

EÖTVÖS LORÁND UNIVERSITY  
Faculty of Science

NUMERICAL SOLUTION OF  
BOUNDARY VALUE PROBLEMS:  
A NEURAL NETWORK-BASED APPROACH

MASTER'S THESIS

**Written by**

Miskei Ferenc István  
applied mathematics MSc student

**Supervisor**

Izsák Ferenc, Ph.D.  
associate professor



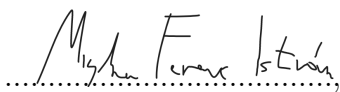
Budapest, 2023.

# Author's declaration

I, the undersigned **Miskei Ferenc István** (ELTE Neptun ID: **YC62WJ**) hereby declare that I am the sole author of this thesis. To the best of my knowledge, this thesis contains no material previously published by any other person except where due acknowledgment has been made. This thesis contains no material which has been accepted as part of the requirements of any other academic degree or non-degree program, in English or any other language.

This is a true copy of the thesis, including final revisions as accepted by my examiners.

I understand that my thesis may be made available to the public, electronically or otherwise.



Budapest, June 7, 2023

Miskei Ferenc István

# Acknowledgements

First, I would like to express my deepest gratitude towards my supervisor, Dr. Izsák Ferenc, for his diligent guidance in my studies through this challenging topic. He never missed an opportunity to pose interesting new questions for me to ponder, make mathematical and stylistic suggestions, or refer me to relevant literature, thus helping me along every step of the thesis writing process. Without his insightful feedback and constructive criticism, this thesis could not have become what it is.

I am also grateful to all my other professors and other faculty members during this program for sharing some of their fascinating knowledge with me and my peers and always being the most helpful that they could be.

Last but certainly not least, I thank my friends and family for their unwavering support and reassurance during this difficult period. Special acknowledgment is due to my faithful and loving companion and soon-to-be wife, Zsóka. I cannot express in words how much I value her loving care and emotional support during the last several months: from listening to my complaints to nudging me to do my job, and for this, I will always be grateful to her.

# Contents

<b>1</b>	<b>Theoretical background</b>	<b>7</b>
1.1	Prerequisites from analysis . . . . .	7
1.1.1	Lebesgue integral and Lebesgue spaces . . . . .	10
1.1.2	Generalized derivatives and distribution theory . . . . .	12
1.2	Prerequisites from PDE theory . . . . .	16
1.2.1	Fundamental solutions and PDEs . . . . .	17
1.2.2	Laplace's equation and Poisson's equation . . . . .	18
1.3	Prerequisites from functional analysis . . . . .	22
1.3.1	Single variable Sobolev spaces . . . . .	22
1.3.2	Multivariable Sobolev spaces . . . . .	23
1.3.3	Existence and uniqueness theorems . . . . .	27
1.4	Basics of artificial neural networks . . . . .	28
1.4.1	Structure of neural networks . . . . .	28
<b>2</b>	<b>The method of fundamental solutions and artificial neural networks</b>	<b>30</b>
2.1	Introducing the MFS scheme through an example . . . . .	30
2.2	Solving the MFS scheme for Laplace's equation with Dirichlet-type boundary condition using a primitive neural network . . . . .	33
2.3	Introducing a generalized version of the MFS algorithm . . . . .	35
2.4	Solving the generalized MFS scheme for Poisson's equation using a primitive neural network . . . . .	38
<b>3</b>	<b>Experiments</b>	<b>41</b>
3.1	Experiment 1: Laplace's equation with inhomogeneous Dirichlet-type boundary condition . . . . .	41
3.2	Experiment 2: Poisson's equation with inhomogeneous Dirichlet-type boundary condition . . . . .	42
3.3	Experiment proposal: Laplace's equation with non-linear boundary conditions . . . . .	44
<b>4</b>	<b>Conclusions</b>	<b>47</b>
4.1	Summary . . . . .	47
4.2	Further directions . . . . .	47

# Introduction

*“It was as though applied mathematics was my spouse,  
and pure mathematics was my secret lover.”*  
– Edward Frenkel

Solving partial differential equations (PDEs) is a central field in modern applied sciences. Some prominent applications of PDEs in science appear in physics, chemistry, epidemiology, and engineering. However, finding exact solutions of PDEs that emerge in real-life applications is often impossible, and therefore, we must resort to numerical methods. The two most traditional approaches to solving PDEs are finite difference methods (FDM) and finite element methods (FEM). In this thesis, we are going to discuss two numerical methods that do not belong to either of these categories: the method of fundamental solutions and another one of its variants.

A new approach that got some attention in recent years is solving PDEs with the application of neural networks (NN or neural nets), as they may function as universal approximators. Hornik, Stinchcombe, and White (1989) [16] established “*multilayer feedforward [neural] networks as a class of universal approximators [...] provided sufficiently many hidden units are available*” in the sense that “*feedforward networks are capable of arbitrarily accurate approximation to any real-valued continuous function over a compact set.*” This positive theoretical result gives us confidence in the capability of multilayer feedforward neural networks of sufficient complexity to approximate PDE’s classical (and thus continuous) solutions. Moreover, since the space of continuous functions is dense in the Sobolev spaces in some sense that weak solutions are usually sought in (see Meyers–Serrin theorem), we can conclude that the same result holds for the approximation of weak solutions as well. Therefore, the use of multilayer feedforward neural networks is established as a valid approach to finding the weak solution of well-posed PDEs.

Applying neural networks to solve various PDE problems has shown promising results in recent literature. Prominent approaches to solving PDEs using neural networks include the so-called Physics-Informed Neural Networks (PINN models), Deep Galerkin Method (DMG), Finite Element Neural Networks (FENN), Residual Networks (RN), the Method of Fundamental Solutions (MFS), and even some adaptations of Convolutional Neural Networks (CNNs). In this thesis, we examine the MFS applied to boundary value problems and implemented via rather primitive neural networks and discuss applicability to more general cases of PDEs.

The method of fundamental solutions (MFS) approximates the solution of a boundary value problem as a linear combination of the operator’s fundamental solutions. The collocation points where the singularities are placed may either be predetermined by the user or may be also a part of the problem, see [6].

One advantage of the MFS approach implemented via neural networks compared to FDM and FEM models is that an already trained NN serves as a solver for the PDE, and provides new solutions as a matter of function evaluation when changing the auxiliary conditions, such as boundary conditions, force terms, etc. In other words, it is akin to inverting the coefficient matrix of a system of linear equations rather than solving it over and over again when new right-hand side vectors are given.

The goals of this thesis are as follows.

1. Lay a firm foundation of the theoretical elements and basic tools used in PDE theory and functional analysis that are used in these methods, including many examples.
2. Introduce the method of fundamental solutions and its proposed modification through concrete examples, and then examine their relationship to primitive neural networks.
3. Conduct numerical experiments that serve as proof of concept for the viability of the methods discussed.
4. Discuss directions of further generalization and fine-tuning of the methods introduced.

Each chapter corresponds to one of these goals. Throughout the thesis, we assume a reader that is familiar with bachelor-level mathematics but perhaps has little or no specific training in numerical analysis per se. In other words, we attempt to be more educational than technical in our approach and will devote time to give concrete examples of the theoretical elements discussed throughout the thesis.

# Chapter 1

## Theoretical background

*“A house built on granite and strong foundations, not even the onslaught of pouring rain, gushing torrents, and strong winds will be able to pull down.”*  
– Haile Selassie

Before diving into the nitty-gritty of applying neural networks to solve partial differential equations numerically, we believe it is worth spending some time to lay out a firm theoretical foundation for both partial differential equations and neural networks.

We shall only work out examples, constructions, and proofs where we believe it provides insight, and we will generally prioritize clarity over technicality. We will omit technical details, such as most proofs, to avoid turning this chapter into a textbook. We refer the reader to sources such as [5], [17], and [2] for more detailed explanations and proofs.

### 1.1 Prerequisites from analysis

We begin with some definitions and basic examples related to measure theory, which are essential for understanding the concepts of the weak derivative, Lebesgue spaces, and, ultimately, Sobolev spaces. The letter  $\Omega$  shall denote a non-empty, open subset of  $\mathbb{R}^n$ .

**Definition 1.1.1.** (Support of a function)

Let  $f: \Omega \rightarrow \mathbb{R}$ . Let  $Z$  be the broadest open subset of  $\Omega$  for which  $f$  restricted to  $Z$  is zero, that is

$$Z := \bigcup \{U \subset \Omega \text{ open, such that } f|_U \equiv 0\} \quad (1.1)$$

Then, the support of the function  $f$  is the set  $\text{supp}(f) := \Omega \setminus Z$ .

*Remark 1.1.2.* If  $\mathbf{x} \in (\Omega \setminus \text{supp}(f))$ , then  $f(\mathbf{x}) = 0$ . Hence, if  $S \subset (\Omega \setminus \text{supp}(f))$ , then  $\int_S f = 0$ .

*Remark 1.1.3.* If  $f^{-1}[\{0\}]$  has no interior points, then  $Z = \emptyset$ , so  $\text{supp}(f) = \Omega$ . For example, for the function  $\sin: \mathbb{R} \rightarrow \mathbb{R}$ ,  $Z$  is the largest open subset of  $\sin^{-1}[\{0\}] = \{k\pi \mid k \in \mathbb{Z}\}$ . Since this is a set of isolated points, and  $Z$  is open, and therefore  $\text{supp}(\sin) = \mathbb{R} \setminus \emptyset = \mathbb{R}$ , even though the sine function has infinitely many zeroes.

The importance of the following concept lies in the fact that (Lebesgue) integrals are in some sense “blind” to changes of the function on certain sets. This is expected, as the Riemann integral also does not change if we change the function at finitely many points. Since we solve differential equations by “integration” – or methods that mimic integration – it is expected that in the general theory of differential equations, uniqueness theorems can only apply on sets that (Lebesgue) integrals “can see.”

**Definition 1.1.4.** (Null sets in  $\mathbb{R}^n$ )

We say that  $N$  is a null set if, for every  $\varepsilon > 0$ , there exists a countable collection of balls  $\mathcal{B} = \{B_i(\mathbf{x}_i, r_i)\}_{i \in I}$ , ( $I \subset \mathbb{N}$ ) such that

$$N \subset \bigcup_{B_i \in \mathcal{B}} B_i \quad \text{and} \quad \sum_{i \in I} r_i^n \leq \varepsilon.$$

*Remark 1.1.5.* The second condition of this definition is equivalent to the sum of the  $n$ -dimensional “volumes” (measure) of the balls being at most a constant multiple of  $\varepsilon$ . (This constant is the volume of the  $n$ -dimensional unit ball.)

*Remark 1.1.6.* If the set  $S$  contains an open ball  $B(\mathbf{x}, r)$  – for instance, if  $S$  is an open set – then for  $\varepsilon = \frac{r^n}{2}$ , there can be no sufficient covering, and thus  $S$  is not a null set. Therefore, the domain  $\Omega$  is never a null set.

**Example 1.1.7.** If  $S$  is a countable set, i.e.,  $S \subset \{\mathbf{x}_i\}_{i \in \mathbb{N}^+}$ , then  $S$  is a null set. Indeed,

$$\mathcal{B} = \left\{ B_i := B\left(\mathbf{x}_i, \frac{\sqrt[n]{\varepsilon}}{2^{i/n}}\right) \right\}_{i \in \mathbb{N}^+} \quad (1.2)$$

satisfies the definition, as

$$S \subset \underbrace{\bigcup_{i=1}^{+\infty} \{\mathbf{x}_i\}}_{\mathbf{x}_i \in B_i} \subset \bigcup_{B_i \in \mathcal{B}} B_i \quad \text{and} \quad \sum_{i=1}^{+\infty} \left(\frac{\sqrt[n]{\varepsilon}}{2^{i/n}}\right)^n = \sum_{i=1}^{+\infty} \frac{\varepsilon}{2^i} = \varepsilon \leq \varepsilon. \quad (1.3)$$

*Remark 1.1.8.* The countable union of null sets is also countable.

*Remark 1.1.9.* Since  $S = \mathbb{Q}$  is countable, it is a null set in  $\mathbb{R}$ .

*Remark 1.1.10.* There exist sets of measure zero in  $\mathbb{R}$  that are uncountable. An important example of this phenomenon is the Cantor set  $\mathcal{C} := \bigcap_{n=1}^{+\infty} C_n$ , where  $C_0 = [0, 1]$ , and we obtain  $C_{n+1}$  from  $C_n$  by removing the middle third of every line segment. Hence,  $2^n$  line segments of length  $1/3^n$  units can cover  $\mathcal{C}$ , implying that  $\mathcal{C}$  is a null set in  $\mathbb{R}$ .

*Remark 1.1.11.* If  $S \subset \mathbb{R}^{n-1}$ , and  $\varphi: S \rightarrow \mathbb{R}$  is a continuous function, then the graph of  $\varphi$  is a null set in  $\mathbb{R}^n$ . Hence the boundary of “sufficiently nice” subsets of  $\mathbb{R}^n$  are null sets in  $\mathbb{R}^n$ .

**Definition 1.1.12.** (Almost everywhere)

A proposition  $P(\mathbf{x})$  that depends on the point  $\mathbf{x}$  is true almost everywhere if the set where it is not true is a null set. We often denote this property by the abbreviation “a.e.”.

The following three examples should illustrate the meaning of the null set property more concretely – at least in  $\mathbb{R}$ .

**Example 1.1.13.** The so-called Dirichlet function  $D: [0, 1] \rightarrow \mathbb{R}$ ,

$$D(x) := \begin{cases} 1, & \text{if } x \in [0, 1] \cap \mathbb{Q} \\ 0, & \text{if } x \in [0, 1] \setminus \mathbb{Q} \end{cases} \quad (1.4)$$

is equal to zero almost everywhere, because  $\{x \in \mathbb{R} \mid D(x) \neq 0\} = \mathbb{Q}$  is a null set in  $\mathbb{R}$ . The plot of  $D$  would appear to be the line segment  $[0, 1] \times \{0\}$ , the same as the plot of  $f := 0$ .



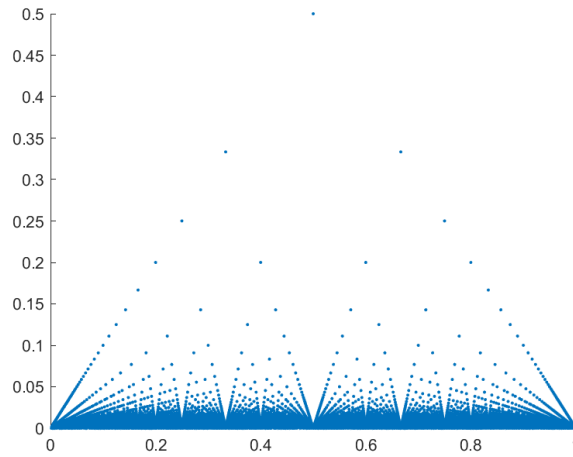


Figure 1.1: Thomae function on  $[0, 1]$ .

**Example 1.1.14.** Consider the function

$$T(x) := \begin{cases} \frac{1}{q}, & \text{if } x = \frac{p}{q}, \text{ where } p, q \in \mathbb{Z}, \gcd(p, q) = 1 \text{ and } q > 0, \\ 0, & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

named after German mathematician Karl Johannes Thomae and plotted below. One can show that this function is continuous at  $x$  if and only if  $x$  is irrational, and thus by Remark 1.1.9, this function is continuous almost everywhere.

**Example 1.1.15.** The Cantor ternary function  $c$  plotted below – also known as “the Devil’s staircase” – is continuous on  $[0, 1]$ , has a derivative  $c'(x) = 0$  almost everywhere, and yet  $c$  is not constant. The sum of the lengths of the constant segments is  $\sum_{k=0}^{+\infty} \frac{2^k}{3^{k+1}} = 1$ . The derivative  $c'(x)$  does not exist if and only if  $x \in \mathcal{C}$ .

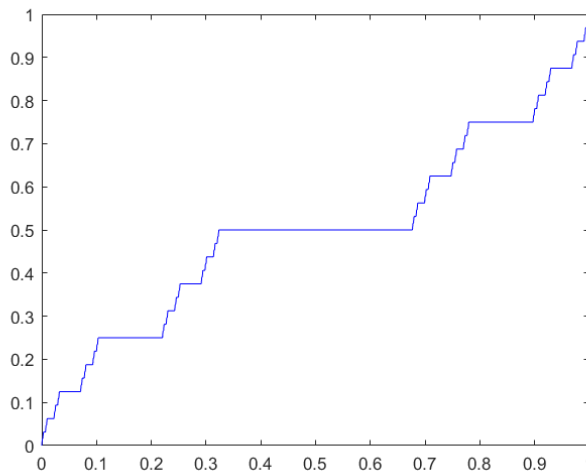


Figure 1.2: Cantor function  $c$  on  $[0, 1]$ .

*Remark 1.1.16.* The functions  $D$ ,  $T$ , and  $c'$  are equal to each other almost everywhere, even though they seem to be very different objects.

The following class of functions will prove to be helpful when we define weak derivatives and generalized functions.

**Definition 1.1.17.** (Smooth functions of compact support)

Let  $k \in \mathbb{N}^+$ . Then  $C^k(\Omega)$  denotes the set of real-valued functions that are at least  $k$  times continuously differentiable everywhere, and  $C^\infty(\Omega) := \bigcap_{k=1}^{+\infty} C^k(\Omega)$ . Furthermore,

$$C_0^k(\Omega) := \{\varphi \in C^k(\Omega) \mid \text{supp}(\varphi) \text{ is a compact set}\}, \quad (1.5)$$

and

$$C_0^\infty(\Omega) := \{\varphi \in C^\infty(\Omega) \mid \text{supp}(\varphi) \text{ is a compact set}\}. \quad (1.6)$$

*Remark 1.1.18.* We define the spaces  $C^k(\overline{\Omega})$  and  $C_0^k(\overline{\Omega})$  such that all the partial derivatives of at most order  $k$  must continuously extend to the boundary of  $\Omega$ . Accordingly, we always consider the extended functions. In  $C_0^k(\overline{\Omega})$ , we also require that  $\text{supp}(\varphi)$  is a compact subset of the open set  $\Omega$ , which implies that  $\varphi|_{\partial\Omega} \equiv 0$ , and all partial derivatives also vanish on  $\partial\Omega$ .

*Remark 1.1.19.* These sets form vector spaces under the usual operations defined on functions.

The following formulas are quintessential in the study of PDEs and Sobolev spaces. They may be understood as multivariable generalizations of the integration by parts formula.

**Theorem 1.1.20.** (*Green–Gauß formulas*)

Let  $\Omega \subset \mathbb{R}^n$  be a bounded set with piecewise smooth boundary,  $f \in C^1(\overline{\Omega})$  and  $g \in C^2(\overline{\Omega})$ . Then

$$\int_{\Omega} f(\mathbf{x}) \Delta g(\mathbf{x}) \, d\mathbf{x} = \int_{\partial\Omega} f(\mathbf{x}) \partial_{\mathbf{n}} g(\mathbf{x}) \, d\mathbf{S} - \int_{\Omega} \nabla f(\mathbf{x}) \cdot \nabla g(\mathbf{x}) \, d\mathbf{x}, \quad (1.7)$$

or simply, without variables

$$\int_{\Omega} f \Delta g = \int_{\partial\Omega} f \partial_{\mathbf{n}} g - \int_{\Omega} \nabla f \cdot \nabla g.$$

Furthermore, if  $f, g \in C^2(\overline{\Omega})$ , and  $p \in C^1(\overline{\Omega})$ , then

$$\int_{\Omega} \left( f(\mathbf{x}) \operatorname{div} (p(\mathbf{x}) \nabla g(\mathbf{x})) - g(\mathbf{x}) \operatorname{div} (p(\mathbf{x}) \nabla f(\mathbf{x})) \right) \, d\mathbf{x} = \int_{\partial\Omega} \left( f(\mathbf{x}) \partial_{\mathbf{n}} g(\mathbf{x}) - g(\mathbf{x}) \partial_{\mathbf{n}} f(\mathbf{x}) \right) \, d\mathbf{S}, \quad (1.8)$$

or simply

$$\int_{\Omega} \left( f \operatorname{div} (p \nabla g) - g \operatorname{div} (p \nabla f) \right) = \int_{\partial\Omega} (f \partial_{\mathbf{n}} g - g \partial_{\mathbf{n}} f),$$

which reduces to the following formula if  $p \equiv 1$

$$\int_{\Omega} (f \Delta g - g \Delta f) = \int_{\partial\Omega} (f \partial_{\mathbf{n}} g - g \partial_{\mathbf{n}} f). \quad (1.9)$$

For the proof, we refer to [5].

### 1.1.1 Lebesgue integral and Lebesgue spaces

We take it for granted that the reader is familiar with Lebesgue integrals. An elegant and brief construction – based on Riesz Frigyes’ approach – by Komornik Vilmos can be found in [12]. Otherwise, it is an appropriate conceptual approximation to think of Lebesgue integrable functions as the completion of the space of Riemann integrable functions and the Lebesgue integral as a continuous extension of the Riemann integral functional to the space of Lebesgue integrable functions. See [3] for the details of this approach.

One remark on the Lebesgue integral that is worth making explicit immediately is that if  $f = g$  a.e., then  $\int_{\Omega} f = \int_{\Omega} g$ . This fact limits the sense in which we can even expect uniqueness theorems and is thus built into the very foundation of the whole theory.

**Example 1.1.21.** Dirichlet's function  $D: [0, 1] \rightarrow \mathbb{R}$  defined in Example 1.1.13 is equal to  $f \equiv 0$  a.e, and therefore  $\int_{[0,1]} D = \int_{[0,1]} 0 = 0$  in the Lebesgue integral sense.

