NTNU – Trondheim
Norwegian University of
Science and Technology

Distributions, Schwartz Space and Fractional Sobolev Spaces

Fredrik Joachim Gjestland

Master of Science in Physics and Mathematics
Submission date: July 2013
Supervisor: Mats Ehrnstrøm, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4900
Industrial Mathematics, Master's Thesis: Distributions, Schwartz Space and Fractional Sobolev Spaces

Fredrik Joachim Gjestland

July 8, 2013

Abstract

This thesis derives the theory of distributions, starting with test functions as a basis. Distributions and their derivatives will be analysed and exemplified. Schwartz functions are introduced, and the Fourier transform of Schwartz functions is analysed, creating the basis for Tempered distributions on which we also analyse the Fourier transform. Weak derivatives and Sobolev spaces are defined, and from the Fourier transform we define Sobolev spaces of non-integer order. The theory presented is applied to an initial value problem with a derivative of order one in time and an arbitrary differentiation operator in space, and we take a look at conditions for well-posedness under different differnetiation operators and present some minor results. The Riesz representation theorem and the Lax–Milgram theorem are presented in order to offer a different perspective on the results from the initial value problem.

Sammendrag

Denne oppgaven utleder teori om distribusjoner ved å definere testfunksjoner og analysere hvordan kontinuerlige lineære funksjonaler påvirker disse. Schwartz-funksjoner og Fouriertransformasjonen blir så introdusert, og egenskapene til Fouriertransformasjonen anvendt på Schwartz-funksjoner blir gjennomgått. Temperære distribusjoner blir introdusert som elementer i dual-rommet til Schwartz-rommet, og vi ser hvordan Fouriertransformasjonen påvirker disse. Svake deriverte og Sobolevrom blir definert, og gjennom Fouriertransformasjonen finner vi en alternativ måte å definere svake deriverte, som lar oss introdusere Sobolevrom av fraksjonell orden. Teorien blir så anvendt på et initialverdiproblem med førsteordens tidsderivert og en vilkårlig derivasjonsoperator i rom. Det undersøkes under hvilke tilefeller problemer er velstilt og noen mindre resultater presenteres. Så presenteres Riesz representasjonsteorem og Lax–Milgram teoremet som i noen grad oppsumerer resultatene fra initialverdiproblemet i form av at de begge påviser unik løsning under gitte vilkår.

Preface

This master's thesis is written during the spring semester of 2013 to complete the degree Master of Science through the course TMA4900, "Matematikk, Masteroppgave" at the Norwegian University of Science and Technology (NTNU), where I have studied "Industriell Matematikk". The course has a value of 30 units and is the culmination of five years of studying mathematics, physics and engineering.

When embarking on this thesis, I knew surprisingly little about the theory I were to write about, and as such it has been a tremendous learning process in the subject itself, mathematical rigour, discipline and structure. In trying to convey the theory to the reader I had to spend at least twice as much time and energy accepting and understanding the theory myself, making the work process challenging, but rewarding. The work started in February 2013, and has, apart from the month of May 2013 when it was put on hold in order to prepare for two exams, been an ongoing process until July eight 2013.

I would like to express my gratefulness to my supervisor, førsteamanuensis/Associate Professor Mats Harald Andreas Ehrnström for his inputs, ideas and helpful comments. I would particularly like to thank him for having the time to meet me on a daily basis during the final week before submission of the thesis.

Contents

1 Distributions **11**
1.1 Test Functions 12
1.2 Distributions 12
1.3 Distributional Derivative 15

2 Schwartz Space **17**
2.1 The Fourier Transformation on $S(\mathbb{R}^n)$ 18
2.2 Convolutions of Continuous Functions 23
2.3 Tempered Distributions 27
2.3.1 Example of Fourier Transformation of Distributions 28
2.3.2 Convolutions with Tempered Distributions 30

3 Sobolev Spaces **33**

4 A Family of Initial Value Problems **37**
4.0.3 Example of a Homogeneous Solution 37
4.0.4 A Short Description of the General Homogeneous Solution . . 38
4.1 Polynomial Differential Operator 39
4.1.1 Polynomials of Order Four and Higher 44
4.1.2 Superpositioning 45
4.2 Duhamel's principle 45
4.2.1 Example of Duhamel's Principle 46
4.3 Non-Polynomial Differentiation Operators 47

5 The Riesz Representation Theorem **49**
5.1 Example of the Riesz Representation Theorem 51

6 The Lax–Milgram Theorem **53**

7 Bibliography **59**

Chapter 1

Distributions

Distributions, or "generalized functions", is a concept which, as implied, allows us to expand the notion of a function. Take for instance the impulse created by a baseball bat hitting a baseball, which, when we consider the force of impact to occur at a singular instance in time, can not be described by a function, as it is a multiple of Dirac's delta, δ, which, as we shall see, is a distribution.

One of the main motivations for distributions is the way an integrable function acts on bounded functions, when their product is integrated over Euclidean space — such integrable functions define regular distributions. However, when one wants to generalise this, allowing for derivatives of, say, non-integrable functions, one gets inspiration from the chain rule, which for smooth and compactly supported functions yields the equality

$$\int (D^\alpha \phi)\psi \, dx = (-1)^{|\alpha|} \int \phi(D^\alpha \psi) dx, \tag{1.1}$$

where support, as mentioned in the previous paragraph, is taken in the usual sense:

Definition Let Ω be a domain[1]. For $f \in C(\Omega)$,

$$\text{supp} f = \overline{\{x \in \Omega : f(x) \neq 0\}}$$

is called the support of f.

The "exchange of derivatives" in (1.1) can be used to define derivatives of functions that are not differentiable, but only if we require that the test functions (ψ in (1.1)) have a corresponding degree of regularity and decay at infinity. This is the reason for the definition of test functions in section 1.1, in which we will use a multi-index notation:

[1]Throughout this paper, a domain will mean an open set in $\mathbb{R}^n$.

Definition When writing $\alpha \in \mathbb{N}_0^n$, we use a multi-index notation where α is the n-tuple $(\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_n)$. The notation D^α is defined as

$$D^\alpha = \frac{\partial^{\alpha_1}}{(\partial x_1)^{\alpha_1}} \frac{\partial^{\alpha_2}}{(\partial x_2)^{\alpha_2}} \frac{\partial^{\alpha_3}}{(\partial x_3)^{\alpha_3}} \cdots \frac{\partial^{\alpha_n}}{(\partial x_n)^{\alpha_n}},$$

and the size of α is defined as

$$|\alpha| = \alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_n.$$

Remark As we shall see, it is possible to extend this notion from integrable functions to general objects (distributions) which act in essentially the same way on test functions as integrable functions do. The space of test functions will therefore always determine the corresponding space of distributions.

1.1 Test Functions

Definition Let Ω be a domain in $\mathbb{R}^n$. Then the functions contained in

$$D(\Omega) = \{\phi \in C^\infty(\Omega) : \text{supp } \phi \text{ compact in } \Omega\}$$

are called test functions. A sequence $\{\phi_j\}_{j=1}^\infty \subset D(\Omega)$ is said to be convergent in $D(\Omega)$ to $\phi \in D(\Omega)$ if there is a compact set $K \subset \Omega$ with

$$\text{supp } \phi_j \subset K, \quad j \in \mathbb{N}$$

and

$$\sup_K |D^\alpha \phi_j - D^\alpha \phi| \to 0 \quad \text{for all } \alpha \text{ in } \mathbb{N}_0^n.$$

The notation $\phi_j \underset{D}{\longrightarrow} \phi$ means "convergent in $D(\Omega)$".

1.2 Distributions

As stated in the introduction to the chapter, we can define distributions by how they interact with test functions:

Definition Let Ω be a domain in $\mathbb{R}^n$ and let $D(\Omega)$ be defined as above. $D'(\Omega)$ is the collection of all complex-valued linear continuous functionals T over $D(\Omega)$:

$$T : D(\Omega) \to \mathbb{C}, \quad T : \phi \mapsto T(\phi), \quad \phi \in D(\Omega)$$

$$T(\lambda_1\phi_1 + \lambda_2\phi_2) = \lambda_1 T(\phi_1) + \lambda_2 T(\phi_2), \quad \lambda_1, \lambda_2 \in \mathbb{C}, \quad \phi_1, \phi_2 \in D(\Omega)$$

and

$$T(\phi_j) \to T(\phi) \text{ whenever } \phi_j \underset{D}{\longrightarrow} \phi.$$

Any $T \in D'(\Omega)$ is called a distribution. By

$$T_j \to T \text{ in } D'(\Omega), \quad T_j, T \in D'(\Omega), \quad j \in \mathbb{N}$$

we mean that

$$T_j(\phi) \to T(\phi) \text{ in } \mathbb{C}, \quad \text{if } j \to \infty \text{ for any } \phi \in D(\Omega).$$

Remark Instead of writing $T(\phi)$, where $\phi \in D(\Omega)$, it is conventional to write $\langle T, \phi \rangle$. A simple example is Dirac's delta:

$$\delta(\phi) = \langle \delta, \phi \rangle = \phi(0).$$

Remark If f is locally integrable in a domain Ω, then the functional T_f defined by

$$T_f(\phi) = \int_\Omega f(x)\phi(x)dx$$

is said to be a regular distribution. A distribution that is not regular is said to be singular.

Example Dirac's delta, $\delta(\phi)$, is singular

Proof. We consider the point distribution at $x = 0$: $\delta(\phi) = \phi(0)$. Let us first check that δ really is a distribution. Is is clear that $\delta : D(\Omega) \to \mathbb{C}$, as for linearity, we can check

$$\begin{aligned} \delta(\lambda_1\phi_1 + \lambda_2\phi_2) &= \lambda_1\phi_1(0) + \lambda_2\phi_2(0) \\ &= \lambda_1\delta(\phi_1) + \lambda_2\delta(\phi_2). \end{aligned}$$

Lastly,

$$\begin{aligned} \delta(\phi_j) &= \phi_j(0) \\ &\xrightarrow{j\to\infty} \phi(0) \\ &= \delta(\phi) \end{aligned}$$

whenever

$$\sup_K |\phi_j - \phi| \to 0.$$

We need to show that there does not exist a locally integrable function, f, such that

$$\int_\mathbb{R} f\phi = \phi(0) \tag{1.2}$$

for all $\phi \in D(\Omega)$. Consider the test function $\rho(x) = \begin{cases} e^{-\frac{1}{1-|x|^2}} & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1. \end{cases}$

It is clear that $\rho(x)$ has compact support, and through some calculation it can be seen that it is infinitely differentiable. Thus, $\rho(x) \in D(\Omega)$. If we assume that the locally integrable function in (1.2) exists, we would have

$$\int_{\mathbb{R}} f(x)\rho(nx)dx = \rho(0), \qquad \forall \quad n \in \mathbb{N}.$$

However,

$$\begin{aligned}\frac{1}{e} &= |\rho(0)| \\ &= |\int_{\mathbb{R}} f(x)\rho(nx)dx| \\ &\leq \int_{\mathbb{R}} |f(x)||\rho(nx)|dx \\ &= \int_{-\frac{1}{n}}^{\frac{1}{n}} |f(x)||\rho(nx)|dx \\ &\leq \int_{-\frac{1}{n}}^{\frac{1}{n}} |f(x)|dx \\ &\xrightarrow{n\to\infty} 0,\end{aligned}$$

where the second inequality follows from $|\rho(x)| \leq 1$ for all x. So, by contradiction such an f can not exist. □

Definition We define $\eta \in D(\Omega)$ to be

$$\eta(x) = \begin{cases} c_\eta e^{-\frac{1}{1-|x|^2}} & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

where the positive constant c_η is chosen such that $\int_{\mathbb{R}^n} \eta(x)dx = 1$. For every $\epsilon > 0$, let

$$\eta_\epsilon(x) = \frac{1}{\epsilon^n}\eta(\frac{x}{\epsilon}).$$

We call η the standard mollifier. The functions η_ϵ are infinitely differentiable, and

$$\int_{\mathbb{R}^n} \eta_\epsilon dx = 1, \quad \text{supp}(\eta_\epsilon) = B(0,\epsilon),$$

where $B(0,\epsilon)$ denotes the n-dimentional sphere centred at the origin and with radius ϵ. Thus, $\eta_\epsilon \in D(\mathbb{R}^n)$.

Definition If u is a locally integrable function in $\mathbb{R}^n$, we define its mollification

$$u_\epsilon(x) = \int_{\mathbb{R}^n} \eta_\epsilon(y)u(x-y)dy = \int_{\mathbb{R}^n} \eta_\epsilon(x-y)u(y)dy.$$

Lemma 1.2.1. *For any open set $\Omega \in \mathbb{R}^n$, $D(\Omega)$ is dense in $L^p(\Omega)$ for $1 \leq p < \infty$.*

Proof. As any function in $L^p(\Omega)$ can be approximated by a step function,

$$g = \sum_{j=1}^{m} a_j \chi_{Q_j}, \quad a_j \in \mathbb{C},$$

where χ_{Q_j} are the characteristic functions of open cubes Q_j with $\overline{Q_j} \in \Omega$, it is enough to show that χ_Q, the characteristic function of an arbitrary cube in Ω can be approximated by functions in $D(\Omega)$.

