LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

MASTER

MATHEMATICAL FINANCE

MASTER´S FINAL WORK

DISSERTATION

AN INTRODUCTION TO DISTRIBUTION THEORY, FOURIER TRANSFORM, SOBOLEV SPACES AND ITS APPLICATIONS TO THE BLACK-SCHOLES EQUATION

BRENO LUCAS DA COSTA GONÇALVES

OCTOBER - 2019

LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

MASTER

MATHEMATICAL FINANCE

MASTER´S FINAL WORK

DISSERTATION

AN INTRODUCTION TO DISTRIBUTION THEORY, FOURIER TRANSFORM, SOBOLEV SPACES AND ITS APPLICATIONS TO THE BLACK-SCHOLES EQUATION

BRENO LUCAS DA COSTA GONÇALVES

SUPERVISOR:
FILIPE SERRA DE OLIVEIRA

OCTOBER - 2019

Acknowledgements

I want to express my full gratitude and appreciation to all those who, directly or indirectly, have contributed to this work. To all, my sincere recognition.

First, I thank my supervisor, Professor Filipe Oliveira, for the challenge proposed, availability, support, bits of advice and patience. Thank you for being a true mentor during this project.

To FCT and its professors, namely Professor Elvira Coimbra, Fernanda Cipriano, Júlia Vaz de Carvalho and Luís Trabucho, for the inspiration, correction and privilege that was to attend their classes.

To ISEG for the opportunities offered. To Professor Maria do Rosário Grossinho and my colleagues Daniel de Alcântara, Inês Rodrigues and Maria Inês Correia.

The friendships I have created and I will always take with me. A special word to Luís Marques and to the initial Grossos: André Sá, Marco António, Miguel Catela, Ricardo Correia and Tomás Lópes. To everyone who joined later and now after six years are part of this family. To Catarina Félix, Flávia Moreira and special to Alina Teles, for being present friends since my first day in the Bachelors.

To Carla Tavares who lighted up the bad moments and made the good ones even brighter. Thank you for the company, love and caring!

Finally, I thank my parents, for the strength, unconditional support, encouragement, friendship, love and patience demonstrated and total help to overcome all obstacles in my life. This final work is dedicated to them!

Resumo

A maior parte da teoria citada em cursos introdutórios de análise matemática e cálculo foi elaborada ainda antes do século XVIII. No entanto, em meados de 1950, *Laurent Schwartz* desenvolveu um novo conceito que mudou e redefiniu a noção de função.

A definição de derivada e transformada de Fourier são dois dos conceitos mais importantes em análise e essenciais para o estudo de equações diferenciais parciais. No entanto, note-se que nem todas as funções são diferenciáveis ou possuem uma transformada de Fourier. A teoria de distribuições permite esta correção ao incorporar as funções clássicas numa classe maior de objetos matemáticos.

Esta dissertação tem como objetivo completar o artigo publicado por *D. da Silva, K. Igibayeva, A. Khoroshevskay* e *Z.Sakayeva* [17]. Em particular, pretende-se demonstrar que, em certas condições, para uma dada função $F : \mathbb{R} \to \mathbb{R}$, o problema $u_t - u_{xx} = u_x + F(u)$ com a condição inicial $u(x, 0) = f(x)$ é localmente bem posto no espaço de Sobolev $\mathbb{H}^1$. De modo a alcançar o nosso objectivo apresentam-se inicialmente as ferramentas necessárias para uma introdução à teoria de distribuições, transformadas de Fourier e espaços de Sobolev que serão usadas para o cálculo explícito da solução da equação de difusão do calor.

Palavras-chave: Funções Teste, Distribuições, Transformadas de Fourier, Difusão do Calor, Espaços de Sobolev, Equação de Black-Scholes.

Abstract

Most of the mathematical theory study in standard courses of calculus were developed even before the eighteen century. However, around the year of 1950, *Laurent Schwartz* came up with a new concept that would change and redefine the concept of function.

Differentiability and the Fourier transform are two of the most important notions in analysis and genuinely essential when working with partial differential equations. It is well-known that not all functions are differentiable or have a Fourier transform. The theory of distributions allows us to correct this issue by embedding classical functions into a larger class of mathematical objects.

This dissertation aims to complete the article published by *D. Da Silva, K. Igibayeva, A. Khoroshevskay* and *Z.Sakayeva (2018)* [17]. In particular, we intend to prove that under certain conditions on the real-valued functions f and F, the problem $u_t - u_{xx} = u_x + F(u)$ with initial condition $u(x,0) = f(x)$ is locally well-posed in the Sobolev space $\mathbb{H}^1$. To accomplish our goal, we provide and develop the necessary tools for an introductory course in distribution theory, Fourier transforms, Sobolev spaces and use them to solve the heat diffusion equation.

Keywords: Test functions, Distributions, Fourier transform, Heat Diffusion, Sobolev Spaces, Black-Scholes Equation.

Contents

List of Figures

Chapter 1

Introduction

Although the concept of function was created a long time ago, it is still today one of the most crucial notions in science. A partial differential equation, PDE for future references, as the name indicates, is a equation that expresses relationships between a function and its partial derivatives. In this thesis we concern ourselves with linear second-order differential equations of parabolic type, mainly with constant coefficients. That is, an equation of the type

$$Au_{xx} + 2Bu_{xt} + Cu_{tt} + au_x + bu_t + cu = 0 \qquad (A, B, C) \neq (0, 0, 0),$$

that verifies

$$B^2 - 4AC = 0$$

for $A, B \cdots c$ real numbers. For the most curious reader, if $B^2 - 4AC < 0$ then the above PDE is said to be of the *elliptic type*, while if $B^2 - 4AC > 0$ then it said to be of the *hyperbolic type*.

Over the last two centuries, PDEs proved to be a powerful tool for solving an enormous array of problems in biological, earth, social and surely physical sciences. Here, we will concentrate on the dynamic motion of heat. However, the study of such notion cannot be made without the name of Fourier. The conditions that led to the development of the heat equation and the reasons why it had a significant influence is itself instructive. Given that, it is of interest to understand a bit of the historical evolution.

Let us dive into the scientific reality of the eighteen century. Given the nature of the heat equation, notions such as temperature, heat transport and heat-temperature ratio were crucial. We shall remember that when Fourier began his work, such notions were still in a development phase.

The first step that engaged the heat conduction study occur ed in 1714 with Gabriel Fahrenheit (1686-1736). He developed an instrument with the capability of observe temperature changes with rigour (the closed-tube mercury thermometer) and also introduced the first scale for temperature measurement (the Fahrenheit scale).

Joseph Black (1728-1799) noticed that, when ice melts, it loses heat without changing temperature. This observation defined terms such as *latent heat* (thermal energy released or absorbed by a system, during a constant-temperature process) and *specific heat* (heat capacity per unit mass of material). In plain English, different materials need a different quantity of heat to raise their temperature by the same value.

Alongside Black's work, in 1783, one of the most important papers of modern chemistry and thermodynamic was published by Antoine Lavoisier (1743-1794) and Pierre Laplace (1749-1827). Named *"Mémoire sur la Chaleur"*, this paper provided details on the study of latent heat of ice in a melting process and specific heat measurements of different materials.

Jean Baptiste Biot (1774-1862) directed the problem of heat conduction in a thin bar heated at one end, believing that the temperature of a point in the bar was influenced by the temperature of all the points of its neighbourhood. In Fourier's work the distance between the points was also considered. This allowed involving temperature gradients and the formulation of a differential equation modelling this phenomena.

From the mathematical point of view, science witnessed the development of differential equations by Laplace and George Green (1793-1841), with significant influence of Joseph-Louis Lagrange (1739-1813), Leonhard Euler (1707-1783), Jean d'Alembert (1717-1783) and Daniel Bernoulli (1700-1782), having the latter presented the solution in terms of trigonometric series of the harmonic movement equation.

With all this baggage, Fourier started his work around 1803 and soon began the innovations. He conceived the problem in terms of three ingredients — heat transport in space, heat storage and boundary conditions. The differential equation is only valid inside its domain. The relationship with the exterior is performed by the boundaries and the initial temperature is known a priori. Fourier formulated the equation linearly. The parameters involved were constants, not depending on the time, position or the temperature itself.

He tried to submit his first paper entitled *"Théorie de la Propagation de la Chaleur dans les Solides"* in 1807. However, it was in 1822 with the publication of his masterpiece *"Théorie Analytique de la Chaleur"* that his work took a new dimension.

Shortly after the publication of his work Fourier influenced several fields of physics as, for instance, electricity and magnetism with names like Geor Ohm (1787-1854), Michael Faraday (1791 - 1867) and James Clerk Maxwell (1831-1879). In particular, Ohm assumed the flow of electricity as precisely analogous to the flow of heat.

In the later 19th, earlier 20th century, the problem of heat conduction was extended to a new area with the introduction of the concept of "randomness". Such concept led to the area of stochastic processes. This new science was based on four prominent names. John Strutt (1842-1919), Francis Edgeworth (1845-1926), Louis Bachelier (1870-1946) and Albert Einstein (1879-1955). Paul Langevin (1872-1946) also had a significant influence in this subject with formulation of the stochastic differential equations. Strutt (theory of sound) was the one to make the bridge between what is deterministic and what is stochastic. Edgeworth (law of error) together with Bachelier (theory of speculation) played the leading role in the introduction of mathematical finance. Edgeworth was the one who combined stochastic theory and social economics analysis. Bachelier proved that option prices could be described as diffusion equations with the right assumption on the randomness of the stock prices. His work motivated many others, for example Paul Samuelson (1915-1994) and one of the most famous and widely used models in the academical environment introduced by Fisher Black (1938-1995), Myron Scholes and Robert Merton. Einstein and his brilliant mind took all of these concepts and expanded them to a different level with the molecular-kinetic theory of heat.

Now, having understood a bit of the history of PDEs and some of their applications, we come across with the unpleasant fact that not every function is differentiable. Laurent Schwartz (1915-2002) and Sergei Sobolev (1908-1989) came up with an elegant solution.

Along with Émile Borel (1871-1956) and Henri Lebesgue (1875 -1941) measure theory, Schwartz distribution theory was one of the major evolution in mathematics on the 20th century. The motivation came largely from the limitations of classical calculus. Quoting Treves, "Schwartz gave a strong functional analysis slant to the theory $(\cdots)$. But the enduring merit of distribution theory has been the basic operations of analysis, differentiation and convolution, and the Fourier/Laplace transforms and their inversion, which demanded so much care in the classical framework, could now be carried out without qualms by obeying purely algebraic rules."

The aim of this thesis is, therefore, introduce the reader to distribution theory, Fourier transforms and Sobolev spaces, show how to overcome the major flaws of classical analysis and also present an alternative method to solve the Cauchy problem of heat diffusion. Plus, we focus on the article published by *D. Da Silva, K. Igibayeva, A. Khoroshevskay* and *Z.Sakayeva*. The authors proved on Theorem 1 that, under certain assumptions on the real-valued functions f and F, the problem $u_t - u_{xx} = u_x + F(u)$ with initial condition $u(x, 0) = f(x)$ is locally well-posed in $\mathbb{L}^p$ for $p \geq 1$. Our approach consists in using the results obtained in the main chapters, as well as the results proved by the authors in the article above [17] to prove the aforementioned theorem for the Sobolev space $\mathbb{H}^1$.

Chapter 2

Distribution Theory in $\mathbb{R}^n$

2.1 Test Functions

The main idea that inspires the study of distributions is to generalize the notion of function used in classical calculus. Such objects, the distributions, will be defined as continuous linear maps from the so-called test functions to $\mathbb{R}$. With this in mind and before going any further with the description and analysis of our main subject, a few concepts concerning these test functions are introduced. They are essential to the comprehension of the remaining text. Only the major results will be highlighted. For a more detailed study, the reading of standard textbooks as [7], [10] and [15] is recommended. For future references, Ω is an arbitrary open subset of $\mathbb{R}^n$.

Definition 2.1.1. Let ϕ: $\Omega \to \mathbb{R}$ a real-valued function. The *support* of ϕ, denoted by *supp* ϕ, is the closure in Ω of the subset formed by the points where ϕ does not vanish:

$$\mathit{supp}\,\phi = \overline{\left\{x \in \mathbb{R}^n : \phi(x) \neq 0\right\}}. \tag{2.1}$$

A *test function* on Ω is an infinitely differentiable real-valued function on Ω whereof support is a compact subset of Ω. The linear space of all test functions on Ω is denoted by $C_0^\infty(\Omega)$ or $\mathscr{D}(\Omega)$.

The subscript 0 is to emphasize the fact that outside the support, the value of the function is zero. A few trivialities about the support of a function are described in the following lemma.

Lemma 2.1.1. *Let f, g continuous functions on Ω. We have*

- $f \equiv 0 \Leftrightarrow \operatorname{supp} f = \emptyset$;
- $\operatorname{supp} f$ *is closed in* Ω;
- $\operatorname{supp}(fg) \subseteq \operatorname{supp} f \cap \operatorname{supp} g$;
- *If h is a differentiable function then* $\operatorname{supp} h' \subset \operatorname{supp} h$.

Now, a question arises naturally: is the null function the only one to satisfy both prerequisites (to be compactly supported and infinitely differentiable)? All smooth functions that might come promptly to our mind, e.g. polynomials, trigonometric and exponential functions, fail to be test functions, since they do not have compact support. As a matter of fact, due to the analytic uniqueness continuation theorem, no other analytic function, rather than $\phi \equiv 0$, belongs to C_0^∞. Yet, the answer of the above question is negative. The following Cauchy's celebrated example is often used to proved that the space of test functions is, indeed, non trivial.

Lemma 2.1.2. *The function $\alpha : \mathbb{R} \to \mathbb{R}$ defined by*

$$\alpha(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ e^{-\frac{1}{x}}, & \text{if } x > 0 \end{cases}$$

belongs to $C^\infty(\mathbb{R})$.

Proof. It should be obvious that the only issue here is the differentiability at $x = 0$. One may check by induction on the order of derivative $n \in \mathbb{N}_0$, that for $x > 0$

$$\alpha^{(n)}(x) = \frac{e^{-\frac{1}{x}} P_{n-1}(x)}{x^{2n}},$$

where $P_{-1} = 1$ and P_{n-1} is a polynomial of degree $n - 1$. By simple use of calculus,

$$\lim_{x \to 0^+} \alpha^{(n)}(x) = \lim_{x \to 0^-} \alpha^{(n)}(x) = \alpha^{(n)}(0) = 0.$$

■

The above lemma guarantees the existence of non-negative test functions such that the value at zero is positive. Furthermore, for $a, b \in \mathbb{R}$ such that $a \leq b$, the β function defined by

$$\beta(x) = \beta_{a,b}(x) := \alpha(x-a)\alpha(b-x)$$

is infinitely differentiable with *supp* $\beta = [a, b]$. β is, therefore, a test function. Similarly, more test functions can be built based on this β function. In fact, it is not hard to build test functions ϕ such that for any given $a \in \mathbb{R}^n$, $r > 0$:

- $0 \leq \phi \leq 1$,
- $\phi \equiv 1$ in $B_a(r)$,
- $\phi \equiv 0$ in $B_a(2r)^c$.

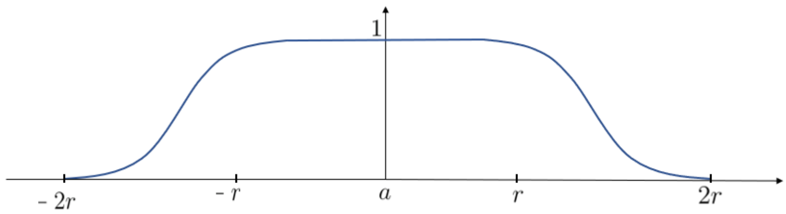

Figure 2.1. *An one-dimensional representation of a test function.*

Let us introduce some basic notations required for the following definition and future chapters. As it is somewhat standard in the literature, a *multi-index* is a sequence

$$\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$$

of n non-negative integers. The sum:

$$|\alpha| := \sum_{j=1}^{n} \alpha_j$$

is called *order or length of the multi-index*. Plus, for a given multi-index α, we define the derivative

$$\partial_x^\alpha := \frac{\partial^\alpha}{\partial x^\alpha} = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}, \text{ where } \partial_j := \frac{\partial}{\partial x_j}.$$

The concept of convergence in the space of test functions is at least vital for the study of distributions.