**Definition 1.1.22.** ( $\mathcal{L}^1(\Omega)$  and  $L^1(\Omega)$ )

Let  $\mathcal{L}^1(\Omega)$  be the set of functions  $f: \Omega \rightarrow \mathbb{R}$  for which the Lebesgue integral  $\int_{\Omega} |f|$  exists and is finite

$$\mathcal{L}^1(\Omega) := \left\{ f: \Omega \rightarrow \mathbb{R} \mid \int_{\Omega} |f| < +\infty \right\}. \quad (1.10)$$

Let us take the equivalence relation  $f \sim g \iff f = g$  almost everywhere. Now, we define the Lebesgue space  $L^1(\Omega)$  as

$$L^1(\Omega) := \mathcal{L}^1(\Omega) / \sim, \quad (1.11)$$

i.e,  $L^1(\Omega)$  consists of equivalence classes of  $\mathcal{L}^1(\Omega)$  with respect to the equivalence relation  $\sim$ . Finally, this set is equipped with a vector space structure and the norm

$$\|f\|_{L^1(\Omega)} := \int_{\Omega} |f|. \quad (1.12)$$

Henceforth, we always use these equivalence classes instead of the individual functions.

*Remark 1.1.23.* Considering equivalence classes of individual functions instead of functions is necessary so that the property  $\|f\|_{L^1(\Omega)} = 0 \implies f = 0$  holds, which is a requirement for  $\|\cdot\|_{L^1(\Omega)}$  to be a norm. However, this raises several problems that need resolution, such as verifying that vector operations are well-defined. Sorting out these technicalities is beyond the scope of this thesis.

*Remark 1.1.24.* The function  $f(x) = \frac{\sin(x)}{x}$  does not belong to the space  $L^1(\mathbb{R})$ , because  $\int_{-\infty}^{+\infty} |f(x)| dx = +\infty$ , even though  $\int_{-\infty}^{+\infty} f(x) dx = \pi < +\infty$  in the (Riemann) improper integral sense.

**Definition 1.1.25.** ( $L^p(\Omega)$  spaces for  $1 < p < +\infty$ )

Let the sets  $L^p(\Omega)$  contain all functions  $f$  for which  $|f|^p \in L^1(\Omega)$ . Defining the norm on this vector space

$$\|f\|_{L^p} := \left( \int_{\Omega} |f|^p \right)^{\frac{1}{p}} \quad (1.13)$$

results in a normed vector space  $L^p(\Omega)$ .

**Definition 1.1.26.** ( $L^\infty(\Omega)$  space and essential supremum)

Let

$$\|f\|_{L^\infty(\Omega)} := \text{ess sup}\{|f|\} := \inf\{M \in \mathbb{R} \mid |f(\mathbf{x})| \leq M \text{ a.e.}\}.$$

Then  $L^\infty(\Omega)$  denotes the vector space of functions for which  $\text{ess sup}\{|f|\} < +\infty$  equipped with the norm  $\|f\|_{L^\infty(\Omega)}$ .

*Remark 1.1.27.* Even though  $L^p(\Omega)$  contain classes of functions, we say that “ $f \in L^p(\Omega)$ ” or “ $f = g$ ”, by which we mean that “ $f$  belongs to one of the equivalence classes of  $L^p(\Omega)$ ” or “ $f$  and  $g$  belong to the same equivalence class of  $L^p(\Omega)$ ” respectively.

**Theorem 1.1.28.** (*Riesz–Fischer theorem*)

*The spaces  $L^p(\Omega)$  defined above are complete normed vector spaces – also known as Banach spaces – for  $1 \leq p \leq +\infty$ . That is, if  $(f_n) \subset L^p(\Omega)$  is a Cauchy sequence with respect to  $\|\cdot\|_{L^p(\Omega)}$ , then there exists  $f \in L^p(\Omega)$  such that  $\|f - f_n\|_{L^p(\Omega)} \rightarrow 0$ . See [17].*

**Theorem 1.1.29.** (*Relationship of  $L^p(\Omega)$  spaces*)

*For any  $1 \leq p < q \leq +\infty$ ,  $L^q(\Omega) \subset L^p(\Omega)$  in the sense that the identity map from  $L^q(\Omega)$  to  $L^p(\Omega)$  is bounded. In particular,  $L^\infty(\Omega) \subset L^p(\Omega) \subset L^1(\Omega)$  ( $p \geq 1$ ). See [17].*

A key result in the theory of Lebesgue integrals is a sufficient condition for the exchange of integration and taking the limit. See [17] for details.

**Theorem 1.1.30.** (*Lebesgue's dominated convergence theorem*)

Let  $g_n \in L^1(\Omega)$ , and suppose that  $g_n(x) \rightarrow g(x)$  for almost every  $x \in \Omega$ . Suppose that there exists  $m \in L^1(\Omega)$  – a so-called integrable majorant – such that for every  $n \in \mathbb{N}^+$   $|g_n| \leq m$  a.e. Then it follows that  $g \in L^1(\Omega)$ , and  $g_n \rightarrow g$  in  $L^1(\Omega)$ , and consequently

$$\lim_{n \rightarrow +\infty} \int_{\Omega} g_n = \int_{\Omega} g. \quad (1.14)$$

The following two classes of functions will play an important role in distribution theory.

**Definition 1.1.31.** (Locally integrable functions)

The function  $f: \Omega \rightarrow \mathbb{R}$  belongs to  $L^p_{\text{loc}}(\Omega)$ , if for every compact set  $C \subset \Omega$ ,  $f \in L^p(C)$ . That is,

$$L^p_{\text{loc}}(\Omega) := \left\{ f: \Omega \rightarrow \mathbb{R} \mid \forall C \subset \Omega \text{ compact set, } f \in L^p(C) \right\}. \quad (1.15)$$

Finally, we state the following important relationship between the  $C^\infty(\Omega)$  and  $L^p$  spaces.

**Theorem 1.1.32.** For every  $1 \leq p \leq +\infty$ , the space  $C^\infty(\Omega)$  is a dense subspace of  $L^p(\Omega)$  in  $\|\cdot\|_{L^p}$  norm in the following sense. For every  $f \in L^p(\Omega)$  and  $\varepsilon > 0$ , there exists  $g \in C^\infty(\Omega)$  such that  $\|f - g\|_{L^p(\Omega)} < \varepsilon$ .

*Remark 1.1.33.* Consequently, for every  $1 \leq p \leq +\infty$  and  $k \in \mathbb{N}^+$ ,  $C^k(\Omega)$  is dense in  $L^p(\Omega)$  and  $C(\Omega)$  is dense in  $L^p(\Omega)$ . See [5] for details on both claims.

## 1.1.2 Generalized derivatives and distribution theory

In this subsection, we introduce weak derivatives and generalized functions called distributions. Much of what we present here relies on Besenyei Ádám's introductory PDE course and the course's official (Hungarian) lecture notes [2]. A similar and considerably longer introduction in English is by Lawrence C. Evans [5]. To define general partial derivatives, we first define the weak derivative of single variable functions. A well-known fact from single-variable calculus is the so-called integration by parts formula. Weak versions of further differential operators (such as the gradient, divergence, curl, etc.) are defined in the same spirit.

**Theorem 1.1.34.** (*Integration by parts in single variable*)

Let  $f, \varphi \in C^1([a, b])$ . Then

$$\int_a^b f'(x)\varphi(x) \, dx = f(b)\varphi(b) - f(a)\varphi(a) - \int_a^b f(x)\varphi'(x) \, dx \quad (1.16)$$

**Corollary 1.1.35.** Let  $f \in C^1([a, b])$ ,  $\varphi \in C_0^1([a, b])$ . Then

$$\int_a^b f'(x)\varphi(x) \, dx = - \int_a^b f(x)\varphi'(x) \, dx \quad (1.17)$$

This opens a door to defining the weak derivative of integrable functions via the principle of permanence.

**Definition 1.1.36.** (Weak derivatives of integrable functions)

Let  $f \in L^1_{\text{loc}}((a, b))$ , with  $a < b \in \overline{\mathbb{R}}$ . We say that  $f' = g$  in the weak sense, if

$$\int_a^b g(x)\varphi(x) \, dx = - \int_a^b f(x)\varphi'(x) \, dx \quad \left( \forall \varphi \in C_0^1((a, b)) \right). \quad (1.18)$$

**Example 1.1.37.** Let  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) := \max\{0, x\}$ , which we know to be non-differentiable at  $x = 0$ . Then  $g(x) := \begin{cases} 1, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$  is a suitable weak derivative of  $f$ . Indeed, for any  $\varphi \in C_0^1(\mathbb{R})$ , we have that

$$\int_{-\infty}^{+\infty} g(x)\varphi(x) dx = \int_0^{+\infty} \varphi(x) dx := I, \quad (1.19)$$

and

$$\int_{-\infty}^{+\infty} f(x)\varphi'(x) dx = \int_0^{+\infty} x\varphi'(x) dx = [x\varphi(x)]_{x=0}^{x \rightarrow +\infty} - \int_0^{+\infty} \varphi(x) dx = \quad (1.20)$$

$$= \lim_{x \rightarrow +\infty} (x\varphi(x)) - (0 \cdot \varphi(0)) - \int_0^{+\infty} \varphi(x) dx = - \int_0^{+\infty} \varphi(x) dx = -I, \quad (1.21)$$

because  $\varphi(x) = 0$  for sufficiently large values of  $x$ .

*Remark 1.1.38.* Notice that we did not use the specific value  $g(0)$ . Indeed, the weak derivative is unique up to a difference on a null set only.

Of course, if  $f' = g$  in the classical sense, then  $f' = g$  in the weak sense, also. Notice that we have not specified yet what set  $g$  belongs to. The following example demonstrates that some locally integrable functions have weak derivatives that are not even locally integrable functions.

**Example 1.1.39.** Consider  $f(x) := \begin{cases} 1, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0 \end{cases}$ . This function is of course integrable on every bounded interval. Let us see what criteria its weak derivative  $g = f'$  would have to satisfy. Let  $\delta > 0$ . Then, we need that for every  $\varphi \in C_0^1([-\delta, \delta])$ ,

$$\int_{-\delta}^{\delta} g(x)\varphi(x) dx = - \int_{-\delta}^{\delta} f(x)\varphi'(x) dx = - \int_0^{\delta} \varphi'(x) dx = -(\underbrace{\varphi(\delta)}_{=0} - \varphi(0)) = \varphi(0). \quad (1.22)$$

However, there is no locally integrable function  $g$  with this property.

To verify this statement, let us take the sequence  $\delta_n := \frac{1}{n}$  and functions  $\varphi_n \in C_0^1([-\delta_n, \delta_n])$  such that  $\varphi_n(0) \equiv 1$  as

$$\varphi_n(x) := \begin{cases} \exp\left(1 + \frac{1}{(nx)^2 - 1}\right), & \text{if } |x| < \delta_n, \\ 0, & \text{if } |x| = \delta_n. \end{cases} \quad (1.23)$$

It is easy to see that  $\varphi_n$  satisfies all the conditions, hence

$$\int_{-\delta_n}^{\delta_n} g(x)\varphi_n(x) dx = \varphi_n(0) \equiv 1 \quad (\forall n \in \mathbb{N}^+). \quad (1.24)$$

Taking the limit ( $n \rightarrow +\infty$ ) will lead to the contradiction  $0 = 1$ , and thus  $g$  cannot have been a member of  $L_{\text{loc}}^1([-1, 1])$ . If  $g \in L_{\text{loc}}^1([-1, 1])$ , then  $|g|$  is an integrable majorant of the sequence of functions  $g_n = |g| \cdot \chi_{[-\delta_n, \delta_n]}$ , and thus Lebesgue's dominated convergence theorem and  $|\varphi| \leq 1$  implies

$$\left| \lim_{n \rightarrow +\infty} \int_{-\delta_n}^{\delta_n} g(x)\varphi_n(x) dx \right| \leq \lim_{n \rightarrow +\infty} \int_{-\delta_n}^{\delta_n} |g| = \lim_{n \rightarrow +\infty} \int_{-1}^1 g_n = \int_{-1}^1 \lim_{n \rightarrow +\infty} g_n = \int_{-1}^1 0 = 0. \quad (1.25)$$

*Remark 1.1.40.* The function  $f$  from the previous example is often called the Heaviside function after the English mathematician and physicist Oliver Heaviside. The supposed  $g$  is called the Dirac delta “function” and plays a central role in PDE theory.

Including objects such as the Dirac delta lead to the development of distribution theory, which defines generalized functions among which differentiation can be done without restriction. Furthermore, distributions contain as a subset a sufficiently large class of “ordinary” functions, namely, the locally integrable functions.

Next, we give the main concepts and results of distribution theory necessary for our discussion of PDE theory.

**Definition 1.1.41.** (Multi-indices)

Let  $n \in \mathbb{N}^+$ . Vectors of the form  $\alpha := (\alpha_1, \dots, \alpha_n) \in (\mathbb{N}_0)^n$  are called multi-indices. Given a multi-index  $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{N}_0)^n$ , the partial derivative  $\partial^\alpha u$  shall denote the partial derivative  $\partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n} u$ . Furthermore the order of  $\alpha$  is defined as  $|\alpha| := \alpha_1 + \alpha_2 + \dots + \alpha_n$ .

**Definition 1.1.42.** ( $\mathcal{D}(\Omega)$  space of test functions)

Let  $n \in \mathbb{N}^+$ ,  $\Omega$  be a non-empty, open subset of  $\mathbb{R}^n$ . Let us equip the vector space  $C_0^\infty(\Omega)$  with the following notion of convergence. We say that  $\varphi_n \xrightarrow{\mathcal{D}} \varphi$ , if

- there exists a  $C \subset \Omega$  compact set such that for every  $n \in \mathbb{N}^+$   $\text{supp } \varphi, \text{supp } \varphi_n \subset C$ ,
- for every multi-index  $\alpha$ , we have that  $\partial^\alpha \varphi_n \rightrightarrows \partial^\alpha \varphi$  (uniformly).

$\mathcal{D}(\Omega)$  denotes  $C_0^\infty(\Omega)$  equipped with the above notion of convergence. Elements of  $\mathcal{D}(\Omega)$  are called test functions.

**Definition 1.1.43.** (Distributions)

A linear functional  $u: \mathcal{D}(\Omega) \rightarrow \mathbb{R}$  is called a distribution if it is continuous in the sense that  $\varphi_n \xrightarrow{\mathcal{D}} \varphi$  implies that  $u(\varphi_n) \xrightarrow{\mathbb{R}} u(\varphi)$ . The set of all distributions on  $\mathcal{D}(\Omega)$  is denoted by  $\mathcal{D}'(\Omega)$ .

**Theorem 1.1.44.** *The space  $\mathcal{D}'(\Omega)$  is a vector space with the operations*

- $(u + v)(\varphi) := u(\varphi) + v(\varphi)$ , and
- $(\lambda u)(\varphi) := u(\lambda\varphi)$ . See [2] and [5].

Every locally integrable function  $f \in L_{\text{loc}}^1(\Omega)$  can be associated with a distribution  $T_f$ .

**Definition 1.1.45.** (Regular distributions)

Let  $f \in L_{\text{loc}}^1(\Omega)$ , and consider the linear map  $T_f: \mathcal{D}(\Omega) \rightarrow \mathbb{R}$  defined as

$$T_f(\varphi) := \int_{\Omega} f\varphi. \quad (1.26)$$

The map  $T_f$  is indeed a distribution [2], and is called the regular distribution associated with the locally integrable function  $f$ .

This association between locally integrable functions and regular distributions is one-to-one.

**Theorem 1.1.46.** *Let  $f, g \in L_{\text{loc}}^1(\Omega)$  and suppose that the associated regular distributions satisfy  $T_f = T_g$ . Then  $f = g$  in  $L^1(\Omega)$ . See [2] for the proof.*

Therefore,  $L_{\text{loc}}^1(\Omega) \subset \mathcal{D}'(\Omega)$  in embedding sense. Furthermore, Remark 1.1.33 implies that for every  $p \geq 1$ ,  $L_{\text{loc}}^p(\Omega) \subset \mathcal{D}'(\Omega)$  in embedding sense.

**Definition 1.1.47.** (Dirac delta distribution)

Let  $\mathbf{a} \in \mathbb{R}^n$ . Then the distribution  $\delta_{\mathbf{a}}: \mathcal{D}(\mathbb{R}^n) \rightarrow \mathbb{R}$  defined by  $\delta_{\mathbf{a}}(\varphi) := \varphi(\mathbf{a})$  is called the Dirac delta distribution centered at  $\mathbf{a}$ .

It is true that if  $\varphi_n \xrightarrow{\mathcal{D}} \varphi$  then  $\delta_{\mathbf{a}}(\varphi_n) \xrightarrow{\mathbb{R}} \delta_{\mathbf{a}}(\varphi)$ , i.e.,  $\delta_{\mathbf{a}} \in \mathcal{D}'(\mathbb{R}^n)$ . Furthermore, Example 1.1.37 shows that the Dirac delta distribution is not a regular distribution.

To define PDEs among distributions, one must first define the partial derivatives of distributions. Furthermore, to define the fundamental solution of differential operators – which is a requirement to talk about the MFS – it is necessary to define distributions' direct product and convolution. The working hypothesis of these definitions is that one looks at the analogous results among regular distributions and invokes the principle of permanence to define analogous concepts among all distributions.

The following is the generalization to distributions of the iterated application of (1.18).

**Definition 1.1.48.** (Partial derivatives of distributions)

Let  $u \in \mathcal{D}'(\Omega)$ . Then, for every multi-index  $\alpha$ , we define the partial derivative  $\partial^\alpha u$  as

$$(\partial^\alpha u)(\varphi) := (-1)^{|\alpha|} u(\partial^\alpha \varphi) \quad (\varphi \in \mathcal{D}(\Omega)). \quad (1.27)$$

**Definition 1.1.49.** (Direct product of distributions)

Let  $u \in \mathcal{D}'(\mathbb{R}^n)$ ,  $v \in \mathcal{D}'(\mathbb{R}^m)$ . Then, we say that the direct product of  $u$  and  $v$  is the distribution

$$(u \times v)(\varphi) := u \left( \overbrace{\mathbf{x} \mapsto v \left( \underbrace{\mathbf{y} \mapsto \varphi(\mathbf{x}, \mathbf{y})}_{\in \mathcal{D}(\mathbb{R}^m)} \right)}^{\in \mathcal{D}(\mathbb{R}^n)} \right) \quad (\varphi \in \mathcal{D}(\mathbb{R}^n \times \mathbb{R}^m)). \quad (1.28)$$

Again, verifying that the above relationship holds for regular distributions is easy. Therefore, it is sensible to define the direct product in the general case. Of course, one must check that this is indeed a distribution on  $\mathcal{D}(\mathbb{R}^n \times \mathbb{R}^m)$ . See [2] for details.

The definition of convolution raises the following issue. For  $f, g \in L^1_{\text{loc}}(\mathbb{R}^n)$ ,

$$T_{f*g}(\varphi) = \int_{\mathbb{R}^n} (f * g)(\mathbf{x}) \varphi(\mathbf{x}) \, d\mathbf{x} = \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} f(\mathbf{y}) g(\mathbf{x} - \mathbf{y}) \, d\mathbf{y} \right) \varphi(\mathbf{x}) \, d\mathbf{x} = \quad (1.29)$$

$$= \int_{\mathbb{R}^n} f(\mathbf{y}) \underbrace{\int_{\mathbb{R}^n} g(\mathbf{x} - \mathbf{y}) \varphi(\mathbf{x}) \, d\mathbf{x}}_{\mathbf{x} \rightsquigarrow \mathbf{x} + \mathbf{y}} \, d\mathbf{y} = \int_{\mathbb{R}^{2n}} f(\mathbf{x}) g(\mathbf{y}) \varphi(\mathbf{x} + \mathbf{y}) \, d(\mathbf{x}, \mathbf{y}), \quad (1.30)$$

however, it is not true that if  $\varphi$  is compactly supported in  $\mathbb{R}^n$ , then  $\psi: (\mathbf{x}, \mathbf{y}) \mapsto \varphi(\mathbf{x} + \mathbf{y})$  is also compactly supported in  $\mathbb{R}^{2n}$ . For instance, if  $\text{supp}(\varphi) = B(\mathbf{0}, 1)$ , then the support of  $\psi$  is an unbounded subset (infinite cylinder) in  $\mathbb{R}^{2n}$ :  $\text{supp}(\psi) = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{2n} \mid \|\mathbf{x} + \mathbf{y}\| \leq 1\}$ . Nonetheless, this partial result shows us why we define the direct product in the first place, and that thinking about  $(u * v)(\varphi)$  as  $(u \times v)(\psi)$  with  $\psi$  as defined above gets us into the ballpark of understanding the more technical definition of convolutions. The rigorous resolution to this problem is to multiply  $\psi$  by a sequence of compactly supported functions  $(h_k)$  that tend to the function 1 in some very strong convergence sense – enough to ensure that the convolution is also a distribution – and then to take the limit as  $k \rightarrow +\infty$ . See [2] for the technical details.

**Definition 1.1.50.** (Convolution of distributions)

Let  $(h_k)$  be a sequence of functions in  $C_0^\infty(\mathbb{R}^{2n})$  such that for every  $(2n)$  long multi-index  $\alpha$

- $\partial^\alpha (1 - h_k(\mathbf{x}, \mathbf{y})) \rightrightarrows 0$  uniformly on every compact subset of  $\mathbb{R}^{2n}$ ,
- there exists a constant  $C_\alpha$  such that  $\sup_{k \in \mathbb{N}_0} \sup_{\mathbb{R}^{2n}} |\partial^\alpha h_k| \leq C_\alpha$ .

Suppose that  $u, v \in \mathcal{D}'(\mathbb{R}^n)$  are such that the following limit exists, is finite and continuous for every  $\varphi \in \mathcal{D}(\mathbb{R}^n)$

$$\lim_{k \rightarrow +\infty} (u \times v)(h_k \psi), \quad (1.31)$$

where  $\psi(\mathbf{x}, \mathbf{y}) := \varphi(\mathbf{x} + \mathbf{y})$ . Then we define this value to be  $(u * v)(\varphi)$ .

**Theorem 1.1.51.** *Let  $u, v, w \in \mathcal{D}'(\mathbb{R}^n)$  distributions,  $\alpha, \beta \in \mathbb{R}$  scalars,  $\boldsymbol{\alpha} \in \mathbb{R}^n$  a multi-index and  $\delta_{\mathbf{0}}$  be the Dirac delta concentrated to  $\mathbf{0}$ . Then the following properties hold (see [2] for details)*

- (Commutativity:)  $u * v = v * u$ ,
- (Identity:)  $\delta_{\mathbf{0}} * u = u * \delta_{\mathbf{0}} = u$ ,
- (Linearity:)  $(\alpha u + \beta v) * w = \alpha u * w + \beta v * w$ .
- (Relationship with partial derivatives:)  $\partial^{\boldsymbol{\alpha}}(u * v) = (\partial^{\boldsymbol{\alpha}} u) * v = u * (\partial^{\boldsymbol{\alpha}} v)$ . Indeed, for any linear partial differential operator of finite order  $L: \mathcal{D}'(\mathbb{R}^n) \rightarrow \mathcal{D}'(\mathbb{R}^n)$ , we have that

$$L(u * v) = (Lu) * v = u * (Lv) \quad (1.32)$$

However, the convolution of distributions is non-associative.