For $h < \infty$, $(\chi_Q)_h(x) \in D(\Omega)$, and

$$\begin{aligned}
\|(\chi_Q)_h - \chi_Q\|_{L^p(\Omega)} &= \| \int_{\mathbb{R}^n} \eta_h(x-y)(\chi_Q(y) - \chi_Q(x))dy\|_{L^p(\Omega)} \\
&= \| \frac{c_\eta}{h^n} \int_{B(x,h)} \eta(\frac{x-y}{h})(\chi_Q(y) - \chi_Q(x))dy\|_{L^p(\Omega)} \\
&\leq \| \frac{c_\eta}{h^n} \int_{B(x,h)} |\chi_Q(y) - \chi_Q(x)|dy\|_{L^p(\Omega)} \\
&\leq \| \frac{1}{|B(x,h)|} \int_{B(x,h)} c|\chi_Q(y) - \chi_Q(x)|dy\|_{L^p(\Omega)}, \qquad (1.3)
\end{aligned}$$

where $|B(x,h)|$ denotes the size of the ball B. By Lebesgue's differentiation theorem, (1.3) tends to the value of its integrand for every x as h goes to zero, and we obtain

$$\|(\chi_Q)_h - \chi_Q\|_{L^p(\Omega)} \to 0,$$

where $(\chi_Q)_h \in D(\Omega)$, proving the lemma. □

1.3 Distributional Derivative

While functions are limited in the sense that they do not necessarily have derivatives which are functions (i.e. the derivative of the Heavyside function is not a function), all distributions have derivatives which are in turn distributions. Inspired by integration by parts we make the following definition for the derivative of a distribution:

Definition Let $\alpha \in \mathbb{N}_0^n$ and $T \in D'(\Omega)$. Then the derivative $D^\alpha T$ is given by

$$D^\alpha T(\phi) = (-1)^{|\alpha|} T(D^\alpha \phi)$$

or, by inner-product notation:

$$\langle D^\alpha T, \phi \rangle = (-1)^{|\alpha|} \langle T, D^\alpha \phi \rangle.$$

Example of a distributional derivative

Let $g(x) = |x|$. We will look at $\frac{dg}{dx}$.

$$\begin{aligned}
\langle g', \phi \rangle &= -\langle g, \phi' \rangle \\
&= -\int_{\mathbb{R}} g(x)\phi'(x)dx \\
&= -\int_{\mathbb{R}} |x|\phi'(x)dx \\
&= \int_{-\infty}^{0} x\phi'(x)dx - \int_{0}^{\infty} x\phi'(x)dx \\
&= x\phi(x)\Big|_{-\infty}^{0} - \int_{-\infty}^{0} \phi(x)dx - x\phi(x)\Big|_{0}^{\infty} + \int_{0}^{\infty} \phi(x)dx \\
&= -\int_{-\infty}^{0} \phi(x)dx + \int_{0}^{\infty} \phi(x)dx \\
&= \langle \mathrm{sgn}, \phi \rangle.
\end{aligned}$$

Where

$$L^1_{loc} \ni \mathrm{sgn}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 1 & \text{if } x > 0. \end{cases}$$

Thus, the derivative of the absolute value of x is the signum function, as one would expect.

Chapter 2

Schwartz Space

When working with the Fourier transformation we require $\Omega = \mathbb{R}^n$. However, $D(\mathbb{R}^n)$ is, in a sense, too small for for the Fourier transform, which also makes $D'(\mathbb{R}^n)$ too large. For the purpose of the Fourier transformation, the Schwartz spaces, $S(\mathbb{R}^n)$ (named in honour of Laurent Schwartz) and $S'(\mathbb{R}^n)$, as introduced below, are optimally adapted in the sense that they are both closed under the Fourier transform.

Definition For $n \in \mathbb{N}$,

$$S(\mathbb{R}^n) = \{\phi \in C^\infty(\mathbb{R}^n) : \sup_{x \in \mathbb{R}^n} (1 + |x|^2)^{\frac{k}{2}} \sum_{|\alpha| \leq l} |D^\alpha \phi(x)| < \infty, \quad \text{for all } k, l \in \mathbb{N}_0\} \tag{2.1}$$

is called the Schwartz space of all rapidly decreasing infinitely differentiable functions, or Schwartz space for short.

A sequence $\{\phi_j\}_{j=1}^\infty \subset S(\mathbb{R}^n)$ is said to converge in $S(\mathbb{R}^n)$ to $\phi \in S(\mathbb{R}^n)$ if

$$\|\phi_j - \phi\|_{k,l} \to 0 \text{ for } j \to \infty \text{ and all } k, l \in \mathbb{N}_0.$$

Where

$$\|\phi\|_{k,l} = \sup_{x \in \mathbb{R}^n} (1 + |x|^2)^{\frac{k}{2}} \sum_{|\alpha| \leq l} |D^\alpha \phi(x)|.$$

Proposition 2.0.1. *The Schwartz space, $S(\mathbb{R}^n)$ is a subspace of $L^p(\mathbb{R}^n)$ for every p in $\mathbb{N}$.*

Proof. For every function ϕ in $S(\mathbb{R}^n)$, there exists a constant K such that

$$|\phi(x)| \leq \frac{K}{1 + |x|^2}$$

for every x in $\mathbb{R}^n$. Thus,

$$\int_{\mathbb{R}^n} |\phi(x)|^p dx \leq K^p \int_{\mathbb{R}^n} \frac{1}{(1 + |x|^2)^p} dx < \infty$$

for every $p < \infty$. For $p = \infty$, it follows from the definition, (2.1), that

$$\sup_{x\in\mathbb{R}^n} |\phi(x)| < \infty,$$

and we conclude that $S(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for every p in $\mathbb{N}$. $\square$

Proposition 2.0.2. *If $\phi \in S(\mathbb{R}^n)$, so are both $x^\alpha\phi$ and $D^\alpha\phi$ for $\alpha \in \mathbb{N}_0^n$.*

Proof. This follows directly from the definition of $S(\mathbb{R}^n)$, (2.1). All functions $\phi \in S(\mathbb{R}^n)$ are rapidly decreasing (i.e. go to zero when multiplied with an arbitrary polynomial), and so do all of their derivatives. $\square$

2.1 The Fourier Transformation on $S(\mathbb{R}^n)$

The Fourier transform, named after Joseph Fourier is an important tool, and it has several applications in physics and engineering. As we shall see, it allows us to, amongst others, transform differentiation operators in our regular dimension (usually time in the applied sense) into polynomials in another dimension (usually frequency in the applied sense), a property of great use in solving differential equations.

Definition For $\phi \in S(\mathbb{R}^n)$, the Fourier transform, $\mathcal{F}$, and the inverse Fourier, $\mathcal{F}^{-1}$, are given by

$$(\mathcal{F}\phi)(\xi) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix\xi}\phi(x)dx,$$

$$(\mathcal{F}^{-1}\phi)(\xi) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{ix\xi}\phi(x)dx,$$

for $\xi \in \mathbb{R}^n$.

Remark With the purpose of shortening notation, we will sometimes write

$$\mathcal{F}(\phi(x))(\xi) = \hat{\phi}(\xi).$$

Proposition 2.1.1. *The Fourier transformation of $D^\alpha\phi(x)$ is given by*

$$\mathcal{F}(D_x^\alpha\phi(x))(\xi) = i^{|\alpha|}\xi^\alpha\hat{\phi}(\xi). \tag{2.2}$$

Proof. The proof for $n = 1$ is straightforward by calculation:

$$\begin{aligned}
\mathcal{F}(D_x^\alpha \phi(x))(\xi) &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} e^{-ix\xi} D_x^\alpha \phi(x) dx \\
&= D_x^{\alpha-1}\phi(x) e^{-ix\xi}\big|_{\partial\mathbb{R}} + (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} i\xi e^{-ix\xi} D_x^{\alpha-1}\phi(x) dx \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} i\xi e^{-ix\xi} D_x^{\alpha-1}\phi(x) dx \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} (i\xi)^2 e^{-ix\xi} D_x^{\alpha-2}\phi(x) dx \\
&\vdots \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} (i\xi)^\alpha e^{-ix\xi} \phi(x) dx \\
&= (i\xi)^\alpha (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} e^{-ix\xi} \phi(x) dx \\
&= (i\xi)^\alpha \hat{\phi}(\xi).
\end{aligned}$$

To extend to $n > 1$ we note that

$$\begin{aligned}
\mathcal{F}(D_x^\alpha \phi(x))(\xi) &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix\xi} D_x^\alpha(x) dx \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} e^{-ix\xi} D_x^\alpha(x) dx \\
&= (i\xi_1)^{\alpha_1} (i\xi_2)^{\alpha_2} \cdots (i\xi_n)^{\alpha_n} \hat{\phi}(\xi) \\
&= i^{|\alpha|} \xi^\alpha \hat{\phi}(\xi)
\end{aligned}$$

□

Proposition 2.1.2. *The Fourier transform of $x^\alpha \phi(x)$ is given by*

$$\mathcal{F}(x^\alpha \phi(x)) = i^{|\alpha|} D_\xi^\alpha \hat{\phi}(\xi). \tag{2.3}$$

Proof. Again, the proof for $n = 1$ straightforward by calculation, but we start with the right hand side of (2.3).

$$
\begin{aligned}
(iD_\xi)^\alpha \hat{\phi}(\xi) &= i^\alpha D_\xi^\alpha (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} e^{-ix\xi}\phi(x)dx \\
&= i^\alpha (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} D_\xi^\alpha e^{-ix\xi}\phi(x)dx \\
&= i^\alpha (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} -ixD_\xi^{\alpha-1} e^{-ix\xi}\phi(x)dx \\
&= i^\alpha (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} (-ix)^2 D_\xi^{\alpha-2} e^{-ix\xi}\phi(x)dx \\
&\vdots \\
&= i^\alpha (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} (-ix)^\alpha e^{-ix\xi}\phi(x)dx \\
&= i^\alpha (-i)^\alpha (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}} e^{-ix\xi} x^k \phi(x)dx \\
&= \mathcal{F}(x^k\phi(x)).
\end{aligned}
$$

We extend to $n > 1$ in the same manner as the previous preposition:

$$
\begin{aligned}
\mathcal{F}(x^k\phi(x)) &= (iD_{\xi_1})^{\alpha_1}(iD_{\xi_2})^{\alpha_2}\cdots(iD_{\xi_n})^{\alpha_n}\hat{\phi}(\xi) \\
&= i^{|\alpha|}D_\xi^\alpha \hat{\phi}(\xi).
\end{aligned}
$$

□

Theorem 2.1.3. *If $\phi \in S(\mathbb{R}^n)$, so is $\mathcal{F}\phi$ and $\mathcal{F}^{-1}\phi$.*

Proof. We need to show that if $\phi \in S(\mathbb{R}^n)$, then

$$\sup_{\xi\in\mathbb{R}^n} (1+|\xi|^2)^{\frac{k}{2}} \sum_{\alpha\leq l} |D^\alpha \hat{\phi}(\xi)| < \infty \quad \text{for every } k, l.$$

We look at each term in the series individually, so for every k, l in $\mathbb{N}_0^n$:

$$
\begin{aligned}
\sup_{\xi\in\mathbb{R}^n} \xi^k |D_\xi^l \hat{\phi}(\xi)| &= \sup_{\xi\in\mathbb{R}^n} |\frac{\xi^k}{i^l}\mathcal{F}(x^l\phi(x))(\xi)| \quad \text{by (2.3)} \\
&= \sup_{\xi\in\mathbb{R}^n} |\frac{1}{i^{l+k}}\mathcal{F}(D_x^k x^\alpha \phi(x))| \quad \text{by (2.1.1)} \\
&= \sup_{\xi\in\mathbb{R}^n} |\frac{1}{i^{l+k}}(2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix\xi} D_x^k (x^l\phi(x))dx| \\
&\leq \sup_{\xi\in\mathbb{R}^n} |\frac{1}{i^{l+k}}(2\pi)^{-\frac{n}{2}}| \int_{\mathbb{R}^n} |e^{-ix\xi} D_x^l (x^l\phi(x))dx| \\
&\leq \sup_{\xi\in\mathbb{R}^n} (2\pi)^{-\frac{n}{2}} \int |D_x^k (x^\alpha\phi(x))|dx.
\end{aligned}
$$

Now, by proposition 2.0.2, since $\phi(x) \in S(\mathbb{R}^n)$, so is $x^\alpha \phi(x)$, and thus $D_x^k(x^\alpha \phi(x))$, and since $S(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$ by proposition 2.0.1, we conclude that $\hat{\phi} \in S(\mathbb{R}^n)$. □

Proposition 2.1.4. *Let* $\phi(x) = e^{-\frac{\epsilon^2|x|^2}{2}}$ *where* ϵ *is a positive, real constant. The Fourier transform of this Gaussian is given by*

$$\hat{\phi}(\xi) = \epsilon^{-n} e^{-\frac{|\xi|^2}{2\epsilon^2}}.$$

Proof. Let us first consider the case $n = 1$, by propositions 2.1.1 and 2.3 we have

$$\begin{aligned} D_\xi \hat{\phi}(\xi) &= \mathcal{F}(\frac{x}{i} e^{-\frac{\epsilon^2 \xi^2}{2}})(\xi) \\ &= \frac{1}{i}\mathcal{F}(-\epsilon^{-2}\frac{d}{dx} e^{-\frac{\epsilon^2 x^2}{2}})(\xi) \\ &= -\frac{1}{i\epsilon^2} i\xi\hat{\phi}(\xi) \\ &= -\frac{1}{\epsilon^2}\xi\hat{\phi}(\xi). \end{aligned}$$

