(x), \mathbf{n}(x))$ is a positively-oriented orthonormal basis of $\mathbb{R}^2$. Considering the arc-length parametrization of Γ given by the map

$$s : [0, L) \to \mathbb{R}^2, \quad s(t) = \int_0^t \|r'(\sigma)\| d\sigma$$

where L represents the arc-length of Γ and r is a parametrization of Γ, we denote by $\kappa : \Gamma \to \mathbb{R}$ the signed curvature of Γ where the *Frenet-Serret* formula (we dropped the dependency of s in the parameter t)

$$\frac{\partial \tau}{\partial s} = \kappa(s)\mathbf{n}(s)$$

holds. Similar to the Theorem 2.4.4, some general results regarding the spectrum of the Dirac operator can also be stated.

Proposition 3.2.1. *Consider the eigenvalue problem stated in* (3.5)*. Then, the following results hold,* [2]

- *The eigenvalues are real and the spectrum of the Dirac operator is discrete. Also, the spectrum is symmetric, and the eigenvalues can be arranged as follows*

$$-\infty \leftarrow \cdots \leq -\lambda_3 \leq -\lambda_2 \leq -\lambda_1 < 0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \to \infty;$$

- *The principal (first) eigenvalue can be described using the variational form*

$$\lambda_1^2 = \min_{0 \neq u \in D(\hat{H})} \frac{\|\nabla u\|^2_{L^2(\Omega)} + m^2\|u\|^2_{L^2(\Omega)} + m\|\gamma_0 u\|^2_{L^2(\Gamma)}}{\|u\|^2_{L^2(\Omega)}},$$

  where $\gamma_0 : H^1(\Omega, \mathbb{C}^2) \to H^{\frac{1}{2}}(\Gamma, \mathbb{C}^2)$ denotes the trace of u;

- *Let $m = 0$ and Ω be the unit disk $\mathbb{D}$. Then, we have that the first eigenvalue is the solution to the equation*

$$J_0(\lambda_1) = J_1(\lambda_1),$$

  and the associated eigenfunction is (in polar coordinates)

$$u(r, \theta) = \begin{pmatrix} J_0(\lambda_1 r) \\ ie^{i\theta} J_1(\lambda_1 r) \end{pmatrix}.$$

  For future comparison, its numerical approximation is $\lambda_1 \approx 1.434695650819$, where we recall that J_p is the Bessel function of the first kind of order p.

[2] Notice that we changed the eigenvalue notation E to λ. Not only for coherence reasons, but also because we are mainly interested in the mathematical description of the problem, and not in the physical intuition behind it.

The following proposition regarding the lack of separable solutions of the Dirac operator will have important consequences in the numerical approach to solve the Dirac equation. However, we start by stating and proving the following auxiliary lemma:

Lemma 3.2.2. *Let $u \in H^2(\Omega)$ a solution of (3.5) and $u \in D(\hat{H})$. Then we have*

$$\|\hat{H}u\|^2_{L^2(\Omega)} = \|\nabla u\|^2_{L^2(\Omega)} + m^2\|u\|^2_{L^2(\Omega)} + m\|\gamma u\|^2_{L^2(\Gamma)} - \frac{1}{2}\int_\Gamma \kappa|u|^2 d\sigma \tag{3.6}$$

Proof. Recalling the $L^2(\Omega)$ inner product for complex functions

$$(f,g)_{L^2(\Omega)} = \int_\Omega f\bar{g}dx,$$

(3.6) can be rewritten as

$$\|\hat{H}u\|^2_{L^2(\Omega)} = m^2\|u_1\|^2_{L^2(\Omega)} + m^2\|u_2\|^2_{L^2(\Omega)} \tag{3.7}$$

$$+ im\int_\Omega u_1(\partial_1 + i\partial_2)\bar{u}_2 dx - im\int_\Omega \bar{u}_1(\partial_1 - i\partial_2)u_2 dx \tag{3.8}$$

$$+ im\int_\Omega \bar{u}_2(\partial_1 + i\partial_2)u_1 dx - im\int_\Omega u_2(\partial_1 - i\partial_2)\bar{u}_1 dx \tag{3.9}$$

$$+ \|(\partial_1 - i\partial_2)u_2\|^2_{L^2(\Omega)} + \|(\partial_1 + i\partial_1)u_1\|^2_{L^2(\Omega)}. \tag{3.10}$$

We now address each line of the expression above individually:

- For (3.7) we directly have

$$m^2\|u_1\|^2_{L^2(\Omega)} + m^2\|u_2\|^2_{L^2(\Omega)} = m^2\|u\|^2_{L^2(\Omega)}.$$

- For (3.8) we integrate by parts the first term:

$$\int_\Omega u_1(\partial_1 + i\partial_2)\bar{u}_2 dx = \int_\Gamma u_1\bar{u}_2(1+i)d\sigma - \int_\Omega \bar{u}_2(\partial_1 + i\partial_2)u_1 dx$$

  where the last term cancels with the first term of (3.9).

- Analogously, for (3.9) we obtain a similar result for the last term

$$\int_\Omega u_2(\partial_1 - i\partial_2)\bar{u}_1 dx = \int_\Gamma u_2\bar{u}_1(1-i)d\sigma - \int_\Omega \bar{u}_1(\partial_1 - i\partial_2)u_2 dx$$

  where the last term cancels with the last term of (3.8).

- For (3.10) we firstly deduced the following property:

$$\mathrm{Im}\Big(\int_\Omega \partial_1 v\partial_2\bar{v}dx\Big) = \frac{1}{2i}\int_\Gamma \bar{v}\partial_\tau v d\sigma,\ \forall v \in H^2(\Omega),$$

  where $\partial_\tau v = \tau\cdot\nabla v$, which can be obtained using integration by parts. Then, for each term we have

$$\begin{aligned}\|(\partial_1 - i\partial_2)u_2\|^2_{L^2(\Omega)} &= \|\nabla u_2\|^2_{L^2(\Omega)} + i\Big(\int_\Omega \partial_1 u_2\partial_2\bar{u}_2 dx - \int_\Omega \partial_2 u_2\partial_1\bar{u}_2 dx\Big)\\ &= \|\nabla u_2\|^2_{L^2(\Omega)} + i\int_\Gamma \bar{u}_2\partial_\tau u_2 d\sigma\end{aligned}$$

$$\|(\partial_1 + i\partial_2)u_1\|^2_{L^2(\Omega)} = \|\nabla u_1\|^2_{L^2(\Omega)} - i\Big(\int_\Omega \partial_1 u_1 \partial_2 \bar{u}_1 dx - \int_\Omega \partial_2 u_1 \partial_1 \bar{u}_1 dx\Big)$$
$$= \|\nabla u_1\|^2_{L^2(\Omega)} - i\int_\Gamma \bar{u}_1 \partial_\tau u_1 d\sigma$$

where we used the property above.

As such, we can write everything as

$$\|\hat{H}u\|^2_{L^2(\Omega)} = m^2\|u\|^2_{L^2(\Omega)} + im\Big(\int_\Gamma u_1\bar{u}_2(1+i)d\sigma - \int_\Gamma u_2\bar{u}_1(1-i)d\sigma\Big)$$
$$+ \|\nabla u\|_{L^2(\Omega)} + i\Big(\int_\Gamma \bar{u}_2\partial_\tau u_2 d\sigma - \int_\Gamma \bar{u}_1\partial_\tau u_1 d\sigma\Big).$$

Finally, using the boundary conditions $u_2 = i(n_1 + in_2)u_1$, we conclude that

$$im\Big(\int_\Gamma u_1\bar{u}_2(1+i)d\sigma - \int_\Gamma u_2\bar{u}_1(1-i)d\sigma\Big) = \|\gamma u\|^2_{L^2(\Gamma)}$$

while

$$i\int_\Gamma \bar{u}_2\partial_\tau u_2 d\sigma - i\int_\Gamma \bar{u}_1\partial_\tau u_1 d\sigma = -\frac{1}{2}\int_\Gamma \kappa|u|^2 d\sigma$$

where we used the Frenet-Serret formula above and the fact that at Γ we have $|u_1|^2 = |u_2|^2$. □

Proposition 3.2.3. *Let $u \in H^2(\Omega)$ a solution of (3.5) and $u \in D(\hat{H})$. Then u cannot be written using separable solutions, both in cartesian coordinates and polar coordinates.*

Proof. We start by showing that $|\lambda| > m$ for any eigenvalue λ if $\kappa = 0$ a.e. Assuming that there exists an eigenvalue λ associated with an eigenfunction u such that $|\lambda| \leq m$, by Lemma 3.2.2 we get that

$$\|\nabla u\|^2_{L^2(\Omega)} + m\|\gamma u\|^2_{L^2(\Gamma)} \leq 0 \implies \|\nabla u\|_{L^2(\Omega)} = 0 \wedge m\|\gamma u\|_{L^2(\Gamma)} = 0$$

and u must be a constant, which does not satisfy the boundary conditions (unless $u = 0$, which would satisfy the conditions above, but it is not considered).

Since $u \in H^2(\Omega)$, using (3.5), we can express u_2 as

$$u_2 = \frac{-i(\partial_1 + i\partial_2)u_1}{\lambda + m}$$

allowing us to rewrite Dirac equation (3.5) using the Helmholtz equation with Cauchy–Riemann oblique boundary conditions:

$$\begin{cases} -\Delta u_1 = (\lambda^2 - m^2)u_1, & \text{in } \Omega \\ i(\partial_1 + i\partial_2)u_1 + (\lambda + m)i(n_1 + in_2)u_1 = 0, & \text{on } \Gamma. \end{cases} \tag{3.11}$$

The rest of the proof will have into consideration two domain types: triangular (on a wedge) and rectangular.

1. Consider Ω to be a wedge domain with angle θ of maximum amplitude Θ (see Figure 4.1):

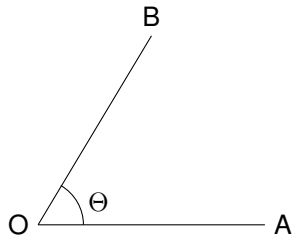


Figure 3.1: A wedge domain with an interior angle θ.

In this case, since the outward unit normal on $\overline{OA}$ is $n = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ and on $\overline{OB}$ is $n = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$, using polar coordinates system (3.11) transforms into

$$\begin{cases} \left(\partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2\right)u_1 = (\lambda^2 - m^2)u_1, & \text{in } \Omega \\ i(\cos\theta\partial_r - \frac{1}{r}\sin\theta\partial_\theta + i(\sin\theta\partial_r + \frac{1}{r}\cos\theta\partial_\theta))u_1 + (\lambda + m)u_1 = 0, & \text{on } \overline{OA} \cdot \\ i(\cos\theta\partial_r - \frac{1}{r}\sin\theta\partial_\theta + i(\sin\theta\partial_r + \frac{1}{r}\cos\theta\partial_\theta))u_1 + (\lambda + m)i(-\sin\theta + i\cos\theta)u_1 = 0, & \text{on } \overline{OB} \end{cases} \tag{3.12}$$

Assume that there exists a solution $u(r,\theta) = R(r)T(\theta)$ of (3.12). Then, from the PDE in the interior Ω we have

$$u(r,\theta) = J_k\left(r\sqrt{\lambda^2 - m^2}\right)(A\cos(k\theta) + B\sin(k\theta))$$

for some $A, B, k \in \mathbb{C}$, were A and B are not simultaneously zero. Applying the condition on $\overline{OA}$, substituting $\theta = 0$, one finds that

$$\frac{J_k\left(r\sqrt{\lambda^2 - m^2}\right)(r(m+\lambda)A - kB)}{r} + \frac{1}{2}i\sqrt{\lambda^2 - m^2}A\left(J_{-1+k}\left(r\sqrt{\lambda^2 - m^2}\right) - J_{1+k}\left(r\sqrt{\lambda^2 - m^2}\right)\right) = 0$$

From the recurrence relations for the derivatives of Bessel functions

$$2\frac{dJ_\alpha(r)}{dr} = J_{\alpha-1}(r) - J_{\alpha+1}(r)$$

the equation above can be rewritten as

$$\frac{J_k\left(r\sqrt{\lambda^2 - m^2}\right)(r(m+\lambda)A - kB)}{r} + iA\frac{d}{dr}\left[J_k\left(r\sqrt{\lambda^2 - m^2}\right)\right] = 0$$

where $J_k\left(r\sqrt{\lambda^2 - m^2}\right)$ is the solution of the differential equation

$$\frac{\tilde{R}(r)(r(m+\lambda)A - kB)}{r} + iA\frac{d}{dr}\tilde{R}(r) = 0,$$

whose solutions are of the form

$$\tilde{R}(r) = Ce^{ir(m+\lambda)}r^{-ik\frac{B}{A}}$$

for every $C \in \mathbb{C}$. Since they form a basis of solutions, there must be some $C \neq 0$ such that

$$J_k\left(r\sqrt{\lambda^2 - m^2}\right) = Ce^{ir(m+\lambda)}r^{-ik\frac{B}{A}}.$$

Analogously, considering the condition on $\overline{OB}$, one must have

$$\begin{aligned} &ir\Big(A\cos(k\Theta) + B\sin(k\Theta)\Big)\frac{d}{dr}J_k\left(r\sqrt{\lambda^2 - m^2}\right) \\ &- J_k\left(r\sqrt{\lambda^2 - m^2}\right)\Big(\cos(k\Theta)(Bk + Amr + A\lambda r) + \sin(k\Theta)(Br(\lambda + m) - Ak)\Big) = 0. \end{aligned}$$

Once again, $J_k\left(r\sqrt{\lambda^2 - m^2}\right)$ is the solution of the differential equation

$$\begin{aligned} &ir\Big(A\cos(k\Theta) + B\sin(k\Theta)\Big)\frac{d}{dr}\tilde{\tilde{R}}(r) \\ &- \tilde{\tilde{R}}(r)\Big(\cos(k\Theta)(Bk + Amr + A\lambda r) + \sin(k\Theta)(Br(\lambda + m) - Ak)\Big) = 0, \end{aligned}$$

whose solutions are given by

$$\tilde{\tilde{R}}(r) = De^{-ir(\lambda+m)}r^{-\frac{ik(B\cos(k\Theta) - A\sin(k\Theta))}{A\cos(k\Theta) + B\sin(k\Theta)}},\ \forall D \in \mathbb{C}.$$

Then, for some $D \neq 0$,

$$J_k\left(r\sqrt{\lambda^2 - m^2}\right) = De^{-ir(\lambda+m)}r^{-\frac{ik(B\cos(k\Theta) - A\sin(k\Theta))}{A\cos(k\Theta) + B\sin(k\Theta)}}$$

which implies that

$$Ce^{ir(m+\lambda)}r^{-ik\frac{B}{A}} = De^{-ir(\lambda+m)}r^{-\frac{ik(B\cos(k\Theta) - A\sin(k\Theta))}{A\cos(k\Theta) + B\sin(k\Theta)}}.$$

Since C and D are non-zero $\mathbb{C}$ constants, this implies that

$$\begin{cases} -2ir(m + \lambda) = 0 \\ -ik\left(-\frac{(B\cos(k\Theta) - A\sin(k\Theta))}{A\cos(k\Theta) + B\sin(k\Theta)} - \frac{A}{B}\right) = 0 \end{cases} \implies \begin{cases} \lambda = -m \\ k = 0, \end{cases}$$

a contradiction.

2. For rectangular domains, we point the proof in [40].

□

Remark 3.2.4. *A careful reader will note that two important details are being overlooked: triangular domains or quadrilaterals do not have the required smoothness to obtain the formula* (3.6) *(and signed curvature is not defined everywhere, only on each edge where $\kappa = 0$), neither u has enough regularity to be integrated by parts while expanding* (3.10)*. We point to [42], where such details can be found for triangles, but are any two-dimensional polygon.*

We now present some open conjectures that we try numerically address in this work:

Conjecture 3.2.5 (A Faber-Krahn type inequality)**.** *Let $m \geq 0$ and $\Omega \subset \mathbb{R}^2$ an open and Lipschitz domain. Then,*

$$\lambda_1(\Omega) \geq \lambda_1(\Omega^*)$$

where Ω^* is the disk of the same area or perimeter as Ω.

The conjecture above is regarded as a hot open problem in spectral geometry [5]. In [43] a geometric lower bound for the first (non-negative) eigenvalue was found, while in [39] a sharp upper bound (a reverse Faber-Krahn type inequality, like in Theorem 3.1.5) was proved to hold for convex domains with C^3 boundary. However, due to the difficulty of the problem, simpler versions with domain restrictions to triangles and rectangles are being studied. For example, in [40] a study for rectangles was conducted, where the conjectures below were proved under some extra hypothesis:

Conjecture 3.2.6 (Shape optimization in rectangles)**.** *Given $m \geq 0$, let $\lambda_1(a,b) = \lambda_1(\Omega_{a,b})$ denote the first eigenvalue of the Dirac Operator with infinite-mass boundary conditions in a rectangle with sides a and b. Then,*

1. *Area constraint:* $\lambda_1(a, \frac{1}{a}) \geq \lambda_1(1,1),\ \forall a > 0$*;*
2. *Perimeter constraint:* $\lambda_1(a, 2-a) \geq \lambda_1(1,1),\ \forall a \in (0,2)$.

In the same vain, in [42] very similar results were proven for triangles:

Conjecture 3.2.7 (Shape optimization in triangles)**.** *Consider the triangle $\Omega_{a,b}$ defined by the points $O = (0,0), A = (a,0), B = (0,b)$ for $a, b > 0$. Then, given any $m \geq 0$,*

1. *Area constraint:* $\lambda_1(a,b) \geq \lambda(k,k),\ \forall a,b,k > 0$ *such that* $ab = k^2$*;*
2. *Perimeter constraint:* $\lambda_1(a,b) \geq \lambda(k,k),\ \forall a \in (0,(2+\sqrt{(2)}k))$ *and* $\forall b,k > 0$ *such that* $a + b + \sqrt{a^2+b^2} = (2+\sqrt{2})k$.

3.3 A domain decomposition problem

Consider the polygonal domain $\Omega \subset \mathbb{R}^2$ which we divide into two non-overlapping regions Ω_1 and Ω_2 such that $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$. We denote their common boundary by $\gamma = \partial\Omega_1 \cap \partial\Omega_2$ and denote by $\Gamma_i = \partial\Omega_1 \setminus \gamma$ the boundary of each domain Ω_1 minus the common boundary. The problem we address in this section is to find functions u_1, u_2 such that

$$\begin{cases} -\nabla k_i \nabla u_i = f_i, & \text{in } \Omega_i \\ u_1 - u_2 = 0, & \text{on } \gamma \\ k_1 \frac{\partial u_1}{\partial n_1} + k_2 \frac{\partial u_2}{\partial n_2} = 0, & \text{on } \gamma \\ u_i = 0, & \text{on } \Gamma_i \end{cases} \tag{3.13}$$

where $k_1 \geq k_2 > 0$ are constants, $f_i \in L^2(\Omega_i)$ is a source function on each domain, and n_i is the (normalized) outward normal to each domain subdomain $\Omega_i, i = 1, 2$. Finally, we will write $n = n_1 = -n_2$ when we are restricted to the interface.

In what follows we mainly used the reference [44]. Equations (3.13) can be used to study a system of two bodies with different material parameters (contact resistance or thermal conductivity) connected by an interface γ. If we set

$$f = \begin{cases} \frac{f_1}{k_1}, & \Omega_1 \\ \frac{f_2}{k_2}, & \Omega_2 \end{cases},$$

then the problem above can be seen as a natural reformulation of the Poisson equation

$$\begin{cases} -\Delta u = f, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega \end{cases} \tag{3.14}$$

where f is (possibly) discontinuous on the interface γ. In order to keep the equivalence between both problems, we enforce some transmission conditions in Equations (3.13) using the continuity of the solutions and the continuity of their normal derivative on γ. To establish the equivalence between both problems we write the variational weak form associated with them. For (3.14) it's an easy process: we multiply the equation in Ω by a test function $v \in C_0^\infty(\Omega)$ and using Green's Identity we find that

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v = \int_\Omega fv$$

where a is the associated bilinear form. As such, enlarging our functional space to the Hilbert Space $V^0 = H_0^1(\Omega)$, our problem can be rewritten as

$$\text{find } u \in V^0 : a(u, v) = (f, v),\ \forall v \in V^0 \tag{3.15}$$

which has a unique solution in V^0 by virtue of Lax-Milgram lemma. In fact, by regularity results[3] (see Section 6.3.1 of [12]), $u \in V^0 \cap H^2_{\text{loc}}(\Omega'),\ \forall \Omega' \Subset \Omega$.

For (3.13) the process is not so direct. Given the subdomain Ω_1, we multiply the first equation with a test function in $v_1 \in C_0^\infty(\Omega_1)$ and integrate by parts where we get

$$a(u_1, v_1) = \int_{\Omega_1} \nabla u_1 \cdot \nabla v_1 = \left(\frac{f_1}{k_1}, v_1\right)$$

and analogously, for the subdomain Ω_2, considering $v_2 \in C_0^\infty(\Omega_2)$ we find

$$a(u_2, v_2) = \int_{\Omega_2} \nabla u_2 \cdot \nabla v_2 = \left(\frac{f_2}{k_2}, v_2\right).$$

To add the interface condition on the normal derivative notice that if $v \in C_0^\infty(\Omega)$, one can define $v_1 = v_{|\Omega_1}$ and $v_2 = v_{|\Omega_2}$, where $v_i \in C^\infty(\Omega_i)$ and $v_i(x) = 0,\ \forall x \in \Gamma_i$ for $i = 1, 2$. This allows us to write that

$$\begin{aligned} -\int_{\Omega_1} k_1 \Delta u_1 v_1 - \int_{\Omega_2} k_2 \Delta u_2 v_2 = \int_{\Omega_1} k_1 \nabla u_1 \cdot \nabla v_1 + \int_{\Omega_2} k_2 \nabla u_2 \cdot \nabla v_2 \\ - k_1 \int_\gamma \frac{\partial u_1}{\partial n_1} v_1 - k_2 \int_\gamma \frac{\partial u_2}{\partial n_2} v_2 \end{aligned}$$

[3] We write that $A \Subset B$ when $\overline{A}$ is compact and $\overline{A} \subset B$.