**Example 1.1.52.** (Convolution of distributions is non-associative)

Let  $u = T_H$ , the regular distribution belonging to the single-variable Heaviside function,  $v = \delta'_0$ , the derivative of the Dirac delta concentrated at  $x = 0$ , and  $w = T_1$ , the regular distribution belonging to the constant function 1 ( $u, v, w \in \mathcal{D}'(\mathbb{R})$ ). Then both  $(u * v) * w$  and  $u * (v * w)$  exist, but they are not equal to each other as

$$(T_H * \delta'_0) * T_1 = (T'_H * \delta_0) * T_1 = (\delta_0 * \delta_0) * T_1 = T_1, \text{ but} \quad (1.33)$$

$$T_H * (\delta'_0 * T_1) = T_H * (\delta_0 * T'_1) = T_H * (\delta_0 * T_0) = T_0. \quad (1.34)$$

Furthermore, it is important to emphasize that it is possible for a convolution not to exist.

**Example 1.1.53.** (Convolution of distributions does not always exist)

The convolution  $T_1 * T_1$  does not exist, because

$$(T_1 * T_1)(\varphi) = \int_{\mathbb{R}^{2n}} \varphi(\mathbf{x} + \mathbf{y}) d(\mathbf{x}, \mathbf{y}) \geq \int_{\mathbf{x} - \mathbf{y} \in S} p = p \cdot \text{Vol}_{2n}(\underbrace{\{\mathbf{x}, \mathbf{y} \mid \mathbf{x} - \mathbf{y} \in S\}}_{\text{infinite strip in } \mathbb{R}^{2n}}) = +\infty \quad (1.35)$$

for any  $\varphi \geq 0$  that takes a strictly positive value at least  $p$  on an open set  $S$ .

## 1.2 Prerequisites from PDE theory

In this section, we introduce fundamental solutions of linear partial differential operators and illustrate their use to obtain general solutions of PDEs.



### 1.2.1 Fundamental solutions and PDEs

The following theorem is invoked in the method of fundamental solutions.

**Theorem 1.2.1.** *Let  $L: \mathcal{D}'(\mathbb{R}^n) \rightarrow \mathcal{D}'(\mathbb{R}^n)$  be a linear differential operator of finite order and suppose that  $LE = \delta_{\mathbf{0}}$  for some  $E \in \mathcal{D}'(\mathbb{R}^n)$ . Furthermore, let  $F$  be a fixed distribution, and suppose that the convolution  $E * F$  exists. Then  $u = E * F$  is the unique solution of the equation  $Lu = F$ .*

*Proof.* The distribution  $u = E * F$  solves the PDE

$$Lu = L(E * F) = (LE) * F = \delta_{\mathbf{0}} * F = F. \quad (1.36)$$

For the uniqueness claim, suppose that  $Lu_1 = Lu_2 = F$  and let  $h := u_1 - u_2$ . By the linearity of  $L$ , we have that  $Lh = 0$ , and we want to show that  $h = 0$ . Indeed, by 1.32,

$$h = \delta_{\mathbf{0}} * h = (LE) * h = E * (Lh) = E * 0 = 0. \quad (1.37)$$

□

**Definition 1.2.2.** (Fundamental solutions)

Let  $L: \mathcal{D}'(\mathbb{R}^n) \rightarrow \mathcal{D}'(\mathbb{R}^n)$  be a linear differential operator. Solutions of the equation  $Lu = \delta_{\mathbf{0}}$  are called the fundamental solutions of the differential operator  $L$ .

**Example 1.2.3.** (Fundamental solution of the one dimensional wave operator by [2])

Consider the function

$$E(t, x) := \frac{1}{2}H(t - |x|), \quad (1.38)$$

where  $H$  is the Heaviside function, and the corresponding regular distribution  $T_E$ . Then the distribution  $T_E$  is the fundamental solution of the wave-equation  $Lu := \partial_t^2 u - \partial_x^2 u = \delta_{\mathbf{0}}$  in one-dimension.

To verify this, we use the fact that all partial derivatives of  $\varphi$  are sufficiently smooth, and that  $\text{supp}(\varphi)$  is bounded,

$$\begin{aligned} (LT_E)(\varphi) &= (\partial_t^2 T_E)(\varphi) - (\partial_x^2 T_E)(\varphi) = T_E(\partial_t^2 \varphi) - T_E(\partial_x^2 \varphi) = \\ &= \frac{1}{2} \int_{\mathbb{R}^2} \left( H(t - |x|) (\partial_t^2 \varphi(t, x) - \partial_x^2 \varphi(t, x)) \right) d(t, x) = \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} \int_{|x|}^{+\infty} (\partial_t^2 \varphi(t, x) - \partial_x^2 \varphi(t, x)) dt dx = \\ &= \frac{1}{2} \left( \int_{-\infty}^{+\infty} \int_{|x|}^{+\infty} \partial_t^2 \varphi(t, x) dt dx - \int_0^{+\infty} \int_{-t}^t \partial_x^2 \varphi(t, x) dx dt \right) = \\ &= \frac{1}{2} \left( \int_{-\infty}^{+\infty} \left[ \partial_t \varphi(t, x) \right]_{t=|x|}^{t \rightarrow +\infty} dx - \int_0^{+\infty} \left[ \partial_x \varphi(t, x) \right]_{x=-t}^{x=t} dt \right) = \\ &= -\frac{1}{2} \left( \int_{-\infty}^{+\infty} \partial_t \varphi(|x|, x) dx + \int_0^{+\infty} (\partial_x \varphi(t, t) - \partial_x \varphi(t, -t)) dt \right) = \\ &= -\frac{1}{2} \left( \int_0^{+\infty} (\partial_t \varphi(x, x) + \partial_t \varphi(x, -x)) dx + \int_0^{+\infty} (\partial_x \varphi(t, t) - \partial_x \varphi(t, -t)) dt \right) = \\ &= -\frac{1}{2} \left( \int_0^{+\infty} (\partial_t \varphi(t, t) + \partial_t \varphi(t, -t)) dt + \int_0^{+\infty} (\partial_x \varphi(x, x) - \partial_x \varphi(x, -x)) dx \right) = \\ &= -\frac{1}{2} \left( \left[ \varphi(t, t) + \varphi(t, -t) \right]_{t=0}^{t \rightarrow +\infty} + \left[ \varphi(x, x) - \varphi(x, -x) \right]_{x=0}^{x \rightarrow +\infty} \right) = \\ &= -\frac{1}{2} \left( (0 - \varphi(0, 0)) + (0 - \varphi(0, 0)) + (0 - \varphi(0, 0)) - (0 - \varphi(0, 0)) \right) = \varphi(0, 0) = \delta_{(0,0)}(\varphi), \end{aligned}$$

which is what we need. Similarly,  $H(ct - |x|)$  is the fundamental solution of the operator  $L = \partial_t^2 - c^2 \partial_x^2$ .

The “physical” content of the solution – as much as one may call a discontinuous wave physical – is that given an infinite string at rest with no outer forces, except an instantaneous hit of unit momentum at  $t = x = 0$ , will result in a “wave” propagation in both the positive and negative directions, resulting in a growing domain  $[-ct, +ct]$  where the displacement is now 1, instead of the initial 0, where  $c > 0$  is the speed of the wave propagation.

The concrete formula that Example 1.2.3 and Theorem 1.2.1 imply is as follows.

**Corollary 1.2.4.** *Suppose that  $f \in L^1_{loc}(\mathbb{R}^2)$  and consider the PDE*

$$\partial_t^2 u(t, x) - c^2 \partial_x^2 u(t, x) = f(t, x) \quad ((t, x) \in \mathbb{R}^2). \quad (1.39)$$

*Then the solution  $u(t, x)$  is of the form*

$$\begin{aligned} u(t, x) &= \int_{\mathbb{R}^2} E(t - \tau, x - \xi) f(\tau, \xi) \, d(\tau, \xi) + \psi_1(ct + x) + \psi_2(ct - x) = \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} H(c(t - \tau) - |\xi - x|) f(\tau, \xi) \, d\xi \, d\tau + \psi_1(ct + x) + \psi_2(ct - x) = \\ &= \frac{1}{2} \int_{-\infty}^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} f(\tau, \xi) \, d\xi \, d\tau + \psi_1(ct + x) + \psi_2(ct - x) \end{aligned}$$

for some  $\psi_1, \psi_2 \in C^2(\mathbb{R})$ .

## 1.2.2 Laplace’s equation and Poisson’s equation

Suppose that  $\Omega \subset \mathbb{R}^n$  is a non-empty, open set. Let  $u: \Omega \rightarrow \mathbb{R}$  denote the concentration or density of some material or physical quantity (such as heat) and describe the equilibrium distribution of the material in the presence of a source term  $f: \Omega \rightarrow \mathbb{R}$ . If we assume that

- the flux of this material  $\mathbf{F}$  is proportional to the gradient of  $u$ , and points in the opposite direction, i.e.,  $\mathbf{F}(\mathbf{x}) = -p(\mathbf{x})\nabla u(\mathbf{x})$  (Fick’s first law for diffusion, Fourier’s law of heat conduction or Ohm’s law of electrical conduction),
- in equilibrium, for every smooth, small domain  $V \subset \Omega$ , the overall flux through the boundary of  $V$  is equal to the integral of some source function  $f: \Omega \rightarrow \mathbb{R}$ , and that
- Gauß’s law holds,

then

$$\int_V f(\mathbf{x}) \, d\mathbf{x} = \int_{\partial V} \mathbf{F}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{S} = \int_{\partial V} -p(\mathbf{x})\nabla u(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{S} = \int_V -\operatorname{div}(p(\mathbf{x})\nabla u(\mathbf{x})) \, d\mathbf{x}.$$

Since this holds for every “sufficiently nice”  $V \subset \Omega$ , we may conclude that the functions must be equal, thus obtaining a general elliptic PDE

$$-\operatorname{div}(p(\mathbf{x})\nabla u(\mathbf{x})) = f(\mathbf{x}) \quad (\mathbf{x} \in \Omega) \quad (1.40)$$

The first condition in all three cases captures the idea that the material tends to diffuse from higher concentration domains towards lower concentration domains, and this proportionality is linear with respect to the concentration gradient. The dependence of  $p$  on  $\mathbf{x}$  captures that the proportionality constant may vary across the domain  $\Omega$ . This law may not be assumed when

there is a non-negligible fluid motion or the material is porous. The other two assumptions require the smoothness of  $u$ . See [5].

If  $p$  is constant, we may assume  $p \equiv 1$ . Then

$$-\operatorname{div}(\nabla u(\mathbf{x})) = -\sum_{i=1}^n \partial_{x_i}(\partial_{x_1}u(\mathbf{x}), \partial_{x_2}u(\mathbf{x}), \dots, \partial_{x_n}u(\mathbf{x})) = -\sum_{i=1}^n \partial_{x_i}^2 u(\mathbf{x}) := -\Delta u(\mathbf{x}). \quad (1.41)$$

*Laplace's equation* on the domain  $\Omega$  is the second order homogeneous linear PDE

$$\Delta u(\mathbf{x}) = 0 \quad (\mathbf{x} \in \Omega), \quad (1.42)$$

and *Poisson's equation* on the domain  $\Omega$  is the second order inhomogeneous linear PDE

$$-\Delta u(\mathbf{x}) = f(\mathbf{x}) \quad (\mathbf{x} \in \Omega), \quad (1.43)$$

where  $f: \Omega \rightarrow \mathbb{R}$  is given. If  $\partial\Omega$  is not empty – i.e.,  $\Omega \neq \mathbb{R}^n$  and  $\Omega \neq \emptyset$  – then there are usually extra boundary conditions imposed as well.

*Remark 1.2.5.* Both Laplace's equation and Poisson's equation are the  $\partial_t u = 0$  (stationary) special cases of the more general forced heat (diffusion) equation in homogenous material

$$\partial_t u(t, \mathbf{x}) - \alpha \Delta_{\mathbf{x}} u(t, \mathbf{x}) = F(t, \mathbf{x}), \quad (1.44)$$

where  $F: \mathbb{R} \times \Omega \rightarrow \mathbb{R}$  is a given source,  $\alpha$  is the (constant) diffusivity of the material, and  $\Delta_{\mathbf{x}}$  means the Laplacian is taken only in the spatial variables. In other words, Poisson's equation describes the equilibrium heat distribution in a material given some sources of heat,  $F$ , (which may be negative), and Laplace's equation is a special case when there are no sources in this interpretation.

The following example is another application of Laplace's equation and Poisson's that we shall see later.

**Example 1.2.6.** In electrostatics (when there are no moving charges), we assume that the curl of the electric field  $\mathbf{E}$  is zero. It is assumed that  $\mathbf{E}$  is the gradient of some scalar field  $\phi$  (electric potential field) since (by Young's theorem)  $\operatorname{curl}(\nabla\phi) = 0$  for sufficiently a smooth scalar field  $\phi$ . Then substituting  $\mathbf{E} = \nabla\phi$  into the electrostatic Gauß's law  $\operatorname{div}(\mathbf{E}) = -\frac{\rho}{\varepsilon}$  yields Poisson's equation for electrostatics, which is

$$-\Delta\phi(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\varepsilon(\mathbf{x})}. \quad (1.45)$$

If the total volume charge density  $\rho = 0$  – which may be assumed for instance, in metals –, then we get Laplace's equation for electrostatics.

**Example 1.2.7.** (Fundamental solution of the two-dimensional Laplace operator)

Consider the two-dimensional PDE below

$$\Delta u(x, y) = 0 \quad ((x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}) \quad (1.46)$$

where  $\Delta := \partial_x^2 + \partial_y^2$ . It will become clear why we exclude the origin soon. Let us find solutions of this PDE. The first thing to notice is that if  $u$  is a solution, and  $R$  is an orthogonal matrix (rotation), i.e.,  $R^T = R^{-1}$ , then  $v(\mathbf{x}) = u(R\mathbf{x})$  is also a solution. Indeed, setting  $\mathbf{y} = R\mathbf{x}$ , we have that  $\nabla_{\mathbf{y}} = \nabla_{\mathbf{x}} R^{-1}$ , and therefore, using a bit of operator algebra, we have

$$\Delta_{\mathbf{y}} = \nabla_{\mathbf{y}} \cdot \nabla_{\mathbf{y}}^T = (\nabla_{\mathbf{x}} R^{-1}) \cdot (\nabla_{\mathbf{x}} R^{-1})^T = (\nabla_{\mathbf{x}} R^{-1}) \cdot (R \nabla_{\mathbf{x}}^T) = \nabla_{\mathbf{x}} (R^{-1} R) \cdot \nabla_{\mathbf{x}}^T = \nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}}^T = \Delta_{\mathbf{x}}.$$

Therefore, it makes sense to look for radially symmetric solutions. Set  $r := \sqrt{x^2 + y^2}$  and let us calculate  $\Delta v(r)$

$$\begin{aligned}\Delta v(\sqrt{x^2 + y^2}) &= \partial_x^2 \left( v(\sqrt{x^2 + y^2}) \right) + \partial_y^2 \left( v(\sqrt{x^2 + y^2}) \right) = \\ &= \partial_x \left( \frac{xv'(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}} \right) + \partial_y \left( \frac{yv'(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}} \right) = \\ &= \frac{x^2 v''(r)}{r^2} + \frac{y^2 v'(r)}{r^3} + \frac{y^2 v''(r)}{r^2} + \frac{x^2 v'(r)}{r^3} = v''(r) + \frac{1}{r} v'(r).\end{aligned}$$

Solving the ordinary differential equation (ODE)  $v''(r) + \frac{1}{r}v'(r) = 0$  is fairly easy. Introducing the new variable  $w = v'$ , we get the first order linear ODE  $w'(r) + \frac{1}{r}w(r) = 0$ , and hence

$$\begin{aligned}w'(r) + \frac{1}{r}w(r) &= 0 \\ \frac{w'(r)}{w(r)} &= -\frac{1}{r} \\ \ln |w(r)| &= -\ln(r) + c_0 \\ w(r) &= \frac{c_1}{r} \\ v(r) &= C \ln(r) + D\end{aligned}$$

All the radially symmetric solutions of Laplace's equation in two dimensions are of the form  $C \ln \sqrt{x^2 + y^2} + D$  for some  $C, D \in \mathbb{R}$ . These are not the only solutions, however. For example, the polynomials  $\Re((x + y \cdot i)^n)$  and  $\Im((x + y \cdot i)^n)$  – the so-called harmonic polynomials – also satisfy (1.46).

The fundamental solution of the Laplace operator is the regular distribution belonging to one of the radially symmetric solutions. To simplify the already quite complicated calculation, we will only consider  $E_2(x, y) = \frac{1}{2\pi} \ln(\sqrt{x^2 + y^2})$  the pair of parameters that yields the fundamental solution. Of course, one could carry out a similar but more complicated calculation with these parameters and find that  $\Delta T_{E_2} = 2C\pi\delta_0 + z$  for some  $z \in \mathcal{D}'(\mathbb{R}^2)$ .

$$\begin{aligned}(\Delta T_{E_2})(\varphi) &= \partial_x^2 T_{E_2}(\varphi) + \partial_y^2 T_{E_2}(\varphi) = (-1)^2 T_{E_2}(\partial_x^2 \varphi) + (-1)^2 T_{E_2}(\partial_y^2 \varphi) = T_{E_2}(\Delta \varphi) = \\ &= \frac{1}{2\pi} \int_{\mathbb{R}^2} \ln(\sqrt{x^2 + y^2}) \Delta \varphi(x, y) \, d(x, y)\end{aligned}$$

We calculate this integral as a limit of integrals defined on compact sets. Consider the sequence of domains  $\Omega_n := \{(x, y) \in \mathbb{R}^2 \mid \frac{1}{n} \leq \sqrt{x^2 + y^2} \leq n\}$  and the sequence of integral values

$$I_n := \int_{\Omega_n} \ln(\sqrt{x^2 + y^2}) \Delta \varphi(x, y) \, d(x, y).$$

By the *symmetric Green's formula* (1.9) and the fact that  $\Delta \ln \sqrt{x^2 + y^2} = 0$  on  $\Omega_n$ ,

$$\begin{aligned}I_n &= \int_{\Omega_n} \left( \ln(\sqrt{x^2 + y^2}) \Delta \varphi(x, y) - \underbrace{\varphi(x, y) \Delta \ln(\sqrt{x^2 + y^2})}_{\equiv 0 \text{ on } \Omega_n} \right) d(x, y) = \\ &= \int_{S(\mathbf{0}, n) \cup S(\mathbf{0}, \frac{1}{n})} \left( \ln(\sqrt{x^2 + y^2}) \partial_{\mathbf{n}} \varphi(x, y) - \varphi(x, y) \partial_{\mathbf{n}} \ln(\sqrt{x^2 + y^2}) \right) ds\end{aligned}$$

Suppose that  $N$  is large enough, such that for every  $n > N$  we have  $\text{supp}(\varphi) \subsetneq B(\mathbf{0}, n)$ . Such an  $N$  exists because  $\varphi$  is compactly supported. Then for every  $n > N$

$$\begin{aligned} I_n &= \int_{S(\mathbf{0}, \frac{1}{n})} \left( \ln(\sqrt{x^2 + y^2}) \partial_{\mathbf{n}} \varphi(x, y) - \varphi(x, y) \partial_{\mathbf{n}} \ln(\sqrt{x^2 + y^2}) \right) ds = \\ &= \int_{S(\mathbf{0}, \frac{1}{n})} \left( \ln(\sqrt{x^2 + y^2}) \nabla \varphi(x, y) \cdot \mathbf{n}(x, y) - \varphi(x, y) \nabla(\ln(\sqrt{x^2 + y^2})) \cdot \mathbf{n}(x, y) \right) ds. \end{aligned}$$

A suitable choice for  $\mathbf{n}$  is  $\mathbf{n}(x, y) = -\frac{(x, y)}{|(x, y)|} = \left(\frac{-x}{r}, \frac{-y}{r}\right) = (-nx, -ny)$ . Using the fact that if  $(x, y) \in S(\mathbf{0}, \frac{1}{n})$ , then  $\sqrt{x^2 + y^2} = r = \frac{1}{n}$  and  $\nabla(\ln(\sqrt{x^2 + y^2})) = \left(\frac{x}{r^2}, \frac{y}{r^2}\right) = (n^2x, n^2y)$ , we have

$$\begin{aligned} I_n &= \int_{S(\mathbf{0}, \frac{1}{n})} \left( \ln\left(\frac{1}{n}\right) \nabla \varphi(x, y) \cdot (-nx, -ny) - \varphi(x, y) (n^2x, n^2y) \cdot (-nx, -ny) \right) ds = \\ &= \int_{S(\mathbf{0}, \frac{1}{n})} \left( \ln(n) \nabla \varphi(x, y) \cdot (nx, ny) + \varphi(x, y) \underbrace{(n^2x, n^2y) \cdot (nx, ny)}_{=n} \right) ds = \\ &= \ln(n) \int_{S(\mathbf{0}, \frac{1}{n})} \nabla \varphi(x, y) \cdot (nx, ny) ds + \int_{S(\mathbf{0}, \frac{1}{n})} n \varphi(x, y) ds. \end{aligned}$$