We can rewrite this to

$$\frac{d}{d\xi} \ln \phi(\xi) = -\frac{1}{\epsilon^2}\xi.$$

By integrating both sides with respect to ξ and taking exponents, we obtain

$$\hat{\phi}(\xi) = Ce^{-\frac{\xi^2}{2\epsilon^2}}.$$

The constant C is obviously equal to $\hat{\phi}(0)$, which we can find:

$$\begin{aligned} \hat{\phi}(0) &= (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} e^{-\frac{\epsilon^2|x|^2}{2}} e^{-ix0} dx \\ &= (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} e^{-\frac{\epsilon^2|x|^2}{2}} dx \\ &= (2\pi)^{-\frac{1}{2}} \sqrt{\frac{2\pi}{\epsilon^2}} \qquad (2.4) \\ &= \frac{1}{\epsilon} \end{aligned}$$

where (2.4) follows from taking a contour integral and calculating residues. Thus, we have for $n = 1$:

$$\hat{\phi}(\xi) = \frac{1}{\epsilon} e^{-\frac{\xi^2}{2\epsilon^2}}.$$

For $n > 1$ we use the fact that $|x|^2 = \sum_{j=1}^n (x_j)^2$ together with the properties of the exponential function:

$$\begin{aligned}
\hat{\phi}(\xi) &= \int_{\mathbb{R}^n} e^{-\frac{\epsilon^2|x|^2}{2}} e^{-ix\xi} dx \\
&= \int_{\mathbb{R}^n} e^{-\frac{\epsilon^2}{2}\sum_{j=1}^n (x_j)^2} e^{-i\sum_{k=1}^n \xi_j^k x_j^k} dx \\
&= \int_{\mathbb{R}^n} \prod_{j=1}^n \left(e^{-\frac{\epsilon^2}{2}(x_j)^2} e^{-i\xi_j x_j} \right) dx \\
&= \prod_{j=1}^n \left(\int_{\mathbb{R}} e^{-\frac{\epsilon^2}{2}(x_j)^2} e^{-i\xi_j x_j} dx \right) \\
&= \prod_{j=1}^n \epsilon^{-1} e^{-\frac{\xi_j^2}{2\epsilon^2}} \\
&= \epsilon^{-n} e^{-\frac{1}{2\epsilon^2}\sum_{j=1}^n \xi_j^2} \\
&= \epsilon^{-n} e^{\frac{|\xi|^2}{2\epsilon^2}}.
\end{aligned}$$

Proposition 2.1.5. *Both $\mathcal{F}\mathcal{F}^{-1}$ and $\mathcal{F}^{-1}\mathcal{F}$ are identity operators on $S(\mathbb{R}^n)$,*

$$\phi = \mathcal{F}^{-1}\mathcal{F}\phi = \mathcal{F}\mathcal{F}^{-1}\phi. \tag{2.5}$$

Proof. Assuming $\phi, \psi \in S(\mathbb{R}^n)$, Fubini's theorem gives

$$\begin{aligned}
\int_{\mathbb{R}^n} (\mathcal{F}\phi)(\xi) e^{ix\xi} \psi(\xi) d\xi &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \phi(y) \int_{\mathbb{R}^n} e^{-i(y-x)\xi} \psi(\xi) d\xi dy \\
&= \int_{\mathbb{R}^n} \phi(y)(\mathcal{F}\psi)(y-x) dy \\
&= \int_{\mathbb{R}^n} \phi(x+y)(\mathcal{F}\psi)(y) dy. \qquad (2.6)
\end{aligned}$$

By letting

$$\psi(x) = e^{-\frac{\epsilon^2|x|^2}{2}}, \quad \epsilon > 0, \quad x \in \mathbb{R}^n$$

we obtain

$$\begin{aligned}
(\mathcal{F}\psi)(y) &= \epsilon^{-n}(\mathcal{F}e^{-\frac{|x|^2}{2}})(\frac{y}{\epsilon}) \\
&= \epsilon^{-n} e^{-\frac{|y|^2}{2\epsilon^2}} \qquad (2.7)
\end{aligned}$$

by proposition 2.1.4. We then insert (2.7) into (2.6) and substitute $y = \epsilon z$:

$$\int_{\mathbb{R}^n} (\mathcal{F}\psi)(\xi) e^{ix\xi} e^{-\frac{\epsilon^2|\xi|^2}{2}} d\xi = \int_{\mathbb{R}^n} \phi(x+\epsilon z) e^{-\frac{|z|^2}{2}} dz.$$

Finally we let $\epsilon \to 0$ and obtain

$$\begin{aligned}(\mathcal{F}^{-1}\mathcal{F}\phi)(x) &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} (\mathcal{F}\psi)(\xi)e^{ix\xi}d\xi \\ &= (2\pi)^{-\frac{n}{2}}\phi(x)\int_{\mathbb{R}^n} e^{-\frac{|z|^2}{2}}dz \\ &= (2\pi)^{-\frac{n}{2}}(2\pi)^{\frac{n}{2}}\phi(x) \\ &= \phi(x).\end{aligned}$$

Thus, $\mathcal{F}^{-1}\mathcal{F}$ is an identity operator on $S(\mathbb{R}^n)$, the same property can be shown for $\mathcal{F}\mathcal{F}^{-1}$ in a similar fashion. □

Theorem 2.1.6. *Both $\mathcal{F}$ and $\mathcal{F}^{-1}$ map $S(\mathbb{R}^n)$ one-to-one onto itself,*

$$\mathcal{F}S(\mathbb{R}^n) = S(\mathbb{R}^n), \quad \mathcal{F}^{-1}S(\mathbb{R}^n) = S(\mathbb{R}^n). \tag{2.8}$$

Proof. By applying (2.5) to $\phi = \mathcal{F}^{-1}\psi$,

$$\begin{aligned}\phi &= \mathcal{F}^{-1}\mathcal{F}\phi \\ &= \mathcal{F}\psi,\end{aligned}$$

thus, $\mathcal{F}S(\mathbb{R}^n) = S(\mathbb{R}^n)$, and similarly one obtains $\mathcal{F}^{-1}S(\mathbb{R}^n) = S(\mathbb{R}^n)$. In addition, if $\mathcal{F}\phi_1 = \mathcal{F}\phi_2$, (2.5) yields $\phi_1 = \phi_2$, hence $\mathcal{F}$ and $\mathcal{F}^{-1}$ are one-to-one mappings of $S(\mathbb{R}^n)$ onto itself. □

2.2 Convolutions of Continuous Functions

Definition For two functions ϕ and ψ, both in $C(\mathbb{R}^n)$ and at least one of them having compact support, the convolution $(\phi, \psi) \mapsto \phi * \psi$ is defined through the continuous function

Lemma 2.2.1. *If either $\phi * \psi$ or $\psi * \phi$ exist, $\phi * \psi = \psi * \phi$.*

$$(\phi * \psi)(x) = \int_{\mathbb{R}^n} \psi(x-y)\phi(y)dy.$$

Proof. The proof follows from a simple substitution:

$$\begin{aligned}(\phi * \psi)(x) &= \int_{\mathbb{R}^n} \phi(x-y)\psi(y)dy, \qquad u = x-y, \quad du = -dy \\ &= -\int_{\mathbb{R}^n} \phi(u)\psi(x-u)du \\ &= \int_{\mathbb{R}^n} \psi(x-u)\phi(u)du \\ &= (\psi * \phi)(x).\end{aligned}$$

□

Theorem 2.2.2. *Whenever $f, g \in L^1(\mathbb{R}^n)$, and their convolution is defined, $\widehat{f * g} = (2\pi)^{\frac{n}{2}} \hat{f}\hat{g}$.*

Proof. Because $f*g \in L^1(\mathbb{R})$ and $|e^{-ix\xi}| = 1$, Fubini's theorem implies that the Fourier transformation is well-defined, and

$$\begin{aligned}
\mathcal{F}(f * g)(\xi) &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(x-y)g(y)dy \right) e^{-ix\xi}dx \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(x-y)e^{-ix\xi}dx \right) g(y)dy \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(z)e^{-i(z+y)\xi}dz \right) g(y)dy \\
&= (2\pi)^{\frac{n}{2}} \left((2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(z)e^{-iz\xi}dz \right) \left((2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} g(y)e^{-iy\xi}dy \right) \\
&= (2\pi)^{\frac{n}{2}} \mathcal{F}f(\xi)\mathcal{F}g(\xi) \\
&= (2\pi)^{\frac{n}{2}} \hat{f}\hat{g}.
\end{aligned}$$

□

Lemma 2.2.3. *With $\phi \in C^j$ and $\psi \in C^k$, we can, in the derivative of a convolution as defined above, and with $|\alpha| \leq j$, $|\beta| \leq k$ interchange the derivative and the convolution in the following way:*

$$D^{\alpha+\beta}(\phi * \psi) = (D^\alpha \phi * D^\beta \psi).$$

Proof. From theorem 2.2.2, we have

$$\mathcal{F}(\phi * \psi) = (2\pi)^{\frac{n}{2}} \mathcal{F}(\phi)\mathcal{F}(\psi),$$

and further,

$$\begin{aligned}
\mathcal{F}(D^{\alpha+\beta}(\phi * \psi)) &= i^{|\alpha+\beta|} x^{\alpha+\beta} \mathcal{F}(\phi * \psi)(x) \\
&= i^{|\alpha+\beta|} x^{\alpha+\beta} (2\pi)^{\frac{n}{2}} \mathcal{F}(\phi)\mathcal{F}(\psi) \\
&= (2\pi)^{\frac{n}{2}} i^{|\alpha|} x^\alpha \mathcal{F}(\phi) i^{|\beta|} x^\beta \mathcal{F}(\psi) \\
&= (2\pi)^{\frac{n}{2}} \mathcal{F}(D^\alpha \phi)\mathcal{F}(D^\beta \psi) \\
&= \mathcal{F}(D^\alpha \phi * D^\beta \psi).
\end{aligned}$$

The proof is completed by the uniqueness of the Fourier-transform. □

Theorem 2.2.4. *Let $j, k \geq 0$. If $\phi \in C_0^j$ and $\psi \in L^1_{loc}$, $\phi * \psi \in C^{j+k}$ if $\psi \in C^k$.*

Proof. Firstly, let $f, g \in C^0(\mathbb{R}^n), g \in L^1_{loc}(\mathbb{R}^n =$ and either f or g have compact support. We define $h(x) = f * g(x)$. We obtain

$$|h(x) - h(x+\delta)| = |\int_{\mathbb{R}^n} (f(x-y) - f(x-y+\delta))g(y)dy| \tag{2.9}$$

$$\leq \int_{\mathbb{R}^n} (|f(x-y) - f(x-y+\delta))||g(y)|dy \tag{2.10}$$

$$\leq \int_{\mathbb{R}^n} \epsilon|g(y)|dy \tag{2.11}$$

$$\leq k\epsilon, \tag{2.12}$$

for some k, thus the convolution of two continuous functions is a continuous function.

Next, let $f \in C^1(\mathbb{R}^n), g \in L^1_{loc}(\mathbb{R}^n) \cap C^0(\mathbb{R}^n)$ and either f or g have compact support. Now,

$$\frac{h(x) - h(x_0)}{x - x_0} = \int_{\mathbb{R}^n} \frac{f(x-y) - f(x_0 - y)}{x - x_0} g(x)dx.$$

We need to show that $\frac{f(x-y)-f(x_0-y)}{x-x_0}$ converges uniformly to $f'(x_0 - y)$. By the mean value theorem, we have

$$\begin{aligned} f(x-y) - f(x_0-y) &= \int_0^1 \frac{df(x_0 - y + t(x-x_0))}{dt} dt \\ &= \left(\int_0^1 f'(x_0 - y + t(x-x_0))dt \right)(x - x_0). \end{aligned}$$

Thus,

$$\frac{f(x-y) - f(x_0) - y)}{x - x_0} - f'(x_0 - y) = \int_0^1 \left(f'(x_0 - y + t(x-x_0)) - f'(x_0 - y)\right) dt. \tag{2.13}$$

Because f' is continuous by assumption and has compact support, it is uniformly continuous. Thus, for any $\epsilon > 0$ there exists a $\delta > 0$ such that if $|x - x_0| < \delta$, then

$$|f'(x_0 - y + t(x - x_0)) - f'(x_0 - y)| < \epsilon. \tag{2.14}$$

Thus, (2.13) tends to zero uniformly for all x_0 and y and $h \in C^1(\mathbb{R}^n)$. The assertion then follows from induction on j or k fixed. □

Theorem 2.2.5. *If $0 \leqq \phi \in C_0^\infty$, $\int_{\mathbb{R}^n} \phi(x)dx = 1$, and $u \in C_0^j(\mathbb{R}^n)$, then $u_\phi = u * \phi \in C_0^\infty(\mathbb{R}^n)$ by theorem 2.2.4. Further,*

$$\sup |D^\alpha u - D^\alpha u_\phi| \to 0 \tag{2.15}$$

whenever supp $\phi \to \{0\}$ and $|\alpha| \leq j$.

Proof. By using theorem 2.2.4 and lemma 2.2.3 and, it is sufficient to prove (2.15) for $\alpha = 0$. Let $|y| < \delta$ in supp ϕ, and we obtain

$$\begin{aligned}|u(x) - u_\phi(x)| &= |\int (u(x) - u(x-y))\phi(y)dy| \\ &\leq \sup_{|y|<\delta} |u(x) - u(x-y)| \\ &\xrightarrow{\delta \to 0} 0.\end{aligned}$$

□

Theorem 2.2.6. *If $\phi(x, y) \in C^\infty(X, Y)$ where Y is an open set in $\mathbb{R}^n$, and if there is a compact set $K \subset X$ such that $\phi(x, y) = 0$ when $x \notin K$, then*

$$y \mapsto u(\phi(\cdot, y))$$

is a C^∞ function of y if $u \in D'(X)$, and

$$D_y^\alpha u(\phi(\cdot, y)) = u(D_y^\alpha(\cdot, y)).$$

Proof. We fix $y \in Y$ and use Taylor's formula to obtain

$$\phi(x, y + h) = \phi(x) + \sum h_j \frac{\partial}{\partial y_j}\phi(x, y) + \psi(x, y, h),$$

where

$$\sup_x |D_x^\alpha \psi(x, y, h)| = \mathcal{O}(|h|^2), \qquad \text{as } h \to 0.$$

Hence,

$$u(\phi(\cdot, y + h)) = u(\phi(\cdot, y)) + \sum h_j u(\frac{\partial}{\partial y_j}\phi(\cdot, y)) + \mathcal{O}(|h|^2).$$

Thus, $y \mapsto u(\phi(\cdot, y))$ is differentiable and

$$\frac{\partial}{\partial y_j} u(\phi(\cdot, y)) = u(\frac{\partial}{\partial y_j}\phi(\cdot, y))$$

and the theorem is proved by iterating this argument. □

2.3 Tempered Distributions

In the same fashion as defining Distributions to be functionals acting on test functions, we define a smaller set of functionals acting on Schwartz functions:

Definition Let $S(\mathbb{R}^n)$ be as previously defined. $S'(\mathbb{R}^n)$ is the collection of all complex-valued linear continuous functionals T over $S(\mathbb{R}^n)$:

$$T : S(\mathbb{R}^n) \to \mathbb{C}, \quad T : \phi \mapsto T(\phi), \quad \phi \in S(\mathbb{R}^n).$$

$$T(\lambda_1\phi_1 + \lambda_2\phi_2) = \lambda_1 T(\phi_1) + \lambda_2 T(\phi_2), \quad \lambda_1, \lambda_2 \in \mathbb{C}, \quad \phi_1, \phi_2 \in S(\mathbb{R}^n).$$

We furnish $S'(\mathbb{R}^n)$ with the the simple convergence topology

$$T_j \to T \text{ in } S'(\mathbb{R}^n), \quad T_j \in S'(\mathbb{R}^n), \quad j \in \mathbb{N}, \quad T \in S'(\mathbb{R}^n),$$

means that

$$T_j(\phi) \to T(\phi) \text{ in } \mathbb{C} \text{ if } j \to \infty \text{ for any } \phi \in S(\mathbb{R}^n).$$

Any $T \in S'(\mathbb{R}^n)$ is called a tempered distribution.