Definition 2.1.2. Let $(\phi_i)_{i\in\mathbb{N}}$ a sequence of $C_0^\infty(\Omega)$ and $\phi \in C_0^\infty(\Omega)$. *The sequence* $(\phi_i)_{i\in\mathbb{N}}$ *converges to* ϕ *in* $C_0^\infty(\Omega)$ if both conditions are satisfied:

1. There exists a fix compact subset K of Ω such that $\forall i$ *supp* $\phi_i \subseteq$ K,

2. For every multi-index α, the sequence $(\partial^\alpha \phi_i)_{i\in\mathbb{N}}$ converges uniformly to $\partial^\alpha \phi$ (i.e $\|\partial_x^\alpha \phi_i - \partial_x^\alpha \phi\|_\infty \to 0$).

From the definition, we can deduce that *supp* $\phi \subseteq$ K. Generally speaking, to define which sequences converge in $C_0^\infty(\Omega)$ is equivalent to define a topology in this space. Due to its great complexity, it could be a dissertation theme itself. For such a reason, we will not analyze it in detail in this work. For us, it will be enough to know which sequences converge. As a reference for this subject, we suggest the reading of [20].

2.2 Distributions

All the challenges in the theory of partial differential equations, harmonic and Fourier analysis force the extension of the usual functions to distributions. In this section, we present the formal definition of distribution, its main properties and some of the most common examples.

Definition 2.2.1. A linear functional $T : C_0^\infty(\Omega) \to \mathbb{R}$ is said to be a *distribution* if and only if, for every $\phi \in C_0^\infty(\Omega)$

$$\forall K \subset\subset \Omega, \exists C = C(K) : |T(\phi)| \leq C(K) \sup_{|\alpha|\leq p} \|\partial^\alpha \phi\|_\infty. \tag{2.2}$$

If there exists a smallest $p_0 \in \mathbb{N}_0$, independent of K, such that (2.2) holds with $p = p_0$, p_0 is called the *order of the distribution* T.

Here, the double contained symbol is to highlight the fact that K is a compact set in Ω. Actually, condition (2.2) is equivalent to the continuity of T if one provides the topology defined by definition 2.1.2 (see, for instance, Proposition 1.2.2 of [13]). Schwartz himself used the notation $\mathscr{D}(\Omega)$ to denote the space of test functions on Ω with the notion of convergence mentioned above. The space of distributions is, in truth, the topological dual of $\mathscr{D}$, represented by $\mathscr{D}^*$. As it is common in most of the literature, we will employ the notation $\langle T, \phi\rangle$ to represent $T(\phi)$, the value of T at ϕ. Hence we have

Theorem 2.2.1. *A linear functional T on $\mathscr{D}(\Omega)$ is a distribution if and only if for any sequence ϕ_i that converges to ϕ in $\mathscr{D}(\Omega)$ implies the convergence of $\langle T, \phi_i\rangle$ to $\langle T, \phi\rangle$ in $\mathbb{R}$ (by linearity, it is enough to require $\phi_i \to 0$ in $\mathscr{D}(\Omega)$ implies $\langle T, \phi_i\rangle \to 0$ in $\mathbb{R}$).*

Example 2.2.1. The *Dirac distribution* (δ-function) at $x \in \mathbb{R}$. The functional

$$\langle \delta_x, \phi\rangle := \phi(x) \tag{2.3}$$

defines a distribution of order 0.

Proof. For $a \in \mathbb{R}$ and $\phi, \gamma \in \mathscr{D}(\mathbb{R})$:

- Linearity: $\langle \delta_x, a(\phi + \gamma)\rangle := a(\phi + \gamma)(x) = a\phi(x) + a\gamma(x) = a\langle \delta_x, \phi\rangle + a\langle \delta_x, \gamma\rangle$;
- Continuity: $|\langle \delta_x, \phi\rangle| := |\phi(x)| \leq \|\phi\|_\infty$.

∎

Several functions can be seen as distributions but not all the distributions are functions. In measure theory, a function $f\colon\ \Omega \to \mathbb{R}$ is *locally integrable* (point-function) if it is measurable and for all compact $K \subset \Omega$

$$\int_K |f(x)| d(x) < \infty.$$

The space of locally integrable functions on Ω is denoted by $\mathbb{L}^1_{loc}(\Omega)$.

Theorem 2.2.2. *Regular distributions: For any $f \in \mathbb{L}^1_{loc}(\Omega)$, the functional*

$$\langle T_f, \phi \rangle := \int_\Omega f(x)\phi(x)dx \tag{2.4}$$

defines a distribution of order 0.

Proof. For $a \in \mathbb{R}$ and $\phi, \gamma \in \mathscr{D}(\Omega)$:

- Linearity:

$$\begin{aligned}\langle T_f, a(\phi + \gamma)\rangle &:= \int_\Omega f(x)a(\phi(x) + \gamma(x))dx \\ &= a \int_\Omega f(x)\left(\phi(x) + \gamma(x)\right)dx \\ &= a \int_\Omega f(x)\phi(x)dx + a \int_\Omega f(x)\gamma(x)dx \\ &= a\langle T_f, \phi\rangle + a\langle T_f, \gamma\rangle;\end{aligned}$$

- Continuity: Assume that $supp\, \phi \subset K$. Then,

$$|\langle T_f, \phi\rangle| := \Big|\int_\Omega f(x)\phi(x)dx\Big| \leq \int_K |f(x)\phi(x)|dx \leq \underbrace{\int_K |f(x)|dx}_{C(K)} \|\phi\|_\infty.$$

■

The previous theorem has embedded several results. The map $f \in \mathbb{L}^1_{loc}(\Omega) \to T_f \in \mathscr{D}^*(\Omega)$ is injective. Indeed, if $\displaystyle\int f\phi\, dx = \int g\phi\, dx$ for all $\phi \in \mathscr{D}(\Omega)$ then $f = g\, a.e.$ For this reason it is common to identify f with the associated distribution T_f, writing $\langle T_f, \phi\rangle = \langle f, \phi\rangle$. Also with this identification, $\mathbb{L}^1_{loc}(\Omega) \subset \mathscr{D}^*(\Omega)$. For any test function ϕ, the map $\phi \to \displaystyle\int \phi(x)dx$ defines a distribution of order 0. Furthermore, by means of the Hölder inequality, for any $p > 1$ one has $\mathbb{L}^p_{loc}(\Omega) \subseteq \mathbb{L}^1_{loc}(\Omega)$. As a consequence, (2.4) holds for any $\mathbb{L}^p_{loc}$ function. In [12], we can find a very instructive inclusion scheme. For $1 \leq p \leq q < \infty$, the following inclusions hold:

$$\begin{array}{ccccccccccc}
L^\infty_c(\Omega) & \subset & L^q_c(\Omega) & \subset & L^p_c(\Omega) & \subset & L^1_c(\Omega) & \subset & \mathcal{M}_c(\Omega) & & \\
\cap & & \cap & & \cap & & \cap & & \cap & & \\
L^\infty(\Omega) & & L^q(\Omega) & & L^p(\Omega) & & L^1(\Omega) & \subset & \mathcal{M}^1(\Omega) & & \\
\cap & & \cap & & \cap & & \cap & & \cap & & \\
L^\infty_{loc}(\Omega) & \subset & L^q_{loc}(\Omega) & \subset & L^p_{loc}(\Omega) & \subset & L^1_{loc}(\Omega) & \subset & \mathcal{M}(\Omega) & \subset & \mathcal{D}'(\Omega)
\end{array}$$

Figure 2.2. *Subsets of* $\mathscr{D}^*(\Omega)$

Here, $\mathcal{M}, \mathcal{M}^1$ and $\mathcal{M}_c$ denote the spaces of Radon measures, bounded measures and measures with compact support on Ω respectively.

Now, we can look into some more exotic examples.

Example 2.2.2. *Cauchy Principal Value.* The functional $vp\frac{1}{x} : \mathscr{D}(\mathbb{R}) \to \mathbb{R}$

$$\langle vp\frac{1}{x}, \phi\rangle := \lim_{\epsilon\to 0}\int_{|x|\geq\epsilon}\frac{\phi(x)}{x}dx \tag{2.5}$$

defines a distribution of order 1.

Proof. Let $a \in \mathbb{R}$ and $\phi, \gamma \in \mathscr{D}(\mathbb{R})$:

- Linearity:

$$\begin{aligned}
\langle vp\frac{1}{x}, a(\phi+\gamma)\rangle &:= \lim_{\epsilon\to 0}\left(\int_{|x|\geq\epsilon}\frac{a(\phi(x)+\gamma(x))}{x}dx\right) \\
&= \lim_{\epsilon\to 0}\left(\int_{|x|\geq\epsilon} a\frac{\phi(x)}{x}dx\right) + \lim_{\epsilon\to 0}\left(\int_{|x|\geq\epsilon} a\frac{\gamma(x)}{x}dx\right) \\
&= a\langle vp\frac{1}{x}, \phi\rangle + a\langle vp\frac{1}{x}, \gamma\rangle.
\end{aligned}$$

- Continuity: Assume that *supp* $\phi \subseteq K \subset [-M, M]$. By Taylor's formula

$$\phi(x) = \phi(0) + x\phi'(\theta), \qquad \theta \in]0, x[.$$

Hence,

$$\int_{|x|\geq\epsilon}\frac{\phi(x)}{x}dx = \int_{\epsilon\leq|x|\leq M}\frac{\phi(x)}{x}dx = \underbrace{\int_{\epsilon\leq|x|\leq M}\frac{\phi(0)}{x}dx}_{0} + \int_{\epsilon\leq|x|\leq M}\phi'(\theta)dx$$

Moreover, using for example Lebesgue's dominated convergence theorem, we have

$$\langle vp\frac{1}{x}, \phi\rangle = \lim_{\epsilon\to 0}\int_{\epsilon\le|x|\le M}\phi'(\theta)dx = \int_{|x|\le M}\phi'(\theta)dx = \int_K \phi'(\theta)dx.$$

Therefore,

$$\left|\langle vp\frac{1}{x}, \phi\rangle\right| = \Big|\int_K \phi'(\theta)dx\Big| \le \mu(K)\|\phi'(x)\|_\infty \le C(K)\sup_{p\le 1}\|\partial^p\phi\|_\infty.$$

■

Example 2.2.3. *Finite Part of* $\frac{1}{x^2}$. The functional $fp\dfrac{1}{x^2} : \mathscr{D}(\mathbb{R}) \to \mathbb{R}$

$$\langle fp\frac{1}{x^2}, \phi\rangle := \lim_{\epsilon\to 0}\left(\int_{|x|\ge\epsilon}\frac{\phi(x)}{x^2}dx - 2\frac{\phi(0)}{\epsilon}\right) \tag{2.6}$$

defines a distribution of order 2.

Proof. For $a \in \mathbb{R}$ and $\phi, \gamma \in \mathscr{D}(\mathbb{R})$:

- Linearity:

$$\begin{aligned}\langle fp\frac{1}{x^2}, a(\phi+\gamma)\rangle &:= \lim_{\epsilon\to 0}\left(\int_{|x|\ge\epsilon}\frac{a(\phi(x)+\gamma(x))}{x^2}dx - 2\frac{a(\phi(0)+\gamma(0))}{\epsilon}\right)\\ &= \lim_{\epsilon\to 0}\left(\int_{|x|\ge\epsilon}\frac{a\phi(x)}{x^2} + \frac{a\gamma(x)}{x^2}dx - 2\frac{a\phi(0)}{\epsilon} - 2\frac{a\gamma(0)}{\epsilon}\right)\\ &= a\cdot\lim_{\epsilon\to 0}\left(\int_{|x|\ge\epsilon}\frac{\phi(x)}{x^2}dx - 2\frac{\phi(0)}{\epsilon}\right) + a\cdot\lim_{\epsilon\to 0}\left(\int_{|x|\ge\epsilon}\frac{\gamma(x)}{x^2}dx - 2\frac{\gamma(0)}{\epsilon}\right)\\ &= a\langle fp\frac{1}{x^2}, \phi\rangle + a\langle fp\frac{1}{x^2}, \gamma\rangle.\end{aligned}$$

- Continuity: Assume that *supp* $\phi = K \subset [-M, M]$. By Taylor's formula

$$\phi(x) = \phi(0) + x\phi'(0) + \frac{x^2}{2}\phi''(\theta), \qquad \theta \in]0, x[.$$

  Hence,

$$\begin{aligned}\int_{|x|\ge\epsilon}\frac{\phi(x)}{x^2}dx - 2\frac{\phi(0)}{\epsilon} &= \int_{\epsilon\le|x|\le M}\frac{\phi(x)}{x^2}dx - 2\frac{\phi(0)}{\epsilon}\\ &= \int_{\epsilon\le|x|\le M}\frac{\phi(0)}{x^2} + \frac{x\phi'(0)}{x^2} + \frac{x^2\phi''(\theta)}{2x^2}dx - 2\frac{\phi(0)}{\epsilon}\\ &= \int_{\epsilon\le|x|\le M}\frac{\phi(0)}{x^2}dx + \underbrace{\int_{\epsilon\le|x|\le M}\frac{\phi'(0)}{x}dx}_{0}\\ &+ \frac{1}{2}\int_{\epsilon\le|x|\le M}\phi''(\theta)dx - 2\frac{\phi(0)}{\epsilon}.\end{aligned}$$

But,

$$\int_{\epsilon\leq|x|\leq M}\frac{\phi(0)}{x^2}dx=\phi(0)\int_{\epsilon\leq|x|\leq M}\frac{1}{x^2}dx=2\frac{\phi(0)}{\epsilon}-2\frac{\phi(0)}{M}.$$

As done before, using for example Lebesgue's dominated convergence theorem, we have

$$\begin{aligned}\langle fp\frac{1}{x^2},\phi\rangle&=\lim_{\epsilon\to 0}\left(2\frac{\phi(0)}{\epsilon}-2\frac{\phi(0)}{M}+\frac{1}{2}\int_{\epsilon\leq|x|\leq M}\phi''(\theta)dx-2\frac{\phi(0)}{\epsilon}\right)\\&=\lim_{\epsilon\to 0}\left(\frac{1}{2}\int_{\epsilon\leq|x|\leq M}\phi''(\theta)dx-2\frac{\phi(0)}{M}\right)\\&=\frac{1}{2}\int_K\phi''(\theta)dx-2\frac{\phi(0)}{M}.\end{aligned}$$

Therefore,

$$\begin{aligned}\left|\langle fp\frac{1}{x^2},\phi\rangle\right|&=\left|\frac{1}{2}\int_K\phi''(\theta)dx-2\frac{\phi(0)}{M}\right|\leq\frac{1}{2}\mu(K)\|\phi''(x)\|_\infty+\frac{2}{M}\|\phi(x)\|_\infty\\&\leq C(K)\sup_{p\leq 2}\|\partial^p\phi\|_\infty.\end{aligned}$$

■

Up to now, we have considered only distributions of finite order. We must alert the reader that exist distributions which have no order at all, as shown in the following theorem.

Theorem 2.2.3. *The application* $\langle T,\phi\rangle:=\sum_{n=0}^{\infty}\phi^{(n)}(n)$ *defines a distribution in* $\mathbb{R}$ *which is not of finite order.*

Proof. See exercise 17 of [21]. ■

As previously mentioned, the heat equation plays an essential role in this thesis. With that in mind, the next theorem has great importance. The reason why will be evident in future chapters.

Theorem 2.2.4. *The function* Φ *in* $\mathbb{R}^2$ *defined by*

$$\Phi(t,x)=\begin{cases}\frac{H(t)}{\sqrt{4\pi t}}\exp\left[-\frac{x^2}{4t}\right], & \textit{if } x>0\\ \delta_0, & \textit{if } x=0\\ 0, & \textit{if } x<0\end{cases}$$

(where $H(t)$ is the Heaviside function: $H(t) = 1$ if $t > 0, H(t) = 0$ if $t < 0$) defines a distribution in $\mathbb{R}^2$.

Proof.

$$\Phi(t,x) = \frac{H(t)}{\sqrt{4\pi t}} \exp\left[-\frac{x^2}{4t}\right] \leq \frac{H(t)}{\sqrt{4\pi t}}$$

which is locally integrable in $\mathbb{R}^2$. ■

A *Radon measure* μ is a Borel measure that is finite on all compact sets. For $\phi \in \mathscr{D}(\Omega)$, the map $T_\mu : \phi \to \langle T_\mu, \phi \rangle := \displaystyle\int \phi \, d\mu$ defines a distribution of order 0. It is clear that T_μ is linear. Furthermore,

$$\Big| \int_\Omega \phi \, d\mu \Big| = \Big| \int_K \phi \, d\mu \Big| \leq \mu(K) \, \|\phi\|_\infty.$$

Note that if $\phi \geq 0$, then $\langle T_\mu, \phi \rangle \geq 0$. Such distributions are often called *positive distributions*. Any Radon measure defines a distribution of order zero. Even further, Schwartz proved that every positive distribution is a Radon measure and therefore as order zero (found, for instance, in [5]). However one of his challenges was to answer if these results still hold when considering distributions and measures in general.