Here, the first term tends to 0 as  $n \rightarrow +\infty$

$$\begin{aligned} \left| \ln(n) \int_{S(\mathbf{0}, \frac{1}{n})} \nabla \varphi(x, y) \cdot (nx, ny) ds \right| &\leq \ln(n) \int_{S(\mathbf{0}, \frac{1}{n})} \left| \nabla \varphi(x, y) \cdot (nx, ny) \right| ds \leq \\ &\leq \ln(n) \int_{S(\mathbf{0}, \frac{1}{n})} |\nabla \varphi(x, y)| ds \leq \ln(n) \|\nabla \varphi\|_{\max} \int_{S(\mathbf{0}, \frac{1}{n})} 1 ds = 2\pi \|\nabla \varphi\|_{\max} \frac{\ln(n)}{n} \rightarrow 0 \end{aligned}$$

and we can deal with the second integral using polar substitution

$$\int_{S(\mathbf{0}, \frac{1}{n})} n \varphi(x, y) ds = \int_0^{2\pi} n \varphi\left(\frac{1}{n} \cos \theta, \frac{1}{n} \sin \theta\right) \frac{1}{n} d\theta = \int_0^{2\pi} \varphi\left(\frac{1}{n} \cos \theta, \frac{1}{n} \sin \theta\right) d\theta.$$

The function  $\varphi$  is continuous, and therefore, it is a suitable integrable majorant in Lebesgue's dominated convergence theorem

$$\begin{aligned} \lim_{n \rightarrow +\infty} \int_0^{2\pi} \varphi\left(\frac{1}{n} \cos \theta, \frac{1}{n} \sin \theta\right) d\theta &= \int_0^{2\pi} \lim_{n \rightarrow +\infty} \varphi\left(\frac{1}{n} \cos \theta, \frac{1}{n} \sin \theta\right) d\theta = \\ &= \int_0^{2\pi} \varphi(0, 0) d\theta = 2\pi \varphi(0, 0), \end{aligned}$$

which finally yields

$$(\Delta T_{E_2})(\varphi) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \ln(\sqrt{x^2 + y^2}) \Delta \varphi(x, y) d(x, y) = \frac{1}{2\pi} \lim_{n \rightarrow +\infty} I_n = \frac{1}{2\pi} 2\pi \varphi(0, 0) = \varphi(0, 0).$$

**Corollary 1.2.8.** *Let  $f \in L^1_{\text{loc}}(\mathbb{R}^2)$  and consider Poisson's equation*

$$-\Delta u(x, y) = f(x, y) \quad ((x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}) \quad (1.47)$$

*Then by Theorem 1.2.1, the solution  $u(x, y)$  is of the form*

$$\begin{aligned} u(x, y) &= - \int_{\mathbb{R}^2} E_2(x - \xi, y - \eta) f(\xi, \eta) d(\xi, \eta) + u_0(x, y) = \\ &= - \frac{1}{2\pi} \int_{\mathbb{R}^2} \ln(\sqrt{(x - \xi)^2 + (y - \eta)^2}) f(\xi, \eta) d(\xi, \eta) + u_0(x, y), \end{aligned}$$

*where  $u_0$  is an arbitrary harmonic function.*

*Remark 1.2.9.* Likewise, one can show with similar calculations that the fundamental solution of the Laplace operator in one dimension (second derivative operator) is the regular distribution that belongs to the locally integrable function

$$E_1(x) = \frac{1}{2}|x|.$$

In three dimensions, it is the regular distribution that belongs to

$$E_3(x, y, z) = \frac{-1}{4\pi\sqrt{x^2 + y^2 + z^2}}, \quad (1.48)$$

and in general, the regular distribution that belongs to

$$E_n(\mathbf{x}) = \frac{-1}{(n-2)S_{n-1}\|\mathbf{x}\|^{n-2}} \quad (n \geq 3), \quad (1.49)$$

where  $S_{n-1}$  is the  $(n-1)$ -dimensional surface measure of the  $n$ -dimensional unit ball.

Possible physical interpretations of the fundamental solution of Laplace's equation in  $n$  dimensions include the electric potential generated by a point charge if the field obeys Gauß's law or the gravitational potential generated by a point particle of unit mass in  $n$ -dimensional space.

Notice these potentials  $E_n$  are bounded from above if and only if  $n \geq 3$ . This suggests that a finite amount of energy can overcome the electric potential of a unit charge or the gravitational potential of a unit mass if and only if  $n \geq 3$ , meaning that ionization of atoms or space travel would not be possible if we lived in a less than three-dimensional space.

## 1.3 Prerequisites from functional analysis

### 1.3.1 Single variable Sobolev spaces

Here we give a brief description of single-variable Sobolev spaces. The reason for taking this excursion is that single-variable Sobolev spaces may be described in a more constructive fashion than their multi-variable counterparts, which are usually defined as certain special subsets of  $\mathcal{D}'(\Omega)$  and  $\mathcal{D}'(\partial\Omega)$ .

**Definition 1.3.1.** ( $W^{1,p}([a, b])$  Sobolev space)

Let  $1 \leq p \leq +\infty$ . We say that  $f \in L^p([a, b])$  belongs to the  $W^{1,p}([a, b])$  Sobolev space if the following conditions hold:

- $f$  is differentiable almost everywhere,
- $f' \in L^p([a, b])$ , and
- $\int_x^y f'(t) dt = f(y) - f(x)$  for almost every  $x, y \in [a, b]$ , i.e., the fundamental theorem of calculus holds.

Furthermore, we define  $\|f\|_{W^{1,p}(\Omega)}$  as

$$\|f\|_{W^{1,p}} := \begin{cases} (\|f\|_{L^p}^p + \|f'\|_{L^p}^p)^{\frac{1}{p}} = \left( \int_{\Omega} (|f|^p + |f'|^p) \right)^{\frac{1}{p}} & \text{if } p < +\infty, \\ \max \{ \|f\|_{L^\infty}, \|f'\|_{L^\infty} \} = \max \left\{ \operatorname{ess\,sup}_{\Omega} |f|, \operatorname{ess\,sup}_{\Omega} |f'| \right\} & \text{if } p = +\infty. \end{cases}$$

*Remark 1.3.2.* The idea behind the definition for  $p = +\infty$  is that  $\lim_{p \rightarrow +\infty} (x^p + y^p)^{\frac{1}{p}} = \max\{x, y\}$ .

*Remark 1.3.3.* In the special case  $p = 1$ , we get that  $f \in W^{1,1}([a, b])$  is equivalent to  $f$  being absolutely continuous. Therefore, all Lipschitz-continuous functions are also elements of  $W^{1,1}$ .

**Theorem 1.3.4.** *The above-defined spaces  $W^{1,p}([a, b])$  are Banach spaces (complete normed vector spaces) with respect to  $\|\cdot\|_{W^{1,p}}$  for every  $1 \leq p \leq +\infty$ . See [5] for details.*

*Remark 1.3.5.* We define higher order single variable Sobolev spaces  $W^{k,p}([a, b])$  analogously. There,  $f \in C^{k-1}([a, b])$ ,  $k$  times differentiable almost everywhere,  $f^{(k)} \in L^p([a, b])$ , and the fundamental theorem of calculus  $\int_x^y f^{(k)}(t) dt = f^{(k-1)}(y) - f^{(k-1)}(x)$  must hold for almost every  $x, y \in [a, b]$ . The norm is either the sum of the  $L^p$  norms of the derivatives ( $p < +\infty$ ), or the maximum of the  $L^\infty$  norms ( $p = +\infty$ ).

*Remark 1.3.6.* A simpler notation in the case of  $p = 2$  is  $H^k := W^{k,2}$ , where the letter  $H$  is suggestive of the fact that  $H^k$  are Hilbert spaces. [5]

The significance of these spaces is that we may state that a function satisfies a boundary value problem or a Cauchy problem with fewer restrictions.

**Example 1.3.7.** The function  $u(x) = x + H(x)x$  is an element of  $H^1([-1, 1])$ , where  $H$  is the Heaviside step function. Indeed,  $u$  is differentiable everywhere, except at  $x = 0$ , and its derivative is  $1 + H(x)$ , which is an element of  $L^2([-1, 1])$ . Therefore,  $u$  is a solution of the one-dimensional boundary value problem

$$\left. \begin{aligned} u'(x) &= 1 + H(x), \text{ if } x \in (-1, 1), \\ u(-1) &= -1, \\ u(1) &= 2. \end{aligned} \right\}$$

**Example 1.3.8.** Consider the following non-linear ODE with initial conditions

$$\left. \begin{aligned} -m\ddot{x}(t) &= kx(t) + |F_s| \operatorname{sgn}(\dot{x}(t)) \quad (0 < t < T), \\ x(0) &= x_0, \\ \dot{x}(0) &= v_0, \end{aligned} \right\}$$

which describes the displacement  $x$  of a body of mass  $m$  from its equilibrium point, suspended on a spring with spring constant  $k$ , while the motion is damped by slipping friction  $F_s$ . Solutions of this initial value problem are not twice continuously differentiable at points where  $\dot{x}$  changes sign. That is, the solution  $x \in H^2([0, T]) \setminus C^2([0, T])$ .

## 1.3.2 Multivariable Sobolev spaces

Assume that  $\Omega \subset \mathbb{R}^n$  is a non-empty bounded open set with a piecewise smooth boundary. The following definitions include distributions, and when we mention concrete functions, we mean the associated regular distributions.

**Definition 1.3.9.** ( $W^{k,p}(\Omega)$  Sobolev space)

Let  $1 \leq p \leq +\infty$ ,  $k \in \mathbb{N}^+$ . The Sobolev space  $W^{k,p}(\Omega)$  consists of all distributions  $u \in \mathcal{D}'(\Omega)$  for which every partial derivative  $\partial^\alpha u$  of at most order  $k$  exists in distribution sense and belongs to  $L^p(\Omega)$ . The norms defined on these spaces are as follows. For  $p < +\infty$

$$\|u\|_{W^{k,p}(\Omega)} := \sum_{|\alpha| \leq k} \|\partial^\alpha u\|_{L^p(\Omega)} = \sum_{|\alpha| \leq k} \left( \int_{\Omega} |\partial^\alpha u|^p \right)^{\frac{1}{p}}$$

and for  $p = +\infty$

$$\|u\|_{W^{k,\infty}(\Omega)} := \max_{|\alpha| \leq k} \{ \|\partial^\alpha u\|_{L^\infty(\Omega)} \} = \max_{|\alpha| \leq k} \left\{ \operatorname{ess\,sup}_\Omega |\partial^\alpha u| \right\}$$

**Theorem 1.3.10.** *For every  $1 \leq p \leq +\infty$ ,  $k \in \mathbb{N}^+$ ,  $W^{k,p}(\Omega)$  is a Banach space. See [5].*

*Remark 1.3.11.* Similarly to the single-variable case,  $H^k(\Omega) := W^{k,2}(\Omega)$ .

Defining a boundary value problem among elements of  $L^p$  spaces is problematic for the following reason. Remark 1.1.11 shows that the boundary  $\partial\Omega$  is a null set with respect to the  $n$ -dimensional Lebesgue measure. Therefore, we may arbitrarily redefine any element  $u \in L^p$  on  $\partial\Omega$ . Hence, we need to introduce a generalization of the restriction operator such that it makes sense to talk about the ‘‘restriction’’ of an element of  $L^p(\Omega)$  to  $\partial\Omega$ .

The idea of this approach is that we approximate an element  $u \in W^{1,p}(\Omega)$  by a sequence of smooth functions  $u_n \xrightarrow{W^{1,p}} u$  from  $W^{1,p}(\Omega) \cap C(\bar{\Omega})$ , which have a sensible notion of a restriction. This is possible because  $C(\bar{\Omega})$  is a dense subspace of  $L^p_{\text{loc}}(\bar{\Omega})$ , and  $W^{1,p}(\Omega) \subset L^p_{\text{loc}}(\bar{\Omega})$ . Then we look at the limit of the functions  $u_n|_{\partial\Omega}$  in  $L^p(\partial\Omega)$ , which exists and thus may serve as a definition of the restriction. The following theorem formalizes this process. See [5] for details.

**Theorem 1.3.12.** *(Trace theorem)*

*There exists a unique bounded linear operator*

$$\operatorname{tr}: W^{1,p}(\Omega) \rightarrow L^p(\partial\Omega),$$

*such that for every  $u \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$*

$$\operatorname{tr}(u) = u|_{\partial\Omega}.$$

*Here, boundedness is understood in the sense that there exists a constant  $C_{p,\Omega}$  such that for every  $u \in W^{1,p}(\Omega)$ , we have that  $\|\operatorname{tr}(u)\|_{L^p(\partial\Omega)} \leq C_{p,\Omega} \|u\|_{W^{1,p}(\Omega)}$ . See [5].*

This is an extension of the restriction operator, so we keep writing  $u|_{\partial\Omega}$  instead of  $\operatorname{tr}(u)$ .

Finally, we define two more spaces of fractional index and their dual spaces that play a role in weak boundary value problems and thus in the upcoming existence and uniqueness theorems. These are definitions of fractional index Sobolev spaces that are specific to their application in boundary value problems. See [4] for details.

**Definition 1.3.13.**  $W_0^{k,p}(\Omega)$  denotes the subspace of  $W^{k,p}(\Omega)$ , where the trace of all the at most  $(k-1)$ <sup>st</sup> order partial derivatives of the elements vanish, that is

$$W_0^{k,p}(\Omega) := \{u \in W^{k,p}(\Omega) \mid \forall \alpha : |\alpha| \leq k-1, \partial^\alpha u|_{\partial\Omega} = 0 \text{ a.e.}\}.$$

*Remark 1.3.14.* One may obtain another equivalent norm on this space by only including  $|\alpha| = k$  order partial derivatives. See [17].

**Definition 1.3.15.**  $(H^{-1}(\Omega)$  Sobolev space)

Let  $H^{-1}(\Omega)$  be the dual space of  $H_0^1(\Omega)$ , i.e.,

$$H^{-1}(\Omega) := \{ \phi : H_0^1(\Omega) \rightarrow \mathbb{R} \mid \phi \text{ is a bounded linear functional} \},$$

where boundedness is understood in the sense that there exists a bound  $B > 0$  such that for every  $u \in H_0^1(\Omega)$ ,  $|\phi u| := |\langle \phi, u \rangle| \leq B \|u\|_{H_0^1(\Omega)}$ . The associated norm is

$$\|\phi\|_{H^{-1}(\Omega)} := \sup \{ \langle \phi, u \rangle \mid u \in H_0^1(\Omega), \|u\|_{H_0^1(\Omega)} = 1 \}$$



**Definition 1.3.16.** ( $H^{\frac{1}{2}}(\partial\Omega)$  Sobolev space)

We say that  $H^{\frac{1}{2}}(\partial\Omega)$  is the trace space of  $H^1(\Omega)$ , that is

$$H^{\frac{1}{2}}(\partial\Omega) := \{g \in L^2(\partial\Omega) \mid \text{for which } \exists u \in H^1(\Omega) \text{ such that } u|_{\partial\Omega} = g\}.$$

The norm defined on this space is

$$\|g\|_{H^{\frac{1}{2}}(\partial\Omega)} := \inf \{ \|u\|_{H^1(\Omega)} \mid u \in H^1(\Omega), u|_{\partial\Omega} = g \}.$$

**Definition 1.3.17.** ( $H^{\frac{3}{2}}(\partial\Omega)$  Sobolev space)

We say that  $H^{\frac{3}{2}}(\partial\Omega)$  is the trace space of  $H^2(\Omega)$ , that is

$$H^{\frac{3}{2}}(\partial\Omega) := \{g \in L^2(\partial\Omega) \mid \text{for which } \exists u \in H^2(\Omega) \text{ such that } u|_{\partial\Omega} = g\}.$$

The norm defined on this space is

$$\|g\|_{H^{\frac{3}{2}}(\partial\Omega)} := \inf \{ \|u\|_{H^2(\Omega)} \mid u \in H^2(\Omega), u|_{\partial\Omega} = g \}.$$

**Definition 1.3.18.** ( $H^{-\frac{1}{2}}(\partial\Omega)$  and  $H^{-\frac{3}{2}}(\partial\Omega)$  Sobolev spaces)

Analogously to the  $H^{-1}(\Omega)$  case, the  $H^{-\frac{1}{2}}(\partial\Omega)$  and  $H^{-\frac{3}{2}}(\partial\Omega)$  Sobolev spaces may be defined as the dual spaces of  $H^{\frac{1}{2}}(\partial\Omega)$  and  $H^{\frac{3}{2}}(\partial\Omega)$  respectively. The associated norms are

$$\|\phi\|_{H^{-k}(\Omega)} := \sup \{ \langle \phi, u \rangle \mid u \in H_0^k(\Omega), \|u\|_{H_0^k(\Omega)} = 1 \} \quad \left( k \in \left\{ \frac{1}{2}, \frac{3}{2} \right\} \right)$$

*Remark 1.3.19.* Elements of  $H^{-k}(\Omega)$  spaces ( $k \in \mathbb{N}^+$ ) may also be understood as distributions who are  $k^{\text{th}}$  order partial derivatives of some function in  $L^2(\Omega)$  in the distribution sense.

*Remark 1.3.20.* In general,  $W^{-k,p}(\Omega)$  is defined as the dual space of  $W_0^{k,q}(\Omega)$ , where  $q$  is the conjugate index of  $p$ , that is  $\frac{1}{p} + \frac{1}{q} = 1$  or one of them is 1 and the other is  $\infty$ . See [4] for details.

**Example 1.3.21.** Some of the most important elements in  $W^{1,p}(\Omega)$  – and in particular,  $H^1(\Omega) = W^{1,2}(\Omega)$  – are piecewise linear functions. Let  $\Omega$  be a polytope (polygon in  $n = 2$ , polyhedron in  $n = 3$  dimensions), and  $T_j$  be simplexes (triangles in  $n = 2$ , tetrahedra in  $n = 3$ ) such that  $T_j$  are almost disjoint (they may meet in vertices, edges, faces, etc.), and  $\bigcup_{j=1}^N T_j = \Omega$ . Then, if  $u|_{T_j}$  is linear for every index  $j$ , and  $u \in C(\Omega)$ , then  $h \in H^1(\Omega)$ . Indeed, all the partial derivatives of  $u$  are (perhaps different) constants on all the simplexes  $T_j$ , and piecewise constant functions are elements of  $L^p(\Omega)$  for every  $1 \leq p \leq +\infty$ . These functions play a key role in the theory of finite element methods [4]. See two examples in Figure 1.3 below, where  $D_{12}$  denotes the regular 12-gon of unit radius centered at the origin.

These functions, however, are not elements of  $H^2(\Omega)$  because their second partial derivatives along the edges usually do not belong to  $L^2(\Omega)$ .

**Example 1.3.22.** Analogously to the previous example, some of the most important elements in  $H^2(\Omega) = W^{2,2}(\Omega)$  are functions that are piecewise quadratic polynomials that continuously differentiably join across component domains. The simplest case of this example is quadratic splines in one dimension.

*Remark 1.3.23.* The domain  $\Omega$  need not be a polytope in either case, only the triangulation is easier if  $\Omega$  is a polytope.

The space  $H^{-1}(\Omega)$  may be fully characterized in line with Remark 1.3.19. That is, for every  $\phi \in H^{-1}(\Omega)$  we may find a function in  $L^2(\Omega)$  that we can think of as the “antiderivative” of  $\phi$ .

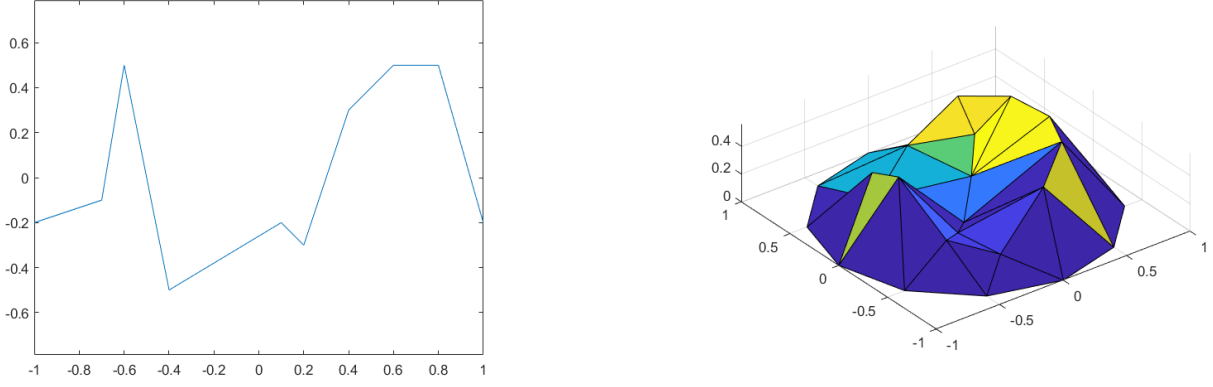


Figure 1.3: A function from  $H^1((-1, 1))$  and another from  $H_0^1(D_{12})$ .

**Theorem 1.3.24.** (Characterization of the space  $H^{-1}(\Omega)$ )

Suppose that  $\phi \in H^{-1}(\Omega)$ . Then there exists a finite set of functions  $\{\phi_j\}_{j=0}^J \subset L^2(\Omega)$  such that the following two properties hold:

1. for every  $u \in H_0^1(\Omega)$ , the value of  $\langle \phi, u \rangle$  is

$$\langle \phi, u \rangle = \int_{\Omega} \left( \phi_0 u + \sum_{j=1}^J \phi_j \partial_{x_j} u \right),$$

2. the norm of  $\phi$  is

$$\|\phi\|_{H^{-1}(\Omega)} = \inf \left\{ \left( \sum_{j=0}^J \|\phi_j\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \mid \{\phi_j\}_{j=0}^J \subset L^2(\Omega), \text{ and 1. holds} \right\}.$$

Then, we write  $\phi = \phi_0 - \sum_{j=1}^J \partial_{x_j} \phi_j$ , see [5].

**Example 1.3.25.** Let  $\Omega = (-1, 1) \subset \mathbb{R}$ . The Dirac delta concentrated to  $x = 0$  is in  $H^{-1}(\Omega)$ . To verify this, we use the triangle inequality, Hölder's inequality, and Remark 1.3.14

$$|\langle \delta_0, u \rangle| = \left| \int_{-1}^0 u' \right| \leq \int_{-1}^0 |u'| \leq \int_{-1}^1 |u'| = \|1 \cdot u'\|_{L^1} \leq \|1\|_{L^2} \|u'\|_{L^2} \leq \sqrt{2} \|u\|_{H_0^1} \quad (u \in H_0^1(\Omega)).$$

Here,  $\phi_0 = 0$  and  $\phi_1 = 1 - H$ , where  $H \in L^2((-1, 1))$  is the Heaviside step function, because

$$\langle \delta_0, u \rangle = \int_{-1}^1 (0 \cdot u + (1 - H)u') = \int_{-1}^0 u'.$$

Therefore, we might write that

$$\delta_0 = \phi_0 - \phi_1' = 0 - (1 - H)' = H'.$$

We conclude this subsection with two boundary value problems formulated using Sobolev spaces and are satisfied in the weak sense.