Remark All test functions are elements in the Schwartz space: $D(\mathbb{R}^n) \subset S(\mathbb{R}^n)$. It follows from this that $S'(\mathbb{R}^n) \subset D'(\mathbb{R}^n)$. Thus, as implied by the name, every tempered distribution is a distribution.

Definition Let $T \in S'(\mathbb{R}^n)$. Then the Fourier transform $\mathcal{F}T$ and its inverse $\mathcal{F}^{-1}T$ are given by

$$(\mathcal{F}T)(\phi) = T(\mathcal{F}\phi), \quad \text{and } (\mathcal{F}^{-1}T)(\phi) = T(\mathcal{F}^{-1}\phi), \quad \phi \in S(\mathbb{R}^n). \tag{2.16}$$

Theorem 2.3.1. *The Fourier transform is a continuous linear one-to-one and onto mapping on $S'(\mathbb{R})^n$, and for all T in $S'(\mathbb{R}^n)$,*

$$T = \mathcal{F}^{-1}\mathcal{F}T = \mathcal{F}\mathcal{F}^{-1}T \tag{2.17}$$

Proof. The mapping $T \mapsto \mathcal{F}T$ from $S'(\mathbb{R}^n)$ is clearly linear. And if $T_n \to 0$ in $S'(\mathbb{R}^n)$, then

$$\langle \mathcal{F}T_n, \phi\rangle = \langle T_n, \mathcal{F}\phi\rangle \to 0 \quad \text{as } n \to \infty.$$

The same proof works for $\mathcal{F}^{-1}$. For all $\phi \in S(\mathbb{R}^n)$.

$$\begin{aligned}\langle \mathcal{F}\mathcal{F}^{-1}T, \phi\rangle &= \langle T, \mathcal{F}^{-1}\mathcal{F}\phi\rangle \\ &= \langle T, \mathcal{F}\mathcal{F}^{-1}\phi\rangle \\ &= \langle \mathcal{F}^{-1}\mathcal{F}T, \phi\rangle \\ &= \langle T, \phi\rangle,\end{aligned}$$

which proves (2.17). Thus, $\mathcal{F}$ and $\mathcal{F}^{-1}$ being one-to-one onto $S'(\mathbb{R}^n)$ follows from $\mathcal{F}$ and $\mathcal{F}^{-1}$ being one-to-one on $S(\mathbb{R}^n)$. □

2.3.1 Example of Fourier Transformation of Distributions

Example 1
Let us first use our standard example: Dirac's delta function. If we insert the "function" into the classical definition of the Fourier transformation we obtain

$$\begin{aligned}\mathcal{F}(\delta(x))(\xi) &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix\xi}\delta(x)dx \\ &= (2\pi)^{-\frac{n}{2}} e^{-i0\xi} \\ &= (2\pi)^{-\frac{n}{2}}.\end{aligned}$$

If we on the other hand use (2.16) we obtain

$$\begin{aligned}(\mathcal{F}\delta)(\phi) &= \delta(\mathcal{F}(\phi)) \\ &= \mathcal{F}(\phi)(0) \\ &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \phi(x)e^{-ix0}dx \\ &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \phi(x)dx \\ &= \langle (2\pi)^{-\frac{n}{2}}, \phi\rangle\end{aligned}$$

Thus, we obtain

$$\langle \mathcal{F}\delta, \phi\rangle = \langle (2\pi)^{-\frac{n}{2}}, \phi\rangle,$$

obtaining the same result as in our more classical approach.

Example 2
Consider the constant function $f(x) = 1$. In the sense of functions, the Fourier transform

$$\mathcal{F}(1)(\xi) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix\xi}1dx$$

does not make sense, as $e^{-ix\xi} \notin L^1(\mathbb{R}^n)$. In the sense of distributions, however, we have

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

giving us

$$\begin{aligned}
\langle \mathcal{F}1, \phi \rangle &= \langle 1, \mathcal{F}\phi \rangle \\
&= \int_{\mathbb{R}^n} \hat{\phi}(\xi) d\xi \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-ix\xi} \phi(x) dx d\xi \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{i0\xi} \int_{\mathbb{R}^n} e^{-ix\xi} \phi(x) dx d\xi \\
&= (2\pi)^{\frac{n}{2}} \mathcal{F}^{-1}(\mathcal{F}(\phi))(0) \\
&= (2\pi)^{\frac{n}{2}} \phi(0) \\
&= \langle (2\pi)^{\frac{n}{2}} \delta, \phi \rangle.
\end{aligned}$$

Thus,

$$\mathcal{F}(1)(\xi) = (2\pi)^{\frac{n}{2}} \delta(\xi) \tag{2.18}$$

<u>Example 3</u>
In this example we will consider the distribution $e^{i\alpha x}$.

$$\begin{aligned}
\mathcal{F}(e^{i\alpha x})(\xi) &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix\xi} e^{i\alpha x} dx \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{ix(\alpha-\xi)} dx \\
&= \mathcal{F}(1)(\xi - \alpha) \\
&= (2\pi)^{\frac{n}{2}} \delta(\xi - \alpha),
\end{aligned}$$

where we used (2.18).

<u>Example 4</u>
In this example we will use the result from example 3 to find the Fourier transform of the trigonometric functions $\sin(\alpha x)$ and $\cos(\alpha x)$:

$$\begin{aligned}
\mathcal{F}(\sin(\alpha x)) &= \mathcal{F}\left(\frac{e^{i\alpha x} - e^{-i\alpha x}}{2i} \right) \\
&= \frac{1}{2i} \left(\mathcal{F}(e^{i\alpha x}) - \mathcal{F}(e^{-i\alpha x}) \right) \\
&= \frac{(2\pi)^{\frac{n}{2}}}{2i} \left(\delta(\xi - \alpha) - \delta(\xi + \alpha) \right).
\end{aligned}$$

And in similar fashion one finds

$$\mathcal{F}(\cos(\alpha x)) = \frac{(2\pi)^{\frac{n}{2}}}{2} \left(\delta(\xi - \alpha) + \delta(\xi + \alpha) \right).$$

Example 5

Let $T \in S'(\mathbb{R}^n)$, then $D^\alpha T \in S'(\mathbb{R}^n)$. Let us find $\mathcal{F}(D^\alpha T)$ and $D^\alpha \mathcal{F}T$, analogous to (2.1.1) and (2.3), and properties we will need in chapter 4.

$$\begin{aligned}
\langle \widehat{D^\alpha T}, \phi \rangle &= \langle D^\alpha T, \hat{\phi} \rangle \\
&= (-1)^{|\alpha|} \langle T, D^\alpha \hat{\phi} \rangle \\
&= (-1)^{|\alpha|} \langle T, \mathcal{F}\left((-i)^{|\alpha|} \xi^\alpha \phi \right) \rangle \\
&= (-1)^{|\alpha|} \langle \hat{T}, (-i)^{|\alpha|} \xi^\alpha \phi \rangle \\
&= (-1)^{|\alpha|} \langle (-i)^{|\alpha|} \xi^\alpha \hat{T}, \phi \rangle \\
&= \langle i^{|\alpha|} \xi^\alpha \hat{T}, \phi \rangle.
\end{aligned}$$

$$\begin{aligned}
\langle D^\alpha \hat{T}, \phi \rangle &= (-1)^{|\alpha|} \langle \hat{T}, D^\alpha \phi \rangle \\
&= (-1)^{|\alpha|} \langle T, \widehat{D^\alpha \phi} \rangle \\
&= (-1)^{|\alpha|} \langle T, i^{|\alpha|} \xi^\alpha \hat{\phi} \rangle \\
&= (-1)^{|\alpha|} \langle i^{|\alpha|} \xi^\alpha T, \hat{\phi} \rangle \\
&= (-1)^{|\alpha|} \langle \widehat{i^{|\alpha|} \xi^\alpha T}, \phi \rangle \\
&= \langle (-i)^{|\alpha|} \widehat{\xi^\alpha T}, \phi \rangle
\end{aligned}$$

Thus,

$$\widehat{D^\alpha T} = i^{|\alpha|} \xi^\alpha \hat{T}, \quad D^\alpha \hat{T} = (-i)^{|\alpha|} \widehat{\xi^\alpha T}. \tag{2.19}$$

2.3.2 Convolutions with Tempered Distributions

In order to define the convolution for tempered distributions we start with a fixed $\psi \in S(\mathbb{R}^n)$, we do this because products are not defined for all distributions, but with one factor in $S(\mathbb{R}^n)$ this will not be a problem. A convolution with ψ is then an operation which preserves $S'(\mathbb{R}^n)$, so to define $\psi * f$ for $f \in S'(\mathbb{R}^n)$ we find

$$\int \psi * \phi_1(x) \psi_2(x) dx = \int \int \psi(x-y) \phi_1(y) \phi_2(x) dy dx.$$

Note that

$$\int \psi(x-y) \phi_2(x) dx = \tilde{\psi} * \phi_2(y), \qquad \tilde{\psi}(x) = \psi(-x).$$

Thus,

$$\int \psi * \phi_1(x) \phi_2(x) dx = \int \phi_1(y) \tilde{\psi} * \phi_2(y) dy.$$

Definition We define $\psi * f$, where $\psi \in S(\mathbb{R}^n)$ and $f \in S'(\mathbb{R}^n)$ by

$$\langle \psi * f, \phi \rangle = \langle f, \tilde{\psi} * \phi \rangle.$$

Lemma 2.3.2. *The Fourier transform of the convolution between a tempered distribution and a Schwartz function is given by the product of their Fourier tranforms:*

$$\mathcal{F}(\psi * f) = (2\pi)^{\frac{n}{2}} \hat{\psi} \hat{f}.$$

Proof.

$$\begin{aligned} \langle \mathcal{F}(\psi * f), \phi \rangle &= \langle \psi * f, \hat{\phi} \rangle \\ &= \langle f, \tilde{\psi} * \hat{\phi} \rangle \\ &= \langle \hat{f}, \mathcal{F}^{-1}(\tilde{\psi} * \hat{\phi}) \rangle \\ &= \langle \hat{f}, (2\pi)^{\frac{n}{2}} (\mathcal{F}^{-1}\tilde{\psi})\phi \rangle \\ &= \langle (2\pi)^{\frac{n}{2}} \hat{f}, \hat{\psi}\phi \rangle \\ &= \langle (2\pi)^{\frac{n}{2}} \hat{\psi}\hat{f}, \phi \rangle. \end{aligned}$$

□

Example of convolution with a tempered distribution

If $f = \delta$ and $\psi \in S(\mathbb{R}^n)$,

$$\begin{aligned} \langle \psi * f, \phi \rangle &= \langle \psi * \delta, \phi \rangle \\ &= \langle \delta, \tilde{\psi} * \phi \rangle \\ &= (\tilde{\psi} * \phi)(0) \\ &= \int_{\mathbb{R}^n} \psi(y)\phi(y) dy \\ &= \langle \psi, \phi \rangle \end{aligned}$$

Thus,

$$\psi * \delta = \psi.$$

We also note that

$$\mathcal{F}(\psi * \delta) = (2\pi)^{\frac{n}{2}} \hat{\psi}\hat{\delta} = \hat{\psi},$$

as $\hat{\delta} = (2\pi)^{-\frac{n}{2}}$.

Theorem 2.3.3. *If $u \in D'(\mathbb{R}^n)$ and $\phi \in D(\mathbb{R}^n)$, then $u * \phi \in C^\infty(\mathbb{R}^n)$ and*

$$\textit{supp } (u * \phi) \subset \textit{supp } u + \textit{supp } \phi.$$

For any multi-index α we have

$$D^\alpha(u * \phi) = (D^\alpha u) * \phi = u * (D^\alpha \phi).$$

Proof. By theorem 2.2.6, $u * \phi \in C^\infty$, and

$$D^\alpha(u * \phi) = u * D^\alpha \phi.$$

proving the second equality.

Note that $u * \phi(x) = 0$ unless $x - y \in \operatorname{supp} \phi$ for some $y \in \operatorname{supp} u$. Thus, $x \in \operatorname{supp} y + \operatorname{supp} \phi$. And this is a closed set because $\operatorname{supp} \phi$ is compact. □

Theorem 2.3.4. *Let* $0 \leqq \phi \in D(\mathbb{R}^n)$ *and* $\int_{\mathbb{R}^n} \phi dx = 1$. *If* $u \in D'(\mathbb{R}^n)$, $u_\phi = u * \phi \in C^\infty(\mathbb{R}^n)$ *and* $u_\phi \to u$ *in* $D'(\mathbb{R}^n)$ *as supp* $\phi \to \{0\}$.