As we know, classical probabilities are particular cases of positive Radon measures. If in (2.4) we have $f \geq 0$ with $\displaystyle\int f(x)dx = 1$, the distribution T_f is called a *probability measure* and f is called a *probability density function*. This is a simple example of the distributional interpretation of probabilities.

2.3 Product of distributions by a function

In the following sections we will study with some detail three important operations using distributions. First, the concept of multiplication. In the previous section, we saw that the definition of distributions implies particularly that for $T_1, T_2 \in \mathscr{D}^*(\Omega), \phi \in \mathscr{D}(\Omega), \alpha \in \mathbb{R}$, the operations $T_1 + T_2$ and αT_1 are of course well defined by

$$\Big\langle \alpha(T_1 + T_2), \phi \Big\rangle = \alpha \Big\langle (T_1 + T_2), \phi \Big\rangle = \alpha \Big(\langle T_1, \phi \rangle + \langle T_2, \phi \rangle \Big).$$

What about the multiplication of distributions? Is there any meaning in $\langle T_1T_2, \phi\rangle$? The answer is generally negative. The product is not defined for arbitrary distributions. However, there is a simple way to define the product of distributions by C^∞ functions which enlarge the idea of product between functions of the classical calculus. Let us begin with a few remarks.

The notion of convergence in $C^\infty(\Omega)$ is analogous to the concept introduced for the space of test functions.

Definition 2.3.1. Let $(\varphi_i)_{i\in\mathbb{N}}$ a sequence of $C^\infty(\Omega)$ and $\varphi \in C^\infty(\Omega)$. *The sequence* $(\varphi_i)_{i\in\mathbb{N}}$ *converges to* φ *in* $C^\infty(\Omega)$ if for every multi-index α and for all compact $K \subset \Omega$ the sequence $(\partial^\alpha\varphi_i)_{i\in\mathbb{N}}$ converges uniformly to $\partial^\alpha\varphi$ (i.e $\|\partial_x^\alpha\varphi_i - \partial_x^\alpha\varphi\|_{\infty,K} \to 0$).

From now on, the space of infinitely many times differentiable functions on Ω endowed with the notion of convergence above defined will be denoted by $\mathscr{E}(\Omega)$. Clearly, $\mathscr{D}(\Omega) \subset \mathscr{E}(\Omega)$. In addition, one can prove that $\mathscr{D}(\Omega)$ is dense in $\mathscr{E}(\Omega)$ (see, for instance, in [10]).

As will be standard procedure in this thesis, first we present the motivation that drives to the formal definition. Consider f, g continuous functions on Ω. (2.4) suggest that for $\phi \in \mathscr{D}(\Omega)$,

$$\langle fg, \phi\rangle = \int_\Omega (fg)(x)\phi(x)dx = \int_\Omega g(x)(f\phi)(x)dx = \text{“}\langle g, \phi f\rangle\text{”}. \tag{2.7}$$

However, we may have a problem here, since there is no guarantee that the product $f\phi \in \mathscr{D}(\Omega)$. But, if it is possible to ensure that, we can remove the quotes. This lead us to the following definition.

Definition 2.3.2. Let $T \in \mathscr{D}^*(\Omega)$ and $f \in \mathscr{E}(\Omega)$. The formula

$$\langle fT, \phi\rangle := \langle T, \phi f\rangle \qquad (\forall\phi \in \mathscr{D}(\Omega)) \tag{2.8}$$

defines a distribution $fT \in \mathscr{D}^*(\Omega)$.

Indeed, with the assumption that $f \in \mathscr{E}(\Omega)$ and $\phi \in \mathscr{D}(\Omega)$, we have corrected the issue on the right-hand side of (2.7). But we still have to check that (2.8) defines a distribution. The linearity comes right away. Furthermore, with $supp\,\phi \subseteq K$,

$$\begin{aligned}\langle f \cdot T, \phi\rangle := \langle T, \phi f\rangle &\leq C(K) \sup_{|\alpha|\leq p} \|\partial^\alpha (f\cdot\phi)\|_\infty \\ &\leq C(K) \sup_{|\alpha|\leq p} \left\| \sum_{|\beta|\leq|\alpha|} \binom{\alpha}{\beta} \partial^{\alpha-\beta} f \partial^\beta \phi \right\|_\infty \\ &\leq C(K) \left\| \sum_{|\beta|\leq|\alpha|} \partial^{\alpha-\beta} f \right\|_\infty \sup_\beta \left\|\partial^\beta \phi\right\|_\infty \\ &\leq C_1(K) \sup_\beta \left\|\partial^\beta\phi\right\|_\infty.\end{aligned}$$

It is clear now that if T has finite order $\leq p_0$ then $f\cdot T$ has also finite order $\leq p_0$. We can similarly define the product when $g \in C^m$ and T is a distribution of order $\leq m$. As a result, $g \cdot T_\mu$ defines a measure for every continuous function g and measure μ.

We next present typical examples of multiplication of distributions by smooth functions.

Example 2.3.1. Let δ_a the Dirac distribution at $a \in \mathbb{R}$, $\varphi \in \mathscr{E}(\mathbb{R})$ and $\phi \in \mathscr{D}(\mathbb{R})$. Then we have

$$\langle \varphi \cdot \delta_a, \phi\rangle := \langle \delta_a, \phi\varphi\rangle = \varphi(a)\phi(a).$$

Example 2.3.2. For $f \in \mathscr{E}(\Omega), T = T_g \in \mathscr{D}^*(\Omega)$ with $g \in \mathbb{L}^1_{loc}(\Omega)$, we have

$$fT_g = T_{gf}.$$

Proof. For $\phi \in \mathscr{D}(\Omega)$:

$$\begin{aligned}\langle fT_g, \phi\rangle := \langle g, \phi f\rangle &:= \int_\Omega g(x)(f\phi)(x)dx \\ &= \int_\Omega (gf)(x)\phi(x)dx \\ &= \langle T_{gf}, \phi\rangle.\end{aligned}$$

■

Example 2.3.3. The equality $x \cdot vp\frac{1}{x} = 1$ holds in $\mathscr{D}^*$.

Proof. For $\phi \in \mathscr{D}(\Omega)$:

$$\begin{aligned}\langle x \cdot vp\frac{1}{x}, \phi\rangle := \langle vp\frac{1}{x}, \phi x\rangle &= \lim_{\epsilon \to 0} \int_{|x|\geq \epsilon} \frac{\phi(x)x}{x} dx \\ &= \int_{|x|\geq 0} \phi(x)dx \\ &= \int_{\Omega} \phi(x)dx \\ &= \langle 1, \phi\rangle.\end{aligned}$$

■

Using this equality, we can now show that, in general, the product of distributions is not associative. With simple computations, one may check that, in the distribution sense $\left(vp\frac{1}{x} \cdot x\right) \cdot \delta_0 = \delta_0$ while $vp\frac{1}{x} \cdot (x \cdot \delta_0) = 0$.

2.4 Differentiation of Distributions

As in the case of multiplication, our aim here is to define the concept of differentiation of distributions that is compatible with the same concept in the classical analysis. In a nutshell, for a distribution $f \in \mathscr{D}^*(\mathbb{R})$, can we come up with a definition for f' that coincides with the classical notion of derivative whenever $f \in C^1(\mathbb{R})$? The answer is positive. Not only such definition exists but it is also compatible with the usual algebraic properties (including the well-known Leibniz formula and the chain rule). We can even say that distributions are the natural completion of differential calculus. So, as we will see, distributions always possess partial derivatives of arbitrary orders. This implies particularly that the derivative of some function f may not exist as a function but will exist as a distribution.

The motivation for the formal definition is as it follows. Let $f \in C^1(\Omega)$. Thus $f' \in C(\Omega)$. Plus assume for a test function ϕ, *supp* $\phi \subseteq K$. From (2.4),

$$\langle f', \phi\rangle = \int_K f'(x)\phi(x)dx = \underbrace{[f\phi]_K}_{0} - \int_K f\phi' dx = -\int_{\mathbb{R}^n} f\phi' dx = -\langle f, \phi'\rangle. \tag{2.9}$$

There is no issue with the use of the integration by parts formula given that

f' is continuous and ϕ is a test function. We want to expand (2.9) to any other distribution.

Definition 2.4.1. Let $T \in \mathscr{D}^*(\Omega)$ and $i \in \{1, \dots, n\}$. The *distributional derivative* of T with respect to the $i-th$ variable is defined by

$$\langle \partial_i T, \phi \rangle := -\langle T, \partial_i \phi \rangle, \qquad (\forall \phi \in \mathscr{D}(\Omega)). \tag{2.10}$$

When $n = 1$, we denote the derivative by the shorthand T' instead of $\partial_1 T$.

The fact that equation (2.10) certainly defines distribution follows from

$$\begin{aligned} \langle \partial_i T, \phi \rangle &:= -\langle T, \partial_i \phi \rangle \\ &\leq C(K) \sup_{|\alpha| \leq p} \|\partial^\alpha \partial_i \phi\|_\infty \\ &\leq C(K) \sup_{|\alpha| \leq p+1} \|\partial^\alpha \phi\|_\infty. \end{aligned}$$

Hence, if T is a distribution of finite order p_0, $\partial_i T$ has finite order p_0+1. By applying (2.10) iteratively (which is possible since $\phi \in \mathscr{D}(\Omega)$), we may define the derivative $\partial^\alpha T$ by

$$\langle \partial^\alpha T, \phi \rangle := (-1)^{|\alpha|} \langle T, \partial^\alpha \phi \rangle \qquad (\forall \phi \in \mathscr{D}(\Omega)). \tag{2.11}$$

As expected, if T is a distribution of finite order p_0, $\partial^\alpha T$ is a distribution of order $p_0 + |\alpha|$. We can, of course, define the anti-derivative of a distribution.

Definition 2.4.2. Let $T \in \mathscr{D}^*(\Omega)$. A distribution S such that $\langle S', \phi \rangle = \langle T, \phi \rangle$ for all $\phi \in \mathscr{D}(\Omega)$ is called *anti-derivative of* T.

As in the case of classical functions, one may prove that any distribution $T \in \mathscr{D}^*(\Omega)$ has an anti-derivative which is unique up to a constant.

Now, let us compute some examples of derivatives of some important distributions.

Example 2.4.1. Derivatives of the *Dirac distribution* δ_x.

$$\langle \delta_x', \phi \rangle = -\langle \delta_x, \phi' \rangle = -\phi'(x).$$

With $|\langle \delta'_x, \phi\rangle| = |-\phi'(x)| \leq \|\phi'\|_\infty$. Thus, the first derivative of the Dirac δ_x is a distribution of order 1. By recurrence, it is easy to prove that

$$\langle \delta_x^{(k)}, \phi\rangle = (-1)^k \phi^{(k)}(x) \tag{2.12}$$

is a distribution of order k, with k $\in \mathbb{N}$.

Example 2.4.2. Consider the *Heaviside function* H. In the distribution sense, $H' = \delta_0$.

Proof. For $\phi \in \mathscr{D}(\mathbb{R})$, with $supp\,\phi \subset [-M, M]$:

$$\begin{aligned}\langle H', \phi\rangle &:= -\langle H, \phi'\rangle = -\int_{\mathbb{R}} H(x)\phi'(x)dx = -\int_0^{+\infty} \phi'(x)dx = -\int_0^M \phi'(x)dx \\ &= -\underbrace{\phi(M)}_{0} + \phi(0) = \phi(0) \\ &= \langle \delta_0, \phi\rangle.\end{aligned}$$

■

Example 2.4.3. Let $f = |x|$. In the distribution sense, $|x|' = sgn(x)$, where $sgn(x)$ is the *sign function* ($sgn(x) = 2H(x) - 1$)).

Proof. For $\phi \in \mathscr{D}(\mathbb{R})$, with $supp\,\phi \subset [-M, M]$:

$$\begin{aligned}\langle |x|', \phi\rangle &:= -\langle |x|, \phi'\rangle = -\int_{-M}^{M} |x|\phi'(x)dx = \int_{-M}^{0} x\phi'(x)dx - \int_0^M x\phi'(x)dx \\ &= \Big[x\phi(x)\Big]_{-M}^{0} - \int_{-M}^{0} \phi(x)dx - \Big[x\phi(x)\Big]_0^M + \int_0^M \phi(x)dx \\ &= \int_0^M \phi(x)dx - \int_{-M}^0 \phi(x)dx \\ &= \int_{\mathbb{R}} \big(\mathbb{1}_{[0,M]} - \mathbb{1}_{[-M,0]}\big)\phi(x)dx \\ &= \langle sgn, \phi\rangle.\end{aligned}$$

■

Example 2.4.4. Let $f = \ln(|x|)$. In the distribution sense, $(ln(|x|))' = vp\frac{1}{x}$.

Proof. For $\phi \in \mathscr{D}(\mathbb{R})$, with $supp\,\phi \subset [-M, M]$:

$$\langle \log(|x|)', \phi \rangle = \int_{-M}^{M} \log(|x|)\phi'(x) = \lim_{\epsilon \to 0} \Big(\underbrace{\int_{-M}^{-\epsilon} log(-x)\phi'(x)dx}_{A} + \underbrace{\int_{\epsilon}^{M} log(x)\phi'(x)dx}_{B} \Big).$$

- A:

$$\int_{-M}^{\epsilon} log(-x)\phi'(x)dx = \Big[\log(-x)\phi(x) \Big]_{-M}^{-\epsilon} - \int_{-M}^{-\epsilon} \frac{\phi(x)}{x} dx = \log(\epsilon)\phi(-\epsilon) - \int_{-M}^{-\epsilon} \frac{\phi(x)}{x} dx.$$

- B:

$$\int_{\epsilon}^{M} log(x)\phi'(x)dx = \Big[\log(x)\phi(x) \Big]_{\epsilon}^{M} - \int_{\epsilon}^{M} \frac{\phi(x)}{x} dx = -(\log(\epsilon)\phi(\epsilon)) - \int_{\epsilon}^{M} \frac{\phi(x)}{x} dx.$$

Thus,

$$\begin{aligned}
\langle \log(|x|)', \phi \rangle &= \lim_{\epsilon \to 0} \Big(\log(\epsilon)\phi(\epsilon) - \int_{-M}^{-\epsilon} \frac{\phi(x)}{x} dx - \log(\epsilon)\phi(\epsilon) - \int_{\epsilon}^{M} \frac{\phi(x)}{x} dx \Big) \\
&= \lim_{\epsilon \to 0} \Big(\log(\epsilon)(\phi(-\epsilon) - \phi(\epsilon)) - \int_{-M}^{-\epsilon} \frac{\phi(x)}{x} dx - \int_{\epsilon}^{M} \frac{\phi(x)}{x} dx \Big) \\
&= \lim_{\epsilon \to 0} \Big(\log(\epsilon)\epsilon\phi'(c(\epsilon)) - \int_{-M}^{-\epsilon} \frac{\phi(x)}{x} dx - \int_{\epsilon}^{M} \frac{\phi(x)}{x} dx \Big) \\
&= \lim_{\epsilon \to 0} \int_{\epsilon \leq |x| \leq M} \frac{\phi(x)}{x} = \langle vp\frac{1}{x}, \phi \rangle.
\end{aligned}$$

■

Just as for the smooth functions, the mixed partial derivatives are independent of the order of differentiation.

Lemma 2.4.1. *For any distribution T and for all indices such that $1 \leq i, j \leq n$, it come: $\partial_i(\partial_j T) = \partial_j(\partial_i T)$.*

Proof. For every $\phi \in \mathscr{D}(\Omega)$: $\langle \partial_i \partial_j T, \phi \rangle = \underbrace{(-1)^j(-1)^i \langle T, \partial_j \partial_i \phi \rangle = (-1)^i(-1)^j \langle T, \partial_i \partial_j \phi \rangle}_{Schwartz\ Theorem} = \langle \partial_j \partial_i T, \phi \rangle$. ■

2.5 Support of a Distribution

In section 2.1, the support of a test function was defined as the closure of the subset formed by the points where the value of the test function is not zero. This concept, however, can be often extended. Roughly speaking, the support of something is the set where it values are non-trivial.

To illustrate this idea, consider, for instance, the Dirac distribution δ_0, the zero-distribution T_0 and a test function ϕ such that $supp\,\phi \subseteq \mathbb{R}\backslash\{0\}$. By definition, $\langle \delta_0, \phi \rangle = \phi(0)$, which is zero since the point $x = 0$ does not belong to the support of ϕ. Also, $\langle T_0, \phi \rangle = 0$ for any ϕ, regardless of its support. Intuitively, $supp\,\delta_0 = \{0\}$ and $supp\,T_0 = \emptyset$.