Observe that if $v_{1_{|\gamma}} = v_{2_{|\gamma}} = \eta$, then using the condition on the normal derivative we would find that

$$\int_{\Omega_1} k_1 \nabla u_1 \cdot \nabla v_1 + \int_{\Omega_2} k_2 \nabla u_2 \cdot \nabla v_2 = (f_1, v_1) + (f_2, v_2).$$

As such, consider the continuous extension operator P_i from the interface to each domain Ω_i such that $(P_i \eta)_{|\gamma} = \eta$. In that case, given a function μ defined in γ[4], the identity above holds if we rewrite it in the form

$$\int_{\Omega_1} k_1 \nabla u_1 \cdot \nabla P_1 \mu + \int_{\Omega_2} k_2 \nabla u_2 \cdot \nabla P_2 \mu = (f_1, P_1\mu) + (f_2, P_2\mu).$$

Finally, for the continuity on the interface, enforce that $u_1 = u_2$ on γ. The process above allows us to state the weak form of problem (3.13) in the result below.

Proposition 3.3.1. *Consider the set of equations* (3.13) *and the bilinear form* $a_i = (w_i, v_i) = (\nabla w_i, \nabla v_i)$. *Let*

$$\begin{aligned}
&V_i = \{v_i \in H^1(\Omega_i) : v_{i_{|\partial\Omega \cap \partial\Omega_i = 0}}\};\\
&V_i^0 = H_0^1(\Omega_i)\\
&\Lambda = \{\eta \in H^{\frac{1}{2}}(\gamma) : \eta = v_{|\gamma} \text{ for some } v \in V^0\}\\
&a_i(u_i, v_1) = \int_{\Omega_i} \nabla u_i \cdot \nabla v_1
\end{aligned}$$

where $V^0 = H_0^1(\Omega)$ *as above. Then the weak formulation of* (3.13) *reads as*

$$\text{find } u_1 \in V_1, u_2 \in V_2 \text{ such that } \begin{cases} a_1(u_1, v_1) = (\frac{f_1}{k_1}, v_1), & \forall v_1 \in V_1^0 \\ a_2(u_2, v_2) = (\frac{f_2}{k_2}, v_1), & \forall v_2 \in V_2^0 \\ u_1 = u_2, & \text{on } \gamma \\ a_1(k_1 u_1, P_1\mu) + a_2(k_2 u_2, P_2\mu) = (f_1, P_1\mu) + (f_2, P_2\mu), & \forall \mu \in \Lambda \end{cases} \tag{3.16}$$

where the operator $P_i : \Lambda \to V_i$ *is continuous for* $i = 1, 2$.

Remark 3.3.2. *The existence of the (continuous) extension operators* P_1 *and* P_2 *is not obvious. The theorem 2.3.10 states the existence of a right inverse of the trace operator. See the presented references for more details.*

We are now ready to prove the equivalence between both problems:

Theorem 3.3.3. *Problem* (3.15) *is equivalent to* (3.16).

Proof. $(\implies)$:

[4] A description of such function will be given below, when we formalize the weak form of the problem. For now, assume that such extension has the regularity that we need.

Let $u \in V^0$ be a solution of problem (3.15). Define $u_1 = u_{|\Omega_1}$ and $u_2 = u_{|\Omega_2}$. It's clear that u_1 and u_2 satisfy equations 1 and 2 of (3.16).

For equation 3, we use the fact that $u \in H^1(\Omega)$ and therefore $u \in H^2_{\text{loc}}(\Omega)$. Given Ω' such that $\Omega' \Subset \Omega$ and $\gamma \subset \Omega'$, we have that $u \in H^2(\Omega')$ and $u \in C^{0,\lambda}(\overline{\Omega'})$ (u belongs to the Hölder Space of exponent λ; see Theorem 4.12 of [18] for more details.) As such, u is continuous on the interface and $u_1 = u_2$ on γ.

Finally, given $\mu \in \Lambda$, we can define the function

$$P\mu = \begin{cases} P_1\mu \\ P_2\mu \end{cases}$$

which satisfies $P\mu \in V^0$ and the equality 4 in (3.16).

$(\Longleftarrow)$:

Let $u_1 \in V_1^0, u_2 \in V_2^0$ be the two solutions of problem (3.15). We now set

$$u = \begin{cases} u_1, \text{ in } \Omega_1 \\ u_2, \text{ in } \Omega_2 \end{cases}.$$

In this case, again notice that for $i = 1, 2$ we have $u_i \in H^2_{\text{loc}}$ which implies that u_1 and u_2 is Hölder continuous and therefore u is also Hölder continuous since $u_1 = u_2$ on the interface, by assumption. Therefore, we also have that $u \in V^0$.

Defining $P\mu$ like above, we take $\mu \in \Lambda$ for some $v \in V$ such that $\mu = v_{|\gamma}$. As such, we have that the difference $(v_{|\Omega_i} - P_i\mu) \in V_i^0$ for each $i = 1, 2$ and

$$\begin{aligned} a(ku, v) &= \Big[a_1(k_1u_1, v_{|\Omega_1} - P_1\mu) + a_1(k_1u_1, P_1\mu)\Big] + \Big[a_2(k_2u_2, v_{|\Omega_2} - P_2\mu) + a_2(k_2u_2, P_2\mu)\Big] \\ &= \Big[(f_1, v_{|\Omega_1} - P_1\mu) + (f_1, P_1\mu)\Big] + \Big[(f_2, v_{|\Omega_2} - P_2\mu) + (f_2, P_2\mu)\Big] \\ &= (\tilde{f}, v), \implies a(u, v) = (f, v) \end{aligned}$$

where we take $\tilde{f} = \begin{cases} f_1, \text{ in } \Omega_1 \\ f_2, \text{ in } \Omega_2 \end{cases}$ as an auxiliary function and used the assumptions 1, 2, and 4 of equations (3.16). □

4

The Method of Fundamental Solutions

Contents

4.1 Density and linear independence results

The method of Fundamental Solutions is the base method to be implemented throughout this work. As a meshless method, it does not involve any kind of domain discretization (into some mesh) like in finite differences or finite elements methods. Instead, one faces a point placement problem. Since mesh generation is one of the most computationally expensive parts of the methods above, one of the main advantages of the MFS is exactly the lack of it.

As the name implies, MFS is based on the fundamental solutions of a previously known PDE. Consider the Elliptic linear differential operator $\mathcal{L}$ with fundamental solution Φ such that $\mathcal{L}\Phi(x) = \delta,\ \forall x \in \mathbb{R}^d$. Intuitively, we can consider the approximation

$$\tilde{u}(x) = \sum_{j=1}^{N} \alpha_j \Phi(x - y_j)$$

to the partial differential equation with a linear boundary operator $\mathcal{B}$

$$\begin{cases} \mathcal{L}u(x) = 0, & x \in \Omega \\ \mathcal{B}u(x) = 0, & x \in \partial\Omega \end{cases}.$$

By definition and using the linearity of the operator $\mathcal{L}$, $\tilde{u}$ satisfies the equation in Ω, and the coefficients α_j can be determined by imposing the boundary conditions

$$\mathcal{B}\tilde{u}(x) = \sum_{j=1}^{N} \alpha_j \mathcal{B}\Phi(x - y_j) = 0,$$

where $y_j \in \hat{\Gamma} \subset \mathbb{R}^d \setminus \overline{\Omega}$ with $j = 1, \ldots, N$, are the so-called source points to be chosen. We chose them to be outside our domain since, by translation, the Fundamental Solution Φ has a singularity at each y_j because, otherwise, it would render our approximation full of singularities inside Ω.

Remark 4.1.1. *While the motivation above might seem far-fetched, one can observe that the approximation $\tilde{u}$ resembles a convolution between some density function $\alpha(x)$ and the fundamental solution Φ. As we will see below, it can be proven that the fundamental solutions of the operator $\mathcal{L}$ are dense in the functional space defined in $\partial\Omega$ (which is enough since, as we already saw, the interior condition is satisfied by construction). For example, considering Dirichlet boundary conditions, one can use the single layer potential (which is going to be presented and studied in the next pages) that allows, through a discretization argument, to give a numerical approximation to the Boundary Value Problem (BVP) and*

$$u(x) = \int_{\Gamma} \Phi(x - y)\varphi(y)d\sigma(y) \approx \sum_{j=1}^{N} w_j \varphi(y_j)\Phi(x - y_j) = \tilde{u}(x), \tag{4.1}$$

where $\varphi(y)$ is a layer density to be determined and $y_j \in \hat{\Gamma}$ and w_j are the nodes and weights of some quadrature, respectively. Setting $\mathcal{B} = I$ and $\alpha_j = w_j\varphi(y_j)$, we recover the approximation given above.

First, we introduce the notion of *artificial boundary* (or *pseudo-boundary*), which is analyzed in [45].

Definition 4.1.2. *A source set $\hat{\Gamma}$ is said to be admissible if*

1. *$\hat{\Gamma} \subset \mathbb{R}^d \setminus \overline{\Omega}$ is an open set with components in each external part of Ω;*
2. *$\hat{\Gamma} = \partial\hat{\Omega}$ is the boundary of $\hat{\Omega}$, where $\hat{\Omega} \subset \mathbb{R}^d \setminus \overline{\Omega}$ is an open set with components in each external part of Ω. Note that the problem must be well-posed in $\hat{\Omega}$;*
3. *$\hat{\Gamma} \subset \partial\hat{\Omega}$, when $\partial\hat{\Omega}$ is an analytical boundary set verifying (2.) and $\hat{\Gamma}$ is open in the $\partial\hat{\Omega}$ topology.*

We denote the set of chosen source points by $\mathcal{Y} = \{y_j \in \hat{\Gamma} : j = 1, \dots, N\}$.

Throughout this work, the adopted source set will always be the second option since it provides better numerical results. However, notice that the density results below still hold when considering different types of admissible source sets.

Assume that $\mathcal{L} = -\Delta$ and Φ is the fundamental solution of the Laplace equation. Until further notice, we are working with Dirichlet boundary conditions. Of course, the results stated below are also valid with different boundary conditions. In the appropriate functional space and given an admissible source set $\hat{\Gamma}$, consider the approximation space

$$\mathcal{S}(\Gamma, \hat{\Gamma}) = \operatorname{span}\{\Phi(x - y)_{|x\in\Gamma} : y \in \hat{\Gamma}\}.$$

Some preliminary results are going to be needed in order to present the desired density proofs. We start by introducing the concept of analytic continuation. For further details see [46].

Definition 4.1.3. *Let f be a complex-valued function defined on $\Omega \subset \mathbb{C}$. We say that f is holomorphic on Ω if for every $a \in \Omega$, there exists a neighborhood U of a and $(c_n)_{n\in\mathbb{N}} \subset \mathbb{C}$ such that the power series*

$$\sum_{n=0}^{\infty} c_n (z - a)^n$$

converges to $f(z)$ for every $z \in U$.

Theorem 4.1.4 (Analytic continuation)**.** *Let f be a holomorphic function in the connected subset $\Omega \subset \mathbb{C}$. If there exists a non-empty $U \subset \Omega$ such that $f = 0$ in U, then $f = 0$ in Ω.*

As seen in the above remark, the study of layer potentials plays an important role in the Method of Fundamental Solutions, which are now formalized, and some results are stated. Since this is by itself a large topic see [19], [47] and [48] for more details.

Definition 4.1.5. *Let Ω be a bounded domain of class C^2 and $\varphi \in H^{\frac{1}{2}}(\Omega)$. The functions*

$$S\varphi(x) = \int_{\partial\Omega} \Phi(x - y)\varphi(y) d\sigma(y)$$

and

$$M\varphi(x) = \int_{\partial\Omega} \frac{\partial\Phi(x-y)}{\partial n}\varphi(y)d\sigma(y)$$

are called the single and double layer potentials with density φ, respectively.

Proposition 4.1.6. *Let Ω be a bounded domain of class C^2 and $\varphi \in H^{\frac{1}{2}}(\Omega)$. Then the single layer potential is harmonic in $\mathbb{R}^d \setminus \overline{\partial\Omega}$, continuous across $\partial\Omega$, and for every $x \in \partial\Omega$ we have the following jump relations for the normal derivative:*

$$\lim_{z\to x} \frac{\partial S\varphi}{\partial n^-}(z) = \int_{\partial\Omega} \frac{\partial\Phi(x-y)}{\partial n}\varphi(y)d\sigma(y) - \frac{1}{2}\varphi(x), \quad z \in \mathbb{R}^d \setminus \overline{\Omega}$$

and

$$\lim_{z\to x} \frac{\partial S\varphi}{\partial n^+}(z) = \int_{\partial\Omega} \frac{\partial\Phi(x-y)}{\partial n}\varphi(y)d\sigma(y) + \frac{1}{2}\varphi(x), \quad z \in \Omega.$$

In particular, we have

$$\varphi(x) = \frac{\partial S\varphi}{\partial n^+}(x) - \frac{\partial S\varphi}{\partial n^-}(x), \ \forall x \in \partial\Omega.$$

Analogously, the double layer potential is harmonic in $\mathbb{R}^d \setminus \overline{\partial\Omega}$, its normal derivative is continuous across $\partial\Omega$, and for every $x \in \partial\Omega$ we have the following jump relations:

$$\lim_{z\to x} M^+\varphi(z) = \int_{\partial\Omega} \frac{\partial\Phi(x-y)}{\partial n}\varphi(y)d\sigma(y) + \frac{1}{2}\varphi(x), \quad z \in \mathbb{R}^2 \setminus \overline{\Omega}$$

and

$$\lim_{z\to x} M^-\varphi(z) = \int_{\partial\Omega} \frac{\partial\Phi(x-y)}{\partial n}\varphi(y)d\sigma(y) - \frac{1}{2}\varphi(x), \quad z \in \Omega$$

In particular, we have

$$\varphi(x) = M^+\varphi(x) - M^-\varphi(x), \ \forall x \in \partial\Omega.$$

Lastly, it will be useful to study the well-posedness of the exterior Dirichlet problem for the Laplace Equation, c.f [49].

Theorem 4.1.7 (Well-Posedness of the Exterior Dirichlet problem)**.** *Let Ω be a bounded and open subset of $\mathbb{R}^2$. Then, there exists a unique solution $u \in C^2(\Omega^c) \cap C(\overline{\Omega^c})$ of the exterior Dirichlet Laplacian problem given by*

$$\begin{cases} \Delta u = 0, \ \textit{in} & \mathbb{R}^2 \setminus \overline{\Omega} \\ u = 0, \ \textit{on} & \partial\Omega \\ u(x) = \mathcal{O}(1), \ \textit{for} & |x| \to \infty. \end{cases}$$

Remark 4.1.8. *Notice that the condition at infinity must be enforced to have uniqueness. Otherwise, one could easily find a family of solutions up to a multiplicative constant (if $u(x)$ is a solution, then $\alpha u(x)$ would also be a solution for any $\alpha \in \mathbb{R}$).*

The main result which justifies the MFS for the Laplace equation is now stated. The proof given here is slightly different from the ones in [13], [45]. It is also influenced by the proofs in [50] and [51].

Theorem 4.1.9. *Let Ω be an open and bounded set with C^2 boundary $\Gamma = \partial\Omega$ such that $\overline{\Omega} \subset \hat{\Omega} \subset \mathbb{R}^2$, where $\hat{\Omega}$ is an open and bounded set and $\hat{\Gamma} = \partial\hat{\Omega}$ is an admissible source set. Then, $\mathcal{S}(\Gamma, \hat{\Gamma}) \oplus \mathbb{R}$ is dense in $H^{\frac{1}{2}}(\Gamma)$ and in $H^{-\frac{1}{2}}(\Gamma)$.*

Proof. Given Lemma 2.1.10 we start by fixing some notation. Let $E = H^{\frac{1}{2}}(\Gamma)$. For every (fixed) $y \in \hat{\Gamma}$, the maps

$$\Phi(\cdot - y) : \varphi \mapsto \int_\Gamma \Phi(x - y)\varphi(x)d\sigma(x)$$
$$1 : \varphi \mapsto \int_\Gamma \varphi(x)d\sigma(x)$$

are linear and continuous in $H^{\frac{1}{2}}(\Gamma)$, and $1, \Phi(\cdot - y) \in H^{-\frac{1}{2}}(\Gamma)$ (notice that $1, \Phi(\cdot - y) \in L^1_{\text{loc}}(\mathbb{R}^2)$ and $1, \Phi(\cdot - y) \in L^2(\hat{\Gamma})$).

Let $N = \mathcal{S}(\Gamma, \hat{\Gamma}) \oplus \mathbb{R} \subset H^{-\frac{1}{2}}(\Gamma)$. Using the Definition (2.1.9),

$$N^\perp = \{\varphi \in H^{\frac{1}{2}}(\Gamma) : \langle \psi, \varphi \rangle = 0,\ \forall \psi \in N\}$$

our goal is to prove that $N^\perp = \{0\}$. Let $\varphi \in N^\perp$ and consider $w(y) = \int_\Gamma \Phi(x - y)\varphi(x)d\sigma(x),\ y \in \mathbb{R}^2$. Then,

$$\int_\Gamma \Phi(x - y)\varphi(x)d\sigma(x) = 0,\ \forall y \in \hat{\Gamma} \tag{4.2}$$

and

$$\int_\Gamma \varphi(x)d\sigma(x) = 0. \tag{4.3}$$

In order to verify that $w(y)$ satisfies the exterior Laplace problem with Dirichlet boundary conditions, one can use the fact that w exhibits the asymptotic behavior

$$w(y) = -\frac{1}{2\pi}\int_\Gamma \varphi(x)d\sigma(x)\log|y| + \mathcal{O}(1),\ |y| \to \infty$$

and condition (4.3), to check that w is bounded at infinity. Therefore, by condition (4.2)

$$\begin{cases} \Delta w = 0, \text{ in } \mathbb{R}^2 \setminus \hat{\Omega} \\ w(y) = 0, \text{ on } \hat{\Gamma} \\ w(y) = \mathcal{O}(1), |y| \to \infty. \end{cases}$$

Since the problem above is well-posed, its unique solution is $w(y) = 0,\ \forall y \in \mathbb{R}^2 \setminus \overline{\hat{\Omega}}$. By (a unique) analytic continuation (see Theorem 4.1.4), we can extend w by zero in $\mathbb{R}^2 \setminus \overline{\Omega}$. Since w is a single layer potential over Γ, w is continuous on Γ and therefore, by continuity, $w = 0$ on Γ. Once again, using the fact that the single layer potential is harmonic in Ω, w satisfies the (inner) Laplace problem

$$\begin{cases} \Delta w = 0, \text{ in } \Omega \\ w(y) = 0, \text{ on } \Gamma \end{cases}$$

which, by uniqueness, implies that $w = 0$ in $\hat{\Omega}$. Finally, we can conclude that $\varphi = 0$ in Γ by Proposition 4.1.6 since the normal derivate jump is zero.

Therefore, we can write $N^{\perp} = \{0\}$ and by Lemma 2.1.10

$$\mathcal{S}(\Gamma, \hat{\Gamma}) \oplus \mathbb{R} = \{0\}^{\perp}$$

given the fact that $H^{\frac{1}{2}}(\Omega)$ is reflexive. Since $0 \in H^{\frac{1}{2}}(\Gamma)$ (and $0 \in H^{-\frac{1}{2}}(\Gamma)$) then $\mathcal{S}(\Gamma, \hat{\Gamma}) \oplus \mathbb{R}$ is dense in $H^{\frac{1}{2}}(\Gamma)$ and in $H^{-\frac{1}{2}}(\Gamma)$ (this is to be expected since $H^s(\Omega)$ is dense in $H^{-s}(\Omega)$ for $s > 0$). □

Remark 4.1.10. *The proof above guarantees the existence of a sequence of density layers $\{\varphi_n\} \subset H^{\frac{1}{2}}(\Gamma)$ and a sequence of constants $\{c_n\} \subset \mathbb{R}$ such that the modified single layer potential*

$$\hat{\mathcal{S}}\varphi_n(y) = \int_{\hat{\Gamma}} \Phi(x - y)\varphi_n(x) d\sigma(x) + c_n$$

converges to the Dirichlet boundary data $g(y)$ in $H^{\frac{1}{2}}(\Gamma)$, i.e,

$$\left\| \hat{\mathcal{S}}\varphi_{n|\Gamma} - g \right\|_{H^{\frac{1}{2}}(\Gamma)} \to 0, \ n \to \infty.$$

Since $\hat{\mathcal{S}}\varphi_n$ is harmonic for each $n \in \mathbb{N}$, every interior Dirichlet BVP can be approximated using the modified single layer potential. Conversely, any $\varphi \in H^{\frac{1}{2}}(\Gamma)$ and $c \in \mathbb{R}$ define a BVP whose solution is given by the associated modified single layer potential $\hat{\mathcal{S}}\varphi(y)$, with boundary data given by its restriction to Γ.

The density proof for the transmission/decomposition problem follows directly from the Theorem above and the equivalence formulation presented in Theorem 3.3.3. Consider the source function

$$f = \begin{cases} \frac{1}{k_1}, \text{ in } & \Omega_1 \\ \frac{1}{k_2}, \text{ in } & \Omega_2 \end{cases} \tag{4.4}$$

and let $u \in H_0^1(\Omega)$ be the associated unique weak solution to the Poisson equation with Dirichlet boundary conditions and a (discontinuous) source function f. By Theorem 3.3.3, the restriction of u to each subdomain uniquely solves the weak form of

$$\begin{cases} -\Delta u_i = \frac{1}{k_i}, & \text{in } \Omega_i \\ u_1 - u_2 = 0, & \text{on } \gamma \\ k_1 \frac{\partial u_1}{\partial n_1} + k_2 \frac{\partial u_2}{\partial n_2} = 0, & \text{on } \gamma \\ u_i = 0, & \text{on } \Gamma_i \end{cases}, \qquad \text{where } u = \begin{cases} u_1, \text{ in } & \Omega_1 \\ u_2, \text{ in } & \Omega_2, \end{cases} \tag{4.5}$$

and $u_i \in H^1(\Omega)$.