Proof. We have $u(\psi) = u * \tilde{\psi}(0)$ if $\psi \in D(\mathbb{R}^n)$, where again, $\tilde{\psi}(x) = \psi(-x)$, giving

$$\begin{aligned} u_\phi(\psi) &= u_\phi * \tilde{\psi}(0) \\ &= u * \phi * \tilde{\psi}(0) \\ &= u(\tilde{\phi} * \psi). \end{aligned}$$

Theorem 2.2.5 gives us that $\tilde{\phi} * \psi \to \psi$ in C_0^∞ as $\operatorname{supp} \phi \to \{0\}$. Thus, $u_\phi(\psi) \to u(\psi)$. □

Chapter 3

Sobolev Spaces

With the foundation of distributions and the Fourier transform we are almost equipped with the tools necessary to define Sobolev spaces. A Sobolev space is a vector space which is a combination of L^p norms of the function itself and its derivatives up to a given order. However, for the obtained space to be a Banach space, we need to look at the norm of its so-called weak derivatives:

Definition A function f in $L^1_{loc}(\mathbb{R})$ is called weakly differentiable if there exists a function $\partial_x f$ in $L^1_{loc}(\mathbb{R})$ such that

$$\int_{\mathbb{R}^n} (\partial_x f)\phi dx = -\int_{\mathbb{R}^n} f\partial_x \phi, \text{ for all } \phi \in D(\mathbb{R}).$$

Furthermore, if for every $k = 0, 1, \cdots, n$, there exists $\partial_x^k f \in L^1_{loc}(\mathbb{R})$ with

$$\int_{\mathbb{R}^n} (\partial_x^k f)\phi dx = (-1)^k \int_{\mathbb{R}^n} f\partial_x^k \phi, \text{ for all } \phi \in D(\mathbb{R})$$

we say that f is n times weakly differentiable with corresponding weak derivatives $\partial_x^k f$.

Definition For $k \in \mathbb{N}_0^n$ and $1 \leq p < \infty$,

$$W_p^k(\mathbb{R}^n) = \{f \in L^p(\mathbb{R}^n) : D^\alpha f \in L^p(\mathbb{R}^n) \forall \alpha \in \mathbb{N}_0^n, |\alpha| \leq k\},$$

where the derivatives are taken in the weak sense. W_k^p is then called a Sobolev spaces. When furnished with the norm

$$\|f\|_{W_p^k(\mathbb{R}^n)} = \left(\sum_{|\alpha|k} \|D^\alpha f\|^p_{L^p(\mathbb{R}^n)} \right)^{\frac{1}{p}}$$

$W_p^k(\mathbb{R}^n)$ becomes a Banach space. With the inner product

$$\langle f, g \rangle_{W_2^k(\mathbb{R}^n)} = \sum_{|\alpha| \leq k} \int_{\mathbb{R}^n} D^\alpha f(x) \overline{D^\alpha g(x)} dx$$

$W_2^k(\mathbb{R}^n)$ become Hilbert spaces.

Remark The spaces mentioned thus far are connected in the following way:

$$D(\mathbb{R}^n) \subset S(\mathbb{R}^n) \subset W_p^k(\mathbb{R}^n) \subset L^p(\mathbb{R}^n) \subset S'(\mathbb{R}^n) \subset D'(\mathbb{R}^n). \tag{3.1}$$

Theorem 3.0.5. *$D(\mathbb{R}^n)$ is dense in $D'(\mathbb{R}^n)$.*

Proof. We begin by choosing a sequence $\chi_j \in D(\Omega)$ such that on any compact subset of Ω we have $\chi_j = 1$ for all sufficiently large j. Then we choose $\phi_j \in D(\mathbb{R}^n)$ satisfying theorem 2.3.4 and with small enough support to satisfy

$$\text{supp } \phi_j + \text{supp } \chi_j \subset \Omega, \qquad |x| < \frac{1}{j} \text{ if } x \in \text{supp } \phi_j.$$

Since $\chi_j u$ is a compactly supported distribution we can form

$$u_j = (\chi_j u) * \phi_j,$$

thus obtaining a function in $D(\Omega)$ by theorem 2.3.3 and theorem 2.3.4, and we have as in the proof of theorem 2.3.4

$$\begin{aligned} u_j(\psi) &= (\chi_j u)(\tilde{\phi}_j * \psi) \\ &= u(\chi_j(\tilde{\phi}_j * \psi)). \end{aligned}$$

Now, because supp $\tilde{\phi}_j * \psi$ belongs to any neighborhood of supp ψ whenever j is large enough, and we obtain $\chi_j(\tilde{\phi}_j\psi) = \tilde{\phi}_j\psi$ for those same j. It follows that $u_j(\psi) \to u(\psi)$ as required. □

Remark Because $D(\mathbb{R}^n)$ is dense in $D'(\mathbb{R}^n)$, and all inclusions in (3.1) are continuously embedded, every inclusion in (3.1) is dense.

Lemma 3.0.6. *If*

$$(1 + |\xi|^2)^{\frac{n}{2}} g(\xi) \in L_2(\mathbb{R}^n), \qquad n \in \mathbb{N}_0,$$

there exists an n times weakly differentiable function f in $L_2(\mathbb{R}^n)$ with $\hat{f} = g$ and weak derivatives in $L_2(\mathbb{R}^n)$ such that

$$\partial_x^k f \xrightarrow{\mathcal{F}} (i\xi)^k g(\xi) \in L_2(\mathbb{R}^n), \qquad k = 0, 1, \cdots, n.$$

Proof. It is clear that g is in $L_2(\mathbb{R}^n)$, as $\xi^n g$ is in $L_2(\mathbb{R}^n)$, so there exists $f \in L_2(\mathbb{R}^n)$ and a sequence $\{\phi_j\}_{j \in \mathbb{Z}_{>0}} \subset S(\mathbb{R}^n)$ such that

$$\lim_{j\to\infty} \|f - \phi_j\|^2_{L^2(\mathbb{R}^n} = \frac{1}{2\pi} \lim_{j\to\infty} \|g - \hat{\phi}_j\|^2_{L_2(\mathbb{R}^n)} = 0.$$

Because $(1 + |\xi|^2)^{\frac{n}{2}} g(\xi)$ is in $L_2(\mathbb{R}^n)$, so is $|\xi|^k g(\xi)$ for $k = 1, 2, \cdots, n$. From Lebesgue's dominated theorem we obtain

$$\lim_{j\to\infty} \int |\xi|^{2k} |g(\xi) - \hat{\phi}_j(\xi)| d\xi = 0$$

since $\hat{\phi}_j \to g$ and the convergent series $\{\hat{\phi}_j\}_j$ is bounded. Thus $(i\xi)^k\hat{\phi}_j \to (i\xi)^k g$ in $L_2(\mathbb{R}^n)$ and

$$\begin{aligned}\int_{\mathbb{R}^n} \phi_j \partial_x^k \psi dx &= (-1)^k \int_{\mathbb{R}^n} \left(\partial_x^k \phi_j\right) \psi dx \\ &= (-1)^k \int_{\mathbb{R}^n} \mathcal{F}^{-1}\left((i\xi)^k g\right) \psi dx,\end{aligned}$$

which by definition make $\partial_x^k \phi_j$, and hence also $\mathcal{F}^{-1}((i\xi)^k g)$, weak derivatives. □

Now that we have introduced a notation for derivatives that do not require integer-order we can define Sobolev spaces of non-integer order.

Definition

$$H^s(\mathbb{R}^n) = \{f \in S'(\mathbb{R}^n) : (1+|x|^2)^{\frac{s}{2}}\mathcal{F}f \in L_2(\mathbb{R}^n)\}, \quad s \in \mathbb{R}$$

With the scalar product

$$\langle f, g\rangle_{H^s(\mathbb{R}^n)} = \int_{\mathbb{R}^n} (1+|x|^2)^{\frac{s}{2}}\mathcal{F}f(x)\overline{(1+|x|^2)^{\frac{s}{2}}\mathcal{F}g(x)}dx$$

$H^s(\mathbb{R}^n)$ are Hilbert spaces. The associated norm is given by

$$\begin{aligned}\|f\|_{H^s(\mathbb{R}^n} &= \sqrt{\langle f, f\rangle_{H^s(\mathbb{R}^n)}} \\ &= \left(\int_{\mathbb{R}^n} (1+|x|^2)^s \hat{f}(x)\overline{\hat{f}}(x)dx\right)^{\frac{1}{2}} \\ &= \left(\int_{\mathbb{R}^n} (1+|x|^2)^s |\hat{f}(x)|^2 dx\right)^{\frac{1}{2}} \\ &= \|(1+|x|^2)^{\frac{s}{2}}\hat{f}(x)\|_{L^2(\mathbb{R}},\end{aligned}$$

which is equivalent, in the sense of norms, to

$$\|f\|_{H^s(\mathbb{R}^n)} = \|(1+|x|^s)\hat{f}\|_{L^2(\mathbb{R}^n},$$

which is the H^s-norm we will use in the next chapter.

Proposition 3.0.7. *For natural numbers k, we have:*

$$H^k(\mathbb{R}^n) = W_2^k(\mathbb{R}^n), \quad k \in \mathbb{N}_0$$

Proof. If f is in $H^s(\mathbb{R}^n)$, then by Lemma 3.0.6, it is also on in $W_2^k(\mathbb{R}^n)$. Giving us

$$H^s(\mathbb{R}^n) \subseteq W_2^k(\mathbb{R}^n).$$

On the other hand, if f is in $W_2^k(\mathbb{R}^n)$, f and all its derivatives up to order k are in $L^2(\mathbb{R}^n)$. Thus,

$$\begin{aligned}
\mathcal{F}f &\in L^2(\mathbb{R}^n) \\
\mathcal{F}(D_x f) = (2\pi i \xi)\hat{f} &\in L^2(\mathbb{R}^n) \\
\mathcal{F}(D_x^2 f) = (2\pi i \xi)^2\hat{f} &\in L^2(\mathbb{R}^n),
\end{aligned}$$

and so on. We conclude that

$$W_2^k(\mathbb{R}^n) \subseteq H^s(\mathbb{R}^n).$$

And the sets are equal. □

Chapter 4

A Family of Initial Value Problems

In this chapter we will attempt to apply the theory from the preceding chapters in solving the following initial value problem:

$$u_t(t,x) + f(D)u(t,x) = g(t,x), \quad u(0,x) = u_0(x) \tag{4.1}$$

where $t \in [0,\infty), x \in \mathbb{R}$, and $f(D)$ is an arbitrary differentiation operator in spatial direction defined by

$$\mathcal{F}(f(D)u(t,x)) = f(i\xi)\hat{u}(t,\xi).$$

Remark Note that for polynomials this definition of f is consistent with proposition 2.1.1.

We will refer to this equation as *the initial value problem*, and will often leave out the initial value conditions in our notation to avoid writing them repeatedly.

4.0.3 Example of a Homogeneous Solution

For the sake of an example, and to obtain an understanding of the initial value problem, let us work in $\mathbb{R}^1$, let $g(t,x) = 0$, and $f(D) = 1 - D_x^2$, and let u_0 be a triangle with nodes in $(-1,0),(0,1)$ and $(1,0)$. This gives us the following initial value problem:

$$u_t(t,x) + (1 - D_x^2)u(t,x) = 0,$$

$$u(0,x) = u_0(x) = \text{Tri}(x) = \begin{cases} 1+x & \text{if } -1 < x < 0 \\ 1-x & \text{if } 0 \leq x < 1. \end{cases}$$

By taking the Fourier transformation we obtain

$$\hat{u}_t(t,\xi) + (1+\xi^2)\hat{u}(t,\xi) = 0.$$

We now have a first order partial differential equation with solution (with respect to t)

$$\hat{u}(t,\xi) = c(\xi)e^{-(1+\xi^2)t},$$

where c is some function of ξ. By the initial value condition we find

$$\begin{aligned}\hat{u}(0,\xi) &= c(\xi)\\ &= \hat{u}_0(\xi).\end{aligned}$$

Giving us

$$\hat{u}(t,\xi) = \hat{u}_0(\xi)e^{-(1+\xi^2)t}.$$

This is a solution on the Fourier side, in order to obtain a transformation, we note that it is a product of two functions. By letting $\mathcal{F}(K)(t,\xi) = e^{-(1+\xi^2)t}$ we can use theorem 2.2.2 to obtain a solution:

$$\begin{aligned}u(t,x) &= \mathcal{F}^{-1}\left(\mathcal{F}(u_0(0,x))\mathcal{F}(K(t,x))\right)\\ &= \mathcal{F}^{-1}\left(\mathcal{F}(u_0 * K))(t,\xi)\right)\\ &= (u_0 * K)(t,x).\end{aligned} \tag{4.2}$$

Where the convolution is done in the spatial direction and

$$\begin{aligned}K(x,t) &= \mathcal{F}^{-1}(e^{-(1+\xi^2)t})\\ &= \frac{1}{\sqrt{2t}^n}e^{-\frac{x^2}{4t}-t},\end{aligned}$$

by proposition 2.1.4. As K is a Schwartz-function and u_0 is a compactly support continuous function, the classical convolution yields

$$\begin{aligned}u(t,x) &= (u_0 * K)(t,x)\\ &= \int_{\mathbb{R}} u_0(y)K(x-y)dy\\ &= \int_{-1}^{0}(1+y)\frac{1}{\sqrt{2t}}e^{-\frac{(y-x)^2}{4t}-t}dy + \int_{0}^{1}(1-y)\frac{1}{\sqrt{2t}}e^{-\frac{(y-x)^2}{4t}-t}dy.\end{aligned}$$

We are left with a non-elementary function, u. In the proceeding sections of this chapter we will pay more attention to the properties of the solution u, and less attention to the explicit solution.