Definition 2.5.1. Let $T \in \mathscr{D}^*(\Omega)$. The *support of* T is the complement of the set of points x such that $\langle T, \phi \rangle$ vanishes if ϕ has support in a neighbourhood of x:

$$supp\,T = \overline{\big\{x :\, \forall V_x\, \exists \phi \in \mathscr{D}(V_x) : supp\,\phi(x) \subseteq V_x \,\wedge\, \langle T, \phi \rangle \neq 0\big\}}. \tag{2.13}$$

It results, almost immediately, that if $supp\,T \cap supp\,\phi = \emptyset$ then $\langle T, \phi \rangle = 0$. Furthermore, Lemma 2.1.1 still holds true. The definition is consistent with what we have so far, namely it is not hard to prove that if $f \in \mathbb{L}^1_{loc}$, $supp\,T_f = supp\,f$.

Until now, the compactness of the support of a test function was essential to define distributions. The reason why is quite obvious. Assume, for instance $g \equiv 1$ and $\phi \equiv 1$. According to (2.4),

$$\langle T_g, \phi \rangle = \langle 1, 1 \rangle = \int_{\mathbb{R}^n} 1dx$$

which is a divergent integral and thus $\langle g, \phi \rangle$ has no meaning.

Definition 2.5.2. The space $\mathscr{E}^*(\Omega)$ of distributions in Ω with compact support is defined as

$$\mathscr{E}^*(\Omega) = \{T \in \mathscr{D}^*(\Omega) \,:\, supp\,T \text{ is compact}\}.$$

Distributions with compact support make an important class of distributions. From the notation, as the reader may expect, $\mathscr{E}^*(\Omega)$ is the topological dual of $\mathscr{E}(\Omega)$. Simply put, if E is a distribution with compact support then $\langle E, \varphi \rangle$ is well defined for every $\varphi \in \mathscr{E}(\Omega)$. In some sense the compactness condition on the support of the test functions is transfer to the support of the distribution. In addition, one can prove that $\mathscr{E}^*(\Omega)$ is dense in $\mathscr{D}^*(\Omega)$.

Theorem 2.5.1. *Let* $E : \mathscr{E}(\Omega) \to \mathbb{R}$ *a linear functional. It comes that* $E \in \mathscr{E}^*(\Omega)$ *if and only if*

$$\exists K \subset\subset \Omega, \exists C = C(K) > 0,\ \exists p \in \mathbb{N}_0\ : |\langle E, \varphi\rangle| \leq C(K) \sup_{|\alpha|\leq p} \|\partial^\alpha \varphi\|_{\mathbb{L}^\infty(supp\, E)}. \tag{2.14}$$

for every $\varphi \in \mathscr{E}(\Omega)$.

Proof. See Theorem 1.62 of [8]. ∎

From (2.14), it results that every distribution with compact support has finite order. The converse is false (just consider the constant distribution $T \equiv 1$). Every distribution whose support is a single point is a finite linear combination of the Dirac distribution and its derivatives at that single point. Mathematically, fix $x_0 \in \Omega$. If $E \in \mathscr{E}^*(\Omega)$ with $supp\, E = \{x_0\}$ then $\exists p \in \mathbb{N}_0$ and $c_\alpha \in \mathbb{R}$ such that:

$$\langle E, \varphi\rangle = \sum_{|\alpha|\leq p} c_\alpha \partial^\alpha \varphi(x_0) \qquad (\forall \varphi \in \mathscr{E}(\Omega)).$$

The most curious reader may consult, for instance, [5], [7] and [18] for the formal proof.

2.6 Convolution

Now, we shall focus on one last operation regarding distributions. We will briefly explain the convolution operation, which will have significant importance for our purpose, namely when we define the fundamental solution of a differential operator and the Fourier transform in the next chapter. The term itself is relatively new. It was named only around 1950 despite one of its first appearances in the eighteen century with D'Alembert, Laplace, Fourier and Poisson. The concept has applications in many areas of science (probability, statistics, computer vision, natural language, signal processing, engineering, etc.). For a better understanding, we suggest the reading of [3], [9] and [11].

This section is organized in the following manner. First, we present the definition and basic properties of the convolution between classical functions. Next, we study the case where the convolution is applicable to distributions and test functions. Finally, the notion and properties of the operation are extended to distributions.

2.6.1 Convolution between functions

Consider f and g continuous functions in $\mathbb{R}^n$ and assume that one, say f, has compact support K. The convolution of f and g is the continuous function defined by

$$(f * g)(x) := \int_K f(t)g(x-t)dt. \tag{2.15}$$

The operation has several algebraic properties. For adequate given functions f, g, h and $a \in \mathbb{R}$, the convolution satisfies

- Commutativity: $(f * g) = (g * f)$,
- Associativity: $f * (g * h) = (f * g) * h$,
- Distributivity: $f * (g + h) = (f * g) + (f * h)$,
- Associativity with scalar multiplication: $a(f * g) = (af) * g = f * (ag)$,
- Support property: $\mathit{supp}\, f * g \subset \mathit{supp}\, f + \mathit{supp}\, g$.

The relationship between convolution and differentiation is also very interesting and deserves a special note.

Theorem 2.6.1. *For every $f \in C^j(\Omega)$, $g \in C^k(\Omega)$ and for every multi-indices α and β such that $|\alpha| \leq j$, $|\beta| \leq k$:*

1. $f * g \in C^{j+k}(\Omega)$,
2. $\partial^{\alpha+\beta}(f * g) = \partial^\alpha f * \partial^\beta g$. *Particularly,* $\partial^\alpha(f * g) = (\partial^\alpha f) * g = f * (\partial^\alpha g)$.

The next very useful theorem is commonly referred as the *Young's Inequality.*

Theorem 2.6.2. *Let $p, q, r \in \mathbb{N}$ such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$. Consider $f \in \mathbb{L}^p(\Omega)$, $g \in \mathbb{L}^q(\Omega)$. Then,*

$$\|f * g\|_{\mathbb{L}^r} \leq \|f\|_{\mathbb{L}^p} \|g\|_{\mathbb{L}^q}.$$

Corollary 2.6.2.1. *Let $p \in \mathbb{N}$. Consider $f \in \mathbb{L}^p(\Omega)$, $g \in \mathbb{L}^1(\Omega)$. Then,*

1. $f * g \in \mathbb{L}^p(\Omega)$,
2. $\|f * g\|_{\mathbb{L}^p} \leq \|f\|_{\mathbb{L}^p}\|g\|_{\mathbb{L}^1}$.

2.6.2 Convolution between distributions and test functions

Here, we enlarge the concept of convolution to distributions and test functions. Toward this end, we first need some preceding definition. For a given function f, its *reflection* $\check{f}$ is defined as $\check{f}(x) := f(-x)$.

The approach is very similar to what we have done so far. Let $g \in \mathbb{L}^1_{loc}(\Omega)$. By direct application of (2.4), for $\phi \in \mathscr{D}(\Omega)$

$$(g * \phi)(x) := \int g(t)\phi(x-t)dt = \langle g, \phi(x - \cdot)\rangle. \tag{2.16}$$

Thus, it is natural to define the *convolution between a distribution $T \in \mathscr{D}^*(\Omega)$ and a test function $\phi \in \mathscr{D}(\Omega)$* by:

$$(T * \phi)(x) := \langle T, \phi(x - \cdot)\rangle. \tag{2.17}$$

The properties enunciated for the convolution between functions still holds in this case.

Theorem 2.6.3. *If $T \in \mathscr{D}^*(\Omega)$ and $\phi, \gamma \in \mathscr{D}(\Omega)$ then:*

1. $T * \phi \in \mathscr{E}(\Omega)$,
2. $\operatorname{supp} T * \phi \subset \operatorname{supp} T + \operatorname{supp} \phi$,
3. $\partial^\alpha(T * \phi) = \partial^\alpha(T) * \phi = T * \partial^\alpha(\phi)$
4. $T * (\phi * \gamma) = (T * \phi) * \gamma$
5. $(T * \phi)(0) = \langle T, \check{\phi}\rangle$.

2.6.3 Convolution between distributions

Now, we expand to the more general case, the convolution between distributions. The question is fairly simple. Given two distributions, T_1, T_2 it is possible to define $T_1 * T_2$? The answer is positive, but only when at least one of the distributions has compact support. The idea is much like to the first approach in 3.6.2. Assume, without loss of generality, that $f, g \in \mathbb{L}^1_{loc}(\Omega)$ and $\phi \in \mathscr{D}(\Omega)$. Then formally,

$$\begin{aligned}\langle g * f, \phi\rangle &= \int (g * f)(x)\phi(x)dx = \int\int g(t)f(x-t)dt\phi(x)dx \\ &= \int\int g(t)f(x-t)\phi(x)dtdx \\ &= \int g(t)\int f(y)\phi(y+t)dydt \\ &= \langle g, \langle f, \phi(y+t)\rangle\rangle.\end{aligned}$$

The reason why g requires compact support is very clear now since, by Theorem 2.6.3, $\langle f * \phi(y+t)\rangle \in \mathscr{E}(\Omega)$. One may also consider $f \in \mathscr{E}^*(\Omega)$ rather than g. Because the convolution is commutative, the results will be analogous. This naive approach suggest the following definition of convolution between two distributions.

Definition 2.6.1. Let $E \in \mathscr{E}^*(\Omega), T \in \mathscr{D}^*(\Omega)$. For $\phi \in \mathscr{D}(\Omega)$,

$$\langle T * E, \phi\rangle := \Big\langle E_x, \langle T_y, \phi(x+y)\rangle\Big\rangle. \tag{2.18}$$

The subscripts x and y pointing out in which variable the distribution is applied. The reader may find in the literature, other ways to define the convolution between distributions. For instance, [7] and 13 define the convolution between T_1 and T_2 where at least one has compact support, as the unique distribution T that verifies the associativity property $T_1 * (T_2 * \phi) = (T_1 * T_2) * \phi$. That is, the unique distribution T such that

$$T_1 * (T_2 * \phi) = T * \phi \qquad (\forall \phi \in \mathscr{D}(\Omega)). \tag{2.19}$$

The construction of such T relies on continuity theorems, which are not studied in this work. A few notes about (2.19). If T_2 has compact support, then $T_2 * \phi \in \mathscr{D}(\Omega)$ and $T_1 * (T_2 * \phi)$ is well-defined. On the other hand, if T_1 has compact support, then $T_2 * \phi \in \mathscr{E}(\Omega)$ and $T_1 * (T_2 * \phi)$ is also well-defined. One may even find the definition of convolution between distributions based on tensor products as in [12]. Such notion, however, is out of scope of this work. As referred before, the properties stated on Theorem 2.6.3 remain valid for distributions. In addition we have the following theorem.

Theorem 2.6.4. *The Dirac distribution is the identity operator of the convolution.*

Proof. Consider δ_0, which we already know that has compact support. For any $T \in \mathscr{D}^*(\Omega)$ and $\phi \in \mathscr{D}(\Omega)$:

$$\langle T * \delta_0, \phi \rangle := \left\langle \delta_{0_x}, \langle T_y, \phi(x+y) \rangle \right\rangle := \langle T_y, \phi(y) \rangle = \langle T, \phi \rangle. \tag{2.20}$$

■

We can use (2.20) to express differentiation using convolution. For any $T \in \mathscr{D}^*(\Omega)$ and every multi-index α,

$$\partial^\alpha T = \partial^\alpha (T * \delta_0) = \partial^\alpha \delta_0 * T.$$

Plus, using the associativity of the operation, for any $\phi \in \mathscr{D}(\Omega)$,

$$(\partial^\alpha T) * \phi = T * (\partial^\alpha \phi) = (T * \delta_0) * (\partial^\alpha \phi) = T * (\partial^\alpha \delta_0) * \phi.$$

Chapter 3

Fourier Transform

3.1 Fundamental Solutions

One important application of distributions and convolution theories to partial differential equations relies on the so-called fundamental solutions of linear differential operators, which we will soon define. It displays a significant role in the study of existence and regularity of solutions of such operators. Although the first appearance of the concept happened in 1911 with N. Zeilon, the definition as we know it today was given by Schwartz in 1950. In natural sciences, physical events are mainly described by vector or tensor fields instead of scalar quantities. Thus, it is more usual to use systems of differential equations to model physical processes. In this framework, Schwartz defined fundamental matrices and fundamental solutions for systems of differential operators. In this thesis, we will often consider the one-dimensional case. For a more complete study of fundamental solutions, we indicate the reading of [12] and surely of [16]. We start by fixing the notation for differential operators.

Definition 3.1.1. Consider α a multi-index and let

$$P(\xi) = \sum_{|\alpha| \leq m} c_\alpha \xi^\alpha$$

be a polynomial in n variables of degree m with coefficients $c_\alpha \in \mathbb{C}$. By replacing ξ

by ∂_x we define

$$P(\partial_x) = \sum_{|\alpha|\leq m} c_\alpha \partial_x^\alpha = \sum_{|\alpha|\leq m} c_\alpha \frac{\partial^\alpha}{\partial x^\alpha} \tag{3.1}$$

a *linear partial differential operator with constant coefficients of order m.*

From now on, $P(\partial_x)$ will represent such linear partial differential operators with constant coefficients, as in (3.1). For such operators, we can guarantee the existence of fundamental solutions. This was first proved by Ehrenpreis and Malgrange around the years of 1954 to 1956. Despite being very technical, the proof can be found, for instance, in [10]. The linear partial differential operators with variable coefficients require more complex tools. For such reason and to maintain simplicity, we will not consider them.

Definition 3.1.2. A distribution $T \in \mathscr{D}^*(\Omega)$ such that $P(\partial_x)T = \delta_0$ is called a *fundamental solution of* P.

Example 3.1.1. In section 2.4, we proved that the Heaviside function H is a fundamental solution of the operator $P(\partial_x) = \partial_x$, since $H' = \delta_0$.

Once we have found a fundamental solution of an operator $P(\partial_x)$, we can use it to find solutions of other equations involving $P(\partial_x)$. To formalize the importance of fundamental solutions, we present a first very simple theorem.

Theorem 3.1.1. *Assume T a fundamental solution of $P(\partial_x)$. Then, for every $E \in \mathscr{E}^*(\Omega)$ we have*

$$P(\partial_x)(T * E) = T * P(\partial_x)E = P(\partial_x)T * E = \delta_0 * E = E. \tag{3.2}$$

The equalities in (3.2) mean in particular that for every distribution f of Ω, the in-homogeneous differential equation $P(\partial_x)u = f$ has a solution of the form $T * f$. By this time, it is natural to ask if the fundamental solution E of an operator $P(\partial_x)$ is unique. The answer is negative and very intuitive. Suppose that $u \in \mathscr{D}^*(\Omega)$ is a solution of the homogeneous equation $P(\partial_x)u = 0$. Then $\tilde{T} = T + u$ is also a fundamental solution of $P(\partial_x)$. Indeed, $P(\partial_x)\tilde{T} = P(\partial_x)(T + u) = \underbrace{P(\partial_x)T}_{=\delta_0} + \underbrace{P(\partial_x)u}_{0} = \delta_0$.

3.2 Fourier Transform

Before moving to a more in-depth theory, perhaps it is a good idea to describe first some elementary results and properties about the Fourier transform in the classical sense. In order to keep simplicity, for now, consider the one-dimension case. As we know, any continuous function, periodic, admits a representation in terms of Fourier series. For such a function, we may consider a period of 2π by using a scaling argument.

If g is a function in the above conditions, the *discrete Fourier series representation* is:

$$g(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos{(nx)} + b_n \sin{(nx)}, \tag{3.3}$$

with Fourier coefficients given by:

- $$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g(x) \cos{(nx)}\ dx$$
- $$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g(x) \sin{(nx)}\ dx$$
- $$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} g(x)\, dx. \tag{3.4}$$

The first step to reach the final formula of the Fourier transform is to write the Fourier series in the complex form. By consequences of the Euler formula ($e^{ix} = \cos(x) + i\sin(x)$) we can write

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}, \qquad \sin(x) = \frac{e^{ix} - e^{-ix}}{2i}.$$

After substitution and some simple computations, (3.3) can be re-written as:

$$g(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \tag{3.5}$$

where

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx. \tag{3.6}$$

Unfortunately, most functions are not periodic. To pass from periodic to aperiodic functions, we let the period get very large, i.e., $\tau \to \infty$. Note that c_n is the amplitude of the exponential of frequency n. By considering continuous frequencies $\xi \in \mathbb{R}$, with some algebraic computations (see [18]) we define the more general Fourier transform.