Since $H^{\frac{1}{2}}(\partial\Omega_i)$ is the trace space of $H^1(\Omega_i)$, and the normal derivatives of u_i belong to the space $H^{-\frac{1}{2}}(\partial\Omega_i)$ (as seen at the end of subchapter 2.3), by imposing the compatibility conditions on the interface γ, one can apply Theorem 4.1.9 to each subdomain in equation (4.5), and find a sequence of

density layers $\{\varphi_i^{(n)}\} \subset H^{\frac{1}{2}}(\partial\Omega_i)$ such that

$$\left\| \hat{\mathcal{S}}\varphi_i^{(n)}{}_{|\partial\Omega_i} - u_i \right\|_{H^{\frac{1}{2}}(\partial\Omega_i)} \to 0,\ n \to \infty,$$
$$\left\| M\varphi_i^{(n)}{}_{|\partial\Omega_i} - \frac{\partial u_i}{\partial n_i} \right\|_{H^{-\frac{1}{2}}(\partial\Omega_i)} \to 0,\ n \to \infty,$$

where $\hat{\mathcal{S}}$ is defined in Remark 4.1.10 and M is the double layer potential in Definition 4.1.5 (see Remark 4.1.15 for more details how different boundary conditions can be handled). Therefore, given the equivalence between both problems, it is possible to approximate the solution u of the Poisson equation with discontinuous source term using the MFS via the decomposition approach.

Finally, the discretization argument to be presented follows from the fact that given a set of source points $\mathcal{Y} = \{y_1, \dots, y_N\} \subset \mathbb{R}^2 \setminus \overline{\Omega}$, the fundamental solutions $\Phi(\cdot - y_1), \dots, \Phi(\cdot - y_N)$ are linearly independent on $\partial\Omega$ and therefore in Ω.

Theorem 4.1.11. *Let $\mathcal{Y}$ be a set of source points, as defined above. Then, the restriction of the functions $\Phi(\cdot - y_1), \dots, \Phi(\cdot - y_N)$ to $\partial\Omega$ are linearly independent.*

Proof. Assume that $\tilde{u}(x) = \sum_{j=1}^{N} \alpha_j \Phi(x - y_j) = 0,\ \forall x \in \partial\Omega$. We prove that $\alpha_1 = \cdots = \alpha_N = 0$. Since, by construction, $\tilde{u}$ satisfies the Laplace equation and by assumption $\tilde{u}(x) = 0,\ \forall x \in \partial\Omega$, by the well-posedness of the interior Dirichlet problem, $\tilde{u} = 0$ in $\overline{\Omega}$. Again, by analytic continuation, $\tilde{u} = 0$ in $\mathbb{R}^2 \setminus \mathcal{Y}$. Applying the Laplace operator to $\tilde{u}$, by linearity

$$\sum_{j=1}^{N} \alpha_j \delta y_j = 0$$

which implies that $\alpha_1 = \cdots = \alpha_N = 0$ by the linear independence of the Dirac deltas. □

Consider now the operator $\mathcal{L} = -(\Delta + k^2)$, and assume that k is **not** an eigenfrequency of the Helmholtz equation. The results presented for the fundamental solution of the Laplace Equation still hold for the fundamental solution of the Helmholtz equation. However, a different type of infinity conditions must be considered, the so-called *Sommerfeld Radiation Conditions*, c.f [48].

Theorem 4.1.12 (Well-Posedness of the Exterior Dirichlet problem of the Helmholtz Equation)**.** *Let Ω be a bounded and open subset of $\mathbb{R}^2$. Then, there exists a unique solution $u \in C^2(\Omega^c) \cap C(\overline{\Omega^c})$ of the exterior Dirichlet Helmholtz problem given by*

$$\begin{cases} -\Delta u = k^2 u, & \text{in } \mathbb{R}^2 \setminus \overline{\Omega} \\ u = 0, & \text{on } \partial\Omega \\ |x| \left(\frac{x}{|x|} \nabla u(x) - ik \right) u(x) = 0, & \text{for } |x| \to \infty. \end{cases}$$

Remark 4.1.13. *Just like the exterior Dirichlet problem for the Laplace Equation, the Well-Posedness of the exterior Helmholtz Problem depends on the conditions at infinity. In this case, they are known as Sommerfeld Radiation Conditions and are of the form*

$$|x|^{\frac{d-1}{2}}\left(\frac{x}{|x|}\nabla u(x) - ik\right)u(x) = 0, \text{ for } |x| \to \infty,$$

where d stands for the space dimension. We also notice that the single layer potential given by

$$\int_\Gamma \Phi_k(x-y)\varphi(x)d\sigma(x)$$

satisfies the Sommerfeld Radiation Condition, when $|y| \to \infty$.

Analogously to the Laplace problem, consider the space

$$\mathcal{S}(\Gamma, \hat{\Gamma}) = \text{span}\{\Phi_k(x-y)_{|x\in\Gamma} : y \in \hat{\Gamma}\}.$$

Again, like in Theorem 4.1.9, we point the reader to [45], [50] and [52], where slightly different proofs are stated.

Theorem 4.1.14. *Let Ω be an open and bounded set with C^2 boundary $\Gamma = \partial\Omega$ such that $\overline{\Omega} \subset \hat{\Omega} \subset \mathbb{R}^2$, where $\hat{\Omega}$ is an open and bounded set and $\hat{\Gamma} = \partial\hat{\Omega}$ is an admissible source set. Then, $\mathcal{S}(\Gamma, \hat{\Gamma})$ is dense in $H^{\frac{1}{2}}(\Gamma)$ and in $H^{-\frac{1}{2}}(\Gamma)$.*

Proof. This proof follows the same steps as in the proof of Theorem 4.1.9. Let $E = H^{\frac{1}{2}}(\Gamma)$. For every (fixed) $y \in \hat{\Gamma}$, the map

$$\Phi_k(\cdot - y) : \varphi \mapsto \int_\Gamma \Phi_k(x-y)\varphi(x)d\sigma(x)$$

is linear and continuous in $H^{\frac{1}{2}}(\Gamma)$ and $\Phi_k(\cdot - y) \in H^{-\frac{1}{2}}(\Gamma)$. Let $N = \mathcal{S}(\Gamma, \hat{\Gamma})\}$ and

$$N^\perp = \{\varphi \in H^{\frac{1}{2}}(\Gamma) : \langle \psi, \varphi\rangle = 0, \psi \in N\}.$$

Once again, we prove that $N^\perp = \{0\}$, i.e, it suffices to prove that given $\varphi \in H^{\frac{1}{2}}(\Gamma)$ the implication

$$\forall y \in \hat{\Gamma}, \int_\Omega \Phi_k(x-y)\varphi(x)d\sigma(x) = 0 \implies \varphi(x) = 0, \ \forall x \in \mathbb{R}^2$$

holds. Define

$$w(y) = \int_\Gamma \Phi_k(x-y)\varphi(x)d\sigma(x).$$

Given that w satisfies the Sommerfeld Radiation Conditions and, by assumption, $w(y) = 0$ in $\hat{\Gamma}$, then $w = 0$ in Ω is the unique solution of the exterior Dirichlet problem of the Helmholtz equation

$$\begin{cases} -\Delta w = k^2 w, & \text{in } \mathbb{R}^2 \setminus \overline{\Omega} \\ w = 0, & \text{on } \partial\Omega \\ |y|(\frac{y}{|x|}\nabla w(y) - ik)u(x) = 0, & \text{for } |y| \to \infty, \end{cases}$$

since k is not an eigenfrequency. By analytic continuation, we can extend w by zero to $\mathbb{R}^2 \setminus \overline{\Omega}$. The rest of the proof is the same as in the Theorem 4.1.9, using the fact that k is not an eigenfrequency and the interior Dirichlet problem is well-posed. □

Remark 4.1.15. *Both in Theorem 4.1.9 and Theorem 4.1 the density proof can generalize to $H^s(\Gamma)$, for $s \geq \frac{1}{2}$. However, in applications, we are only interested in the case $s = \frac{1}{2}$. Notice that if $s = 0$, it is not required to invoke the Hahn-Banach Theorem since $H^0(\Gamma) = L^2(\Gamma)$ which is a Hilbert Space, and it would suffice to use Corollary (2.2.3). For general boundary conditions, the proofs above follow the same argument, where one should consider the appropriate approximating set $\mathcal{S}(\Gamma, \hat{\Gamma})$ and the appropriate integral operator (for example, for Neumann boundary conditions, one should consider the set*

$$\mathcal{S}(\Gamma, \hat{\Gamma}) = \text{span}\{\partial_n \Phi(x - y)_{|x \in \Gamma} : y \in \hat{\Gamma}\}$$

and the double layer potential $M\varphi$, and recall that $\partial_n u = g \in H^{-\frac{1}{2}}(\Gamma)$).

Once again, the discretization argument follows from the linear independence of the functions $\Phi_k(\cdot - y_1), \ldots, \Phi_k(\cdot - y_1)$, where $y_1, \ldots, y_N \in \mathbb{R}^2 \setminus \overline{\Omega}$ are distinct source points. The proof is identical to the one presented in Theorem 4.1.11, where we use the fact that k is not an eigenfrequency.

Before stating some results regarding the convergence and the stability of the MFS for the Laplace and Helmholtz equations, we address the problem of the source points placement. Although different methods can be considered, e.g. [45], throughout this work we place the artificial boundary $\hat{\Gamma}$ over the boundary of Ω. To be more precise, consider $\Gamma = \partial\Omega$ and the (equally spaced) colocation points $x_1, \ldots, x_M \in \Gamma$. Then, we approximate the (outward) normal vector $\tilde{n}_i$ to the point x_i, which is given by

$$\tilde{\mathbf{n}}_i = \frac{(x_i - x_{i-1})^\perp}{2} + \frac{(x x_{i+1} - x_i)^\perp}{2},$$

with the orthogonal notation $z^\perp = (-z_2, z_1)$. This way, one can approximate the unit normal vector by $\mathbf{n} = \frac{\tilde{\mathbf{n}}}{|\tilde{\mathbf{n}}|}$ and define the source point y_i by

$$y_i = x_i + \eta \mathbf{n},$$

where $\eta > 0$ is some small coefficient that controls the distance from each point in Γ and $\hat{\Gamma}$. As we shall see below, this coefficient has an important impact on the convergence of the MFS, where bigger values of η produce better approximations. However, this is only feasible for simple geometries, because it also increases the condition number of matrix A, denoted by $\kappa(A)$.

4.2 Numerical approach for the Laplace Equation

Given the results above, it is possible to numerically solve the Laplace and Helmholtz equations for any boundary data $g(x)$ if we can find the coefficients in the discretization of the single layer potential. For

simplicity, we still assume Dirichlet boundary conditions. Let N be the number of source points and M the number of colocation points on the boundary, denoted by $x_1, \ldots, x_i, \ldots, x_M$ with $i = 1, \ldots, M$. Then, we solve the discretized equation

$$\tilde{u}(x_i) = \sum_{j=1}^{N} \alpha_j \Phi(x_i - y_j) + \alpha_{N+1} = g(x_i)$$

with respect to the coefficients α_j. Defining $g_i := g(x_i)$, $i = 1, \ldots, M$, notice that the equation above can be rewritten in the matricial form

$$\underbrace{\begin{bmatrix} \Phi(x_1, y_1) & \cdots & \Phi(x_1, y_N) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \Phi(x_M, y_1) & \cdots & \Phi(x_M, y_N) & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \\ \alpha_{N+1} \end{bmatrix}}_{\alpha} = \underbrace{\begin{bmatrix} g_1 \\ \vdots \\ g_M \end{bmatrix}}_{g} \tag{4.6}$$

$M \times (N+1)$ $(N+1) \times 1$ $M \times 1$

This problem can be framed in two different ways:

- if $N + 1 = M$, then we are faced with an interpolation problem, where we solve a linear system of N equations and N unknowns;
- if $M > N + 1$, then we must solve a least-squares problem. Observe that the case $M < N + 1$ is an under-determined system of equations and therefore the number of colocation points must be greater than the number of source points. In any case, every numerical linear algebra software can solve this type of problem efficiently. This is the method used in this work, since it avoids interpolation instabilities, forces the boundary condition to hold at some specific points, and is particularly robust when dealing with non-regular boundary data. After several numerical experiments, e.g. [45], it was concluded that $N = 2M$ is a good choice for the number of source points.

Notice that different boundary conditions can be considered: for example, if solving the Laplace equation with Neumann boundary conditions, one should replace the entries $\Phi(x_i, y_j)$ with $\partial_{n_x} \Phi(x_i, y_j) = \nabla_x \Phi(x_i, y_j) \cdot n$, where n is the unit normal vector to the boundary in the point x_i.

We now state some results regarding the convergence and the stability of the MFS for the Laplace equation with Dirichlet boundary conditions. Here, we state this results when the domain Ω is a disk and the artificial boundary $\hat{\Gamma} = \partial\hat{\Omega}$ that involves Ω is also a circle. Let ρ be the radius of Ω, R the radius of $\hat{\Omega}$ and assume that $N = M$.

Theorem 4.2.1. *Assume that $R^N - \rho^N \neq 1$. Then,*

1. *the matrix A is non-singular;*
2. *if $R \neq 1$ and the boundary data g is real and analytic we can also prove that the exact solution u of the Laplace equation admits a harmonic extension to some neighborhood of $\overline{\Omega}$. Therefore, we*

may assume that u is harmonic in $0 \leq r \leq r_0$ for some $r_0 \geq \rho$. In this case, there exists $C > 0$ and $c \in (0,1)$ which are independent of N and u such that

$$\sup_{x \in \overline{\Omega}} |u(x) - \tilde{u}(x)| \leq C c^N \sup_{|x| \leq r_0} |u(x)|.$$

The Theorem above provides some important insights regarding the MFS. First, we cannot fail to notice that this method displays *exponential convergence* in the number of source points, which is quite remarkable. In fact, the term c depends on the distance between Γ and the artificial boundary $\hat{\Gamma}$, which is controlled by the coefficient η, and

$$c = \begin{cases} \frac{\rho}{R}, \text{ if } r_0 > \frac{R^2}{\rho} \\ \sqrt{\frac{\rho}{R}}, \text{ if } r_0 < \frac{R^2}{\rho}. \end{cases}$$

Unfortunately, one of the main drawbacks of the MFS is the ill-conditioning of the matrix A and the fact that the matrix A is very dense, and we cannot use optimized sparse software solvers on the system (4.6). In particular, while bigger values of the parameter η allow for better numerical approximations it also implies an exponential growth[1] of the condition number $\kappa(A)$, see [55], [56] and [57].

Theorem 4.2.2. *In the conditions of the Theorem 4.2.1, the condition number can be estimated by*

$$\kappa(A) \sim \frac{\log R}{2} N \left(\frac{R}{\rho}\right)^{\frac{N}{2}}.$$

Another interesting remark, is the condition $R \neq 1$, which is in direct connection with the space $\mathcal{S}(\Gamma, \hat{\Gamma}) \oplus \mathbb{R}$ that was proven to be dense in $H^{\frac{1}{2}}(\Gamma)$ in Theorem 4.1.9. Assume that $R = 1$ and $\hat{\Omega}$ is a disk with radius R that contains the origin. Then, if one does not consider the constant basis function 1,

$$\tilde{u}(0) = \sum_{j=1}^{N} \alpha_j \Phi(0 - y_j) = -\frac{1}{2\pi} \sum_{j=1}^{N} \alpha_j \log(R) = 0,$$

no matter the choice of the source points over $\hat{\Gamma}$. Therefore, it is impossible to approximate any harmonic function which does not vanish on the origin. However, this is not the case if we add the basis function 1, which was used to prove the density result. While this rarely interferes with the numerical approximations in the next chapters, it will be considered nevertheless by a reason of coherence. This is the reason why a column of ones was added in the matrix A.

4.2.1 An enrichment technique

Before diving into the numerical approach for the Helmholtz equation, we introduce some modifications to the classical MFS method presented above. There is another drawback in our method: our basis

[1] It is very important to remark that the ill-conditioning of the matrix A is associated with the classical basis functions used in the MFS. However, it may be possible to derive basis functions that, after certain manipulations, do not exhibit the same limitations and have a condition number $\kappa(A)$ of $\mathcal{O}(1)$. We refer the reader to [53] and [54] for more information.

functions are analytical and might lose precision when approximating functions that display singularities, for example near a corner if the domain is not smooth. In what follows we introduce an enrichment technique that allows for singularity treatment. In the same vein as in Proposition 3.2.3, we reintroduce the notion of a wedge domain and consider some wedge-like domain with interior angle Θ.

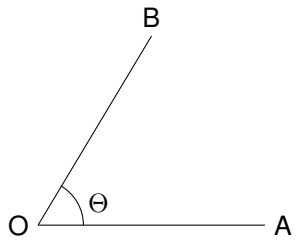


Figure 4.1: A wedge domain with an interior angle Θ.

Consider the Laplace equation in polar coordinates, given by

$$\left(\partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2\right) u(r,\theta) = 0, \quad r > 0,\ 0 \leq \theta \leq \Theta. \tag{4.7}$$

Then, by separation of variables $u(r,\theta) = R(r)T(\theta)$, one can find two different families of particular solutions given by,

$$u(r,\theta) = \left(c_1 r^{\alpha} + c_2 r^{-\alpha}\right) \times \left(c_3 \cos(\alpha\theta) + c_4 \sin(\alpha\theta)\right),\ \alpha > 0$$

and

$$u(r,\theta) = \left(c_1 \log(r) + c_2\right) \times \left(c_3 \cos(\alpha\theta) + c_4 \sin(\alpha\theta)\right),\ \alpha = 0$$

where $c_1, c_2, c_3, c_4 \in \mathbb{C}$. In order to find α, one must consider the amplitude of the angle Θ and the boundary conditions at each segment $\overrightarrow{OA}$ and $\overrightarrow{OB}$. Below we summarize the asymptotic harmonic solutions of (4.7). For more details, we point the reader to [1].

- For Dirichlet-Dirichlet boundary conditions given by $u(r,0) = A, u(r,\Theta) = B$, then $\alpha_k = \frac{k\pi}{\Theta}$ and

$$u(r,\theta) = A(B-A)\frac{\theta}{\Theta} + \sum_{k=0}^{\infty} \alpha_k r^{\alpha_k} \sin(\alpha_k \theta);$$

- For Dirichlet-Neumann boundary conditions given by $u(r,0) = A, \partial_n u(r,\Theta) = B$, then $\alpha_k = \frac{\left(k+\frac{1}{2}\right)\pi}{\Theta}$ and

  - If $\Theta \neq \frac{\pi}{2}, \frac{3\pi}{2}$,

$$u(r,\theta) = A + \frac{B}{\cos(\Theta)} r \sin(\theta) + \sum_{k=0}^{\infty} \alpha_k r^{\alpha_k} \sin(\alpha_k \theta);$$

– If $\Theta = \frac{\pi}{2}, \frac{3\pi}{2}$,

$$u(r,\theta) = A + (-1)^{l+1}\frac{Br}{\Theta}\left(\log(r)\sin(\theta) + \theta\cos(\theta)\right) + \sum_{k=0}^{\infty}\alpha_k r^{\alpha_k}\sin(\alpha_k\theta),$$

with $l = 0$ if $\Theta = \frac{\pi}{2}$ and $l = 1$ if $\Theta = \frac{3\pi}{2}$;

- For Neumann-Dirichlet boundary conditions given by $\partial_n u(r,0) = A, u(r,\Theta) = B$, then $\alpha_k = \frac{\left(k+\frac{1}{2}\right)\pi}{\Theta}$ and

  – If $\Theta \neq \frac{\pi}{2}, \frac{3\pi}{2}$,

$$u(r,\theta) = B - Ar\sin(\theta) + \frac{A\sin(\Theta)}{\cos(\Theta)}r\cos(\theta) + \sum_{k=0}^{\infty}\alpha_k r^{\alpha_k}\cos(\alpha_k\theta);$$

  – If $\Theta = \frac{\pi}{2}, \frac{3\pi}{2}$,

$$u(r,\theta) = B - \frac{Ar}{\Theta}\left(\log(r)\cos(\theta) - \theta\sin(\theta)\right) - Ar\sin(\theta) + \sum_{k=0}^{\infty}\alpha_k r^{\alpha_k}\cos(\alpha_k\theta);$$

- For Neumann-Neumann boundary conditions given by $\partial_n u(r,0) = A, \partial_n u(r,\Theta) = B$, then $\alpha_k = \frac{k\pi}{\Theta}$ and

  – If $\Theta \neq \pi, 2\pi$,

$$u(r,\theta) = -Ar\sin(\theta) - \frac{B + A\cos(\Theta)}{\sin(\Theta)}r\cos(\theta) + \sum_{k=0}^{\infty}\alpha_k r^{\alpha_k}\cos(\alpha_k\theta);$$

  – If $\Theta = \pi, 2\pi$,

$$u(r,\theta) = -Ar\sin(\theta) + \frac{(-1)^l B - A}{\Theta}r\left(\log(r)\cos(\theta) - \theta\sin(\theta)\right) + \sum_{k=0}^{\infty}\alpha_k r^{\alpha_k}\cos(\alpha_k\theta);$$

with $l = 0$ if $\Theta = \pi$ and $l = 1$ if $\Theta = 2\pi$.

Remark 4.2.3. *Observe that if the wedge domain is rotated by some angle θ_1 (see Figure (4.2)) one can consider the translation $\theta^\star = \theta - \theta_1$, where $\theta^\star$ is the angle on the "correct" wedge domain, see Figure (4.1).*

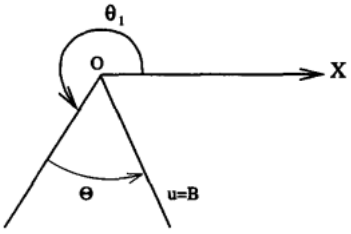


Figure 4.2: Rotation of the wedge domain. Image taken from [1].