4.0.4 A Short Description of the General Homogeneous Solution

By general we refer to tempered distributions, the largest class for which we have defined a Fourier transform. We will find a solution to the case $g(t,x) = 0$, where u_t and $f(D)u(t,x)$ are tempered distributions:

$$u_t(t,x) + f(D_x)u(t,x) = 0. \tag{4.3}$$

As in the example in the previous section we take the Fourier transformation to obtain the partial differential equation

$$\hat{u}_t(t,\xi) + f(i\xi)\hat{u}(t,\xi) = 0,$$

which has the solution

$$\hat{u}(t,\xi) = c(\xi)e^{-f(i\xi)t}.$$

And again, we see that

$$\hat{u}(0,\xi) = c(\xi) = \hat{u}_0(\xi),$$

such that

$$\hat{u}(t,\xi) = \hat{u}_0(\xi)e^{-f(i\xi)t}.$$

Now that we have an expression for $\hat{u}$, we can write the solution,u, as either the inverse Fourier transform of $\hat{u}$:

$$u(t,x) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{ix\xi}\hat{u}_0(\xi)e^{-f(i\xi)t}d\xi,$$

or we can write it as a convolution, like we did in (4.2):

$$u(t,x) = u_0 * K(t,x),$$

where K is the inverse-Fourier transform of $e^{-f(i\xi)t}$. Of course we have to make assumptions on the Fourier transform or the convolution being defined, but even then there are few properties we can observe, but, when analysing the general case, it is hard to obtain strong results. As we shall see, when we make stronger assumptions on f, we can find stronger results.

4.1 Polynomial Differential Operator

The more classical case of the problem occurs when f is a polynomial, as it is in a lot of famous equations such as the heat equation or the Schrödinger equation. For clarity in notation we will denote $p(D) = f(D)$ when f is a polynomial, and although we will work in $\mathbb{R}$ to avoid confusion between vectors and scalar derivatives, and to allow ourselves to not use multi-index notation, the theory can be extended to $\mathbb{R}^n$. We are interested in how the initial value conditions affect the solution, and for what polynomials, p, the problem is well-posed:

Definition We say that the problem is well-posed when

- A solution exists.
- The solution is unique.
- The solution is continuously dependent on the initial value conditions.

As have seen, the existence of a solution follows from taking the Fourier transform, solving for t and transforming back, the uniqueness of the solution then follows from the Fourier transform being one-to-one for the spaces we work in, and from the solution on the Fourier side being unique. The continuous dependency on initial value conditions can be stated as

$$\|u - v\|_X \to 0 \quad \text{if} \quad \|u_0 - v_0\|_X \to 0.$$

We will often come across an even stronger dependency on initial conditions, namely Lipschitz continuity:

$$\|u - v\|_X \leq C\|u_0 - v_0\|_X \tag{4.4}$$

for elements in the space X. In a less general case, we can say that the solution is continuously dependent on initial value conditions if

$$u_0 \in X \implies u \in X,$$

Proposition 4.1.1. *Let*

$$f(\lambda) = \mp\lambda.$$

1. *If u_0 is in $S(\mathbb{R}^n)$, $L^2(\mathbb{R}^n)$ or $S'(\mathbb{R}^n)$, the solution u will be in the same class as u_0.*
2. *For $u_0 \in L^2(\mathbb{R})$ and $u_0 \in H^s(\mathbb{R})$, the problem is well-posed.*

Proof. The initial value problem now takes form

$$u_t \mp u_x = 0, \tag{4.5}$$

and we obtain solution on the Fourier-side

$$\hat{u}(t, x) = e^{\pm i\xi t} u_0(\xi).$$

1.
We make use of the fact that $|e^{\pm i\xi t}| \leq 1$. If u_0 is a Schwartz function, we have

$$\begin{aligned} u_0 \in S(\mathbb{R}) &\implies \hat{u_0} \in S(\mathbb{R}) \\ &\implies e^{\pm i\xi t} u_0 \in S(\mathbb{R}) \\ &\implies \hat{u} \in S(\mathbb{R}) \\ &\implies u \in S(\mathbb{R}). \end{aligned}$$

If $u_0 \in L^2(\mathbb{R})$, we have

$$\begin{aligned} u_0 \in L^2(\mathbb{R}) &\implies \hat{u_0} \in L^2(\mathbb{R}) \\ &\implies e^{\pm i\xi t} u_0 \in L^2(\mathbb{R}) \\ &\implies \hat{u} \in L^2(\mathbb{R}) \\ &\implies u \in L^2(\mathbb{R}). \end{aligned}$$

And if u_0 is a tempered distribution we have

$$\begin{aligned} u_0 \in S'(\mathbb{R}) &\implies \hat{u_0} \in S'(\mathbb{R}) \\ &\implies e^{\pm i\xi t} u_0 \in S'(\mathbb{R}) \\ &\implies \hat{u} \in S'(\mathbb{R}) \\ &\implies u \in S'(\mathbb{R}). \end{aligned}$$

2.
The equation

$$\hat{u} = e^{\pm i\xi t}\hat{u}_0(\xi)$$

has solution

$$\begin{aligned} u &= \int_{\mathbb{R}} e^{ix\xi} e^{\pm i\xi t}\hat{u}_0(\xi)d\xi \\ &= \int_{\mathbb{R}} e^{i\xi(x\pm t)}\hat{u}_u(\xi)d\xi \\ &= u_0(x \pm t). \end{aligned}$$

Thus,

$$\|u - v\|_{L^2(\mathbb{R})} = \|u_0 - v_0\|_{L^2(\mathbb{R})},$$

and the problem is continuously dependent on initial value conditions.

In $H^s(\mathbb{R})$,

$$\begin{aligned} \|u - v\|_{H^s(\mathbb{R})} &= \|(1 + |\xi|^s)(\hat{u} - \hat{v})\|_{L^2(\mathbb{R})} \\ &= \|(1 + |\xi|^s)(e^{\pm i\xi t}\hat{u}_0 - e^{\pm i\xi t}\hat{v}_0)\|_{L^2(\mathbb{R})} \\ &\leq \|(1 + |\xi|^s)(\hat{u}_0 - \hat{v}_0)\|_{L^2(\mathbb{R})}. \\ &= \|u_0 - v_0\|_{H^s(\mathbb{R})}. \end{aligned}$$

Existence and uniqness of the solution follows from (4.1.1) being the unique solution of (4.5), and the Fourier transformation being one-to-one on $S'(\mathbb{R})$. □

Proposition 4.1.2. *If*

$$p(\lambda) = -\lambda^2,$$

u is a C^∞ function for $u_0 \in S'(\mathbb{R})$, and $u \in H^{s'}(\mathbb{R})$ whenever $u_0 \in H^s(\mathbb{R})$ and $s' > s$.

Proof. The initial value problem takes the form

$$u_t - D_x^2 u = 0,$$

with solution on the Fourier-side:

$$\hat{u}(t, \xi) = e^{-\xi^2 t}\hat{u}_0(\xi). \tag{4.6}$$

As $e^{-\xi^2 t}$ is a Schwartz function, u can be expressed as a convolution between a Schwartz function and a tempered distribution. As all Schwartz-functions are C^∞, so is the convolution, u. As for $u \in H^{s'}(\mathbb{R})$ whenever $u_0 \in H^s(\mathbb{R})$, $s' > s$,

$$\begin{aligned}
\|u - v\|_{H^{s'}(\mathbb{R})} &= \|(1 + |\xi|^{s'})(\hat{u} - \hat{v})\|_{L^2(\mathbb{R})} \\
&= \|(1 + |\xi|^{s'})e^{-\xi^2 t}(\hat{u}_0 - \hat{v}_0)\|_{L^2(\mathbb{R})} \\
&= \left(\int_{\mathbb{R}} |(1 + |\xi|^s)(\hat{u}_0 - \hat{v}_0) \frac{1 + |\xi|^{s'}}{1 + |\xi|^s} e^{-2\xi^2 t}|^2 d\xi \right)^{\frac{1}{2}} \\
&\leq \left(M \int_{\mathbb{R}} |(1 + |\xi|^s)(\hat{u}_0 - \hat{v}_0)|^2 d\xi \right)^{\frac{1}{2}},
\end{aligned}$$

where $M = \max_{\xi \in \mathbb{R}} \frac{1+|\xi|^{s'}}{1+|\xi|^s} e^{-\xi^2 t} < \infty$ for all $t > 0$. Thus,

$$\begin{aligned}
\|u - v\|_{H^{s'}(\mathbb{R})} &\leq M\|(1 + |\xi|^s)(\hat{u}_0 - \hat{v}_0)\|_{L^2(\mathbb{R})} \\
&= M\|u_0 - v_0\|_{H^s(\mathbb{R})}.
\end{aligned}$$

□

Remark Note that if we instead had written $p(\lambda) = \lambda^2$ we would require $\hat{u}_0$ to have extremely rapid decay, meaning we would have to put strict requirements on u_0, and the inverse Fourier transformation would not have made sense unless $\hat{u}_0 e^{x^2}$ was integrable. By a Paley-Wiener theorem, we would require u_0 to be an entire function, see ([8], p. 120-122). In his 1935 paper, Tikhonov proves uniqueness of the heat equation if the solutions are not too large, and through a counter-example of a solution growing extremely fast for x showed non-uniqueness in general [9].

Proposition 4.1.3. *If*

$$p(\lambda) = \mp\lambda^3 \qquad \textit{and} \qquad u_0 \in H^3(\mathbb{R}),$$

the problem is well-posed and

$$u \in C([0, \infty), H^3(\mathbb{R})) \cap C^1([0, \infty), L^2(\mathbb{R})). \tag{4.7}$$

Proof. The initial value takes the form

$$u_t \mp D_x^3 u = 0.$$

By taking the Fourier transformation we obtain

$$\hat{u}_t \mp i\xi^3 \hat{u} = 0,$$

with solution

$$\hat{u}(t, \xi) = e^{\pm i\xi^3 t}\hat{u}_0(\xi).$$

By taking the inverse Fourier transformation, we obtain

$$u(t,x) = \int_{\mathbb{R}} e^{i\xi(x\pm\xi^2 t)} \hat{u}_0(\xi) d\xi. \tag{4.8}$$

Now, as

$$|e^{i\xi(x\pm\xi^2 t)}| \leq 1,$$

we have

$$u_0 \in L^2(\mathbb{R}) \iff e^{i\xi^3 t}\hat{u}_0 \in L^2(\mathbb{R}). \tag{4.9}$$

And because the Fourier transformation is a unitary operation on $L^2(\mathbb{R}^n)$, we also have $u \in L^2(\mathbb{R}^n)$. Next, we have

$$\begin{aligned}
\widehat{\partial_t u} &= \partial_t \hat{u} \\
&= \partial_t e^{\pm i\xi^3 t}\hat{u}_0(\xi) \\
&= \pm(i\xi)^3 e^{\pm i\xi^3 t}\hat{u}_0(\xi),
\end{aligned}$$

so

$$\partial_t u \in L^2(\mathbb{R}) \iff \pm\xi^3 e^{\pm i\xi^3 t}\hat{u}_0(\xi) \in L^2(\mathbb{R}).$$

And,

$$|\pm \xi^3 e^{\pm i\xi^3 t}\hat{u}_0(\xi)| \leq |\xi^3 \hat{u}_0|.$$

As for the function being C^1 and its time-derivative being L^2 whenever u_0 is H^3,

$$\begin{aligned}
\|\partial_t u(t_1) - \partial_t u(t_2)\|_{L^2(\mathbb{R})} &= \|\mathcal{F}(\partial_t u(t_1) - \partial_t u(t_2))\|_{L^2(\mathbb{R})} \\
&= \|\partial_t \hat{u}(t_1) - \partial_t u(t_1)\|_{L^2(\mathbb{R})} \\
&= \|\partial_t e^{\pm i\xi^3 t_1}\hat{u}_0 - \partial_t e^{\pm i\xi^3 t_2}\hat{u}_0\|_{L^2(\mathbb{R})} \\
&= \|\pm (i\xi^3)e^{\pm i\xi^3 t_1}\hat{u}_0 \mp (i\xi)^3 e^{\pm i\xi^3 t_2}\hat{u}_0\|_{L^2(\mathbb{R})} \\
&= \|\xi^3\hat{u}_0(e^{\pm i\xi^3 t_1} - e^{\pm i\xi^3 t_2})\|_{L^2(\mathbb{R})} \qquad (4.10) \\
&\leq \|2\xi^3\hat{u}_0\|_{L^2(\mathbb{R})} \\
&= \|2\partial_x^3 u_0\|_{L^2(\mathbb{R})},
\end{aligned}$$

which is bounded as $u_0 \in H^3(\mathbb{R})$. By Lebesgue's dominated convergence theorem (4.10) tends to zero as $t_2 \to t_1$, and thus $\partial_t u$ is a continuous function from $[0,\infty)$ to $L^2(\mathbb{R})$:

$$u_0 \in H^3(\mathbb{R}) \implies u \in C_t^1([0,\infty), L^2(\mathbb{R})).$$

For the other part of the intersection, we look at

$$\begin{aligned}
\widehat{\partial_x^3 u} &= \mp i\xi^3 \hat{u} \\
&= \mp i\xi^3 e^{\pm i\xi^3 t}\hat{u}_0(\xi).
\end{aligned}$$