**Example 1.3.26.** (Picture hanging on the wall)

Let  $\Omega = (-2, 2) \subset \mathbb{R}$ ,  $F = -\frac{1}{2}(\delta_{-1} + \delta_1) \in H^{-1}(\Omega)$ , and consider the boundary value problem

$$\left. \begin{aligned} u'' &= F \\ u|_{\partial\Omega} &= 0, \end{aligned} \right\}$$

which describes the shape of a string that supports a unit weight in two pieces, each attached to the string at  $x = -1$  and  $x = 1$ . It is easy to verify based on  $\frac{1}{2}|x - a|'' = H'(x - a) = \delta_a$  that the solution of this BVP is the function

$$u(x) = \frac{|x + 1| + |x - 1|}{2} - 2, \quad (x \in [-2, 2]),$$

( $u \in H_0^1$ ) whose plot, of course, is the shape that one expects in this setup.

**Example 1.3.27.** (Unit charge in a unit ball with Dirichlet boundary condition)

Let  $\Omega = B(\mathbf{0}, 1) \subset \mathbb{R}^3$ ,  $F = \delta_{\mathbf{0}} \in H^{-2}(\Omega)$ , and  $g \equiv 0 \in H^{\frac{3}{2}}(\Omega)$ . Consider the boundary value problem

$$\left. \begin{aligned} \Delta \phi &= F \\ \phi|_{\partial\Omega} &= g, \end{aligned} \right\}$$

which describes a unit charge placed in the middle of a spherical cavity of radius one, where the boundary of the cavity has fixed zero electric potential (earthing/grounding), and the unknown function  $\phi$  corresponds to the electric potential.

One might guess that the solution is a linear function of the fundamental solution  $E_3$  defined in Remark 1.2.9. In fact,  $\phi$  is of the form  $E_3 + b$ , where  $b$  is determined such that  $u$  satisfies the boundary condition.

*Remark 1.3.28.* It is also true that if we impose the Neumann-type homogeneous boundary condition  $\partial_{\mathbf{n}}\phi|_{\partial\Omega} = 0$ , we do not have a solution.

### 1.3.3 Existence and uniqueness theorems

In this subsection, we state the existence and uniqueness theorems that apply to the handful of PDEs that we are going to study later. First, we state a stronger version of the Green–Gauß formulas that we are going to apply.

**Theorem 1.3.29.** (*General Green–Gauß formula*)

Let  $p \in L^\infty(\Omega)$ ,  $u \in W^{1,p}(\Omega)$  such that  $\operatorname{div}(p\nabla u) \in L^p(\Omega)$ . Then  $p\partial_{\mathbf{n}}u \in W^{-\frac{1}{p},p}(\partial\Omega)$  and

$$\int_{\Omega} -\operatorname{div}(p\nabla u)\varphi = \int_{\Omega} p\nabla u \cdot \nabla \varphi - \underbrace{\int_{\partial\Omega} \varphi(p\partial_{\mathbf{n}}u)}_{\langle \varphi, p\partial_{\mathbf{n}}u \rangle} \quad (\forall \varphi \in W^{1,p'}(\Omega)), \quad (1.50)$$

in particular, if  $\varphi|_{\partial\Omega} = 0$ , then the boundary term vanishes, and

$$\int_{\Omega} -\operatorname{div}(p\nabla u)\varphi = \int_{\Omega} p\nabla u \cdot \nabla \varphi \quad (\forall \varphi \in W_0^{1,p'}(\Omega)). \quad (1.51)$$

*Outline of the proof.* Ern and Guermond [4] suggest the following proof. Let us apply divergence theorem  $\int_{\Omega} \operatorname{div} \mathbf{F} = \int_{\partial\Omega} \mathbf{F} \cdot \mathbf{n}$  ( $\mathbf{F} \in C^1(\Omega, \mathbb{R}^n)$ ) to  $\mathbf{F} = (p\nabla u)\varphi$  when it is a smooth function

$$\int_{\Omega} \operatorname{div}((p\nabla u)\varphi) = \int_{\partial\Omega} ((p\nabla u)\varphi) \cdot \mathbf{n} \quad (p, u, \varphi \in C^2(\Omega)). \quad (1.52)$$

Here, the right-hand side is the desired boundary term, while the left-hand side may be rewritten using the identity

$$\operatorname{div}((p\nabla u)\varphi) = p\nabla u \cdot \nabla \varphi + \operatorname{div}(p\nabla u)\varphi,$$

which enables us to rearrange (1.52) into (1.51). A density argument concludes the proof.  $\square$

Now, let us examine the general elliptic boundary value problem with inhomogenous Dirichlet-type boundary conditions. Let  $\Omega \subset \mathbb{R}^n$  be a non-empty, bounded open domain and consider the following elliptic partial differential equation with Dirichlet-type boundary condition

$$\left. \begin{aligned} Lu &:= -\operatorname{div}(p\nabla u) = f \\ u|_{\partial\Omega} &= g, \end{aligned} \right\} \quad (1.53)$$

where  $p \in L^\infty(\Omega)$ ,  $p(\mathbf{x}) \geq m > 0$  a.e., and  $f \in L^2(\Omega)$  are given, and we seek  $u$  in the most general space possible. Since both  $L$  and the trace operator are linear, it is reasonable to look for  $u$  in the form of  $u_1 + u_2$ , where

$$\left. \begin{aligned} Lu_1 &= f \\ u_1|_{\partial\Omega} &= 0, \end{aligned} \right\} \quad \text{and} \quad \left. \begin{aligned} Lu_2 &= 0 \\ u_2|_{\partial\Omega} &= g, \end{aligned} \right\} \quad (1.54)$$

Multiplying both sides of the left equation by an arbitrary test function from  $\varphi \in C_0^1(\Omega)$  and then integrating over the domain  $\Omega$  yields the integral equation

$$\int_{\Omega} -\operatorname{div}(p\nabla u_1)\varphi = \int_{\Omega} f\varphi \quad (\forall \varphi \in C_0^1(\Omega)).$$

By applying the general Green's formula (1.51) to the left-hand side, we may rewrite the above equation as

$$\int_{\Omega} p\nabla u \cdot \nabla \varphi = \int_{\Omega} f\varphi \quad (\forall \varphi \in C_0^1(\Omega)), \quad (1.55)$$

which is called the weak form of the left BVP. Ern and Guermond [4] give the following existence theorem to these problems.

**Theorem 1.3.30.** (*Convex polytope*)

Let  $\Omega \subset \mathbb{R}^n$  be a convex polytope, and suppose that  $p \in C^1(\Omega)$ ,  $f \in L^p(\Omega)$ . Then

1. The unique solution  $u_1$  of (1.54) is in  $W_0^{2,p}(\Omega)$ , and
2. if  $\Omega \subset \mathbb{R}^2$ ,  $f \in L^2(\Omega)$  and  $g \in H^{\frac{3}{2}}(\partial\Omega)$  then the unique solution  $u_2$  of the right-hand side BVP (1.54) is in  $H^2(\Omega)$ .

*Remark 1.3.31.* If  $\Omega$  is a concave polytope, then it has a concave interior angle, then we cannot have second-order regularity of  $u$ . See [4] for details.

## 1.4 Basics of artificial neural networks

Artificial neural networks are parametrized functions whose structure is motivated by the human brain. They have a wide and rapidly broadening range of applications, such as image recognition, natural language processing, speech recognition, health care, operations research, and numerical analysis.

### 1.4.1 Structure of neural networks

A neural network consists of a large number of interconnected processing units that hold activation values. These units represent neurons of the artificial neural network. Neurons are organized into layers, with each layer performing a specific function. The input layer receives the input data, which is then processed through a series of hidden layers before producing the final output.

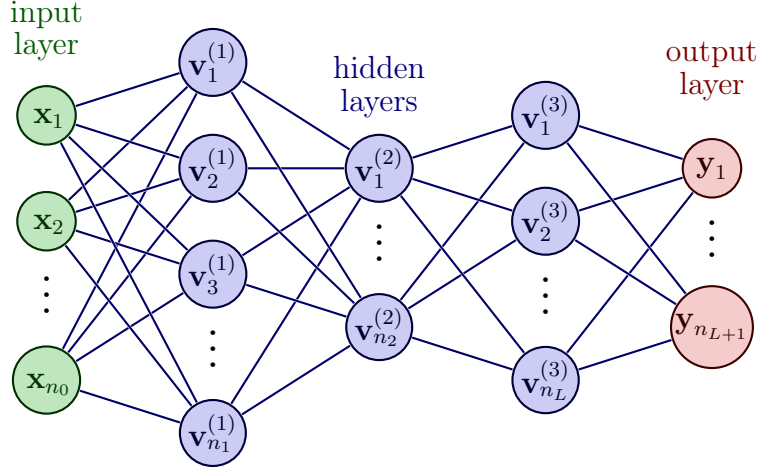


Figure 1.4: The structure of a neural network.

Mathematically, we have vectors  $\mathbf{v}^{(j)} \in \mathbb{R}^{n_j}$  that contain the activation values of each layer ( $j = 0, \dots, L + 1$ ), where  $L$  is the number of hidden layers,  $\mathbf{x} = \mathbf{v}^0$  is the input vector (data),  $\mathbf{y} = \mathbf{v}^{(L+1)}$  is the output. The  $j^{\text{th}}$  layer contains  $n_j \in \mathbb{N}^+$  neurons. Then, we define functions  $f_j: \mathbf{v}^{(j)} \mapsto \mathbf{v}^{(j+1)}$  ( $j = 0, \dots, L$ ) between the layers as

$$f_j(\mathbf{x}) := \boldsymbol{\sigma}_j(\mathbf{A}_j \mathbf{x} + \mathbf{b}) \quad (1.56)$$

with  $\mathbf{A}_j \in \mathbb{R}^{n_{j+1} \times n_j}$  weight matrix,  $\mathbf{b} \in \mathbb{R}^{n_{j+1}}$  bias vector and  $\boldsymbol{\sigma}_j: \mathbb{R}^{n_{j+1}} \rightarrow \mathbb{R}^{n_{j+1}}$  activation function. Typically,  $\boldsymbol{\sigma}$  is the same real function coordinate-wise for any given  $j$ . Finally, the neural network is the function

$$f_{\boldsymbol{\theta}} := f_L \circ f_{L-1} \circ \dots \circ f_0. \quad (1.57)$$

Here  $\boldsymbol{\theta}$  stands for all the parameters that are the elements of all the weight matrices and bias vectors. This means, that  $\boldsymbol{\theta}$  may have up to  $\sum_{j=0}^L (n_j + 1)n_{j+1} = \mathcal{O}(L \cdot N^2)$  free parameters, where  $N = \max_j \{n_j\}$ .

To train the network, one needs to define a loss function that measures how well the network performs on a given task. For example, when solving a PDE numerically, it makes sense to pick a loss function that measures some norm (perhaps, the discrete  $\|\cdot\|_{H^1}$ -norm) of the residual error of the numerical estimation. During training, the weights and biases in the neural network – i.e., elements of the matrices  $\mathbf{A}_j$  and the bias vectors  $\mathbf{b}_j$  – are fine-tuned to minimize this loss function. This may be done using several optimization algorithms, such as variations of the gradient descent algorithm, etc. [1]

The patterns of connections between layers – i.e., the patterns of indices for which the elements of the weight matrices and bias vectors may be non-zero – may vary depending on the task the network is built to accomplish. In a dense (or fully connected) neural network, each neuron in one layer is connected to every neuron in the neighboring layers. In a convolutional neural network, the connections are more localized, with each neuron in one layer only connected to a subset of neurons in the next layer. This feature attempts to resemble neural pathways and assumes some form of a locality to the input data, thus trading a great deal of generality for reducing the number of parameters that must be fine-tuned. Each connection between neurons is associated with a weight, representing the connection's strength. See [1] for many more details on artificial neural networks.

# Chapter 2

## The method of fundamental solutions and artificial neural networks

*“The heart of mathematics consists of concrete examples and concrete problems. Big general theories are usually afterthoughts based on small but profound insights; the insights themselves come from concrete special cases.”*  
– Halmos Pál

The method of fundamental solutions was first proposed by V. D. Kupradze and M. A. Aleksidze in 1963, see [15]. A general proof of convergence is currently not available. Cheng and Hong [7] claim that *“if the boundary is sufficiently smooth, and the solution has analytic continuation outside the domain, the [MFS] approximation has the form”*  $\varepsilon \sim \left(\frac{\text{diam}(\Omega)}{\text{diam}(\Sigma)}\right)^{N/n}$ , where  $\Sigma$  is the fictitious boundary (to be discussed later),  $N$  is the number of collocation points, and  $n \in \{1, 2, 3\}$  depending on different proofs. Furthermore, an abundance of publications discussed in [6] and [7] greatly support the numerical effectiveness of the MFS for second-order elliptic problems. In this chapter, we introduce the MFS through a concrete example and discuss how one might apply neural networks that are inspired by this method as Izsák and Haffner have already done [8]. Finally, we propose modifications to this approach that are numerically tested in Chapter 3.

### 2.1 Introducing the MFS scheme through an example

Consider Laplace’s equation with inhomogeneous Dirichlet-type boundary condition

$$\left. \begin{aligned} \Delta u &= 0 && (\text{in } \Omega) \\ u &= g && (\text{on } \partial\Omega), \end{aligned} \right\} \quad (2.1)$$

where  $\Omega = (0, 2) \times (0, 1)$  and  $g: \partial\Omega \rightarrow \mathbb{R}$  is defined as

$$g(x, y) = \begin{cases} x^2 & \text{if } (x, y) \in [0, 2] \times \{0\} \\ 1 - y^2 & \text{if } (x, y) \in \{1\} \times [0, 2] \\ x^2 - 1 & \text{if } (x, y) \in [0, 2] \times \{1\} \\ -y^2 & \text{if } (x, y) \in \{1\} \times [0, 2]. \end{cases}$$

Theorem 1.3.30 ensures that there exists a unique solution  $u \in H^2(\Omega)$ . Indeed, notice that  $g(x, y) = x^2 - y^2$  on  $\partial\Omega$ , which is the restriction of the harmonic polynomial function  $x^2 - y^2$  to  $\partial\Omega$ . Hence, the exact solution of this BVP is  $u: \Omega \rightarrow \mathbb{R}$ ,  $u(x, y) = x^2 - y^2$ . We are going

to demonstrate the method of fundamental solutions (MFS) by approximating this solution as if it were unknown. Then, we are going to compare the resulting  $\tilde{u}$  with the exact solution to draw conclusions.

The basic idea of the MFS is to take copies of the fundamental solution of the operator and place them on the boundary of a fictitious domain  $\Sigma$  that encompasses  $\Omega$ . Let us define the following auxiliary sets:

$$Y = \{\mathbf{y}_j\}_{j=0}^{J-1} \subset \text{ext } \Omega \quad \text{and} \quad Z = \{\mathbf{z}_k\}_{k=0}^{K-1} \subset \partial\Omega, \quad \text{as}$$

$$Y = \left\{ (-1, -1), (1, -1), (3, -1), \left(3, \frac{1}{2}\right), (3, 2), (1, 2), (-1, 2), \left(-1, \frac{1}{2}\right) \right\}, \quad \text{and}$$

$$Z = \left\{ (0, 0), (1, 0), (2, 0), \left(2, \frac{1}{2}\right), (2, 1), (1, 1), (0, 1), \left(0, \frac{1}{2}\right) \right\}.$$

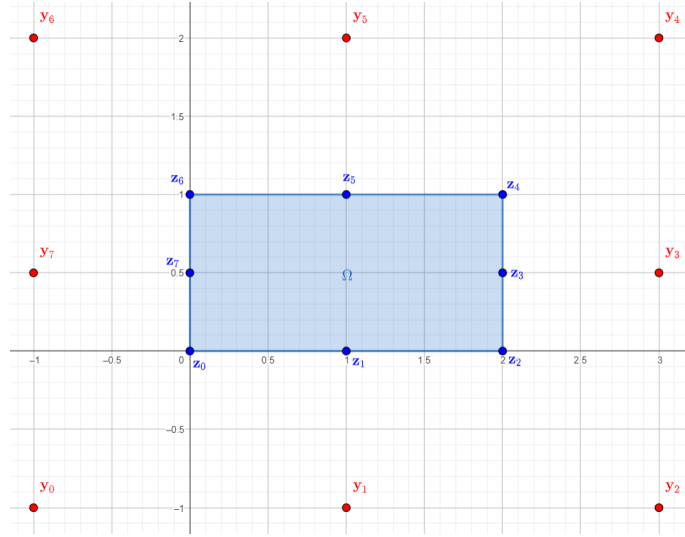


Figure 2.1:  $\Omega$ ,  $Y$ , and  $Z$  as defined above

Note that the external points are further away from the domain, as the error approximation discussed in the introduction of this chapter suggests that it is beneficial for the diameter of the fictitious domain to be large compared to the diameter of  $\Omega$ .

According to Example 1.2.7, the fundamental solution of the two-dimensional Laplace operator is the function  $E_2(\mathbf{x}) = \frac{1}{2\pi} \ln \|\mathbf{x}\|_2$ . Let

$$\Psi_j(\mathbf{x}) := E_2(\mathbf{x} - \mathbf{y}_j) = \frac{1}{2\pi} \ln \|\mathbf{x} - \mathbf{y}_j\|_2 \quad (j = 0, \dots, J-1). \quad (2.2)$$

Then, we seek the numerical approximation  $\tilde{u}$  as a linear combination of these functions

$$\tilde{u} := \sum_{j=0}^{J-1} \alpha_j \Psi_j \quad (2.3)$$

such that the error vector  $\mathbf{e} := (\tilde{u}(\mathbf{z}_k) - g(\mathbf{z}_k))_{k=0}^{K-1} \in \mathbb{R}^K$  is minimal in some norm (ideally  $\mathbf{0}$ ).

*Remark 2.1.1.* The intuition for this approach comes from the fact that according to the boundary integral equations discussed in [14], for every boundary function  $g \in H^{\frac{1}{2}}(\partial\Omega)$ , there exists a unique *Green function*  $\mathcal{G} \in H^{-\frac{1}{2}}(\partial\Omega)$  such that the unique solution of the boundary value problem 2.1 is of the form of the surface integral

$$u(\mathbf{x}) = \int_{\partial\Omega} \mathcal{G}(\mathbf{y}) E_2(\mathbf{x} - \mathbf{y}) \, dS(\mathbf{y}) \quad (\mathbf{x} \in \Omega), \quad (2.4)$$

and the sum in (2.3) can be thought of as the discrete approximation of the integral in (2.4), where values of  $G(\mathbf{y})$  correspond to the coefficients in the sum.

Substituting  $\tilde{u}$  from (2.3) into the boundary condition of (2.1) shows that finding the appropriate coefficients  $\alpha_j$  comes down to solving the following linear system of equations

$$\sum_{j=0}^{J-1} \alpha_j \Psi_j(\mathbf{z}_k) = g(\mathbf{z}_k) \quad (k = 0, \dots, K-1)$$

for  $\alpha_j$ , which is equivalent to the matrix equation

$$\begin{bmatrix} \Psi_0(\mathbf{z}_0) & \Psi_1(\mathbf{z}_0) & \dots & \Psi_{J-1}(\mathbf{z}_0) \\ \Psi_0(\mathbf{z}_1) & \Psi_1(\mathbf{z}_1) & \dots & \Psi_{J-1}(\mathbf{z}_1) \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_0(\mathbf{z}_{K-1}) & \Psi_1(\mathbf{z}_{K-1}) & \dots & \Psi_{J-1}(\mathbf{z}_{K-1}) \end{bmatrix} \cdot \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{J-1} \end{bmatrix} = \begin{bmatrix} g(\mathbf{z}_0) \\ g(\mathbf{z}_1) \\ \vdots \\ g(\mathbf{z}_{K-1}) \end{bmatrix}. \quad (2.5)$$

After solving the above equation for the coefficient vector  $\boldsymbol{\alpha}$ , it becomes possible to evaluate the numerical solution  $\tilde{u}$  through Equation (2.3) for any  $\mathbf{x} \in \bar{\Omega}$ . In our example,

$$\boldsymbol{\alpha} = (6.62, 7.53, -8.17, -13.21, -0.528, 11.41, 14.26, -7.05)^T \in \mathbb{R}^8,$$

which gives us the numerical solution

$$\tilde{u}(x, y) = 1.05 \ln \sqrt{(x+1)^2 + (y+1)^2} + 1.2 \ln \sqrt{(x-1)^2 + (y+1)^2} - \quad (2.6)$$

$$- 1.3 \ln \sqrt{(x-3)^2 + (y+1)^2} - 2.1 \ln \sqrt{(x-3)^2 + (y-0.5)^2} - \quad (2.7)$$

$$- 0.084 \ln \sqrt{(x-3)^2 + (y-2)^2} + 1.82 \ln \sqrt{(x-1)^2 + (y-2)^2} + \quad (2.8)$$

$$+ 2.27 \ln \sqrt{(x+1)^2 + (y-2)^2} - 1.12 \ln \sqrt{(x+1)^2 + (y-0.5)^2}. \quad (2.9)$$

Notice that the coefficients in (2.6)-(2.9) are  $\frac{\alpha_j}{2\pi}$  due to (2.2).

The accuracy of this estimation in the maximum norm is  $\|\tilde{u} - u\|_{\max} \approx 0.0514$ , and in  $L^2$ -norm it is  $\|\tilde{u} - u\|_{L^2(\Omega)} \approx 0.0178$  with only 8 collocation points. Adding 8 more collocation points results in an accuracy of  $\|\tilde{u} - u\|_{\max} \approx 0.003$ , and  $\|\tilde{u} - u\|_{L^2(\Omega)} \approx 0.0005$ , which is an excellent approximation for a method as cheap as the solution of a  $16 \times 16$  system of linear equations.