In applications, we are mostly concerned with Dirichlet-Neumann and Neumann-Dirichlet boundary conditions. Just like stated above, some of these boundary conditions have different expansions for the angles $\Theta = \frac{\pi}{2}$ *and* $\Theta = \frac{3\pi}{2}$*. However, we are only concerned with the expansion*

$$v(r,\theta) = \sum_{k=0}^{\infty} \alpha_k r^{\alpha_k} \psi(\alpha_k \theta)$$

where $\psi = \sin$ *or* $\psi = \cos$*. In these cases, if we neglect the other terms, for the special angle* $\Theta = \frac{\pi}{2}$ *above we would find that* $\alpha_k \in \mathbb{N}$*. Without going into much dep2t in singularity analysis, then its partial derivative* $\partial_r v(r,\theta)$ *would be of the form*

$$\partial_r v(r,\theta) = \sum_{k=0}^{\infty} \alpha_k^2 r^{\alpha_k - 1} \psi(\alpha_k \theta)$$

where $\alpha_k - 1 \in \mathbb{N}$*. In general, if* $\alpha_k \in \mathbb{N}$ *for some angle* Θ *then all of its derivatives are continuous and* $v(r,\theta)$ *is analytical. More precisely, there is no singularity in these cases. Therefore, one does not need to enrich the set of basis functions since the fundamental solutions correctly reproduce the behavior near the corner's tip. Such corners are called regular. On the other hand, corners that present singularities are called singular and are the ones that we are interested to approximate.*

Also notice that, in the expansions above, the term $r^{-\alpha_k}$ *does not appear. This has to do with the fact we are dealing with an interior problem: when considering the exterior problem, the terms* r^{α_k} *are replaced with* $r^{-\alpha_k}$ *(observe that it satisfies the asymptotic conditions prescribed in order to have well-posedness!).*

To incorporate the singular behavior near a singular corner, first assume, without loss of generality, that the domain Ω has just one corner and that the solution of our BVP can be decomposed in regular and singular parts,

$$u(x) = u_R(x) + u_S(x),\ x \in \overline{\Omega},$$

where u_R is the regular part approximated by the MFS basis functions and the singular part u_S is approximated by the expansions above, having the boundary conditions into account. Let $\phi_s(r,\theta)$ be one of those expansions centered at the corner's tip, where s is the order of the expansion. Then, the numerical approximation can be written as

$$\tilde{u}(x) = \sum_{j=1}^{N} \alpha_j \Phi(x - y_j) + \alpha_{N+1} + \sum_{s=1}^{p} \beta_s \phi_s(r(x),\theta(x)),\ x \in \overline{\Omega}. \tag{4.8}$$

Considering the collocation points $x_1, \ldots, x_M \in \partial\Omega$, the linear system of equations (4.6) can be generalized to

$$\underbrace{\begin{bmatrix} A_1 \mid B_1 \end{bmatrix}}_{A}_{M\times(N+1+P)} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{(N+1+P)\times 1} = \begin{bmatrix} g \end{bmatrix}_{M\times 1}, \tag{4.9}$$

where the block matrix A_1 is the matrix A in (4.6) and the B_1 block matrix is given by

$$B_1 = \begin{bmatrix} \phi_1(r(x_1),\theta(x_1)) & \cdots & \phi_p(r(x_1),\theta(x_1)) \\ \vdots & \ddots & \vdots \\ \phi_1(r(x_M),\theta(x_M)) & \cdots & \phi_p(r(x_M),\theta(x_M)). \end{bmatrix}$$

4.3 Numerical approach for the Helmholtz Equation

For the Helmholtz equation, the MFS convergence and stability results resemble the previous ones for the Laplace equation with Dirichlet boundary conditions. Once again, the results are stated for identical geometries as before, where Ω is the unit disk and the radius of $\hat{\Omega}$ is $R > 1$. Here is assumed that the boundary data g can be analytically extended to the annulus $\{z \in \mathbb{C} : \frac{1}{\rho} < |z| < \rho\}$, where $\rho > 1$. The following result is due to [58].

Theorem 4.3.1. *Let $R > 1$ and N be an even number. Then the minimum boundary error achieved by the MFS in the unit disk satisfies*

$$\left\| g - \tilde{u}_{|\Gamma} \right\|_{L^2(\Gamma)} \leq \begin{cases} C\rho^{-\frac{N}{2}}, \text{ if } \rho < R^2 \\ C\sqrt{N}R^{-N}, \text{ if } \rho = R^2 \\ CR^{-N}, \text{ if } \rho > R^2 \end{cases}$$

where C is a constant that not depends on N.

To solve the Helmholtz equation with the MFS, one must start by computing the eigenvalues λ (or the eigenfrequencies k, with $\lambda = k^2$) first. In order to achieve that, recall that the Helmholtz equation

$$\begin{cases} -\Delta u = k^2 u, \text{ in } \Omega \\ u = 0, \text{ on } \partial\Omega \end{cases} \tag{4.10}$$

is well-posed when λ is not an eigenvalue, and in that case the nullspace of the single layer potential operator

$$S_k\varphi(y) = \int_{\hat{\Gamma}} \Phi_k(x-y)\varphi(x)d\sigma(x)$$

is trivial. More precisely, one can prove the following result, e.g. [59].

Theorem 4.3.2. *If k is not an eigenfrequency of the interior Dirichlet problem, then* $\dim(N(S_k)) = 0$.

Proof. If k is not an eigenfrequency, then the interior problem is well posed which implies that $\varphi(x) = 0,\ x \in \overline{\Omega}$. By analytical continuation, $\varphi(x) = 0,\ x \in \hat{\Omega}$. Since the single layer potential is continuous, then $\varphi(x) = 0,\ x \in \hat{\Gamma}$. By the well-posedness of the exterior problem (notice that the single layer potential satisfies the Sommerfeld radiation conditions), then $\varphi(x) = 0,\ \forall x \in \mathbb{R}^2$. □

This theorem can be used to search for the eigenvalues/eigenfrequencies of the Laplace operator. By virtue of the discretization of the single layer potential (4.1), one should find the values k such that

the nullspace of the matrix $A(k) = \left[\Phi_k(x_i - y_j)\right]_{M\times N}$ is not trivial. Like it was discussed in the previous section 4.2, that can be done in two different ways:

1. if $A(k)$ is a square matrix (with $M = N$), one can compute the determinant of $A(k)$. Since the components of $A(k)$ are complex numbers, then its determinant is also a complex number, and we consider its absolute value. In any case, instead of working with $|\det A(k)|$, since this value is very small, one must work with its logarithm and consider the function $d(k) = \log|\det A(k)|$;

2. if $A(k)$ is an $M \times N$ rectangular matrix, with $M > N$, one considers the smallest singular value, which we denote by $\sigma_N(k)$, where the singular values of $A(k)$ are assumed to be in decreasing order $\sigma_1(k) \geq \cdots \geq \sigma_N(k)$. We emphasize that we only work with this case.

Therefore, in order to find the eigenvalues/eigenfrequencies of the Laplace operator, one must find the singularities of the functions $d(k)$ or $\sigma_N(k)$ for the first and second cases above, respectively, i.e, study the local minima of the functions $d(k)$ or $\sigma_N(k)$ which, in very regular domains, should approach zero.

To search for these singularities, a simple direct search algorithm was developed to bracket the set of local minima in a given interval. The iterative algorithm used is based on the Golden Ratio Search already used in [59]. First, consider the graph of $\sigma_N(k)$ in a given interval $I = (a, b)$ (which likely has more than one local minima) and fix (a relatively large) step size h. Let $I^0_{M_0} = \{a^0_0, \ldots, a^0_{M_0}\}$ be the discretization of $I^0_{M_0}$ in M_0 points spaced by h in the zeroth iteration, and denote the set of local minimums of I_{M_0} by $X^0_K = \{x^0_0, \ldots, x^0_K\} \subset I_{M_0}$, i.e, $\forall i = 1, \ldots, K$ there exists $c^0, d^0 \in I^0_{M_0}$ such that $c^0 = x^0_i - h$, $d^0 = x^0_i + h$ and $x_i < \min(c^0, d^0)$. Then, for each x^0_i, $I^0_{M_0}$ is enlarged with the middle points between c^0 and x^0_i, d^0 and x^0_i, and it is denoted by $\tilde{I}^1_{M_1}$ (notice that $M_0 < M_1$)[2]. Finally, we sort $\tilde{I}^1_{M_1}$ in increasing order and repeat the process for a specified depth d. When the maximum depth is attained, one was able to successfully find small intervals which bracket *each local minimum* of σ_N in I, and it is now possible to apply a direct search method to find it to any desired precision: in this work, Brent's method was used [60], although Golden Ratio Search is also possible. The general form of the algorithm is given in 4.1.

Remark 4.3.3. *Although the algorithm above is enough, we point out some problems and possible changes to consider. First, there are no guarantees regarding the eigenvalues found: this depends on the chosen interval and the step. For example, if the first eigenvalue is not in the chosen interval, it will never be found. One may also find a "jump" if two eigenvalues are arbitrarily close, but the algorithm just found one of them: this is the case when an eigenvalue has a multiplicity bigger than one, which depends on the domain. However, this last case is easy to spot if one considers the one parameter*

[2] Given the nature of the algorithm, one can also consider the left adjacent point to c^0, which we denote by $e_0 \in I^0_{M_0}$, compute and add their middle point to $\tilde{I}^1_{M_1}$

Algorithm 4.1: Direct Bracketing Algorithm

Set maximum depth d

Set step size h

Set bracketing interval $I = (a, b)$

begin

Discretize I into $I^0_{M_0} = \{a^0_0, \ldots, a^0_{M_0}\}$, where $a_{j+1} = a_j + h$ for each $j = 1, \ldots M_0$

Set $s = 0$

`(Bracketing step)`

while $s < d$ **do**

Compute $X^s_{K_s} = \{x^s_0, \ldots, x^s_{K_s}\}$, the set of local minima of $I^s_{M_s}$

Let E be an empty array

foreach $x \in X^s_{K_s}$ **do**

Consider the adjacent points to x given by $c, d \in I^s_{M_s}$ such that $c < x < d$

Consider the left adjacent point to c, denoted by e

Compute $\tilde{c} = \frac{c+x}{2}$

Compute $\tilde{d} = \frac{d+x}{2}$

Compute $\tilde{e} = \frac{e+c}{2}$

Append $\tilde{c}$, $\tilde{d}$ and $\tilde{e}$ to E

Let $\tilde{I}^{s+1}_{M_{s+1}} = I^s_{M_s} \cup E$

Define $I^{s+1}_{M_{s+1}} = sort(\tilde{I}^{s+1}_{M_{s+1}})$

$s \leftarrow s + 1$

`(Direct search step)`

Let E be an empty array

foreach $x \in X^d_{K_d}$ **do**

Consider the adjacent points to x given by $c, d \in I^d_{M_d}$ such that $c < x < d$

$\lambda = Brent(c, d)$

Append λ to E

return E

transformation between the domain Ω and the unitary disk $\mathbb{D}$ given by

$$\Omega(t) = (1 - t)\mathbb{D} + t\Omega,\ t \in (0, 1),$$

since the graph of the eigenvalues is continuous concerning t. Regarding the complexity of the algorithm, the objective is to evaluate the function $\sigma_N(k)$ the least amount of time: a brute-force approach with a

very small step is not feasible since each evaluation takes some time for big complex-valued matrices. However, we note that the algorithm can be improved by just considering the initial point a and "walking forward" with step size h. In that case, one could bracket each local minimum individually and only break to the main loop when the maximum depth was reached for each local minimum initially found. The loop would ***break*** *when the number of eigenvalues found attains a prescribed value.*

Before introducing a new technique to improve the condition of the system, we state a posteriori estimate based on a result proved by Moler and Payne in [61].

Theorem 4.3.4. *Let $\tilde{k}$ and $\tilde{u} \in C^2(\Omega) \cap C(\overline{\Omega})$ be an approximate eigenfrequency and eigenfunction which satisfy the following problem:*

$$\begin{cases} -\Delta\tilde{u} = k^2 u, & \text{in } \Omega \\ u = \xi(x), & \text{on } \Gamma. \end{cases}$$

Then, there exists an eigenfrequency k_n of (4.10) such that

$$\frac{\left|k_n - \tilde{k}\right|}{|k_n|} \leq \theta$$

where

$$\theta = \frac{\sqrt{|\Omega|}\|\xi\|_{L^\infty(\Gamma)}}{\|\tilde{u}\|_{L^2(\Omega)}}.$$

If, in addition, $\tilde{u}_{L^2(\Omega)}$ and u is the normalized orthogonal projection of $\tilde{u}$ onto the eigenspace of k_n, then

$$\|\tilde{u} - u\|_{L^2(\Omega)} \leq \frac{\theta}{\rho_n}\left(1 + \frac{\theta^2}{\rho_n^2}\right)^{\frac{1}{2}},$$

where

$$\rho_n = \min_{k_n \neq k_p} \frac{\left|k_p^2 - \tilde{k}^2\right|}{k_p^2}, \text{ for } p \in \mathbb{N}.$$

4.3.1 The Subspace Angle Technique

As stated before, one of the drawbacks of this method is the ill-conditioning of the system. In this subsection we introduce the so-called *Subspace Angle Technique*, first presented in [6]. Intuitively, there are two problems at play: firstly, while the MFS only needs the boundary data to approximate the solution of the BVP, the exponential growth of the condition number against the exponential convergence can be seen has the lack of information given by the collocation points on the boundary, which is not enough to decide if an approximate eigenfunction is spurious; secondly, while we proved the linear independence of the basis functions, in practice the columns of the matrix $A(k)$ are almost linear dependent if its number is too large (in fact, this is, once again, associated with the distance from the boundary to the artificial boundary).

To solve the first problem, we add additional interior points in order to over-determine the system; and for the second problem, we construct an orthonormal basis of the column space of $A(k)$, denoted by $\mathcal{C}(A(k))$, using the QR factorization of $A(k)$. Let M_B be the number of boundary points and M_I the number of interior points, such that $M = M_B + M_I$. Then, by adding some interior points the matrix $A(k)$ can be extended to

$$A(k) = \begin{bmatrix} A_B(k) \\ A_I(k) \end{bmatrix},$$

where the indices B and I correspond to the block matrices with the boundary and interior collocation points, respectively. To generate an orthonormal basis of the column space of $A(k)$, consider the QR factorization of $A(k)$, given by $A(k) = Q(k)R$, where $Q(k)$ is a unitary complex matrix (i.e., $Q^{\dagger}(k) = Q^{-1}(k)$) and R is an upper triangular matrix. By partitioning $Q(k)$ in the boundary and interior collocation points, we also have

$$Q(k) = \begin{bmatrix} Q_B(k) \\ Q_I(k) \end{bmatrix},$$

and each unit vector $u \in \mathcal{C}(A(k))$ has the form

$$u = \begin{bmatrix} u_B \\ u_I \end{bmatrix} = Q(k)v = \begin{bmatrix} Q_B(k) \\ Q_I(k) \end{bmatrix} v \tag{4.11}$$

for some $v \in \mathbb{R}^2$, $\|v\| = 1$. Assuming homogeneous Dirichlet boundary conditions, we are interested in non-trivial solutions $v \in \mathbb{R}^2$ to the above problem when $u \approx 0$ at the boundary, i.e, to solve the constrained minimization problem

$$\min_{v \in \mathbb{R}^2, \|v\|=1} \|Q_B(k)v\|.$$

The above problem is easy to solve and has a closed-form solution which can be found using Lagrange multipliers. The solution $\check{v}$ is the right singular vector associated with the smallest singular value σ_N and

$$\sigma_N(k) = \|Q_B(k)\check{v}\|.$$

Let $\check{u} = Q(k)\check{v}$. Taking the norm on both sides of equation (4.11), one can write

$$1 = \|\check{u}\|^2 = \left\| \begin{bmatrix} Q_B(k) \\ Q_I(k) \end{bmatrix} \check{v} \right\|^2 = \sigma_N^2(k) + \|Q_I(k)\check{v}\|^2. \tag{4.12}$$

Notice how equation (4.12) can be used to eliminate spurious solutions: since $0 < \sigma_N < 1$, if $\sigma_N \approx 1$, then $Q_I(k)\check{v} \approx 0 \implies u_I \approx 0$ which is an incorrect solution (is zero on the interior and does not satisfy the boundary constraints); on the other hand, if $\sigma_N \approx 0$, then we found an eigenfunction which is small on the boundary points and whose interior is not null.

The name Subspace Angle Technique comes from the fact that σ_N is related to the angle between the subspaces $\mathcal{C}(A(k))$ and $\mathcal{G}_0$, the space of vectors that are zero at boundary points[3]. The angle $\phi(k) = \angle(\mathcal{C}(A(k)), \mathcal{G}_0)$ between both subspaces is defined by

$$\cos\phi(k) = \sup_{\substack{u \in \mathcal{C}(A(k)),\, \|u\|=1 \\ v \in \mathcal{G}_0,\, \|v\|=1}} (u, v),$$

[3] $\mathcal{G}_0$ can be seen as the discretization of the functions which satisfy the boundary conditions but not the Helmholtz equation.

and one can prove (c.f. [6]) that

$$\sigma_N = \sin \phi(k).$$

Therefore, the discrete problem has a non-trivial solution (i.e. λ is an eigenvalue of the Laplace Operator) if and only if $\mathcal{C}(A(k))$ and $\mathcal{G}_0$ have a non-trivial intersection (i.e. $\phi(k) = l\pi$, $l \in \mathbb{Z}$).

Remark 4.3.5. *While the construction above assumed homogeneous Dirichlet boundary conditions, it can be easily generalized to any type of homogeneous boundary conditions $\mathcal{B}$ by considering the appropriate A matrix.*

Remark 4.3.6. *Neither the enrichment technique nor the Subspace Angle Technique are specific methods only applicable to the Laplace equation and the Helmholtz equation, respectively. They can be used for both equations at the same time. For example, in [62], both methods were used to study the spectrum of the Laplace operator in domains with corners and cracks.*

5

Conducted Numerical Simulations

Contents

5.1 Dirac equation simulations

Recall the system of equations given in (3.5)

$$\begin{bmatrix} m & -i(\partial_1 - i\partial_2) \\ -i(\partial_1 + i\partial_2) & -m \end{bmatrix} \begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix} = \lambda \begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix}, \tag{5.1}$$

which can be reduced to the Helmholtz equation with Cauchy-Riemann oblique boundary conditions

$$\begin{cases} -\Delta u_1 = (\lambda^2 - m^2)u_1, & \text{in } \Omega \\ i(\partial_1 + i\partial_2)u_1 + (\lambda + m)i(n_1 + in_2)u_1 = 0, & \text{on } \Gamma, \end{cases}$$

where $\Gamma = \partial\Omega$ as usual, and the function u_2 depends on the function u_1 by the following equality

$$u_2 = \frac{-i(\partial_1 + i\partial_2)u_1}{\lambda + m}.$$

5.1.1 Method validation

In order to validate the MFS for the Dirac equation with infinite mass boundary conditions, we start by testing it for the unit disk. If $m = 0$, then its value is known like stated in Proposition 3.2.1, where $\lambda_1(\mathbb{D}) = 1.434695650819$ is the solution of the equation

$$J_0(\lambda_1(\mathbb{D})) = J_1(\lambda_1(\mathbb{D})).$$

In the numerical simulations presented below, 158 inner collocation points for the subspace angle technique were considered and the number of source points N is always half of the number of boundary points. Figure 5.1 shows the configuration used. The method described at the end of the subchapter 4.1 was used to place the source points, where $\eta = 0.5$ was used.

While in the other numerical simulations to be presented more eigenvalues are studied, in Table 5.1 only the first three eigenvalues are shown for the sake of brevity.

Eigenvalues	**N=1200**	**N=1000**	**N=800**
$\tilde{\lambda}_1(\mathbb{D})$	1.4346956515	1.4346956481	1.4346956367
$\tilde{\lambda}_2(\mathbb{D})$	2.6298741163	2.6298741147	2.6298741276
$\tilde{\lambda}_3(\mathbb{D})$	3.1128644920	3.1128645083	3.1128645008
Absolute error: $\left\lvert\lambda_1(\mathbb{D}) - \tilde{\lambda}_1(\mathbb{D})\right\rvert$	$6,877 \times 10^{-10}$	$2,693 \times 10^{-9}$	$1,413 \times 10^{-8}$

Table 5.1: Eigenvalues for different values of N and the measured absolute error.

The plot of the bracketing algorithm 4.1 is shown in Figure 5.2. For future reference, the first eigenfunction on the disk with $m = 0$ is also shown in Figure 5.3, where the real and imaginary parts of the spinors u_1 and u_2 are presented. The plots are also normalized, i.e, $\|\mathbf{u}\|_{L^2(\mathbb{D})} = 1$.

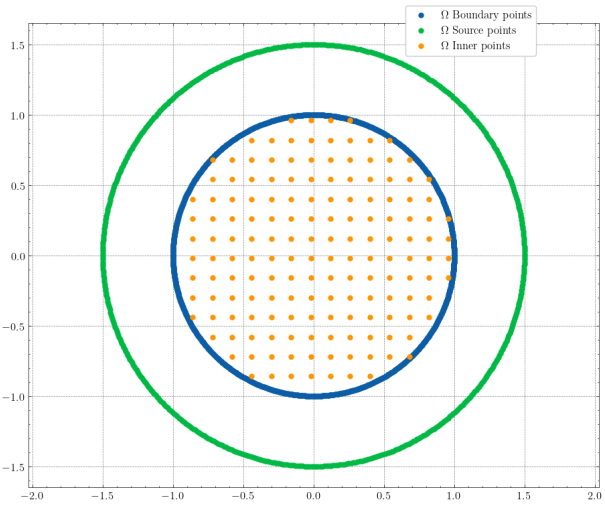


Figure 5.1: Configuration of the boundary, source, and inner points. The number of boundary collocation points used is 1200.