Again,

$$|i\xi^3 e^{\pm i\xi^3 t}\hat{u}_0(\xi)| \leq |\xi^3 \hat{u}_0(\xi)|.$$

And,

$$\begin{aligned} \|u(t_1) - u(t_2)\|_{H^3(\mathbb{R})} &= \|(1 + |\xi|^3)(\hat{u}(t_1) - \hat{u}(t_2)\|_{L^2(\mathbb{R})} \\ &= \|(1 + |\xi|^3)(e^{\pm i\xi^3 t_1} - e^{\pm i\xi^3 t_2})\hat{u}_0\|_{L^2(\mathbb{R})} \\ &\leq \|2(1 + |\xi|^2)\hat{u}_0\|_{L^2(\mathbb{R})} \\ &= \|2u_0\|_{H^s(\mathbb{R})} \end{aligned} \tag{4.11}$$

which is bounded. Again, by Lebesgues dominated convergence theorem, as $t_2 \to t_1$, (4.11) tends to zero. Thus,

$$u \in C([0, \infty), H^3(\mathbb{R})) \tag{4.12}$$

As for well-posedness, the problem is continuously dependent on initial value conditions as

$$\begin{aligned} \|u - v\|_{H^3(\mathbb{R})} &= \|(1 + |\xi|^3)\widehat{(u - v)}\|_{L^2(\mathbb{R})} \\ &= \|(1 + |\xi|^3)(\hat{u} - \hat{v})\|_{L^2(\mathbb{R})} \\ &= \|(1 + |\xi|^3)(e^{\pm i\xi^3 t}\hat{u}_0 - e^{\pm i\xi^3 t}\hat{v}_0)\|_{L^2(\mathbb{R})} \\ &= \|(1 + |\xi|^3)e^{\pm i\xi^3 t}(\hat{u}_0 - \hat{v}_0)\|_{L^2(\mathbb{R})} \\ &\leq \|(1 + |\xi|^3)(\hat{u}_0 - \hat{v})\|_{L^2(\mathbb{R})} \\ &= \|u_0 - v_0\|_{H^3(\mathbb{R})}. \end{aligned}$$

The existence and uniqueness of a solution follows from the Fourier-transform being one-to-one on $S'(\mathbb{R})$ and the solution on the Fourier side being unique. □

4.1.1 Polynomials of Order Four and Higher

Note that because $\widehat{p(D)u} = i\xi\hat{u}$, increasing the exponent by one results in a 90-degree rotation on the complex plane on the Fourier-side. And has we have seen, $p(\lambda) = \pm\lambda$ share properties with $p(\lambda) = \mp\lambda^3$ in the sense that $|K| \leq 1$. In the same sense, one would expect similarities between $p(\lambda) = \lambda^2$ and $p(\lambda) = \lambda^4$. For $p(\lambda) = \lambda^4$, one obtains the equation

$$u_t + D^4 u = 0, \tag{4.13}$$

and by the Fourier transform

$$\hat{u}_t + \xi^4 \hat{u} = 0.$$

Solving in the usual manner, we obtain

$$\hat{u} = \hat{u}_0(\xi)e^{-\xi^4 t}.$$

As with the the quadratic polynomial, $\mathcal{F}^{-1}(e^{-\xi^4 t}) \in S(\mathbb{R})$, and the solution shares properties with the solution from the quadratic case.

If we were to continue increasing the exponent by one, we would obtain similar results, in some sense, for all even exponents, and for all odd exponents. However, results like (4.7) will vary depending on the exponent.

4.1.2 Superpositioning

As we have seen, the properties of the solution can, to some degree of precision, be determined from whether the exponents in p are even or odd, but what happens when p consists of more than one term? The initial value problems can be divided into three different classes for polynomial differentiation operators. Let k be in $\mathbb{N}_{>0}$:

1. If $p(\lambda) = +\lambda^{4k-2}$ or $p(\lambda) = -\lambda^{4k}$, $\hat{u}$ can be expressed as $e^{\xi^{4k+2}t}\hat{u}_0(\xi)$ or $e^{\xi^{4k}t}\hat{u}_0$ respectively.

2. If $p(\lambda) = \pm\lambda^{2k+1}$, $\hat{u}$ can be expressed as $e^{i(\mp\xi)^{2k+1}}\hat{u}_0(\xi)$.

3. If $p(\lambda) = -\lambda^{4k+2}$ or $p(\lambda) = +\lambda^{4k}$, $\hat{u}$ can be expressed as $e^{-\xi^{4k-2}t}\hat{u}_0(\xi)$ or $e^{-\xi^{4k}t}\hat{u}_0$ respectively.

Note that case 1 requires strict requirements in the sense of very rapid decay for u_0. Case 2 does not affect the norm of, in the sense that $|\hat{u}| = |\hat{u}_0|$. Case 3 will cause $\hat{u}$ to decay rapidly enough to yield $u \in S(\mathbb{R})$ for all tempered distributions, u_0. One may think of case 1 to be the "worst" and case 3 to be the "nicest" in terms of putting limitations on u_0. If p consists of more than one term, terms from case 1 will be dominant over cases 2 and 3, and case 3 will dominate case 2.

4.2 Duhamel's principle

We have so far only considered the homogeneous case of the equation, but by using the same methods we obtain a solution for the inhomogeneous equation.

Theorem 4.2.1. *Whenever the homogeneous initial value problem,*

$$u_t - Lu = 0, \quad u(0,x) = u_0(x), \tag{4.14}$$

has a solution operator $T(t)$, such that

$$(T(t)u_0)(x) = u(t,x),$$

the solution of the corresponding inhomogeneous initial value problem,

$$u_t(t,x) - Lu(t,x) = g(t,x), \quad u(0,x) = u_0(x), \tag{4.15}$$

has solution

$$u(t,x) = (T(t)u_0)(x) + \int_0^t (T(t-s)g(s))(x)ds. \tag{4.16}$$

The existence of this solution is known as Duhamel's principle.

Proof. Firstly, note that the initial condition is satisfied, as the integral equals zero at $t = 0$. Also, as $T(t)$ solves the homogeneous equation, $\partial_t - L$ applied to the first term on the right hand side of (4.16) causes is to vanish as the term solves the homogeneous equation. Thus,

$$\begin{aligned}
(\partial_t - L)u(t,x) &= (\partial_t u - Lu)(t,x) \\
&= (\partial_t - L)\left((T(t)u_0)(x) + \int_0^t (T(t-s)g(s))(x)ds\right) \\
&= (\partial_t - L)\left(\int_0^t (T(t-s)g(s))(x)ds\right) \\
&= (T(0)g(t))(x) + \int_0^t (\partial_t - L)(T(t-s)g(s))(x)ds \\
&= g(t,x).
\end{aligned}$$

□

4.2.1 Example of Duhamel's Principle

To illustrate the previous theorem, we look at $p(\lambda) = \mp\lambda^3$, with the homogeneous problem

$$u_t \mp D_x^3 u = 0, \qquad u(0,x) = u_0(x).$$

From proposition 4.1.3, the homogeneous solution has form

$$\begin{aligned}
u_H(t,x) &= \mathcal{F}^{-1}\left(e^{\pm i\xi^3 t}\hat{u}_0\right) \\
&= (T(t)u_0)(x).
\end{aligned} \tag{4.17}$$

The inhomogeneous problem takes form

$$u_t \mp D_x^3 u = g(t,x), \qquad u(0,x) = u_0(x),$$

and from Duhamel's principle we obtain the solution

$$\begin{aligned}
u(t,x) &= (T(t)u_0)(x) + \int_0^t (T(t-s)g(s))(x)ds \\
&= u_H(t,x) + \int_0^t \mathcal{F}^{-1}\left(e^{\pm i\xi^3(t-s)}\hat{g}(\xi)\right)ds \\
&= u_H(t,x) + \int_0^t \int_{\mathbb{R}^n} e^{ix\xi}e^{\pm i\xi^3(t-s)}\hat{g}(\xi)d\xi ds \\
&= u_H(t,x) + \int_0^t \int_{\mathbb{R}^n} e^{i\xi(x\pm(t-s))}\hat{g}(\xi)d\xi ds.
\end{aligned}$$

For the integral to be defined, the Fourier transform of g must exist, which it does for $g \in S'(\mathbb{R})$. If g is in the same class as u_0, u will be in said class, or; u will be in the same class as u_H.

4.3 Non-Polynomial Differentiation Operators

So far we have only considered f as a polynomial, which yields ordinary differential equations, but as we defined f by its Fourier transform, we could just as well use non-polynomials. This does however bring us out of the comfort-zone of the well-known theory of partial and ordinary differential equations. For instance, if we set $f(\lambda) = \sin(\lambda)$, or initial value problem, in the homogeneous case becomes

$$u_t + \sin(D)u = 0, \tag{4.18}$$

and the Fourier transformation yields

$$\hat{u}_t + \sin(i\xi)\hat{u} = 0.$$

This has solution

$$\begin{aligned} \hat{u} &= e^{-\sin(i\xi)t}\hat{u}_0 \\ &= e^{-i\sinh(\xi)t}\hat{u}_0. \end{aligned}$$

From the inverse Fourier transform we obtain

$$\begin{aligned} u(t,x) &= \int_{\mathbb{R}^n} e^{ix\xi} e^{-i\sinh(\xi)t}\hat{u}_0(\xi)d\xi \\ &= \int_{\mathbb{R}^n} e^{i(x\xi - t\sinh(\xi))}\hat{u}_0(\xi)d\xi. \end{aligned}$$

Now, because sinh takes real values for real domains,

$$|e^{-i\sinh(\xi)t}| \leq 1,$$

Thus, the same results as in proposition 4.1.1 hold. To connect non-polynomial differential operators to classical polynomials differential operators, one can consider the Taylor-expansion of the sinus function and consider the initial value problem to take the form of an infinite-order differential equation, with an infinite amount of differentiation terms. As the Taylor expansion of sin only involves terms with odd expansion, it is analogous to case 2 in the superpositioning section, as

$$|\hat{u}| \leq |\hat{u}_0|.$$

In heuristic terms, if we were to look at $f(\lambda) = \cos(\lambda)$, the same property would not be obtained, but a resemblance to polynomials of even-order would be evident as the Taylor-expansion of $f(\lambda) = \cos(\lambda)$ only involves terms of even order, and with signs corresponding to case 1. On the other hand, the Taylor expansion of

$f(\lambda) = -\cos(\lambda)$ consists, with the exception of a constant term, only of terms from case 3, which makes the initial value problem it defines allow for a large class of initial value conditions. Taking this one step further, one could even analyse the initial value problem for non-elementary functions, f, however, strong results and proof of well-posedness would not come easily.

Chapter 5

The Riesz Representation Theorem

The Riesz represnetation theorem guarantees a unique (weak) solution to initial value problems under certain criteria.

Proposition 5.0.1. *A linear functional L on a Banach space B is continuous if and only if it is bounded, or if there exists a constant C such that*

$$|L(v)| \leq C\|v\|_B \qquad \textit{for every} \quad v \in B.$$

Proof. A bounded linear function is Lipschitz continuous:

$$\begin{aligned} |L(u) - L(v)| &= |(L(u - v)| \\ &\leq C\|u - v\|_B \qquad \text{for all} \quad u, v \in B. \end{aligned}$$

Conversely, suppose L is continuous, if it is not bounded there must exists a sequence $\{v_n\}$ in B such that

$$\frac{|L(v_n)|}{\|v_n\|_B} \geq n.$$

Through the normalization $w_n = \frac{v_n}{n\|v_n\|_B}$, so that $|L(w_n)| \geq 1$. But we have $\|w_n\|_B \leq \frac{1}{n}$, so $w_n \to 0$, and by continuity of L we should have $L(w_n) \to 0$, which is a contradiction. □

Corollary 5.0.2. *For a continuous linear functional L on a Banach space B the expression*

$$\|L\|_{B'} = \sup_{0 \neq v \in B} \frac{L(v)}{\|v\|_B}$$

is always finite, and forms a norm on B'. It can be shown that B' is complete with respect to this norm, called the dual norm, so B' is also a Banach space.

Theorem 5.0.3 (The Riesz representation theorem)**.** *Any continuous linear functional L on a Hilbert space H can be represented uniquely as $L(v) = (u, v)$ for some $u \in H$, and furthermore we have*

$$\|L\|_{H'} = \|u\|_H,$$

and thus H is isomorphic to H'.

Proof. First, we show the uniqueness-property:

$$\begin{aligned} 0 &= L(u_1 - u_2) - L(u_1 - u_2) \\ &= (u_1, u_1 - u_2) - (u_2, u_1 - u_2) \\ &= (u_1 - u_2, u_1 - u_2), \end{aligned}$$

and thus $u_1 = u_2$. Next, define

$$M = \{v \in H : L(v) = 0\}.$$

Obviously, $M \subset H$, so

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

(See [2], p. 204-205, "the projection theorem")

If $M^\perp = \{0\}$ we have $M = H$ and we take $u = 0$.

If $M^\perp \neq \{0\}$, choose a non-zero $z \in M^\perp$, then $L(z) \neq 0$. For $v \in H$ and $\beta = \frac{L(v)}{L(z)}$ we have

$$L(v - \beta z) = L(v) - \beta L(z) = 0,$$

or $v - \beta z \in M$. Thus, $v - \beta z \in P_{M^v}$ and $\beta z \in P_{M^\perp} v$. In particular, if $v \in M^\perp$, then $v = \beta z \in P_{M^\perp}$, which proves that $M^\perp$ is one-dimensional. Now choose $u = \frac{L(z)}{\|z\|_H^2} z$. Note that $u \in M^\perp$ and we have

$$\begin{aligned} (u, v) &= (u, (v - \beta z) + \beta z) \\ &= (u, v - \beta z) + (u, \beta z) \\ &= (u, \beta z), \quad (u \in M^\perp, v - \beta z \in M) \\ &= \beta \frac{L(z)}{\|z\|_H^2} (z, z) \\ &= \beta L(z) \\ &= L(v). \end{aligned}$$

So we choose $u = \frac{L(z)}{\|z\|_H^2}$ as our element in H.