Adding further collocation points and looking at the graphs of the errors does indeed suggest convergence. Cheng and Hong [7] claim that “*traditional error analysis for the MFS shows that when the boundary condition is prescribed by a harmonic function, the convergence is exponential either by increasing the number of terms in the approximation, or by increasing the radius of the fictitious boundary.*” See the plot of  $\tilde{u} - u$  below for all our experiments.

*Remark 2.1.2.* Note that in the 60 collocations point case, we have achieved that the error is mere random noise coming from computational inaccuracy.

*Remark 2.1.3.* The uneven nature of the error on the boundary is particularly visible in the 16 and 36 collocation points cases. This suggests that we could perhaps get better estimates in various norms if we also let the collocation points vary around the domain. Allowing this results in a non-linear system of equations instead of (2.5), see [6] for details.

*Remark 2.1.4.* Applying the MFS algorithm in practice – in particular, solving (2.5) – has several numerical difficulties.

1. The resulting system of equations is typically very poorly conditioned: in our two cases, the condition numbers of the matrices were 86 and 15,396 in the 8 and 16 collocation point cases respectively. This may result in wildly different solutions  $\boldsymbol{\alpha}$  given a small perturbation of the input data  $\mathbf{g}$ .



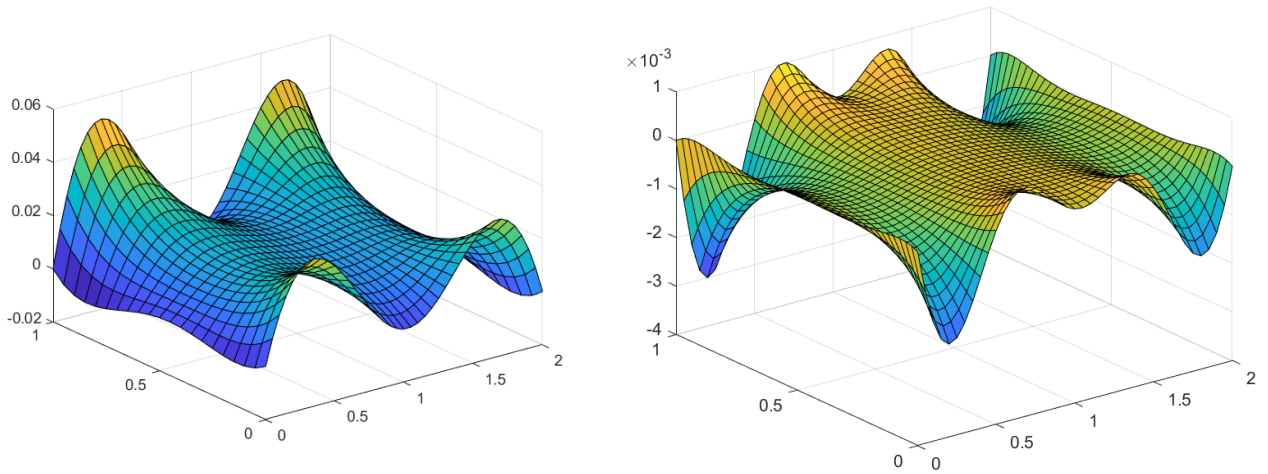


Figure 2.2: Plots of  $\tilde{u} - u$  over  $\Omega$  with 8 and 16 collocation points.

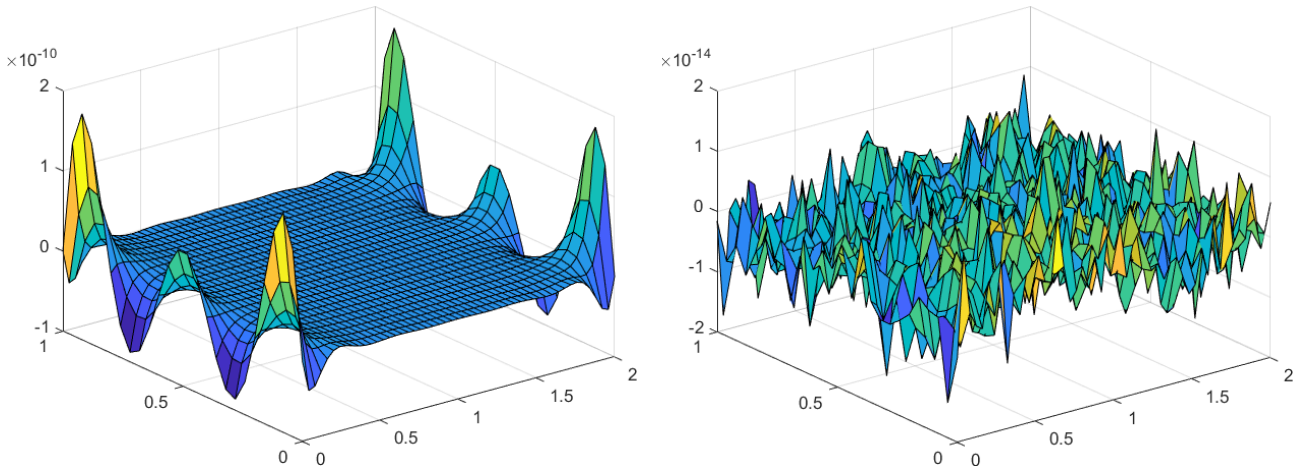


Figure 2.3: Plots of  $\tilde{u} - u$  over  $\Omega$  with 36 and 60 collocation points.

2. The matrix  $\mathbf{G}$  has no clear structure that may be typical in other numerical methods. Furthermore, it is not even guaranteed to be invertible, if we choose setups that are “too symmetric”. For all intents and purposes,  $\mathbf{G}$  is a poorly conditioned, full matrix with random entries. [7]

For these reasons, it is sensible to look for a neural network-based numerical method that circumvents the solution of (2.5) – especially in the case where it is replaced by a non-linear system of equations – but still mimics the approach of finding  $\tilde{u}$  as a linear combination of shifted copies of fundamental solutions.

## 2.2 Solving the MFS scheme for Laplace’s equation with Dirichlet-type boundary condition using a primitive neural network

The following implementation was proposed by Izsák and Haffner [8]. Suppose that we are only interested in the values of the function at points  $X = \{\mathbf{x}_j\}_{j=0}^{I-1} \subset \Omega$ . Let us train a neural network – whose architecture will be specified later – such that

$$\text{NN} : \underbrace{(\Psi_j(\mathbf{z}_k))_{k=0}^{K-1}}_{\in \mathbb{R}^K} \mapsto \underbrace{(\Psi_j(\mathbf{x}_i))_{i=0}^{I-1}}_{\in \mathbb{R}^I} \quad (j = 0, \dots, J-1). \quad (2.10)$$

Then, we define the numerical approximation as

$$\tilde{u}: \mathbb{R}^K \rightarrow \mathbb{R}^I, \quad (\tilde{u}(\mathbf{x}_i))_{i=0}^{I-1} := \text{NN} \left( (g(\mathbf{z}_k))_{k=0}^{K-1} \right) \quad (2.11)$$

The NN is trained to return the function values in the inner points for function values on the boundary. In other words, the NN learns to predict the values inside the domain given the values on the boundary, which is really the task in (2.1). The results of this and two other MFS algorithms applied to elliptic boundary value problems shall be discussed in Chapter 3.

**Theorem 2.2.1.** *Suppose that*

- the sets of points  $Y$  and  $Z$  are such that the set of vectors

$$G = \left\{ (\Psi_j(\mathbf{z}_k))_{k=0}^{K-1} \in \mathbb{R}^K \right\}_{j=0}^{J-1}$$

is a generating set in  $\mathbb{R}^K$  ( $K \leq J$ ), and

- all layers in the neural network are linear, i.e. NN is a linear map from  $\mathbb{R}^K$  to  $\mathbb{R}^I$ .

Then, the algorithm defined by (2.10) and (2.11) is equivalent to the MFS algorithm.

*Proof.* It is sufficient to show that the algorithm defined by (2.10) and (2.11) also produces the estimation

$$\tilde{u}(\mathbf{x}_i) = \sum_{j=0}^{J-1} \alpha_j \Psi_j(\mathbf{x}_i) \quad (i = 0, \dots, I-1) \quad (2.12)$$

with some coefficients  $\alpha_j$ . Since  $G$  is a generating set in  $\mathbb{R}^K$ , there exist coefficients  $\alpha_0, \dots, \alpha_{J-1}$  such that we can express the vector of boundary values as

$$(g(\mathbf{z}_k))_{k=0}^{K-1} = \sum_{j=0}^{J-1} \alpha_j (\Psi_j(\mathbf{z}_k))_{k=0}^{K-1}.$$

Applying the linear map NN to both sides of the equation yields

$$(\tilde{u}(\mathbf{x}_i))_{i=0}^{I-1} = \text{NN} \left( (g(\mathbf{z}_k))_{k=0}^{K-1} \right) = \text{NN} \left( \sum_{j=0}^{J-1} \alpha_j (\Psi_j(\mathbf{z}_k))_{k=0}^{K-1} \right) = \quad (2.13)$$

$$= \sum_{j=0}^{J-1} \alpha_j \text{NN} \left( (\Psi_j(\mathbf{z}_k))_{k=0}^{K-1} \right) = \sum_{j=0}^{J-1} \alpha_j (\Psi_j(\mathbf{x}_i))_{i=0}^{I-1}, \quad (2.14)$$

that is,  $\tilde{u} = \sum_{j=0}^{J-1} \alpha_j \Psi_j$ , the same as the one provided by the algorithm in (2.10) and (2.11).  $\square$

*Remark 2.2.2.* Of course, this is only a valid approach if it is guaranteed that there exist such  $\alpha_j$  scalars for which

$$(\tilde{u}(\mathbf{x}_i))_{i=0}^{I-1} = \sum_{j=0}^{J-1} \alpha_j (\Psi_j(\mathbf{x}_i))_{i=0}^{I-1}.$$

That is if the vectors  $\left\{ (\Psi_j(\mathbf{x}_i))_{i=0}^{I-1} \in \mathbb{R}^I \right\}_{j=0}^{J-1}$  form a generating system of the vector space  $\mathbb{R}^I$ .

The difference between the applicability of this approach versus the standard MFS algorithm is akin to the difference between solving a system of linear equations once by Gauß–Jordan Elimination in  $\mathcal{O}(K \cdot J)$  steps, versus finding the inverse of a  $K \times K$  matrix ( $J = K$ ) in  $\mathcal{O}(K^3)$  steps. Furthermore, once this algorithm is implemented, it only finds the numerical estimation in a finite set of points, and therefore, if we are interested in  $\tilde{u}(\mathbf{x})$  where  $\mathbf{x} \notin X$ , we either have to teach the neural network all over again, or interpolate from the existing estimations. In other words, this is no longer a mesh-free method.

Another – perhaps more natural – extension of the MFS with neural networks is to let the network guess the unknown scalars  $\alpha_j$  that belong to specific boundary conditions. We have not found a reference that attempts to do exactly this.

## 2.3 Introducing a generalized version of the MFS algorithm

Just like in the case of the MFS, we present the following modified version of the MFS through a concrete example that is so elementary that the reader can conceivably follow along and replicate the experiment with little effort. However, for the sake of brevity, we are not presenting all the numerical outputs as in Section 2.1, only the qualitatively important ones.

Consider Poisson’s equation with inhomogeneous Dirichlet-type boundary condition

$$\left. \begin{aligned} \Delta u &= f && (\text{inside } \Omega) \\ u &= g && (\text{on } \partial\Omega), \end{aligned} \right\} \quad (2.15)$$

where  $\Omega = B(\mathbf{0}, 1) \subset \mathbb{R}^2$ ,  $f(x, y) = 24y$  and  $g(\cos \theta, \sin \theta) = \cos 3\theta$  for  $\theta \in [0, 2\pi)$ . By Theorem 1.3.30, there exists a unique solution  $u \in H^2(\Omega)$ . Indeed, one can check that the exact solution of this BVP is  $u(x, y) = -3x + 4y^3$ , as

$$\left. \begin{aligned} \Delta u(x, y) &= \Delta(-3x + 4y^3) = 24y = f(x, y) \\ u(\cos \theta, \sin \theta) &= -3 \cos \theta + 4 \sin^3 \theta = \cos 3\theta = g(\cos \theta, \sin \theta) \end{aligned} \right\} \quad (2.16)$$

We find a numerical estimation for this function as if it were unknown with a modified version of the MFS. Of course, taking (2.15) apart into Laplace’s boundary value problem with inhomogeneous Dirichlet-type boundary condition and Poisson’s boundary value problem with homogeneous Dirichlet-type boundary condition is a valid approach. However, this relies on the boundary value condition and the governing operator being linear, and we wish to develop a method that can deal with non-linear problems utilizing the power of neural networks as well.

To approximate the effects of the source function  $f$ , we need to include data from new functions  $\psi_m$ , for which  $\Delta\psi_m \neq 0$ . Let us take the following auxiliary sets as plotted below.

Here, the new set  $W$  serves as the set of measurement points, where  $\Delta\psi_m$  values are measured, and  $T$  is a set of collocation points that are used to define the functions  $\psi_m$ . Note that  $|Y_1| = |Z_1| = 11$ ,  $|W_1| = |T_1| = 14$ ,  $|Y_2| = |Z_2| = 17$ , and  $|W_2| = |T_2| = 21$ , as seen in Figure 2.4. The radius of the fictitious domain is 1.5 in both cases. The points of the auxiliary sets  $W_{1,2}$  and  $T_{1,2}$  lay on circles of radii  $\frac{1}{3}$ ,  $\frac{2}{3}$ , and 0.9. Let

$$\Psi_j(\mathbf{x}) := E_2(\mathbf{x} - \mathbf{y}_j) = \frac{1}{2\pi} \ln \|\mathbf{x} - \mathbf{y}_j\|_2, \quad (2.17)$$

as in the MFS scheme. We can pick functions  $\psi_m$  from a wide variety of functions, however, it is not easy to pick functions that work well for the following approximation. We managed to pick a class of functions that appears to work relatively well and therefore serves as proof

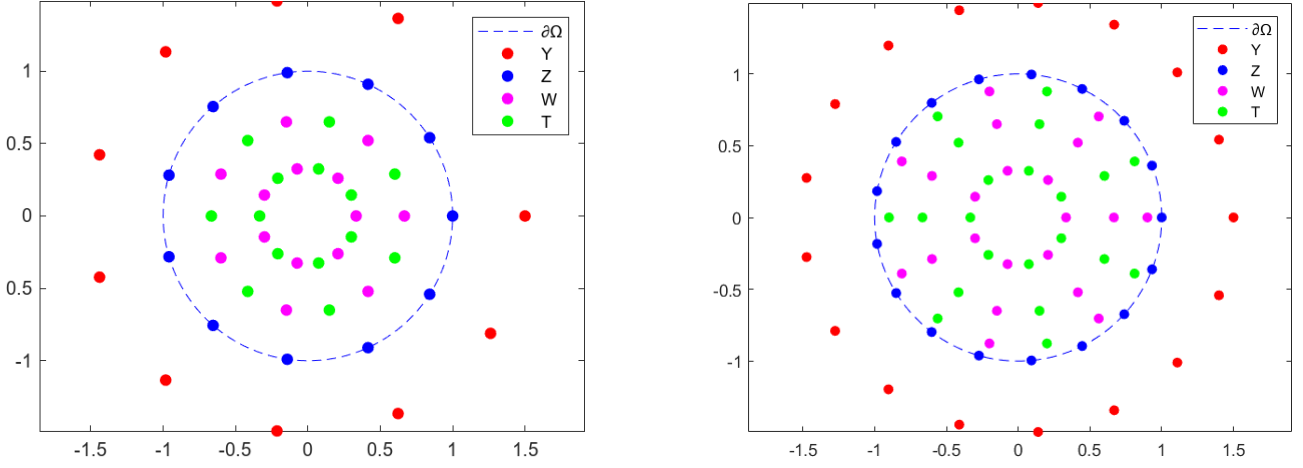


Figure 2.4: The sets  $\Omega$ ,  $Y$ ,  $Z$ ,  $W$ ,  $T$  as defined above.

of concept. However, further research toward finding a set of functions that work as well as fundamental solutions work in the case of the MFS is still needed.

The reasoning behind our choice goes as follows.

- We approximate  $f$  by a linear combination of  $\Delta\psi_m$ . Heuristically, it makes sense to do that by shifting around the same function  $\psi$  for which  $\Delta\psi_m$  takes small values everywhere except on a relatively small patch. This way, we should be able to put  $f$  together as a patchwork of different functions.
- If we look for a radially symmetric function  $\psi$ , solving  $\Delta_{(x,y)}\psi(r) = h(r)$  comes down to solving a relatively simple ordinary differential equation, similar to what we did in Example 1.2.7.
- Picking  $h(r)$  to be an exponentially decaying bump such as  $e^{-r}$  or  $e^{-r^2}$  does achieve what we want in the first point. Setting  $h(r) = e^{-r}$  would result in a non-differentiable point on  $\psi$ , so we look for transformed versions of the function  $e^{-r^2}$ . Given the parameters of the domain, we chose  $h(r) = 8e^{-4r^2}$ .
- The corresponding possible  $\psi: \mathbb{R} \rightarrow \mathbb{R}$  functions are solutions of the ODE

$$\psi''(r) + \frac{1}{r}\psi'(r) = 8e^{-4r^2}, \quad (2.18)$$

which we can easily solve:

$$\begin{aligned} (r \cdot \psi'(r))' &= 1 \cdot \psi'(r) + r\psi''(r) = r\left(\psi''(r) + \frac{1}{r}\psi'(r)\right) = 8re^{-4r^2} \\ r \cdot \psi'(r) &= \int 8re^{-4r^2} dr = -e^{-4r^2} + c_1 \\ \psi(r) &= \int \left(-\frac{e^{-4r^2}}{r} + \frac{c_1}{r}\right) dr = -\frac{1}{2} \int -8r \frac{e^{-4r^2}}{-4r^2} dr + c_1 \int \frac{1}{r} dr = \\ &= -\frac{1}{2} \text{Ei}(-4r^2) + c_1 \ln(r) + c_2. \end{aligned}$$

Here, Ei is the exponential integral function  $\text{Ei}(r) := \int_{-\infty}^r \frac{e^\varrho}{\varrho} d\varrho$  for  $(r < 0)$ .

Both the natural logarithm and  $\text{Ei}(-4r^2)$  have a singularity at  $r = 0$ . However, the arbitrary constant  $c_1$  can be chosen such that  $\psi$  has a removable singularity at 0. Using the asymptotic

behavior of Ei around 0 – namely that  $\text{Ei}(x) = \gamma + \ln|x| + \mathcal{O}(x)$ , where  $\gamma$  is the Euler–Mascheroni constant, – we can pick  $c_1$  by calculating that

$$\lim_{r \rightarrow 0^+} \frac{\text{Ei}(-4r^2)}{\ln(r)} = \lim_{r \rightarrow 0^+} \frac{\gamma + \ln|-4r^2| + \mathcal{O}(r^2)}{\ln(r)} = \lim_{r \rightarrow 0^+} \frac{\gamma + \ln 4}{\ln(r)} + \lim_{r \rightarrow 0^+} \frac{2 \ln(r)}{\ln(r)} + \lim_{r \rightarrow 0^+} \frac{\mathcal{O}(r^2)}{\ln(r)} = 2,$$

and therefore,  $c_1 = -2 \cdot \frac{-1}{2} = 1$  is a good choice for  $\psi$  to have a removable singularity at  $r = 0$ . Moreover,

$$\begin{aligned} \lim_{r \rightarrow 0^+} \left( -\frac{1}{2} \text{Ei}(-4r^2) + \ln(r) \right) &= \lim_{r \rightarrow 0^+} \left( -\frac{1}{2} \left( \gamma + \ln|-4r^2| + \mathcal{O}(r^2) \right) + \ln(r) \right) = \\ &= -\frac{1}{2} \lim_{r \rightarrow 0^+} \left( \gamma + \ln(4r^2) - 2 \ln(r) \right) = -\frac{\gamma}{2} - \frac{1}{2} \lim_{r \rightarrow 0^+} \ln \frac{4r^2}{r^2} = -\frac{\gamma}{2} - \ln 2. \end{aligned}$$

See the lecture notes [13] for the asymptotic behavior of the Ei function. Therefore, we can finally define the smooth functions  $\psi_m: \mathbb{R}^2 \rightarrow \mathbb{R}$  as

$$\psi_m(\mathbf{x}) := \begin{cases} -\frac{1}{2} \text{Ei}(-4\|\mathbf{x} - \mathbf{t}_m\|^2) + \ln\|\mathbf{x} - \mathbf{t}_m\| - 2, & \mathbf{x} \neq \mathbf{t}_m, \\ -2 - \frac{\gamma}{2} - \ln 2, & \mathbf{x} = \mathbf{t}_m \end{cases} \quad (2.19)$$

for which  $\Delta\psi_m(x, y) = 8e^{-4\|\mathbf{x} - \mathbf{t}_m\|^2}$ .