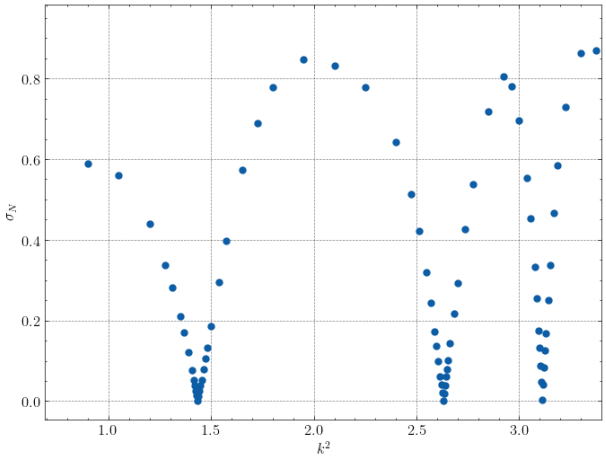


Figure 5.2: Direct search algorithm for the first three eigenvalues of the disk with $m = 0$. Empirically, better approximations are obtained when smaller values of σ_N are found in each singularity.

5.1.2 Quadrilateral results

In this subchapter, several numerical results regarding quadrilateral polygons are shown and discussed. First, numerical evidence for the Conjecture 3.2.6 is presented both for $m = 1$ and $m = 5$. Then, some simulations show that an analogous to the Ashbaugh-Benguria inequality 3.1.13 for the Laplace operator also holds for the Dirac Operator with infinite-mass boundary conditions. Finally, we will also show that some disagreement between the spectrum of both operators is already evident for the third eigenvalue, whose optimal shape is not the disk, contrary to what is conjectured for the Laplacian (which is still an open problem).

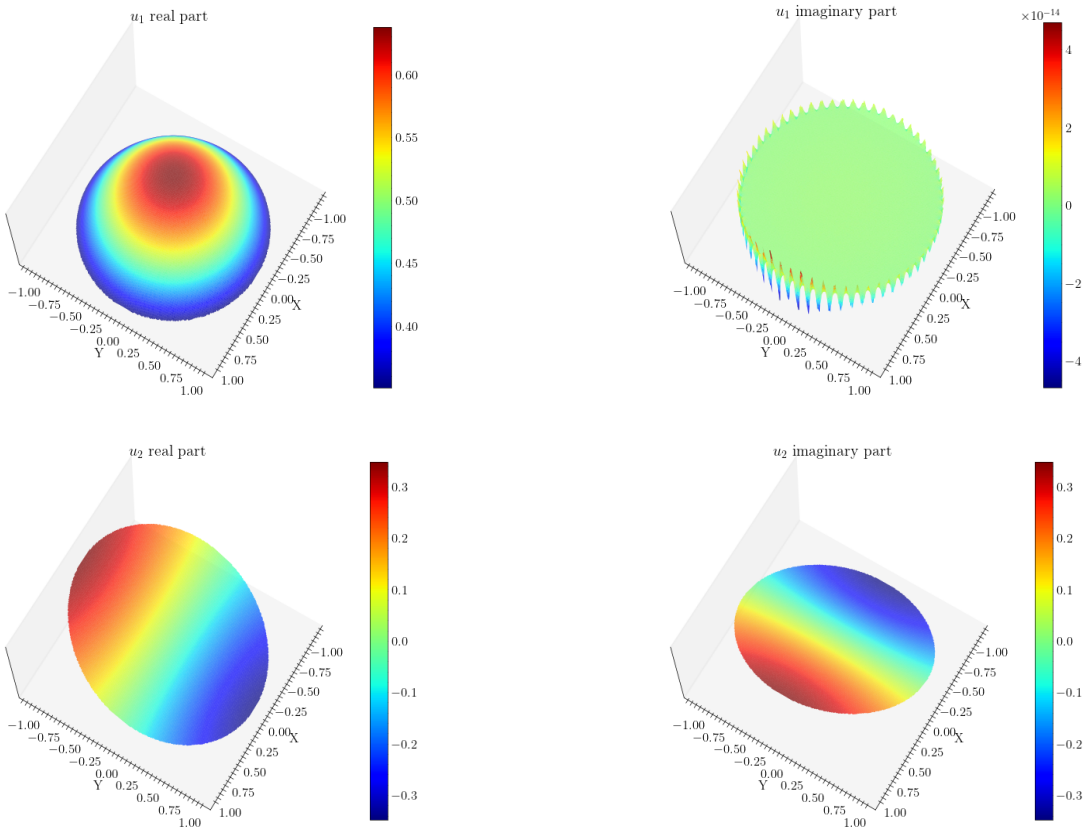


Figure 5.3: Plots of the real and imaginary parts of u_1 and u_2 of the first eigenfunction $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$. Observe that the imaginary part of u_1 is zero and the artifacts presented are due to precision lost.

Consider a rectangle with a width $a > 0$ and unitary area, where the sides are a and $\frac{1}{a}$. Figures 5.4 and 5.5 illustrate the behavior of the first 5 eigenvalues for $m = 1$ and $m = 5$, respectively. Several interesting observations can be made:

- The behavior of the spectrum remains qualitatively consistent for different values of m. Notably, the most prominent difference lies in the decreasing gap between the numerical eigenvalues and m as m increases. While this trend is discernible, it becomes more apparent as m takes on even larger values;

- Starting from the third eigenvalue and beyond, certain spikes become evident in the plot. These spikes correspond to rectangular domains where the eigenvalue has multiplicity two. For instance, between the values 1 and 1.5, the third and fourth eigenvalues appear to converge. This behavior becomes more frequent as the order of the eigenvalues increases. However, only multiplicity 2 is observed which is not consistent with the behavior seen in the Laplace operator, where multiplicity 3 is already achieved in the third eigenvalue and only the first eigenvalue is simple;

- By varying the parameter a, linear growth in the eigenvalues is observed, with no change in their multiplicity. As a increases, the eigenvalues approach one another, as expected. This phenomenon occurs because the rectangle transforms into an unbounded line, resulting in a non-discrete spectrum.

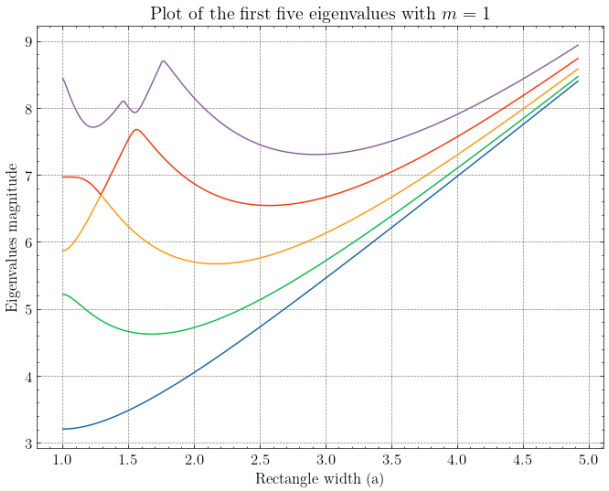


Figure 5.4: Behavior of the first five eigenvalues for rectangles with unit area, width a and $m = 1$.

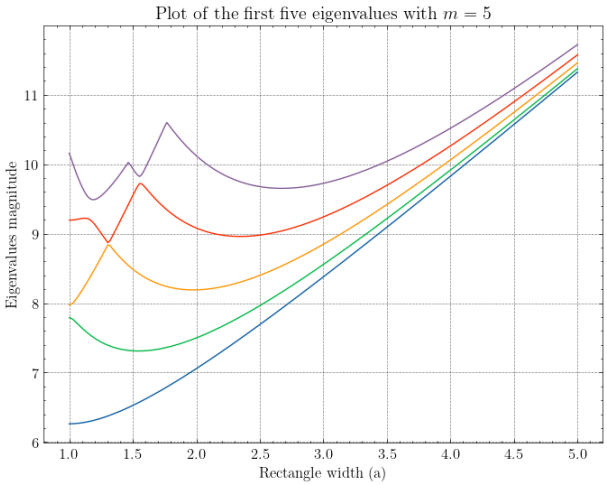


Figure 5.5: Behavior of the first five eigenvalues for rectangles with unit area, width a and $m = 5$.

In Figures 5.6 and 5.7 the results for the perimeter are shown. While Conjecture 3.2.6 is framed for $a \in (0, 2)$ observe that is enough to consider $a \in (0, 1)$ given the symmetry of the problem. Here, one considers the perimeter to be 4, and by varying the rectangle width the area is not constant. The conclusions presented earlier are still valid.

Given the results for rectangles with fixed area or perimeter, it is possible to state that the Conjecture 3.2.6 should hold, although it is still not possible to ascertain that the results presented here hold for every m. In any case, it is known that the eigenvalues are continuous with domain perturbations (here these perturbations can be seen as stretching the rectangles), exactly what has been found in these numerical simulations.

To finish this subsection, we present the results for general quadrilaterals. Since no (practical) parameterization can describe every quadrilateral, we have resorted to a random sampling of quadrilaterals and studied the results against rectangles and rhombuses. Figure 5.8 presents the results of the numerical simulations for $m = 1$. Notice how the rectangles and rhombuses clearly define a region of "acceptable" quadrilaterals and this appears to be particularly true for the first eigenvalue. For the second and third eigenvalues some quadrilaterals behave differently, and it is already possible to find some domains whose third eigenvalue is smaller than the eigenvalue of the unit disk. In Figures 5.9 and 5.10

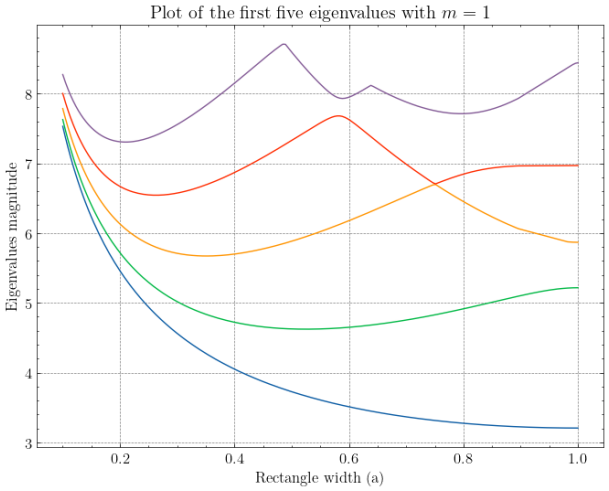


Figure 5.6: Behavior of the first five eigenvalues for rectangles with unit area, width a and $m = 1$.

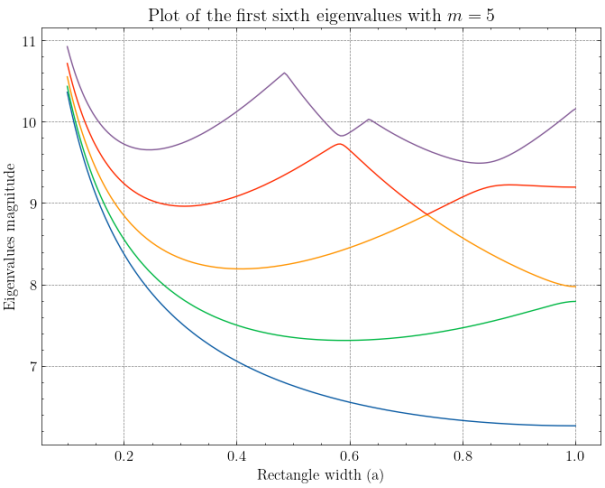


Figure 5.7: Behavior of the first five eigenvalues for rectangles with unit area, width a and $m = 5$.

the ratio between the first eigenvalues is studied, where a version of the Ashbaugh-Benguria Theorem 3.1.13 appears to hold.

5.1.3 Triangle results

In this subchapter, we tackle the Conjectures 3.2.7 stated on [42]. Instead of considering random triangles, its study can be done systematically given that, up to congruence, three parameters completely define every triangle. The approach presented here is based on the work of Antunes & Freitas [4]. Consider the region R defined by

$$R = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y > 0, (x + 1)^2 + y^2 \leq 4\},$$

and its piecewise boundary $\partial R = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2$, where

$$\Gamma_0 = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, y = 0\}$$
$$\Gamma_1 = \{(x, y) \in \mathbb{R}^2 : 0 \leq x < 1, y = \sqrt{4 - (x + 1)^2}\}$$
$$\Gamma_2 = \{(x, y) \in \mathbb{R}^2 : x = 0, 0 < y < \sqrt{3}\}.$$

A triangle T is said to be *admissible* if its basis vertices are $(0, 0)$ and $(1, 0)$, and the other vertex coordinates[1] are (x, y) such that $(x, y) \in \overline{R}$. A triangle T is said to be *subequilateral* if $(x, y) \in \Gamma_1$ and

[1] Of course, one does not need to enforce such constraints to the triangle: as said before, every triangle is unique up to congruence. Here, we just emphasize that is enough to consider triangles defined using the region R. Observe that this is just a model to numerically exhaust all possible triangles up to congruence, one still needs to normalize them to have a unitary area.

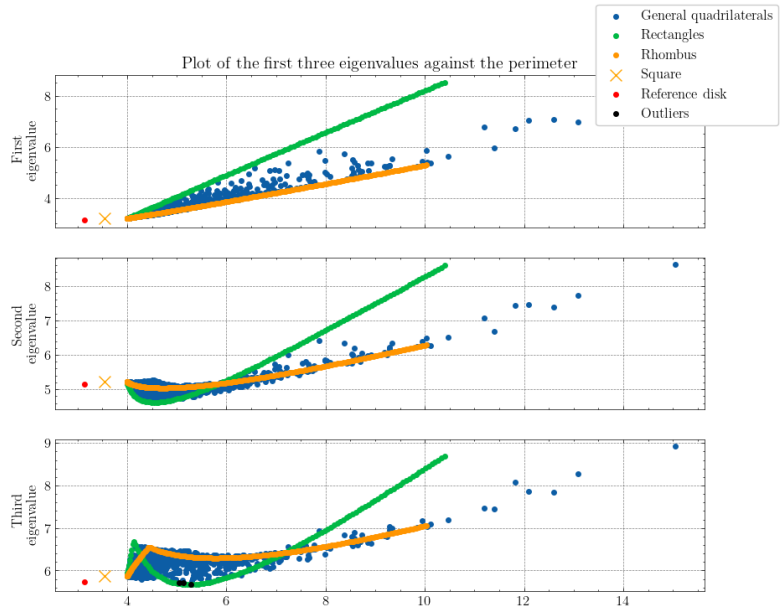


Figure 5.8: Plot of the first three eigenvalues against the perimeter. The "outliers" marked in black represent the domains in which the third eigenvalue is less than the third eigenvalue of the disk.

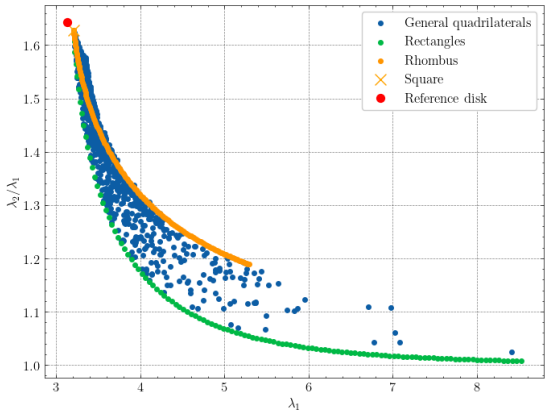


Figure 5.9: Ratio between the first two eigenvalues $\frac{\lambda_2}{\lambda_1}$.

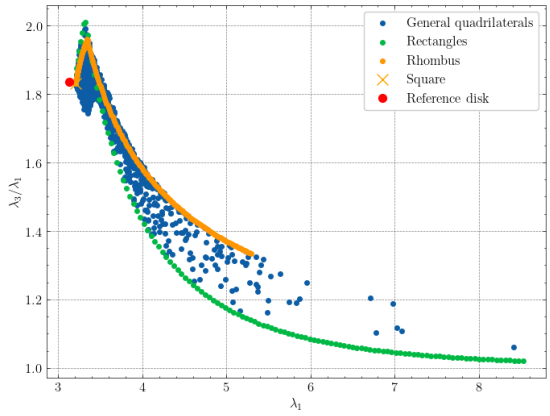


Figure 5.10: Ratio between the third and first eigenvalues $\frac{\lambda_3}{\lambda_1}$.

superequilateral if $(x, y) \in \Gamma_2$; if $(x, y) = (0, \sqrt{3})$ then it is equilateral (obviously). The plot of the region R and its boundary ∂R is in Figure 5.11. In Figure 5.12 we present numerical evidence for Conjecture 3.2.7 with $m = 1$. Notice how the eigenvalues of the superequilateral and subequilateral triangles define a region that includes every other triangle whose (x, y) vertex belongs to R. In Figure 5.13 a version of the Ashbaugh-Benguria (see Theorem 3.1.13) inequality is presented and in Figure 5.14 a 3D plot of the first eigenvalue for every admissible triangle considered against the coordinates (x, y) of the vertex

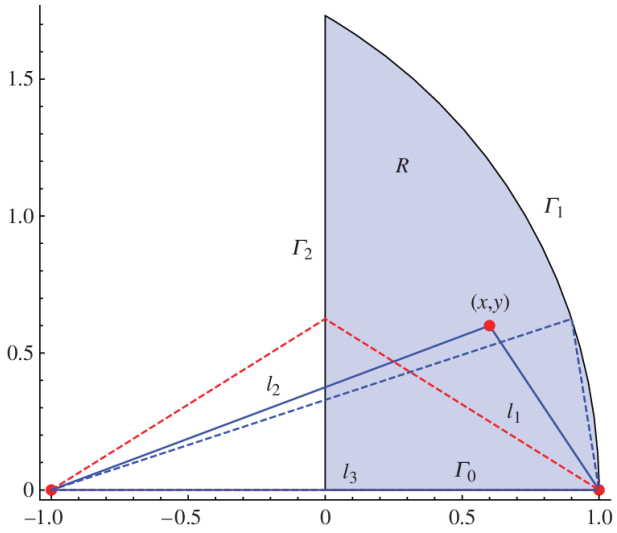


Figure 5.11: Configuration space of the admissible triangles. In a dashed red line is a superequilateral triangle; in a dashed blue line a subequilateral triangle is also represented.

in $\overline{R}$.

While in the previous results for quadrilaterals (and for smooth domains in the next subsection) random domains were considered, it is remarkable that for the triangle problem, one can consider *every* type of triangle just by varying (x, y) in $\overline{R}$. Since the eigenvalues are continuous when considering domain perturbations, it means that it is very unlikely that the conjectures studied for triangles do not hold, otherwise some jump would probably be found, for example in Figure 5.14.

5.1.4 Smooth domain results

In this last subsection, the behavior of the spectrum for smooth (connected) domains with unit area is studied. Here, we fix $m = 1$. The objective of this study is twofold: first, these type of domains is studied since the numerical approximations are more reliable, and it allows for the study of arbitrary domains; second, one can use them to also study the domain which minimizes the third eigenvalue.

In order to generate random smooth domains, periodic B-spline interpolation for each component of an $\mathbb{R}^2$ vector was used. One starts by generating five random points (using a uniform distribution), fitting a periodic B-spline in each component, drawing the two-dimensional B-spline, and checking for auto-intersections. If it does not auto-intersect, then it is a valid domain. Figure 5.15 presents one of these domains.

The figures below are analogous to the ones presented before. In Figure 5.16 a scatter plot of the first three eigenvalues as a function of the perimeter. Once again, analogously to the Laplacian, a Faber-Krahn type result appears to hold for the Dirac operator with infinite-mass boundary conditions. As we

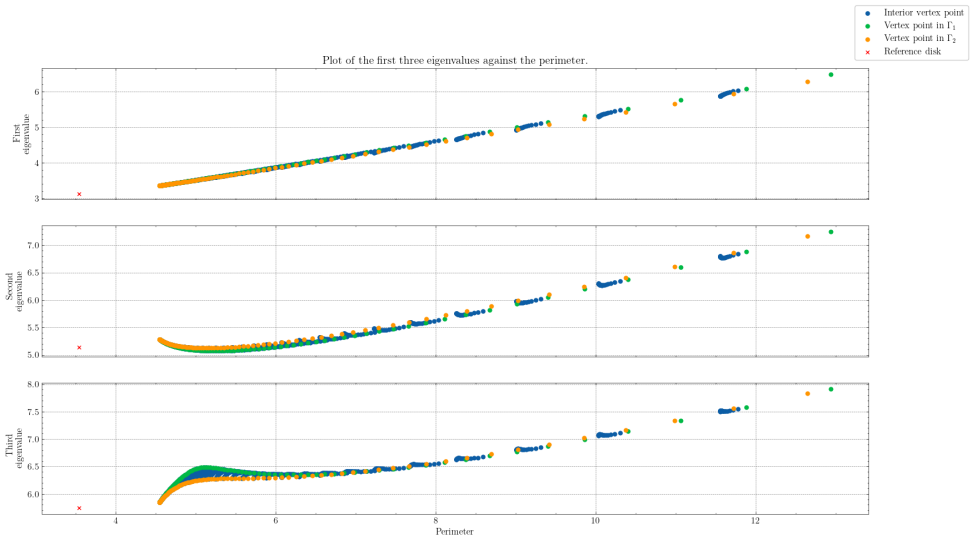


Figure 5.12: Plot of the first three eigenvalues against the perimeter. The "outliers" marked in black represent the domains in which the third eigenvalue is less than the third eigenvalue of the disk.

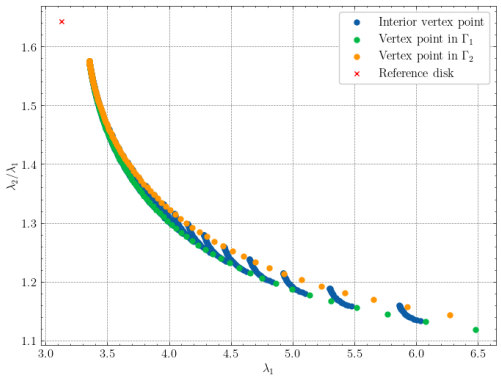


Figure 5.13: Ratio between the first two eigenvalues $\frac{\lambda_2}{\lambda_1}$.