3.2.1 $\mathbb{L}^1$ theory of Fourier transform

Definition 3.2.1. The *Fourier transform* of a function $f \in \mathbb{L}^1(\mathbb{R}^n)$ is defined as the function $\mathscr{F} : \mathbb{R}^n \to \mathbb{C}$ given by:

$$\mathscr{F}(f)(\xi) := \int_{\mathbb{R}^n} f(x)e^{-ix\cdot\xi}\,dx, \tag{3.7}$$

where $x \cdot \xi$ denotes the standard inner product of x and ξ in $\mathbb{R}^n$. The value of the integral is finite since $\int_{\mathbb{R}^n} |f(x)e^{-ix\cdot\xi}|\,dx = \int_{\mathbb{R}^n} |f(x)|\,dx < \infty$ for every ξ. We will often use the standard notation $\widehat{f}$ instead of $\mathscr{F}(f)$ to represent the Fourier transform of a function f. Morally, decomposing f in a sum of periodic functions, $\widehat{f}$ is the amplitude of the wave of frequency ξ.

Right here, a warning must be done. The reader will find, surely, in literature different definitions of the Fourier transform. All somehow equivalent differing only by the constant 2π and/or a minus sign. For instance, [1] and [6] use as definition $\widehat{f}(\xi) = \int f(x)e^{-2\pi i\,x\cdot\xi}\,dx$, [4] uses $\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}}\int f(x)e^{-ix\xi}dx$ while [18] uses $\widehat{f}(\xi) = \int f(x)e^{i\,x\cdot\xi}\,dx$. There are even more exotic formulas. So, one should remember when working with Fourier transform, to always be careful with the definition that is being used. Employing the theory that we have studied so far, it is not hard to prove the following theorem.

Theorem 3.2.1. *Let $f, g \in \mathbb{L}^1(\mathbb{R}^n)$. Then,*

1. *The mapping $f \to \widehat{f}$ is continuous from $\mathbb{L}^1(\mathbb{R}^n)$ to $\mathbb{L}^\infty(\mathbb{R}^n)$. In fact*

$$|\widehat{f}(\xi)| \le \|f(x)\|_{\mathbb{L}^1}, \qquad (\forall \xi \in \mathbb{R}^n). \tag{3.8}$$

2. *$\widehat{f}$ is uniformly continuous.*

3. *Riemann-Lebesgue:* $\widehat{f}(\xi) \to 0$ *as* $|\xi| \to \infty$.

4.
$$\int_{\mathbb{R}^n} f(x)\widehat{g}(x)\,dx = \int_{\mathbb{R}^n} \widehat{f}(\xi)g(\xi)\,d\xi, \tag{3.9}$$

5.
$$\widehat{f * g} = \widehat{f} \cdot \widehat{g}. \tag{3.10}$$

The relationship between Fourier transform and differentiation is also quite interesting. Applying the Fourier transform after multiplying by the $j-th$ variable is equivalent to taking the partial derivative w.r.t the $j-th$ variable of the Fourier transform. To avoid playing with factors of the form $(-i)^{|\alpha|}$, it is very common to define

$$D_j := \frac{1}{i}\partial_j, \qquad (1 \leq j \leq n), \qquad (i^2 = -1).$$

Theorem 3.2.2. *For every* $1 \leq j \leq n$, $\xi \in \mathbb{R}^n$ *and* α *a multi-index:*

1.
$$\widehat{D_j^\alpha f}(\xi) = \xi_j^\alpha \widehat{f}(\xi), \tag{3.11}$$

2.
$$\widehat{x_j^\alpha f}(\xi) = (-1)^{|\alpha|} D_j^\alpha \widehat{f}(\xi). \tag{3.12}$$

Now, let us focus on the problem of inverting the Fourier transform. In other words, how for a given Fourier transform $\widehat{f}$ of some integrable function f, we obtain f back again from $\widehat{f}$. Given the definitions presented above, we should expect something similar to

$$\mathscr{F}^{-1}\widehat{f} = \widehat{\mathscr{F}^{-1}f} = f,$$

and therefore

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(\xi)e^{ix\cdot\xi}\,d\xi. \tag{3.13}$$

However, there is an issue with (3.13). How do we guarantee that $\widehat{f}$ is, in fact, integrable? With the aim of trying to solve the above question, [6] makes use of summability methods for integrals (Abel and Gaussian summability methods). We will return to this integrability point in a moment. Of course, the expression of $\mathscr{F}^{-1}$ depends on the definition adopted for $\widehat{f}$.

Theorem 3.2.3. *Under the assumption that f and $\widehat{f} \in \mathbb{L}^1(\mathbb{R}^n)$ and knowing (3.7), the Fourier Inversion Formula is given by*

$$f(x) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{f}(\xi) e^{ix\cdot\xi}\, d\xi. \tag{3.14}$$

This implies particularly that

$$\widehat{\widehat{f}} = (2\pi)^n \check{f}. \tag{3.15}$$

Note that if $\widehat{f}(\xi) = 0$ for all ξ then $f(x) = 0$ a.e. This is the basis of the following injectivity result for the Fourier transform.

Theorem 3.2.4. *If f and g belong to $\mathbb{L}^1(\mathbb{R})$ and $\widehat{f}(\xi) = \widehat{g}(\xi)$ for all $\xi \in \mathbb{R}$ then $f(x) = g(x)$ a.e.*

3.2.2 $\mathbb{L}^2$ theory of Fourier transform

The theory presented in the last section depends on the assumptions that f and its Fourier transform $\widehat{f}$ are integrable functions. The space $\mathbb{L}^2$ deserves our special attention. For the next results, consider $\mathbb{R}$ instead of $\mathbb{R}^n$. Note that an arbitrary function $f \in \mathbb{L}^2(\mathbb{R})$ may not belong to $\mathbb{L}^1(\mathbb{R})$. Since the inclusion is not verified, the definition of $\widehat{f}$, as in (3.7), may not be well-defined. The Fourier transform has a natural definition in this space and a particularly elegant theory. One reason that $\mathbb{L}^2$ is a special setting of the Fourier transform is, as we will see, the fact that this space is preserved under this operation. If we assume f to be square-integrable then $\widehat{f}$ is also square-integrable. The Fourier transform definition on the $\mathbb{L}^2$ space relies on the next two essential theorems.

Theorem 3.2.5. *The space $\mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$ is a dense subset of $\mathbb{L}^2(\mathbb{R})$.*

Theorem 3.2.6. *Plancherel Theorem: If $f \in \mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$ then $\widehat{f} \in \mathbb{L}^2(\mathbb{R})$. Moreover,*

$$\int f(x)\widehat{g}(x)\, dx = \int \widehat{f}(\xi) g(\xi)\, d\xi,$$

holds for all $f, g \in \mathbb{L}^2(\mathbb{R})$. Plus,

$$\|f\|_{\mathbb{L}^2} = \frac{1}{\sqrt{2\pi}} \left\|\widehat{f}\right\|_{\mathbb{L}^2}. \tag{3.16}$$

Definition 3.2.2. For a function $f \in \mathbb{L}^2(\mathbb{R})$, Theorem 3.2.5 ensures the existence of a sequence $(f_n)_{n\in\mathbb{N}} \in \mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$ such that $(f_n)_{n\in\mathbb{N}}$ converges to f under the norm $\mathbb{L}^2$, that is $\|f_n - f\|_{\mathbb{L}^2} \to 0$. By linearity of the Fourier transform, Plancherel's theorem and since that $\mathbb{L}^2(\mathbb{R})$ is a complete space, the sequence $(\widehat{f_n})_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{L}^2(\mathbb{R})$,

$$\left\|\widehat{f_n} - \widehat{f_m}\right\|_{\mathbb{L}^2} = \left\|\widehat{f_n - f_m}\right\|_{\mathbb{L}^2} = \sqrt{2\pi}\|f_n - f_m\|_{\mathbb{L}^2} \to 0.$$

The *Fourier transform of f* is defined as the unique limit of $(\widehat{f_n})$ in $\mathbb{L}^2(\mathbb{R})$.

This limit is of course independent of the approximating sequence. For any other convergent sequence $(g_n)_{n\in\mathbb{N}} \in \mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$,

$$\lim_{n\to+\infty} \left\|\widehat{f} - \widehat{g}_n\right\|_{\mathbb{L}^2} = 0.$$

It is very common to find in standard textbooks, for instance in [13], the definition of the Fourier transform of a function $f \in \mathbb{L}^2(\mathbb{R})$ as

$$\widehat{f}(\xi) = \lim_{n\to+\infty} \int_{|x|\leq n} f(x)e^{-ix\xi}dx \qquad (\mathbb{L}^2 \textit{ sense}). \tag{3.17}$$

It is no more that considering the sequence $f_n = f|_{[-n,n]}$ in Definition 3.2.2.

Plancherel's theorem asserts that the Fourier transform is an *isometry* (minus the constant factor) in $\mathbb{L}^2(\mathbb{R})$. In our case, a *similarity.* A linear operator on $\mathbb{L}^2$ that is an isometry and maps onto $\mathbb{L}^2$ is called a *unitary operator.* The next result is commonly referred as the *Parceval Theorem.*

Theorem 3.2.7. *The Fourier transform is a unitary operator on $\mathbb{L}^2$.*

Corollary 3.2.7.1. *For any $g \in \mathbb{L}^2(\mathbb{R})$ there exists a unique $f \in \mathbb{L}^2(\mathbb{R})$ such that $\widehat{f} = g$.*

3.2.3 The space of rapidly decreasing functions

We want to extend the definition of Fourier transform much further. If $f \in \mathbb{L}^1(\mathbb{R}^n)$ and $\phi \in \mathscr{D}(\mathbb{R}^n)$ from (3.9) we have

$$\langle \widehat{f}, \phi \rangle = \langle f, \widehat{\phi} \rangle.$$

Thus, it is tempting to define the Fourier transform of any distribution $T \in \mathscr{D}^*(\mathbb{R}^n)$ by:

$$\langle \widehat{T}, \phi \rangle = \langle T, \widehat{\phi} \rangle.$$

But, if $\phi \in \mathscr{D}(\mathbb{R}^n)$, then $\widehat{\phi}$ cannot have compact support unless $\phi = 0$. (It is possible to show, see Lemma 4.3 of [10], that if $\phi \in \mathscr{D}(\mathbb{R}^n)$ then $\widehat{\phi}$ is an analytic function). So, with a convenient abuse of notation, $\widehat{\mathscr{D}} \not\subset \mathscr{D}$. We are looking for an explicit function space, say $\mathscr{S}$, such that $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n) \subset \mathbb{L}^1(\mathbb{R}^n)$ with the property of $\widehat{\mathscr{S}} \subseteq \mathscr{S}$. In other words, closed under the Fourier transform. Theorem 3.2.2 tells us that one basic requirement of this space $\mathscr{S}$ should be invariant under differentiation and multiplication by a polynomial. That is,

$$x^\alpha \partial^\beta \mathscr{S} \subset \mathscr{S}, \qquad \forall \alpha, \beta \in \mathbb{N}_0^n. \tag{3.18}$$

In this section we will discuss in some detail the space $\mathscr{S}$. As we will see, this appropriate space is given by considering functions which are smooth and well behaved at ∞. Namely, smooth functions φ such that $x^\alpha \partial^\beta \varphi(x)$ is bounded.

Definition 3.2.3. A function f is said to be *rapidly decreasing* if for every multi-index α, the application $x \to x^\alpha f(x)$ is bounded. Simply, f is rapidly decreasing if after multiplication by a polynomial x^α, $x^\alpha f(x)$ still goes to zero when $|x| \to \infty$. The space $\mathscr{S}(\mathbb{R}^n)$, often called *Schwartz Space*, is defined as the space of all $\varphi \in \mathscr{E}(\mathbb{R}^n)$ such that φ and all its partial derivatives are rapidly decreasing. That is

$$\forall \alpha, \beta \in \mathbb{N}_0^n \ \sup_{x \in \mathbb{R}} |x^\alpha \partial^\beta \varphi(x)| < \infty. \tag{3.19}$$

We will use ψ whenever we refer to a function of the $\mathscr{S}(\mathbb{R}^n)$ space. *The sequence* $(\psi_i)_{i \in \mathbb{N}}$ *converges to* ψ *in* $\mathscr{S}(\mathbb{R}^n)$ if for all multi-index α, β and for all compact $K \subset \mathbb{R}^n$ the sequence $(x^\alpha \partial^\beta \psi_i)_{i \in \mathbb{N}}$ converges uniformly to $x^\alpha \partial^\beta \psi$ in K (i.e $\|\psi_i - \psi\|_{\alpha,\beta} = \left\|x^\alpha \partial^\beta \psi_i - x^\alpha \partial^\beta \psi\right\|_{\infty,K} \to 0$).

It is clear from the definition that $\mathscr{S}(\mathbb{R}^n)$ is closed under multiplication by polynomial and differentiation as desired. Furthermore,

- $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n) \subset \mathscr{E}(\mathbb{R}^n)$;

- $\mathscr{D}(\mathbb{R}^n)$ is dense in $\mathscr{S}(\mathbb{R}^n)$;
- $\mathscr{S}(\mathbb{R}^n)$ is dense in $\mathbb{L}^p(\mathbb{R}^n)$ for $p \in [1, +\infty)$.

A classical example to show that $\mathscr{S}(\mathbb{R}^n) \setminus \mathscr{D}(\mathbb{R}^n) \neq \emptyset$ is the function $\psi = e^{-|x|^2}$. Note that the derivative of ψ is a polynomial times ψ. The exponential decreases fast enough to compensate any polynomial growth. Once again, the notion of convergence in $\mathscr{S}(\mathbb{R}^n)$ will not be a subject in this thesis. Nevertheless, it is of great importance to all the theory and we could not left it behind. For example, it is the key to prove that $\mathscr{S}$, as a metric space, is complete. Note that for $\psi \in \mathscr{S}(\mathbb{R}^n)$ we have $\psi \in \mathbb{L}^1(\mathbb{R}^n)$, so $\widehat{\psi}$, as in (3.7), is well defined.

Definition 3.2.4. The *Fourier transform* of a function $\psi \in \mathscr{S}(\mathbb{R}^n)$ is defined as the function in $\mathbb{L}^\infty(\mathbb{R}^n)$ given by

$$\widehat{\psi}(\xi) := \int_{\mathbb{R}^n} \psi(x) e^{-ix\cdot\xi}\, dx, \qquad (\xi \in \mathbb{R}^n), \tag{3.20}$$

The reason that $\mathscr{S}(\mathbb{R}^n)$ is such an important function space in the study of Fourier transform is that the operation preserves the class. With our standard abuse of notation, $\widehat{\mathscr{S}} \subset \mathscr{S}$. So we can work with Fourier transform and its inversion formula without worrying with technical details, since $\widehat{\mathscr{S}}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n) \subset \mathbb{L}^1(\mathbb{R}^n)$. We summarize the last two paragraphs in the following two results found, for instance, in [6] and [7].

Theorem 3.2.8. *If $\psi \in \mathscr{S}(\mathbb{R}^n)$ then $\widehat{\psi} \in \mathscr{S}(\mathbb{R}^n)$.*

Theorem 3.2.9. *The Fourier transform is an isomorphism on $\mathscr{S}$ with inverse given by (3.14). Hence, $\forall \psi \in \mathscr{S}(\mathbb{R}^n)$, $\widehat{\widehat{\psi}} = (2\pi)^n \check{\psi}$.*

In addition to Theorem 3.2.2, we can synthesize the most import properties of the Fourier transform in the following theorem.

Theorem 3.2.10. *For $\psi, \upsilon \in \mathscr{S}(\mathbb{R}^n)$:*

1.

$$\int_{\mathbb{R}^n} \widehat{\psi}(x)\upsilon(x)\, dx = \int_{\mathbb{R}^n} \psi(\xi)\widehat{\upsilon}(\xi)\, d\xi, \tag{3.21}$$

2.

$$\int_{\mathbb{R}^n} \psi(x)\upsilon(x)\,dx = (2\pi)^{-n}\int_{\mathbb{R}^n} \widehat{\psi}(\xi)\widehat{\upsilon}(\xi)\,d\xi, \tag{3.22}$$

3.

$$\psi * \upsilon \in \mathscr{S} \quad and \quad \widehat{\psi * \upsilon} = \widehat{\psi}\cdot\widehat{\upsilon}, \tag{3.23}$$

4.

$$\widehat{\psi\upsilon} = (2\pi)^{-n}\widehat{\psi} * \widehat{\upsilon}. \tag{3.24}$$

Lemma 3.2.11. *The Fourier transform of the Gaussian function $x \to exp(-|x|^2/2)$ in $\mathscr{S}(\mathbb{R}^n)$ is given by*

$$(\widehat{e^{-|x|^2/2}})(\xi) = (2\pi)^{n/2}e^{-|\xi|^2/2}. \tag{3.25}$$

Proof. See, for instance, [8] and [18]. ■

3.2.4 Tempered Distributions

By this time, we have collected enough information to construct an appropriate extension of Fourier transform to distribution theory. In the previous section, we familiarized the reader with the similarities between the spaces $\mathscr{D}(\mathbb{R}^n)$ and $\mathscr{S}(\mathbb{R}^n)$. Distributions were defined as continuous linear functions of $\mathscr{D}(\mathbb{R}^n)$. We may expect that equivalent objects of $\mathscr{S}(\mathbb{R}^n)$ should be of interest. And, in fact, they are. The dual space of $\mathscr{S}(\mathbb{R}^n)$, $\mathscr{S}^*(\mathbb{R}^n)$, that is, the set of all linear continuous funtionals $\mathscr{S}(\mathbb{R}^n) \to \mathbb{R}$ is called space of *tempered distributions.*

From the definition, $\mathscr{S}^*(\mathbb{R}^n)$ is closed under multiplication by polynomials and by differentiation, i.e (3.18) holds for every $S \in \mathscr{S}^*(\mathbb{R}^n)$. Furthermore, we have the following inclusions.