Now, let us seek an approximation  $\tilde{u} \approx u$  as some linear combination of  $\Psi_j$  and  $\psi_l$

$$\tilde{u} = \sum_{j=0}^{J-1} \alpha_j \Psi_j + \sum_{m=0}^{M-1} \beta_m \psi_m.$$

Substituting (2.3) back into (2.15) yields the system of equations

$$\sum_{m=0}^{M-1} \Delta\beta_m \psi_m(\mathbf{w}_l) = f(\mathbf{w}_l) \quad (l = 0, \dots, L-1) \quad (2.20)$$

$$\sum_{j=0}^{J-1} \alpha_j \Psi_j(\mathbf{z}_k) + \sum_{m=0}^{M-1} \beta_m \psi_m(\mathbf{z}_k) = g(\mathbf{z}_k) \quad (k = 0, \dots, K-1), \quad (2.21)$$

which is equivalent to the matrix equation

$$\underbrace{\begin{bmatrix} 0 & \dots & 0 & \Delta\psi_0(\mathbf{w}_0) & \dots & \Delta\psi_{M-1}(\mathbf{w}_0) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \Delta\psi_0(\mathbf{w}_{L-1}) & \dots & \Delta\psi_{M-1}(\mathbf{w}_{L-1}) \\ \Psi_0(\mathbf{z}_0) & \dots & \Psi_{J-1}(\mathbf{z}_0) & \psi_1(\mathbf{z}_0) & \dots & \psi_{M-1}(\mathbf{z}_0) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \Psi_0(\mathbf{z}_{K-1}) & \dots & \Psi_{J-1}(\mathbf{z}_{K-1}) & \psi_1(\mathbf{z}_{K-1}) & \dots & \psi_{M-1}(\mathbf{z}_{K-1}) \end{bmatrix}}_{\in \mathbb{R}^{(L+K) \times (J+M)}} \cdot \underbrace{\begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_{J-1} \\ \beta_0 \\ \vdots \\ \beta_{M-1} \end{bmatrix}}_{\in \mathbb{R}^{(J+M) \times 1}} = \underbrace{\begin{bmatrix} f(\mathbf{w}_0) \\ \vdots \\ f(\mathbf{w}_{L-1}) \\ g(\mathbf{z}_0) \\ \vdots \\ g(\mathbf{z}_{K-1}) \end{bmatrix}}_{\in \mathbb{R}^{(L+K) \times 1}}.$$

After solving the system of equations for the coefficients  $\alpha_j$  and  $\beta_m$ , it is possible to evaluate the numerical solution  $\tilde{u}$  through Equation (2.3) for any  $\mathbf{x} \in \bar{\Omega}$ . Like in the MFS case, this is often a very ill-conditioned system, and one must take great care to ensure that the solution is accurate enough.

Further tests show that increasing only the number of  $W$  and  $T$  collocation points in and of themselves does not necessarily decrease the error in the maximum norm. However, cleverly increasing the number of all the collocation points can push the error further down. See Figures 2.5 and 2.6.

*Remark 2.3.1.* Notice that we broke the symmetry of the auxiliary sets in the last case because otherwise, the resulting matrix would have had identical rows otherwise.

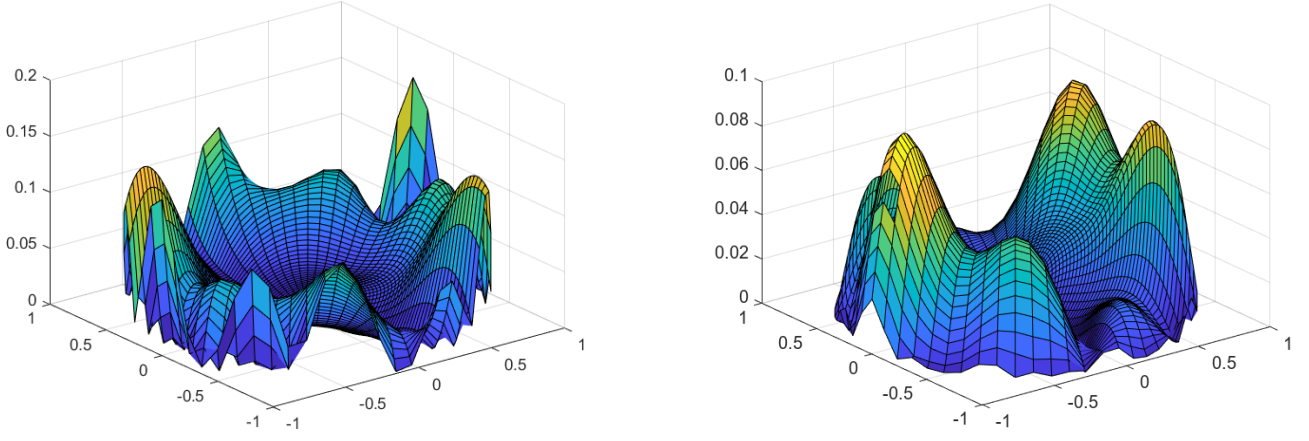


Figure 2.5: Plots of  $|\tilde{u} - u|$  over  $\Omega$  in both cases.

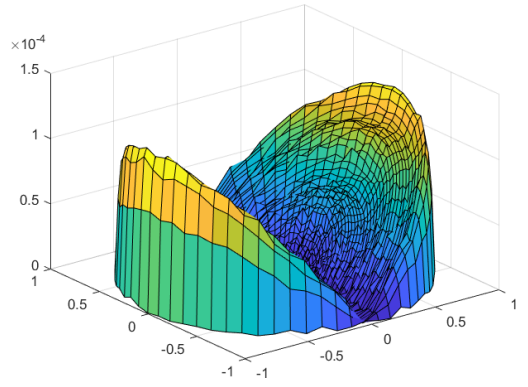
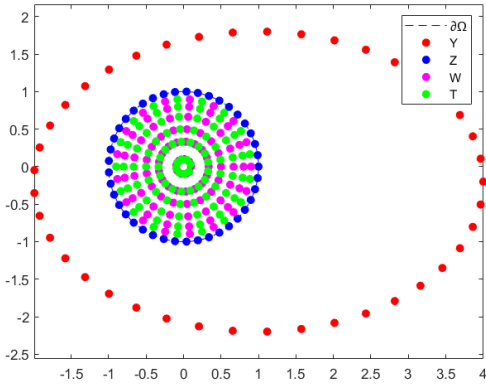


Figure 2.6: The auxiliary sets and the plot of  $|\tilde{u} - u|$  with  $|Y| = |Z| = 41$  and  $|W| = |T| = 102$ .

## 2.4 Solving the generalized MFS scheme for Poisson's equation using a primitive neural network

Now, we turn our attention to implementing the algorithm discussed in the previous section using a primitive neural network, just as in the case of the MFS scheme. We shall see that a very similar theorem holds for this approach also.

Let us consider the boundary value problem

$$\left. \begin{aligned} \Delta u &= f & (\text{in } \Omega) \\ u &= g & (\text{on } \partial\Omega). \end{aligned} \right\} \quad (2.22)$$

Suppose that we are interested in the values  $\{u(\mathbf{x}_i)\}_{i=0}^{I-1}$ , where  $X = \{\mathbf{x}_i\}_{i=0}^{I-1} \subset \Omega$ . To approximate these values, let us define the following auxiliary sets:

$$Y = \{\mathbf{y}_j\}_{j=0}^{J-1} \subset \text{ext } \Omega, \quad Z = \{\mathbf{z}_k\}_{k=0}^{K-1} \subset \partial\Omega, \quad \text{and} \quad W = \{\mathbf{w}_l\}_{l=0}^{L-1} \subset \Omega.$$

The task of the numerical approximation is to find a map  $A: \mathbb{R}^L \times \mathbb{R}^K \rightarrow \mathbb{R}^I$ , for which

$$A\left(\left(f(\mathbf{w}_l)\right)_{l=0}^{L-1}, \left(g(\mathbf{z}_k)\right)_{k=0}^{K-1}\right) \approx \left(u(\mathbf{x}_i)\right)_{i=0}^{I-1}.$$

Let us train a neural network using training data of the form

$$\text{NN} : \underbrace{\left(\mathbf{0}, \left(\Psi_j(\mathbf{z}_k)\right)_{k=0}^{K-1}\right)}_{\in \mathbb{R}^L \times \mathbb{R}^K} \mapsto \left(\Psi_j(\mathbf{x}_i)\right)_{i=1}^I \quad (j = 0, \dots, J-1), \quad (2.23)$$

as before. However, we cannot only use functions whose Laplacian vanishes everywhere inside the domain, since a linear combination of such functions also has a Laplacian that vanishes everywhere inside the domain. We also need to take input-output pairs with some functions  $\psi_m$  for which  $\Delta\psi_m \neq 0$  on  $\Omega$ . Let us include training data of the form

$$\text{NN}: \underbrace{\left( (\Delta\psi_m(\mathbf{w}_l))_{l=0}^{L-1}, (\psi_m(\mathbf{z}_k))_{k=0}^{K-1} \right)}_{\in \mathbb{R}^L \times \mathbb{R}^K} \mapsto (\psi_m(\mathbf{x}_i))_{i=1}^I \quad (m = 0, \dots, M-1), \quad (2.24)$$

and define the numerical approximation as

$$(\tilde{u}(\mathbf{x}_i))_{i=0}^{I-i} := \text{NN} \left( (f(\mathbf{w}_l))_{l=0}^{L-1}, (g(\mathbf{z}_k))_{k=0}^{K-1} \right). \quad (2.25)$$

**Theorem 2.4.1.** *Suppose that  $L \leq M$ ,  $K \leq J$ . Furthermore, suppose that*

(1) *the set of vectors  $\left\{ (\Delta\psi_m(\mathbf{w}_l))_{l=0}^{L-1} \right\}_{m=0}^{M-1}$  form a generating set in the vector space  $\mathbb{R}^L$ , and*

(2) *the set of vectors  $\left\{ (\Psi_j(\mathbf{z}_k))_{k=0}^{K-1} \right\}_{j=0}^{J-1}$  form a generating set in the vector space  $\mathbb{R}^K$ .*

*Then the algorithm defined by (2.23)-(2.25) with a linear neural network finds an approximation to (2.22) in the form of  $\tilde{u} = \sum_j \alpha_j \Psi_j + \sum_m \beta_m \psi_m$ .*

*Proof.* We omit writing the ranges of indices in the case of the vectors for the sake of simpler notation. First, let us express the vector containing the input data as

$$\left( (f(\mathbf{w}_l)), (g(\mathbf{z}_k)) \right) = \left( (f(\mathbf{w}_l)), \mathbf{0}_K \right) + \left( \mathbf{0}_L, (g(\mathbf{z}_k)) \right) := \mathbf{f} + \mathbf{g}.$$

Due to assumptions (1) and (2),  $\mathbf{f}$  and  $\mathbf{g}$  can be expressed as linear combinations

$$\mathbf{f} = \sum_{m=0}^{M-1} a_m \left( (\Delta\psi_m(\mathbf{w}_l)), \mathbf{0}_K \right), \quad \text{and} \quad (2.26)$$

$$\mathbf{g} = \sum_{j=0}^{J-1} b_j \left( \mathbf{0}_L, (\Psi_j(\mathbf{z}_k)) \right). \quad (2.27)$$

Here,  $\mathbf{g}$  is of the form required in (2.23), and so

$$\text{NN}(\mathbf{g}) = \text{NN} \left( \sum_{j=0}^{J-1} b_j \left( \mathbf{0}_L, (\Psi_j(\mathbf{z}_k)) \right) \right) = \sum_{j=0}^{J-1} b_j \text{NN} \left( \mathbf{0}_L, (\Psi_j(\mathbf{z}_k)) \right) = \sum_{j=0}^{J-1} b_j (\Psi_j(\mathbf{x}_i)). \quad (2.28)$$

However,  $\mathbf{f}$  is not of the form required by (2.24). Adding  $(\Delta\psi_m(\mathbf{w}_l))$  in the second coordinate, and subtracting it separately resolves this issue.

$$\text{NN}(\mathbf{f}) = \text{NN} \left( \sum_{m=0}^{M-1} a_m \left( (\Delta\psi_m(\mathbf{w}_l)), \psi_m(\mathbf{z}_k) \right) - \sum_{m=0}^{M-1} a_m \left( (\mathbf{0}_K, \psi_m(\mathbf{z}_k)) \right) \right) = \quad (2.29)$$

$$= \sum_{m=0}^{M-1} a_m \text{NN} \left( (\Delta\psi_m(\mathbf{w}_l)), \psi_m(\mathbf{z}_k) \right) - \sum_{m=0}^{M-1} a_m \text{NN} \left( (\mathbf{0}_K, \psi_m(\mathbf{z}_k)) \right) = \quad (2.30)$$

$$= \sum_{m=0}^{M-1} a_m (\psi_m(\mathbf{x}_i)) - \sum_{m=0}^{M-1} a_m \text{NN}(\mathbf{h}_m), \quad (2.31)$$

To complete the proof, we need to calculate  $\text{NN}(\mathbf{h}_m)$ . Again, we may use assumption (2) for  $\mathbf{h}_m$  to conclude the existence of scalars  $c_{mj}$  for which

$$\mathbf{h}_m = \sum_{j=0}^{J-1} c_{mj} \left( \mathbf{0}_L, (\Psi_j(\mathbf{z}_k)) \right), \quad (2.32)$$

and therefore for every  $m = 0, \dots, M-1$ , we have that

$$\text{NN}(\mathbf{h}_m) = \text{NN} \left( \sum_{j=0}^{J-1} c_{mj} \left( \mathbf{0}_L, (\Psi_j(\mathbf{z}_k)) \right) \right) = \sum_{j=0}^{J-1} c_{mj} \text{NN} \left( \mathbf{0}_L, (\Psi_j(\mathbf{z}_k)) \right) = \sum_{j=0}^{J-1} c_{mj} \Psi_j(\mathbf{x}_i). \quad (2.33)$$

Putting (2.28), (2.29)-(2.31) and (2.33) together, and using the linearity of NN, we finally have

$$\begin{aligned} (\tilde{u}(\mathbf{x}_i)) &= \text{NN} \left( (f(\mathbf{w}_l)), (g(\mathbf{z}_k)) \right) = \text{NN}(\mathbf{f} + \mathbf{g}) = \text{NN}(\mathbf{f}) + \text{NN}(\mathbf{g}) = \\ &= \sum_{m=0}^{M-1} a_m (\psi_m(\mathbf{x}_i)) - \sum_{m=0}^{M-1} a_m \left( \sum_{j=0}^{J-1} c_{mj} \Psi_j(\mathbf{x}_i) \right) + \sum_{j=0}^{J-1} b_j \Psi_j(\mathbf{x}_i) = \\ &= \sum_{j=0}^{J-1} \underbrace{\left( b_j + \sum_{m=0}^{M-1} a_m c_{mj} \right)}_{:=\alpha_j} \Psi_j(\mathbf{x}_i) + \sum_{m=0}^{M-1} \underbrace{a_m}_{:=\beta_m} (\psi_m(\mathbf{x}_i)) \end{aligned}$$

Therefore, we got that the linear neural network is approximating the exact solution  $u$  in the form of  $\tilde{u} = \sum_j \alpha_j \Psi_j + \sum_m \beta_m \psi_m$ .  $\square$



# Chapter 3

## Experiments

*“It doesn’t matter how beautiful your theory is...  
If it doesn’t agree with experiment, it’s wrong.”*  
– Richard Feynman

### 3.1 Experiment 1: Laplace’s equation with inhomogeneous Dirichlet-type boundary condition

Let

$$\Omega = (-5, 5)^2 \setminus \left( ([-3, -1] \cup [1, 3]) \times [-5/3, 5] \right), \quad (3.1)$$

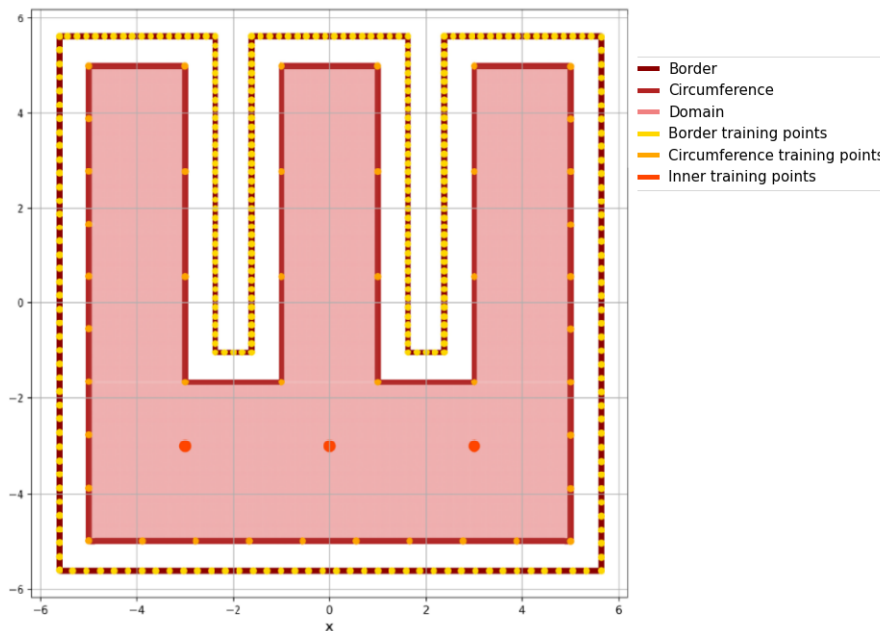


Figure 3.1: The plot of the domain  $\Omega$ , credit: Izsák & Haffner [8]

an open concave polygon (see Figure 3.1), and consider Laplace’s equation

$$\left. \begin{aligned} \Delta u &= 0 \quad (\text{in } \Omega) \\ u(x, y) &= (x + 5)^2 - (y - 2)^2, \quad (\text{on } \partial\Omega). \end{aligned} \right\} \quad (3.2)$$

Notice that  $g$  is the restriction of a harmonic polynomial to the boundary, and therefore the exact solution of the boundary value problem is

$$u(x, y) = (x + 5)^2 - (y - 2)^2 \quad ((x, y) \in \Omega). \quad (3.3)$$

Let us apply the scheme with a primitive neural network discussed in Section 2.2. All the tests were done with  $I = 3$ ,  $J = 95$ ,  $K = 18$ , train-test split ratio = 0.9, batch size 32 (which is  $\approx 15\%$  of the total number of data), and a mean-squared error loss function. Further parameters of the NN include the optimizer method (and its own specific parameters), learning rate, and number of epochs. The approach of [8] achieves the following numerical estimation properties with respect to these three variables. The relative error is calculated as:

$$e_p = \frac{\|(u(\mathbf{x}_0) - \tilde{u}(\mathbf{x}_0), u(\mathbf{x}_1) - \tilde{u}(\mathbf{x}_1), u(\mathbf{x}_2) - \tilde{u}(\mathbf{x}_2))\|_p}{\|(u(\mathbf{x}_0), u(\mathbf{x}_1), u(\mathbf{x}_2))\|_p} \quad (p \in \{1, 2, \infty\}).$$

optimizer	learning rate	epochs	$e_1$	$e_2$	$e_\infty$
Adam	0.1	1,000	0.7358	0.7345	0.7851
Adam	0.01	1,000	$8.4109 \times 10^{-2}$	$7.8732 \times 10^{-2}$	$7.3069 \times 10^{-2}$
Adam	0.01	5,000	$6.9011 \times 10^{-2}$	$7.0638 \times 10^{-2}$	$7.5214 \times 10^{-2}$
Adam	0.01	10,000	$7.9641 \times 10^{-2}$	$7.2514 \times 10^{-2}$	$7.5266 \times 10^{-2}$
Adam	0.001	5,000	$7.005 \times 10^{-3}$	$7.285 \times 10^{-3}$	$7.6783 \times 10^{-3}$
Adam	0.001	10,000	$5.032 \times 10^{-3}$	$4.342 \times 10^{-3}$	$4.018 \times 10^{-3}$
SGD	0.001	1,000	1.2421	0.9722	0.6726
SGD	0.001	10,000	$1.1992 \times 10^{-2}$	$1.21523 \times 10^{-2}$	$1 \times 10^{-2}$

The table above includes the results of some of the experiments. As we can see, the model does in fact seem to approach the desired solution, even if this convergence is not optimal yet. Manuscript [9] shows that the method does indeed approximate the exact solution.

**Theorem 3.1.1.** *Suppose that the points  $Y = \{\mathbf{y}_j\}_{j=0}^{J-1}$  equally-spaced around the domain, and their distance from the domain is above a particular positive constant. Then, the linear combination*

$$\tilde{u}(\mathbf{x}) := \sum_{j=0}^{J-1} a_j \Psi_j(\mathbf{x})$$

provides the approximation rate

$$\min_{\mathbb{R}^J} \left\{ \sup_{\mathbf{x} \in \Omega} |u(\mathbf{x}) - \tilde{u}(\mathbf{x})| + \sup_{\mathbf{x} \in \Omega} |\nabla(u(\mathbf{x}) - \tilde{u}(\mathbf{x}))| \right\} \lesssim h. \quad (3.4)$$

We also have

$$\min_{\mathbb{R}^J} \|u - \tilde{u}\|_{H^1(\Omega)} \lesssim h. \quad (3.5)$$

## 3.2 Experiment 2: Poisson's equation with inhomogeneous Dirichlet-type boundary condition

In this experiment, we solved Poisson's equation with inhomogeneous Dirichlet-type boundary condition on a trapezoidal boundary in two dimensions using the MFS implemented with a linear neural network.

Let

$$\Omega = (0, 5)^2 \setminus \{(x, y) \in \mathbb{R}^2 \mid 5x + 2y \geq 25\} \quad (3.6)$$

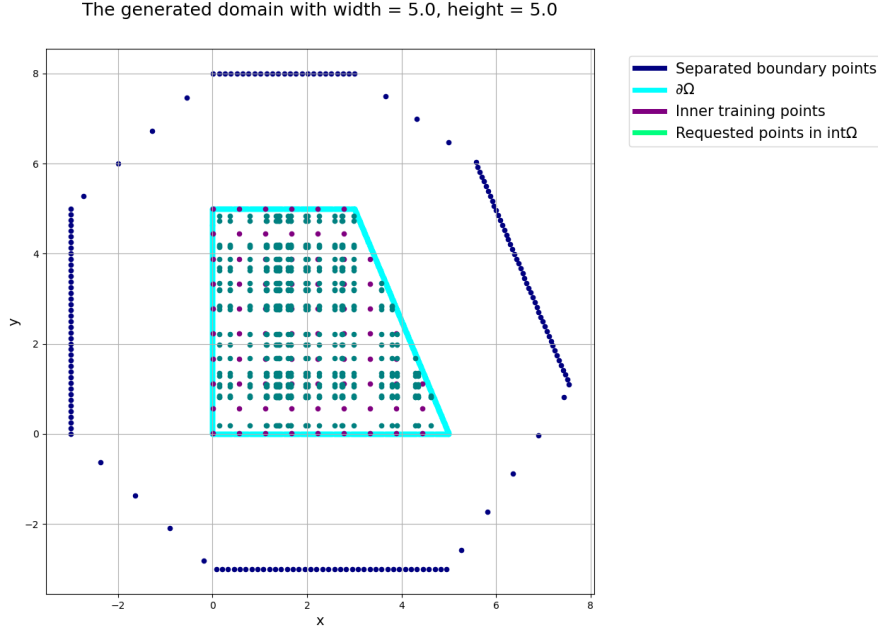
be an open convex polygon as plotted below, and

$$g(x, y) = (x + 5)^2 - (y - 2)^2 \quad ((x, y) \in \partial\Omega) \quad (3.7)$$

a polynomial function restricted to the boundary of  $\Omega$ . Notice that  $g$  is the restriction of a harmonic polynomial function to the boundary, and therefore, the exact solution of the boundary value problem is

$$u(x, y) = (x + 5)^2 - (y - 2)^2 \quad ((x, y) \in \Omega). \quad (3.8)$$

Let us apply the scheme with a primitive neural network discussed in section 2.2.