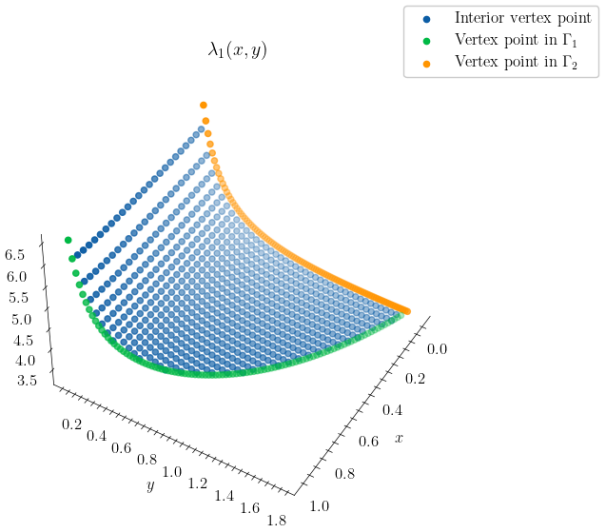


Figure 5.14: Ratio between the third and first eigenvalues $\frac{\lambda_2}{\lambda_1}$.

saw before, some "outlier" domains have a third eigenvalue that is smaller than the third eigenvalue on the disk, which contradicts the conjecture for the Laplace operator. Figures 5.17 and 5.18 are related to the ratio of the first eigenvalues and the spectral gap between them. An analogous to the Ashbaugh-Benguria Theorem for the Laplacian also seems to hold for the Dirac operator, as well its generalization for the ratio $\frac{\lambda_3}{\lambda_1}$.

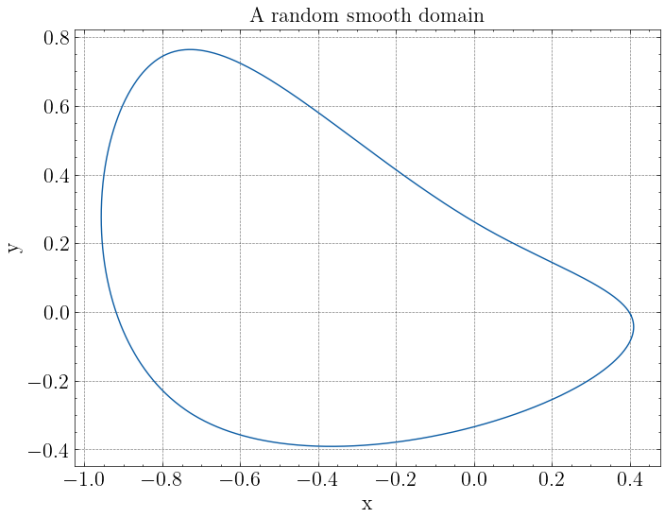


Figure 5.15: Some smooth domain generated by B-splines.

Next, we investigate the domain with the smallest third eigenvalue, and we look for the minimizer of the functional

$$\mathcal{F}(\Omega) = \lambda_3(\Omega).$$

In general, one can address minimization problems in Banach spaces using the notion of *Fréchet*-derivative. In shape optimization problems, for the eigenvalues of an elliptic operator, one can use the variational formula for its eigenvalues. For example, the formula proved in Theorem 2.4.4 can be used to derive some results for the Laplacian Dirichlet problem. In fact, consider $\Omega_t = \varphi(t)(\Omega)$ a small perturbation of Ω in the parameter t, where $\varphi(t)$ is some diffeomorphism for small values of t,

$$\varphi(t) = I + tV$$

for some fix vector field V and $\Omega_0 = \Omega$. Let $\lambda_k(t)$ be an eigenvalue of the Laplace operator with Dirichlet boundary conditions on the domain Ω_t and $u_t^{(k)}$ its associated eigenfunction in $H_0^1(\Omega_t)$. If λ_k has multiplicity one (if it is simple) and Ω is of class C^2 or convex, then

$$\lambda_k'(0) = -\int_{\partial\Omega} \left(\frac{\partial u_0^{(k)}}{\partial n}\right)^2 V \cdot n d\sigma.$$

For more details, we point the reader to [25] and [63]. As of the moment of writing, the author is not aware of any closed form for the derivative of the eigenvalues of the Dirac operator. In that case, two different strategies were considered to solve the unconstrained minimization problem

$$\min_{\substack{\Omega\subset\mathbb{R}^2 \\ |\Omega|=1}} \mathcal{F}(\Omega). \tag{5.2}$$

Starting from a given domain Ω, $\mathcal{F}$ can be minimized using the Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) algorithm, a quasi-Newton method that uses the local curvature of $\mathcal{F}$ to find the de-

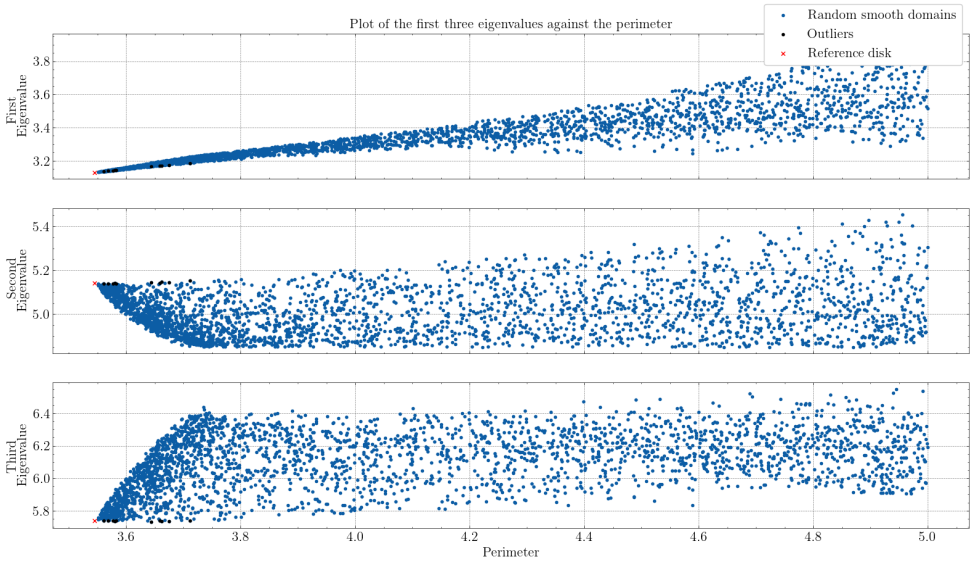


Figure 5.16: Plot of the first three eigenvalues against the perimeter. The "outliers" marked in black represent the domains in which the third eigenvalue is less than the third eigenvalue of the disk.

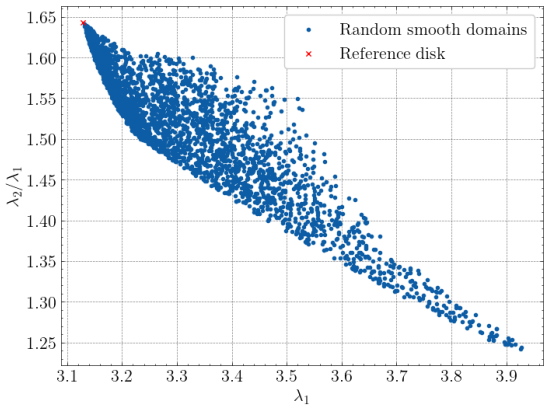


Figure 5.17: Ratio between the first two eigenvalues $\frac{\lambda_2}{\lambda_1}$.

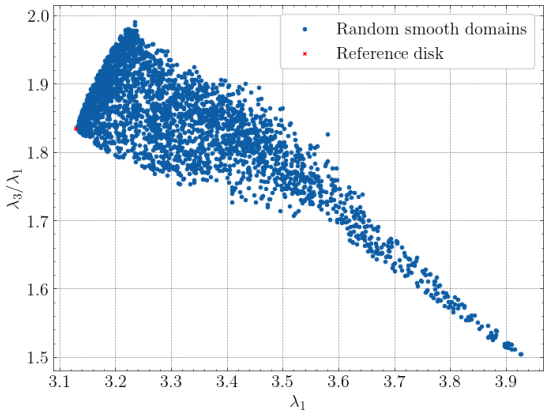


Figure 5.18: Ratio between the third and first eigenvalues $\frac{\lambda_3}{\lambda_1}$.

scent direction given by an approximation of the Hessian matrix. Unfortunately, the BFGS method also needs the derivative at the point, which can only be numerically found using finite-difference methods. The second method is the multidimensional Nelder-Mead direct search method. Just like the bracketing algorithm used to find the singularities on the graph of the smallest singular value, the Nelder-Mead method does not use any information on the derivative and relies on evaluations of the loss function $\mathcal{F}$ to bracket the local minima: in this case, a heuristic strategy with a multidimensional simplex is used to approximate it.

To apply both the BFGS and the Nelder-Mead algorithms, one starts by approximating the domain Ω_0 with the lowest third eigenvalue by a polar parameterization. This is achieved by considering a sample of N boundary points from the domain, considering its polar coordinates, and finding the coefficients of the trigonometric interpolation. More precisely, let M be the order of the trigonometric interpolation. Assume that the radial part of each boundary point of Ω_0 can be parametrized by $r(\theta)$, where $\theta \in (0, 2\pi)$. In that case, one uses the approximation

$$r(\theta_i) \approx a_0 + \sum_{m=1}^{M} a_m \cos(m\theta_i) + \sum_{m=1}^{M} b_m \sin(m\theta_i),$$

where θ_i is the polar part of the sample point i with $i = 1, \ldots, N$. Then, the system

$$\begin{bmatrix} 1 & \cos(1x_1) & \cos(2x_1) & \ldots & \cos(Mx_1) & \sin(1x_1) & \ldots & \sin(Mx_1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \cos(1x_N) & \cos(2x_N) & \ldots & \cos(Mx_N) & \sin(1x_N) & \ldots & \sin(Mx_N) \end{bmatrix}_{N\times(2M+1)} \begin{bmatrix} a_0 \\ \vdots \\ b_M \end{bmatrix} = \begin{bmatrix} r(\theta_1) \\ \vdots \\ r(\theta_N) \end{bmatrix}$$

can be solved by least squares when considering the over-determined system with $N > 2M + 1$. Notice that the problem (5.2) is now discretized into a finite-dimensional problem since every domain is now a vector of coefficients in $\mathbb{R}^{2M+1}$. Of course, one must still consider the domain generated by the found coefficients with unitary area.

Figure 5.19 presents an (almost) optimal domain $\Omega^\star$ shape which minimizes the third eigenvalue of the Dirac operator with infinite-mass boundary conditions. This plot was obtained through the Nelder-Mead algorithm. The third eigenvalue of this domain is approximately $\lambda_3 \approx 5.63787728$ and its perimeter L is $L \approx 5.2650031$. In Figure 5.20 we validate our findings by plotting the (continuous) family of one-parameter transformations (also known as a Minkowski sum) from the unit disk to $\Omega^\star$, given by

$$\Omega_t = (1-t)\mathbb{D} + t\Omega^\star,\ 0 \leq t \leq 1.$$

As said above, the BFGS method was also used. However, no meaningful results were found using this method, since any descent direction produced not

Remark 5.1.1. *One can not fail to point out a valid criticism of this method: since we only parametrized the radial part of the domain's boundary, given the periodicity of the trigonometric interpolation one always end up with star-like shapes. Particularly, in our case, the order of the trigonometric interpolation was low, with $M = 4$. However, by increasing the order of interpolation the dimension of the optimization space would increase, and, in this case, we are already working on $\mathbb{R}^9$, and is very hard for a direct search method like Nelder-Mead to find a local minimum in such a "high" dimensional space. The option to only work with the radial part, instead of both cartesian components, was also a way to reduce the dimensionality of the problem.*

Finally, the normalized eigenfunction associated with the optimal third eigenvalue of domain $\Omega^\star$ is shown in Figure 5.21.

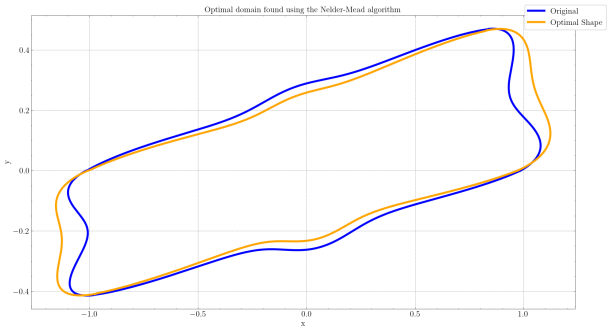


Figure 5.19: Optimal domain $\Omega^\star$ (on orange) against the original domain in the first iteration of the Nelder-Mead algorithm.

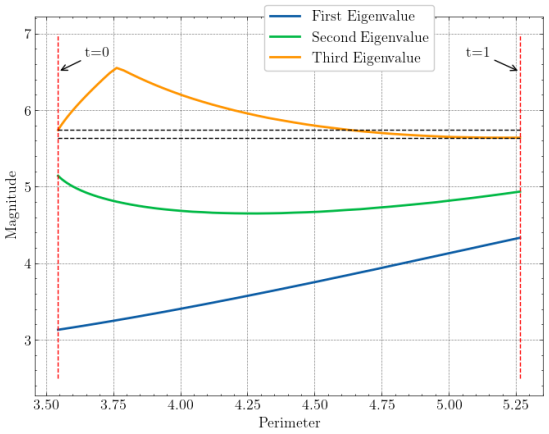


Figure 5.20: Plot of the first three eigenvalues of the Minkowski sum Ω_t for each increasing value of t.

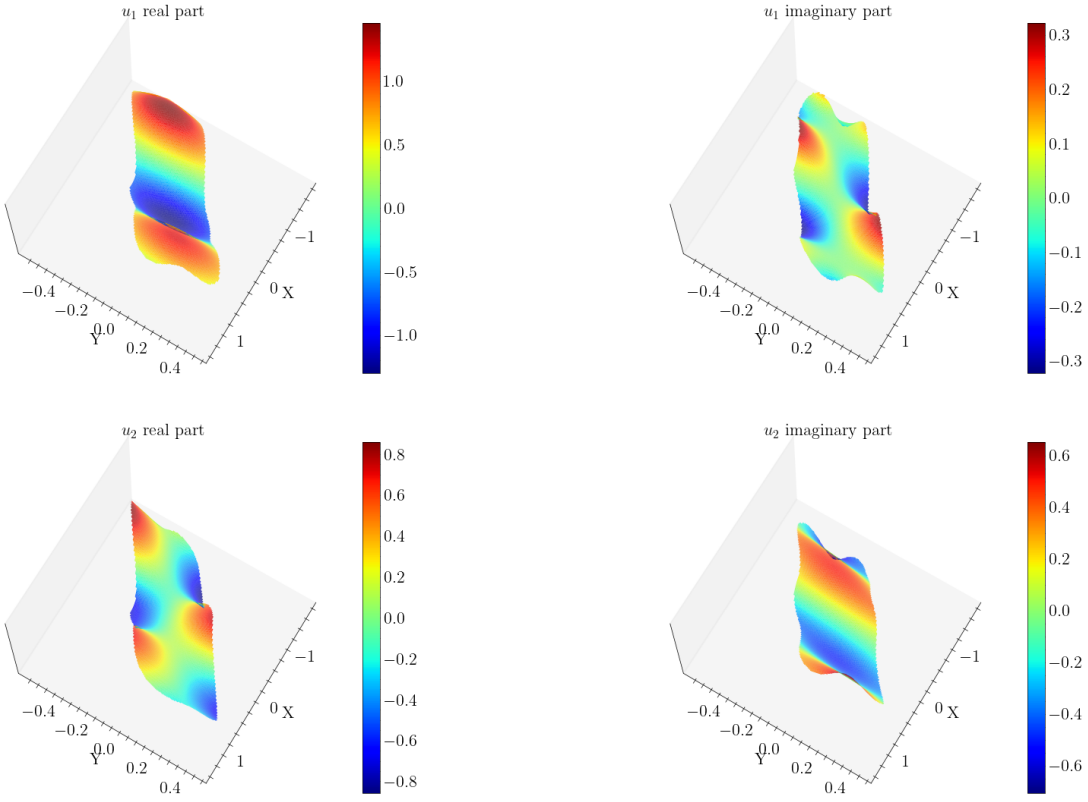


Figure 5.21: Plots of the real and imaginary parts of u_1 and u_2 of the third eigenfunction $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ associated with the optimal domain $\Omega^\star$.

5.2 Transmission problem simulations

For the Transmission problem, we now consider the set of equations studied previously in the subchapter 3.3, given by

$$\begin{aligned} -\nabla k_i \nabla u_i &= f_i, \text{ in } \Omega_i \\ u_1 - u_2 &= 0, \text{ on } \gamma \\ k_1 \frac{\partial u_1}{\partial n_1} + k_2 \frac{\partial u_2}{\partial n_2} &= 0, \text{ on } \gamma \\ u_i &= 0, \text{ on } \Gamma_i. \end{aligned} \tag{5.3}$$

Like before, the domain $\Omega \subset \mathbb{R}^2$ is divided into two non-overlapping regions Ω_1 and Ω_2 such that $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$. Their common boundary is denoted by $\gamma = \partial\Omega_1 \cap \partial\Omega_2$ and the boundary of each domain (minus the common boundary) is also denoted by $\Gamma_i = \partial\Omega_1 \setminus \gamma$. In what follows, the source functions f_i of each domain are constant[2], and we took $f_i = 1$. Recall that $k_1 \geq k_2 > 0$ are constants and n_i is the (normalized) outward normal to each domain subdomain $\Omega_i, i = 1, 2$, where we shall write $n = n_1 = -n_2$ when we are restricted to the interface.

The procedure presented here is based on [52] and [64]. Given that $f_i = 1$ for each $i = 1, 2$, a solution of (5.3) can be found by taking the following steps:

1. Find a solution for the non-homogeneous problem

$$\begin{cases} -\Delta u_1^{NH} = \frac{1}{k_1} \\ -\Delta u_2^{NH} = \frac{1}{k_2}. \end{cases}$$

   This can easily be done, and we have

$$\begin{cases} u_1^{NH} = -\frac{x_1^2 + x_2^2}{4k_1} \\ u_2^{NH} = -\frac{x_1^2 + x_2^2}{4k_2}; \end{cases}$$

2. Then we solve the homogeneous problem

$$\begin{aligned} -\Delta u_i^H &= 0, \text{ in } \Omega_i \\ u_1^H - u_2^H &= u_2^{NH} - u_1^{NH}, \text{ on } \gamma \\ k_1 \frac{\partial u_1}{\partial n_1} - k_2 \frac{\partial u_2}{\partial n_1} &= k_2 \frac{\partial u_2^{NH}}{\partial n_1} - k_1 \frac{\partial u_1^{NH}}{\partial n_1}, \text{ on } \gamma \\ u_i &= -u_i^{NH}, \text{ on } \Gamma_i; \end{aligned} \tag{5.4}$$

3. Finally, the solution of (5.3) is $u_i = u_i^H + u_i^{NH}$.

[2] In this work we only considered $f_i = 1$. In any case, we are still working with a general discontinuous source function (if $k_1 \neq k_2$). Working with different *continuous* source functions should make no difference in the result. We will also present how to deal with general and continuous source functions.

For general source functions, the steps above can also be used: however, it may not be possible to find an analytical solution in the first step. In this case, **1.** must be solved numerically. A popular choice is to use radial basis functions (Radial Basis Functions (RBF)s) (see [65] for example), like the thin plate spline

$$\varphi(r) = r^2 \log r,$$

and find the coefficients α_j in $f_i(x_k) = \tilde{f}_i(x_k) = \sum_{j=1}^{n} \alpha_j \varphi_j(x_k)$ using least square methods, where x_k are collocation points. Finally, one can analytically solve the equation $-\Delta\Psi_j = \varphi_j$ to recover the non-homogeneous solutions u_1^{NH} and u_2^{NH}. In [52] and [64] a different approach was suggested using the fundamental solutions of the Helmholtz equation instead of the classical RBFs; that method is known today as *Kansa MFS method*.

In what follows, the numerical results illustrate the accuracy of the method in simply connected 2D domains. Let N_i denote the number of source points for each domain i, such that $N = N_1 + N_2$. We denote the approximate solution by

$$\tilde{u} = \begin{cases} \tilde{u}_1, & \text{in } \Omega_1 \\ \tilde{u}_2, & \text{in } \Omega_2, \end{cases}$$

where

$$\tilde{u}_1(x) = \sum_{j=1}^{N_1} \alpha_j^{(1)} \Phi\left(x - y_j^{(1)}\right)$$
$$\tilde{u}_2(x) = \sum_{j=1}^{N_2} \alpha_j^{(2)} \Phi\left(x - y_j^{(1)}\right).$$

Let M_i be the number of boundary collocation points $x_m^{(i)}$ for each Ω_i and $M = M_1 + M_2$. We also consider Q interface points $z_q \in \gamma$ with $q = 1, \ldots, Q$. Then, a full block system is written as

$$\begin{bmatrix} \left[\Phi\left(x_m^{(1)} - y_j^{(1)}\right)\right] & [0] \\ [0] & \left[\Phi\left(x_m^{(2)} - y_j^{(2)}\right)\right] \\ \left[\Phi\left(z_q - y_j^{(1)}\right)\right] & \left[-\Phi\left(z_q - y_j^{(2)}\right)\right] \\ \left[k_1\partial_n\Phi\left(z_q - y_j^{(1)}\right)\right] & \left[-k_2\partial_n\Phi\left(z_q - y_j^{(2)}\right)\right] \end{bmatrix} \begin{bmatrix} \left[\alpha_j^{(1)}\right] \\ \left[\alpha_j^{(2)}\right] \end{bmatrix} = \begin{bmatrix} \left[-u_1^{NH}(x_m^{(1)})\right] \\ \left[-u_2^{NH}(x_m^{(2)})\right] \\ \left[u_2^{NH}(z_q) - u_1^{NH}(z_q)\right] \\ \left[k_2\partial_n u_2^{NH}(z_q) - k_1\partial_n u_1^{NH}(z_q)\right] \end{bmatrix} \tag{5.5}$$

Most of the examples below do not have an analytical solution: only in the subsection 5.2.1, when we test the results against a known solution, we can find the absolute error. In the other cases, we are only interested in the relative error, i.e, the boundary error (against which we can compare since $u_i = 0$ in Γ_i) and the interface error by checking the transmission conditions on γ:

- $\|\tilde{u}_i - 0\|_{L^2(\Gamma_i)}$, $i = 1, 2$: boundary collocation error;
- $\|\tilde{u}_1 - \tilde{u}_2\|_{L^2(\gamma)}$, $i = 1, 2$: continuity error (C^0) of $\tilde{u}$ across γ;
- $\|k_1\partial_n\tilde{u}_1 - k_2\partial_n\tilde{u}_2\|_{L^2(\gamma)}$, $i = 1, 2$: continuity error C^1 of $\partial_n\tilde{u}$ across γ.