- $\mathscr{E}^*(\mathbb{R}^n) \subset \mathscr{S}^*(\mathbb{R}^n) \subset \mathscr{D}^*(\mathbb{R}^n)$,
- $\mathbb{L}^p(\mathbb{R}^n) \subset \mathscr{S}^*(\mathbb{R}^n)$ for $p \geq 1$.

At this stage, it seems logical to ask why are tempered distributions so important to us. The answer relies on the fact that it is possible to define the Fourier transform of a tempered distribution as a tempered distribution and it is not possible to define the Fourier transform of a more general distribution as a distribution (if $\phi \in \mathscr{D}(\mathbb{R})$ and $\widehat{\phi} \in \mathscr{D}(\mathbb{R})$, then $\phi \equiv 0$). We shall finally introduce the generalization of the Fourier transform to distributions.

Definition 3.2.5. Let $S \in \mathscr{S}^*(\mathbb{R})$. Its *Fourier transform*, $\widehat{S}$, is defined by

$$\langle \widehat{S}, \psi \rangle := \langle S, \widehat{\psi} \rangle, \qquad (\forall \psi \in \mathscr{S}(\mathbb{R}^n)). \tag{3.26}$$

In other words, $\widehat{S}$ is the functional on $\mathscr{S}(\mathbb{R}^n)$ that assign to ψ the value $\langle S, \widehat{\psi} \rangle$. The definition is, of course, set up in such a way that it is consistent with the other cases that we have already studied. It should be clear that, for instance, in the case of $S \in \mathbb{L}^1(\mathbb{R}^n)$, that $\widehat{S}$ coincides with the classical notion of Fourier transform as a $\mathbb{L}^\infty(\mathbb{R}^n)$ function. Note also that for $S \in \mathbb{L}^1(\mathbb{R}^n)$ we have $S \in \mathscr{S}^*(\mathbb{R}^n)$. By (3.8), $\widehat{S} \in \mathbb{L}^\infty(\mathbb{R}^n) \subset \mathscr{S}^*(\mathbb{R}^n)$. With the usual abuse of notation, $\widehat{\mathscr{S}^*} \subset \mathscr{S}^*$.

Theorem 3.2.12. *The Fourier transform $\mathscr{F} \colon \mathscr{S}^* \to \mathscr{S}^*$ is an isomorphism with Fourier inversion formula*

$$\widehat{\widehat{S}} = (2\pi)^n \check{S}. \tag{3.27}$$

We have seen that the Fourier transform defines an isomorphism of $\mathscr{S}$ onto $\mathscr{S}$, of $\mathbb{L}^2$ onto $\mathbb{L}^2$ and now of $\mathscr{S}^*$ onto $\mathscr{S}^*$. Also $\widehat{\mathbb{L}^1} \subset \mathbb{L}^\infty$ and $\widehat{\mathscr{D}} \nsubseteq \mathscr{D}^*$. The reader may ask about the $\widehat{\mathscr{E}}$. This last case is known in literature as *Fourier-Laplace transform.*

Theorem 3.2.13. *For a distribution with compact support E, $\widehat{E}(\xi)$ is defined for all $\xi \in \mathbb{C}^n$ and coincides with the function*

$$\widehat{E}(\xi) := \langle \underbrace{E}_{\in \mathscr{E}^*}, \underbrace{e^{-ix\cdot\xi}}_{\in \mathscr{E}} \rangle. \tag{3.28}$$

From complex analysis, one can show that $\widehat{E}(\xi)$ satisfies the Cauchy-Riemann equations and thus it is a holomorphic function of $\xi_j \in \mathbb{C}$, *ceteris paribus.* It follows here that the support of $\widehat{E}$ cannot be compact unless $E \equiv 0$. The spaces of holomorphic functions $\widehat{\mathscr{D}}$ (resp.($\widehat{\mathscr{E}}$)), may be characterized by the growth properties in ξ. For the study of such spaces, we recommend the reading of the The Paley-Wiener theory (see e.g. the book of L. Hörmander [7]).

Example 3.2.1. For $\psi \in \mathscr{S}(\mathbb{R}^n)$,

$$\begin{aligned}\langle \widehat{\delta}, \psi \rangle &:= \langle \delta, \widehat{\psi} \rangle = \widehat{\psi}(0) \\ &= \int \psi(x) \underbrace{e^{-i0\xi}}_{=1} \, d\xi \\ &= \langle 1, \psi \rangle.\end{aligned}$$

Therefore,

$$\widehat{\delta} = 1.$$

Moreover, since $\check{\delta} = \delta$ and $\check{u} = (2\pi)^{-n}\widehat{\widehat{u}}$, we may further deduce that

$$\begin{aligned}\langle \delta, \psi \rangle &= \langle \delta, \check{\psi} \rangle = (2\pi)^{-n} \langle \widehat{\delta}, \widehat{\psi} \rangle \\ &= (2\pi)^{-n} \langle 1, \widehat{\psi} \rangle \\ &= \langle (2\pi)^{-n} \widehat{1}, \psi \rangle.\end{aligned}$$

Hence, in the distribution sense,

$$\widehat{1} = (2\pi)^n \delta.$$

The properties of the Fourier transform that we have seen so far, namely in Theorem 3.2.2 and Theorem 3.2.10, naturally hold on $\mathscr{S}^*(\mathbb{R}^n)$.

Example 3.2.2. Consider the Heaviside distribution $H \in \mathscr{S}^*(\mathbb{R})$.
Bearing in mind that $H' = \delta_0$, we have $i\xi \widehat{H} = \widehat{\delta} = 1$ and hence $\langle \widehat{H}, \phi \rangle = \langle \frac{-i}{\xi}, \phi \rangle$ when 0 does not belong to $supp\,\phi$. Furthermore, since the equality $\xi \cdot vp(1/\xi) = 1$ holds for distributions,

$$\xi \left(\widehat{H}(\xi) + ivp\left(\frac{1}{\xi}\right) \right) = 0.$$

Here we can use the result that if $\xi u(\xi) = 0$ then $u = c\delta$ with c a complex constant. So, it remains to determine the constant c in the equation $\widehat{H} + i \cdot vp(1/\xi) = c\delta$. Recall that

$$\check{\delta} = \delta, \; \check{H} = 1 - H, \; \check{vp}\left(\frac{1}{\xi}\right) = -vp\left(\frac{1}{\xi}\right).$$

Thus we can compute

$$\begin{aligned} c\delta = c\check{\delta} &= \widehat{\check{H}} + i \cdot \widecheck{vp}\,(1/\xi) \\ &= \widehat{(1-H)} \, - i \cdot vp\,(1/\xi) \\ &= \widehat{1} - (\underbrace{\widehat{H} + i \cdot vp(1/\xi)}_{c\delta}) = 2\pi\delta - c\delta = (2\pi - c)\delta. \end{aligned}$$

Therefore, $c = \pi$ and

$$\widehat{H} = \pi\delta - i \cdot vp\left(\frac{1}{\xi}\right). \tag{3.29}$$

Example 3.2.3. Since $\widehat{\widehat{H}} = 2\pi\check{H} = 2\pi(1-H)$, we compute $\widehat{vp(1/\xi)}$ from (3.29) as it follows:

$$\begin{aligned} i \cdot \widehat{vp}\,(1/\xi) &= \pi\widehat{\delta} - \widehat{\widehat{H}} = \pi - 2\pi\check{H} \\ &= \pi - 2\pi(1-H) \\ &= \pi\,\underbrace{(2H-1)}_{sgn(\xi)}. \end{aligned}$$

Therefore, in the distribution sense, $\widehat{vp(1/\xi)} = -i\pi sgn$.

Example 3.2.4. *Heisenberg Uncertainty Principle*

Suppose that ψ is a function in $\mathbb{L}^2(\mathbb{R})$ with $\|\psi\|^2_{\mathbb{L}^2} = 1$. For $x_0, \xi_0 \in \mathbb{R}$ we have

$$\int_{\mathbb{R}} (x - x_0)^2 |\psi(x)|^2 \, dx \cdot \int_{\mathbb{R}} (\xi - \xi_0)^2 |\widehat{\psi}(\xi)|^2 \, d\xi \geq c > 0. \tag{3.30}$$

Proof. Assume, without any lost of generality that $x_0 = \xi_0 = 0$.

$$1 = \|\psi\|^2_{\mathbb{L}^2} = \int_{\mathbb{R}} \psi^2(x) dx = \int_{\mathbb{R}} x' \cdot \psi^2(x) dx = -2 \int_{\mathbb{R}} x \cdot \psi(x) \cdot \psi_x(x) \, dx,$$

which implies

$$\begin{aligned} \frac{1}{2} = \left| \int_{\mathbb{R}} x \cdot \psi(x) \cdot \psi_x(x) dx \right| &\leq \int_{\mathbb{R}} |x \cdot \psi(x)| \cdot |\psi_x(x)| \, dx \\ &\leq \|x \cdot \psi\|_{\mathbb{L}^2} \cdot \|\psi_x\|_{\mathbb{L}^2} \\ &\leq \|x \cdot \psi\|_{\mathbb{L}^2} \cdot \frac{1}{\sqrt{2\pi}} \cdot \left\| \xi \cdot \widehat{\psi} \right\|_{\mathbb{L}^2} \end{aligned}$$

Thus,

$$c = \frac{\pi}{2} \leq \int_{\mathbb{R}} x^2 \cdot |\psi(x)|^2 \, dx \cdot \int_{\mathbb{R}} \xi^2 \cdot |\widehat{\psi}(\xi)|^2 \, d\xi$$

∎

The fact that $\|f\|_{\mathbb{L}^2}^2 = \int_{\mathbb{R}} |f(x)|^2\, dx = 1$ motivates one to think in $|f(x)|^2$ as a probability density function which lead to the interpretation of the first left-hand side integral as the standard deviation. With the proper definition of the Fourier transform, Plancherel's theorem asserts that $\left\|\widehat{f}\right\|_{\mathbb{L}^2}^2 = \int_{\mathbb{R}} |\widehat{f}(\xi)|^2\, d\xi = 1$. Thus, $|\widehat{f}(\xi)|^2$ is also a probability density function and the correspondent integral, as in (3.30), is interpreted as the standard deviation. It is common the use of the square of a function, rather than the function itself, to guarantee that we have a positive function. As we know, the smaller the standard deviation of a probability distribution is, the more precisely one can predict the outcome of the associated random event.

Equation (3.30) shows that if we have ψ concentrated more and more around x_0, the first quantity in the inequality goes to zero, which implies that the second quantity of the inequality must necessarily "explode".

The Heisenberg Uncertainty Principle, formulated in 1927, is a key principle in quantum mechanics. It says basically and quoting Aladin "A statement about the limitations of one's ability to perform measurements on a system without disturbing it". In a nutshell, the principle says the more we know about the particle's position (uncertainty of the position is small) the less we know about the particle's momentum (uncertainty of the momentum is large), and vice-versa. Position and momentum of a particle cannot be measured with high precision at the exactly same time. Thus, it is no more than (3.30) applied to quantum physics.

In classical mechanics, the pair $(\vec{x}, \vec{p})$ where $\vec{x}$ is the position vector and $\vec{p} = m\vec{v}$ is the momentum, it is enough to describe the state of a system. In quantum theory, it is quite different. One may assert that it is possible for the same particle (for instance, an electron e^-) to be in two different places at once. Without further details (we leave that for the most curious reader) the system is described by a function $\psi(x,t)$ called *wave function* which satisfies the famous *Schrödinger equation* $i\psi_t + \Delta_x \psi = 0$ (compare with the heat equation). Physicists assert that in the quantum world, the equivalent quantity to momentum is given by the Fourier transform of the wave function $\psi(x,t)$.

Classic mechanics		Quantum mechanics
$\vec{x}$	$\underbrace{\leftrightarrow}_{Position}$	$\psi(x,t)$
$\vec{p}$	$\underbrace{\leftrightarrow}_{Momentum}$	$\widehat{\psi}(\xi,t)$

Max Born was the first to relate the wave function with probability amplitudes. In fact, the square of the magnitude of the wave function describes the probability of the electron e^- to exist in a particular region of the domain $[a, b]$.

$$P\left(e^- \in [a,b]\right) = \int_a^b |\psi(x,t)|^2 dx$$

Since the electron must be somewhere, we have of course

$$\forall t, \int_{\mathbb{R}} |\psi(x,t)|^2 dx = 1.$$

Thus, $|\psi(x,t)|^2$ is the probability density function of the system and therefore, the wave function satisfies the conditions of the above theorem.

If we consider the uncertainty of the position and momentum (Δ_x, Δ_p) as standard deviations then, with some adjustments in the constants, equation (3.30) can be re-written and lead us to the following theorem.

Theorem 3.2.14. *Heisenberg Uncertainty Principle.*

$$\Delta_x \cdot \Delta_\xi \geq \frac{h}{4\pi}, \tag{3.31}$$

where h is the Planck constant.

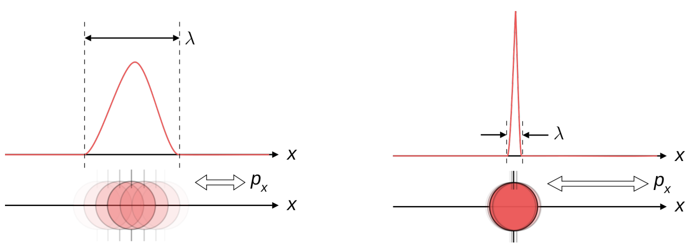

Figure 3.1. *An Illustration of Heisenberg's uncertainty principle*

Chapter 4

Sobolev Spaces

In this chapter, we introduce the Sobolev Spaces, which consist of functions that fulfill certain integrability properties jointly with their derivatives. Distributions theory and Fourier transform allow us to find solutions of problems that a priori had no solutions. However, we would like to be able to distinguish among these solutions the ones that are smooth. As we saw, the regularity of a distribution is connected with the behaviour of its Fourier transform at high frequencies. Roughly speaking, the more derivatives f has the faster $\widehat{f}$ decreases to zero at infinity while the faster f decreases to zero at infinity the more derivatives $\widehat{f}$ has. Quoting [18]: "If doing "ping" to f does "pong" to $\widehat{f}$ then we should expect that doing "pong" to f does "ping" to $\widehat{f}$."

$f(x)$		$\widehat{f}(\xi)$
smooth ($\mathscr{E}$)	$\leftrightarrow$	decays rapidly as $\lvert\xi\rvert \to \infty$
decays rapidly as $\lvert x\rvert \to \infty$	$\leftrightarrow$	smooth ($\mathscr{E}$)

In the mid 1930's, Sergei Sobolev provided his theory about what would be the appropriated functional spaces to work and solve PDEs. The existing literature on Sobolev Spaces and their generalization is vast. For a more complete study on this subject, we suggest the reading of [14] and [19].