All the tests were done with  $I = 488$ ,  $J = 169$ ,  $K = 1209$ , train-test split ratio = 0.9. The batch size was 15% of the total number of data, and a mean-squared error loss function was applied. Further parameters of the neural network include the optimizer method (and its own specific parameters), learning rate, and number of epochs.

The approach of [8] achieves the following numerical estimation properties with respect to these three variables. The relative error is calculated as:

$$e_p = \frac{\| (u(\mathbf{x}_i)) - (\tilde{u}(\mathbf{x}_i)) \|_p}{\| (u(\mathbf{x}_i)) \|_p}$$

The following table displays the results of some of the experiments:

optimizer	learning rate	epochs	$e_1$	$e_2$	$e_\infty$
Adam	0.001	10,000	0.413	0.2452	0.1644
Adam	0.0001	10,000	0.234	0.142	0.091

Increasing the batch size proved to be beneficial with respect to the accuracy, therefore, we then increased the batch size to be two thirds of all the training data. We got the following results at this batch size.

optimizer	learning rate	epochs	$e_1$	$e_2$	$e_\infty$
Adam	0.0001	10,000	0.199	0.119	0.090
Adam	0.0001	100,000	0.137	0.077	0.044

As we can see, the model does in fact seem to approach the desired solution, even if this convergence is not optimal yet.

### 3.3 Experiment proposal: Laplace's equation with non-linear boundary conditions

The following experiment proposal is inspired by an application from civil engineering, see [10], Section 3.2.2. As mentioned in Example 1.2.6, the value of the electric potential in a solid piece of metal can be modeled by Laplace's equation with certain types of boundary conditions depending on the situation. Let us consider a rectangular cross-section of a metal beam that is embedded in concrete. Suppose that three sides of the rectangle are electrically insulated, which corresponds to a homogeneous Neumann-type boundary condition on  $\Gamma_0 \subset \partial\Omega$ , but the fourth side is exposed to corrosion. The polarization of the exposed surface results in two nonlinear boundary conditions, one along the positive polarization site (anodic boundary  $\Gamma_a$ ) and one along the negative polarization site (cathodic boundary  $\Gamma_c$ ), as explained in [10] and [11].

Consider the boundary value problem

$$\left. \begin{aligned} \Delta\phi &= 0 && (\text{in } \Omega) \\ \partial_{\mathbf{n}}\phi &= 0 && (\text{on } \Gamma_0) \\ \phi &= \phi_a && (\text{on } \Gamma_a) \\ \partial_{\mathbf{n}}\phi &= g_c(\phi) && (\text{on } \Gamma_c), \end{aligned} \right\} \quad (3.9)$$

where  $\phi$  denotes the electric potential, and

$$g_c(\phi) = \frac{i_L}{\frac{i_L}{i_{OC}} \exp\left(\frac{2.303}{\beta_c}(\phi - \phi_{O_2}) + 1\right)},$$

with  $i_{OC} = 6.25 \cdot 10^{-12} \frac{\text{A}}{\text{mm}^2}$ , which is the cathodic exchange current density,  $\beta_c = 160 \text{ mV}$ , which is the cathodic Tafel slope, and  $\phi_{O_2} = 400 \text{ mV}$ , which is the non-standard half-cell potential of  $O_2$ . For all mathematical purposes, we have

$$g_c(\phi) = \frac{1}{\frac{1}{6.25 \cdot 10^{-12}} \exp\left(\frac{2.303}{160}(\phi - 400) + 1\right)} = 6.25 \cdot 10^{-12} e^{4.7575 - \phi} = \underbrace{\text{const.}}_k e^{-\phi}. \quad (3.10)$$

Ge Ji and O. Burkan Isgor [11] suggest that the ratio of the length of the anodic boundary to the cathodic boundary be around 1 : 9.

We propose the solution of a similar boundary value problem that contains numbers that are easier to deal with. Let  $\Omega := \left(-\frac{3}{2}, \frac{3}{2}\right) \times \left(-\frac{1}{2}, \frac{1}{2}\right)$ ,  $\Gamma_a = \left[-\frac{3}{2}, -\frac{6}{5}\right] \times \left\{-\frac{1}{2}\right\}$ ,  $\Gamma_c = \left[-\frac{6}{5}, \frac{3}{2}\right] \times \left\{-\frac{1}{2}\right\}$ ,  $\Gamma_0 = \partial\Omega \setminus (\Gamma_a \cup \Gamma_b)$ , and  $g_c(\phi) = k \cdot e^{-\phi}$ , as shown below in Figure 3.2.

Setting up the MFS-based solver is not significantly more difficult in this case than in Experiment 1. In fact, since  $\Delta\phi = 0$ , we need not include functions  $\psi_m$  for which  $\Delta\psi_m \neq 0$ .

Let us take a number of external points  $Y = \{\mathbf{y}_j\}_{j=0}^{J-1} \subset \text{ext } \Omega$ , internal points  $X = \{\mathbf{x}_i\}_{i=0}^{I-1}$  and a number of boundary points  $Z = \{\mathbf{z}_k\}_{k=0}^{K-1}$ , the latter of which we distinguish by the type of boundary condition they belong to. Let us define

$$\Psi_j(\mathbf{x}) := \frac{1}{2\pi} \ln \|\mathbf{x} - \mathbf{y}_j\|_2 \quad (i = 0, \dots, J-1),$$

and take a number of  $J$  input-output pairs of the form

$$\underbrace{\left( \left( \Psi_j(\mathbf{z}_a), (-\partial_y \Psi_j(\mathbf{z}_c) - g_c(\Psi_j(\mathbf{z}_c))), (\partial_x \Psi_j(\mathbf{z}_r)), (\partial_y \Psi_j(\mathbf{z}_t)), (-\partial_x \Psi_j(\mathbf{z}_l)) \right) \right)}_{\text{input}_j}, \underbrace{\left( \Psi_j(\mathbf{x}_i) \right)}_{\text{output}_j}, \quad (3.11)$$

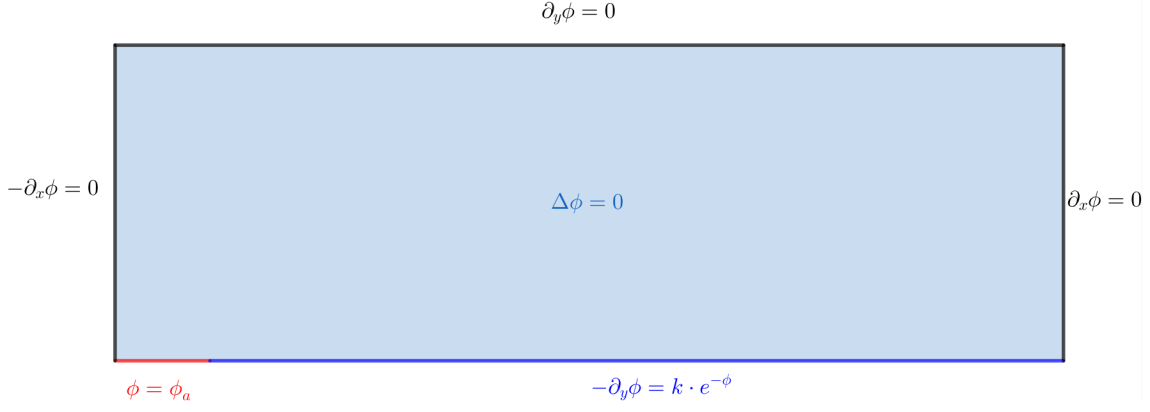


Figure 3.2:  $\Omega$ , and the BVP as defined above

where  $\mathbf{z}_a \in \Gamma_a$ ,  $\mathbf{z}_c \in \Gamma_c$ ,  $\mathbf{z}_r \in \Gamma_r$ ,  $\mathbf{z}_t \in \Gamma_t$ , and  $\mathbf{z}_l \in \Gamma_l$ . Here,  $\Gamma_r, \Gamma_t$ , and  $\Gamma_l$  denotes is the right, top, and left side, respectively.

The structure of the training set defined in 3.11 means that for every function  $\Psi_j$ , we evaluate the right-hand side of the boundary condition, just like we did in the standard MFS case. Note that on the nonlinear part of the boundary, we evaluate the difference between the left-hand side and the right-hand side. This is done so that we are able to evaluate the numerical approximation by setting those coordinates to zero.

In (3.11), we have

$$\begin{aligned}\Psi_j(\mathbf{z}_a) &= \frac{1}{2\pi} \ln \|\mathbf{z}_a - \mathbf{y}_j\|_2 \\ -\partial_y \Psi_j(\mathbf{z}_c) - g_c(\Psi_j(\mathbf{z}_c)) &= -\frac{1}{2\pi} \frac{(\mathbf{z}_c - \mathbf{y}_j)_2}{\|\mathbf{z}_c - \mathbf{y}_j\|_2^2} - k \cdot \|\mathbf{z}_c - \mathbf{y}_j\|_2^{\frac{1}{2\pi}} \\ \partial_x \Psi_j(\mathbf{z}_r) &= \frac{1}{2\pi} \frac{(\mathbf{z}_r - \mathbf{y}_j)_1}{\|\mathbf{z}_r - \mathbf{y}_j\|_2^2} \\ \partial_y \Psi_j(\mathbf{z}_t) &= \frac{1}{2\pi} \frac{(\mathbf{z}_t - \mathbf{y}_j)_2}{\|\mathbf{z}_t - \mathbf{y}_j\|_2^2} \\ -\partial_x \Psi_j(\mathbf{z}_l) &= -\frac{1}{2\pi} \frac{(\mathbf{z}_l - \mathbf{y}_j)_1}{\|\mathbf{z}_l - \mathbf{y}_j\|_2^2}\end{aligned}$$

Let us train a neural network on this training data. Then, the numerical approximation is defined to be

$$(\tilde{u}(\mathbf{x}_i)) := \text{NN}\left((\phi_a(\mathbf{z}_a)), \mathbf{0}_C, \mathbf{0}_R, \mathbf{0}_T, \mathbf{0}_L\right), \quad (3.12)$$

where  $C, R, T$ , and  $L$  are the number of cathodic, right, top, and left boundary collocation points, respectively. Notice that due to the nonlinearity of the boundary condition, we most definitely need to use nonlinear activation functions or biases in the neural network. Otherwise, we would get that given  $\phi_a \equiv 0$ , the function  $\phi \equiv 0$  is a solution of the boundary value problem, which does not satisfy the nonlinear boundary condition.

The verification in this case, is quite challenging because we do not know of any choice  $\phi_a$  and  $k$  for which an exact analytic solution of (3.9) is known. Instead, we suggest the following method of verification. See Figure 3.3 for the geometric setup.

Suppose that the exact solution  $u$  analytically extends to a small neighborhood of  $\Omega$ , and let us include outside points in the set  $X$ , as in Figure 3.3. Since the numerical method builds approximation values in the points of  $X$ , we obtain numerical approximations that are useful for approximating  $\partial_{\mathbf{n}} \tilde{u}$  in boundary points of  $\Omega$ .

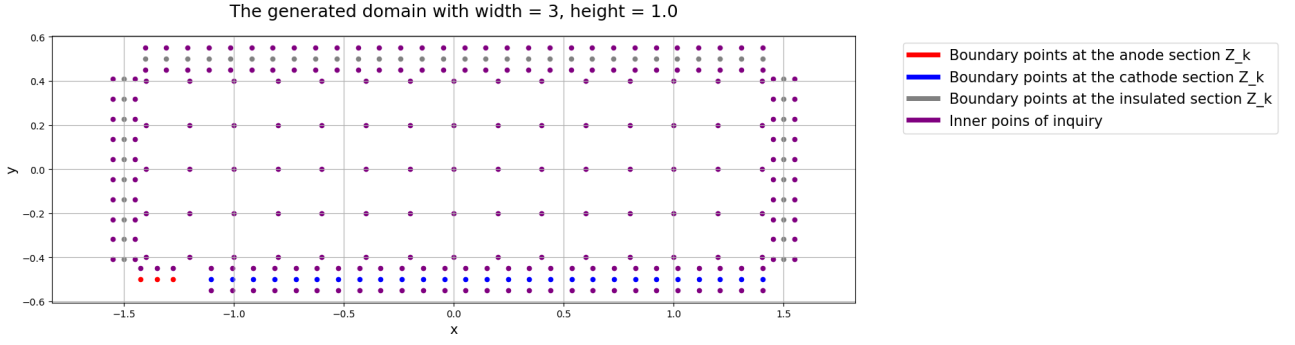


Figure 3.3: The domain and the auxiliary sets in the nonlinear problem

Suppose that  $\mathbf{x} \in X$  and  $\mathbf{x}' \in X$  are each other's reflections to the boundary of  $\Omega$  such that  $\mathbf{x}' \notin \Omega$  and  $\text{dist}(\mathbf{x}, \mathbf{x}') = 2h$ . Then, using the boundary point  $\boldsymbol{\xi} = \frac{\mathbf{x} + \mathbf{x}'}{2}$ ,

- if  $\boldsymbol{\xi} \in \Gamma_c$ , we can approximate  $\phi(\boldsymbol{\xi}) \approx \tilde{\phi}(\boldsymbol{\xi}) := \frac{\tilde{u}(\mathbf{x}') + \tilde{u}(\mathbf{x})}{2}$ , and

$$\partial_{\mathbf{n}} \tilde{u}(\boldsymbol{\xi}) = -\partial_y \tilde{u}(\boldsymbol{\xi}) \approx \frac{\tilde{u}(\mathbf{x}') - \tilde{u}(\mathbf{x})}{2h} \implies e_{c,\boldsymbol{\xi}} \approx \frac{\tilde{u}(\mathbf{x}') - \tilde{u}(\mathbf{x})}{2h} - g_c(\tilde{\phi}(\boldsymbol{\xi}))$$

- if  $\boldsymbol{\xi} \in \Gamma_r$ , we can approximate

$$\partial_{\mathbf{n}} \tilde{u}(\boldsymbol{\xi}) = \partial_x \tilde{u}(\boldsymbol{\xi}) \approx \frac{\tilde{u}(\mathbf{x}') - \tilde{u}(\mathbf{x})}{2h} \implies e_{r,\boldsymbol{\xi}} \approx \frac{\tilde{u}(\mathbf{x}') - \tilde{u}(\mathbf{x})}{2h} - 0$$

- if  $\boldsymbol{\xi} \in \Gamma_t$ , we can approximate

$$\partial_{\mathbf{n}} \tilde{u}(\boldsymbol{\xi}) = \partial_y \tilde{u}(\boldsymbol{\xi}) \approx \frac{\tilde{u}(\mathbf{x}') - \tilde{u}(\mathbf{x})}{2h} \implies e_{t,\boldsymbol{\xi}} \approx \frac{\tilde{u}(\mathbf{x}') - \tilde{u}(\mathbf{x})}{2h} - 0$$

- finally, if  $\boldsymbol{\xi} \in \Gamma_l$ , we can approximate

$$\partial_{\mathbf{n}} \tilde{u}(\boldsymbol{\xi}) = -\partial_x \tilde{u}(\boldsymbol{\xi}) \approx \frac{\tilde{u}(\mathbf{x}') - \tilde{u}(\mathbf{x})}{2h} \implies e_{l,\boldsymbol{\xi}} \approx \frac{\tilde{u}(\mathbf{x}') - \tilde{u}(\mathbf{x})}{2h} - 0.$$

Showing that these error values become negligibly small would constitute evidence that the scheme indeed converges. However, we have not been able to construct a neural network model that achieves this in a timely fashion yet.

# Chapter 4

## Conclusions

*“In science, there are no shortcuts to truth.”*  
– Carl Sagan

### 4.1 Summary

In this thesis, we gave a rather detailed introduction to some of the main concepts and results of analysis, functional analysis, and partial differential equations necessary to understand boundary value problems, and in particular the method of fundamental solutions, such that we are confident that even a reader with minimal numerical analysis training should be able to follow along. The architecture of this theoretical summary focused on the idea of fundamental solutions of operators, as well as function spaces that the modern theory typically deals with. We concluded the first chapter with a brief overview of the basic architecture and working mechanism of artificial neural networks.

In Chapter 2, we turned our attention to the method of fundamental solutions (MFS), devised an educationally motivated classical setup, and ran some tests to convince the reader of the effectiveness of the MFS on simple setups. Then, we proposed a modification of the MFS, found a suitable auxiliary function for its application to Poisson’s equation, and ran some further tests to show that this variety of the MFS also works relatively efficiently on simple models. Finally, we proved that these methods are transferable to machine-learning problems via primitive (linear) neural network structures.

Finally, in Chapter 3, we showed that the transfer proposed in Chapter 2 is feasible. We recreated the test run by Izsák and Haffner [8], and used it as inspiration to build our own test for the modified version of the MFS also. Finally, we devised a test for a real-world-application-inspired Laplace’s equation with nonlinear mixed boundary conditions and ran a few tests that did not yield satisfactory results.

### 4.2 Further directions

In light of the successful outcomes and insights gained from the theoretical discussions on the MFS and the modified version of the MFS, the successful implementations in Experiment 1 and Experiment 2, as well as being in possession of the algorithm that can run Experiment 3, we wish to see the last experiment through, find the variables that make it work at least as well as the algorithm in Experiment 2. The success of Experiment 3 should conceivably provide us with a new frontier to attacking non-linear boundary-value problems. We also believe that an important direction of generalization could be to implement some variation of these algorithms to time-dependent problems. We also note that from Theorem 2.2.1 and Theorem 2.4.1, it

is clear that proofs on the classical MFS will be immediately applicable via artificial neural networks, and therefore the classical analysis of the MFS method remains relevant in this sense.



# Bibliography

- [1] Charu C Aggarwal. Neural Networks and Deep Learning. en. 1st ed. Cham, Switzerland: Springer International Publishing, Sept. 2018.
- [2] Ádám Besenyei, Vilmos Komornik, and László Simon. Parciális differenciálegyenletek. Typotex, 2013. URL: [http://etananyag.ttk.elte.hu/Files/downloads/\\_Besenyei\\_Parc\\_diffegyenlet.pdf](http://etananyag.ttk.elte.hu/Files/downloads/_Besenyei_Parc_diffegyenlet.pdf).
- [3] Nicolas Bourbaki. Integration I. en. Berlin, Germany: Springer, Feb. 2012.
- [4] Alexandre Ern and Jean-Luc Guermond. Theory and practice of finite elements. en. Applied Mathematical Sciences. New York, NY: Springer, Apr. 2004.
- [5] Lawrence C Evans. Partial differential equations. en. Providence, RI: American Mathematical Society, Mar. 2022.
- [6] Graeme Fairweather and Andreas Karageorghis. “The method of fundamental solutions for elliptic boundary value problems”. In: Advances in Computational Mathematics 9 (Sept. 1998), pp. 69–95. DOI: 10.1023/A:1018981221740. URL: <https://doi.org/10.1023/A:1018981221740>.
- [7] Alexander H.D. Cheng and Yongxing Hong. “An overview of the method of fundamental solutions—Solvability, uniqueness, convergence, and stability”. In: Engineering an. w b.e. 120 (2020), pp. 118–152. ISSN: 0955-7997. DOI: <https://doi.org/10.1016/j.enganabound.2020.08.013>. URL: <https://www.sciencedirect.com/science/article/pii/S0955799720302150>.
- [8] Domonkos Haffner and Ferenc Izsák. “Solving the Laplace equation by using neural networks”. In: URL: <https://m2.mtmt.hu/api/publication/32625402>.
- [9] T. Hieu Hoang, Ferenc Izsák, and Gábor Maros. “Approximation properties of fundamental solutions: a three-dimensional study with Sobolev norms”. In: proceedings.
- [10] Ge Ji. On the num. sol. of Laplace’s eq. w nonlin. b.c. for corrosion of steel in concrete. en. 2006.
- [11] Ge Ji and O. Isgor. “On the numerical solution of Laplace’s equation with nonlinear boundary conditions for corrosion of steel in concrete”. In: (Jan. 2006). URL: [https://www.researchgate.net/publication/228834887\\_ON\\_THE\\_NUMERICAL\\_SOLUTION\\_OF\\_LAPLACE’S\\_EQUATION\\_WITH\\_NONLINEAR\\_BOUNDARY\\_CONDITIONS\\_FOR\\_CORROSION\\_OF\\_STEEL\\_IN\\_CONCRETE](https://www.researchgate.net/publication/228834887_ON_THE_NUMERICAL_SOLUTION_OF_LAPLACE’S_EQUATION_WITH_NONLINEAR_BOUNDARY_CONDITIONS_FOR_CORROSION_OF_STEEL_IN_CONCRETE).
- [12] Vilmos Komornik. A simplified construction of the Lebesgue integral. 2018. DOI: 10.48550/ARXIV.1805.07289. URL: <https://arxiv.org/abs/1805.07289>.
- [13] Enrico Masina. Useful review on the Exponential-Integral special function. 2019. eprint: arXiv:1907.12373.
- [14] Stefan A. Sauter and Christoph Schwab. Boundary Element Methods. Springer Berlin Heidelberg, 2011. DOI: 10.1007/978-3-540-68093-2. URL: <https://doi.org/10.1007/978-3-540-68093-2>.

- [15] M. A. Aleksidze V. D. Kupradze. “An approximate method of solving certain boundary-value problems (in Russian)”. In: Soobšč. Akad. Nauk Gruzin. SSR 30 (1963), pp. 529–536.
- [16] Kurt Hornik, Maxwell Stinchcombe, Halbert White. “Multilayer Feedforward Networks are Universal Approximators”. In: Neural Networks Vol. 2 (1989), pp. 359–366.
- [17] Michel Willem. Functional analysis. en. 2013th ed. Cornerstones. Secaucus, NJ: Birkhauser Boston, Aug. 2013.