From a numerical point of view, let $\mathcal{I}$ be the sample of test points. The L^2 norm is discretized into the Root Mean Squared Error (Root Mean Squared Error (RMSE)) which is equivalent to the l^2 norm and is given by

$$\|u - \tilde{u}\| = \sqrt{\frac{1}{\#\mathcal{I}} \sum_{z \in \mathcal{I}} |u(z) - \tilde{u}(z)|^2}.$$

For every result below the number of sample test points is 5 times larger than the sample used to find the coefficients of the MFS, and we fix $k_2 = 1$ since the ratio $\frac{k_1}{k_2}$ is responsible for the behavior of the solutions near the interface.

5.2.1 Numerical validation of the method

First, we start by testing the numerical algorithm for the unit disk $\mathbb{D}$, with $k_1 = k_2 = 1$. From Theorem 3.3.3, we know that the system of differential equations (5.3) is equivalent to the Poisson equation

$$\begin{aligned} -\Delta u &= 1, \text{ in } \mathbb{D} \\ u &= 0, \text{ on } \partial\mathbb{D}, \end{aligned} \tag{5.6}$$

and is easy to see that the exact solution of Equation (5.6) is given in polar coordinates by $u(r, \theta) = \frac{1-r^2}{4}$.

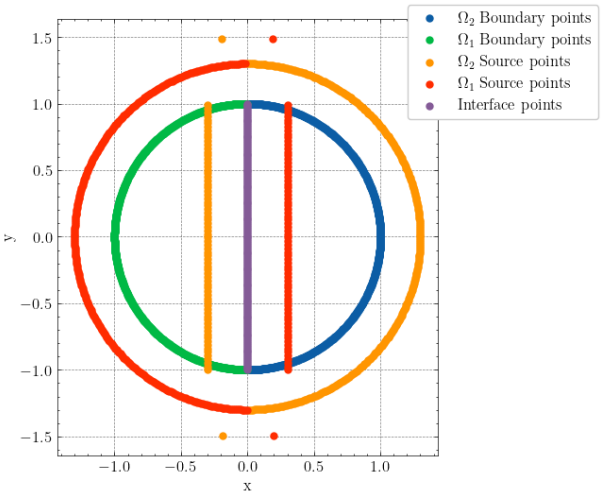


Figure 5.22: Configuration of the boundary, source, and interface points. Each domain has 600 boundary points, 377 source points and the common interface has 100 points.

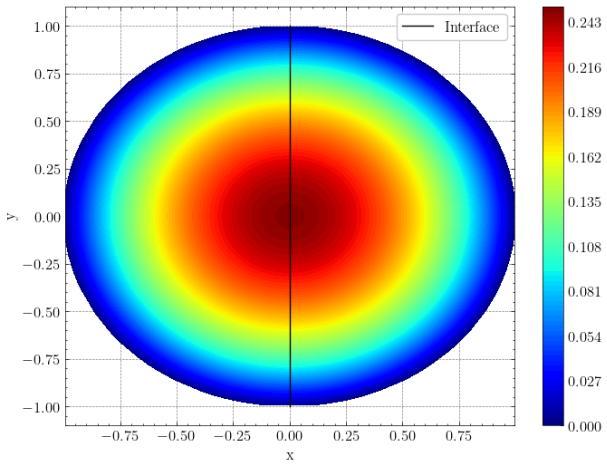


Figure 5.23: Numerical approximation of the BVP (5.6) under the conditions presented in Figure 5.22

The absolute error between the approximate solution and the exact solution can then be calculated for each domain point. The sample points to compute the absolute error were generated in a uniform

Boundary/Interface Points	Boundary Error		Absolute Error	
	Domain 1	Domain 2	Domain 1	Domain 2
600/150	9.759×10^{-12}	9.541×10^{-12}	1.465×10^{-11}	1.439×10^{-11}
500/100	3.667×10^{-11}	3.945×10^{-11}	9.382×10^{-11}	9.310×10^{-11}
412/100	3.721×10^{-11}	5.036×10^{-11}	9.652×10^{-11}	9.584×10^{-11}

Table 5.2: Numerical errors for the boundary and the whole Domains Ω_1 and Ω_2

Boundary/Interface Points	Interface C^0 Error	Interface C^1 Error	Condition number
600/150	6.945×10^{-11}	1.841×10^{-11}	2.528×10^{19}
500/100	3.910×10^{-10}	1.100×10^{-10}	2.382×10^{18}
412/100	4.035×10^{-10}	1.342×10^{-10}	7.597×10^{17}

Table 5.3: Numerical error on the interface γ. The condition number of the matrix is also presented.

grid and were also used to plot Figure 5.23. The method described at the end of the subchapter 4.1 was used to place the source points, where $\eta = 0.3$.

The numerical results presented in Tables 5.2 and 5.3 are not yet optimal. When considering a larger value of η, the results increase by more than two orders of magnitude, but this also leads to a significant increase in the already very high condition number (see Table 5.3), as expected. Despite this, the results show great promise, which was anticipated due to the domain's analyticity.

As mentioned previously, the Method of Fundamental Solutions (MFS) yields better results in highly regular domains, even when considering curved geometries. However, increasing the number of boundary and interface collocation points improves the accuracy of the solution. It is important to note that a larger number of points also escalates the condition number, making the problem more challenging to solve accurately.

The inclusion of the "corner" source points, as depicted in Figure 5.22, also significantly impacts the method's accuracy. These source points, strategically added to capture the behavior near the interface corner, can be inside one of the domains as the solution will be split into two parts.

5.2.2 Results for the rectangle

In this subsection a rectangular domain $[-1, -0.5] \times [1, 0.5]$ with a vertical interface along the line $x = 0$ is considered. We are now interested to study the problem for $k_1 \neq k_2$ where $k_2 = 1$ is fixed. The results below were conducted with 600 boundary points and 404 source points for each domain. The number of interface points is 150 and $\eta = 0.08$.

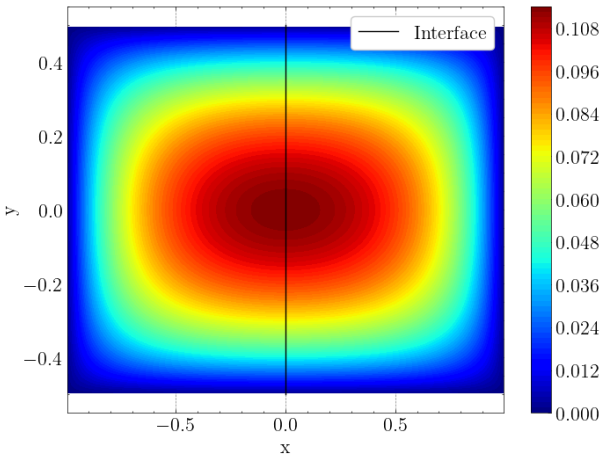


Figure 5.24: Numerical simulation with $k_1 = 1$.

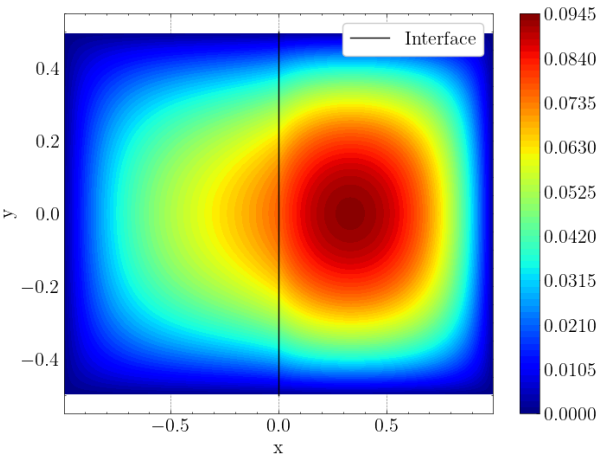


Figure 5.25: Numerical simulation with $k_1 = 2$.

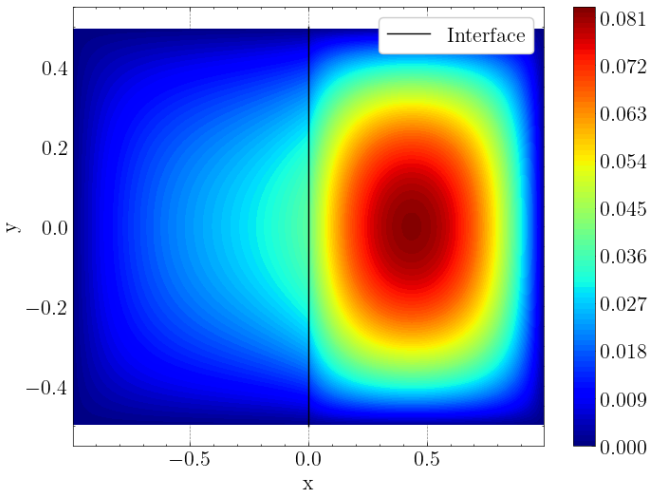


Figure 5.26: Numerical simulation with $k_1 = 5$.

Numerical approximations of the BVP for a rectangular domain with different k_1 values.

In Figures 5.24, 5.25, and 5.26, we present the numerical approximation for different k_1 values. Observe that increasing k_1 breaks the symmetry of the solution, which shifts from Ω_1 (the domain on the left) to Ω_2 (the domain on the right). Table 5.4 summarizes the results for different k_1 values. While the results are slightly worse than the previous ones due to the worse domain regularity, they still preserve high accuracy. It is worth noting that for different values of k_1 and k_2, the accuracy of the method decreases. This is also to be expected, as we are dealing with a discontinuous source function, which decreases the regularity of the solution.

Notice that the condition number for the rectangle is smaller than the one presented in Table 5.3 for the unit disk. This is a consequence of a smaller value of η, which in this case appears to be optimal

k_1 value	Boundary Error		Interface Errors		Condition number
	Domain 1	Domain 2	C^0	C^1	
1	7.775×10^{-8}	7.779×10^{-8}	4.732×10^{-9}	7.589×10^{-9}	2.331×10^{13}
2	4.398×10^{-8}	8.614×10^{-8}	2.499×10^{-6}	7.868×10^{-8}	3.623×10^{13}
3	2.181×10^{-8}	1.036×10^{-7}	3.838×10^{-6}	1.551×10^{-7}	8.182×10^{13}

Table 5.4: Numerical relative error on the boundary and in the interface γ

since increasing its value decreases the overall accuracy.

5.2.3 Results for an L-shape domain with enrichment

In the previous subsection, a domain with corners was analyzed. However, it still preserved some regularity, and the MFS with classical basis functions was able to capture its corner's behavior, as explained in Remark 4.2.3. In what follows, we are going to study the case of an L-shape, a non-convex domain with singular corners. Two different interfaces will be considered: first, the usual interface along the line $x = 0$; then, along its symmetry axis with the line $y = \frac{1}{2}x$.

Consider the L-shape given in Figure 5.27. The left and right subdomains are denoted by Ω_1 and Ω_2, respectively. The number of interface points is 300, and the number of source points for each domain is 428. The number of boundary collocation points for Ω_1 and Ω_2 is 710 and 639, respectively.

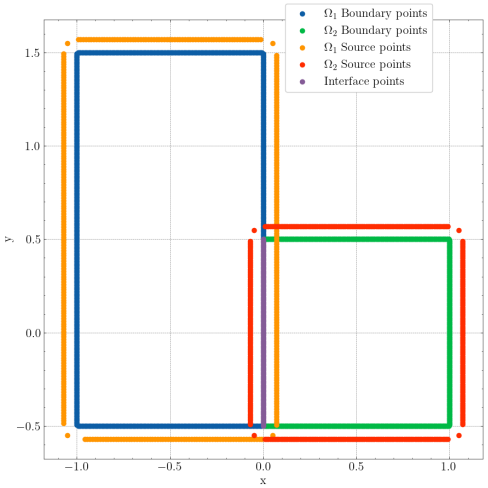


Figure 5.27: L-shape domain with a vertical interface. Configuration of the boundary, source, and interface points.

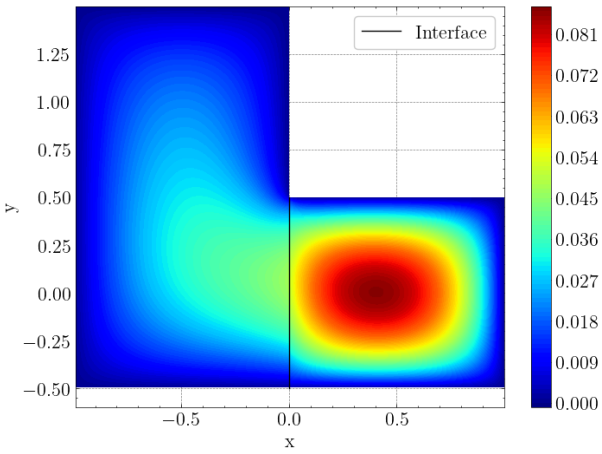


Figure 5.28: Numerical approximation of the BVP for an L-shape domain with interface along $x = 0$ and $k_1 = 5$.

In Table 5.5, the results without resorting to enrichment are presented. It is evident that the method

yields poorer results due to the lower regularity of the domain. Particularly, the interface error (mainly the C^0 error) is significantly higher compared to previous cases, even when considering more collocation points on the interface. In fact, it appears that the domain itself poses more challenges than the discontinuous source function when considering different values for k_1 and k_2. In Figure 5.28, it is even possible to see that there already exists some small jump near the edges of the interface.

k_1 **value**	**Boundary Error**		**Interface Errors**		**Condition number**
	Domain 1	**Domain 2**	C^0	C^1	
1	7.853×10^{-5}	1.155×10^{-4}	2.916×10^{-3}	2.155×10^{-5}	5.587×10^{12}
2	8.152×10^{-5}	1.350×10^{-4}	3.835×10^{-3}	1.149×10^{-5}	8.161×10^{12}
5	7.079×10^{-5}	1.378×10^{-4}	4.085×10^{-3}	5.411×10^{-5}	1.776×10^{13}

Table 5.5: Numerical relative error on the boundary and in the interface γ

One of the major problems for the method is the behavior of the solution near the degenerate corner with π radians in Ω_1, where some singularity may exist due to the different boundary conditions imposed there. Notice that for Ω_2 there exists no problem since the interface edges make a right angle with the adjacent edges. Therefore, we consider some particular solutions which describe the solution in the domain Ω_1. In this case, we are going to use Dirichlet-Neummann particular solutions, like the ones presented in 4.2.1, centered in the singular corner. Let

$$v_{p_1}(r,\theta) = \alpha_{p_1} r^{\alpha_{p_1}} \sin(\alpha_{p_1}(\theta - \theta_1)) \tag{5.7}$$

where

$$\alpha_{p_1} = \frac{(p+\frac{1}{2})\pi}{\Theta},$$

$\theta_1 = \frac{\pi}{2}$ is the angle shift, $\Theta = \pi$ is the total angle amplitude, and the coordinates r and θ are given in polar coordinates by

$$r(x,y) = \sqrt{x^2+y^2} \quad \theta(x,y) = \begin{cases} \arctan\left(\frac{y}{x}\right), & \text{if } \arctan\left(\frac{y}{x}\right) > 0 \\ \arctan\left(\frac{y}{x}\right) + 2\pi, & \text{if } \arctan\left(\frac{y}{x}\right) \leq 0. \end{cases}$$

By differentiating Equation (5.7) in cartesian coordinates and substituting polar coordinates again we find that

$$\nabla v_{p_1}(r,\theta) = \left(-\frac{(2\pi p_1+\pi)^2 r^{\frac{2\pi p_1+\pi}{2\Theta}} \sin\left(\theta + \frac{\pi\left(p_1+\frac{1}{2}\right)(s-\theta)}{\Theta}\right)}{4\Theta^2 r}, \frac{(2\pi p_1+\pi)^2 r^{\frac{2\pi p_1+\pi}{2\Theta}} \cos\left(\theta + \frac{\pi\left(p_1+\frac{1}{2}\right)(s-\theta)}{\Theta}\right)}{4\Theta^2 r} \right).$$

Finally, considering the truncated expansion

$$\phi(r,\theta) = \sum_{p_1=0}^{P_1} \beta_{p_1} v_{p_1}(r,\theta), \tag{5.8}$$

and expanding the matrix in (5.5), one can write

$$\begin{bmatrix} \left[\Phi\left(x_m^{(1)} - y_j^{(1)}\right)\right] & [0] & [\phi(r\left(x_m\right), \theta\left(x_m\right))] \\ [0] & \left[\Phi\left(x_m^{(2)} - y_j^{(2)}\right)\right] & [0] \\ \left[\Phi\left(z_q - y_j^{(1)}\right)\right] & \left[-\Phi\left(z_q - y_j^{(2)}\right)\right] & [\phi(r\left(z_k\right), \theta\left(z_q\right))] \\ \left[k_1 \partial_n \Phi\left(z_q - y_j^{(1)}\right)\right] & \left[-k_2 \partial_n \Phi\left(z_q - y_j^{(2)}\right)\right] & [k_1 \partial_n \phi(r\left(z_q\right), \theta\left(z_q\right))] \end{bmatrix} \tag{5.9}$$

Table 5.7 presents the results after applying the enrichment technique for the previous k_1 values. In the expansion (5.8), different values for P_1 were considered. For example, for the first section of the Table, we set $P_1 = 1$. After some simulations, it became clear that increasing P_1 would not give better results. Furthermore, since the form of Equation (5.8) is also valid for the exterior problem, negative values of p_1 were considered. Interestingly, when adding solutions for the exterior problem, better results were achieved. Not only did the error decrease for the boundary Γ_i of each domain, but better results were also achieved on the interface.

p_1 **values**	k_1 **value**	**Boundary Error**		**Interface Errors**		**Condition number**
		Domain 1	**Domain 2**	C^0	C^1	
0, 1	1	2.965×10^{-5}	7.907×10^{-5}	2.94×10^{-3}	2.624×10^{-5}	5.584×10^{12}
	2	2.203×10^{-5}	6.657×10^{-5}	1.86×10^{-3}	2.068×10^{-5}	8.156×10^{12}
	5	2.203×10^{-5}	6.657×10^{-5}	1.86×10^{-3}	2.068×10^{-5}	8.156×10^{12}
-1, 0, 1	1	8.627×10^{-6}	4.132×10^{-5}	7.68×10^{-4}	6.876×10^{-6}	5.585×10^{12}
	2	7.333×10^{-6}	2.791×10^{-5}	6.01×10^{-4}	2.555×10^{-5}	8.157×10^{12}
	5	4.271×10^{-6}	1.118×10^{-5}	2.69×10^{-4}	4.166×10^{-5}	1.775×10^{13}
-2, -1, 0, 1	1	3.898×10^{-6}	5.975×10^{-6}	1.44×10^{-3}	2.505×10^{-6}	4.156×10^{14}
	2	2.584×10^{-6}	2.514×10^{-6}	9.69×10^{-4}	1.048×10^{-5}	4.156×10^{14}
	5	1.119×10^{-6}	6.838×10^{-7}	4.89×10^{-4}	1.106×10^{-5}	4.156×10^{14}

Table 5.7: Numerical relative error on the boundary and in the interface γ after considering particular (angular) solutions

To finish this subsection, and as stated in the beginning, we present a more complicated L-shape domain where the interface is drawn along its axis of symmetry. In this case, notice that there are two singular corners, both with $\frac{3}{4}\pi$ radians[3]. In Figure 5.29 the configuration is presented, where the left domain is Ω_1 and the right domain Ω_2. The number of interface points and source points for each domain

[3]The other two corners have an angle of $\frac{\pi}{4}$ radians which cause no problem for the classical MFS basis functions.

is still 300 and 428, respectively. We also chose 628 boundary collocation points for both domains. Table 5.8 summarizes the results without considering particular solutions.