It remains to prove that $\|L\|_{H'} = \|u\|_H$. We observe that

$$\|u\|_H = \frac{|L(z)|}{\|z\|_H}.$$

From the definition of the dual norm we have

$$\begin{aligned}\|L\|_{H'} &= \sup_{0\neq v\in H} \frac{|L(v)|}{\|v\|_H} \\ &= \sup_{0\neq v\in H} \frac{|(u,v)|}{\|v\|_H} \\ &\leq \|u\|_H \\ &= \frac{|L(z)}{\|z\|_H} \\ &\leq \|L\|_{H'}.\end{aligned}$$

Thus, $\|u\|_H = \|L\|_{H'}$. □

5.1 Example of the Riesz Representation Theorem

In order to understand the application of the Riesz representation theorem, an example is applied to the following Dirichlet problem:

$$\begin{aligned}-\Delta u + u &= g \qquad \text{in } \mathbb{R}^n \\ u &= 0 \qquad \text{on } \partial\mathbb{R}^n.\end{aligned}$$

Then $u \in H^1(\mathbb{R}^n)$ is a weak solution if

$$\int_{\mathbb{R}^n} (Du \cdot Dv + uv)dx = \langle g, v\rangle_{L^2(\mathbb{R}^n)} \qquad \text{for all} \quad v \in H^1(\mathbb{R}^n).$$

By defining $(u,v)_1 = \int_{\mathbb{R}^n}(Du \cdot Dv + uv)dx$, this can be rewritten to

$$(u,v)_1 = \langle g, v\rangle_{L^2(\mathbb{R}^n)} \qquad \text{for all} \quad v \in H^1(\mathbb{R}^n).$$

Now, $(g, \cdot)$ is a continuous linear functional on $H^1(\mathbb{R}^n)$ as

$$\begin{aligned}\int_{\mathbb{R}^n} \hat{g}(\xi)\hat{\bar{v}}(\xi)d\xi &\leq \|\hat{g}\|_{L^2(\mathbb{R}^n)}\|\hat{v}\|_{L^2(\mathbb{R}^n} \\ &\leq \|g\|_{L^2(\mathbb{R}^n)}\|v\|_{H^1(\mathbb{R}^n)}.\end{aligned}$$

Which follows from $1 \leq (1 + |\xi|^s)$. Thus, by the Riesz representation theorem, there exists a unique $u \in H^1(\mathbb{R}^n)$ solving this Dirichlet problem.

Chapter 6

The Lax–Milgram Theorem

In an attempt to generalize the results found in chapter 5, we will prove the Lax-Milgram theorem, but in order to get there we need to establish some definitions and results.

Definition A bilinear form, $a(\cdot, \cdot)$, is a function on a normed linear space, H that is linear in both arguments:

- $a(u + v, w) = a(u, w) + a(v, w)$.
- $a(u, v + w) = a(u, v) + a(u, w)$.
- $a(\lambda u, w) = a(u, \lambda v) = \lambda u(v, w)$.

We say that the bilinear form is bounded and continuous if there exists a constant $c < \infty$ such that

$$|a(v, w)| \leq c\|v\|_H\|w\|_H \qquad \text{for all} \quad v, w \in H.$$

We say that the bilinear form is coercive on the subset $V \subset H$ if there exists an $\alpha > 0$ such that

$$a(v, v) \geq \alpha\|v\|_H^2,$$

and call α the coercivity of a on V.

We say that the bilinear form if symmetric on V if

$$a(v, w) = a(w, v) \qquad \text{for all} \quad v, w \in H.$$

Proposition 6.0.1. *If H is a Hilbert space and $a(\cdot, \cdot)$ is a symmetric bilinear form on H which is coersive on a closed subset $V \subset H$, then $(V, a(\cdot, \cdot))$ is a Hilbert space.*

Proof. Because $a(\cdot,\cdot)$ is coercive on V it is an an inner-product on V. Let $\|v\|_E = \sqrt{a(v,v)}$ and let $\{v_n\}$ be a Cauchy-sequence in $(V, \|\cdot\|_E)$. By coersivity, v_n is also a Cauchy-sequence in $(H, \|\cdot\|_H)$ and because H is complete there exists a $v \in H$ such that $v_n \to v$ in $\|\cdot\|_H$. Because V is closed in H, $v \in V$. Next, because $a(\cdot,\cdot)$ is bounded we have that

$$\|v - v_n\|_H \leq \sqrt{c_1}\|v - v_n\|_H,$$

and thus $v_n \to v$ in $\|\cdot\|_E$, so $(V, \|\cdot\|_E)$ is complete. □

Definition If

1. H is a Hilbert space
2. V is a closed subspace of H
3. $a(\cdot,\cdot)$ is a bounded, symmetric bilinear form on H that is coercive on V

the symmetrical variational problem is as follows:

Given $F \in V'$, find $u \in V$ such that

$$a(u, v) = F(v) \qquad \text{for all} \quad v \in V. \tag{6.1}$$

Theorem 6.0.2. *Suppose points* $1-3$ *in the previous definition holds, then there exists a unique* $u \in V$ *solving the symmetrical variational problem.*

Proof. By proposition 6.0.1, $a(\cdot,\cdot)$ is an inner-product on V, and $(V, a(\cdot,\cdot))$ is a Hilbert space, so by the Riesz representation theorem the theorem holds. □

Definition Given a finite-dimensional subspace $V_n \subset V \subset H$, where H is a Hilbert space and $F \in V'$, the problem of finding u_h in V_h such that

$$a(u_h, v) = F(v) \qquad \text{for all} \quad v \in V_h,$$

is known as the Ritz–Galerkin approximation problem.

Theorem 6.0.3. *Under conditions* $1-3$ *of the symmetrical variational problem there exists a unique* u_h *that solves the Ritz–Galerkin approximation problem.*

Proof. We know that $(V_h, a(\cdot,\cdot))$ is a Hilbert space and that $F \in V_h'$, the result follow from the Riesz representation theorem. □

Remark If u solves the symmetrical variatonal problem and u_n solves the Ritz–Galerkin approcimation problem, then we observe that

$$a(u - u_h, v) = 0 \qquad \text{for all} \quad v \in V_h.$$

And in the symmetric case, u_h minimizes the quadratic functional

$$Q(v) = a(v, v) - 2F(v) \qquad \text{for all} \quad v \in V_h.$$

Theorem 6.0.4. *For a Banach space, V, and a mapping $T : V \to V$ satisfying*

$$\|Tv_1 - Tv_2\| \le M\|v_1 - v_2\| \qquad \textit{for all} \quad v_1, v_2 \in V \ \textit{and fixed} \quad M : 0 \le M < 1,$$

there exsts a unique $v \in V$ such that

$$u = Tu.$$

In other words: The contraction mapping T has a unique fixed point, u.

Proof. We start by showing the existence of the fixed point: Let $v_0 \in V$ and define

$$v_1 = Tv_0, v2 = Tv_1, \cdots, v_{k+1} = Tv_k.$$

Note that

$$\begin{aligned} \|v_{k+1} - v_k\| &= \|Tv_k - Tv_{k-1}\| \\ &\le M\|v_k - v_{k-1}\|. \end{aligned}$$

By induction,

$$\|v_k - v_{k-1}\| \le M^{k-1}\|v_1 - v_0\|.$$

Thus, for every $N > n$ we have

$$\begin{aligned} \|v_N - v_n\| &= \| \sum_{k=n+1}^{N} v_k - v_{k-1} \| \\ &\le \|v_1 - v_0\| \sum_{k=n+1}^{N} M^{k-1} \\ &\le \frac{M^n}{1-M}\|v_1 - v_0\| \\ &= \frac{M^n}{1-M}\|Tv_0 - v_0\|. \end{aligned}$$

From this it follows that $\{v_n\}$ is a Cauchy sequence. Because V is complete and $\lim_{n\to\infty} v_n = v$, we have

$$\begin{aligned} v &= \lim_{n\to\infty} v_{n+1} \\ &= \lim_{n\to\infty} Tv_n \\ &= T(\lim_{n\to\infty} v_n) \\ &= Tv, \end{aligned}$$

so there exists a fixed point.

Next, we want to prove that the fixed point is unique. Let $Tv_1 = v_1$ and $Tv_2 = v_2$, since T is a contraction mapping we have

$$\|Tv_1 - Tv_2\| \leq M\|v_1 - v_2\|$$

for some M between 0 and 1. But

$$\|Tv_1 - Tv_2\| = \|v_1 - v_2\|,$$

therefore

$$\|v_1 - v_2\| \leq M\|v_1 - v_2\|,$$

giving us $\|v_1 - v_2\| = 0$ (otherwise we would require $M \geq 1$). □

Definition Under the conditions:

1. $(H, (\cdot,\cdot))$ is a Hilbert space
2. V is a closed subspace of H
3. $a(\cdot,\cdot)$ is a continuous bilinear form on V
4. $a(\cdot,\cdot)$ is continuous (bounded) on V
5. $a(\cdot,\cdot)$ is coercive on V

The following problem is known as the non symmetric variational problem:

Given $F \in V'$, find $u \in V$ such that

$$a(u, v) = F(v) \text{ for all } \quad v \in V.$$

Theorem 6.0.5 (The Lax–Milgram theorem)**.** *Let $(V, (\cdot,\cdot))$ be a Hilbert space and $a(\cdot,\cdot)$ a continuous, coercive bilinear form on V. For $F \in V'$, there exists a unique $u \in V$ such that*

$$a(u, v) = F(v) \qquad \text{for all } v \in V. \tag{6.2}$$

Proof. For any $u \in V$ we define a functional $Au(v) = a(u, v)$ for all $v \in V$. Au is linear because

$$\begin{aligned} Au(\lambda_1 v_1 \lambda_2 v_2) &= a(u, \lambda_1 v_1 + \lambda_2 v_2) \\ &= \lambda_1 a(u, v_1) + \lambda_2 a(u, v_2) \\ &= \lambda_1 Au(v_1) + \lambda_2 Au(v_2) \qquad \text{for all} \quad v1, v2 \in V, \quad \lambda_1, \lambda_2 \in \mathbb{R}. \end{aligned}$$

Au is continuous because

$$\begin{aligned} |Au(v)| &= |a(u, v)| \\ &\leq C\|u\|\|v\| \qquad \text{for all} \quad v \in V. \end{aligned}$$

Thus,

$$\begin{aligned}\|Au(v)\|_{V'} &= \sup_{v\neq 0}\frac{Au(v)}{\|v\|}\\ &\leq C\|u\|\\ &< \infty.\end{aligned}$$

Yielding $Au \in V'$. In similar fashion one can show that the mapping $u \to Au$ is a linear map $V \to V'$.

By the Riesz representation theorem, for every $\phi \in V'$ there exists a unique $\tau\phi \in V$ such that

$$\phi(v) = (\tau\phi, v) \qquad \text{for all} \quad v \in V.$$

We need to find a unique u such that

$$Au(v) = F(v) \qquad \text{for all} \quad v \in V,$$

in other words: We want to find a unique u such that

$$Au = f$$

in V', or

$$\tau Au = \tau F$$

in V, since $\tau : V' \to V$ is a one-to-one mapping. By the contraction mapping principle, we want to find a $\rho \neq 0$ such that the mapping $T : V \to V$ is a contraction mapping when T is defined by

$$Tv = v - \rho(\tau Av - \tau F) \qquad \text{for all} \quad v \in V.$$

If T is a contraction mapping, then by the contraction mapping principle there exists a unique $u \in V$ such that $Tu = u - \rho(\tau Au - \tau F) = u$, that is: $\rho(\tau Au - \tau F) = 0$ or $\tau Au = \tau F$. It remains to show that such a $\tau \neq 0$ exists.

For every $v1, v2 \in V$, let $v = v_1 - v_2$, then

$$\begin{aligned}\|Tv_1 - Tv_2\|^2 &= \|v_1 - v_2 - \rho(\tau Av_1 - \tau Av_2)\|^2\\ &= \|v - \rho(\tau Av)\|^2\\ &= \|v\|^2 - 2\rho(\tau Av, v) + \rho^2\|\tau Av\|^2\\ &= \|v\|^2 - 2\rho Av(v) + \rho^2 Av(\tau Av)\\ &= \|v\|^2 - 2\rho a(v, v) + \rho^2 a(v, \tau Av)\\ &\leq \|v\|^2 - 2\rho\alpha\|v\|^2 + \rho^2 C\|v\| \cdot \|\tau Av\|\\ &\leq (1 - 2\rho\alpha + \rho^2 C^2)\|v\|^2\\ &\leq (1 - 2\rho\alpha + \rho^2 C^2)\|v_1 - v_2\|^2\\ &= M^2\|v_1 - v_2\|^2.\end{aligned}$$

Thus, we require $1 - 2\rho\alpha + \rho^2 C^2 < 1$ for some ρ, by choosing $\rho \in (0, \frac{2\alpha}{C^2})$ we obtain $M < 1$ and the proof is complete. □

The Lax–Milgram theorem has great value in the field of differential equations, and together with the Riesz representation theorem it does, to some extent, generalize the results obtained in chapter 4. When applying the Lax–Milgram theorem, one needs to check properties of the bilinear form and the functional defining a weak solution, which will often be a lot easier than showing the existence of a unique solution directly. For examples and applications of the Lax–Milgram theorem, see [1].

Chapter 7

Bibliography

1. L. C. Evans; Partial Differential Equations, ISBN: 0-8218-0772-2, 1998.

2. A. Friedman; Foundations of modern analysis, ISBN 0-486-64062-0, 1982.

3. C. Gasquet, & P. Witomski; Fourier Analysis and Applications, ISBN: 0-387-98485-2, 1998.

4. L. Hörmander; The analysis of Linear Partial Differential Operators, Volume $\mathbb{I}$: Distribution Theory and Fourier Analysis, ISBN: 3-540-00662-1, 2007.

5. D. D. Haroske & H. Triebel; Distributions, Sobolev Spaces, Elliptic Equations, ISBN: 978-3-03719-042-5, 2007.

6. L. Hörmander; The Analysis of Linear Partial Differential Operators, Volume $\mathbb{III}$: Differential Operators with Constant Coefficients, ISBN: 3-540-12139-0, 1983.

7. D. W. Jordan & P. Smith; Nonlinear Ordinary Differential Equations, ISBN; 978-0-19-920825-8, 2008.

8. R. S. Strichartz; A Guide to Distribution Theory and Fourier Transforms, ISBN: 9812384219, 2003.

9. A. Tychonoff; Théorèmes d'unicité pour l'équation de la chaleur Mat. Sb., 1935 42, p 199–216, http://mi.mathnet.ru/msb6410, JFM 61.1203.05