4.1 Sobolev spaces of natural order m

Definition 4.1.1. Let $p \geq 1$ and $m \in \mathbb{N}_0$. The *Sobolev space* $W^{m,p}(\Omega)$ consists of the functions $f \in \mathbb{L}^p(\Omega)$ such that for every multi-index α with $|\alpha| \leq m$, the derivative $\partial_x^\alpha f$ exists and belongs to $\mathbb{L}^p(\Omega)$. Thus,

$$W^{m,p}(\Omega) := \big\{ f \in \mathbb{L}^p(\Omega) : \partial_x^\alpha f \in \mathbb{L}^p(\Omega), |\alpha| \leq m \big\}, \tag{4.1}$$

endowed with the norm

$$\|f\|_{W^{m,p}} = \left(\sum_{|\alpha| \leq m} \int_\Omega |\partial_x^\alpha f|^p dx \right)^{\frac{1}{p}}, 1 \leq p < \infty, \tag{4.2}$$

and

$$\|f\|_{W^{m,\infty}} = \sum_{|\alpha| \leq m} \|\partial_x^\alpha f\|_\infty. \tag{4.3}$$

There are several equivalent norms that may be used as well. For example, [4] and [14] uses

$$\|f\|_{W^{m,p}} = \sum_{|\alpha| \leq m} \|\partial_x^\alpha f\|_{\mathbb{L}^p}, \qquad \|f\|_{W^{m,\infty}} = \max_{|\alpha| \leq m} \|\partial_x^\alpha f\|_{\mathbb{L}^\infty}. \tag{4.4}$$

A few notes about this space. There should be no doubt that $W^{m,p}(\Omega)$ is a linear space and for $m = 0$, $W^{0,p}(\Omega) = \mathbb{L}^p(\Omega)$. The derivatives that appear in the definition are, of course, derivatives in the distributional sense. They are often called *weak derivatives*. Thus, the statement $\partial_x^\alpha f \in \mathbb{L}^p(\Omega)$ means that the distribution $\partial_x^\alpha f \in \mathscr{D}^*(\Omega)$ can be represented by a $\mathbb{L}^p$ function. Also, it is clear that if $f \in W^{m,p}(\Omega)$ then $\partial_x^\alpha f \in W^{m-|\alpha|,p}(\Omega)$. The next theorem states some of the basic properties of the Sobolev spaces.

Theorem 4.1.1. *Let $1 < p < \infty$. The space $W^{m,p}(\Omega)$ is a separable and reflexive Banach space.*

Furthermore, one may prove that $\mathscr{E}(\Omega) \cap W^{m,p}(\Omega)$ is dense in $W^{m,p}(\Omega)$ in a result known as the *Meyers-Serrin theorem.* As we mentioned, the space $\mathbb{L}^2$ is a natural setting for operations like convolution and Fourier transform. In this context, we will denote

$$\mathbb{H}^m(\Omega) = W^{m,2}(\Omega) = \big\{ f \in \mathbb{L}^2(\Omega) : \partial_x^\alpha f \in \mathbb{L}^2(\Omega), |\alpha| \leq m \big\} \tag{4.5}$$

Theorem 4.1.2. *The spaces $\mathbb{H}^m(\Omega)$ are Hilbert spaces when endowed with the scalar product*

$$(f|g)_{\mathbb{H}^1} = \sum_{|\alpha|\leq m} \int_\Omega \partial_x^\alpha f \cdot \partial_x^\alpha g \, dx.$$

We are particularly interested in the case $m = 1$. To formalize it, the *Sobolev space of order 1* is the Banach space

$$\mathbb{H}^1(\Omega) = W^{1,2}(\Omega) = \left\{ f \in \mathbb{L}^2(\Omega) : \partial_x f \in \mathbb{L}^2(\Omega) \right\} \tag{4.6}$$

with

$$\|f\|_{\mathbb{H}^1} = \left(\int_\Omega |f|^2 dx + \int_\Omega |\partial_x f|^2 dx \right)^{\frac{1}{2}}. \tag{4.7}$$

Corollary 4.1.2.1. *We have that $\mathscr{E}(\Omega)$ is dense in $\mathbb{H}^m(\Omega)$ and, for $\Omega = \mathbb{R}^n$, $\mathscr{D}(\mathbb{R}^n)$ is dense in $\mathbb{H}^1(\mathbb{R}^n)$.*

If $\Omega \subset \mathbb{R}$ is a bounded interval, then each function of $\mathbb{H}^1(\Omega)$ is continuous and bounded in Ω. For higher dimensions ($\Omega \subset \mathbb{R}^n$ with $n \geq 2$) the functions of $\mathbb{H}^1(\Omega)$ are, in general, not continuous.

4.2 The Gagliardo-Nirenberg Inequality

With strong assumptions, it is possible to prove relationships between some Sobolev spaces. These theorems are usual called *Sobolev imbedding theorems*. The reader will find, for sure, innumerous versions of such theorems, each one specific to the case of study. Our goal in this section is to analyze the relation between the space $\mathbb{H}^1$ and other functional spaces of interest. But first, a couple of introdutory results. As already seen, for $p \in [1, \infty)$ and $k \leq m$, we have $W^{m,p}(\Omega) \subset W^{k,p}(\Omega)$. Furthermore, for a finite measure domain, $p, q \in [1, \infty]$ with $q > p$ and $m \geq 0$, $W^{m,q}(\Omega) \subseteq W^{m,p}(\Omega)$.

In order to keep simplicity, let us work from now on and until the end of this section in $\mathbb{R}$. Consider $u \in \mathbb{H}^1(\mathbb{R})$. By simple use of calculus

$$u^2(x) = \int_{-\infty}^x \left(u^2(t)\right)' dt = 2 \int_{-\infty}^x u'(t) u(t) dt.$$

Hölder's inequality implies

$$\begin{aligned}\left|u^2(x)\right| = |u(x)|^2 = 2\left|\int_{-\infty}^{x} u'(t)u(t)dt\right| \\ \leq 2\left|\int_{\mathbb{R}} u'(t)u(t)dt\right| \\ \leq 2\cdot \|u\|_{\mathbb{L}^2}\cdot \|u'\|_{\mathbb{L}^2},\end{aligned}$$

and therefore

$$\|u\|_{\mathbb{L}^\infty} \leq \sqrt{2}\cdot \|u\|_{\mathbb{L}^2}^{\frac{1}{2}}\cdot \|u'\|_{\mathbb{L}^2}^{\frac{1}{2}}. \tag{4.8}$$

The norm defined in (4.7) allow us to write,

$$\begin{cases} \|u\|_{\mathbb{L}^2} \leq \|u\|_{\mathbb{H}^1} \\ \|u'\|_{\mathbb{L}^2} \leq \|u\|_{\mathbb{H}^1}. \end{cases} \tag{4.9}$$

Substituting on (4.8), we may conclude

$$\|u\|_{\mathbb{L}^\infty} \leq \sqrt{2}\|u\|_{\mathbb{H}^1}. \tag{4.10}$$

We have the *Sobolev injection* $\mathbb{H}^1(\mathbb{R}) \hookrightarrow \mathbb{L}^\infty(\mathbb{R})$. Similar computations can be done for an arbitrary $p \geq 2$.

$$\begin{aligned}\|u\|_{\mathbb{L}^p}^p = \int_{\mathbb{R}} |u(x)|^p\,dx = \int_{\mathbb{R}} \left|u^{p-2}(x)u^2(x)\right|dx \\ \leq \|u\|_{\mathbb{L}^\infty}^{p-2}\int_{\mathbb{R}} u^2(x)dx \\ \leq \left(\sqrt{2}\cdot \|u\|_{\mathbb{L}^2}^{\frac{1}{2}}\cdot \|u'\|_{\mathbb{L}^2}^{\frac{1}{2}}\right)^{p-2}\|u\|_{\mathbb{L}^2} \\ \leq \sqrt{2}^{p-2}\cdot \|u\|_{\mathbb{L}^2}^{\frac{p+2}{2}}\cdot \|u'\|_{\mathbb{L}^2}^{\frac{p-2}{2}}.\end{aligned}$$

This implies

$$\|u\|_{\mathbb{L}^p} \leq (\sqrt{2})^{\frac{p-2}{2p}}\cdot \|u\|_{\mathbb{L}^2}^{\frac{p+2}{2p}}\cdot \|u'\|_{\mathbb{L}^2}^{\frac{p-2}{2p}} \tag{4.11}$$

By (4.9), we conclude

$$\|u\|_{\mathbb{L}^p} \leq (\sqrt{2})^{\frac{p-2}{2p}}\cdot \|u\|_{\mathbb{H}^1} \tag{4.12}$$

Therefore, for $p \geq 2$ we have the *Sobolev injection* $\mathbb{H}^1(\mathbb{R}) \hookrightarrow \mathbb{L}^p(\mathbb{R})$. Inequality (4.11) is known as the *Gagliardo-Nirenberg Inequality.* To see the generalization of such inequality for $\Omega \subset \mathbb{R}^n$ we recommend the reading of the article of Haïm Brezis and Petru Mironescu, *Gagliardo-Nirenberg inequalities and non-inequalities: the full story* - [2].

4.3 Sobolev Spaces of arbitrary order

We have already defined the Sobolev spaces of natural order m. Actually, one may define it for any $m = s \in \mathbb{R}$. In this section, we aim to give the reader just a little taste of the universe that constitutes Sobolev spaces.

Definition 4.3.1. The *Sobolev spaces with zero boundary value* denoted by $H_0^m(\Omega)$ is the closure of $\mathscr{D}(\Omega)$ in $\mathbb{H}^m(\Omega)$, endowed with the norm of $\mathbb{H}^m$.

The properties of 4.1.1 also holds for the space $H_0^m(\Omega)$ which is indeed a close subspace of $H^m(\Omega)$. With the previous argumentation, it is clear that $H_0^m(\mathbb{R}) = H^m(\mathbb{R})$, that is, the standard Sobolev space and the Sobolev space with zero boundary value coincide in the whole space. Furthermore, we have

$$\mathscr{D}(\Omega) \subset \mathbb{H}_0^1(\Omega) \subset \mathbb{H}^1(\Omega) \subset \mathbb{H}^0(\Omega) = \mathbb{L}^2(\Omega) \subset \mathscr{D}^*(\Omega)$$

Morally, the difference between $\mathbb{H}_0^1(\Omega)$ and $\mathbb{H}^1(\Omega)$ is that the functions of $\mathbb{H}_0^1(\Omega)$ may be approximated by $\mathscr{D}(\Omega)$ functions instead of $\mathscr{E}(\Omega)$ functions.

Definition 4.3.2. *Sobolev spaces of negative order.* For $m \in \mathbb{N}$, the Sobolev space $H^{-m}(\Omega)$ is defined as the dual space of $H_0^m(\Omega)$, i.e. $H^{-m}(\Omega) := H^m(\Omega)^*$, equiped with the norm

$$\|u\|_{H^{-m}} := \sup_{v \in \mathbb{H}^m : \|v\|_{\mathbb{H}^m} \neq 0} \frac{(u|v)_{\mathbb{H}^1}}{\|v\|_{\mathbb{H}^m}}.$$

Since $\mathscr{D}(\Omega) \subset \mathbb{H}_0^m(\Omega)$ it comes that $\mathbb{H}^{-m}(\Omega) \subset \mathscr{D}^*(\Omega)$. Thus, the elements of $\mathbb{H}^{-m}(\Omega)$ are distributions. It is common to find in the literature the definition of $\mathbb{H}^{-m}(\Omega)$ as follows.

$$\mathbb{H}^{-m}(\Omega) := \{u \in \mathscr{D}^*(\Omega) : \|u\|_{H^{-m}} < \infty\}.$$

Just as in the case of $\mathbb{H}^m$, $\mathbb{H}^{-m}$ are Hilbert spaces. Furthermore, we have the inclusions

$$\cdots \mathbb{H}^2(\Omega) \subset \mathbb{H}^1(\Omega) \subset \mathbb{L}^2(\Omega) \subset \mathbb{H}^{-1}(\Omega) \subset \mathbb{H}^{-2}(\Omega) \cdots$$

Before moving to the more general Sobolev space, let us introduce the following notation. Let $s \in \mathbb{R}$ and $\xi \in \mathbb{R}^n$. We shall write

$$\lambda \equiv \lambda(\xi) = (1 + |\xi|^2)^{\frac{1}{2}} \therefore \lambda^s = (1 + |\xi|^2)^{\frac{s}{2}}.$$

Definition 4.3.3. Let $s \in \mathbb{R}$. We define the *Sobolev space* $\mathbb{H}^s(\mathbb{R}^n)$ (also known as *Bessel potential space*) by

$$\mathbb{H}^s(\mathbb{R}^n) = \{S \in \mathscr{S}^*(\mathbb{R}^n) : \lambda^s \widehat{S} \in \mathbb{L}^2(\mathbb{R}^n)\}$$

equiped with the norm

$$\|.\|_{\mathbb{H}^s} = (2\pi)^{\frac{-n}{2}} \cdot \left\| \lambda^s \widehat{S} \right\|_{\mathbb{L}^2}.$$

Just like the case before, the space $\mathbb{H}^{-s}$ are defined as the topological dual space of $\mathbb{H}^s$, that is $H^{-s}(\mathbb{R}^n) := H^s(\mathbb{R}^n)^*$.

Theorem 4.3.1. *Sobolev Embedding Theorem. Define by $C^k_{\hookrightarrow 0}(\mathbb{R}^n)$ the set of C^k functions on $\mathbb{R}^n$ that tends to zero at infinity as well as all their derivative of order less or equal to k. We have*

$$For\ s > \frac{n}{2} + k,\ \mathbb{H}^s(\mathbb{R}^n) \hookrightarrow C^k_{\hookrightarrow 0}(\mathbb{R}^n).$$

In particular,

$$For\ s > \frac{1}{2},\ \mathbb{H}^s(\mathbb{R}) \hookrightarrow C_{\hookrightarrow 0}(\mathbb{R}).$$

Example 4.3.1. We have

$$\delta \in \mathbb{H}^{-s}(\mathbb{R}^n) \Leftrightarrow s > \frac{n}{2}.$$

Indeed, $\widehat{\delta} \in \mathbb{L}^1_{loc}$.

$$\delta \in \mathbb{H}^{-s}(\mathbb{R}^n) \Leftrightarrow \lambda^{-s}\widehat{\delta} \in \mathbb{L}^2(\mathbb{R}^n) \underbrace{\Leftrightarrow}_{\widehat{\delta}=1} \lambda^{-s} \leq (1 + |\xi|^2)^{-\frac{s}{2}} \in \mathbb{L}^2(\mathbb{R}^n)$$
$$\Leftrightarrow \int_{\mathbb{R}^n} \frac{1}{(1 + |\xi|^2)^{\frac{s}{2}}}\, d\xi < \infty \Leftrightarrow 2s > n.$$

Chapter 5

Solving the Heat Equation

In this chapter, we will solve the heat diffusion equation using the tools provided by distributions theory and Fourier transform. Under the assumption that the solution is well-behaved so that it satisfies the hypothesis of the Fourier transform theory, our goal is to find the solution of the following Cauchy problem with initial data $h \in \mathbb{L}^1$:

$$\begin{cases} u_t(x,t) - \Delta_x u(x,t) = g(x,t) & x \in \mathbb{R}, t > 0 \\ u(x,0) = h(x) & x \in \mathbb{R} \end{cases}. \tag{5.1}$$

The idea is quite simple. Divide (5.1) in two sub-problems, one homogeneous with non null initial condition and other non homogeneous with the null initial condition. Through linearity, the solution of (5.1) will be the sum of the solution of both sub-problems.