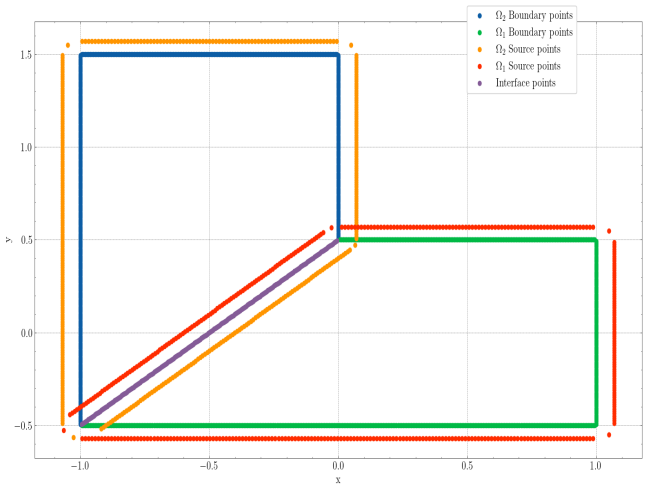


Figure 5.29: L-shape domain with the interface on the symmetry axis. Configuration of the boundary, source, and interface points.

k_1 value	Boundary Error		Interface Errors		Condition number
	Domain 1	Domain 2	C^0	C^1	
1	1.812×10^{-4}	2.060×10^{-4}	7.305×10^{-3}	8.018×10^{-5}	1.050×10^{10}
2	1.398×10^{-4}	9.729×10^{-5}	5.986×10^{-4}	5.505×10^{-5}	1.646×10^{10}
5	7.096×10^{-5}	3.030×10^{-5}	1.528×10^{-3}	4.348×10^{-5}	3.730×10^{10}

Table 5.8: Numerical relative error on the boundary and in the interface γ

Since particular solutions will be added to the singular corners in Ω_2, one should now consider Neummann-Dirichlet particular solutions centered in the singular corner. These particular solutions have the form

$$w_{p_2}(r,\theta) = \alpha_{p_2} r^{\alpha_{p_2}} \cos(\alpha_{p_2}(\theta - \theta_2)),$$

where the α_{p_2} coefficients are the same as before, $\Theta = \frac{3}{4}\pi$ and $\theta_2 = \frac{5}{4}\pi$ is the angle shift. Notice that Θ is the same in both domains. The (polar) gradient of w is now

$$\nabla w(r,\theta)_{p_2} = \left(\frac{(2\pi p_2 + \pi)^2 r^{\frac{2\pi p_2 + \pi}{2\Theta}} \cos\left(\theta + \frac{\pi\left(p_2+\frac{1}{2}\right)(s-\theta)}{\Theta}\right)}{4\Theta^2 r}, \frac{(2\pi p_2 + \pi)^2 r^{\frac{2\pi p_2 + \pi}{2\Theta}} \sin\left(\theta + \frac{\pi\left(p_2+\frac{1}{2}\right)(s-\theta)}{\Theta}\right)}{4\Theta^2 r} \right).$$

Considering the expansion (5.8) for the w_{p_2} particular solutions, one can add more blocks to the matrix (5.9). Table 5.10 shows the results when particular solutions are added to both domains. Once

again, negative values of p_2 were considered. Without intending to show too much data (because we would have to account for every combination of p_1 and p_2 values), we fix $p_2 = 0, 1$. The reason for this choice is that we noticed better results are achieved when increasing the number of particular solutions in the domain where the diffusion coefficient is increasing. In this case, we are only varying the coefficient k_1, and therefore, we can fix the p_2 values.

p_1 values	k_1 value	Boundary Error		Interface Errors		Condition number
		Domain 1	Domain 2	C^0	C^1	
0, 1	1	3.107×10^{-7}	2.100×10^{-7}	1.158×10^{-5}	1.019×10^{-7}	1.054×10^{10}
	2	1.061×10^{-7}	6.892×10^{-7}	6.366×10^{-5}	7.065×10^{-8}	1.648×10^{10}
	5	1.207×10^{-7}	1.141×10^{-6}	9.544×10^{-5}	9.487×10^{-8}	3.732×10^{10}
-1, 0, 1	1	3.319×10^{-7}	2.026×10^{-7}	1.813×10^{-5}	1.048×10^{-7}	1.054×10^{10}
	2	1.676×10^{-7}	3.166×10^{-7}	8.375×10^{-6}	7.951×10^{-8}	1.647×10^{10}
	5	9.871×10^{-8}	4.807×10^{-7}	1.549×10^{-6}	5.285×10^{-8}	3.732×10^{10}
-2, -1, 0, 1	1	3.070×10^{-7}	2.415×10^{-7}	1.610×10^{-5}	9.893×10^{-8}	5.659×10^{12}
	2	1.762×10^{-7}	2.670×10^{-7}	8.385×10^{-6}	8.480×10^{-8}	5.656×10^{12}
	5	9.761×10^{-8}	3.201×10^{-7}	1.993×10^{-6}	6.049×10^{-8}	5.655×10^{12}

Table 5.10: Numerical relative error on the boundary and in the interface γ after considering particular (angular) solutions

Again, we found the same surprising results as before, where considering negative values for p_1 achieves better approximations, particularly when $k_1 \neq k_2 = 1$. An interesting observation was made when considering negative values for p_2: in that case, if p_1 is also negative, the solution in both domains would explode. Intuitively, one can treat the problem numerically as an exterior problem in only one domain. On the other hand, one may consider $p_1 = p_2 = 0, \ldots, P$, where $P = P_1 = P_2$, for $P > 1$ (P may be as large as 8, for example, with 16 particular solutions being added in total) and the results can be as good as the ones presented in Table 5.10 for $p_1 = -2, -1, 0, 1$, but only if $k_1 = k_2 = 1$; otherwise, the approximation on the interface gets worse than what we found.

Figures 5.30 and 5.31 show the absolute value of the errors for each interface point, with $k_1 = 2$ and the p values for the particular solutions are $p_1 = -2, -1, 0, 1$ and $p_2 = 0, 1$. As expected, both errors peak near the edges of the interface with special evidence when near the singular corner.

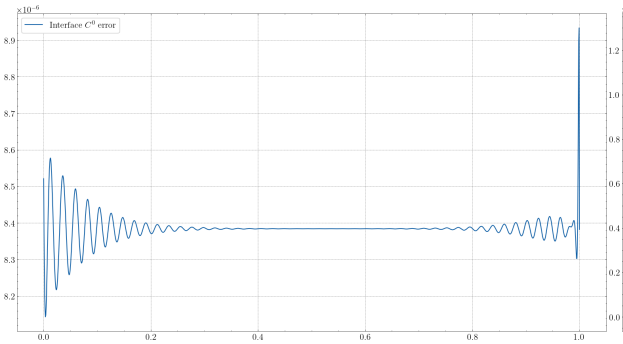


Figure 5.30: Interface C^0 error

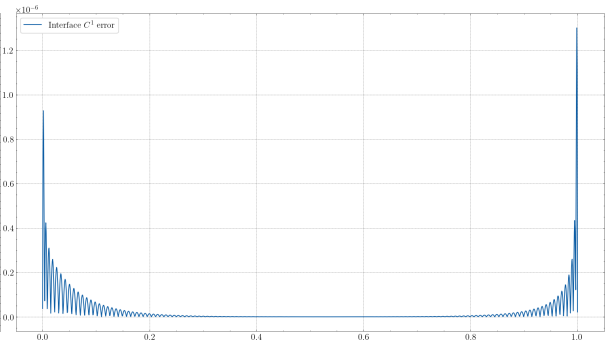
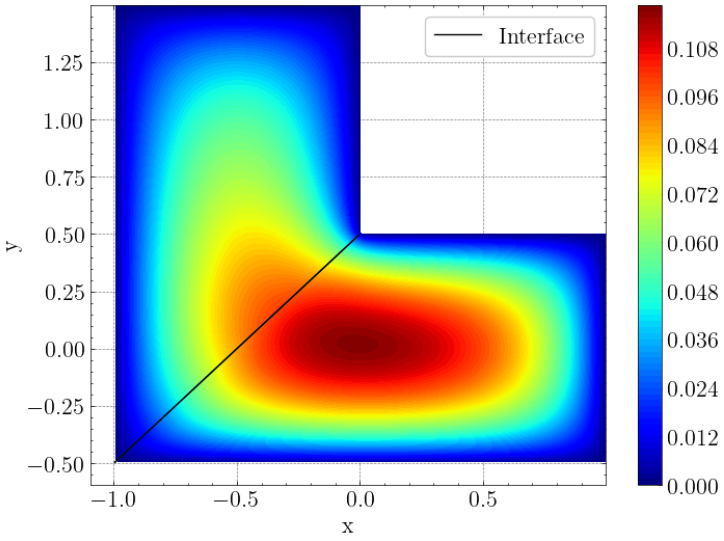


Figure 5.31: Interface C^0 error



Absolute value of the interface errors and numerical approximation for $k_1 = 2$, $p_1 = -2, -1, 0, 1$ and $p_2 = 0, 1$.

6

Conclusion

In this dissertation, the Method of Fundamental Solutions was used within the context of the Dirac operator with infinite mass boundary conditions, as well as in transmission problems involving the Poisson equation. This work also aimed to contribute to the comprehension of the method's capabilities and limitations through these applications while presenting its theoretical background.

Concerning the Dirac operator, the extension of the proof regarding the absence of separation of variables in polar coordinates is interesting in a numerical point of view: although anticipated, this extension complicates the utilization of specific solutions describing angular behavior near corner tips. Numerical simulations proved valuable in validating previous conjectures and generating new ones, reaffirming the MFS's efficacy in addressing specific challenges and solidifying its practical utility within the domain of numerical methods for Partial Differential Equations. Numerical evidence supporting the generalizations of the Faber-Krahn inequality and the Ashbaugh-Benguria theorem for both polygonal and smooth domains was established, further enhancing the credibility of these conjectures. Additionally, the unexpected behavior of the third eigenvalue in the Dirac operator's spectrum was revealed, signaling potential disparities from the behavior of the Laplace operator, particularly at the beginning of the spectrum.

The application of the Method of Fundamental Solutions with particular solutions to transmission problems is also significant, since it is the first study that introduces the use of particular solutions to better describe the solution's behavior near corners in these type of PDEs. Although this technique was previously employed for other types of PDEs, its application to transmission problems represents a novel advancement, expanding the practical utility of the method.

Nonetheless, numerous questions remain unanswered, setting the stage for future work. The study spectrum of the Dirac operator, alongside its distinct behavior compared to the Laplace operator, requires further investigation. This includes the behavior of the second eigenvalue and the optimality of shapes for various eigenvalues. Future explorations may uncover valuable insights that can might be useful for such theoretical problems, advancing our understanding of spectral geometry. Is the optimal shape for the second eigenvalue still two disjoint balls with the same volume, and does this insight provide a basis for understanding higher eigenvalues? Can the Method of Fundamental Solutions be adapted to handle topological changes, like connectedness, and how can the behavior of eigenvalues in such cases be predicted? Notice that for the third eigenvalue we still do not know if the optimal shape is even connected. If it is not, then the shape presented is not the optimal one. For the Laplacian, these topological questions can be addressed using Theorem 3.1.7, but is there any analogous results for the Dirac operator with infinite mass boundary conditions?

Similarly, the transmission problem presents its own set of questions when using the Method of Fundamental Solutions. Is it possible to derive an *a posteriori* error estimate, such that the error between an exact solution and a numerical solution is bounded in some norm? Is it possible to use different particular solutions whose behavior better adapt to the problem? How can these particular solutions

be tailored to yield more accurate results on the interface between domains? Addressing these issues could further refine the Method of Fundamental Solutions and broaden its applicability.

Our objective remains to continue this study, tackle these problems, and adapt these methods to push the boundaries of the already achieved results. This dissertation represents another step towards more accurate numerical methods for complex Partial Differential Equations, with the potential to uncover new insights into the behaviors of these mathematical entities, and to develop techniques that may aid in the solution of these questions in future research.

Bibliography

[1] Z.-C. Li and T.-T. Lu, "Singularities and treatments of elliptic boundary value problems," *Mathematical and Computer Modelling*, vol. 31, no. 8-9, pp. 97–145, 2000.

[2] C. J. Alves and P. R. Antunes, "The method of fundamental solutions applied to some inverse eigenproblems," *SIAM Journal on Scientific Computing*, vol. 35, no. 3, pp. A1689–A1708, 2013.

[3] S. Y. Reutskiy, "The method of fundamental solutions for helmholtz eigenvalue problems in simply and multiply connected domains," *Engineering Analysis with Boundary Elements*, vol. 30, no. 3, pp. 150–159, 2006.

[4] P. R. Antunes and P. Freitas, "On the inverse spectral problem for euclidean triangles," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 467, no. 2130, pp. 1546–1562, 2011.

[5] D. Krejcirik, S. Larson, and V. Lotoreichik, "Problem list of the aim workshop," 2019. [Online]. Available: http://aimpl.org/shapesurface/

[6] T. Betcke and L. N. Trefethen, "Reviving the method of particular solutions," *SIAM review*, vol. 47, no. 3, pp. 469–491, 2005.

[7] W. Rudin, *Functional Analysis*, ser. International series in pure and applied mathematics. McGraw-Hill, 1991. [Online]. Available: https://books.google.pt/books?id=Sh_vAAAAMAAJ

[8] H. Brezis and H. Brézis, *Functional analysis, Sobolev spaces and partial differential equations*. Springer, 2011, vol. 2, no. 3.

[9] W. Arendt and K. Urban, *Partielle differenzialgleichungen*. Springer, 2010.

[10] J. L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications: Vol. 1*. Springer Science & Business Media, 2012, vol. 181.

[11] H. Tavares, "Partial differential equations, lecture notes," https://ulisboa-my.sharepoint.com/:b:/g/personal/ist27898_tecnico_ulisboa_pt/EfizUpSVmCRCo0Nst2a-sacBDwFrkPuh9w8A55EU9BoFmA?e=kY5B2R.

[12] L. C. Evans, *Partial differential equations*. American Mathematical Society, 2022, vol. 19.

[13] A. Bogomolny, "Fundamental solutions method for elliptic boundary value problems," *SIAM Journal on Numerical Analysis*, vol. 22, no. 4, pp. 644–669, 1985.

[14] S. N. Chandler-Wilde, D. P. Hewett, and A. Moiola, "Sobolev spaces on non-lipschitz subsets of r n with application to boundary integral equations on fractal screens," *Integral Equations and Operator Theory*, vol. 87, no. 2, pp. 179–224, 2017.

[15] D. P. Hewett and A. Moiola, "A note on properties of the restriction operator on sobolev spaces," *Journal of Applied Analysis*, vol. 23, no. 1, pp. 1–8, 2017.

[16] G. Geymonat, "Trace theorems for sobolev spaces on lipschitz domains. necessary conditions," in *Annales mathématiques Blaise Pascal*, vol. 14, no. 2, 2007, pp. 187–197.

[17] J. Necas, *Direct methods in the theory of elliptic equations*. Springer Science & Business Media, 2011.

[18] R. A. Adams and J. J. Fournier, *Sobolev spaces*. Elsevier, 2003.

[19] G. Chen and J. Zhou, *Boundary element methods with applications to nonlinear problems*. Springer Science & Business Media, 2010, vol. 7.

[20] L. Hörmander, *The analysis of linear partial differential operators I: Distribution theory and Fourier analysis*. Springer, 2015.

[21] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*. CRC press, 2015.

[22] R. Courant and D. Hilbert, *Methods of mathematical physics: partial differential equations*. John Wiley & Sons, 2008.

[23] D. Borthwick, *Spectral Theory: Basic Concepts and Applications*. Springer Nature, 2020, vol. 284.

[24] M. Reed and B. Simon, *II: Fourier Analysis, Self-Adjointness*. Elsevier, 1975, vol. 2.

[25] A. Henrot, *Extremum problems for eigenvalues of elliptic operators*. Springer Science & Business Media, 2006.

[26] ——, *Shape optimization and spectral theory*. De Gruyter Open, 2017.

[27] E. Krahn, "Über minimaleigenschaften der kugel in drei und mehr dimensionen," *(No Title)*, 1926.

[28] P. Freitas and D. Krejčiřík, "A sharp upper bound for the first dirichlet eigenvalue and the growth of the isoperimetric constant of convex domains," *Proceedings of the American Mathematical Society*, vol. 136, no. 8, pp. 2997–3006, 2008.

[29] S. A. Wolf and J. B. Keller, “Range of the first two eigenvalues of the laplacian,” *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, vol. 447, no. 1930, pp. 397–412, 1994.

[30] A. Henrot and D. Bucur, “Minimization of the third eigenvalue of the dirichlet laplacian,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 456, no. 1996, pp. 985–996, 2000.

[31] D. Bucur, “Minimization of the k-th eigenvalue of the dirichlet laplacian,” *Archive for Rational Mechanics and Analysis*, vol. 206, no. 3, pp. 1073–1083, 2012.

[32] D. Mazzoleni and A. Pratelli, “Existence of minimizers for spectral problems,” *Journal de Mathématiques Pures et Appliquées*, vol. 100, no. 3, pp. 433–453, 2013.

[33] G. Pólya and G. Szegö, *Isoperimetric inequalities in mathematical physics*. Princeton University Press, 1951, no. 27.

[34] B. Bogosel and D. Bucur, “On the polygonal faber-krahn inequality,” *arXiv preprint arXiv:2203.16409*, 2022.

[35] J. Gómez-Serrano and G. Orriols, “Any three eigenvalues do not determine a triangle,” *Journal of Differential Equations*, vol. 275, pp. 920–938, 2021.

[36] M. Kac, “Can one hear the shape of a drum?” *The american mathematical monthly*, vol. 73, no. 4P2, pp. 1–23, 1966.

[37] C. Gordon, D. Webb, and S. Wolpert, “Isospectral plane domains and surfaces via riemannian orbifolds,” *Inventiones mathematicae*, vol. 110, no. 1, pp. 1–22, 1992.

[38] P. A. M. Dirac, “The quantum theory of the electron,” *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 117, no. 778, pp. 610–624, 1928.

[39] V. Lotoreichik and T. Ourmières-Bonafos, “A sharp upper bound on the spectral gap for graphene quantum dots,” *Mathematical Physics, Analysis and Geometry*, vol. 22, pp. 1–30, 2019.

[40] P. Briet and D. Krejčiřík, “Spectral optimization of dirac rectangles,” *Journal of Mathematical Physics*, vol. 63, no. 1, p. 013502, 2022.

[41] P. R. Antunes, R. D. Benguria, V. Lotoreichik, and T. Ourmières-Bonafos, “A variational formulation for dirac operators in bounded domains. applications to spectral geometric inequalities,” *Communications in Mathematical Physics*, vol. 386, no. 2, pp. 781–818, 2021.

[42] T. Vu, “Spectral inequality for dirac right triangles,” *arXiv preprint arXiv:2302.13040*, 2023.

[43] R. D. Benguria, S. Fournais, E. Stockmeyer, and H. Van Den Bosch, “Spectral gaps of dirac operators describing graphene quantum dots,” *Mathematical Physics, Analysis and Geometry*, vol. 20, pp. 1–12, 2017.

[44] A. Quarteroni and A. Valli, *Domain decomposition methods for partial differential equations*. Oxford University Press, 1999, no. BOOK.

[45] C. J. Alves, “On the choice of source points in the method of fundamental solutions,” *Engineering Analysis with Boundary Elements*, vol. 33, no. 12, pp. 1348–1361, 2009.

[46] R. Narasimhan and Y. Nievergelt, *Complex analysis in one variable*. Springer Science & Business Media, 2012.

[47] R. Kress, *Linear Integral Equations*, ser. Applied Mathematical Sciences. Springer New York, 2013.

[48] D. Colton and R. Kress, *Integral equation methods in scattering theory*. SIAM, 2013.

[49] S. Salsa, *Partial differential equations in action: from modelling to theory*. Springer, 2016, vol. 99.

[50] S. Valtchev, “Numerical analysis of methods with fundamental solutions for acoustic and elastic wave propagation problems,” Ph.D. dissertation, Universidade Técnica de Lisboa, Instituto Superior Técnico, 2008.

[51] Y.-S. Smyrlis, “Applicability and applications of the method of fundamental solutions,” *Mathematics of computation*, vol. 78, no. 267, pp. 1399–1434, 2009.

[52] C. Alves and C. Chen, “A new method of fundamental solutions applied to nonhomogeneous elliptic problems,” *Advances in Computational Mathematics*, vol. 23, pp. 125–142, 2005.

[53] P. R. Antunes, “Reducing the ill conditioning in the method of fundamental solutions,” *Advances in Computational Mathematics*, vol. 44, pp. 351–365, 2018.

[54] ——, “A numerical algorithm to reduce ill-conditioning in meshless methods for the helmholtz equation.” *Numer. Algorithms*, vol. 79, no. 3, pp. 879–897, 2018.

[55] S. Christiansen and E. Meister, “Condition number of matrices derived from two classes of integral equations,” *Mathematical Methods in the Applied Sciences*, vol. 3, no. 1, pp. 364–392, 1981.

[56] T. Kitagawa, “On the numerical stability of the method of fundamental solution applied to the dirichlet problem,” *Japan Journal of Applied Mathematics*, vol. 5, pp. 123–133, 1988.

[57] ——, “Asymptotic stability of the fundamental solution method,” *Journal of Computational and Applied Mathematics*, vol. 38, no. 1-3, pp. 263–269, 1991.

[58] A. H. Barnett and T. Betcke, “Stability and convergence of the method of fundamental solutions for helmholtz problems on analytic domains,” *Journal of Computational Physics*, vol. 227, no. 14, pp. 7003–7026, 2008.

[59] C. J. Alves and P. R. Antunes, “The method of fundamental solutions applied to the calculation of eigenfrequencies and eigenmodes of 2d simply connected shapes,” *CMC-TECH SCIENCE PRESS-*, vol. 2, no. 4, p. 251, 2005.

[60] R. P. Brent, “An algorithm with guaranteed convergence for finding a zero of a function,” *The computer journal*, vol. 14, no. 4, pp. 422–425, 1971.

[61] C. Moler and L. Payne, “Bounds for eigenvalues and eigenvectors of symmetric operators,” *SIAM Journal on Numerical Analysis*, vol. 5, no. 1, pp. 64–70, 1968.

[62] P. R. Antunes and S. S. Valtchev, “A meshfree numerical method for acoustic wave propagation problems in planar domains with corners and cracks,” *Journal of Computational and Applied Mathematics*, vol. 234, no. 9, pp. 2646–2662, 2010.

[63] T. Kato, *Perturbation theory for linear operators*. Springer Science & Business Media, 2013, vol. 132.

[64] C. J. Alves, N. F. Martins, and S. S. Valtchev, “Domain decomposition methods with fundamental solutions for helmholtz problems with discontinuous source terms,” *Computers & Mathematics with Applications*, vol. 88, pp. 16–32, 2021.

[65] M. A. Golberg, C.-S. Chen, and S. Karur, “Improved multiquadric approximation for partial differential equations,” *Engineering Analysis with boundary elements*, vol. 18, no. 1, pp. 9–17, 1996.

Glossary

This document is incomplete. The external file associated with the glossary 'main' (which should be called `main.gls`) hasn't been created.

Check the contents of the file `main.glo`. If it's empty, that means you haven't indexed any of your entries in this glossary (using commands like `\gls` or `\glsadd`) so this list can't be generated. If the file isn't empty, the document build process hasn't been completed.

If you don't want this glossary, add `nomain` to your package option list when you load `glossaries-extra.sty`. For example:

```
\usepackage[nomain]{glossaries-extra}
```

Try one of the following:

- Add `automake` to your package option list when you load `glossaries-extra.sty`. For example:

  ```
  \usepackage[automake]{glossaries-extra}
  ```

- Run the external (Lua) application:

  ```
  makeglossaries-lite.lua "main"
  ```

- Run the external (Perl) application:

  ```
  makeglossaries "main"
  ```

Then rerun LaTeX on this document.

This message will be removed once the problem has been fixed.

A

Code of Project

B

A Large Table