In chapter 3, we proved that Φ defined by equation (2.2.4) is a distribution of $\mathbb{R}^2$. In fact, we will prove now that $\Phi(\cdot, t) * h(x)$ is the solution of the one-dimensional initial value problem

$$\begin{cases} u_t(x,t) - \Delta_x u(x,t) = 0 & x \in \mathbb{R}, t > 0 \\ u(x,0) = h(x) & x \in \mathbb{R} \end{cases}. \tag{5.2}$$

Applying the Fourier transform in the spacial variable x,

$$\widehat{u}_t(\xi,t) - \widehat{\Delta_\xi u}(\xi,t) = 0 \underbrace{\Leftrightarrow}_{(4.13)} \widehat{u}_t(\xi,t) + \xi^2 \widehat{u}(\xi,t) = 0.$$

But this is a linear ODE in ξ. The solution is well known and given by

$$\widehat{u}(\xi,t) = ae^{-\xi^2 t},\ a \in \mathbb{R}. \tag{5.3}$$

Employing the initial condition on equation (5.3), it comes straightforward that

$$\widehat{u}(\xi,t) = \widehat{h}(\xi)e^{-\xi^2 t}. \tag{5.4}$$

By this time, it seems that we almost solve our problem. But what we really want is the solution in terms of the initial variable x. We are looking for a function $u(x,t)$ whose Fourier transform is $\widehat{h}(\xi)e^{-\xi^2 t}$. To find such $u(x,t)$ we just apply the inverse of Fourier transform to both sides of equation (5). That is

$$\begin{aligned}\mathscr{F}^{-1}(\widehat{u}) &= \mathscr{F}^{-1}(\widehat{h}(\xi)e^{-\xi^2 t}) \\ &= \mathscr{F}^{-1}(\widehat{h}) * \mathscr{F}^{-1}(e^{-\xi^2 t}) \\ &= h(x) * \mathscr{F}^{-1}(e^{-\xi^2 t}). \end{aligned} \tag{5.5}$$

For now, let us leave (5.5) in standby and compute $\mathscr{F}^{-1}(e^{-\xi^2 t})$. This is a common computation in this kind of exercise. The reader may find several ways to solve it, some easier than others, given the used definition of Fourier transform and its inverse formula. With our definition, it comes that

$$\mathscr{F}^{-1}(e^{-\xi^2 t}) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix\xi}e^{-\xi^2 t}d\xi = \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix\xi-\xi^2 t}d\xi$$

The exponent $ix\xi - \xi^2 t$ is quadratic in ξ. So, it is natural to complete the square. By simple computations,

$$ix\xi - \xi^2 t = -t\Big(\xi^2 - \frac{2ix\xi}{2t} - \frac{x^2}{4t^2}\Big) - \frac{x^2}{4t}.$$

Therefore,

$$\mathscr{F}^{-1}(e^{-\xi^2 t}) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix\xi-\xi^2 t}d\xi = \frac{1}{2\pi}e^{\frac{-x^2}{4t}}\int_{\mathbb{R}} e^{-t(\xi-\frac{ix}{2t})^2}d\xi.$$

Note that the right-hand side integral is the integral of a Gaussian. It can be calculated using, for example, complex analysis and polar coordinates. The result is known explicitly, given by $\sqrt{\frac{\pi}{t}}$. Returning to (5.5), we obtain

$$\begin{aligned} u(x,t) &= \frac{1}{2\pi} e^{\frac{-x^2}{4t}} \mathscr{F}^{-1}(e^{-\xi^2 t}) * h(x) \\ &= \frac{1}{2\pi} e^{\frac{-x^2}{4t}} \sqrt{\frac{\pi}{t}} * h(x) \\ &= \frac{1}{\sqrt{4\pi t}} e^{\frac{-x^2}{4t}} * h(x) := \Phi(x,t) * h(x). \end{aligned} \tag{5.6}$$

The function $\Phi(x,t)$ is commonly referred on literature as the *Heat Kernel*.

5.1 Homogeneous Problem

Theorem 5.1.1. *A solution of the homogeneous heat equation for initial data $h(x) \in \mathbb{L}^1(\mathbb{R})$*

$$\begin{cases} u_t(x,t) - \Delta_x u(x,t) = 0 & x \in \mathbb{R}, t > 0 \\ u(x,0) = h(x) & x \in \mathbb{R} \end{cases}.$$

is given by the convolution of the initial data with the heat kernel

$$\Phi(x,t) * h(x) = \int_{\mathbb{R}} \Phi(x-y,t)h(y)dy = \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} e^{\frac{-(x-y)^2}{4t}} h(y)\, dy. \tag{5.7}$$

Proof. Let $x \in \mathbb{R}$.

1. For $t > 0$ it is immediate from (5.6),
2. For $t = 0$: $u(x,0) = \Phi(x,0) * h(x) = \delta_0 * h(x) = h(x)$.

■

5.2 Non-homogeneous Problem

Consider now the problem

$$\begin{cases} u_t(x,t) - \Delta_x u(x,t) = g(x,t) & x \in \mathbb{R}, t > 0 \\ u(x,0) = 0 & x \in \mathbb{R} \end{cases}. \tag{5.8}$$

Fix s and let us consider $u(x,s,t)$ as the solutions of the following initial-value problem.

$$\begin{cases} u_t(x,s,t) - \Delta_x u(x,s,t) = 0 & x \in \mathbb{R}, t > s \\ u(x,s,t) = g(x,s) & x \in \mathbb{R}, t = s. \end{cases} \tag{5.9}$$

The reader may note that (5.9) is just an initial value problem of the form (5.8) with the start time shifted from $t = 0$ to $t = s$. Therefore, $u(x,t,s) = \Phi(x, t-s) * g(x,s)$. Duhamel's principle asserts that we can build solutions of (5.8) out of solutions of (5.9), by integrating with respect to s. That is,

$$\begin{aligned} u(x,t) &= \int_0^t u(x,t,s)ds = \int_0^t \Phi(x,t-s) * g(x,s)ds \\ &= \int_0^t \int_{\mathbb{R}} \Phi(x-y,t-s)g(y,s)dyds \end{aligned} \tag{5.10}$$

Finally, we can close the solution of (5.1). It is given by

$$\begin{aligned} u(x,t) &= \Phi(x,t) * h(x) + \int_0^t \Phi(x,t-s) * g(x,s)ds \\ &= \int_{\mathbb{R}} \Phi(x-y,t)h(y)dy + \int_0^t \int_{\mathbb{R}} \Phi(x-y,t-s)g(y,s)dyds. \end{aligned} \tag{5.11}$$

Chapter 6

Applications to the Black-Scholes Equation

We finally reach our last chapter which consists, as we mentioned before, in proving that under certain assumptions on the real-valued functions f and F, the problem

$$\begin{cases} u_t - u_{xx} = u_x + F(u) \\ u(x,0) = f(x) \end{cases} . \tag{6.1}$$

is locally well-posed in $\mathbb{H}^1$. As a reference, we strongly recommend the reading of [17], where the authors proved the local well-posedness of 6.1 for initial data in $\mathbb{L}^p, p \geq 1$. We will complete the estimates and theorems in order to show our result. Before that, let us present the financial context.

The Black-Scholes equation is the parabolic PDE given by

$$V_\tau + \frac{\sigma^2}{2} S^2 V_{SS} + rSV_S - rV = F(V)$$

where $V(S,\tau) : [0,\infty] \times [0,T] \to \mathbb{R}$ represents the value of a derivative, usually an european call option. The greek τ is used to represent the time horizon and T the maturity date. S is the current value of the stock, σ is the volatility and finally r represents the risk free interest rate, usually a bank account. The model, however, is quite limited due to its unrealistic assumptions. To overcome its flaws, it is common to find more complex models, as the Barles-Soner and the Platen-Schweizer (both explicit in [17]). These equations are highly non-linear and, therefore, very hard to study. Many authors have proposed semi-linear models.

With the usual changing of variables, $x = ln(S)$ and $\tau = T - t$, and some algebraic computations, many of these semi-linear models may be written as

$$u_t - u_{xx} = u_x + F(u)$$

with F a Lipschitz function. We may now formalize our result.

Theorem 6.0.1. *Let $F_1 : \mathbb{R} \to \mathbb{R}$ a Lipschitz function. Then, the Cauchy problem*

$$\begin{cases} u_t - u_{xx} = u_x + F_1(u) \\ u(x,0) = f(x) \end{cases} .$$

is locally well-posed in $\mathbb{H}^1$.

The argumentation of the proof is entirely analogous to the one done by the authors. Our goal is to use a fixed point theorem to prove the existence of u such that $u = \Psi(u)$ where Ψ is the functional defined by

$$\Psi(u) = \int_{\mathbb{R}} \Phi(x-y,t)f(y)dy + \int_0^t \int_{\mathbb{R}} \Phi(x-y,t-s)G(u(y,s),u_x(y,s))dyds.$$

Recall by Chapter 5 that such fixed point satisfies (6.1). The authors already proved, in particular, that

$$\Psi : C([0,T];\mathbb{L}^2(\mathbb{R})) \to C([0,T];\mathbb{L}^2(\mathbb{R})).$$

and, for some constant $0 < C < 1$

$$\|\Psi(u) - \Psi(v)\|_{\mathbb{L}_t^\infty \mathbb{L}_x^2} \leq C\|u-v\|_{\mathbb{L}_t^\infty \mathbb{L}_x^2} \underbrace{\leq}_{4.9} C\|u-v\|_{\mathbb{L}_t^\infty \mathbb{H}_x^1}.$$

Thus, it remains to prove the analogous estimate

$$\|\partial_x(\Psi(u) - \Psi(v))\|_{\mathbb{L}_t^\infty \mathbb{L}_x^2} \leq C\|u-v\|_{\mathbb{L}_t^\infty \mathbb{H}_x^1}.$$

Before moving to the formal proof, we need the following lemmas.

Lemma 6.0.2. *The heat kernel $\Phi(x,t)$ satisfies*

-
$$\|\Phi(\cdot,t)\|_{\mathbb{L}^p} = \frac{C_p}{(4\pi t)^{\frac{1}{2}(1-1/p)}}, \tag{6.2}$$

-
$$\|\Phi_x(\cdot,t)\|_{\mathbb{L}^p} = \frac{C_p}{(4\pi t)^{1-1/2p}}, \tag{6.3}$$

with $C_p = \frac{1}{p^{1/2p}}$.

Lemma 6.0.3. *The following estimate holds*

$$\left\| \int_0^t \int_{\mathbb{R}} \Phi(x-y,t-s) G_x(y,s)\, dyds \right\|_{\mathbb{L}_t^\infty \mathbb{L}_x^2} \leq CT^{1/2} \|G\|_{\mathbb{L}_t^\infty \mathbb{L}_x^2} \tag{6.4}$$

Proof. See in [17]. ■

We may now move to te proof of Theorem 6.0.1.

Proof.

$$\partial_x(\Psi(u)-\Psi(v)) = \underbrace{\int_0^t \int_{\mathbb{R}} \Phi_x(x-y,t-s)[u_x(y,s)-v_x(y,s)]dyds}_{A} + \underbrace{\int_0^t \int_{\mathbb{R}} \Phi_x(x-y,t-s)[F_1(u(y,s))-F_1(v(y,s))]dyds.}_{B}$$

■

The first integral may be re-written as

$$\begin{aligned} \int_0^t \int_{\mathbb{R}} \Phi_x(x-y,t-s)[u_x(y,s)-v_x(y,s)]dyds \\ = \int_0^t \Phi_x(\cdot,t-s) * [u_x(x,s)-v_x(y,s)]ds \\ = \int_0^t \Phi(\cdot,t-s) * \partial_x[u_x(x,s)-v_x(y,s)]ds \end{aligned}$$

Thus, using estimate (6.4)

$$\begin{aligned} \|A\|_{\mathbb{L}_t^\infty \mathbb{L}_x^2} &\leq CT^{1/2} \|u_x - v_x\|_{\mathbb{L}_t^\infty \mathbb{L}_x^2} \\ &\leq CT^{1/2} \|u-v\|_{\mathbb{L}_t^\infty \mathbb{H}_x^1} \end{aligned}$$

Same way,

$$\begin{aligned}\int_0^t \int_{\mathbb{R}} \Phi_x(x-y,t-s)[F_1(u(y,s)) - F_1(v(y,s))]dyds \\ = \int_0^t \Phi_x(\cdot,t-s) * [F_1(u(x,s)) - F_1(v(x,s))]ds \\ = \int_0^t \Phi(\cdot,t-s) * \partial_x[F_1(u(x,s)) - F_1(v(x,s))]ds\end{aligned}$$

By estimate 4.9, 6.4, and since F_1 is Lipschitz

$$\begin{aligned}\|B\|_{\mathbb{L}_t^\infty \mathbb{L}_x^2} &\leq CT^{1/2}\|F_1(u) - F_1(v)\|_{\mathbb{L}_t^\infty \mathbb{L}_x^2} \\ &\leq CT^{1/2}\|u-v\|_{\mathbb{L}_t^\infty \mathbb{H}_x^1}.\end{aligned}$$

Putting all together,

$$\|\partial_x(\Psi(u) - \Psi(v))\|_{\mathbb{L}_t^\infty \mathbb{L}_x^2} \leq 2CT^{1/2}\|u-v\|_{\mathbb{L}_t^\infty \mathbb{H}_x^1}.$$

Finally, choosing T small enough such that $C(T + T^{1/2}) < 1$ and $2CT^{1/2} < 1$ we conclude that Ψ is a contraction of the space $C([0,T];\mathbb{H}^1(\mathbb{R}))$. Hence, there exists $u \in C([0,T];\mathbb{H}^1(\mathbb{R}))$ such that $\Psi(u) = u$. The proof of the uniqueness of the solution follows the same lines.

Chapter 7

Conclusions & Future Research

In this thesis we prepared the reader to an introductory course on distribution theory.

After introducing the main topics and some basic constructions of test functions, we moved to distributions *per se* by presenting the definition and description of its most vital properties. We took a particular emphasis into three significant operations — product by a smooth function, differentiability and convolution.

With this procedure we were able to take the theory of Fourier and give a modern touch. We introduced the space of rapidly decreasing functions and the Schwartz space. In this space we could easily apply the theory of Fourier transform in distributions without caring about problems as the existence of the integral. It is important to note once again that different versions of Fourier transform formulae may be founded in the literature. To illustrate the importance and applicability of what was described, we presented a real physical example. We proved that the well known Heisenberg's uncertainty principle is no more than a Fourier transform theorem applied to quantum physics.

The next step was to introduce the Sobolev Spaces, in particular, the space $\mathbb{H}^1$ and the Sobolev injections that proved to be crucial to the remaining work.

We used some simple results proved along this thesis to present an alternative method of resolution to the heat diffusion equation with initial condition in the $\mathbb{L}^1$.

Finally, we were able to extend the result proved by D. Da Silva, K. Igibayeva, A. Khoroshevskay and Z.Sakayeva [17] by proving that under certain assumptions on the real-valued functions f and F, the problem $u_t - u_{xx} = u_x + F(u)$ with initial condition $u(x,0) = f(x)$ is locally well-posed in $\mathbb{H}^1$. To do so, we considered results carried out by the authors as well as the results of our core chapters.

The material for a complete course in distribution theory, Fourier transform and Sobolev spaces is incredible vast. Despite the accomplishment attained in this thesis it would be very instructive to continue in the same framework and extend Theorem 2 of [17].

One could extend, if possible, the estimations and proofs to more complex Sobolev spaces.

Bibliography

[1] Bracewell, R. (2000). *The Fourier Transform and Its Applications.* McGraw-Hill series in electrical and computer engineering. Circuits and systems. McGraw Hill, 3rd ed edition.

[2] Brezis, H. and Mironescu, P. (2018). *Gagliardo–Nirenberg inequalities and non-inequalities: The full story*, volume 35.

[3] Chou, C.-C. (1973). *La Transformation de Fourier Complexe et Equation de Convolution.* Lecture Notes in Mathematics. Springer, 1 edition.

[4] Craig, W. (2018). *A course on partial differential equations.* Graduate studies in mathematics volume 197. American Mathematical Society.

[5] Donoghue, W. F. (1969). *Distributions and Fourier transforms.* Pure and Applied Mathematics. Academic Press.

[6] Elias M. Stein, G. W. (1971). *Introduction to Fourier Analysis on Euclidean Spaces.* PMS-32 Princeton Mathematical Series. Princeton Univ Pr.

[7] Hormander, L. (1990). *The Analysis of Linear Partial Differential Operators. I, Distribution Theory and Fourier Analysis.* Grundlehren Der Mathematischen Wissenschaften. Springer, 2 edition.

[8] Hormann, G. (2009). *Lecture Notes on the Theory Of Distributions.*

[9] Isidore Isaac Hirschman, D. V. W. (2005). *The Convolution Transform.*

[10] J.J. Duistermaat, J. K. (2010). *Distributions: Theory and Applications.* Cornerstones. Birkhäuser Boston, 1st edition. edition.

[11] Mukherjea A., T. N. (1976). *Measures on Topological Semigroups: Convolution Products and Random Walks.*

[12] Norbert Ortner, P. W. a. (2015). *Fundamental Solutions of Linear Partial Differential Operators: Theory and Practice.* Springer International Publishing, 1 edition.

[13] Ramond, T. (2015). *Lecture Notes Distributions and Partial Differential Equations.*

[14] Robert A. Adams, J. J. F. F. (2003). *Sobolev spaces.* Pure and Applied Mathematics. Academic Press, 2 edition.

[15] Schwartz, L. (1969). *Application of distributions to the theory of elementary particles in quantum mechanics.* Documents on Modern Physics. Routledge, 1 edition.

[16] Schwartz, L. (1997). *Theorie des distributions.* Hermann.

[17] Silva, D., Igibayeva, K., Khoroshevskaya, A., and Sakayeva, Z. (2018). *Local Well-Posedness for Generalized Semilinear Black-Scholes Equations.*

[18] Strichartz, R. (1994). *A guide to distribution theory and Fourier transforms.* Studies in Advanced Mathematics. CRC-Press, 1 edition.

[19] Tartar, L. (2007). *An Introduction to Sobolev Spaces and Interpolation Spaces.* Lecture Notes of the Unione Matematica Italiana 3. Springer-Verlag Berlin Heidelberg, 1 edition.

[20] Treves, F. (1970). *Topological vector spaces, distributions and kernels.* Academic Press.

[21] Zuily, C. (1988). *Problems in Distributions and Partial Differential Equations*, volume 143. Elsevier.